

1 2 13 24 26 8



Università degli Studi di Padova



Facoltà di Scienze Statistiche

Corso di Laurea Specialistica in Scienze Statistiche, Economiche,  
Finanziarie ed Aziendali

Tesi di Laurea:

**Errori di misura in dati di consumo:  
alcune implicazioni per la stima di curve  
di domanda**

Relatore: Prof. Erich Battistin

Laureanda: Alessandra Riccadonna

Anno Accademico 2007 - 2008



*A mamma e papà.  
A Federica, Francesca e Stefano*

*“Due strade trovai nel bosco, io scelsi la meno battuta,  
per questo sono diverso.”*

*Robert Lee Frost*



# Indice

<b>1</b>	<b>Introduzione</b>	<b>13</b>
1.1	Curve di domanda e notazione . . . . .	13
1.2	Specificazione della curva di domanda . . . . .	14
1.3	Stima . . . . .	16
1.4	I dati . . . . .	18
1.5	Obiettivo della Tesi . . . . .	25
<b>2</b>	<b>Stima della curva di domanda</b>	<b>27</b>
2.1	Introduzione . . . . .	27
2.2	Una possibile soluzione . . . . .	28
2.3	Alcune evidenze da un disegno di simulazione . . . . .	31
2.3.1	Il processo generatore dei dati . . . . .	31
2.3.2	Curva di domanda lineare . . . . .	32
2.3.3	Curva di domanda quadratica . . . . .	35
2.4	Alcuni aspetti critici della soluzione discussa . . . . .	39
<b>3</b>	<b>Una soluzione alternativa</b>	<b>41</b>
3.1	Introduzione . . . . .	41
3.2	Stima del modello alternativo con il “metodo IV classico” e il “metodo di Lewbel” . . . . .	42
3.3	Il processo generatore dei dati . . . . .	42
3.3.1	Curva di domanda lineare e quadratica . . . . .	43
3.4	Distorsione indotta da un errore di misura piccolo . . . . .	50

3.4.1	Correzione della distorsione . . . . .	52
<b>4</b>	<b>Conclusione</b>	<b>57</b>
<b>A</b>	<b>Informazioni sulla Banca d'Italia e l'Indagine sui Bilanci delle Famiglie Italiane</b>	<b>59</b>
<b>B</b>	<b>Distorsione delle stime con errore di misura piccolo</b>	<b>61</b>
B.1	Notazione . . . . .	61
B.2	Relazione tra $E_{W_1^* X^*}[W_1^* x]$ e $E_{W_1 X}[W_1 x]$ . . . . .	61
B.3	Distribuzione di $(W_1^*, X^*)$ . . . . .	62
B.4	Distribuzione di $X^*$ . . . . .	62
B.5	Relazione tra $(Y_1^*, Y_2^*)$ e $(Y_1, Y_2)$ . . . . .	63
B.6	Distribuzione di $(W_1, X)$ . . . . .	63
B.7	Distribuzione di $X$ . . . . .	63
B.8	Relazione tra $(W_1, X)$ e $(W_1^*, X^*)$ . . . . .	63
B.8.1	Sviluppo intorno a $\xi_0 \equiv (x + \ln w_1, x + \ln(1 - w_1))$ . .	64
B.8.2	Sostituzione . . . . .	64
B.9	Relazione tra $X$ e $X^*$ . . . . .	64
B.10	Relazione tra $W_1 X$ e $W_1^* X^*$ . . . . .	64
B.11	Soluzione . . . . .	66
<b>C</b>	<b>Stima della distorsione indotta da errore di misura</b>	<b>67</b>
<b>D</b>	<b>Grafici</b>	<b>71</b>
	<b>Bibliografia</b>	<b>83</b>

# Elenco delle figure

1.1	Curva di domanda per cibo per entrambe le tipologie di nuclei famigliari nello studio di Blundell, Duncan e Pendakur (1998). . .	20
1.2	Curva di domanda per cibo per entrambe le tipologie di nuclei famigliari nel mio studio. . . . .	23
1.3	Curva di domanda per cibo in Colombia, dallo studio di Attanasio, Battistin, Mesnard (2007). . . . .	26
2.1	Distribuzione simulata delle stime standardizzate di $\beta_1$ con i metodi "OLS", "IV" e "Lew" in un modello <i>lineare</i> con struttura dell'errore di Lewbel con $\rho = 1$ (10000 simulazioni, $\beta_1 = -0.1155$ ). .	33
2.2	Distribuzione simulata delle stime standardizzate di $\beta_1$ con i metodi "OLS", "IV" e "Lew" in un modello <i>lineare</i> con struttura dell'errore di Lewbel con $\rho = 0.9$ (10000 simulazioni, $\beta_1 = -0.1155$ ). .	33
2.3	Distribuzione simulata delle stime standardizzate di $\beta_1$ con i metodi "OLS", "IV" e "Lew" in un modello <i>lineare</i> con struttura dell'errore di Lewbel con $\rho = 0.75$ (10000 simulazioni, $\beta_1 = -0.1155$ ). .	34
2.4	Distribuzione simulata delle stime standardizzate di $\beta_1$ con i metodi "OLS", "IV" e "Lew" in un modello <i>quadratico</i> con struttura dell'errore di Lewbel con $\rho = 1$ (10000 simulazioni, $\beta_1 = 0.1572$ ). .	36
2.5	Distribuzione simulata delle stime standardizzate di $\beta_1$ con i metodi "OLS", "IV" e "Lew" in un modello <i>quadratico</i> con struttura dell'errore di Lewbel con $\rho = 0.9$ (10000 simulazioni, $\beta_1 = 0.1572$ ). .	36

2.6	Distribuzione simulata delle stime standardizzate di $\beta_1$ con i metodi "OLS", "IV" e "Lew" in un modello <i>quadratico</i> con struttura dell'errore di Lewbel con $\rho = 0.75$ (700 simulazioni, $\beta_1 = 0.1572$ ).	37
2.7	Distribuzione simulata delle stime standardizzate di $\beta_2$ con i metodi "OLS", "IV" e "Lew" in un modello <i>quadratico</i> con struttura dell'errore di Lewbel con $\rho = 1$ (10000 simulazioni, $\beta_1 = -0.0133$ ).	37
2.8	Distribuzione simulata delle stime standardizzate di $\beta_2$ con i metodi "OLS", "IV" e "Lew" in un modello <i>quadratico</i> con struttura dell'errore di Lewbel con $\rho = 0.9$ (10000 simulazioni, $\beta_1 = -0.0133$ ).	38
2.9	Distribuzione simulata delle stime standardizzate di $\beta_2$ con i metodi "OLS", "IV" e "Lew" in un modello <i>quadratico</i> con struttura dell'errore di Lewbel con $\rho = 0.75$ (700 simulazioni, $\beta_1 = -0.0133$ ).	38
3.1	Distribuzione simulata delle stime standardizzate di $\beta_1$ con i metodi "OLS", "IV" e "Lew" in un modello <i>lineare</i> con struttura dell'errore alternativa a quella di Lewbel con $\rho = 1$ (10000 simulazioni, $\beta_1 = -0.1155$ ).	44
3.2	Distribuzione simulata delle stime standardizzate di $\beta_1$ con i metodi "OLS", "IV" e "Lew" in un modello <i>lineare</i> con struttura dell'errore alternativa a quella di Lewbel con $\rho = 0.9$ (10000 simulazioni, $\beta_1 = -0.1155$ ).	44
3.3	Distribuzione simulata delle stime standardizzate di $\beta_1$ con i metodi "OLS", "IV" e "Lew" in un modello <i>lineare</i> con struttura dell'errore alternativa a quella di Lewbel con $\rho = 0.75$ (10000 simulazioni, $\beta_1 = -0.1155$ ).	45
3.4	Distribuzione simulata delle stime standardizzate di $\beta_1$ con i metodi "OLS", "IV" e "Lew" in un modello <i>quadratico</i> con struttura dell'errore alternativa a quella di Lewbel con $\rho = 1$ (10000 simulazioni, $\beta_1 = 0.1572$ ).	45

- 3.5 Distribuzione simulata delle stime standardizzate di  $\beta_1$  con i metodi "OLS", "IV" e "Lew" in un modello *quadratico* con struttura dell'errore alternativa a quella di Lewbel con  $\rho = 0.9$  (10000 simulazioni,  $\beta_1 = 0.1572$ ). . . . . 46
- 3.6 Distribuzione simulata delle stime standardizzate di  $\beta_1$  con i metodi "OLS", "IV" e "Lew" in un modello *quadratico* con struttura dell'errore alternativa a quella di Lewbel con  $\rho = 0.75$  (10000 simulazioni,  $\beta_1 = 0.1572$ ). . . . . 47
- 3.7 Distribuzione simulata delle stime standardizzate di  $\beta_2$  con i metodi "OLS", "IV" e "Lew" in un modello *quadratico* con struttura dell'errore alternativa a quella di Lewbel con  $\rho = 1$  (10000 simulazioni,  $\beta_2 = -0.0133$ ). . . . . 47
- 3.8 Distribuzione simulata delle stime standardizzate di  $\beta_2$  con i metodi "OLS", "IV" e "Lew" in un modello *quadratico* con struttura dell'errore alternativa a quella di Lewbel con  $\rho = 0.9$  (10000 simulazioni,  $\beta_2 = -0.0133$ ). . . . . 48
- 3.9 Distribuzione simulata delle stime standardizzate di  $\beta_2$  con i metodi "OLS", "IV" e "Lew" in un modello *quadratico* con struttura dell'errore alternativa a quella di Lewbel con  $\rho = 0.75$  (10000 simulazioni,  $\beta_2 = -0.0133$ ). . . . . 48
- 3.10 Distribuzione simulata delle stime standardizzate di  $\beta_1$  con il metodo "OLS" e con la regola di aggiustamento della distorsione, in un modello *lineare* con la struttura dell'errore alternativa a quella da Lewbel con  $\rho = 0.8$  (10000 simulazioni,  $\beta_1 = -0.1155$ ). . . . . 53
- 3.11 Distribuzione simulata delle stime standardizzate di  $\beta_1$  con il metodo "OLS" e con la regola di aggiustamento della distorsione, in un modello *lineare* con la struttura dell'errore alternativa a quella da Lewbel con  $\rho = 0.85$  (10000 simulazioni,  $\beta_1 = -0.1155$ ). . . . . 54

3.12	Distribuzione simulata delle stime standardizzate di $\beta_1$ con il metodo "OLS" e con la regola di aggiustamento della distorsione, in un modello <i>lineare</i> con la struttura dell'errore alternativa a quella da Lewbel con $\rho = 0.9$ (10000 simulazioni, $\beta_1 = -0.1155$ ). . . . .	55
3.13	Distribuzione simulata delle stime standardizzate di $\beta_1$ con il metodo "OLS" e con la regola di aggiustamento della distorsione, in un modello <i>lineare</i> con la struttura dell'errore alternativa a quella da Lewbel con $\rho = 0.95$ (10000 simulazioni, $\beta_1 = -0.1155$ ). . . . .	56
D.1	Distribuzione simulata delle stime standardizzate di $\beta_1$ con il metodo "OLS", in un modello <i>quadratico</i> con struttura dell'errore di Lewbel con $\rho = 0.9$ (10000 simulazioni, $\beta_1 = 0.1572$ ). . . . .	71
D.2	Distribuzione simulata delle stime standardizzate di $\beta_1$ con il metodo "IV", in un modello <i>quadratico</i> con struttura dell'errore di Lewbel con $\rho = 0.9$ (10000 simulazioni, $\beta_1 = 0.1572$ ). . . . .	72
D.3	Distribuzione simulata delle stime standardizzate di $\beta_1$ con il metodo "Lew", in un modello <i>quadratico</i> con struttura dell'errore di Lewbel con $\rho = 0.9$ (10000 simulazioni, $\beta_1 = 0.1572$ ). . . . .	72
D.4	Distribuzione simulata delle stime standardizzate di $\beta_2$ con il metodo "OLS", in un modello <i>quadratico</i> con struttura dell'errore di Lewbel con $\rho = 0.9$ (10000 simulazioni, $\beta_1 = -0.0133$ ). . . . .	73
D.5	Distribuzione simulata delle stime standardizzate di $\beta_2$ con il metodo "IV", in un modello <i>quadratico</i> con struttura dell'errore di Lewbel con $\rho = 0.9$ (10000 simulazioni, $\beta_1 = -0.0133$ ). . . . .	73
D.6	Distribuzione simulata delle stime standardizzate di $\beta_2$ con il metodo "Lew", in un modello <i>quadratico</i> con struttura dell'errore di Lewbel con $\rho = 0.9$ (10000 simulazioni, $\beta_1 = -0.0133$ ). . . . .	74
D.7	Distribuzione simulata delle stime standardizzate di $\beta_1$ con il metodo "OLS", in un modello <i>quadratico</i> con struttura dell'errore di Lewbel con $\rho = 0.75$ (10000 simulazioni, $\beta_1 = 0.1572$ ). . . . .	74

- D.8 Distribuzione simulata delle stime standardizzate di  $\beta_1$  con il metodo "IV", in un modello *quadratico* con struttura dell'errore di Lewbel con  $\rho = 0.75$  (10000 simulazioni,  $\beta_1 = 0.1572$ ). . . . . 75
- D.9 Distribuzione simulata delle stime standardizzate di  $\beta_1$  con il metodo "Lew", in un modello *quadratico* con struttura dell'errore di Lewbel con  $\rho = 0.75$  (10000 simulazioni,  $\beta_1 = 0.1572$ ). . . . . 75
- D.10 Distribuzione simulata delle stime standardizzate di  $\beta_2$  con il metodo "OLS", in un modello *quadratico* con struttura dell'errore di Lewbel con  $\rho = 0.75$  (10000 simulazioni,  $\beta_1 = -0.0133$ ). . . . . 76
- D.11 Distribuzione simulata delle stime standardizzate di  $\beta_2$  con il metodo "IV", in un modello *quadratico* con struttura dell'errore di Lewbel con  $\rho = 0.75$  (10000 simulazioni,  $\beta_1 = -0.0133$ ). . . . . 76
- D.12 Distribuzione simulata delle stime standardizzate di  $\beta_2$  con il metodo "Lew", in un modello *quadratico* con struttura dell'errore di Lewbel con  $\rho = 0.75$  (10000 simulazioni,  $\beta_1 = -0.0133$ ). . . . . 77
- D.13 Distribuzione simulata delle stime standardizzate di  $\beta_1$  con il metodo "OLS", in un modello *quadratico* con struttura dell'errore alternativa a quella di Lewbel con  $\rho = 0.9$  (10000 simulazioni,  $\beta_1 = 0.1572$ ). 77
- D.14 Distribuzione simulata delle stime standardizzate di  $\beta_1$  con il metodo "IV", in un modello *quadratico* con struttura dell'errore alternativa a quella di Lewbel con  $\rho = 0.9$  (10000 simulazioni,  $\beta_1 = 0.1572$ ). . 78
- D.15 Distribuzione simulata delle stime standardizzate di  $\beta_1$  con il metodo "Lew", in un modello *quadratico* con struttura dell'errore alternativa a quella di Lewbel con  $\rho = 0.9$  (10000 simulazioni,  $\beta_1 = 0.1572$ ). 78
- D.16 Distribuzione simulata delle stime standardizzate di  $\beta_2$  con il metodo "OLS", in un modello *quadratico* con struttura dell'errore alternativa a quella di Lewbel con  $\rho = 0.9$  (10000 simulazioni,  $\beta_1 = -0.0133$ ). 79
- D.17 Distribuzione simulata delle stime standardizzate di  $\beta_2$  con il metodo "IV", in un modello *quadratico* con struttura dell'errore alternativa a quella di Lewbel con  $\rho = 0.9$  (10000 simulazioni,  $\beta_1 = -0.0133$ ). 79

- D.18 Distribuzione simulata delle stime standardizzate di  $\beta_2$  con il metodo "Lew", in un modello *quadratico* con struttura dell'errore alternativa a quella di Lewbel con  $\rho = 0.9$  (10000 simulazioni,  $\beta_1 = -0.0133$ ). 80
- D.19 Distribuzione simulata delle stime standardizzate di  $\beta_1$  con il metodo "OLS", in un modello *quadratico* con struttura dell'errore alternativa a quella di Lewbel con  $\rho = 0.75$  (10000 simulazioni,  $\beta_1 = 0.1572$ ). 80
- D.20 Distribuzione simulata delle stime standardizzate di  $\beta_1$  con il metodo "IV", in un modello *quadratico* con struttura dell'errore alternativa a quella di Lewbel con  $\rho = 0.75$  (10000 simulazioni,  $\beta_1 = 0.1572$ ). . 81
- D.21 Distribuzione simulata delle stime standardizzate di  $\beta_1$  con il metodo "Lew", in un modello *quadratico* con struttura dell'errore alternativa a quella di Lewbel con  $\rho = 0.75$  (10000 simulazioni,  $\beta_1 = 0.1572$ ). 81
- D.23 Distribuzione simulata delle stime standardizzate di  $\beta_2$  con il metodo "IV", in un modello *quadratico* con struttura dell'errore alternativa a quella di Lewbel con  $\rho = 0.75$  (10000 simulazioni,  $\beta_1 = -0.0133$ ). 82
- D.24 Distribuzione simulata delle stime standardizzate di  $\beta_2$  con il metodo "Lew", in un modello *quadratico* con struttura dell'errore alternativa a quella di Lewbel con  $\rho = 0.75$  (10000 simulazioni,  $\beta_1 = -0.0133$ ). 82

# Elenco delle tabelle

1.1	Statistiche descrittive per <i>budget share</i> (di spesa per cibo) con i dati dell'Indagine sui Bilanci di Banca d'Italia per l'anno 1995. . . . .	20
1.2	Output delle regressioni nel modello <i>lineare</i> (1.2) e in quello <i>quadratico</i> (1.3) per <i>coppie senza figli</i> . I risultati sono ottenuti utilizzando l'opzione " <i>robust</i> ", per ottenere Standard Errors robusti rispetto ad eteroschedasticità. . . . .	21
1.3	Output delle regressioni nel modello <i>lineare</i> (1.2) e in quello <i>quadratico</i> (1.3) per <i>coppie con figli fino ai 14 anni</i> . I risultati sono ottenuti utilizzando l'opzione " <i>robust</i> ", per ottenere Standard Errors robusti rispetto ad eteroschedasticità. . . . .	22
1.4	Effetti marginali della regressione OLS e Beta per la specificazione <i>lineare</i> (1.2) e per quella <i>quadratica</i> (1.3). . . . .	23
2.1	Media della distribuzione simulata di $\beta_1$ al variare di $\rho$ , in un modello <i>lineare</i> con errore di misura del "metodo di Lewbel" (10000 simulazioni). . . . .	34
2.2	Media delle distribuzione simulata di $\beta_1$ e $\beta_2$ al variare di $\rho$ , in un modello <i>quadratico</i> con errore di misura del "metodo di Lewbel". . . . .	39
3.1	Media della distribuzione simulata di $\beta_1$ al variare di $\rho$ , in un modello <i>lineare</i> con struttura dell'errore alternativa a quella di Lewbel (10000 simulazioni). . . . .	46

3.2	Media della distribuzione simulata di $\beta_1$ e $\beta_2$ al variare di $\rho$ in un modello <i>quadratico</i> , con errore di misura alternativo a quello di Lewbel (10000 simulazioni). . . . .	49
3.3	Media della distribuzione simulata di $\beta_1$ al variare di $\rho$ in un modello <i>lineare</i> per dati sporcati da errore di misura e dati "aggiustati", con tipo di errore di misura alternativo a quello di Lewbel (10000 simulazioni). . . . .	52

# Capitolo 1

## Introduzione

### 1.1 Curve di domanda e notazione

La curva di domanda, o *Engel Curve*, descrive come la spesa per l'acquisto di un bene muti in funzione di variazioni della spesa totale (o del reddito disponibile) del consumatore. Questa tesi verterà prevalentemente sullo studio della curva di domanda per cibo. Riporto di seguito la notazione che sarà utilizzata in questo elaborato:

- $y_1^*$ : spesa per cibo;
- $y_2^*$ : spesa residuale (spesa per non cibo);
- $x^* = y_1^* + y_2^*$ : spesa totale;
- $w_1^* = \frac{y_1^*}{x^*}$ : quota di spesa per cibo o *budget share*.

Definisco la curva di domanda di *spesa per cibo* attraverso la seguente relazione tra il *budget share* (ovvero, la quota di spesa per cibo rispetto alla spesa totale) e il logaritmo della spesa totale:

$$w_1^* = \psi[\beta, \ln(x^*)] + u, \quad (1.1)$$

dove  $\beta$  è il parametro che regola la dipendenza del *budget share* dal logaritmo della spesa totale e l'errore del modello  $u$  ha media zero e, per il momento, è assunto indipendente da  $x^*$ . In base alle specificazioni adottate

per la funzione  $\psi$  in (1.1), la curva descritta può assumere più forme. Tipicamente, nella letteratura economica ed econometrica degli ultimi quarant'anni le scelte adottate sono per funzioni di domanda *lineari* o *quadratiche* in  $\ln(x)$  (si veda ad esempio Leser, 1963, o Banks, Blundell e Lewbel, 1997). L'obiettivo principale di questa tesi è la stima di  $\beta$  nel caso in cui la *spesa per cibo*  $y_1^*$ , e di conseguenza la *spesa totale*  $x^*$ , siano osservate con errore. Tale errore può dipendere da risposte imprecise da parte degli intervistati in indagini sul reddito e/o sulla spesa, ormai disponibili su larga scala per quasi ogni paese (ad esempio, nel nostro paese, nell'Indagine sul Reddito delle Famiglie Italiane di Banca d'Italia o in quella sui Bilanci delle Famiglie dell'ISTAT). Uno studio accurato delle conseguenze di errori di misura sulla stima di  $\beta$  ha senso solo a valle dell'aver discusso alcuni aspetti legati alla specificazione del modello, problema sul quale mi soffermerò nella prossima Sezione.

## 1.2 Specificazione della curva di domanda

La relazione descritta in (1.1), nel caso in cui  $\psi(\cdot)$  sia *lineare* o *quadratica*, è soggetta ad una critica ovvia poichè la variabile dipendente è una percentuale e ha quindi un supporto limitato, mentre  $\ln(x)$  ha supporto illimitato. Le origini di tale specificazione vanno ricercate nella letteratura statistica dei primi anni Quaranta (si veda, ad esempio, Working, 1943) e sono prevalentemente legate ad un problema di bontà di adattamento del modello ai dati economici sulla spesa. Molto pragmaticamente, un modello parsimonioso che descriva la domanda di spesa di un certo bene come funzione del reddito o della spesa totale è quello rappresentato in (1.1). Working (1943) rappresenta la relazione tra *spesa per cibo* e *spesa totale* attraverso la seguente specificazione *lineare*:

$$w_1^* = \beta_0 + \beta_1 \ln(x^*) + u, \quad (1.2)$$

utilizzando i risultati di una analisi effettuata alla metà degli anni Trenta maturati da una collaborazione tra la Works Progress Administration, il

Bureau of Labor Statistics, il Bureau of Home Economics e il Central Statistical Board. Questo studio prende in considerazione un campione di più di 300.000 famiglie rappresentativo delle principali tipologie di comunità degli Stati Uniti.<sup>1</sup> Secondo Working (1943) la legge (1.2) si adatta ragionevolmente bene ai dati anche al variare di caratteristiche demografiche (numerosità del nucleo familiare, occupazione dei famigliari, zona dello studio ecc.).

Successivamente all'evidenza fornita da Working (1943) sono state considerate specificazioni della curva di domanda alternative alla (1.2), tra cui:

$$\begin{aligned}w_1^* &= \beta_0 + \beta_1 x^* + u, \\y_1^* &= \beta_0 + \beta_1 x^* + u, \\w_1^* &= \beta_0 + \beta_1 \frac{1}{x^*} + u.\end{aligned}$$

Leser (1963) studia l'adattamento di queste forme della curva di domanda, riprendendo anche quella proposta da Working, ai dati di spesa raccolti per diversi beni in Irlanda nel 1951-52 e raggruppati per numero di componenti famigliari e per reddito pro-capite. I risultati in tale studio confermano che proprio la (1.2) è la specificazione che si adatta meglio ai dati, soprattutto a quelli di *spesa per cibo*. L'evidenza empirica fornita da Working (1943) e Leser (1963) motivano il fatto che l'Equazione (1.2) per la curva engeliana prenda il nome di specificazione alla Working-Leser.

Come detto, fino alla fine degli anni Settanta la motivazione alla specificazione è stata prevalentemente legata alla bontà di adattamento del modello ai dati di spesa. Solo in anni più recenti, Deaton e Muellbauer (1980) hanno stabilito il legame tra la specificazione alla Working-Leser e la teoria economica del consumatore. Una descrizione accurata del comportamento del consumatore richiede infatti la specificazione sia di una curva di domanda, sia degli effetti dei prezzi sulle decisioni del consumatore in modo

---

<sup>1</sup>I risultati di questa analisi sono pubblicati in 47 volumi, 22 emessi dal Bureau of Labor Statistics, altri 22 dal Bureau of Home Economics e 3 dalla National Resources Planning Board.

coerente con la teoria della massimizzazione dell'utilità. Il sistema di domanda (per un bene) alla Working-Leser che tiene anche in considerazione la specificazione di una funzione di utilità, è chiamato AIDS (*Almost Ideal Demand System*).

Studi recenti sul comportamento del consumatore, come quello di Banks, Blundell e Lewbel (1997) basato sui dati Inglesi dalla Family Expenditure Survey per gli anni 1970-1986, hanno verificato che la curva di domanda che si adatta meglio a certi beni è proprio quella che consegue da un sistema AIDS. Per altri beni, invece è meglio la specificazione *quadratica* dello stesso, definita *Quadratic Almost Ideal Demand System* (QUAIDS), che da luogo alla seguente specificazione della curva di domanda:

$$w_1^* = \beta_0 + \beta_1 \ln(x^*) + \beta_2 \ln(x^*)^2 + u. \quad (1.3)$$

L'utilizzo di un modello di grado più elevato è particolarmente adatto nel caso di beni "necessari".<sup>2</sup>

### 1.3 Stima

In questa tesi mi concentrò sulle stime dei parametri in (1.1) quando una delle variabili coinvolte, e automaticamente, per le relazioni che le legano, anche altre, sono affette da *errore di misura*. In quanto segue, verrà utilizzata la notazione specificata nella *Sezione* (1.1) e, convenzionalmente, saranno indicate con "\*" le quantità non affette da errore di misura e senza "\*" quelle affette da errore di misura. Varrà quindi la seguente notazione in presenza di errore di misura:

- $y_1$ : spesa per cibo osservata;

---

<sup>2</sup>Dallo studio di Banks, Blundell e Lewbel (1997) emerge che nel caso della spesa alimentare e del combustibile per auto, la formulazione *lineare* sembra fornire una buona approssimazione dell'andamento della quota di spesa rispetto alle spesa totale, mentre un comportamento non lineare è particolarmente evidente nel caso del vestiario e della spesa per alcolici.

- $y_2$ : spesa residuale osservata;
- $x = y_1 + y_2$ : spesa totale osservata;
- $w_1 = \frac{y_1}{x}$ : quota di spesa per cibo osservata o *budget share* osservato.

Il problema della stima dei parametri in presenza di errore di misura classico nei regressori nel modello *lineare* (1.2) o in quello *quadratico* (1.3) è stato più volte considerato nella letteratura statistico-econometrica (si veda Griliches e Ringstad, 1970). Nella specificazione (1.2) se  $\ln(x^*)$  fosse osservato con errore di misura classico, la stima di  $\beta_1$  sarebbe distorta verso il basso. Nella specificazione (1.3) Griliches e Ringstad (1970) dimostrano che se l'errore in  $\ln(x^*)$  è classico e se l'errore e  $\ln(x^*)$  sono distribuiti normalmente, allora il coefficiente del termine quadratico è più distorto che il coefficiente del termine lineare.

I risultati ottenuti da studi precedenti, come quello accennato sopra, non sono tuttavia utilizzabili poichè in questa tesi è implicato, dalla struttura della curva di domanda (1.1), un errore di misura non classico che comporta alcune complicazioni nelle stime dei parametri che vedremo successivamente.

Considerando la curva di domanda (1.1), le stime dei parametri possono essere ottenute tramite il metodo dei minimi quadrati ordinari (OLS) oppure attraverso il metodo delle variabili strumentali (IV).

In presenza di errore di misura, la stima OLS dei parametri della (1.1) è distorta, con segno della distorsione non ovvio. Le complicazioni rispetto al caso "classico" (in cui non vi è errore) si possono intuire ammettendo che la *spesa per cibo*  $y_1^*$  sia misurata con errore. Tale errore si riflette su  $x^*$  e su  $w_1^*$ , risultando in un errore sia a destra che a sinistra dell'Equazione in (1.1). Inoltre, l'errore nel regressore, ovvero in  $\ln(x^*)$ , e nella variabile dipendente,  $w_1^*$ , non è classico. Nel primo caso perchè la relazione tra la variabile affetta da errore e regressore passa per una relazione logaritmica. Nel secondo caso perchè coinvolge un rapporto tra variabili entrambe misurate con errore.

La presenza di errore di misura nelle variabili all'interno della curva di domanda crea problemi di stima anche con il metodo delle variabili strumentali. Questo metodo è infatti noto funzionare solo nel caso dei modelli lineari con errore di misura classico (si veda la discussione in Bound, Brown e Mathiowetz 2001).

## 1.4 I dati

I risultati empirici di questa tesi sono ottenuti usando l'Indagine sui Bilanci delle Famiglie Italiane (IBFI) del 1995 condotta da Banca d'Italia<sup>3</sup>. Ho scelto di utilizzare una sola occasione dell'indagine per svincolarmi da variazioni temporali dei prezzi dei beni nel sistema di domanda. Relativamente al 1995, ho considerato le informazioni concernenti le seguenti variabili:

- consumi non durevoli ( $cn$ ), per la spesa totale  $x^*$ ;
- consumi alimentari medi mensili ( $jconsal$ ), per la spesa per cibo  $y_1^*$ ;
- reddito disponibile netto ( $y$ ), come variabile strumentale  $z$  necessaria per effettuare stime col metodo delle variabili strumentali;
- numero di figli con meno di 5 anni e dai 5 ai 14 anni, per analizzare due tipologie di famiglie diverse.

In questa tesi mi sono concentrata su due tipologie di famiglie: coppie senza figli e coppie con figli fino ai 14 anni.

Un'analisi simile a quella che si vuole effettuare in questa Sezione della tesi è presente nello studio di Blundell, Duncan e Pendakur (1998) che è basato sui dati della Family Expenditure Survey per l'intervallo temporale 1980-1982 su famiglie con uno o due figli e considera sei categorie di

---

<sup>3</sup>Informazioni aggiuntive su Banca d'Italia e sull'IBFI sono presenti nell'Appendice A.

beni, tra cui cibo. Di seguito, quindi, confronto i risultati di stima ottenuti con i dati dell'IBFI del 1995 con quelli raggiunti da Blundell, Duncan e Pendakur (1998) per *spesa per cibo*. Supponendo di trovarsi in assenza di errore di misura, osservo se i dati utilizzati per entrambi gli studi conducono a conclusioni economiche simili in termini di parametri stimati per la curva di domanda (1.1). Questa analisi avviene in due passi: al primo passo osservo i valori di *budget share* e affronto il problema di quale sia la specificazione più adatta per la funzione  $\psi$  in (1.1). Seguendo l'approccio di Blundell, Duncan e Pendakur (1998), riporto le statistiche descrittive per *budget share* e nei grafici che seguono la stima semiparametrica della relazione tra  $w_1^*$  e  $\ln(x^*)$  separatamente per le due tipologie di famiglie considerate. L'equazione di stima considerata in questa Sezione è:

$$w_1^* = \beta_0 + g[\ln(x^*)] + \beta_3 v_1 + \beta_4 v_2 + u$$

dove  $v_1$  e  $v_2$  sono dummy che rappresentano la presenza o meno di figli tra 0 e 4 anni e figli tra 5 e 14 anni. La stima semiparametrica dell'equazione avviene in due fasi. Poichè vale:

$$w_1^* - \beta_0 - \beta_3 v_1 - \beta_4 v_2 = g[\ln(x^*)] + u,$$

nella prima fase ottengo stime consistenti di  $\beta_0$ ,  $\beta_3$  e  $\beta_4$  dalla regressione dei residui di  $w_1^*$ ,  $v_1$  e  $v_2$  su  $\ln(x^*)$ . Nella seconda fase, utilizzo la precedente espressione per stimare non-parametricamente la funzione  $g(\cdot)$  (si veda, per ulteriori dettagli, Blundell, Duncan, Pendakur, 1998). Al secondo passo, separatamente per le due tipologie famigliari, mostro le stime della (1.1) che conseguono da una specificazione *lineare*. Per completezza, e poichè ciò sarà rilevante per quanto discusso nel capitolo successivo, presento anche i risultati di stima da una specificazione *quadratica*.

Osservando i valori della Tab.(1.1) si può notare come, sebbene i campioni presi in considerazione da Blundell, Duncan e Pendakur (1998) siano differenti per numerosità, 594 per coppie con un figlio e 925 per coppie con due figli, e in parte anche per le caratteristiche della popolazione, i valori

	coppie senza figli	coppie con figli fino ai 14 anni
Media	0.338	0.357
Deviazione Standard	0.125	0.121
Osservazioni	460	207

Tabella 1.1: Statistiche descrittive per *budget share* (di spesa per cibo) con i dati dell'Indagine sui Bilanci di Banca d'Italia per l'anno 1995.

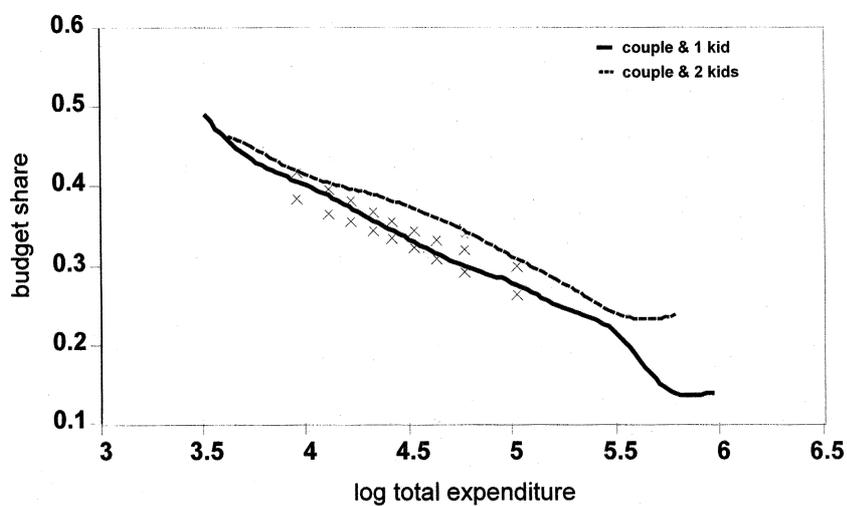


Figura 1.1: Curva di domanda per cibo per entrambe le tipologie di nuclei familiari nello studio di Blundell, Duncan e Pendakur (1998).

Coppie senza figli; curva di domanda lineare.						
N. di oggetti	460					
F(1,458)	94.26					
$R^2$	0.1707					
$\sqrt{MSE}$	0.1144					
$w_1^*$	Coefficiente	Standard Error	t	$P \geq t$	Intervallo al 95%	
$\ln(x^*)$	-0.1155	0.0122	9.45	0	-0.1395	-0.0915
costante	1.5135	0.1248	12.12	0	1.2681	1.7589
Coppie senza figli; curva di domanda quadratica.						
N. di oggetti	460					
F(2,457)	47.41					
$R^2$	0.1718					
$\sqrt{MSE}$	0.1144					
$w_1^*$	Coefficiente	Standard Error	t	$P \geq t$	Intervallo al 95%	
$\ln(x^*)$	0.1572	0.3426	0.43	0.667	-0.5610	0.8755
$\ln(x^*)^2$	-0.0133	0.0166	-0.74	0.458	-0.0484	0.0218
costante	0.1159	1.7585	0.06	0.950	-3.5499	3.7819

Tabella 1.2: Output delle regressioni nel modello *lineare* (1.2) e in quello *quadratico* (1.3) per *coppie senza figli*. I risultati sono ottenuti utilizzando l'opzione "*robust*", per ottenere Standard Errors robusti rispetto ad eteroschedasticità.

Coppie con figli fino ai 14 anni; curva di domanda lineare.						
N. di oggetti	207					
F(3,203)	32.33					
$R^2$	0.3233					
$\sqrt{MSE}$	0.1001					
$w_1^*$	Coefficiente	Standard Error	t	$P \geq t$	Intervallo al 95%	
$\ln(x^*)$	-0.1534	0.0146	-10.45	0.000	-0.1824	-0.1245
<i>figli</i> 0 – 5	0.0313	0.0196	1.59	0.112	-0.0074	0.0700
<i>figli</i> 5 – 14	0.0357	0.0122	2.92	0.004	0.0115	0.0599
costante	1.8733	0.1503	12.46	0.000	1.5768	2.1698
Coppie con figli fino ai 14 anni; curva di domanda quadratica.						
N. di oggetti	207					
F(4,202)	24.21					
$R^2$	0.3241					
$\sqrt{MSE}$	0.1003					
$w_1^*$	Coefficiente	Standard Error	t	$P \geq t$	Intervallo al 95%	
$\ln(x^*)$	0.1049	0.4810	0.22	0.828	-0.8436	1.0535
$\ln(x^*)^2$	-0.0125	0.0231	-0.54	0.590	-0.0579	0.0330
<i>figli</i> 0 – 5	0.0320	0.0198	1.61	0.108	-0.0071	0.0711
<i>figli</i> 5 – 14	0.0363	0.0124	2.92	0.004	0.0118	0.0608
costante	0.5352	2.51067	0.21	0.831	-4.4153	5.4857

Tabella 1.3: Output delle regressioni nel modello *lineare*(1.2) e in quello *quadratico* (1.3) per *coppie con figli fino ai 14 anni*. I risultati sono ottenuti utilizzando l'opzione "*robust*", per ottenere Standard Errors robusti rispetto ad eteroschedasticità.

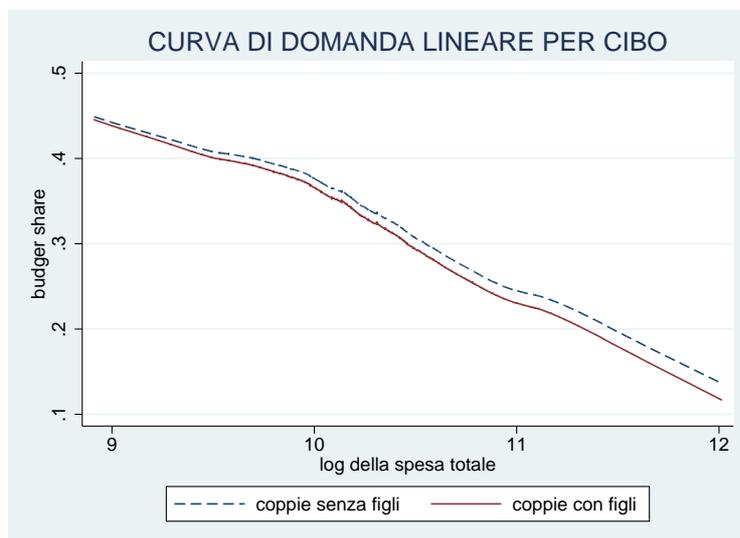


Figura 1.2: Curva di domanda per cibo per entrambe le tipologie di nuclei familiari nel mio studio.

	Effetti marginali Regressione OLS	Effetti marginali Regressione Beta
Modello Lineare $\ln(x^*)$	-0.1155	-0.1185
Modello Quadratico $\ln(x^*)$	-0.1129	-0.1139
$\ln(x^2)$	-0.0055	-0.0056

Tabella 1.4: Effetti marginali della regressione OLS e Beta per la specificazione lineare (1.2) e per quella quadratica (1.3).

delle statistiche invece non sono molto diversi. Infatti i valori delle medie e delle deviazioni standard, ottenuti dallo studio inglese, sono rispettivamente 0.343 e 0.109 per coppie con un figlio e 0.365 e 0.101 per coppie con due figli.

Dalla Figura (1.2) si può notare che, come accade anche nello studio di Blundell, Duncan e Pendakur (si veda Figura (1.1)), la specificazione *lineare* della curva di domanda di Working-Leser sembra l'approssimazione più ragionevole per questa curva di *budget share*. In più dai grafici si può vedere uno spostamento parallelo della curva di domanda per le due diverse tipologie familiari al variare del livello di spesa.

La stima di  $\beta_1$ , coefficiente del logaritmo della spesa totale  $\ln(x^*)$ , ottenuta da Blundell, Duncan e Pendakur è  $-0.1288$ ; mentre quella ottenuta con i campioni presi in considerazione in questa tesi, come si vede dalle Tabelle (1.2) e (1.3) sono  $-0.1155$  per coppie senza figli e  $-0.1533$  per coppie con figli. Anche in questa analisi i valori ottenuti non si discostano molto da quelli dello studio inglese.

Gli effetti marginali implicati dalla stima di massima verosimiglianza dei modelli *lineare*, (1.2), e *quadratico*, (1.3), sono riportati in Tabella (1.4). Per ottenerli ho effettuato la stima di entrambi i modelli con il metodo OLS e tramite una regressione Beta dato il supporto limitato tra zero e uno della variabile dipendente  $w_1^*$ . Appare evidente come le conclusioni inferenziali ottenute con il metodo OLS siano del tutto equivalenti da un punto di vista informativo a quelle ottenute con una regressione Beta, confermando l'approccio pragmatico al problema suggerito nella letteratura degli anni Sessanta.

Nonostante la curva di domanda *lineare* sembri la più adatta per descrivere la funzione  $\psi$  in (1.1) con i dati presi in considerazione in questa tesi, vi sono casi in cui invece la specificazione *quadratica* rappresenta bene la curva di domanda per cibo. Per questo nei prossimi capitoli analizzerò anche il caso in cui  $\psi$  sia di tipo *quadratico*. Riporto di seguito un grafico, Figura (1.3), preso dallo studio di Attanasio, Battistin, Mesnard (2007)

sullo studio delle curve di domanda in Colombia, che supporta l'evidenza nella letteratura recente di una relazione *quadratica* nella domanda per cibo se la popolazione analizzata è povera. Infatti, nell'analisi effettuata da Attanasio, Battistin, Mesnard (2007), vengono utilizzate come campione popolazioni molto povere che vivono in piccole città della Colombia rurale e, come è stato fatto in questa tesi, i prezzi e le caratteristiche demografiche sono considerate fisse. La Figura (1.3) mostra come il *budget share* per cibo varia all'aumentare della spesa totale.

## 1.5 Obiettivo della Tesi

L'obiettivo della tesi è quello di rispondere al seguente quesito: prendendo in considerazione entrambe le specificazioni della curva di domanda di *spesa per cibo* (*lineare e quadratica*), quali sono gli effetti di un errore di misura nella *spesa per cibo* sulle stime dei parametri coinvolti?

Per rispondere a questa domanda presento nel Capitolo 2 e Capitolo 3 due soluzioni alternative. Nel primo descrivo la soluzione proposta da Lewbel (1996); nel secondo, invece, presento un modello basato su una diversa struttura dell'errore di misura. Il principale contributo originale di questa tesi è contenuto nel Capitolo 3, dove caratterizzo la distorsione indotta da un errore di misura "piccolo" nella *spesa per cibo*. In quanto segue, il termine "piccolo" si riferirà a errore classico la cui varianza è molto contenuta rispetto a quella della variabile misurata con errore. Nel Capitolo 4 sono presentate le conclusioni di questa tesi. Nell'Appendice A è riportato un approfondimento sull'indagine sui Redditi delle Famiglie Italiane di Banca d'Italia; nell'Appendice B sono mostrati tutti i conti necessari per poter ricavare la distorsione indotta da errore di misura "piccolo" nella curva di domanda e nell'Appendice C, invece, sono esposti i passaggi per poter ricavare una regola per correggere la distorsione indotta appunto dall'errore di misura. Nell'Appendice D, infine, sono riportati i grafici addizionali rispetto a quelli riportati nel testo principale.

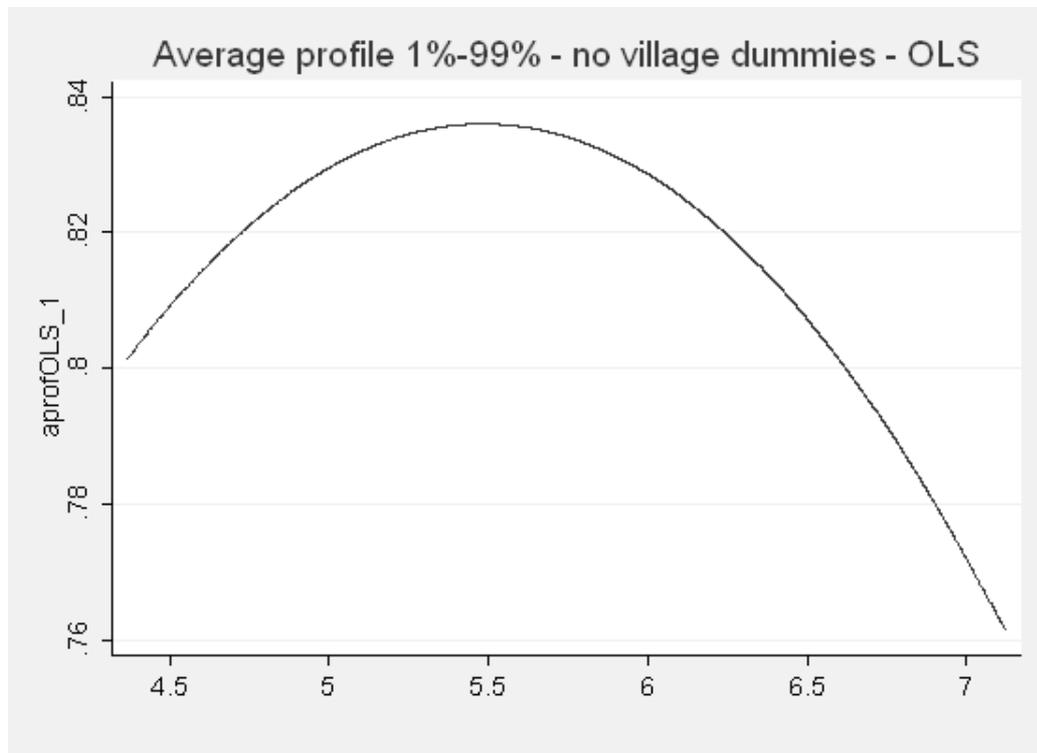


Figura 1.3: Curva di domanda per cibo in Colombia, dallo studio di Attanasio, Battistin, Mesnard (2007).

## Capitolo 2

# Stima della curva di domanda per cibo in presenza di errore di misura

### 2.1 Introduzione

In questo capitolo considero la curva di domanda nella specificazione *lineare* e *quadratica* in  $\ln(x^*)$ , come scritto di seguito:

$$w_1^* = \beta_0 + \beta_1 \ln(x^*) + u \quad (2.1)$$

$$w_1^* = \beta_0 + \beta_1 \ln(x^*) + \beta_2 \ln(x^*)^2 + u. \quad (2.2)$$

Quello che si vuole analizzare è come si possano ottenere stime consistenti dei parametri in (2.1) e (2.2) nel caso in cui vi sia errore di misura in  $y_1^*$ . La soluzione classica al problema di errore di misura passa per la ricerca di una variabile strumentale, ovvero una variabile che corredi con il vero valore della grandezza misurata con errore, ma non con l'errore stesso. In tal caso, Lewbel (1996) mostra che il metodo delle variabili strumentali non produce stime consistenti dei parametri di interesse  $\beta_1$  e  $\beta_2$ . Il motivo è che l'errore di misura in  $y_1$  comporta errore sia in  $\ln(x^*)$  che in  $w_1^*$ , quindi si ha errore di misura sia nella variabile indipendente che in quella

dipendente nell'Equazione (1.1). In più, se anche fosse possibile errore di misura classico in  $y_1^*$ , ciò comporterebbe errore di misura non classico in  $w_1^*$  e in  $\ln(x^*)$ . Nel primo caso poichè  $w_1^*$  contiene errore sia al numeratore che al denominatore; nel secondo caso per la non linearità implicata dal logaritmo. Sono possibili più soluzioni a seconda delle assunzioni che si fanno sulla natura dell'errore di misura; di seguito presento la procedura proposta da Lewbel e nel prossimo capitolo, invece, mi avvalgo di una struttura dell'errore che rappresenta il contributo originale di questa tesi. Di seguito e in tutte le sezioni di applicazione della teoria ai dati prenderò in considerazione solo una tipologia familiare: *coppie senza figli*. Questa decisione è giustificata dal fatto che la struttura del problema per le due tipologie familiari non cambia, infatti la specificazione del modello nel caso di *coppie con figli fino ai 14 anni* può essere sintetizzata dal modello *coppie senza figli* considerando  $w_1^*$  e  $x^*$  come i residui della variabile "figli".

## 2.2 Una possibile soluzione

In quanto segue mi concentrerò sulla situazione in cui la *spesa per cibo* sia misurata con errore, mentre la spesa per non cibo sia misurata correttamente.

Lewbel (1996) propone una procedura per ottenere stime consistenti dei parametri di interesse della curva di domanda sia nella specificazione *lineare* che *quadratica* in  $\ln(x^*)$ . Di seguito mi riferirò a questo metodo con "metodo di Lewbel". La soluzione suggerita consiste nel moltiplicare l'equazione della domanda per la spesa totale e stimare i parametri della nuova equazione usando variabili strumentali, sotto forma di minimi quadrati a due stadi (2SLS).

L'idea alla base della procedura descritta in Lewbel (1996) può essere formalizzata mediante le seguenti assunzioni sulla struttura dell'errore di misura. Come detto nell'introduzione, il *budget share* può contenere errore di misura sia al numeratore che al denominatore. Considerando prima

solo l'errore al numeratore, si può scrivere il *budget share* come:

$$\frac{y_1}{x^*} = w_1^* + v.$$

L'equazione precedente comporta la seguente assunzione circa la struttura dell'errore in  $y_1^*$ :

$$y_1 = y_1^* + vx^*, \quad (2.3)$$

dove  $v$  rappresenta l'errore di misura del *budget share* che consegue da un errore in  $y_1^*$  se la spesa totale  $x^*$  non fosse misurata con errore. Partendo dalla (1.1) e utilizzando la (2.3) si può scrivere:

$$\frac{y_1}{x^*} = \psi(\beta, \ln(x^*)) + u + v, \quad (2.4)$$

dove si assume che  $u$  e  $v$  siano a media nulla, indipendenti fra di loro e da  $x^*$ .

Si consideri ora la complicazione aggiuntiva che segue dall'aver errore di misura nella spesa totale. La presenza dell'errore  $v$  fa sì che  $x \neq x^*$ . Definendo l'errore moltiplicativo come  $V = \frac{x}{x^*}$ , che implica la relazione  $\ln(V) = \ln(x) - \ln(x^*)$ , si ottengono le seguenti equazioni:

$$w_1 = \frac{w_1^* + v}{V},$$

$$w_1 = [\psi(\beta, \ln(x) - \ln(V)) + u + v]/V,$$

che mostrano come il *budget share* osservato dipenda dalla spesa totale osservata  $x$ , dall'errore del modello  $u$  e dagli errori di misura  $v$  e  $V$ , che *non* sono indipendentemente distribuiti data la relazione:

$$V = 1 + v. \quad (2.5)$$

Complessivamente, il modello descritto utilizza due assunzioni fondamentali. La prima è riassunta nell'Equazione (2.3). La seconda è che l'errore  $v$  in (2.4) sia classico, ovvero che non correli con  $u$  nè con alcuna altra

grandezza considerata.

Stabilita la struttura dell'errore di misura la stima dei parametri nella curva di domanda richiede la disponibilità di un vettore di strumenti  $z$  che vengono utilizzati per una stima IV. Si assumano verificate le condizioni standard, cioè che  $z$  correli con  $x^*$  ma non con gli errori  $v$  e  $u$ . Tali assunzioni implicano che anche  $V$  e  $z$  siano incorrelate. L'Equazione (2.5), che determina la relazione di dipendenza tra l'errore di misura del consumo totale  $V$  e l'errore di misura  $v$  nella spesa del consumo per il singolo bene, causa (tuttavia) il fallimento della stima con il metodo IV.

La tecnica suggerita da Lewbel (1996) prevede che si osservino  $y_1, x$  e quindi  $w_1$ ; si moltiplichino l'equazione del *budget share* per  $x$ , e si stimi l'espressione risultante per la curva di domanda con i minimi quadrati a due stadi (2SLS).

In particolare, con riferimento alla specificazione *lineare*, il procedimento suggerito può essere sintetizzato come segue:

$$E(zxw_1) = \beta_0 E(zx) + \beta_1 E(xz \ln(x)).$$

Questa condizione sui momenti implica la seguente relazione:

$$xw_1 = \beta_0 x + \beta_1 x \ln(x) + e, \quad (2.6)$$

dove  $ux = e$ , ha media 0 ed è incorrelato con gli strumenti  $z$ . Precisamente, data la condizione di linearità dell'equazione,  $\beta_0$  e  $\beta_1$  possono essere stimati, regredendo  $xw_1$  sia su  $x$  che  $x \ln(x)$ , usando 2SLS con gli strumenti  $z$ . Dall'Equazione (2.6) l'espressione che si ottiene è:

$$y_1 = \beta_0 x + \beta_1 x \ln(x) + e. \quad (2.7)$$

Stimando questa equazione con 2SLS si ottiene una stima consistente di  $\beta_1$  ma non di  $\beta_0$ <sup>1</sup>, mentre applicando lo stesso metodo di stima all'equazione del *budget share* (2.8) non si ottiene nessuno stimatore consistente (né di  $\beta_0$ , né di  $\beta_1$ ).

---

<sup>1</sup>Sebbene  $z$  sia incorrelato con  $\epsilon$  essi non sono indipendenti e quindi la stima col metodo 2SLS dà stimatori consistenti ma non efficienti.

## 2.3 Alcune evidenze da un disegno di simulazione

In quanto segue verrà presentata l'evidenza sulla validità del "metodo di Lewbel" basato su un esercizio di simulazione. In particolare, nella prossima Sezione descriverò la logica seguita per il disegno della simulazione stessa e, di seguito, mostrerò i risultati relativi alla specificazione di domanda *lineare* e *quadratica*.

### 2.3.1 Il processo generatore dei dati

Di seguito presento i passaggi necessari per poter replicare la soluzione proposta da Lewbel per la stima della curva di domanda con dati reali e riporto le soluzioni ottenute. I dati per poter effettuare queste simulazioni sono quelli descritti nella Sezione 0.4.

Per fissare un processo generatore dei dati per l'Equazione (2.1) o (2.2) serve stabilire i valori di  $\beta_0$ ,  $\beta_1$  e  $\beta_2$ . Il procedimento seguito per generare i dati nell'esercizio di simulazione avviene secondo i seguenti passi:

1. ottengo i valori dei parametri  $\beta_0$ ,  $\beta_1$  e  $\beta_2$  a partire dalle stime OLS delle equazioni (2.1) e (2.2) sui dati osservati. In altre parole, i valori di tali parametri sono calibrati sui dati reali;
2. genero  $\ln(z) \sim N(\mu_z, \sigma_z^2)$ , con  $\mu_z$  e  $\sigma_z^2$  calcolati dai dati reali;
3. genero  $\ln(x^*)$  a partire dalla regressione di primo stadio di  $\ln(x^*)$  su  $\ln(z)$  nei dati osservati;
4. ottengo  $w_1^*$  dall'Equazione (2.1) per il modello *lineare* o dall'Equazione (2.2) per quello *quadratico*;
5. ottengo, di conseguenza,  $y_1^*$  e poi  $y_2^*$ ;
6. avendo a disposizione, a questo punto, tutte le variabili necessarie ( $w_1^*$ ,  $y_1^*$ ,  $x^*$  e  $z$ ), sporco con errore le quantità sopra ottenute, utilizzando la struttura dell'errore descritta nella Sezione 1.2.

La numerosità del campione considerata è 1000 e le simulazioni effettuate sono 10000. Il parametro rispetto al quale “disegno” tali simulazioni è l’intensità dell’errore nel *budget share*, definita come rapporto fra varianza vera e varianza osservata:

$$\rho \equiv \frac{Var(w_1^*)}{Var(w_1)},$$

dove  $Var(w_1^*)$  è calcolata a partire dai dati generati al punto 3. I valori di  $\rho$  considerati sono:  $\rho = 1$ ,  $\rho = 0.9$  e  $\rho = 0.75$ . L’errore  $v$  è assunto distribuito come una variabile casuale normale di media nulla e varianza ottenuta a partire dal valore di  $\rho$ .

Le Figure presentate in quanto segue si riferiranno alla seguente quantità:

$$\frac{\hat{\beta}_j - \beta_j}{\sqrt{Var(\hat{\beta}_j)}} \quad (j = 0, 1, 2.),$$

ovvero, a valori standardizzati degli stimatori “OLS”, “IV” e “Lew” del parametro  $\beta_j$ .

### 2.3.2 Curva di domanda lineare

Riscrivo la curva di domanda come specificazione del modello di *budget share* nella forma di Working-Leser:

$$w_1^* = \beta_0 + \beta_1 \ln(x^*) + u \quad (2.8)$$

dove  $E(u) = 0$  e  $\beta_0$  e  $\beta_1$  sono i parametri da stimare. Dalle equazioni precedenti si ottiene la seguente espressione:

$$w_1 = \frac{\beta_0}{V} + \beta_1 \frac{\ln(x) - \ln(V)}{V} + \frac{v + u}{V} \quad (2.9)$$

### Risultati

Come si può vedere dalle Figure (2.1), (2.2), (2.3) e dalla Tabella (2.1), le proprietà dello stimatore OLS peggiorano al diminuire di  $\rho$ . Al contrario,

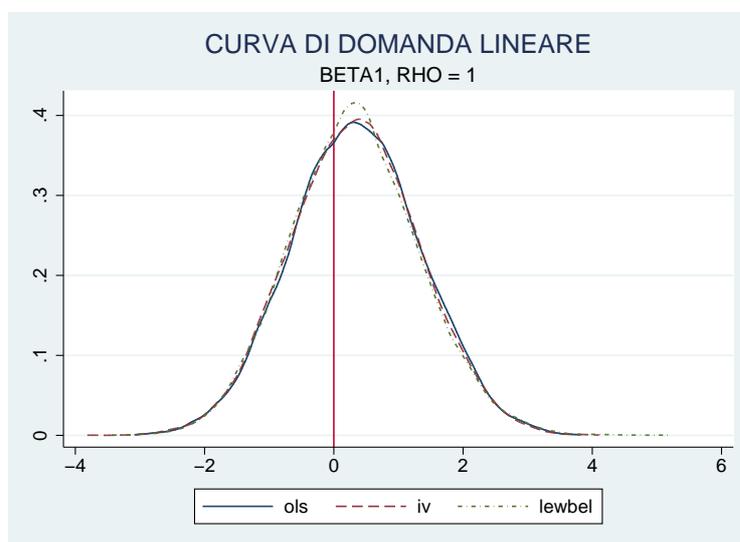


Figura 2.1: Distribuzione simulata delle stime standardizzate di  $\beta_1$  con i metodi "OLS", "IV" e "Lew" in un modello *lineare* con struttura dell'errore di Lewbel con  $\rho = 1$  (10000 simulazioni,  $\beta_1 = -0.1155$ ).

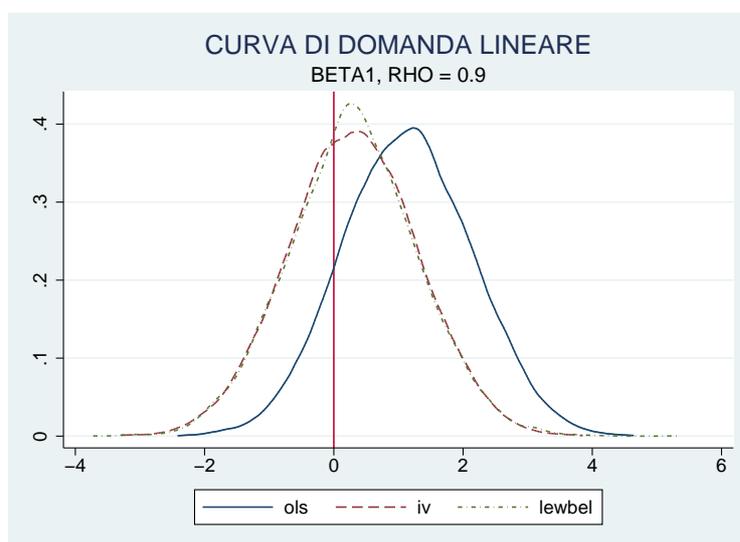


Figura 2.2: Distribuzione simulata delle stime standardizzate di  $\beta_1$  con i metodi "OLS", "IV" e "Lew" in un modello *lineare* con struttura dell'errore di Lewbel con  $\rho = 0.9$  (10000 simulazioni,  $\beta_1 = -0.1155$ ).

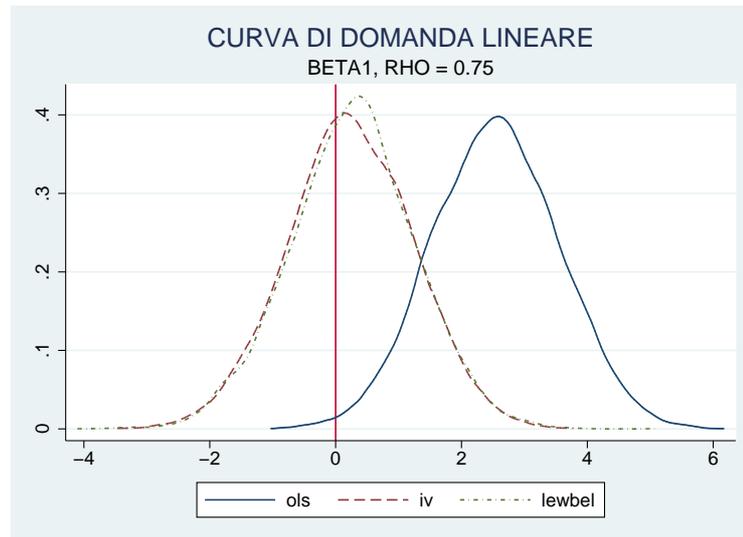


Figura 2.3: Distribuzione simulata delle stime standardizzate di  $\beta_1$  con i metodi "OLS", "IV" e "Lew" in un modello *lineare* con struttura dell'errore di Lewbel con  $\rho = 0.75$  (10000 simulazioni,  $\beta_1 = -0.1155$ ).

Modello lineare, $\beta_1$			
Metodo stima	$\rho = 1$	$\rho = 0.9$	$\rho = 0.75$
OLS	0.3509	1.1340	2.5446
IV	0.3304	0.2967	0.2353
LEW	0.3219	0.3044	0.2695

Tabella 2.1: Media della distribuzione simulata di  $\beta_1$  al variare di  $\rho$ , in un modello *lineare* con errore di misura del "metodo di Lewbel" (10000 simulazioni).

le proprietà dello stimatore ottenute con il “metodo di Lewbel” rimangono buone, e, sorprendentemente, la stessa cosa si verifica per lo stimatore IV. Appare quindi evidente che, almeno per quanto riguarda i disegni di simulazione calibrati sui dati di Banca d’Italia, lo stimatore IV non sia qualitativamente peggiore dello stimatore di Lewbel.

### 2.3.3 Curva di domanda quadratica

Allo stesso modo della procedura per il modello *lineare*, si scrive l’equazione del *budget share* osservato nella forma *quadratica*:

$$w_1 = \beta_0 + \beta_1 \ln(x) + \beta_2 \ln(x)^2 + u. \quad (2.10)$$

Si moltiplica l’equazione a destra e sinistra per  $x$  come propone Lewbel e si ottiene:

$$y_1 = \beta_0 x + \beta_1 x \ln(x) + \beta_2 x \ln(x)^2 + ux. \quad (2.11)$$

Stimando quest’ultima equazione con il metodo 2SLS si può ottenere una stima consistente di  $\beta_2$  ma non di  $\beta_0$  e  $\beta_1$ .

### Risultati

Nella Figure e nelle Tabelle di seguito riportate si può vedere come variano le stime di  $\beta_1$  e di  $\beta_2$  al variare di  $\rho$  in un modello *quadratico*<sup>2</sup>. Osservando i grafici si può vedere che anche con la specificazione *quadratica* le proprietà dello stimatore OLS peggiorano al diminuire di  $\rho$ , mentre le proprietà degli stimatori ottenuti con con il metodo IV e di Lewbel rimangono buone.

---

<sup>2</sup>Ai fini di rendere più visibile le differenze tra le curve presentate in alcuni grafici, nell’Appendice D riporto la distribuzione per le stime utilizzando “OLS”, “IV” e “Lew”, separatamente.

Per la specificazione *quadratica* con  $\rho = 0.75$  il numero le simulazioni sono 700 per mancanza di tempo ad effettuarne di più.

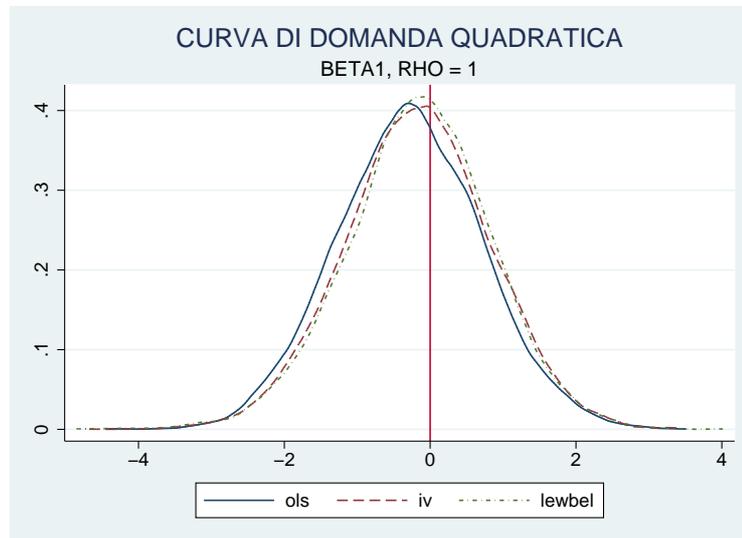


Figura 2.4: Distribuzione simulata delle stime standardizzate di  $\beta_1$  con i metodi "OLS", "IV" e "Lew" in un modello *quadratico* con struttura dell'errore di Lewbel con  $\rho = 1$  (10000 simulazioni,  $\beta_1 = 0.1572$ ).

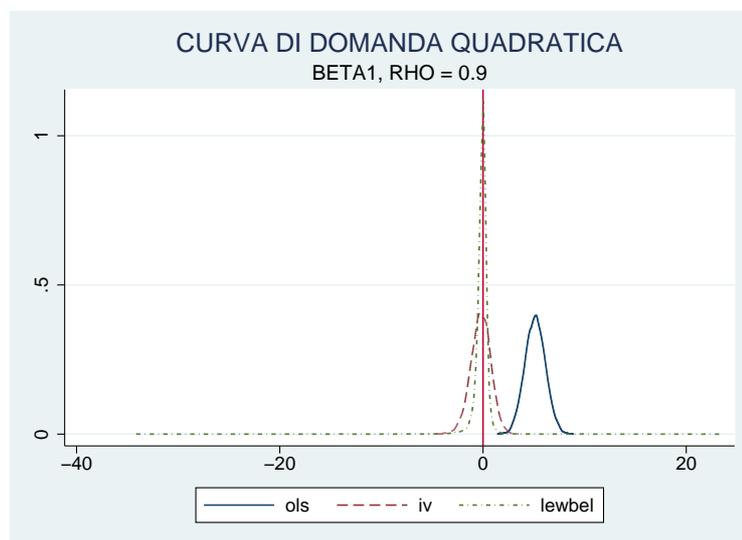


Figura 2.5: Distribuzione simulata delle stime standardizzate di  $\beta_1$  con i metodi "OLS", "IV" e "Lew" in un modello *quadratico* con struttura dell'errore di Lewbel con  $\rho = 0.9$  (10000 simulazioni,  $\beta_1 = 0.1572$ ).

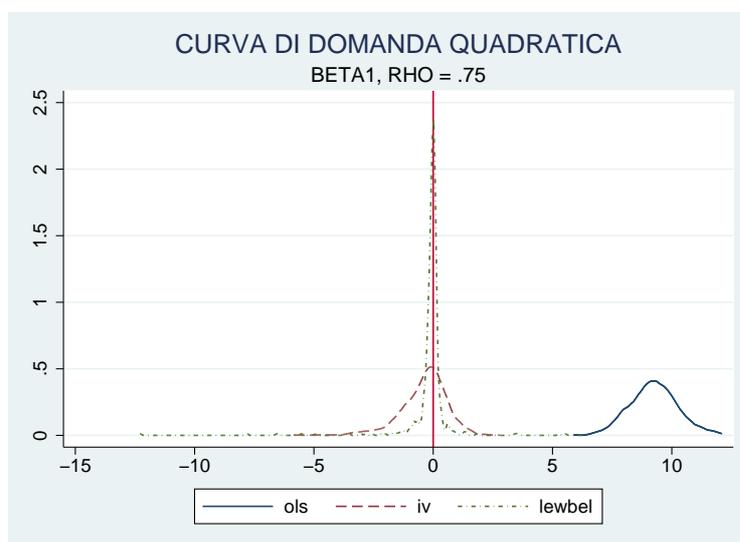


Figura 2.6: Distribuzione simulata delle stime standardizzate di  $\beta_1$  con i metodi "OLS", "IV" e "Lew" in un modello *quadratico* con struttura dell'errore di Lewbel con  $\rho = 0.75$  (700 simulazioni,  $\beta_1 = 0.1572$ ).

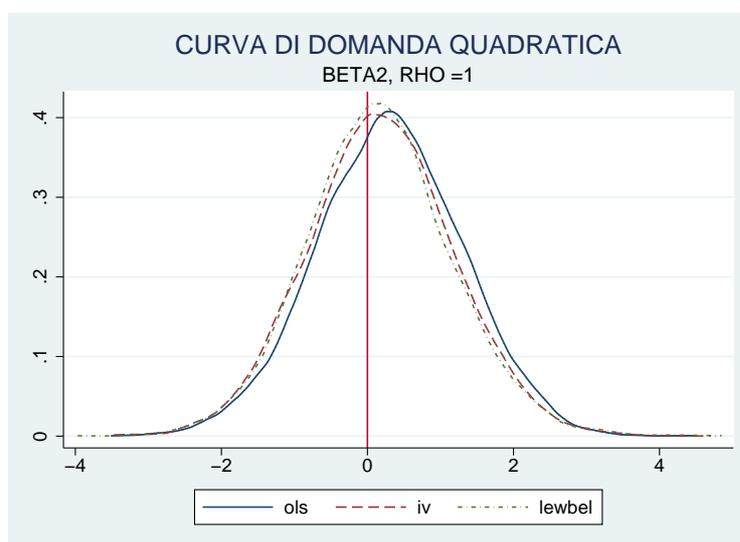


Figura 2.7: Distribuzione simulata delle stime standardizzate di  $\beta_2$  con i metodi "OLS", "IV" e "Lew" in un modello *quadratico* con struttura dell'errore di Lewbel con  $\rho = 1$  (10000 simulazioni,  $\beta_1 = -0.0133$ ).

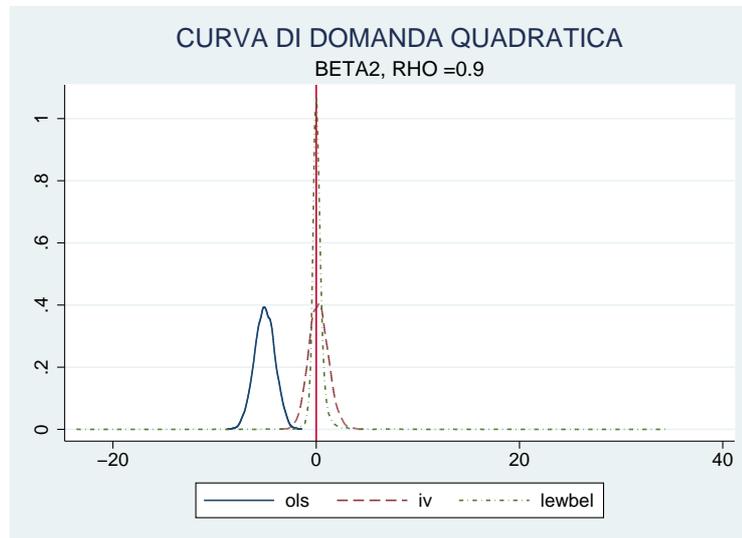


Figura 2.8: Distribuzione simulata delle stime standardizzate di  $\beta_2$  con i metodi "OLS", "IV" e "Lew" in un modello *quadratico* con struttura dell'errore di Lewbel con  $\rho = 0.9$  (10000 simulazioni,  $\beta_1 = -0.0133$ ).

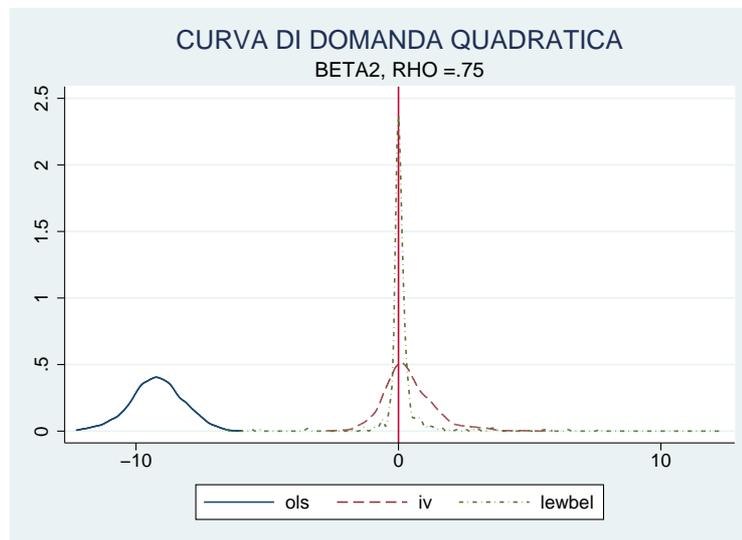


Figura 2.9: Distribuzione simulata delle stime standardizzate di  $\beta_2$  con i metodi "OLS", "IV" e "Lew" in un modello *quadratico* con struttura dell'errore di Lewbel con  $\rho = 0.75$  (700 simulazioni,  $\beta_1 = -0.0133$ ).

<b>Modello quadratico, <math>\beta_1</math></b>			
Metodo stima	$\rho = 1$	$\rho = 0.9$	$\rho = 0.75$
OLS	-0.2877	5.1326	9.2442
IV	-0.1759	-0.2277	-0.3823
LEW	-0.1551	-0.1202	-0.1228
<b>Modello quadratico, <math>\beta_2</math></b>			
Metodo stima	$\rho = 1$	$\rho = 0.9$	$\rho = 0.75$
OLS	0.3002	-5.0644	-9.2008
IV	0.1836	0.2372	0.3934
LEW	0.16024	0.1224	0.1241

Tabella 2.2: Media delle distribuzione simulata di  $\beta_1$  e  $\beta_2$  al variare di  $\rho$ , in un modello *quadratico* con errore di misura del “metodo di Lewbel”.

## 2.4 Alcuni aspetti critici della soluzione discussa

In questo Capitolo è stato presentato l’approccio proposto da Lewbel per poter ottenere stime consistenti dei parametri della curva di domanda per cibo. Questo metodo si basa su una struttura dell’errore ben precisa, che assegna errore di misura classico al *budget share*  $w_1^*$ . Proprio questa scelta di assegnare errore di misura classico a *budget share* è il punto critico del “metodo di Lewbel”. Infatti, questa struttura attribuisce un errore di misura con supporto  $(-\infty, +\infty)$  a una variabile, in questo caso  $w^*$ , che invece è limitata tra  $(0, 1)$ . Di certo, non si può ritenere questa scelta statisticamente corretta.

Nel prossimo Capitolo proporrò una struttura dell’errore alternativa compatibile con il supporto limitato di  $w_1^*$ . In particolare farò vedere come questo nuovo metodo sia sensibile alle proprietà dell’errore di misura e suggerirò una soluzione basata su quelle situazioni in cui l’errore di misura ha una varianza molto piccola.



# Capitolo 3

## Una soluzione alternativa

### 3.1 Introduzione

In questo Capitolo propongo una alternativa per ovviare, in particolare, al punto debole del “metodo di Lewbel”. Nella soluzione vista nel Capitolo precedente, infatti, viene attribuito errore di misura classico, che ha supporto tra  $(-\infty, +\infty)$ , a  $w^*$  che invece ha supporto nell’intervallo  $(0,1)$ . La differenza principale, quindi, tra il procedimento visto nel Capitolo precedente e il modello presentato di seguito, sarà in una struttura dell’errore che mi permette di migliorare tale aspetto critico del “metodo di Lewbel”. Quello che faccio, considerando la curva di domanda per *spesa per cibo*, è assumere affetto da errore di misura non più il *budget share* (come assume Lewbel) ma il logaritmo della *spesa per cibo*. Ne segue che errore di misura e la variabile dipendente hanno lo stesso supporto. La notazione di questo Capitolo riprende esattamente quella presentata nell’introduzione.

Il Capitolo è organizzato come segue. Mi concentrerò dapprima su evidenza da simulazione per far vedere che, con una struttura dell’errore coerente con la natura delle variabili coinvolte, il metodo proposto da Lewbel non conserva le proprietà desiderate. Una correzione che tenga conto dell’errore di misura in fase di stima non è facile da reperire. Per tale motivo deriverò una correzione nel caso di errore di misura “piccolo”. In breve,

considerando le regressioni (2.1) e (2.2) otterrò le espressioni per la distorsione da errore di misura nel caso in cui quest'ultimo sia piccolo, ovvero sia una "perturbazione" normale al logaritmo della *spesa per cibo* a media nulla e varianza piccola. Ciò equivale ad operare uno sviluppo in serie di Taylor per valori della varianza dell'errore prossimi a zero. Tale strategia utilizza delle tecniche già utilizzate da Chesher (1991) e Battistin e Chesher (2007a e 2007b). Infine, partendo dai risultati di questi articoli, proporrò un metodo per correggere la distorsione da errore di misura verificandone le proprietà mediante simulazione.

### 3.2 Stima del modello alternativo con il "metodo IV classico" e il "metodo di Lewbel"

Con questa nuova struttura dell'errore ho effettuato le stime dei parametri di interesse sia con il metodo classico delle variabili strumentali, sia applicando il procedimento proposto da Lewbel, che di seguito, in tabelle e grafici saranno sintetizzati con le sigle "IV" e "Lew".

### 3.3 Il processo generatore dei dati

Le simulazioni presentate in quanto segue sono strutturate nello stesso modo descritto nella Sezione 2.3.1. In questo caso, a differenza del "metodo di Lewbel" descritto nella Sezione 1.2, si sporca con errore non  $y_1^*$  nei livelli, ma il logaritmo della *spesa per cibo*. Si può così scrivere:

$$\ln(y_1) = \ln(y_1^*) + \sigma_1 \epsilon \quad (3.1)$$

con  $\epsilon \sim N(0,1)$  e  $\sigma_1 > 0$ .

### 3.3.1 Curva di domanda lineare e quadratica

Le equazioni della curva di domanda *lineare* e *quadratica* da stimare con metodo "OLS" e "IV" e quelle da stimare con il metodo "Lew" sono le stesse riportate nelle Sezioni 1.3.1 e 1.3.2.

#### Risultati

Nelle Figure e nelle Tabelle che seguono si può osservare come i valori delle stime di  $\beta_1$  per il modello *lineare* e di  $\beta_1$  e  $\beta_2$  per quello *quadratico* cambino al variare di  $\rho$ . Queste stime sono ottenute con i tre metodi di stima utilizzati: "OLS", "IV" e "Lew". Dai grafici si può notare che nella specificazione *lineare* solo le proprietà dello stimatore OLS peggiorano al diminuire di  $\rho$ , mentre le altre rimangono buone. Per quanto riguarda, invece, la specificazione *quadratica*, sia per  $\beta_1$  che per  $\beta_2$  le proprietà dello stimatore OLS peggiorano al diminuire di  $\rho$ , quelle dello stimatore IV rimangono buone per qualsiasi valore di  $\rho$ , mentre quelle dello stimatore di Lewbel peggiorano solo al livello di  $\rho = 0.75^1$ .

---

<sup>1</sup>Ai fini di rendere più visibile le differenze tra le curve presentate in alcuni grafici, nell'Appendice D riporto la distribuzione per le stime utilizzando "OLS", "IV" e "Lew", separatamente.

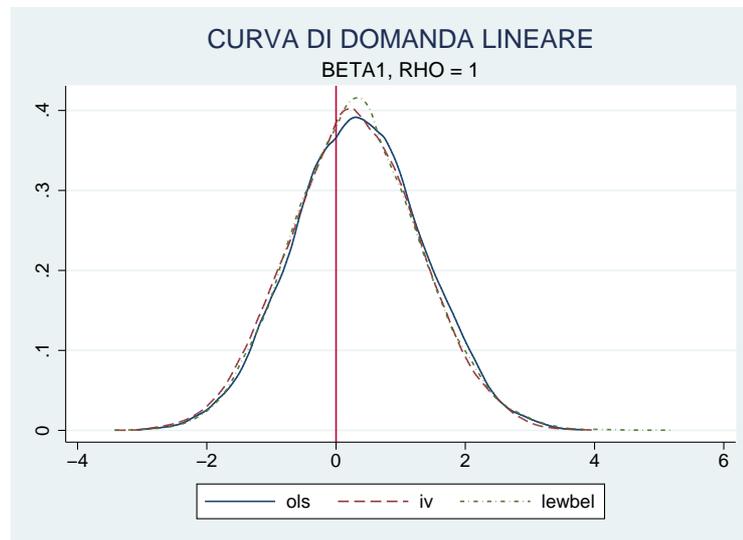


Figura 3.1: Distribuzione simulata delle stime standardizzate di  $\beta_1$  con i metodi “OLS”, “IV” e “Lew” in un modello *lineare* con struttura dell’errore alternativa a quella di Lewbel con  $\rho = 1$  (10000 simulazioni,  $\beta_1 = -0.1155$ ).

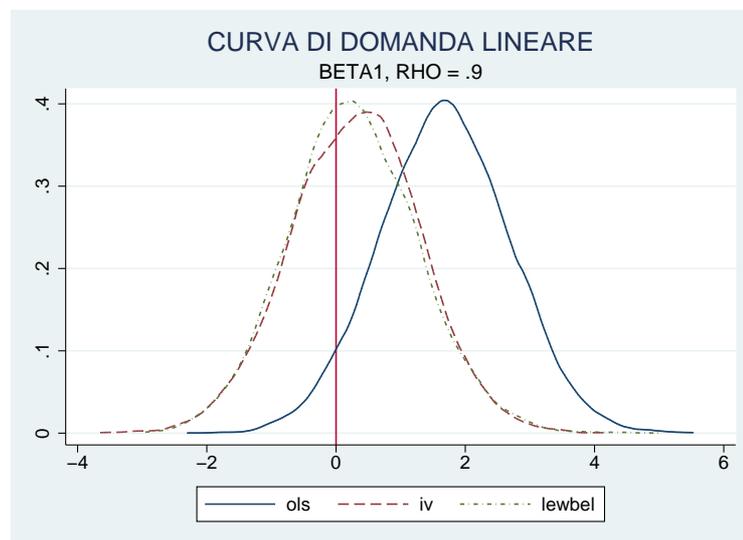


Figura 3.2: Distribuzione simulata delle stime standardizzate di  $\beta_1$  con i metodi “OLS”, “IV” e “Lew” in un modello *lineare* con struttura dell’errore alternativa a quella di Lewbel con  $\rho = 0.9$  (10000 simulazioni,  $\beta_1 = -0.1155$ ).

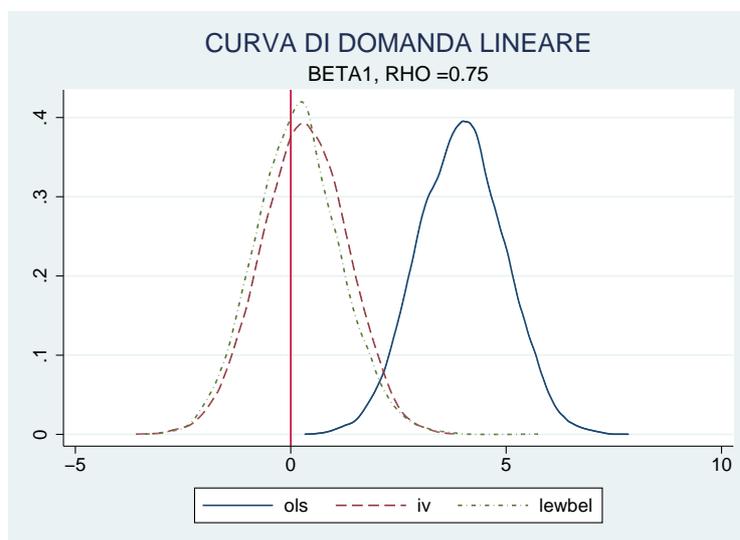


Figura 3.3: Distribuzione simulata delle stime standardizzate di  $\beta_1$  con i metodi "OLS", "IV" e "Lew" in un modello *lineare* con struttura dell'errore alternativa a quella di Lewbel con  $\rho = 0.75$  (10000 simulazioni,  $\beta_1 = -0.1155$ ).

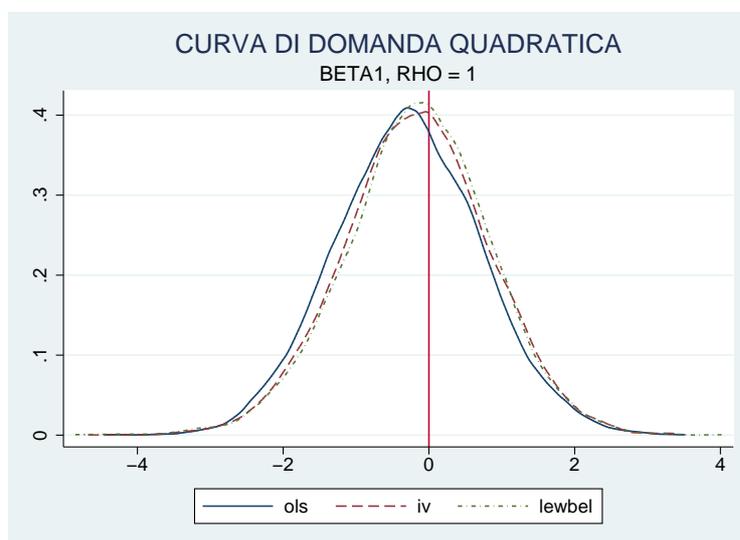


Figura 3.4: Distribuzione simulata delle stime standardizzate di  $\beta_1$  con i metodi "OLS", "IV" e "Lew" in un modello *quadratico* con struttura dell'errore alternativa a quella di Lewbel con  $\rho = 1$  (10000 simulazioni,  $\beta_1 = 0.1572$ ).

Modello lineare, $\beta_1$			
Metodo stima	$\rho = 1$	$\rho = 0.9$	$\rho = 0.75$
OLS	0.3509	1.6801	3.9602
IV	0.2830	0.3052	0.3180
LEW	0.3219	0.2700	0.1641

Tabella 3.1: Media della distribuzione simulata di  $\beta_1$  al variare di  $\rho$ , in un modello *lineare* con struttura dell'errore alternativa a quella di Lewbel (10000 simulazioni).

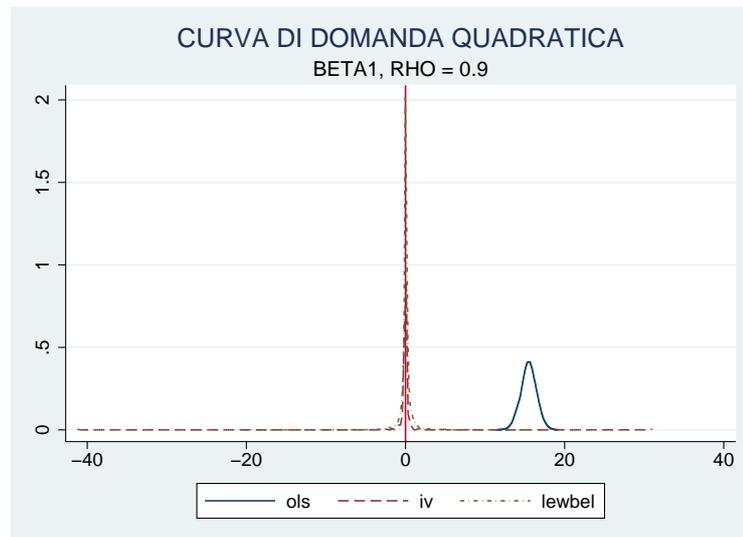


Figura 3.5: Distribuzione simulata delle stime standardizzate di  $\beta_1$  con i metodi "OLS", "IV" e "Lew" in un modello *quadratico* con struttura dell'errore alternativa a quella di Lewbel con  $\rho = 0.9$  (10000 simulazioni,  $\beta_1 = 0.1572$ ).

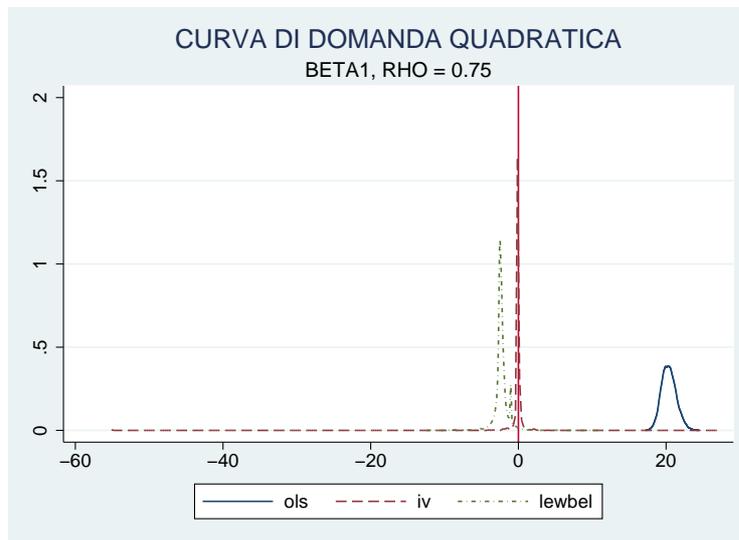


Figura 3.6: Distribuzione simulata delle stime standardizzate di  $\beta_1$  con i metodi "OLS", "IV" e "Lew" in un modello *quadratico* con struttura dell'errore alternativa a quella di Lewbel con  $\rho = 0.75$  (10000 simulazioni,  $\beta_1 = 0.1572$ ).

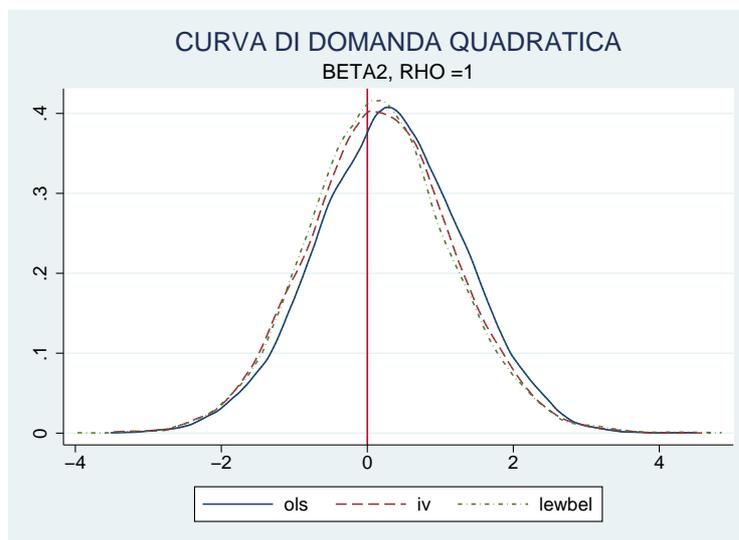


Figura 3.7: Distribuzione simulata delle stime standardizzate di  $\beta_2$  con i metodi "OLS", "IV" e "Lew" in un modello *quadratico* con struttura dell'errore alternativa a quella di Lewbel con  $\rho = 1$  (10000 simulazioni,  $\beta_2 = -0.0133$ ).

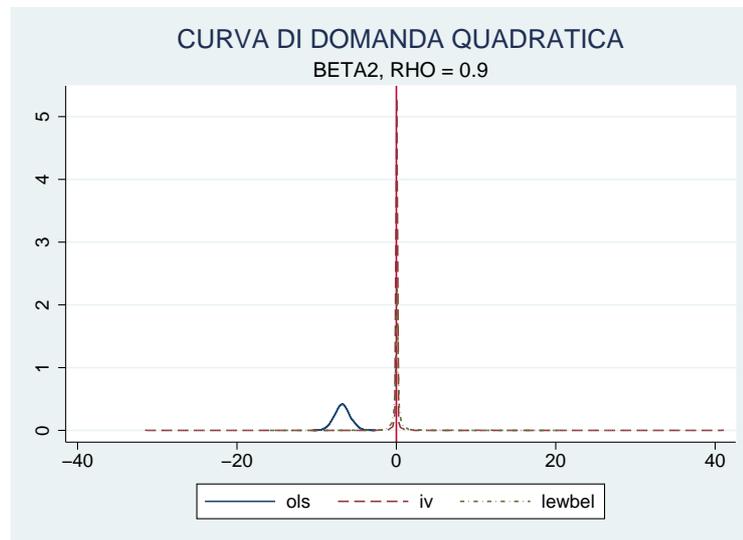


Figura 3.8: Distribuzione simulata delle stime standardizzate di  $\beta_2$  con i metodi “OLS”, “IV” e “Lew” in un modello *quadratico* con struttura dell’errore alternativa a quella di Lewbel con  $\rho = 0.9$  (10000 simulazioni,  $\beta_2 = -0.0133$ ).

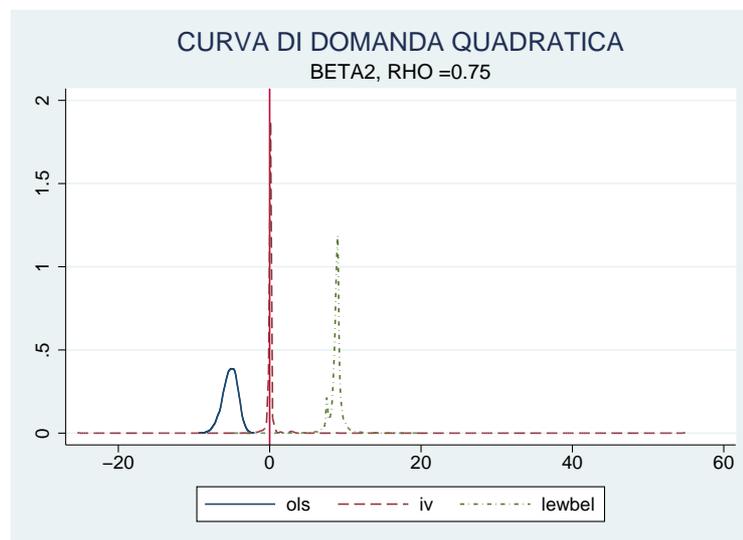


Figura 3.9: Distribuzione simulata delle stime standardizzate di  $\beta_2$  con i metodi “OLS”, “IV” e “Lew” in un modello *quadratico* con struttura dell’errore alternativa a quella di Lewbel con  $\rho = 0.75$  (10000 simulazioni,  $\beta_2 = -0.0133$ ).

<b>Modello quadratico, <math>\beta_1</math></b>			
Metodo stima	$\rho = 1$	$\rho = 0.9$	$\rho = 0.75$
OLS	-0.28940.3019	15.4795	20.3835
IV	-0.1762	-0.017	-0.0541
LEW	-0.1549	0.0300	-2.3106
<b>Modello quadratico, <math>\beta_2</math></b>			
Metodo stima	$\rho = 1$	$\rho = 0.9$	$\rho = 0.75$
OLS	0.3019	-6.7720	-5.2307
IV	0.1841	0.0188	0.6378
LEW	0.1601	0.0694	8.7979

Tabella 3.2: Media della distribuzione simulata di  $\beta_1$  e  $\beta_2$  al variare di  $\rho$  in un modello *quadratico*, con errore di misura alternativo a quello di Lewbel (10000 simulazioni).

### 3.4 Distorsione indotta da un errore di misura “piccolo” nella curva di domanda

In questo paragrafo si vuole calcolare la distorsione che un errore di misura comporta nella stima della curva di domanda quando l'errore intacca il  $\ln(y_1^*)$  come supposto in tutto il Capitolo 3.

La proposizione seguente caratterizza la distorsione indotta da errore di misura nel caso in cui la varianza dell'errore sia piccola.

**Proposizione.** *Si assuma che il logaritmo della spesa per cibo sia contaminato da errore di misura classico:*

$$\ln(y_1) = \ln(y_1^*) + \sigma_1 \epsilon,$$

essendo  $\sigma_1 > 0$ , con  $\epsilon$  indipendente da  $(y_1^*, w_1^*, z)$ .

Vale la seguente approssimazione:

$$E_{W_1|\ln(X)}[W_1|x] \simeq (1 - \sigma_1^2) \int_0^1 \Delta^*(s, x) ds E_{W_1^*|\ln(X^*)}[W_1^*|x] + \sigma_1^2 \int_0^1 s \Delta^*(s, x) ds,$$

dove  $A \simeq B$  significa  $A = B + o(\sigma_1^2)$  e:

$$\Delta^*(w_1, x) \equiv \frac{1}{2w_1(1-w_1)} \frac{\nabla_1^2 f_{Y_1^*, Y_2^*}[x + \ln w_1, x + \ln(1-w_1)]}{f_{\ln(X^*)}[x]}.$$

La proposizione enunciata, la cui dimostrazione è riportata nell'Appendice B, rappresenta il contributo principale di questa tesi. Da tale approssimazione segue che:

$$E_{w_1|\ln(X)}[W_1|x] \simeq E_{W_1^*|\ln(X^*)}[W_1^*|X]$$

se  $\sigma_1 \rightarrow 0^+$  (come c'è da aspettarsi). In altre parole, se la varianza dell'errore di misura tende a zero, cioè se l'errore di misura diventa sempre più

piccolo, la regressione stimata a partire dai dati “sporchi” tende a quella vera. Inoltre, nel caso in cui la curva di domanda sia *lineare*, ovvero quando

$$E_{W_1^*|ln(X^*)}[W_1^*|X] = \beta_0 + \beta_1 x,$$

abbiamo che:

$$E_{W_1|ln(X)}[W_1|X] \simeq \beta_0 + \beta_1 x \quad (3.2)$$

$$-\beta_0 \sigma_1^2 \int_0^1 \Delta^*(s, x) ds \quad (3.3)$$

$$-\beta_1 \sigma_1^2 x \int_0^1 \Delta^*(s, x) ds \quad (3.4)$$

$$+\sigma_1^2 \int_0^1 s \Delta^*(s, x) ds. \quad (3.5)$$

La precedente espressione permette di valutare l'effetto dell'errore di misura sulla stima dei parametri della curva di domanda. La regressione che identificherebbe  $\beta_0$  e  $\beta_1$  dai dati “sporchi” è quella di  $w_1$  su  $x$ ,  $(\int_0^1 \Delta^*(s, x) ds)$ ,  $(\int_0^1 \Delta^*(s, x) ds)x$  e  $\int_0^1 s \Delta^*(s, x) ds$ . Il coefficiente associato a quest'ultima variabile identificherebbe la varianza dell'errore di misura.

Utilizzando i risultati in Battistin e Chesher (2007a e 2007b) è possibile dimostrare che l'approssimazione in (3.2) vale anche quando  $\Delta^*(s, x)$  è sostituito da  $\Delta(s, x)$ , dove:

$$\Delta(w_1, x) \equiv \frac{1}{2w_1(1-w_1)} \frac{\nabla_1^2 f_{Y_1, Y_2^*}[x + \ln w_1, x + \ln(1-w_1)]}{f_{ln(X)}[x]}.$$

Ne segue che la (3.2) e la struttura classica dell'errore implicano la seguente approssimazione (si veda Battistin e Chesher, 2007a e 2007b):

$$E_{W_1|ln(X)}[W_1|X] \simeq \beta_0 + \beta_1 x \quad (3.6)$$

$$-\beta_0 \sigma_1^2 \int_0^1 \Delta(s, x) ds \quad (3.7)$$

$$-\beta_1 \sigma_1^2 x \int_0^1 \Delta(s, x) ds \quad (3.8)$$

$$+\sigma_1^2 \int_0^1 s \Delta(s, x) ds. \quad (3.9)$$

### 3.4.1 Correzione della distorsione indotta da errore di misura

Dalle approssimazioni precedenti si vuole ricavare una regola per correggere la distorsione indotta da errore di misura. L'approssimazione (3.2) non fornisce una regola di correzione operativa in quanto coinvolge quantità definite a partire dai dati non affetti da errore di misura. Al contrario, l'approssimazione (3.6) è calcolabile a partire dai dati osservati. Una regressione di  $w_1$  su  $x$ ,  $(\int_0^1 \Delta(s, x) ds)$ ,  $(\int_0^1 \Delta(s, x) ds)x$  e  $\int_0^1 s\Delta(s, x) ds$  al livello di approssimazione  $o(\sigma_1^2)$ .

La regressione (3.6) coinvolge chiaramente vincoli non lineari nei parametri. Ad esempio, vale che

$$\frac{\beta_0}{\beta_1} = \frac{\beta_0 \sigma_1^2}{\beta_1 \sigma_1^2}.$$

Molto pragmaticamente, per mancanza di tempo, i risultati presentati di seguito si basano su stime non vincolate e si riferiscono soltanto alla specificazione *lineare*.

Si può vedere da questi grafici come la regressione OLS di  $w_1$  su  $x$ ,  $(\int_0^1 \Delta(s, x) ds)$ ,  $(\int_0^1 \Delta(s, x) ds)x$  e  $\int_0^1 s\Delta(s, x) ds$  sia sempre migliore della regressione OLS di  $w_1$  su  $x$ .

Modello lineare, $\beta_1$		
	Media	
	OLS	Aggiustato
$\rho = 0.95$	1.8689	1.1344
$\rho = 0.9$	2.7448	2.0638
$\rho = 0.85$	3.4821	2.8926
$\rho = 0.8$	4.1657	3.6358

Tabella 3.3: Media della distribuzione simulata di  $\beta_1$  al variare di  $\rho$  in un modello *lineare* per dati sporcati da errore di misura e dati "aggiustati", con tipo di errore di misura alternativo a quello di Lewbel (10000 simulazioni).

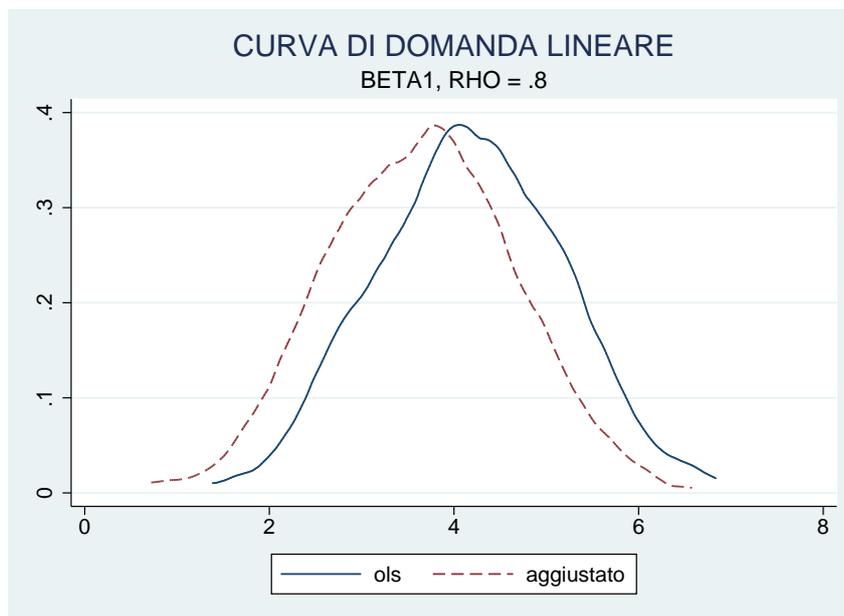


Figura 3.10: Distribuzione simulata delle stime standardizzate di  $\beta_1$  con il metodo "OLS" e con la regola di aggiustamento della distorsione, in un modello *lineare* con la struttura dell'errore alternativa a quella da Lewbel con  $\rho = 0.8$  (10000 simulazioni,  $\beta_1 = -0.1155$ ).

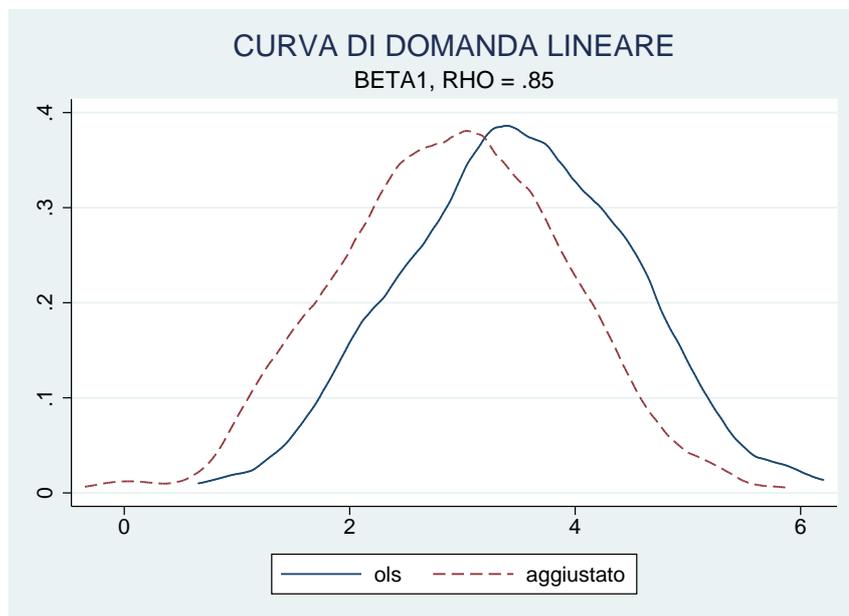


Figura 3.11: Distribuzione simulata delle stime standardizzate di  $\beta_1$  con il metodo "OLS" e con la regola di aggiustamento della distorsione, in un modello *lineare* con la struttura dell'errore alternativa a quella da Lewbel con  $\rho = 0.85$  (10000 simulazioni,  $\beta_1 = -0.1155$ ).

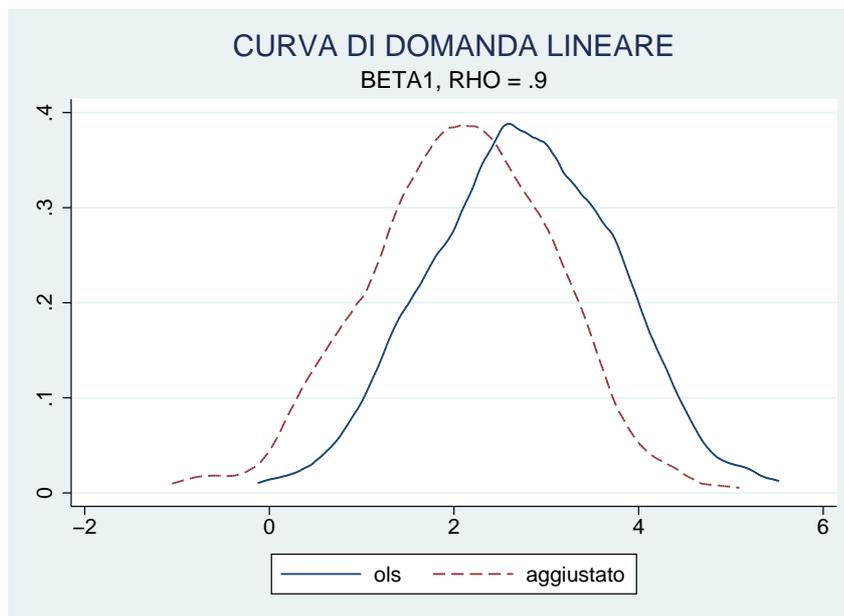


Figura 3.12: Distribuzione simulata delle stime standardizzate di  $\beta_1$  con il metodo "OLS" e con la regola di aggiustamento della distorsione, in un modello *lineare* con la struttura dell'errore alternativa a quella da Lewbel con  $\rho = 0.9$  (10000 simulazioni,  $\beta_1 = -0.1155$ ).

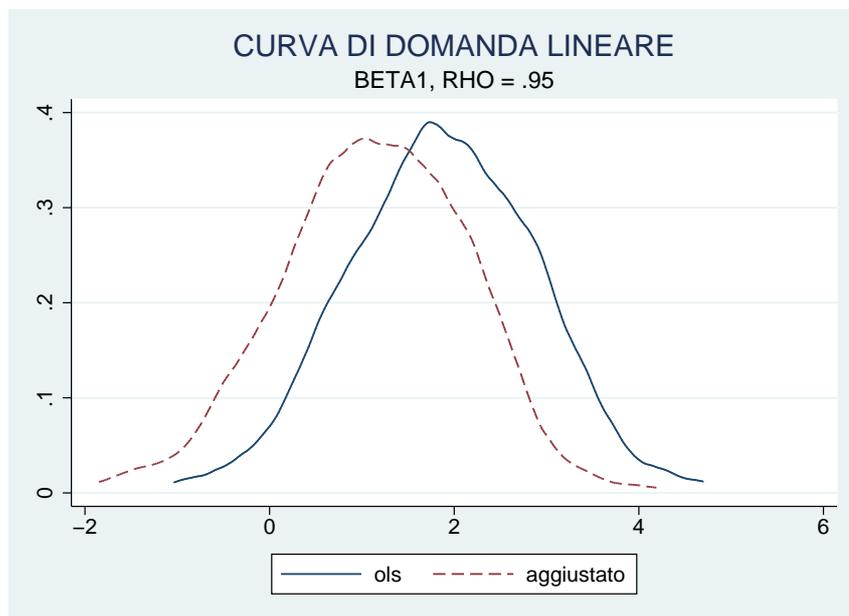


Figura 3.13: Distribuzione simulata delle stime standardizzate di  $\beta_1$  con il metodo "OLS" e con la regola di aggiustamento della distorsione, in un modello *lineare* con la struttura dell'errore alternativa a quella da Lewbel con  $\rho = 0.95$  (10000 simulazioni,  $\beta_1 = -0.1155$ ).

# Capitolo 4

## Conclusione

In questa tesi ho preso in considerazione la stima della curva di domanda per cibo sia con specificazione *lineare* che *quadratica*. Di queste due specificazioni ho studiato come si possano ottenere stime consistenti dei parametri nel caso in cui vi sia errore di misura in *spesa per cibo*. I metodi di stima utilizzati sono il metodo dei minimi quadrati ordinari e delle variabili strumentali. In particolare, ho presentato la procedura suggerita da Lewbel (1996), che però non tiene conto della natura delle variabili coinvolte nella curva di domanda. Per ovviare a questo problema, ho presentato un metodo alternativo a quello di Lewbel che è basato su una diversa struttura dell'errore che sia più coerente con la natura delle variabili coinvolte nella curva di domanda. Sui dati basati su questa nuova struttura dell'errore, quando vi è errore di misura "piccolo" nella *spesa per cibo*, ho calcolato la distorsione delle stime dei parametri e ho cercato una procedura che possa approssimativamente correggere questa distorsione indotta dall'errore di misura. Le simulazioni evidenziano che le stime ottenute con i valori "aggiustati" sono migliori che quelle ottenute con i dati "sporchi".



## Appendice A

# Informazioni sulla Banca d'Italia e l'Indagine sui Bilanci delle Famiglie Italiane

L'Indagine sui Bilanci delle Famiglie Italiane (IBFI), condotta da Banca d'Italia, prende la famiglia come unità di aggregazione delle risorse, definita come "l'insieme di persone conviventi che, indipendentemente da legami di parentela, provvede al soddisfacimento dei bisogni mediante la messa in comune di tutto o parte del reddito percepito dai suoi componenti". Tale indagine è condotta dalla Banca d'Italia a partire dagli anni Sessanta con l'obiettivo di raccogliere informazioni sui redditi e i risparmi delle famiglie italiane. Nel corso degli anni l'oggetto della rilevazione si è esteso includendo anche la ricchezza e altri aspetti inerenti i comportamenti economici e finanziari delle famiglie, come ad esempio l'uso dei mezzi di pagamento. I dati dell'indagine dal 1977 al 2004 sono contenuti nell'archivio storico<sup>1</sup>; tale archivio riporta informazioni sui seguenti aspetti:

- i caratteri anagrafici e lo status occupazionale dei singoli componenti

---

<sup>1</sup>Disponibile su [www.bancaditalia.it](http://www.bancaditalia.it)

## *60 Informazioni sulla Banca d'Italia e l'Indagine sui Bilanci delle Famiglie Italiane*

- i redditi (da lavoro dipendente e indipendente, da pensione, da trasferimenti, da capitale) percepiti dai componenti del nucleo familiare;
- i consumi (durevoli e non durevoli);
- i beni immobiliari abitati o posseduti dai membri della famiglia;
- le attività e le passività finanziarie detenute dalla famiglia.

Negli ultimi anni, circa 8.000 famiglie (24.000 individui) hanno composto il campione.

Poiché il questionario ha subito, nel corso del tempo, numerose modifiche, le variabili di maggior interesse sono state armonizzate per ottenere informazioni omogenee. Il file pubblico messo a disposizione dalla Banca d'Italia contiene pertanto informazione congiunta per reddito e spesa per campioni di famiglie intervistati in diverse occasioni a partire dal 1977. Per quanto riguarda il disegno di campionamento dell'indagine, questo prevede una procedura di selezione a due stadi. Nel primo stadio si procede all'estrazione dei comuni ove effettuare le interviste, suddivisi in base alla regione e alla classe di ampiezza demografica. In un secondo stadio, all'interno dei comuni selezionati, vengono estratte casualmente le famiglie da intervistare. Fino al 1987 l'indagine è stata effettuata sulla base di rilevazioni indipendenti nel tempo, mentre dall'indagine del 1989 è stata introdotta nel campione una quota di unità già intervistate nelle precedenti indagini (famiglie panel); ciò permette di seguire nel tempo i comportamenti di una stessa famiglia e quindi favorire l'analisi dell'evoluzione dei fenomeni oggetto d'interesse. Fino al 1987 l'indagine è stata fatta su base annuale (ad eccezione del 1985), in seguito è stata svolta ogni due anni, notando che nel 1997 l'indagine è stata rinviata all'anno successivo.

# Appendice B

## Distorsione delle stime con errore di misura piccolo

### B.1 Notazione

Riporto di seguito la notazione, che riprende in modo coerente quella presentata nell'introduzione ma le definizioni delle variabili non son più in livelli ma in logaritmi.

- $(Y_1^*, Y_2^*)$ : *spesa per cibo* e spesa residuale in logaritmi
- $X^* = \ln(e^{Y_1^*} + e^{Y_2^*})$ : logaritmo della spesa totale
- $W_1^* = \frac{e^{Y_1^*}}{e^{X^*}} = e^{(Y_1^* - X^*)}$ : quota di spesa per cibo
- $Y_1 = Y_1^* + \sigma_1 U_1$ : *spesa per cibo* osservata, dove  $E_{U_1}[U_1] = 0$ ,  $E_{U_1}[U_1^2] = 1$  e  $U_1 \perp [Y_1^*, Y_2^*]$
- $X = \ln(e^{Y_1} + e^{Y_2^*})$ : logaritmo della spesa totale osservata

### B.2 Relazione tra $E_{W_1^*|X^*}[W_1^*|x]$ e $E_{W_1|X}[W_1|x]$

Definita

$$E_{W_1|X}[W_1|x] = \int_0^1 s f_{W_1|X}[s|x] ds \quad (\text{B.1})$$

dove

$$f_{W_1|X}[w_1|x] = \frac{f_{W_1,X}[w_1, x]}{f_X[x]},$$

nelle sezioni che seguono si studia cosa succeda a (B.1) se  $\sigma_1 \rightarrow 0^+$ .

### B.3 Distribuzione di $(W_1^*, X^*)$

Trasformata di variabili casuali<sup>1</sup> da  $\mathbb{R} \times \mathbb{R}$  a  $(0, 1) \times \mathbb{R}$ :

$$\begin{bmatrix} y_1^* \\ y_2^* \end{bmatrix} \rightsquigarrow \begin{bmatrix} w_1^* = e^{(y_1^* - x^*)} \\ x^* = \ln(e^{y_1^*} + e^{y_2^*}) \end{bmatrix}$$

Trasformata inversa:

$$\begin{bmatrix} y_1^* = \ln w_1^* + x^* \\ y_2^* = \ln(e^{x^*} - e^{y_1^*}) \end{bmatrix} \Rightarrow \begin{bmatrix} y_1^* = \ln w_1^* + x^* \\ y_2^* = \ln[e^{x^*}(1 - w_1^*)] = x^* + \ln(1 - w_1^*) \end{bmatrix}$$

Jacobiano:

$$\left| \begin{array}{cc} \frac{\partial y_1^*}{\partial x^*} = 1 & \frac{\partial y_1^*}{\partial w_1^*} = \frac{1}{w_1^*} \\ \frac{\partial y_2^*}{\partial x^*} = 1 & \frac{\partial y_2^*}{\partial w_1^*} = \frac{-1}{(1-w_1^*)} \end{array} \right| = \frac{1}{(1-w_1^*)w_1^*}$$

Segue che<sup>2</sup>:

$$f_{W_1^*, X^*}[w_1^*, x^*] = \frac{1}{(1-w_1^*)w_1^*} f_{Y_1^*, Y_2^*}[x^* + \ln w_1^*, x^* + \ln(1 - w_1^*)]$$

### B.4 Distribuzione di $X^*$

Integrando l'ultima espressione rispetto a  $w_1^*$  si ottiene:

$$f_{X^*}[x^*] = \int_0^1 \frac{1}{(1-s)s} f_{Y_1^*, Y_2^*}[x^* + \ln s, x^* + \ln(1 - s)] ds \quad (\text{B.2})$$

<sup>1</sup>In quanto segue le quantità di interesse sono calcolate a partire dalla bivariata  $(Y_1^*, Y_2^*)$ .

<sup>2</sup>Si noti come la quota di spesa  $w_1^*$  debba essere strettamente compresa tra 0 e 1.

## B.5 Relazione tra $(Y_1^*, Y_2^*)$ e $(Y_1, Y_2^*)$

Mediante la seguente trasformata di variabili casuali

$$\begin{bmatrix} y_1^* \\ y_2^* \\ u_1 \end{bmatrix} \rightsquigarrow \begin{bmatrix} y_1 = y_1^* + \sigma_1 u_1 \\ y_2^* \\ u_1 \end{bmatrix}$$

si trova

$$f_{Y_1, Y_2^*}[y_1, y_2^*] = \int_{-\infty}^{+\infty} f_{Y_1^*, Y_2^*}[y_1 - \sigma_1 s, y_2^*] f_{U_1}[s] ds$$

## B.6 Distribuzione di $(W_1, X)$

Stesso ragionamento di Sezione B.3, combinando il risultato alla Sezione precedente per ottenere<sup>3</sup>

$$\begin{aligned} f_{W_1, X}[w_1, x] &= \frac{1}{(1-w_1)w_1} f_{Y_1, Y_2^*}[x + \ln w_1, x + \ln(1-w_1)] \\ &= \frac{1}{(1-w_1)w_1} \int_{-\infty}^{+\infty} f_{Y_1^*, Y_2^*}[x + \ln w_1 - \sigma_1 s, x + \ln(1-w_1)] f_{U_1}[s] ds \end{aligned}$$

da cui

$$f_{W_1, X}[w_1, x] = \frac{1}{(1-w_1)w_1} \int_{-\infty}^{+\infty} f_{Y_1^*, Y_2^*}[x + \ln w_1 - \sigma_1 s, x + \ln(1-w_1)] f_{U_1}[s] ds \quad (\text{B.3})$$

## B.7 Distribuzione di $X$

Integrando l'ultima espressione rispetto a  $w_1$  si ottiene:

$$f_X[x] = \int_0^1 \int_{-\infty}^{+\infty} \frac{1}{(1-t)t} f_{Y_1^*, Y_2^*}[x + \ln t - \sigma_1 s, x + \ln(1-t)] f_{U_1}[s] ds dt \quad (\text{B.4})$$

## B.8 Relazione tra $(W_1, X)$ e $(W_1^*, X^*)$

Ottenuta via sviluppo in serie di Taylor di (B.3) se  $\sigma_1 \rightarrow 0^+$ .

<sup>3</sup>La presenza di errore di misura in  $Y_1^*$  non cambia il supporto di  $(W_1, X)$  rispetto a quello di  $(W_1^*, X^*)$  (ovvero  $(0, 1) \times \mathbb{R}$ ).

### B.8.1 Sviluppo intorno a $\xi_0 \equiv (x + \ln w_1, x + \ln(1 - w_1))$

$$\begin{aligned} f_{Y_1^*, Y_2^*}[x + \ln w_1 - \sigma_1 s, x + \ln(1 - w_1)] &\simeq f_{Y_1^*, Y_2^*}[x + \ln w_1, x + \ln(1 - w_1)] \\ &\quad - \sigma_1 u_1 \nabla_1^1 f_{Y_1^*, Y_2^*}[x + \ln w_1, x + \ln(1 - w_1)] \\ &\quad + \frac{1}{2} \sigma_1^2 u_1^2 \nabla_1^2 f_{Y_1^*, Y_2^*}[x + \ln w_1, x + \ln(1 - w_1)] \end{aligned}$$

dove  $A \simeq B$  significa  $A = B + o(\sigma_1^2)$ .

### B.8.2 Sostituzione

$$f_{W_1, X}[w_1, x] \simeq f_{W_1^*, X^*}[w_1, x] + \frac{1}{2w_1(1-w_1)} \sigma_1^2 \nabla_1^2 f_{Y_1^*, Y_2^*}[x + \ln w_1, x + \ln(1 - w_1)] \quad (\text{B.5})$$

## B.9 Relazione tra $X$ e $X^*$

Ottenuta a partire dalla (B.4) integrando la (B.5)

$$f_X[x] \simeq f_{X^*}[x] + \frac{1}{2} \sigma_1^2 \int_0^1 \frac{1}{s(1-s)} \nabla_1^2 f_{Y_1^*, Y_2^*}[x + \ln s, x + \ln(1 - s)] ds \quad (\text{B.6})$$

## B.10 Relazione tra $W_1|X$ e $W_1^*|X^*$

Poichè:

$$f_{W_1|X}[w_1|x] = \frac{f_{W_1, X}[w_1, x]}{f_X[x]}$$

si può usare rispettivamente la (B.5) e la (B.6) come approssimazioni del numeratore e del denominatore, che assumerà forma

$$\frac{A + \sigma_1^2 B}{C + \sigma_1^2 D}$$

dove

$$\begin{aligned} A &\equiv f_{W_1^*, X^*}[w_1, x] \\ B &\equiv \frac{1}{2w_1(1-w_1)} \nabla_1^2 f_{Y_1^*, Y_2^*}[x + \ln w_1, x + \ln(1-w_1)] \\ C &\equiv f_{X^*}[x] \\ D &\equiv \frac{1}{2} \int_0^1 \frac{1}{s(1-s)} \nabla_1^2 f_{Y_1^*, Y_2^*}[x + \ln s, x + \ln(1-s)] ds \end{aligned}$$

Sfruttando la relazione

$$\frac{A + \sigma_1^2 B}{C + \sigma_1^2 D} \simeq \frac{A}{C} + \sigma_1^2 \frac{BC - AD}{C^2}$$

segue che

$$\begin{aligned} \frac{A}{C} &= \frac{f_{W_1^*, X^*}[w_1, x]}{f_{X^*}[x]} = f_{W_1^*|X^*}[w_1|x] \\ \frac{B}{C} &= \frac{\frac{1}{2w_1(1-w_1)} \nabla_1^2 f_{Y_1^*, Y_2^*}[x + \ln w_1, x + \ln(1-w_1)]}{f_{X^*}[x]} \\ \frac{AD}{C^2} &= f_{W_1^*|X^*}[w_1|x] \frac{\frac{1}{2} \int_0^1 \frac{1}{s(1-s)} \nabla_1^2 f_{Y_1^*, Y_2^*}[x + \ln s, x + \ln(1-s)] ds}{f_{X^*}[x]} \end{aligned}$$

da cui

$$\begin{aligned} f_{W_1|X}[w_1|x] &\simeq f_{W_1^*|X^*}[w_1|x] \left( 1 - \frac{\sigma_1^2 \int_0^1 \frac{1}{s(1-s)} \nabla_1^2 f_{Y_1^*, Y_2^*}[x + \ln s, x + \ln(1-s)] ds}{2 f_{X^*}[x]} \right) \\ &+ \frac{\sigma_1^2}{2w_1(1-w_1)} \frac{\nabla_1^2 f_{Y_1^*, Y_2^*}[x + \ln w_1, x + \ln(1-w_1)]}{f_{X^*}[x]} \end{aligned}$$

Segue che

$$\boxed{f_{W_1|X}[w_1|x] \simeq f_{W_1^*|X^*}[w_1|x] \left( 1 - \sigma_1^2 \int_0^1 \Delta(s, x) ds \right) + \sigma_1^2 \Delta(w_1, x)} \quad (\text{B.7})$$

dove utilizzando la (B.2)

$$\begin{aligned} \Delta^*(w_1, x) &\equiv \frac{1}{2w_1(1-w_1)} \frac{\nabla_1^2 f_{Y_1^*, Y_2^*}[x + \ln w_1, x + \ln(1-w_1)]}{f_{X^*}[x]} \\ &= \frac{1}{2w_1(1-w_1)} \frac{\nabla_1^2 f_{Y_1^*, Y_2^*}[x + \ln w_1, x + \ln(1-w_1)]}{\int_0^1 \frac{1}{(1-s)s} f_{Y_1^*, Y_2^*}[x + \ln s, x + \ln(1-s)] ds} \end{aligned}$$

**B.11 Soluzione**

Utilizzando la (B.7) abbiamo

$$E_{W_1|X}[W_1|x] \simeq (1 - \sigma_1^2 \int_0^1 \Delta^*(s, x) ds) E_{W_1^*|X^*}[W_1^*|x] + \sigma_1^2 \int_0^1 s \Delta^*(s, x) ds$$

Da cui segue che  $E_{w_1|X}[W_1|x] \simeq E_{W_1^*|X^*}[W_1^*|X]$  se  $\sigma_1 \rightarrow 0^+$ .

## Appendice C

# Stima della distorsione indotta da errore di misura

In quanto segue assumerò  $(y_1, y_2^*)$  distribuite come una variabile casuale normale bivariata, ovvero:

$$\begin{bmatrix} y_1 \\ y_2^* \end{bmatrix} \sim N \left( \begin{bmatrix} \mu_1 \\ \mu_2^* \end{bmatrix}; \begin{bmatrix} \tau_1^2 & \tau_{12}^* \\ \tau_{12}^* & \tau_2^{*2} \end{bmatrix} \right)$$

dove  $\tau_{12}^* = \gamma\tau_1\tau_2^*$  e  $\gamma$  è il coefficiente di correlazione tra  $y_1$  e  $y_2^*$ .

La funzione di densità congiunta è:

$$f_{Y_1, Y_2^*}(y_1, y_2^*) = ce^{-\frac{1}{2(1-\gamma^2)} \left[ \left[ \frac{y_1 - \mu_1}{\tau_1} \right]^2 + \left[ \frac{y_2^* - \mu_2^*}{\tau_2^*} \right]^2 - 2\gamma \frac{y_1 - \mu_1}{\tau_1} \frac{y_2^* - \mu_2^*}{\tau_2^*} \right]}$$

dove  $c = \frac{1}{2\pi\tau_1\tau_2^*\sqrt{(1-\gamma^2)}}$ .

La densità marginale e condizionata sono rispettivamente:

$$f_{Y_2^*}(y_2^*) = \frac{1}{\tau_2^*} \phi\left(\frac{y_2^* - \mu_2^*}{\tau_2^*}\right)$$

e

$$f_{Y_1|Y_2^*}(y_1|y_2^*) = \frac{1}{\tau_1\sqrt{1-\gamma^2}} \phi\left(\frac{y_1 - \mu_1 - \gamma\frac{\tau_1}{\tau_2^*}(y_2^* - \mu_2^*)}{\tau_1\sqrt{1-\gamma^2}}\right)$$

dove  $\phi()$  rappresenta la densità di una variabile casuale normale standard.

L'assunzione di normalità bivariata di  $(y_1, y_2^*)$  è sufficiente per ottenere una stima di:

$$\Delta(w_1, x) \equiv \frac{1}{2w_1(1-w_1)} \frac{\nabla_1^2 f_{Y_1, Y_2^*}[x + \ln w_1, x + \ln(1-w_1)]}{f_{\ln(X)}[x]}$$

a partire dai dati osservati. In particolare:

- per quanto riguarda il numeratore, vale che

$$\begin{aligned} \nabla_1^2 f_{Y_1, Y_2^*}(y_1, y_2^*) &= \nabla_1^2 f_{Y_1|Y_2^*}(y_1|y_2^*) f_{Y_2^*}(y_2^*), \\ &= f_{Y_2^*}(y_2^*) \nabla_1^2 f_{Y_1|Y_2^*}(y_1|y_2^*) \end{aligned} \quad (C.1)$$

dove si ha

$$\begin{aligned} \nabla_1^2 f_{Y_1|Y_2^*}(y_1|y_2^*) &= \left( \frac{1}{\tau_1 \sqrt{1-\gamma^2}} \right)^2 f_{Y_1|Y_2^*}(y_1|y_2^*) \left( \left[ \frac{y_1 - \mu_1 - \gamma \frac{\tau_1}{\tau_2^*} (y_2^* - \mu_2^*)}{\tau_1 \sqrt{1-\gamma^2}} \right]^2 - 1 \right), \\ &= \left( \frac{1}{\tau_1 \sqrt{1-\gamma^2}} \right)^3 \\ &\quad \left( \left[ \frac{y_1 - \mu_1 - \gamma \frac{\tau_1}{\tau_2^*} (y_2^* - \mu_2^*)}{\tau_1 \sqrt{1-\gamma^2}} \right]^2 - 1 \right) \\ &\quad \phi \left( \frac{y_1 - \mu_1 - \gamma \frac{\tau_1}{\tau_2^*} (y_2^* - \mu_2^*)}{\tau_1 \sqrt{1-\gamma^2}} \right). \end{aligned}$$

Ne segue che l'espressione per il numeratore risulta essere:

$$\begin{aligned} \nabla_1^2 f_{Y_1, Y_2^*}(y_1, y_2^*) &= \frac{1}{\tau_2^*} \phi \left( \frac{y_2^* - \mu_2^*}{\tau_2^*} \right) \\ &\quad \left( \frac{1}{\tau_1 \sqrt{1-\gamma^2}} \right)^3 \\ &\quad \left( \left[ \frac{y_1 - \mu_1 - \gamma \frac{\tau_1}{\tau_2^*} (y_2^* - \mu_2^*)}{\tau_1 \sqrt{1-\gamma^2}} \right]^2 - 1 \right) \\ &\quad \phi \left( \frac{y_1 - \mu_1 - \gamma \frac{\tau_1}{\tau_2^*} (y_2^* - \mu_2^*)}{\tau_1 \sqrt{1-\gamma^2}} \right); \end{aligned}$$

- il denominatore coinvolge la densità della spesa totale in logaritmi, che può essere stimata a partire dai dati osservati mediante una stima non parametrica. Assumendo che tale densità sia approssimativamente normale (si veda l'evidenza fornita in Battistin, Blundell e

Lewbel, 2007) vale che:

$$\hat{f}_{\ln(X)}(x) = \frac{1}{\tau} \phi\left(\frac{x - \mu}{\tau}\right)$$

Quindi, la stima di  $\Delta$  risulta essere:

$$\hat{\Delta}(w_1, x) \equiv \frac{1}{2w_1(1-w_1)} \frac{\left[ \frac{1}{\hat{\tau}_2^*} \phi\left(\frac{x + \ln(1-w_1) - \hat{\mu}_2^*}{\hat{\tau}_2^*}\right) \right] \left[ \left( \frac{1}{\hat{\tau}_1 \sqrt{1-\hat{\gamma}^2}} \right)^3 \left( \left[ \frac{y_1 - \hat{\mu}_1 - \hat{\gamma} \frac{\hat{\tau}_1}{\hat{\tau}_2^*} (y_2^* - \hat{\mu}_2^*)}{\hat{\tau}_1 \sqrt{1-\hat{\gamma}^2}} \right]^2 - 1 \right) \phi\left(\frac{y_1 - \hat{\mu}_1 - \hat{\gamma} \frac{\hat{\tau}_1}{\hat{\tau}_2^*} (y_2^* - \hat{\mu}_2^*)}{\hat{\tau}_1 \sqrt{1-\hat{\gamma}^2}}\right) \right]}{\frac{1}{\hat{\tau}} \phi\left(\frac{x - \hat{\mu}}{\hat{\tau}}\right)}.$$

Le quantità richieste per ottenere l'approssimazione di  $E_{w_1|\ln(X)}[w_1|x]$ , spiegata nel paragrafo 3.4, sono  $(1 - \sigma_1^2 \int_0^1 \Delta(s, x) ds)$  e  $\int_0^1 s \Delta(s, x) ds$ .

La quantità  $\int_0^1 \Delta(s, x) ds$  è calcolata a partire dai dati come una somma delle stime di  $\Delta$  calcolate in piccoli intervalli:

$$\sum_{k \in \{1, \dots, K\}} \hat{\Delta}(k, x).$$

Analogamente, la quantità  $\int_0^1 s \Delta(s, x) ds$  è calcolata nei dati come

$$\sum_{k \in \{1, \dots, K\}} k \hat{\Delta}(k, x).$$

Le classi  $k = 1, \dots, K$  sono ottenute discretizzando l'intervallo  $[0, 1]$  mediante 20 intervalli di ampiezza 0.05.



# Appendice D

## Grafici

Di seguito sono riportati i grafici per la specificazione della curva di domanda *quadratica*, sia nel caso di struttura dell'errore di Lewbel sia nel caso alternativo, quando nel testo non fosse stato possibile vedere chiaramente le differenze tra i tre metodi di stima utilizzati: "OLS", "IV" e "Lewbel".

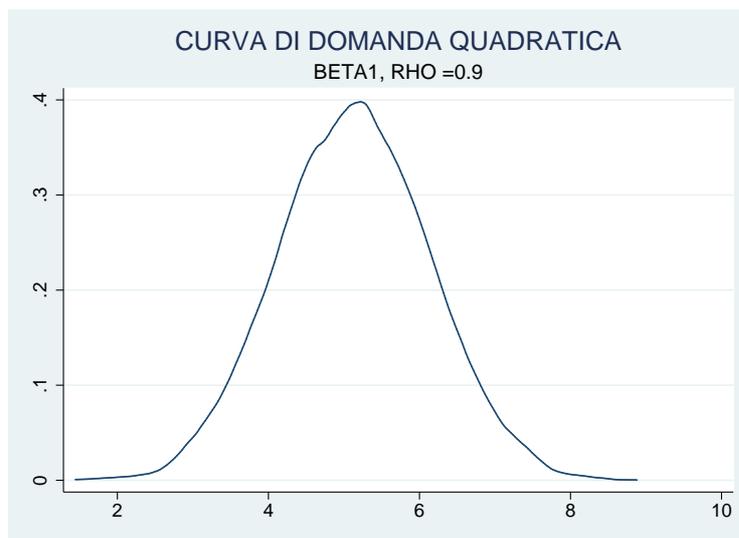


Figura D.1: Distribuzione simulata delle stime standardizzate di  $\beta_1$  con il metodo "OLS", in un modello *quadratico* con struttura dell'errore di Lewbel con  $\rho = 0.9$  (10000 simulazioni,  $\beta_1 = 0.1572$ ).

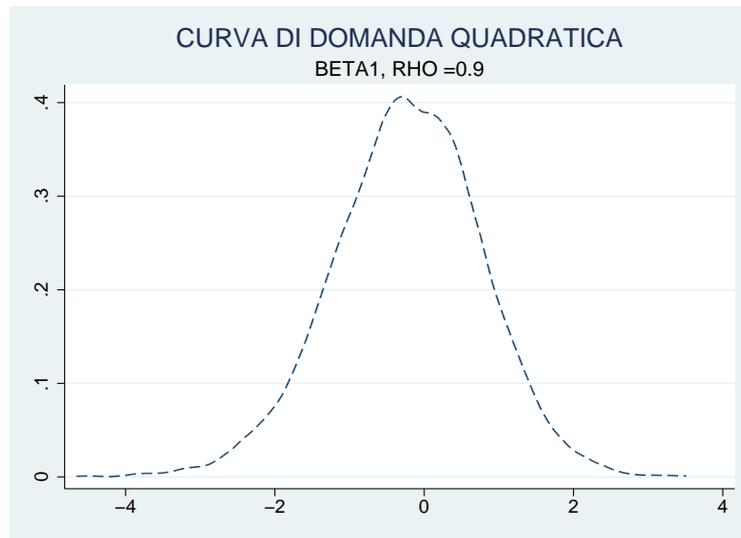


Figura D.2: Distribuzione simulata delle stime standardizzate di  $\beta_1$  con il metodo "IV", in un modello *quadratico* con struttura dell'errore di Lewbel con  $\rho = 0.9$  (10000 simulazioni,  $\beta_1 = 0.1572$ ).

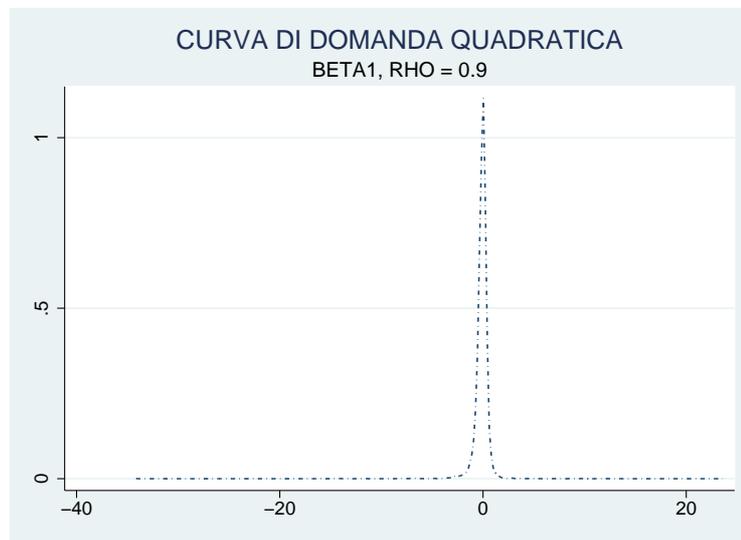


Figura D.3: Distribuzione simulata delle stime standardizzate di  $\beta_1$  con il metodo "Lew", in un modello *quadratico* con struttura dell'errore di Lewbel con  $\rho = 0.9$  (10000 simulazioni,  $\beta_1 = 0.1572$ ).

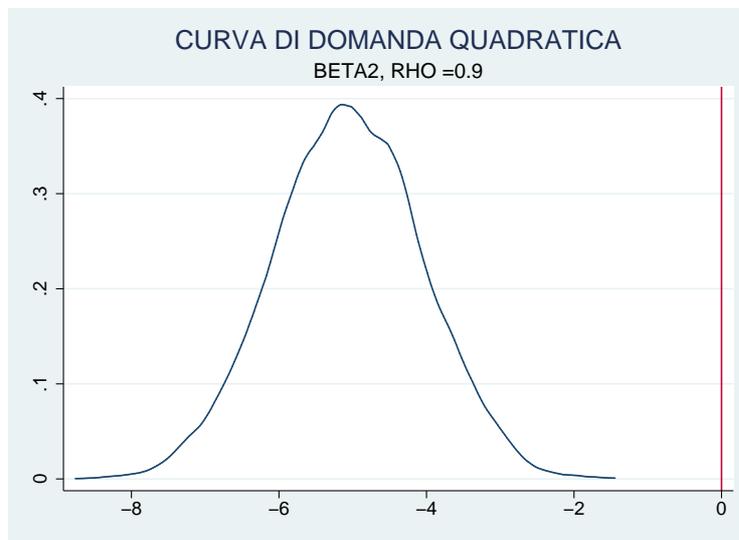


Figura D.4: Distribuzione simulata delle stime standardizzate di  $\beta_2$  con il metodo "OLS", in un modello *quadratico* con struttura dell'errore di Lewbel con  $\rho = 0.9$  (10000 simulazioni,  $\beta_1 = -0.0133$ ).

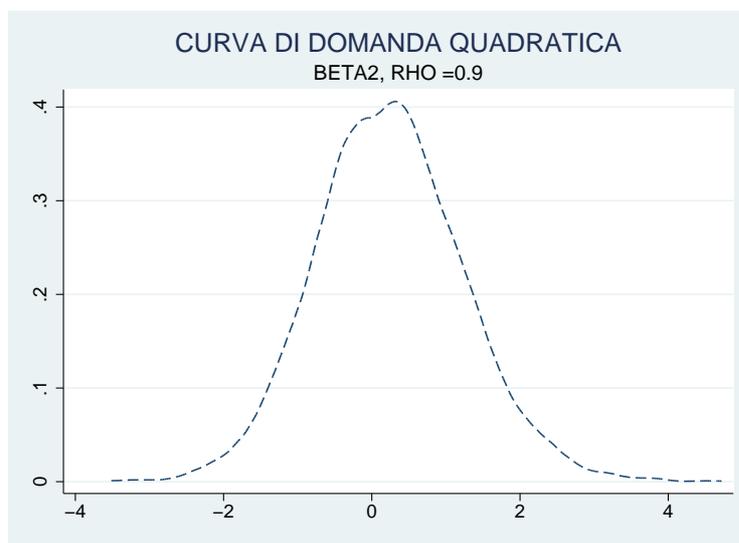


Figura D.5: Distribuzione simulata delle stime standardizzate di  $\beta_2$  con il metodo "IV", in un modello *quadratico* con struttura dell'errore di Lewbel con  $\rho = 0.9$  (10000 simulazioni,  $\beta_1 = -0.0133$ ).

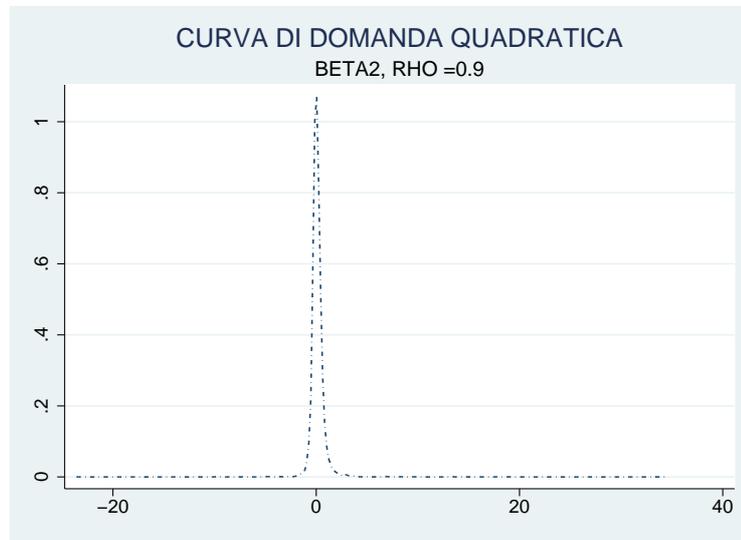


Figura D.6: Distribuzione simulata delle stime standardizzate di  $\beta_2$  con il metodo "Lew", in un modello *quadratico* con struttura dell'errore di Lewbel con  $\rho = 0.9$  (10000 simulazioni,  $\beta_1 = -0.0133$ ).

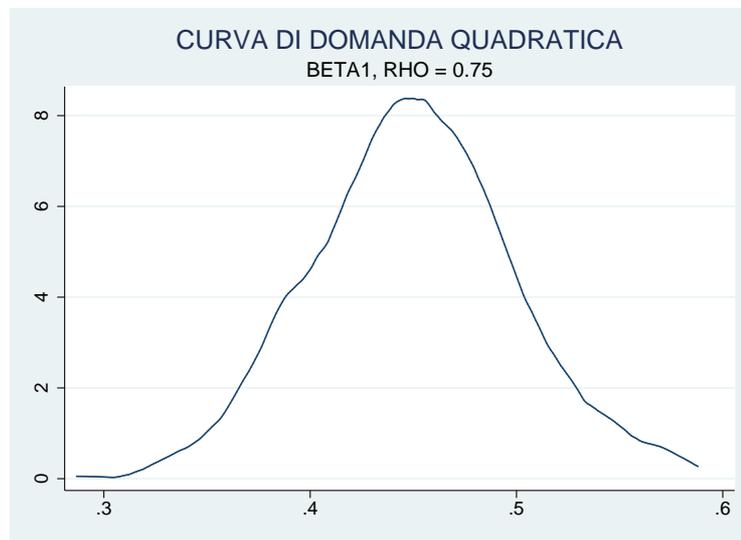


Figura D.7: Distribuzione simulata delle stime standardizzate di  $\beta_1$  con il metodo "OLS", in un modello *quadratico* con struttura dell'errore di Lewbel con  $\rho = 0.75$  (10000 simulazioni,  $\beta_1 = 0.1572$ ).

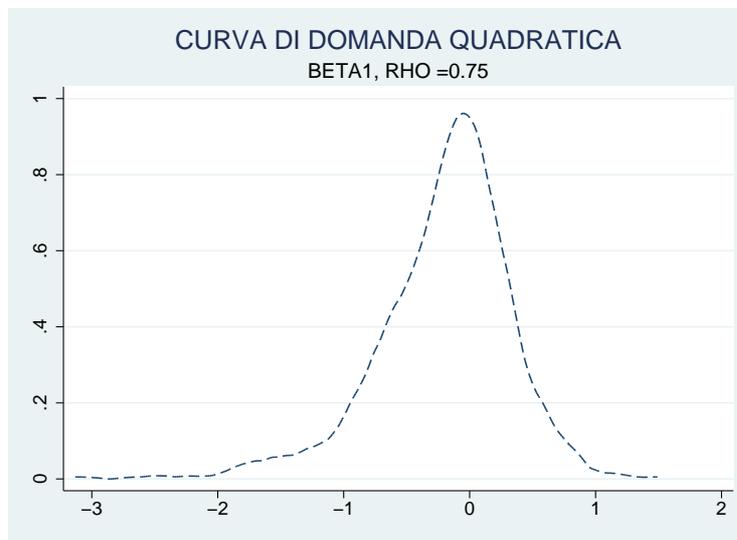


Figura D.8: Distribuzione simulata delle stime standardizzate di  $\beta_1$  con il metodo "IV", in un modello *quadratico* con struttura dell'errore di Lewbel con  $\rho = 0.75$  (10000 simulazioni,  $\beta_1 = 0.1572$ ).

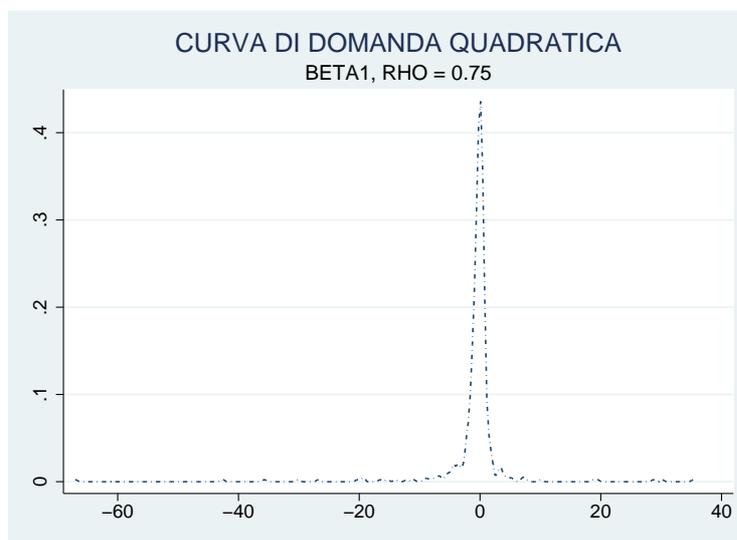


Figura D.9: Distribuzione simulata delle stime standardizzate di  $\beta_1$  con il metodo "Lew", in un modello *quadratico* con struttura dell'errore di Lewbel con  $\rho = 0.75$  (10000 simulazioni,  $\beta_1 = 0.1572$ ).

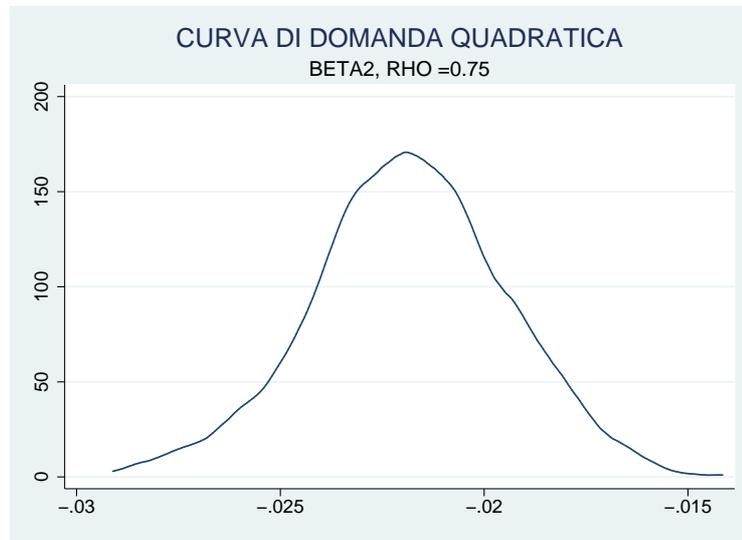


Figura D.10: Distribuzione simulata delle stime standardizzate di  $\beta_2$  con il metodo "OLS", in un modello *quadratico* con struttura dell'errore di Lewbel con  $\rho = 0.75$  (10000 simulazioni,  $\beta_1 = -0.0133$ ).

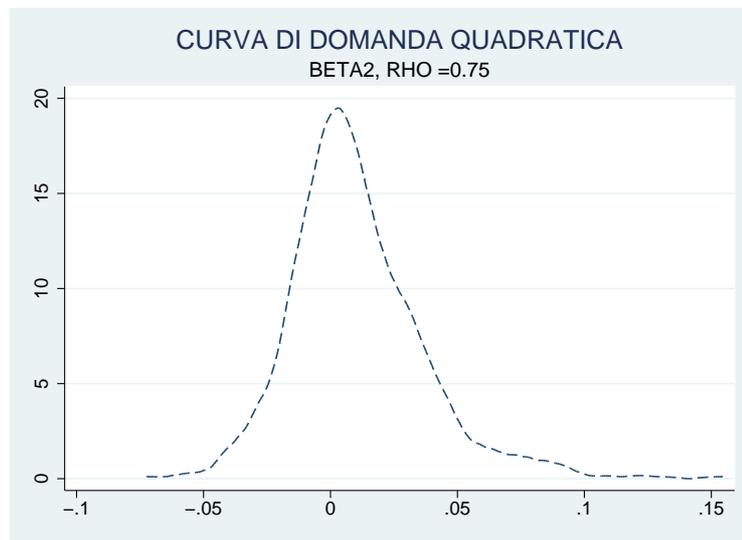


Figura D.11: Distribuzione simulata delle stime standardizzate di  $\beta_2$  con il metodo "IV", in un modello *quadratico* con struttura dell'errore di Lewbel con  $\rho = 0.75$  (10000 simulazioni,  $\beta_1 = -0.0133$ ).

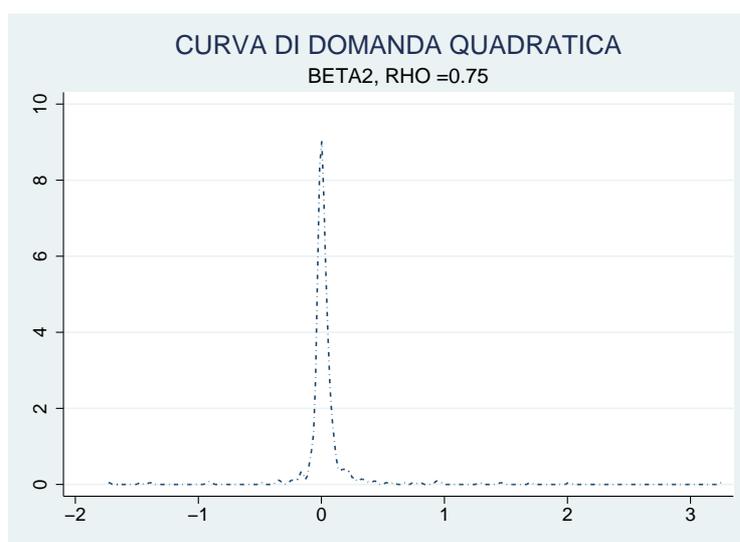


Figura D.12: Distribuzione simulata delle stime standardizzate di  $\beta_2$  con il metodo "Lew", in un modello *quadratico* con struttura dell'errore di Lewbel con  $\rho = 0.75$  (10000 simulazioni,  $\beta_1 = -0.0133$ ).

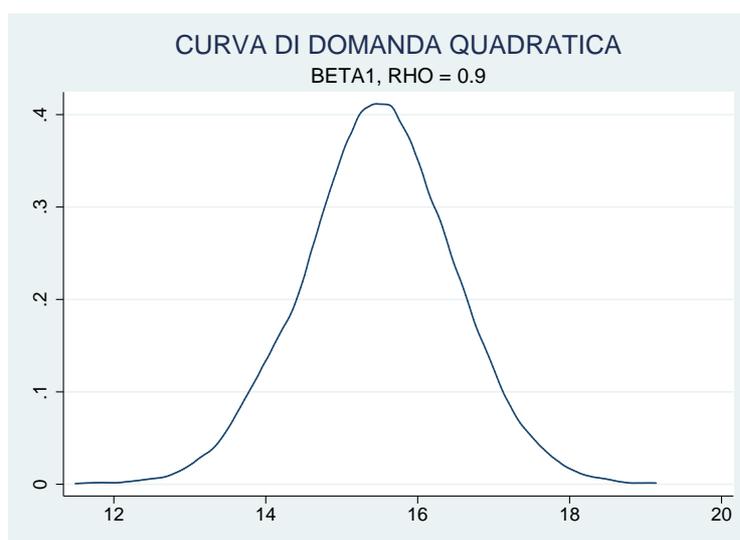


Figura D.13: Distribuzione simulata delle stime standardizzate di  $\beta_1$  con il metodo "OLS", in un modello *quadratico* con struttura dell'errore alternativa a quella di Lewbel con  $\rho = 0.9$  (10000 simulazioni,  $\beta_1 = 0.1572$ ).

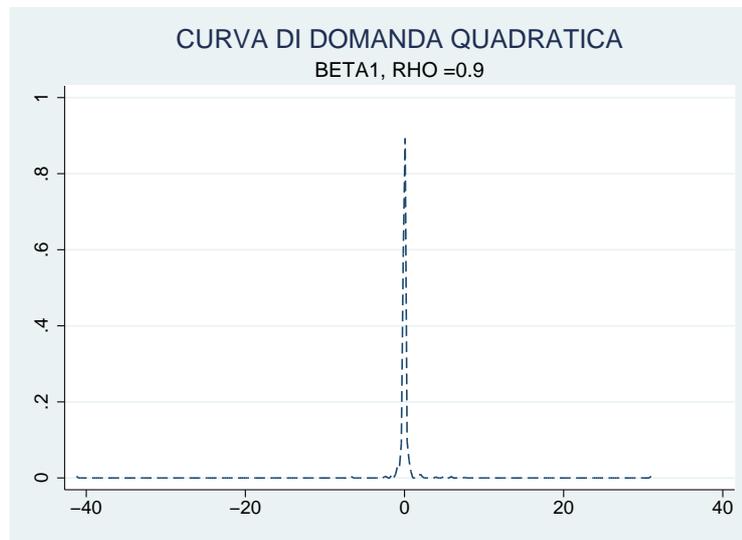


Figura D.14: Distribuzione simulata delle stime standardizzate di  $\beta_1$  con il metodo "IV", in un modello *quadratico* con struttura dell'errore alternativa a quella di Lewbel con  $\rho = 0.9$  (10000 simulazioni,  $\beta_1 = 0.1572$ ).

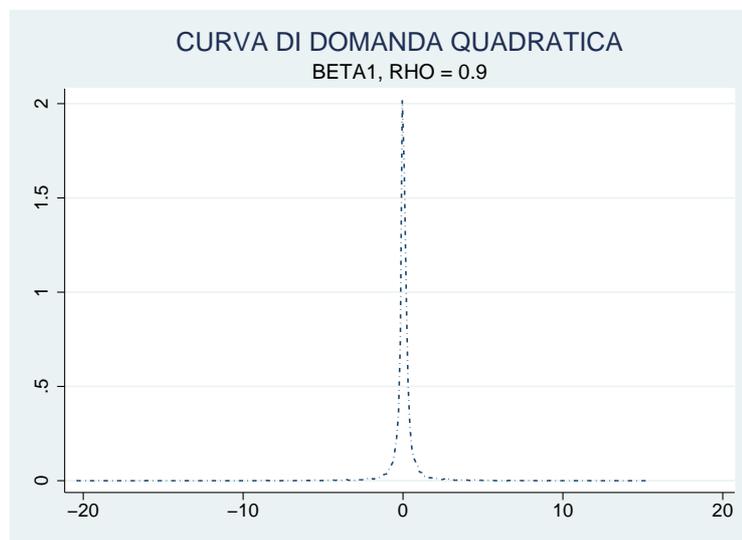


Figura D.15: Distribuzione simulata delle stime standardizzate di  $\beta_1$  con il metodo "Lew", in un modello *quadratico* con struttura dell'errore alternativa a quella di Lewbel con  $\rho = 0.9$  (10000 simulazioni,  $\beta_1 = 0.1572$ ).

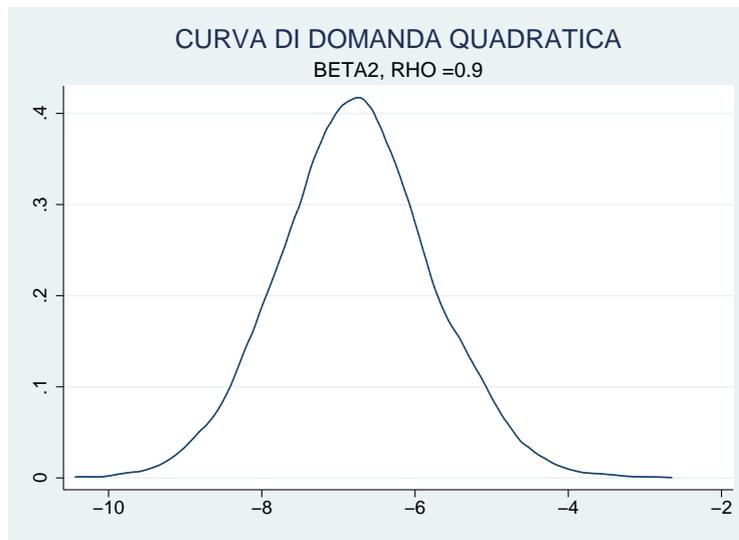


Figura D.16: Distribuzione simulata delle stime standardizzate di  $\beta_2$  con il metodo "OLS", in un modello *quadratico* con struttura dell'errore alternativa a quella di Lewbel con  $\rho = 0.9$  (10000 simulazioni,  $\beta_1 = -0.0133$ ).

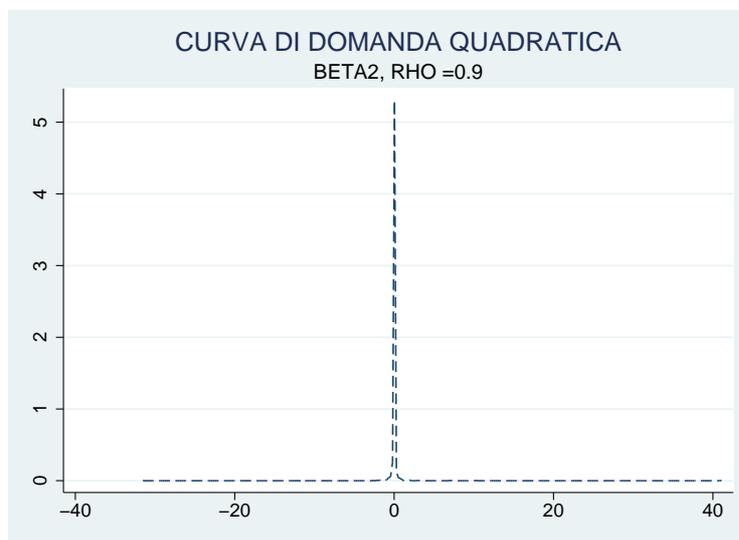


Figura D.17: Distribuzione simulata delle stime standardizzate di  $\beta_2$  con il metodo "IV", in un modello *quadratico* con struttura dell'errore alternativa a quella di Lewbel con  $\rho = 0.9$  (10000 simulazioni,  $\beta_1 = -0.0133$ ).

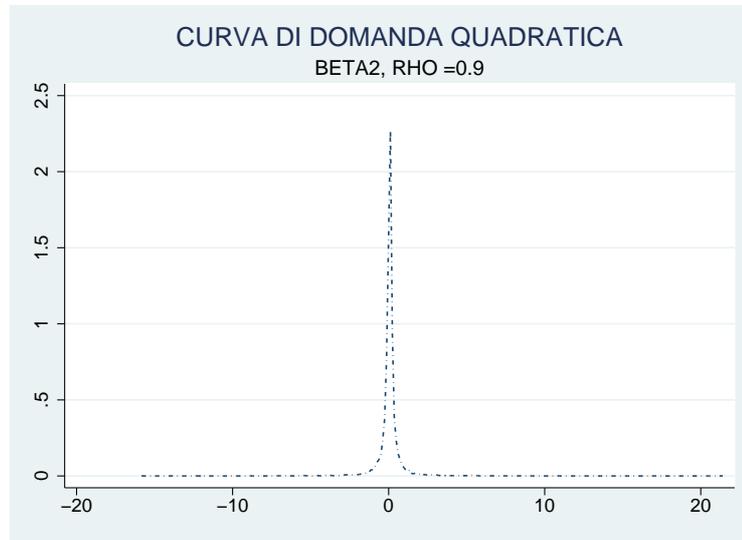


Figura D.18: Distribuzione simulata delle stime standardizzate di  $\beta_2$  con il metodo "Lew", in un modello *quadratico* con struttura dell'errore alternativa a quella di Lewbel con  $\rho = 0.9$  (10000 simulazioni,  $\beta_1 = -0.0133$ ).

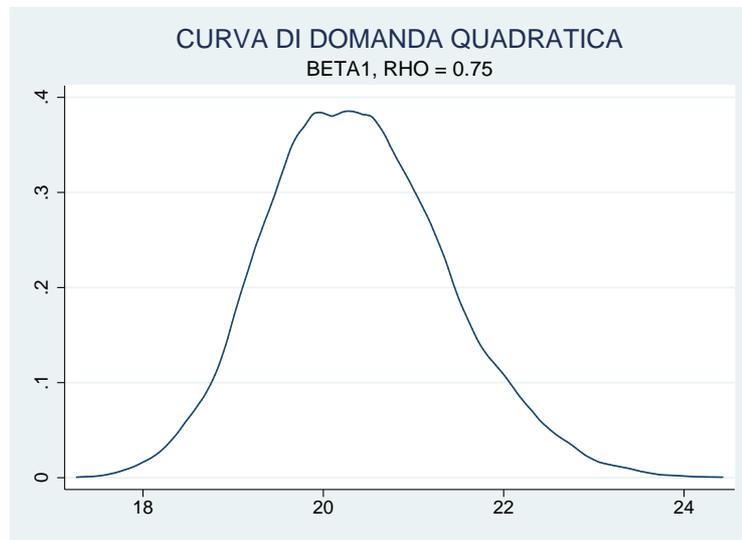


Figura D.19: Distribuzione simulata delle stime standardizzate di  $\beta_1$  con il metodo "OLS", in un modello *quadratico* con struttura dell'errore alternativa a quella di Lewbel con  $\rho = 0.75$  (10000 simulazioni,  $\beta_1 = 0.1572$ ).

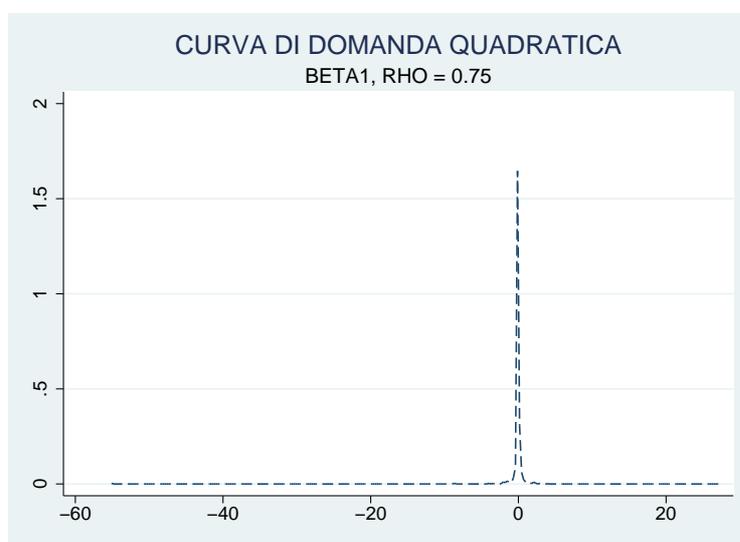


Figura D.20: Distribuzione simulata delle stime standardizzate di  $\beta_1$  con il metodo "IV", in un modello *quadratico* con struttura dell'errore alternativa a quella di Lewbel con  $\rho = 0.75$  (10000 simulazioni,  $\beta_1 = 0.1572$ ).

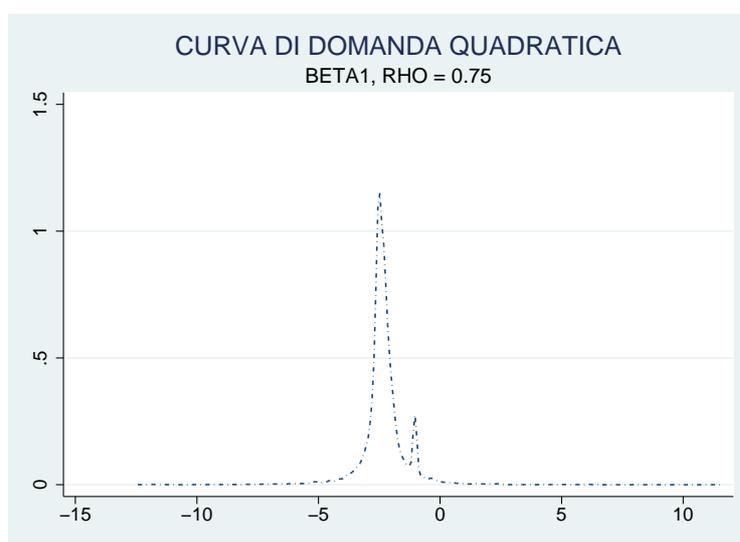


Figura D.21: Distribuzione simulata delle stime standardizzate di  $\beta_1$  con il metodo "Lew", in un modello *quadratico* con struttura dell'errore alternativa a quella di Lewbel con  $\rho = 0.75$  (10000 simulazioni,  $\beta_1 = 0.1572$ ).

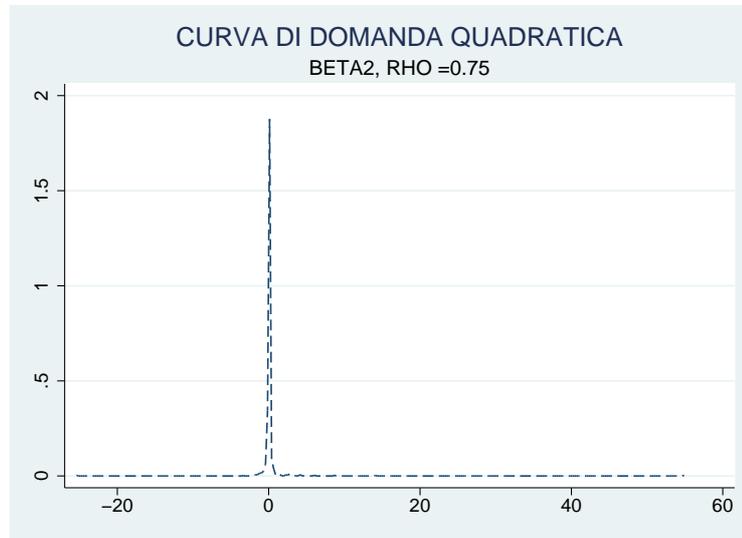


Figura D.23: Distribuzione simulata delle stime standardizzate di  $\beta_2$  con il metodo "IV", in un modello *quadratico* con struttura dell'errore alternativa a quella di Lewbel con  $\rho = 0.75$  (10000 simulazioni,  $\beta_1 = -0.0133$ ).

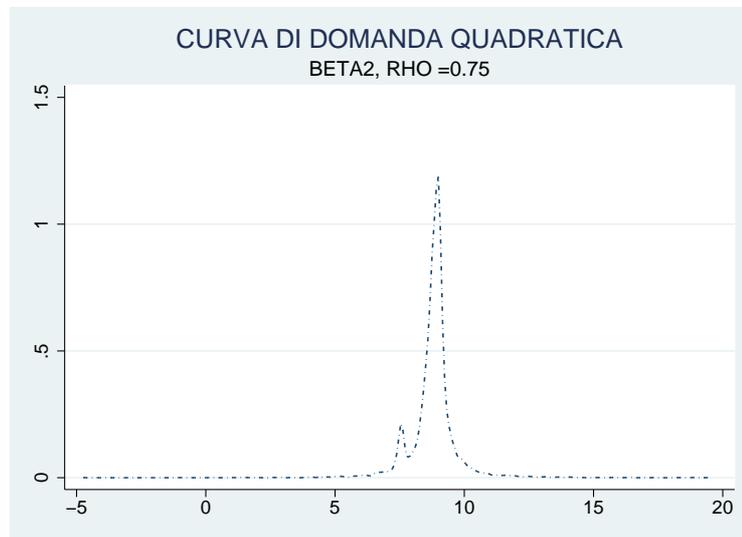


Figura D.24: Distribuzione simulata delle stime standardizzate di  $\beta_2$  con il metodo "Lew", in un modello *quadratico* con struttura dell'errore alternativa a quella di Lewbel con  $\rho = 0.75$  (10000 simulazioni,  $\beta_1 = -0.0133$ ).

# Bibliografia

- [1] Chesher Andrew (1991). The effect of measurement error. *Biometrika*, Vol.78 No.3, 451-462.
- [2] Arthur Lewbel (1996). Demand estimation with expenditure errors on the left and the right hand side. *The review of economics and statistics*, Vol.78 No.3, 451-462.
- [3] Erich Battistin & Andrew Chesher (2007a). Identification and estimation of engel curves with errors in expenditure data, manoscritto.
- [4] Erich Battistin & Andrew Chesher (2007b). Treatment Effect Estimation with Covariate Measurement Error, manoscritto.
- [5] Orazio Attanasio, Erich Battistin, Alice Mesnard (2007). Food Engel Curves.
- [6] Arthur Lewbel (2006). Engel curves. *Entry for the new palgrave dictionary of economics*, 2nd edition.
- [7] James Banks, Richard Blundell, Arthur Lewbel (1997). Quadratic engel curve and consumer demand. *The review of economics and statistics*, Vol. 79, No.4, 527-539.
- [8] Richard Blundell, Alan Duncan, Krishna Pendakur (1998). Semi-parametric estimation and consumer demand. *Journal of applied econometrics*, Vol.13, No.5, 435-461.

- [9] Holbrook Working (1943). Statistical laws of family expenditure. *Journal of the American Statistical Association*, Vol.38, No.221, 43-56.
- [10] Leser (1963). Forms of Engel functions. *Econometrica*, Vol.31, No.4, 694-703.
- [11] John Bound, Charles Brown e Nancy Mathiowetz (2001), Measurement error in survey data. *Handbook of Econometrics*, Vol.5, 3705-3843.

# Ringraziamenti

Non mi sembra ancora vero che ho finito di scrivere la tesi della mia laurea specialistica, il che significa: fine università, fine vita padovana, stop agli spriz, stop allo shopping da stress esame ma via alla ricerca di un lavoro, ritorno alla (vivacissima) vita mantovana e via via alla realizzazione di un futuro tutto ancora da definire.

Prima però che questo percorso da studentessa universitaria si concluda definitivamente devo effettuare dei dovuti ringraziamenti:

come è giusto e naturale ringrazio per prima la mia famiglia: mamma Patrizia e papà Maurizio per il sostegno morale (ed economico) costante durante la mia carriera scolastica anche in periodi di crisi; Federica e Francesca, le mie sorelle, perché insieme ai nostri genitori non hanno mai smesso di credere in me e nelle mie capacità; nonni e zii che e in ogni occasione hanno avuto parole dolci e complimenti da farmi.

Stefano, il mio Stéfano (non Stèfano come dicono in veneto), che mi ha sopportata quando “scleravo”, sostenuta e “rallegrata” quando piangevo, aiutata nelle difficoltà, insomma che mi è sempre stato vicino.

Le amiche di 10 anni fa, del presente e del futuro: Nico, Lara, Bisa, Emy, Ele, Cesca, Pelos che non smetterò mai di ringraziare per le bellissime serate con chiacchierate e risate che abbiamo condiviso. Gli amici di San Biagio e Bagnolo per le grigliate, le cene, i compleanni, le uscite.

Tornando invece alla vita padovana ci sono altre persone a cui devo rivolgere la mia gratitudine: “First of all”, come direbbe lui con i suoi inglesismi qua e là, il prof Erich Battistin, o meglio semplicemente Erich: lo ringrazio

profondamente per avermi seguita durante questo mio lavoro per ben otto mesi con simpatia e dedizione, disponibilità nello spiegarmi e rispiegarmi i concetti a me oscuri e per la professionalità dimostrata. Le mie coinquiline: Francesca, Giovanna, Clara, Elisa, Marika, Cristina e Manu per le lunghe chiacchierate, le pizze a domicilio, le sere sui divani, "Un posto al sole" durante la cena, la sopportazione durante i miei periodi di esame. Per ultimi, ma per tutta l'importanza che riservano nel mio cuore, i miei colleghi, compagni di studio, feste, bevute, spriz, risate, canti nonché miei carissimi amici: Fede, Manu, Fabio, Riccardo, Giulia, Petit, Eris, Roberto, Ombretta, Silvia e Angela. Grazie della collaborazione e per aver reso questi ultimi due anni davvero speciali.

Vi voglio bene!!!