

COME VALUTARE I BENI PUBBLICI? GUIDA AL METODO DELLA VALUTAZIONE CONTINGENTE¹

di **Elisabetta Strazzer**a e **Rinaldo Brau**
DRES e CRENoS, Università di Cagliari

1. Il Metodo della Valutazione Contingente

La valutazione di un investimento pubblico si basa sul computo dei costi e dei benefici sociali connessi all'intervento: la scelta di porre in essere tale intervento dipende dal rapporto tra benefici e costi sociali. Come determinare però il valore di quei beni, come per esempio i beni culturali o ambientali, che non vengono scambiati sul mercato dei beni privati e dunque non assumono un prezzo che ne rifletta l'effettivo valore? In alcuni casi è possibile fare ricorso ai cosiddetti mercati surrogati: si considerano cioè i prezzi di altri beni, che si ritiene possano bene approssimare il valore dell'oggetto dell'analisi. In altri casi ciò non è possibile, e l'unico modo per determinare questo valore è quello di stimare la sua valutazione da parte dei potenziali fruitori del bene. La cosiddetta valutazione contingente è una tecnica che si basa sulla rilevazione diretta delle valutazioni individuali attraverso indagini campionarie, in maniera simile alle indagini di mercato effettuate per i normali beni di consumo attraverso le interviste ai consumatori. La differenza sostanziale è che in questo caso l'indagine ha come oggetto un bene pubblico, il che rende molto più difficile il processo di determinazione del prezzo di riserva individuale.

Questo lavoro ha lo scopo di tracciare una guida al metodo della Valutazione Contingente: si descriveranno le varie fasi, a partire dalla scelta dell'universo campionario fino alla stima del valore da inserire nell'analisi costi-benefici. Ovvii limiti di spazio non permettono in questa sede una trattazione approfondita di tutti gli argomenti: si è perciò deciso di diversificare il livello di dettaglio dell'analisi. Per gli argomenti di carattere più generale, comuni a diversi tipi di indagini campionarie, ci si è limitati ad una breve descrizione del problema, rinviando il lettore a ben consolidati riferimenti bibliografici per gli approfondimenti del caso. Verrà invece concesso maggiore spazio alla trattazione di alcuni problemi specifici delle analisi di Valutazione Contingente, in parte ancora in discussione tra gli studiosi del settore; i numerosi riferimenti bibliografici potranno comunque orientare il lettore in questa letteratura in evoluzione, per l'approfondimento dei temi di maggiore interesse.

2. Definizione della popolazione di riferimento e individuazione del campione.

Le valutazioni dei beni pubblici basate sulla misurazione delle preferenze individuali richiedono che la popolazione di riferimento sia individuata in modo molto preciso. Per una corretta determinazione del valore economico totale del bene in esame, occorrerebbe che l'universo ricomprenda ogni individuo il cui benessere sia influenzato dal bene in questione.

¹ Questo lavoro è stato in parte realizzato grazie al finanziamento della ricerca: "I progetti di riforestazione nei parchi naturali della Sardegna: una valutazione economica degli effetti ambientali" diretta da Sergio Lodde ed Elisabetta Strazzer, con contributo L.R. n.40/99, art. 19; e con contributo L.R. n.2/94, art.37, contratto di formazione in attività di ricerca attribuito a Rinaldo Brau, da parte della Regione Sardegna.

Se in alcuni casi la definizione della popolazione appare relativamente semplice -ad esempio, è ragionevole ritenere che i beneficiari di un programma di difesa nazionale siano individuati dai confini nazionali- in altre situazioni il ricercatore deve utilizzare criteri più soggettivi: quali sono i limiti della popolazione interessata alla conservazione di una foresta pluviale? Quali quelli della popolazione disposta a pagare per il restauro di un'opera d'arte di rilevanza internazionale? In queste condizioni in effetti il rischio è quello di dover ricomprendere l'intera popolazione del pianeta. Questi esempi possono essere considerati dei casi limite, ma danno una chiara idea dei problemi che la valutazione contingente pone anche su scala più modesta.

Una volta definita la popolazione complessiva di riferimento si pone il problema di scegliere il criterio di campionamento. Quando la popolazione è relativamente omogenea rispetto alla variabile di interesse, in questo caso la valutazione del bene, il criterio più indicato è quello del campionamento casuale semplice. Se invece, come spesso accade, il valore attribuito al bene è correlato a specifiche caratteristiche socio-economiche degli individui, è preferibile ricorrere al campionamento stratificato, con diversi gruppi campionari di numerosità proporzionale a quella che si riscontra nella popolazione di riferimento: a parità di numerosità campionaria, le stime saranno più efficienti rispetto al metodo casuale semplice.

Il pericolo maggiore di errore campionario deriva da processi di troncamento della distribuzione della popolazione, o dello stesso campione. Il campione dovrebbe rispecchiare la distribuzione della popolazione oggetto di analisi, ed a tale fine sarebbe perciò necessario utilizzare un elenco completo della popolazione, come per esempio le liste anagrafiche dei comuni: nel caso di popolazioni molto ampie questo metodo risulta però troppo oneroso, e si preferisce ricorrere agli elenchi delle società telefoniche, al prezzo di escludere dal campionamento quella fascia di popolazione (generalmente a basso reddito) che non possiede un impianto telefonico. La preparazione della lista della popolazione può risultare ancora più complessa se invece l'universo statistico di riferimento è molto settoriale. In questi casi occorre spesso svolgere un lavoro di censimento sul campo, il che non esclude il rischio di troncamento della popolazione, dovuto al fatto che certi gruppi possono essere più difficilmente individuati (per esempio, se la popolazione di riferimento è l'insieme dei pescatori in un determinato bacino, è difficile individuare il gruppo dei pescatori di frodo). Ma il maggiore rischio è quello di troncamento della distribuzione campionaria: le fasce di popolazione più anziane e meno alfabetizzate, ad esempio, presentano dei tassi di rifiuto verso l'intervista più alti rispetto alla media, a causa di maggiori difficoltà nella compilazione del questionario, o anche nella comunicazione con l'intervistatore. Viceversa, i soggetti più interessati alla fornitura del bene pubblico oggetto di analisi sono invece più propensi ad accettare l'intervista al fine di manifestare tale interesse: la combinazione dei due elementi crea un processo distorsivo di autoselezione. In questi casi può rendersi necessaria una scomposizione del campione per riuscire a stabilire la forma di contatto ottimale con coloro ai quali male si adattano le procedure di indagine più generali, in modo da garantire che queste fasce di popolazione siano rappresentate. Per una trattazione approfondita delle tecniche di campionamento si rinvia al manuale di Cicchitelli et al. (1992), ed a Mitchell e Carson (1989) per gli aspetti più specificamente connessi alle analisi di Valutazione Contingente.

3. La costruzione del questionario

Nelle analisi di Valutazione Contingente gli intervistati devono dare delle informazioni su quanto sarebbero disposti a pagare affinché venga offerto un determinato bene pubblico. Dato che al momento dell'intervista il consumo del bene, e soprattutto il pagamento, hanno carattere del tutto ipotetico, si corre il rischio di ottenere delle risposte poco realistiche: a questo proposito si parla talvolta di *distorsione da ipoteticità*. La familiarità degli intervistati con gli elementi chiave della situazione delineata può ridurre notevolmente questo rischio: la tecnica di rilevazione² e la forma di pagamento proposta dovrebbero cercare di riprodurre processi di valutazione e di pagamento già sperimentati dagli individui in situazioni reali. Inoltre, gli intervistati dovrebbero essere ben informati sia sul bene oggetto dell'indagine, sia sul livello di fornitura dello stesso: questo anche per ridurre il rischio di un'altra forma di distorsione, detta *parte per il tutto*, per la quale gli individui tendono a considerare come oggetto di valutazione non il bene in esame, ma un insieme che lo contiene: per esempio, si attribuisce alla protezione di una specie il valore che si attribuirebbe alla protezione della biodiversità in generale. Il supporto di materiale illustrativo (documenti, fotografie, grafici) può essere utile in questo senso, a patto che l'informazione sia equilibrata e non distorsiva. Se infatti così non fosse, il rimedio potrebbe essere peggiore del male, dato che la carenza di informazione causerebbe un incremento dell'errore casuale di rilevazione, ma non la distorsione delle stime (Mitchell e Carson (1989:216-17)).

D'altra parte un altro, e forse più grave, motivo di distorsione è dato dal comportamento strategico dei soggetti economici. Se gli individui ritengono che le loro risposte potrebbero produrre degli effetti sul proprio livello di benessere futuro, la CV potrebbe essere esposta al rischio di risposte artatamente falsate, miranti ad influenzare il livello di fornitura del bene pubblico o il livello di contribuzione individuale. In particolare, il carattere pubblico del bene oggetto dell'analisi rende le risposte particolarmente esposte al rischio di *free riding*, anche se questo rischio sembra in qualche modo ridimensionato dai risultati di numerosi esperimenti (si veda Santagata, (1992:71-72) per una discussione su questo punto). Insistere sul carattere ipotetico dell'indagine può rendere meno grave questo problema -tuttavia aggravando il rischio di distorsione da ipoteticità.

Un altro problema legato al comportamento degli intervistati dipende invece dal desiderio, conscio o inconscio, di non essere disapprovati dall'intervistatore, o eventualmente dall'opinione pubblica: in questo caso la dichiarazione dell'individuo tende a riflettere non la propria disponibilità a pagare, bensì la risposta che si ritiene più gradita. Per alleviare questo problema occorre formulare le domande nel modo più neutro possibile, facendo in modo che nessun tipo di risposta possa essere percepita come più o meno accettabile di altre.

Infine consideriamo un particolare problema, che può manifestarsi nel caso in cui le preferenze degli intervistati siano disomogenee circa il segno del valore da attribuire all'oggetto dell'indagine. Quando l'oggetto della valutazione è considerato un "bene" economico, gli individui sono disposti a pagare un prezzo positivo (o nella peggiore delle ipotesi, nullo) perchè tale bene sia reso disponibile. Se una parte dei potenziali fruitori si ritiene invece danneggiata da esso, lo considera cioè un "male", la disponibilità a pagare diventa negativa: in altri termini, dovrebbero essere compensati monetariamente per la

² In inglese, *elicitation method*.

sua fornitura. Se una parte del campione valuta positivamente, ed un'altra valuta negativamente lo stesso oggetto, si vengono a creare delle particolari difficoltà nell'aggregazione dei dati. Supponiamo di chiedere, a coloro che rispondono di essere contrari alla fornitura del bene pubblico, il valore della compensazione monetaria che li può indurre ad accettare questa decisione: una formulazione di questo tipo può indurre gli intervistati a comportamenti strategici, a meno che non si escogitino dei particolari accorgimenti per indurre gli individui a dare risposte veritiere³. Una possibile alternativa consiste nel chiedere agli individui contrari quanto sarebbero disposti a pagare per evitare la fornitura del bene pubblico: il problema qui può nascere dall'eccessiva artificiosità della situazione proposta (si deve pensare a quanto pagare per non acquistare un bene), e quindi determinare una distorsione da ipoteticità. In generale, si può affermare che qualunque strategia impiegata per incentivare l'individuo a rivelare il proprio vero prezzo di riserva produce qualche forma di distorsione: è compito del ricercatore individuare il metodo più adatto nella particolare situazione in esame. Finora, la maggior parte degli studi ha trascurato il problema della valutazione negativa, imponendo un vincolo di non negatività alle risposte sui prezzi di riserva individuali; ma il crescente ricorso alla valutazione contingente come supporto informativo e di orientamento delle scelte pubbliche rende certamente necessario un maggiore impegno per la sua risoluzione.

4. Tecniche di rilevazione del prezzo di riserva

Le diverse tecniche con le quali vengono formulate le domande sul prezzo di riserva del soggetto danno origine ai diversi modelli elaborati nella letteratura empirica.

4.1. Il metodo ad intervallo aperto

Il metodo più semplice prevede che si richieda direttamente agli intervistati di dichiarare quanto si è disposti a pagare per usufruire di un determinato bene pubblico, ed è conosciuto come procedura ad *intervallo aperto*: pur presentando il vantaggio di rendere minima l'influenza dell'intervistatore, essa non è molto usata in quanto il rischio di ottenere valori non realistici è molto elevato. E' infatti piuttosto difficile individuare, con una semplice analisi introspettiva, il proprio prezzo di riserva per un bene: gli intervistati sono invece abituati, in quanto consumatori, a sentirsi proporre un prezzo, che sarà poi pagato o meno a seconda del proprio prezzo di riserva. L'assenza di un prezzo di riferimento nella procedura ad intervallo aperto provoca in genere elevati tassi di mancata risposta: l'impiego di questo metodo è pertanto indicato solo per quei casi in cui il bene in esame abbia dei succedanei il cui prezzo sia ben noto agli intervistati. In questo caso l'esperienza dimostrerebbe che la tecnica ad intervallo aperto funziona in maniera soddisfacente (Mitchell e Carson (1989)). Questo metodo si applica particolarmente bene a quei casi in cui sia già fissato il prezzo di un bene con determinate caratteristiche, e si voglia conoscere il prezzo di riserva per successive modificazioni di tali caratteristiche. Inoltre, viene in genere utilizzato nella fase preliminare delle analisi di valutazione

³ Inoltre, numerosi esperimenti sembrano mostrare che, a prescindere da motivi strategici, la disponibilità ad accettare un compenso per privarsi di un bene è significativamente maggiore della disponibilità a pagare per ottenere lo stesso bene: le scelte dell'individuo sono cioè determinate da funzioni di utilità diverse, condizionate allo stato di possesso o non possesso del bene. Si veda p.es. Adamowicz et al. (1993), e Kahneman et al. (1990).

contingente, o fase di *pretest*, necessaria per definire il *range* di valori che verranno proposti con i metodi ad *intervallo chiuso*.

Nelle procedure ad intervallo chiuso vengono proposte all'intervistato una o più cifre, e questi dovrà indicare se esse corrispondono o meno al suo prezzo di riserva. All'interno di questa categoria si possono individuare tre metodi.

4.2. Il metodo dell'asta

Si tratta di uno dei primi modelli utilizzati negli studi di Valutazione Contingente. Davis (1964), ha introdotto tale tecnica di determinazione del prezzo cercando di riprodurre un meccanismo familiare per gli intervistati: si propone un prezzo, e si chiede a ciascun individuo se sarebbe disposto a spendere quella cifra per consumare il bene pubblico. Se la risposta è positiva si proporranno valori successivamente più elevati fino a quando l'intervistato rifiuta di pagare un certo prezzo; se la risposta è negativa, si proporranno valori più bassi finché l'intervistato dichiarerà di essere disposto a pagare un certo prezzo. Come è stato messo in evidenza da Cummings et al. (1986), questa tecnica permette di identificare l'intero surplus del consumatore, conducendo l'intervistato, attraverso l'approccio iterativo, ad una piena consapevolezza del proprio prezzo di riserva per il bene. Inoltre, il compito dell'intervistato è molto semplice: deve solo rispondere SI o NO alle domande che gli vengono successivamente rivolte.

D'altra parte questo modello comporta anche notevoli problemi, primo fra tutti quello noto come *distorsione del punto di partenza*, dovuto al fatto che a tutti i soggetti venga proposta la stessa offerta iniziale, la quale di fatto condiziona pesantemente il *range* nel quale vanno a collocarsi le risposte degli intervistati. In pratica queste ultime risultano ancorate al valore di partenza, soprattutto quando l'intervistato non ha un'idea molto precisa del valore del bene. Questo problema potrebbe forse essere evitato proponendo diversi prezzi iniziali, variabili da individuo ad individuo, così come accade nel modello a scelta dicotomica che verrà discusso nel prossimo paragrafo.

Ma i problemi principali consistono nel fatto che il processo di determinazione del prezzo, per quanto semplice, può risultare lungo e quindi mal tollerato dagli intervistati, che potrebbero dopo alcune iterazioni dare risposte non ben ponderate; e soprattutto nel fatto che richiede interviste dirette, mal prestandosi all'intervista telefonica (e per nulla a quella postale), ed è quindi notevolmente più costoso di altri metodi. Può tuttavia essere proficuamente impiegato per la fase di *pretest*.

4.3. Il metodo della scelta dicotomica

All'interno del gruppo dei metodi ad intervallo chiuso hanno di recente acquisito ampia diffusione i cosiddetti modelli a scelta dicotomica, nei quali l'intervistato viene posto di fronte ad una cifra indicante la propria disponibilità a pagare, cui egli deve semplicemente rispondere positivamente o negativamente. L'elaborazione e la prima applicazione di tale tecnica "prendere o lasciare" (DC-CVM, da *Dichotomous Choice Contingent Valuation Method*) è dovuta a Bishop e Heberlein (1979) che l'applicarono per lo studio della domanda di attività venatoria in un'area del Wisconsin. Il DC-CVM è attualmente il metodo più diffusamente impiegato nelle indagini di valutazione contingente. Fra i punti di forza di questo modello viene sottolineata l'estrema semplificazione del compito del rispondente, che, come nel caso del metodo dell'asta deve semplicemente rispondere SI o NO. La differenza sostanziale è che viene eliminato

il processo iterativo: ad ogni individuo viene infatti proposta una sola cifra nel caso del modello semplice, e due nel caso del modello doppio. Quest'ultimo, proposto da Carson, Hanemann e Mitchell (1986), incrementa l'informazione campionaria con una seconda domanda nella quale viene proposta un'altra somma (*follow-up*) che sarà maggiore o minore della prima a seconda che la prima risposta sia stata positiva o negativa. E' importante sottolineare la relativa inefficienza statistica dei modelli DC-CVM, dovuta al fatto che i dati campionari forniscono un'informazione molto approssimativa circa la disponibilità massima a pagare dell'intervistato: si può sapere solo se è maggiore o minore della cifra proposta. Per ottenere maggiore informazione sulla forma e sul supporto della distribuzione campionaria si varia la cifra proposta da individuo ad individuo; occorre comunque un più elevato numero di osservazioni rispetto ai modelli visti precedentemente. D'altra parte, come detto, il compito dell'intervistato è molto facilitato; inoltre questa è la tecnica che meglio si presta ai contatti a distanza per via telefonica o postale.

4.4. Il metodo della lista

Un'altra tecnica spesso utilizzata, proposta da Mitchell e Carson (1981), prevede la proposizione di una lista di valori, in un intervallo da zero a quello che si ritiene il limite superiore della disponibilità a pagare individuale. All'intervistato viene chiesto di sbarrare la cifra esatta che rappresenta la sua massima disponibilità a pagare (ed in questo caso come nel metodo dell'asta si arriva alla definizione esatta della WTP), o l'intervallo di valori entro i quali si colloca la sua WTP. Nelle sue applicazioni sul campo il metodo della lista può presentare diverse varianti. E' anzitutto possibile distribuire diverse liste in corrispondenza delle diverse categorie di reddito di appartenenza degli intervistati. Allo scopo di facilitare il processo di valutazione degli intervistati, è talvolta possibile ricorrere all'uso del cosiddetto *benchmark*: tra i valori monetari della lista viene inserito, e segnalato, l'ammontare mediamente pagato per un altro bene pubblico con caratteristiche simili.

Fra le caratteristiche positive di tale tecnica sono da ricordare la relativa facilità d'approccio con l'intervistato. D'altra parte, così come per il metodo dicotomico, il dato campionario non fornisce un'indicazione precisa circa il prezzo di riserva degli individui, ma solo indicazioni sull'intervallo in cui questo valore si colloca. Vedremo nella prossima sezione che la struttura statistica dei due modelli presenta caratteristiche simili.

5. Stima dei modelli e determinazione delle misure di tendenza centrale

Le analisi di valutazione contingente servono non soltanto per determinare il valore che gli intervistati attribuiscono al bene pubblico, ma anche per individuare i fattori socio-economici che determinano tale valutazione: in modo tale da consentire di spiegare e prevedere, date certe modificazioni della struttura economico-sociale di riferimento, eventuali modificazioni nella valutazione stessa. La scelta del modello econometrico da adottare dipende dal tipo di dati a disposizione, che a sua volta dipende dal metodo impiegato per la rilevazione del prezzo di riserva degli intervistati.

5.1. Il modello ad intervallo aperto e il metodo dell'asta

Come visto precedentemente, nel modello ad intervallo aperto gli intervistati dichiarano esplicitamente il prezzo di riserva, mentre con il metodo dell'asta vi sono condotti attraverso un processo iterativo. In entrambi i casi si ottiene un valore puntuale per la disponibilità a pagare dell'individuo, il che rende estremamente semplice la determinazione della disponibilità a pagare media (o mediana), data direttamente dalla media aritmetica (mediana) delle cifre dichiarate dai singoli intervistati. Nella maggior parte degli studi in cui sono stati utilizzati questi metodi, l'analisi di regressione sulle variabili socio-economiche è stata effettuata con il semplice metodo dei minimi quadrati ordinari. Tuttavia, come sottolineato da Donaldson et al. (1996), spesso la variabile dipendente non è in effetti continua, come ipotizzato nel modello OLS, ma è censurata al valore zero: il questionario viene infatti spesso progettato in modo tale da non consentire risposte inferiori allo zero. Occorrerebbe dunque tenerne conto in modo esplicito nell'analisi dei dati.

Il dato sulla disponibilità a pagare con valore zero può avere due motivazioni: può riflettere una genuina indifferenza rispetto al bene pubblico, o viceversa può essere determinato da elementi diversi: il vero valore può essere negativo; oppure molto basso e perciò approssimato dallo stesso intervistato a zero; oppure può esprimere il fastidio per l'intervista, la protesta contro l'idea di assegnare valori monetari a beni culturali o ambientali; o, ancora, può essere determinato da free riding. Una elevata percentuale di risposte con valore zero deve essere considerata come un segnale della presenza di questi elementi distorsivi, ed occorre apportare le necessarie correzioni in sede di analisi econometrica. In questo caso può essere appropriato l'uso di modelli di selettività, o di modelli a due parti (Tobit di tipo II). Nel caso in cui invece si ritenga che il valore zero rifletta effettiva indifferenza rispetto al bene pubblico, il modello econometrico più adatto è il Tobit semplice. Si rinvia ai manuali di Maddala (1983) e Greene (1990) per un approfondimento su questi modelli censurati.

5.2. I modelli a scelta dicotomica.

I modelli a scelta dicotomica presentano maggiori complessità rispetto a quelli visti precedentemente. Infatti in questo caso non si ricavano dall'intervista i valori precisi delle disponibilità a pagare degli intervistati: l'unica informazione riguarda l'intervallo di valori che contiene il vero prezzo di riserva. L'effettivo prezzo di riserva sarà dunque da considerare una variabile casuale, con una distribuzione di probabilità ignota al ricercatore. Come vedremo nei prossimi paragrafi, ciò può creare notevoli problemi nella determinazione delle stime.

5.2.1. Modelli Dicotomici Semplici

Il modello DC-CVM è stato interpretato da Hanemann (1984, 1989) come un modello di scelta probabilistica strutturalmente simile al modello di McFadden (1976).

Un approccio alternativo è stato successivamente proposto da Cameron (Cameron e James, 1987; Cameron, 1988). I due modelli si differenziano anzitutto per l'oggetto della stima. Nell'approccio di Hanemann il prezzo di riserva individuale viene considerato

come una variabile nascosta; la probabilità che la cifra proposta venga accettata o respinta dipende dal differenziale di utilità fra i due stati in cui sia disponibile o meno il bene pubblico, e viene stimata con un modello probabilistico di scelta binaria. Nell'approccio di Cameron il prezzo di riserva individuale viene invece interpretato come una variabile censurata, della quale si rileva solo che è superiore o inferiore ad una certa cifra, ma non il suo valore esatto; la probabilità di ricevere una risposta positiva o negativa dipende dal differenziale tra prezzo di riserva individuale e cifra proposta, che può essere stimato direttamente con un modello di regressione censurato.

I due approcci determinano due diverse strutture del modello DC-CVM, sia in termini microeconomici, che in termini econometrici. Dal punto di vista microeconomico, McConnell (1990) osserva che il modello di Hanemann stima il differenziale delle funzioni di utilità dell'individuo, mentre quello di Cameron stima il differenziale delle funzioni di spesa. McConnell dimostra inoltre che se, e solo se, le funzioni fossero lineari o se non esistesse alcuna componente stocastica, allora i due modelli sarebbero perfettamente duali: nel senso che, manipolando opportunamente le due funzioni di variazione, si potrebbero derivare le stesse funzioni di domanda per il bene pubblico. Se invece queste ipotesi restrittive non sussistono, allora i due modelli non sono duali, e danno risultati diversi: dunque quale modello scegliere? McConnell non dà un giudizio definitivo: poichè entrambi i modelli sono microfondati, la scelta tra i due modelli è lasciata al giudizio soggettivo del ricercatore circa quale dei due interpreti nel modo più convincente lo schema DC-CVM. Passiamo ora all'analisi della struttura econometrica dei due modelli.

Introduciamo la variabile j per rappresentare il bene pubblico, ponendo $j = 1$ nel caso in cui l'individuo possa disporre del bene (e risponda dunque affermativamente alla richiesta di pagamento) e $j = 0$ nel caso contrario.

Nel modello di McFadden (1976) ripreso da Hanemann si considera una funzione di utilità stocastica semi-indiretta, il cui argomento è costituito dal bene pubblico oltre che dal reddito monetario. Introduciamo una variabile dicotomica che assume il valore 1 nel caso che il bene venga fornito, ed entri quindi nelle possibilità di consumo dell'individuo, e che assume valore 0 nel caso opposto. Indichiamo inoltre con y il reddito dell'individuo, e con s un vettore di caratteristiche socio-economiche che possono influenzare le preferenze individuali. La funzione di utilità indiretta dell'individuo avrà dunque la seguente forma:

$$(1) u_1 \equiv u(1, y; s)$$

$$(2) u_0 \equiv u(0, y; s),$$

dove (1) e (2) rappresentano l'utilità dell'individuo nel caso in cui, rispettivamente, il bene pubblico viene fornito o meno. Abbiamo detto che la funzione di utilità viene in questo modello considerata come una variabile stocastica: si presume cioè che ogni individuo, dato un certo livello di reddito e determinate caratteristiche socio-economiche, tragga dal consumo del bene pubblico un certo livello di utilità, con possibili variazioni in positivo o negativo prodotte da una componente di errore stocastico e_j . L'utilità individuale nei due stati è dunque considerata una variabile casuale con media, diciamo, $v_j(j, y; s)$; e si può scrivere

$$(3) u(j, y; s) = v(j, y; s) + e_j \quad \text{con } j = 0, 1.$$

Nel momento in cui per la fornitura del bene pubblico viene richiesto un dato ammontare t di denaro, l'intervistato risponderà di essere disposto a pagare la cifra proposta t se e solo se

$$(4) v_1[(y-t), s] \geq v_0(y, s) + e_0 - e_1 ;$$

l'individuo confronta cioè la propria funzione di utilità in ciascuno dei due stati possibili.

Per il ricercatore la risposta affermativa alla cifra proposta t è una variabile casuale la cui distribuzione di probabilità è data dalla seguente espressione:

$$(5) P_1 = \Pr[v(1, y-t; s) + e_1 \geq v(0, y; s) + e_0] ;$$

ovviamente la probabilità di avere una risposta negativa sarà pari a $P_0 = 1 - P_1$.

La (5) può essere specificata rispetto alla variazione di utilità: ponendo $\eta = e_0 - e_1$, e $\Delta v \equiv v(1, y-t; s) - v(0, y; s)$, possiamo riscrivere la (5) nel modo seguente:

$$(6) P_1 = \Pr(\eta \leq \Delta v) .$$

La stima di massima verosimiglianza rappresenta la massima probabilità di ottenere i dati campionari dalla popolazione considerata: assumendo che le osservazioni siano indipendenti tra loro, la probabilità di osservare il campione sarà data dal prodotto delle probabilità di ogni variabile di assumere la caratteristica osservata. Per semplificare i calcoli si usa generalmente la trasformazione logaritmica, funzione di log-likelihood, che nel nostro caso assume la seguente forma:

$$(7) \text{Log}L = \sum_{i=1}^n I_i \log F_\eta(\Delta v_i) + (1 - I_i) \log[1 - F_\eta(\Delta v_i)] ,$$

dove I_i è una variabile dicotomica che assume valore 1 se l'individuo risponde positivamente alla cifra proposta, e valore zero in caso contrario.

La (7) è un'espressione generale dalla quale, attraverso la specificazione delle funzioni F e Δv , si determina il modello econometrico da stimare. Δv dipende dalla specificazione della funzione di utilità individuale; a questo proposito è interessante notare che alcuni studi (p.es. Bishop e Heberlein (1979)) hanno impiegato delle specificazioni di Δv non ricavabili come differenza di due funzioni di utilità, e che pertanto, come indicato da Hanemann (1984), si devono considerare non fondate da un punto di vista microeconomico. F_η è invece la funzione di distribuzione cumulativa di η , che dipende dalla distribuzione di e_0 e di e_1 . Per esempio, se gli errori sono distribuiti normalmente, $F_\eta(\cdot)$ sarà una distribuzione normale, e la (6) assumerà la seguente forma:

$$(8) P_1 = \Phi(\eta \leq \Delta v) ,$$

e, standardizzando,

$$(9) P_1 = \Phi(z \leq \Delta v / \sigma) ,$$

e la stima della (7) potrà essere effettuata con il modello *probit*; se $F_\eta(\cdot)$ fosse una logistica, il modello sarebbe invece il *logit*, la cui struttura è per il resto del tutto analoga. La (9) mostra chiaramente che in questo modello la stima dei parametri della funzione Δv non può essere isolata dalla stima della varianza della distribuzione $F_\eta(\cdot)$.

Consideriamo adesso l'approccio di Cameron. Il modello econometrico è il seguente:

$$(10) Y = x'\beta + \varepsilon$$

nel quale Y rappresenta l'effettiva disponibilità a pagare individuale, che potrà dipendere da certe caratteristiche individuali, quali il reddito, e altre variabili socioeconomiche contenute nel vettore x ; il termine ε rappresenta l'errore stocastico, con media zero ed una certa varianza v^2 . In questo modello Y viene interpretata come una variabile continua censurata, per la quale non si dispongono di dati completi, ma si sa se è superiore, o inferiore, ad una determinata cifra. Denotiamo con P_1 la probabilità che il prezzo di riserva Y per un certo individuo sia superiore alla cifra proposta t , e con P_0 la probabilità complementare:

$$(11) \begin{aligned} P_1 &= \Pr(Y > t) \\ &= \Pr(x'\beta + \varepsilon > t) \\ &= \Pr(\varepsilon > t - x'\beta) \end{aligned}$$

e, standardizzando,

$$(12) P_1 = \Pr(z > (t - x'\beta) / v).$$

Se ε per esempio si distribuisce secondo una normale, con media zero e varianza σ^2 , la (11) assume la forma:

$$(13) \begin{aligned} P_1 &= 1 - \Phi((t - x'\beta) / \sigma) \\ P_0 &= \Phi((t - x'\beta) / \sigma) \end{aligned}$$

Per un determinato campione di n osservazioni indipendenti, la funzione di log-likelihood assume la forma seguente:

$$(14) \text{Log}L = \sum_{i=1}^n I_i \log(1 - \Phi)((t - x'_i\beta) / \sigma) + (1 - I_i) \log[\Phi((t - x'_i\beta) / \sigma)],$$

dove I_i è una variabile dicotomica che assume valore zero nel caso di risposta negativa, e uno nel caso di risposta positiva. Con tecniche di ottimizzazione non lineare si possono ottenere stime separate per il vettore dei coefficienti β e per la deviazione standard σ . Dalle due espressioni (9) e (11) risulta in modo evidente la differenza tra il modello di utilità stocastica di Hanemann e il modello censurato di Cameron: in quest'ultimo la deviazione standard è identificata separatamente rispetto agli altri parametri, e così si ottengono stime corrette sia per i valori dei parametri che per i loro standard errors. La stima del prezzo di riserva medio $E(Y)$ si ricava dunque direttamente dall'equazione (9),

$$(15) E(Y) = \bar{x}'\beta^*$$

moltiplicando i coefficienti stimati per i valori medi delle variabili indipendenti del modello: cosa che il modello di Hanemann, ed in generale i modelli a scelta discreta (logit o probit) non consentono di fare direttamente, dato che le stime dei parametri includono il valore del parametro di scala della distribuzione.

Nel caso di un modello con funzione di utilità implicita lineare o loglineare e distribuzione del prezzo di riserva della popolazione di tipo simmetrico, il calcolo della disponibilità media (e mediana) risulta comunque piuttosto semplice anche nel modello di

Hanemann. Se si considera una sola variabile esplicativa, si dimostra infatti (Hanemann, 1989; Cooper e Loomis, 1992; Cooper, 1994) che la disponibilità a pagare media è:

$$(16) E(Y) = \frac{\alpha}{-\delta}$$

nel caso di funzione di utilità lineare, dove α è la costante, e δ il coefficiente della variabile esplicativa inserita nel modello;

$$(17) E(Y) = \frac{\ln(1 + \exp^{\alpha})}{-\delta}$$

nel caso di funzione di utilità log-lineare.

Peraltro, negli stessi lavori di Cameron e James (1987), e Cameron (1988), (1991), si indica il modo per trasformare i parametri ottenuti con i modelli qualitativi, nei parametri del modello censurato: se la variabile esplicativa è la cifra proposta t , si può anche in questo caso identificare separatamente la deviazione standard, e ottenere con semplici calcoli i valori delle stime puntuali dei parametri del modello censurato. Non altrettanto immediato è il calcolo degli standard errors dei coefficienti trasformati: come vedremo più avanti, questi sono necessari per il calcolo degli intervalli di confidenza intorno alla stima del prezzo di riserva.

Cooper (1994) generalizza il modello lineare al caso di più variabili esplicative: in questo caso la disponibilità a pagare media sarà pari a

$$(18) E(Y) = \frac{\sum \bar{x}_i \gamma_i}{-\delta};$$

il numeratore, definito da Cooper come *grand mean*, è la sommatoria dei valori medi delle variabili esplicative moltiplicate per i rispettivi coefficienti, e δ è il coefficiente della variabile t .

Con specificazioni meno restrittive, nel modello di Hanemann la disponibilità a pagare media deve essere calcolata integrando la funzione di probabilità cumulativa

$$(19) E(Y) = \int_A [1 - G(t)] dt,$$

dove $G(t)$ è la distribuzione della disponibilità a pagare, i cui parametri vengono ottenuti dalla (7). Purtroppo, questo metodo per la determinazione delle misure di tendenza centrale è fortemente influenzato dalle ipotesi del ricercatore circa il supporto della distribuzione $G(t)$. Quando il dominio di integrazione coincide con R^+ , come per esempio in Hanemann (1989), si ottengono spesso stime troppo elevate per essere realistiche. Per cercare di risolvere questo problema, alcuni ricercatori ricorrono ad un troncamento della coda destra della distribuzione della disponibilità a pagare. Il punto di troncamento è arbitrario: per esempio, Bishop e Heberlein fissano la chiusura del supporto in corrispondenza della cifra più alta fra quelle proposte al campione degli intervistati, ma possono essere adottati altri criteri. Si raccomanda comunque l'impiego della tecnica di normalizzazione proposta da Boyle, Welsh e Bishop (1988). Tali autori hanno infatti dimostrato che le formulazioni originariamente utilizzate nel computo della media troncata del tipo della (20)

$$(20) E(WTP) = \int_0^{X_{max}} [1 - F(x)] dx = X_{max} - \int_0^{X_{max}} F(x) dx,$$

nella quale x è la variabile casuale "disponibilità a pagare", violano le proprietà statistiche della funzione di ripartizione provocando una sistematica sovrastima della media. In luogo della (29) viene perciò proposta la seguente "media troncata normalizzata"

$$(21) E(WTP) = \int_b^{X_{max}} [1 - F(x)/F(X_{max})] dx .$$

Si può notare che la differenza fra i risultati ottenuti con la (20) e quelli ottenuti con la (21) diviene via via minore al crescere di $F(X_{max})$.

I risultati ottenuti da Leòn (1995) sembrano inoltre evidenziare che il ricorso alla media troncata normalizzata con i modelli dicotomici con follow-up permette di ottenere degli intervalli di confidenza della stima della WTP più ristretti.

5.2.2. Modelli dicotomici doppi

Come visto precedentemente, i modelli che prevedono la formulazione di due domande sulla cifra da "prendere o lasciare" sono chiamati modelli dicotomici doppi. Se la risposta alla prima domanda è stata positiva, cioè l'intervistato si è dichiarato disposto a pagare la cifra che gli è stata proposta, gli viene successivamente proposta un'altra cifra, superiore alla prima; se la risposta è stata negativa, la seconda cifra sarà inferiore. In entrambi i casi, l'intervistato risponderà SI o NO a seconda del suo vero prezzo di riserva. Questo metodo di rilevazione dà luogo a quattro possibili risultati, cui sono associate le seguenti probabilità:

$$(22) \begin{aligned} \Pr(si, si) &= \Pr(E \geq t_i^a \geq t_i) = 1 - F(t_i^a) \\ \Pr(si, no) &= \Pr(t_i \leq E \leq t_i^a) = F(t_i^a) - F(t_i) \\ \Pr(no, si) &= \Pr(t_i^b \leq E \leq t_i) = F(t_i) - F(t_i^b) \\ \Pr(no, no) &= \Pr(E \leq t_i^b \leq t_i) = F(t_i^b) \end{aligned}$$

La funzione log-likelihood per questo modello è la seguente:

$$(23) \begin{aligned} \text{Log}L &= \sum_{i=1}^n I_i I_i^a \log[F(t_i^a)] + I_i (1 - I_i^a) \log[F(t_i^a) - F(t_i)] + \\ &I_i^b (1 - I_i) \log[F(t_i) - F(t_i^b)] + (1 - I_i) (1 - I_i^b) \log[F(t_i^b)] \end{aligned}$$

Il simbolo t_i indica il prezzo proposto nella prima domanda, t_i^a : il prezzo proposto nel follow-up qualora la prima risposta sia stata affermativa, t_i^b il prezzo proposto nel follow-up qualora la prima risposta sia stata negativa; le variabili dicotomiche I_i, I_i^a, I_i^b assumono valore uno o zero a seconda che la risposta alla domanda iniziale e al corrispondente follow-up sia stata positiva o negativa.

Data la stessa selezione delle somme offerte inizialmente, Hanemann, Loomis e Kanninen (Hanemann et al. (1991)) hanno dimostrato che il modello DC-CVM doppio è asintoticamente più efficiente del modello semplice, ed i risultati empirici degli stessi autori e di Leòn (1995) sembrano confermare questa proprietà anche per i campioni finiti. D'altra parte si è potuto spesso notare che il modello doppio dà risultati sostanzialmente

diversi rispetto al modello semplice applicato ai dati delle prime risposte. Una possibile interpretazione è che il modello doppio contiene maggiore informazione rispetto al modello semplice, e quindi dia stime meno distorte: Kanninen (1995) calcola la distorsione delle stime su un piccolo campione (100 unità), e trova che questa è maggiore nel modello semplice che in quello doppio. Questo risultato viene attribuito al fatto che con il follow-up si abbia un guadagno di efficienza, in quanto si corregge un eventuale errore nella scelta delle cifre proposte: il problema della determinazione ottimale delle cifre proposte è effettivamente di rilevanza cruciale, e sarà trattato più diffusamente nel prossimo paragrafo. Un'altra interpretazione è che le risposte date alle due domande successive provengano da distribuzioni diverse, per quanto correlate. Herriges e Shogren (1996) attribuiscono la differenza tra le due distribuzioni generate dalle due domande al fatto che la seconda subirebbe un effetto di "ancoraggio" alla prima cifra offerta: ciò determinerebbe quindi una minore attendibilità dell'informazione ottenuta con la seconda domanda. Altri effetti di "disturbo" sui dati del modello doppio sono analizzati da Alberini, Kanninen e Carson (1997). Cameron e Quiggin (1994) sostengono invece che le due risposte vengono determinate da processi decisionali in parte diversi: per questo motivo essi individuano nel probit bivariato il modello econometrico più adatto per il modello di scelta dicotomica con follow-up: si stimano cioè due modelli distinti anche se fra loro correlati, che producono due distinte stime per la disponibilità a pagare media, una basata sui dati delle prime risposte, l'altra sui dati del follow-up. Il ricercatore, sulla base del modello decisionale che ritiene valido per il campione, potrà utilizzare una media ponderata delle due stime. Per un confronto tra il modello univariato ed il modello bivariato si veda Alberini (1995).

5.2.3. *La determinazione ottimale delle cifre proposte*

Nell'implementazione del DC-CVM la scelta dei "prezzi" da proporre agli intervistati è di importanza cruciale. Aggiungere un elemento al vettore dei prezzi da proporre da una parte aumenta l'informazione campionaria sulla distribuzione; ma dall'altra riduce l'efficienza delle stime dei parametri del modello, e questa può essere migliorata solo incrementando la numerosità campionaria. La scelta del vettore dei prezzi è ottimale quando informazione ed efficienza sono massimizzate per una data numerosità campionaria. Gli elementi da definire sono due: la dimensione del vettore di prezzi da proporre, e l'importo di ciascuno di essi.

Cooper e Loomis (1992), e Desvousges et al. (1993) dimostrano che la stima della disponibilità a pagare media è estremamente sensibile ad una riduzione nel numero di cifre proposte: eliminare dal vettore di prezzi proposti anche solo una cifra può modificare pesantemente i risultati dell'analisi. Un metodo per determinare in modo ottimale il livello dei prezzi da proporre viene indicato da Cooper (1993), e da Alberini e Carson (1993). Si suddivide la distribuzione di probabilità ipotizzata in un numero di quantili di equiprobabilità pari al numero di cifre che si intendono proporre; i prezzi saranno i valori dei quantili. L'aspetto più originale del lavoro di Cooper è che data la numerosità campionaria, permette di individuare simultaneamente i diversi ammontari e la numerosità dei sottocampioni a questi associati: un'ulteriore variabile, la numerosità dei sottocampioni cui viene proposto lo stesso prezzo,⁴ viene quindi inserita nella

⁴ In genere le cifre selezionate vengono distribuite uniformemente all'interno del campione.

progettazione dell'indagine di CV. Per quanto riguarda la scelta del numero di cifre iniziali⁵ da proporre, Cooper indica, per un campione di 500 unità, un numero di prezzi che varia dai 10-12 se si ipotizza una distribuzione simmetrica, fino a 40 per distribuzioni asimmetriche. Alberini (1994) effettua delle simulazioni Monte Carlo per analizzare i risultati ottenuti quando la distribuzione ipotizzata è diversa da quella vera: i risultati del suo esperimento contrastano con le indicazioni di Cooper, in quanto suggeriscono, anche nel caso di distribuzioni asimmetriche, un numero ottimale non superiore alle 10 unità nel caso di DC-CVM semplice, e compreso tra le 4 e le 6 unità, a seconda dell'ampiezza del campione, se si impiega il modello doppio. Quest'ultimo è in effetti consigliato dalla stessa Alberini rispetto al modello semplice, in quanto permette di ottenere maggiore informazione circa la forma funzionale della vera distribuzione: permette perciò di evitare con maggiore probabilità un errore nella specificazione del modello, che, come mostrato nello stesso lavoro, non viene facilmente individuato dai comuni test di specificazione. Si rinvia alla stessa Alberini (1994) per l'analisi delle conseguenze di una erronea specificazione del modello sulle stime.

Anche per la scelta dei follow-up alla regola *ad hoc* prevalente (che consiste nel fissarli nella misura del doppio o della metà della prima cifra proposta) si possono sostituire metodologie più sofisticate: si vedano in particolare Kanninen (1993) e Alberini(1994); e Nyquist (1992), per i cosiddetti modelli sequenziali, che prevedono un processo iterativo di aggiustamento del valore del follow-up dopo l'effettuazione di un certo numero di interviste.

5.3 Il metodo della " lista "

Come descritto nel paragrafo 4.4, il metodo della lista prevede che si presenti all'intervistato una tabella contenente diverse cifre, fra le quali dovrà segnalare la cifra esatta che rappresenta la sua massima disponibilità a pagare; oppure l'intervallo in cui ricade la sua massima disponibilità a pagare. In un caso o nell'altro, l'informazione ottenuta si riferisce comunque ad un intervallo nel quale è contenuto il vero prezzo di riserva: indicando con t_k , dove $t = 1, 2, \dots, n$, il valore dichiarato dall'intervistato in una lista ordinata in senso crescente, il dato campionario è il seguente:

$$(24) t_k \leq Y \leq t_{k+1}.$$

In questo caso la variabile Y è una variabile *raggruppata*. Un artificio spesso impiegato per "trasformare" la Y in variabile continua consiste nell'identificarla con il valore intermedio dell'intervallo fra t_k e t_{k+1} . Con questa trasformazione la disponibilità a pagare media può essere determinata calcolando la media aritmetica dei valori intermedi di ciascun intervallo ed analogo procedimento può essere usato per determinare la mediana. L'analisi di regressione per individuare le variabili che influenzano la distribuzione della disponibilità a pagare consiste in questo caso in una semplice stima OLS. Come dimostrato da Cameron (1987), se la distribuzione della disponibilità a pagare è normale, il metodo OLS produce stime relativamente accettabili, anche per intervalli ampi. Se invece si ritiene non troppo robusta l'ipotesi di normalità della

⁵ Si tenga presente che questo numero si moltiplica per tre quando viene impiegato il modello con follow-up. Ricordiamo che per le diverse somme offerte, il follow-up dovrà prevedere per ciascuna di queste sia un ammontare superiore, nel caso di risposta affermativa alla prima domanda, sia una cifra inferiore per il caso di una prima risposta negativa.

distribuzione, allora è senz'altro consigliabile fare ricorso a specifiche tecniche di stima di massima verosimiglianza nelle quali la variabile dipendente è espressa in intervalli.

Il metodo proposto da Cameron e Huppert (1989) ha una struttura simile a quella del modello a scelta dicotomica censurato. Il modello per il vero valore, non osservato, del prezzo di riserva dell'individuo j è il seguente:

$$(25) Y_j = x_j' \beta + e_j,$$

nel quale gli errori e_j sono distribuiti secondo una certa funzione G . Date però le proprietà degli stimatori OLS nel caso di distribuzione normale, è utile effettuare la trattazione con una distribuzione asimmetrica: seguendo l'esempio di Cameron e Huppert (1989, 1991) consideriamo il caso in cui questa sia una Lognormale.

Sotto tale ipotesi infatti il modello che spiega la disponibilità a pagare dei rispondenti è il seguente:

$$(26) \log Y_j = x_j' \beta + u_j$$

Standardizzando, otteniamo la seguente probabilità che la disponibilità a pagare individuale ricada all'interno del k -mo intervallo:

$$\begin{aligned} & \Pr(Y_j \subseteq t_{k,j}, t_{k+1,j}) = \\ (27) & \Pr\left[\log(t_{k,j} - x_j' \beta) / \sigma < z_j < (\log t_{k+1,j} - x_j' \beta) / \sigma\right]. \\ & = \Phi\left[\log(t_{k+1,j} - x_j' \beta) / \sigma\right] - \Phi\left[\log(t_{k,j} - x_j' \beta) / \sigma\right] \end{aligned}$$

La funzione di log-verosimiglianza è la seguente:

$$(28) \log L = \sum_{j=1}^n \left[\Phi\left[\log(t_{k+1,j} - x_j' \beta) / \sigma\right] - \Phi\left[\log(t_{k,j} - x_j' \beta) / \sigma\right] \right].$$

Moltiplicando i valori dei coefficienti stimati β per i valori medi delle variabili esplicative, utilizzando cioè l'equazione (26), si può calcolare il logaritmo della disponibilità a pagare media ⁶.

5.4. Stima degli intervalli di confidenza

Una volta che si è stimata una misura di tendenza centrale, per esempio la disponibilità a pagare media, occorrerebbe calcolare la probabilità che la vera media sia inclusa in un determinato intervallo. Tuttavia, solo di recente si è incominciato ad effettuare la stima degli intervalli di confidenza, in quanto i metodi normalmente impiegati per la determinazione delle misure di tendenza centrale non permettevano di determinare facilmente la loro varianza. Una possibile soluzione al problema è data dall'impiego di tecniche di simulazione: in particolare, sono state adottate delle varianti del noto metodo *bootstrap* (Efron e Tibshirani (1986)): il metodo *jackknife* è stato applicato su dati DC-CVM da McLeod e Bergland (1989); Park et al. (1991) hanno invece impiegato un'altra variante, proposta da Krinsky e Robb (1986).

⁶ Il software LIMDEP fornisce la routine "GROUPED DATA", che permette di effettuare direttamente la precedente stima rispetto ai parametri β e σ .

Alternativamente ai metodi di simulazione, che richiedono procedure lunghe e dispendiose, Cameron (1991) ha proposto un metodo per calcolare analiticamente gli intervalli di confidenza. Se si impiega il suo modello censurato per la stima della disponibilità a pagare media, allora il calcolo degli intervalli di confidenza per la variabile dipendente può essere effettuato direttamente impiegando i coefficienti stimati, e la loro matrice di varianze e covarianze. Se invece si impiega il modello logit o probit, come accennato in precedenza si possono facilmente trasformare i parametri del modello qualitativo e ottenere quelli del modello censurato; con procedimenti più complicati si calcola la matrice di varianze e covarianze dei nuovi parametri come descritto in Cameron (1991). La stima dell'intervallo di confidenza intorno al valore medio della disponibilità a pagare è la seguente:

$$(28) CI_{95}[E(Y)] = \bar{x}'\beta \pm t_{.25} \cdot \sqrt{\bar{x}'\Sigma_{\beta}\bar{x}},$$

dove Σ_{β} è la matrice di varianze e covarianze dei coefficienti, e \bar{x} è il vettore dei valori medi delle variabili esplicative. La validità dei diversi metodi di stima degli intervalli di confidenza cui abbiamo fatto cenno, e cioè il metodo bootstrap, jackknife, Krinsky e Robb, e Cameron, è stata studiata da Cooper (1994) con dati Monte Carlo: i risultati mostrano che i metodi considerati sono tutti soddisfacenti, con piccole variazioni nelle graduatorie di precisione a seconda dell'ampiezza del campione considerato e della forma funzionale della distribuzione.

5.5. Metodi di stima non parametrici

L'approccio non parametrico alla stima su dati DC-CVM è stato introdotto da Kristrom (1990); fra i contributi successivi segnaliamo Carson et al. (1994), McFadden (1994), Haab e McConnell (1996). Il principale vantaggio dei metodi non parametrici è costituito dal fatto che non occorre fare assunzioni circa la distribuzione della disponibilità a pagare.

In generale, la funzione di log-likelihood basata su dati DC-CVM è la seguente:

$$(29) \text{Log}L = \sum_{i=1}^m S_i \log[1 - G(t_i)] + N_i \log[G(t_i)],$$

nella quale il simbolo m indica il numero di cifre proposte, indicate con t , S_i il numero di risposte positive in corrispondenza dell'ammontare t_i e N_i il numero di risposte negative per lo stesso ammontare. I metodi di stima non parametrici permettono di derivare da queste frequenze una stima della media della funzione di distribuzione $G(t)$, senza che sia necessario definire a priori la sua forma funzionale. Alternativamente, la (1) può essere scritta come:

$$(30) \text{Log}L = \sum_{i=1}^m S_i \log\left[1 - \sum_{i=1}^j p_i\right] + N_i \log\left[\sum_{i=1}^j p_i\right],$$

dove i p_i rappresentano la probabilità che il prezzo di riserva individuale cada nell'intervallo (t_{i-1}, t_i) .

Uno stimatore molto semplice per p_i è costituito dalla proporzione di risposte negative associate alla cifra t_i : in simboli

$$(31) \hat{\pi}_i = N_i / (S_i + N_i)$$

Se la sequenza delle proporzioni è (debolmente) crescente, allora la probabilità che la disponibilità a pagare cada nel generico intervallo $[t_{i-1}, t_i]$ è positiva ed uguale alla differenza delle proporzioni riferibili ai due estremi dell'intervallo.⁷

A priori non è però escludibile (ed in pratica, con un numero di cifre proposte relativamente elevato, non infrequente) che la suddetta sequenza risulti decrescente in alcune sue parti. In questo caso lo stimatore $\hat{\pi}_i$ deve essere sostituito dal seguente algoritmo, conosciuto in letteratura con la sigla PAVA (*pool-adjacent-violator algorithm*) (Robertson, Wright e Dykstra (1988)):

$$(32) \bar{\pi} = N_i + N_{i+1} / (S_i + S_{i+1} + N_i + N_{i+1}).$$

Quest'operazione si impone per evitare che lo stimatore $\hat{\pi}_i$ possa presentare un segno negativo e va ripetuta finché l'intera sequenza delle proporzioni diviene non decrescente.

Haab e McConnell (1996) descrivono le operazioni di stima non parametrica nel seguente modo:

- a) Per $i = 1, \dots, n-1$, si calcola la proporzione $\pi_i = N_i / (S_i + N_i)$;
- b) Partendo da $i = 1$, si confronta π_i e π_{i+1} ;
- c) Se $\pi_{i+1} > \pi_i$ il computo delle proporzioni può essere proseguito;
- d) Se $\pi_i \leq \pi_{i+1}$ si unificano i due intervalli adiacenti;
- e) L'operazione va ripetuta fino a che non si ottiene una funzione di distribuzione monotona crescente.
- f) Si calcola la funzione di densità come differenza degli intervalli della funzione di distribuzione.

Per quanto riguarda la stima della disponibilità a pagare media, si può utilizzare lo stimatore Turnbull, con il quale si ottiene una misura di tendenza centrale "conservativa": la *lower bound mean (LBM)*, calcolata come media degli estremi inferiori di ogni intervallo ponderata per la probabilità associata a tale intervallo.

$$(33) E(Y) = 0 \cdot \Pr(0 \leq Y \leq t_1) + t_1 \cdot \Pr(t_1 \leq Y \leq t_2) + \dots + t_m \Pr(t_m \leq Y \leq t_{m+1}),$$

dove t_{m+1} è il limite superiore del supporto della distribuzione: $\Pr(Y \leq t_{m+1}) = 1$.

$$(34) E(Y) = 0(\pi_1 - \pi_0) + t_1(\pi_2 - \pi_1) + \dots + t_{m-1}(\pi_m - \pi_{m-1}) + t_m(1 - \pi_m) = \sum_{i=1}^{m+1} t_{i-1} \pi_i.$$

La stima della LBM ha distribuzione asintotica Normale, con varianza data da:

$$(35) V \left(\sum_{i=1}^{m+1} \pi_i t_{i-1} \right).$$

⁷ Si veda ad esempio Haab e McConnell (1996).

Con la media e la varianza della distribuzione, è possibile costruire degli intervalli di confidenza sulla media campionaria.

E' importante sottolineare che le stime sulla disponibilità a pagare ottenute con procedure non parametriche sono più "grezze" rispetto a quelle ottenute quando viene specificato il modello parametrico. Nel caso di stime non parametriche, la varianza contiene per esempio anche gli elementi sistematici di variabilità che potrebbero essere spiegati dalle variabili esplicative del modello parametrico. Se l'unico scopo dell'analisi è quello di determinare, in un contesto statico, il valore del prezzo medio di riserva, allora la scelta dei metodi nonparametrici può essere opportuna; se però, come normalmente accade, lo scopo dell'analisi è anche quello di individuare alcune relazioni fondamentali tra domanda del bene pubblico e variabili socioeconomiche, a scopo di previsione e inferenza, allora si dovrà optare per i modelli parametrici. In questo caso, i metodi non parametrici potrebbero comunque essere utilmente impiegati per la fase di *pretest*.

6. Conclusioni

La validità del metodo della Valutazione Contingente per la determinazione del valore dei beni pubblici è stato, ed è tuttora, oggetto di dibattito, e talora di accese polemiche, tra gli studiosi del settore. Le critiche riguardano fundamentalmente la presunta incapacità del metodo di produrre dati affidabili, e consistenti con gli assiomi di razionalità della teoria microeconomica (si veda p.es. McFadden (1994)). Allo scopo di dirimere la controversia, la direzione del National Oceanic and Atmospheric Administration (NOAA), USA, aveva istituito nel 1992 una commissione di esperti "eccellenti", presieduta da Kenneth Arrow e Robert Solow: i risultati del loro lavoro sono stati pubblicati in Arrow et al. (1993). Il rapporto conclude con un verdetto favorevole per il metodo della Valutazione Contingente, a condizione che nelle indagini vengano scrupolosamente osservate certe linee guida indicate nel rapporto stesso (da noi riassunte nel *Quadro 1*). Nonostante l'indiscusso prestigio degli autori del rapporto, la questione non sembra però affatto risolta: per un interessante dibattito fra sostenitori e detrattori del metodo si veda Hausman (1993). Alcuni risultati del rapporto Arrow-Solow sono stati sottoposti a verifica empirica da Carson et al. (1996): con un'indagine basata sulle linee guida del rapporto, gli autori dimostrano che i dati ottenuti superano i test di consistenza. I numerosi problemi ancora irrisolti, soprattutto nelle fasi della costruzione del questionario e dell'analisi dei dati, richiedono comunque che i risultati delle analisi di valutazione contingente siano impiegati con molta cautela, anche se a nostro parere non giustificano una totale sfiducia nel metodo.

Quadro 1: Le raccomandazioni del "N.O.A.A. panel" per l'implementazione degli studi di valutazione contingente.

Nel marzo 1989, la superpetroliera Exxon Valdez si infranse nelle scogliere dell'Alaska, riversando in mare alcune decine di milioni di greggio. Il clamore dell'evento spinse il Congresso ad emanare nell'anno successivo l'Oil Pollution Act. Un dispositivo della nuova legge affidava ad un organo del Ministero del Commercio, il National Oceanic and

Atmospheric Administration (NOAA), il compito di predisporre un regolamento che definisse le procedure di valutazione dei danni. In quest'ambito ad un gruppo di esperti guidati dai premi nobel Kenneth Arrow e Robert Solow venne chiesto un pronunciamento circa l'affidabilità della valutazione contingente come strumento di stima dei valori di non uso: pubblicata nel gennaio del 1993, la risposta del "panel" è divenuta rapidamente un punto di riferimento obbligato per le successive applicazioni del metodo della valutazione contingente.⁸

Le principali prescrizioni del rapporto possono essere riassunte riprendendo la distinzione proposta da Pearce e Moran (1994) fra linee guida di ordine generale e linee guida specifiche per la rilevazione del prezzo di riserva.

LINEE GUIDA GENERALI

- Scegliere accuratamente il piano di campionamento.
- Rendere minime le percentuali di mancate risposte.
- Effettuare interviste personali (dirette o telefoniche) e non postali.
- Effettuare un pre-test per valutare l'influenza dell'intervistatore sulle risposte.
- Includere nel rapporto sull'indagine tutte le informazioni utili sul piano di campionamento e sulle percentuali di caduta e di mancata risposta.
- Verificare con un pre-test l'efficacia delle descrizioni sul progetto ambientale.

LINEE GUIDA PER LA RILEVAZIONE DEL PREZZO DI RISERVA

- Rilevare la disponibilità a pagare per godere del bene pubblico piuttosto che la disponibilità ad accettare una compensazione per esserne privati. In generale, adottare criteri che conducano a stime relativamente prudenti.

⁸ Oltre che dai due premi Nobel, del panel facevano parte anche E. Leamer, R. Radner, H. Schuman e P. Portney. Per maggiori dettagli sulla genesi del rapporto NOAA e una breve storia dell'uso della valutazione contingente si veda Portney (1994).

- Usare il formato a scelta dicotomica: la domanda sulla valutazione dovrebbe essere posta come se si trattasse di un referendum, per il quale bisogna votare SI o NO.
- Prevedere anche la possibilità di astensione, includendo in maniera esplicita l'opzione di non risposta.
- Domandare, in seguito ad una risposta SI o NO, le ragioni di tale risposta.
- Descrivere adeguatamente il progetto ambientale in discussione.
- Ricordare che il pagamento per il programma in questione ridurrebbe l'ammontare di denaro disponibile per le altre voci del bilancio personale
- Ricordare l'esistenza di beni sostitutivi direttamente prima dell'effettuazione della domanda sulla valutazione monetaria.
- Verificare l'assenza di trend temporali nelle risposte impiegando diversi campioni indipendenti estratti in diversi periodi di tempo: la presenza di un trend ridurrebbe sostanzialmente la affidabilità dei dati.
- Rilevare, oltre alla variabile relativa alla valutazione, altre informazioni individuali di carattere socio-economico, a scopi di interpretazione e di inferenza.

Riferimenti Bibliografici

- Adamowicz, W.L., Bhardwaj, V. e Macnab, B.** (1993), "Experiments on the Difference between Willingness to Pay and Willingness to Accept", *Land Economics*, 69, pp. 416-27.
- Alberini, A.** (1994), "Optimal Experimental Designs for Applied Work in Environmental Economics, Part 2", *Nota di lavoro Fondazione Eni Enrico Mattei*, 29.94, Milano.
- Alberini, A.** (1995), "Efficiency vs Bias of Willingness to Pay Estimates: Bivariate and Interval-Data Models" *Journal of Environmental Economics and Management*, 29, pp.169-80.
- Alberini, A. e R.T. Carson** (1993), "Choice of Thresholds for Efficient Binary Discrete Choice Estimation, Part 1", *RFF Quality of the Environment Division Discussion Paper QE93-14*, Washington DC.
- Alberini, A., B.J. Kanninen e R.T. Carson** (1997), "Modeling Response Incentive Effects in Dichotomous Choice Contingent Valuation Data", *Land Economics*, forthcoming.
- Arrow, K., R. Solow, P. Portney, R. Radner, e H. Shuman** (1993), "Report of the NOAA Panel on Contingent Valuation", *Federal Register*, 58(10), pp. 4601-14, Washington DC.
- Bishop, R. e T. Heberlein** (1979), "Measuring Values of Extra-market Goods: Are Indirect Measures Biased?", *American Journal of Agricultural Economics*", 61, pp. 926-930.
- Boyle, J.J., M.P. Welsh e R.C. Bishop** (1988), "Validation of empirical measures of welfare change: Comment and extension", *Land Economics*, 64, pp. 94-98.
- Cameron, T. A.** (1987), "The Impact of Grouping Coarseness in Alternative Grouped-data Regression Models", *Journal of Econometrics*, 35, pp. 37-57.
- Cameron, T. A.** (1988), "A New Paradigm for Valuing Non-Market Goods Using Referendum Data: Maximum Likelihood Estimation by Censored Logistic Regression", *Journal of Environmental Economics and Management*, 15, pp. 355-379.
- Cameron, T. A.** (1991), "Interval Estimates of Non-Market Resource Values from Referendum Contingent Valuation Surveys", *Land Economics*, 67, pp. 413-21.
- Cameron, T. A. e M.D. James** (1987), "Efficient Estimation Methods for Closed-Ended Contingent Valuation Surveys", *The Review of Economics and Statistics*, 69, pp. 269-76.

- Cameron, T. A. e D.D. Huppert** (1989), "OLS versus ML Estimation of Non-market Resource Values with Payment Card Interval Data", *Journal of Environmental Economics and Management*, 17, pp. 230-46.
- Cameron, T. A. e J. Quiggin** (1994), "Estimation Using Contingent Valuation Data from a 'Dichotomous Choice with Follow-up' Questionnaire", *Journal of Environmental Economics and Management*, 27, pp. 218-34.
- Carson, R. T., W.M. Hanemann, e R.C. Mitchell** (1986), "Determining the Demand for Public Goods by Simulating Referendums at Different Tax Prices", Dep. Econ. W.P., University of California, San Diego.
- Carson, R.T., L. Wilks e D. Imber** (1994), "Valuing the Preservation of Australia's Kakadu Conservation Zone", *Oxford Economics Papers*, 46, pp. 727-49.
- Cicchitelli, G., A. Herzel e G.E. Montanari** (1992), *Il campionamento statistico*, Il Mulino, Bologna.
- Cooper, J.** (1994), "A Comparison of Approaches to Calculating Confidence Intervals for Benefit Measures from Dichotomous Choice Contingent Valuation Surveys", *Land Economics*, 70, pp. 111-122.
- Cooper, J.** (1993), "Optimal Bid Selection for Dichotomous Choice Contingent Valuation Surveys", *Journal of Environmental Economics and Management*, 24, pp. 25-40.
- Cooper, J. e J. Loomis** (1992), "Sensitivity of Willingness-to-pay to Bid Design in Dichotomous Choice Contingent Valuation Models", *Land Economics*, 68, pp. 221-224.
- Cummings, R., D. Brookshire e W. Schulze** (1986), *Valuing Environmental Goods: An Assessment of the Contingent Valuation Method*, Rowman & Allanheld, Totowa.
- Davis, R.K.** (1964), "The Value of Big Game Hunting in a Private Forest", in *Transactions of the 29th North American Wildlife and Natural Resources Conference*, Wildlife Management Institute, Washington DC.
- Deaton, A. e J. Muellbauer** (1980), *Economics and Consumer Behavior*, Cambridge University Press, Cambridge.
- Desvousges, W.H., F.R. Johnson, R.W. Dunford, S.P. Hudson, K.N. Wilson e K.J. Boyle** (1993), "Measuring Natural Resource Damages with Contingent Valuation: Tests of Validity and Reliability", in **Hausman J.A.** (ed), (1993), *Contingent Valuation: A Critical Assessment*, Cambridge Economics, Inc., Massachusetts.
- Donaldson, C., A.M. Jones, T.J. Mapp, e J.A. Olson** (1997), "Limited Dependent Variables in Willingness to Pay studies: Applications in Health Care", mimeo, University of Aberdeen.

- Efron, B. e R. Tibshirani** (1986), "Bootstrap Methods for Standard Errors, Confidence Intervals and Other Statistical Accuracy", *Statistical Science*, 1, pp. 54-77.
- Freeman, A.M. III** (1983), *Methods for Assessing the Benefits of Environmental Programs*, in Kneese, A. (a cura di), *Handbook of Natural Resources and Energy Economics*, North Holland, Amsterdam.
- Greene, W.H.** (1990), *Econometric Analysis, 2nd ed.*, MacMillan, New York.
- Haab, T.C. and K.E. McConnell** (1996), "Referendum Models and Negative Willingness to Pay: Alternative Solutions", *Journal of Environmental Economics and Management*, forthcoming.
- Hanemann, W.M.** (1984), "Welfare Evaluations in Contingent Valuation Experiments with Discrete Responses", *American Journal of Agricultural Economics*, 66, pp. 332-41.
- Hanemann, W.M.** (1989), "Welfare Evaluations in Contingent Valuation Experiments with Discrete Responses Data: Reply", *American Journal of Agricultural Economics*, 71, pp.1057-61.
- Hanemann, W. M., J. Loomis, e B.J. Kanninen** (1991), "Statistical Efficiency of Double-bounded Dichotomous Choice Contingent Valuation", *American Journal of Agricultural Economics*, 73, pp.1255-63.
- Hausman J.A.** (ed), (1993), *Contingent Valuation: A Critical Assessment*, Cambridge Economics, Inc., Massachusetts.
- Herriges, J.A., e J.F. Shogren** (1994), "Starting Point Bias in Dichotomous Choice Valuation with Follow-up Questioning", *Journal of Environmental Economics and Management*, 30, pp.112-31.
- Howe, C. W., B.J. Lee, B. e L.L. Bennett** (1994), "Design and Analysis of Contingent Valuation Surveys Using the Nested Tobit Model", *Review of Economics and Statistics*, 76, pp.385-389.
- Kahneman, D., J.L. Knetsch, e R.H. Thaler** (1990), "Experimental Tests of the Endowment Effect and the Coase Theorem", *Journal of Political Economy*, 98, pp.1325-48.
- Kanninen, B.J.** (1993), "Optimal Experimental Designs for Double-bounded Dichotomous Choice Contingent Valuation", *Land Economics*, 69, pp.138-146.
- Kanninen, B.J.** (1995), "Bias in Discrete response Contingent Valuation", *Journal of Environmental Economics and Management*, 28, pp.114-25.
- Knetsch, J.L. e J.A. Sinden** (1984), "Willingness to Pay and Compensation Demanded: Experimental Evidence of an Unexpected Disparity in Measures of Value", *Quarterly Journal of Economics*, 99, pp.507-21

- Krinsky, I. e A. Robb** (1986), "Approximating the Statistical Properties of Elasticities", *Review of Economics and Statistics*, 68, pp.715-19.
- Kristrom, B.** (1990), "A Non-Parametric Approach to the Estimation of Welfare Measures in Discrete Response Valuation Studies", *Land Economics*, 66, pp.135-9.
- Leòn, C. J.** (1995), "El metodo dicotomico de valoración contingente: una aplicación a los espacios naturales en Gran Canaria", *Investigaciones Economicas*, 19.
- Maddala, G.S.** (1983), *Limited-dependent and Qualitative Variables in Econometrics*, Cambridge University Press, Cambridge, UK.
- McConnell, K.E.** (1990), "Models for Referendum Data: The Structure of Discrete Choice Models for Contingent Valuation", *Journal of Environmental Economics and Management*, 18, pp.19-34.
- McFadden, D.** (1976), "Quantal Choice Analysis: A Survey", *Ann. Econom. Soc. Measure*, 5, pp.363-90.
- McFadden, D.** (1994), "Contingent Valuation and Social Choice", *American Journal of Agricultural Economics*, 4, pp.689-708.
- McFadden, D. e G. Leonard** (1993), "Issues in the contingent valuation of environmental goods: methodologies for data collection and analysis", in **Hausman J.A.** (ed), (1993), *Contingent Valuation: A Critical Assessment*, Cambridge Economics, Inc., Massachussetts.
- McLeod D. e O. Bergland** (1989), "The Use of Bootstrapping in Contingent Valuation Studies", W.P., Dpt. of Agricultural Economics, Oregon State University, Corvallis.
- Mitchell, R. C. e R.T. Carson** (1989), *Using Surveys to Value Public Goods. The Contingent Valuation Method*, Resources for the Future, Washington.
- Mitchell, R.C. e R.T. Carson** (1981), "An Experiment in Determining Willingness to Pay for National Water Quality Improvements", Resources for the Future, Washington.
- Nyquist, H.** (1992), "Optimal Designs of Discrete Response Experiments in Contingent Valuation Studies", *Review of Economics and Statistics*, 74, pp.559-563.
- Park, T., J. Loomis e M. Creel** (1991), "Confidence Intervals for Evaluating Benefits from Dichotomous Choice Contingent Valuation Studies", *Land Economics*, 67, pp.64-73.
- Pierce, D. e D.Moran** (1994), *The Economic Value of Biodiversity*, Earthscan, London.
- Portney, P.R.** (1994) "The Contingent Valuation Debate: Why Economists Should Care", *Economic Perspectives*, 8, pp.3-18.
- Robertson, T., F.T. Wright e R.L. Dykstra** (1988), *Order Restricted Statistical Inference*, John Wiley and Sons Ltd, New York.

Santagata, W. (1992), “Le preferenze individuali come base di un sistema informativo per l'offerta di beni pubblici. Il problema della loro misurazione”, in **Muraro G.** (ed) (1992), *Economia dell'informazione ed economia pubblica*, Il Mulino, Bologna.