BANCA D'ITALIA

Temi di discussione

del Servizio Studi

Stime regionali con "pochi dati": analisi e simulazione di stimatori alternativi per investimenti, occupazione e fatturato nelle imprese manifatturiere

di Riccardo Cesari e L. Federico Signorini

Numero 152 - Marzo 1991

BANCA D'ITALIA

Temi di discussione

del Servizio Studi

Stime regionali con "pochi dati": analisi e simulazione di stimatori alternativi per investimenti, occupazione e fatturato nelle imprese manifatturiere

di Riccardo Cesari e L. Federico Signorini

Numero 152 - Marzo 1991

SOMMARIO

L'analisi economica a livello regionale o sub-regionale è spesso ostacolata dalla scarsa disponibilità e/o affidabilità di informazioni territorialmente disaggregate. Una delle strategie per estrarre "quanta più informazione possibile" dalle statistiche esistenti è quella dei cosiddetti stimatori sintetici (Steinberg (1979)), che forniscono stime regionali a varianza contenuta, benchè distorte. Il presente lavoro verifica tale strategia con un esperimento di tipo Montecarlo condotto sui dati dell'indagine della Banca d'Italia presso le imprese industriali. La simulazione consente di valutare la performance di vari tipi di stimatori classici e sintetici, in termini di trade-off tra distorsione e variabilità. Si propone, inoltre, un approccio generale alla scelta dello stimatore ottimale a partire da un'ampia classe di funzioni di perdita che hanno come argomenti la distorsione e la varianza degli stimatori.

INDICE

1. Introduzione (*)

Chiunque si occupi di analisi economica a livello regionale o sub-regionale deve spesso fare i conti con un "mercato" della informazione statistica caratterizzato da un forte squilibrio tra domanda e offerta. Infatti, in un sistema economico che destina una quota crescente di risorse all'elaborazione dell'informazione, i soggetti più vari (aziende, enti pubblici, organizzazioni sociali, istituti di ricerca) esprimono una domanda crescente di materia prima informativa in campo economico. Ma tra un censimento industriale e l'altro, l'offerta di informazione economica di tipo quantitativo si basa in massima parte su indagini per campione, la cui precisione è vincolata dalla numerosità campionaria, che dipende, a sua volta, dalle risorse disponibili. Inevitabilmente la qualità e/o la quantità dell'informazione campionaria decresce con l'ambito territoriale di riferimento: non è infatti difficile dimostrare che, per l'operare di certe economie di scala consentite dalle tecniche di campionamento ([1]), al decrescere della dimensione territoriale si determina, ceteris paribus, un deficit di risorse rispetto ai costi di rilevazione come esemplificato nella Fig. 1. In tal modo, via via che si scende dal livello nazionale a quello

---

* Gli autori desiderano ringraziare Luigi Cannari, che ha coordinato un gruppo di studio sugli stimatori regionali, Maurizio Gresti e Francesco Trimarchi. Signorini ha scritto i paragrafi 1, 3 e 4.1; Cesari i paragrafi 2, 4.2, 5 e l'Appendice metodologica. Le simulazioni si basano su dati non pubblicati, elaborati presso la Banca d'Italia.

regionale, a quello comprensoriale o addirittura comunale, esigenze informative sempre più capillari sono destinate ad essere lasciate insoddisfatte da rilevazioni dirette dei dati.

In Italia, in particolare, le principali informazioni infracensuarie sull'economia reale disponibili a livello regionale soffrono di limitazioni di varia natura che ne rendono problematico l'uso. Ad esempio le indagini sulle forze di lavoro sono caratterizzate da un'alta variabilità campionaria; le indagini sul clima congiunturale (come quelle dell'Isco) sono quasi sempre ([2]) di tipo solo qualitativo; i conti economici regionali vengono pubblicati dall'Istat con forti ritardi; le statistiche sul commercio estero soffrono di problemi circa la corretta imputazione geografica dei movimenti di merci; e così via. Ci sono ovviamente differenze tra regione e regione, ma le informazioni disponibili sono considerate largamente insufficienti per quantità e qualità.

Vale la pena di sottolineare, inoltre, che uno dei problemi creati dallo squibibrio tra domanda e offerta di dati regionali consiste nel fatto che gli utilizzatori tendono a trascurarne la natura stocastica. Ad esempio, per l'indagine sulle forze di lavoro l'Istat pubblica informazioni che permettono di ricostruire approssimativamente gli intervalli di confidenza delle stime (e questi sono spesso tali da rendere non significative le variazioni trimestrali o annuali rilevate nelle varie regioni); ma questa informazione spesso non viene considerata dagli utilizzatori dei dati.

Vi sono vari modi ragionevoli per cercare di ovviare in parte a questo tipo di lacune informative.
Una strategia, di tipo 'estensivo', è quella di ricorrere a informazioni indirette, non tradizionali, disponibili a livello regionale: ad esempio il consumo di elettricità come indicatore del livello di attività produttiva; o la collezione di bilanci aziendali come indicatore di macro-performance; o il movimento anagrafico delle imprese come indicatore di sviluppo. Il ricorso a questo tipo di fonti indirette infatti sembra essere molto cresciuto negli ultimi tempi. Per ciascuna di esse, il passaggio dalle quantità effettivamente misurate alle quantità che si vorrebbero conoscere implica problemi logici ed empirici, ma purché questo fatto sia tenuto presente e adeguatamente discusso (con l'aiuto di modelli quantitativi se possibile), in molti casi tali informazioni indirette possono risultare preziose, quanto meno in mancanza di alternative.

Un'altra strategia - di tipo 'intensivo' - consiste nello sviluppare e applicare metodologie statistiche che consentano di estrarre quanta più informazione territoriale sia possibile dalle indagini campionarie (dirette) esistenti. Queste, infatti, come noto, soffrono di regola non solo del problema generale delle 'osservazioni incomplete' (nonresponse e missing data) ma anche, con effetti congiunti negativi, della disponibilità di 'poche osservazioni' a livello regionale con conseguente distorsione e grande variabilità degli usuali stimatori regionali (stimatori "classici"). La strategia di estrarre "quanta più informazione possibile" consente di ovviare in parte a questi problemi anche se, come spesso accade, la metodologia appropriata non è unica ma dipende da quale proprietà degli stimatori si vuole ottimizzare. Poiché in molti casi il problema più serio delle stime regionali consiste nella loro grande variabilità, può essere desiderabile adoperare stimatori che riducano questa variabilità rispetto agli stimatori "classici", anche al costo di un peggioramento di altre caratteristiche (per esempio, un aumento della distorsione).

Il presente lavoro è dedicato ad esplorare questa strategia 'intensiva' in riferimento a stime degli investimenti fissi lordi, dell'occupazione e del fatturato basate sull'indagine campionaria annuale della Banca d'Italia sulle imprese manifatturiere. (Questa indagine è descritta brevemente nel paragrafo 2.1). L'esperimento, che ha natura essenzialmente metodologica, prende come universo di riferimento tale campione. La metodologia impiegata è quella degli stimatori "sintetici" per piccole aree descritta ad esempio da Steinberg (1979) e da Gori e

Marchetti (1988). In estrema sintesi, la metodologia (paragrafo 2.2 e Appendice metodologica) consiste nel supplire all'insufficiente numerosità di un campione regionale utilizzando per la stima della variabile X nella regione R non solo i dati delle imprese di quella regione ma anche quelli delle regioni contermini e/o assimilabili per certe opportune caratteristiche, riponderati (ossia post-stratificati) sulla base della struttura dell'universo nella regione R.

Gli stimatori sintetici così ottenuti sono soggetti a minore variabilità di quelli classici ma al contrario di questi ultimi sono in generale distorti. La scelta fra stimatori classici e sintetici implica pertanto un trade-off tra distorsione e varianza, i cui parametri non sono però determinabili a priori. Nella prospettiva di utilizzare questo tipo di stimatori è necessario disporre di qualche elemento concreto di valutazione dei guadagni che si possono conseguire in termini di precisione delle stime e dei "prezzi" da pagare in termini di bias. Nel presente lavoro si è pertanto adottata una procedura di tipo Montecarlo per (i) verificare l'esistenza del trade-off, (ii) determinarne a posteriori i parametri, (iii) valutare la performance "in laboratorio" di diverse versioni alternative di stimatori sintetici.

I risultati delle simulazioni sono descritti nel paragrafo 3. Essi, tuttavia, costituiscono il punto di partenza (non di arrivo) di una procedura decisionale per la scelta di uno stimatore regionale. Il paragrafo 4 è infatti dedicato allo sviluppo teorico di un approccio generalizzato alla scelta dello stimatore ottimale, di cui si individuano e analizzano i due aspetti fondamentali della funzione di perdita (paragrafo 4.1) e della frontiera degli stimatori efficienti (paragrafo 4.2). Le principali conclusioni, riassunte nel paragrafo 5, indicano che gli stimatori sintetici sperimentati risultano in molti casi (specialmente per quanto riguarda le regioni più grandi) preferibili a quelli classici sulla base di funzioni di perdita ragionevoli.

2. Metodologia

2.1 Il campione della Banca d'Italia

La Banca d'Italia conduce annualmente un'indagine rivolta alle imprese manifatturiere con oltre 50 addetti, che è nota come "Indagine sugli investimenti" ma rileva in realtà un più ampio insieme di dati: gli investimenti fissi lordi, l'occupazione (media e a fine anno), il fatturato, l'utilizzazione desiderata ed effettiva della capacità produttiva, le variazioni della stessa capacità, ed altre informazioni di natura qualitativa e quantitativa. Oltre ai valori dell'anno di riferimento vengono generalmente richiesti i corrispondenti valori di uno o più anni passati e, per alcune variabili, i programmi futuri su un orizzonte temporale di uno o due anni.

Il campione comprende circa 1200 aziende, ottenute da un campionamento stratificato bivariato (dimensione/settore di attività) con distribuzione della numerosità campionaria di cella ottenuta col metodo dell'assegnazione ottima (Hansen, Hurwitz e Madow (1953), vol.II, ch.5 sec.9). Risponde abitualmente circa il 90% delle imprese interpellate; la frazione di campionamento (ex post) varia dal 4% per le imprese con meno di 100 addetti alla quasi totalità per le grandi imprese. Le stime nazionali fornite dal campione per le variabili più importanti sono generalmente ritenute soddisfacenti per correttezza e accuratezza ([3]).

Ovviamente l'indagine della Banca d'Italia non si pone in alternativa rispetto alle stime di contabilità nazionale, ma ne costituisce un complemento per le sue caratteristiche di tempestività e di dettaglio dell'informazione (disaggregazioni, dati qualitiativi, possibilità di seguire unità elementari nel tempo).

In combinazione con le nuove serie di contabilità regionale dell'Istat, la disponibilità di informazioni regionali sul campione della Banca d'Italia svolgerebbe una funzione simile a quella svolta a livello nazionale.

Per la verità alcuni dati regionali dell'indagine della Banca d'Italia vengono già resi noti annualmente nelle "Note sull'andamento economico regionale" che la Banca pubblica per ciascuna regione. Nella maggior parte dei casi tuttavia non si tratta di stime vere e proprie ma semplicemente di dati aggregati riferiti a sub-campioni regionali che non hanno caratteristiche di rappresentatività, e per le regioni più piccole sono numericamente molto limitati. Per la costruzione di stime propriamente dette si potrebbe incrementare la numerosità campionaria e ridefinirne opportunamente la stratificazione, ma recenti calcoli sperimentali hanno provato che la numerosità necessaria perché i sub-campioni regionali forniscano stime ragionevolmente accurate è assai superiore all'attuale e comunque troppo elevata rispetto alle risorse a disposizione per l'indagine. Di qui l'interesse per l'approccio seguito in questo lavoro.

2.2 Caratteristiche dell'esperimento di simulazione

L'alternativa a una costosa estensione del campione è rappresentata dall'adozione di un modello che consenta di sopperire alla mancanza di informazione a livello regionale mediante opportune ipotesi sul legame tra la variabile oggetto di interesse ed altre sulle quali le informazioni a livello regionale sono disponibili (Gori e Marchetti, 1988). Gli stimatori sintetici (Steinberg, 1979), si basano implicitamente o esplicitamente sull'ipotesi che i parametri della distribuzione di probabilità della variabile X in un certo individuo siano determinati dall'appartenenza dell'individuo a un particolare sottoinsieme della popolazione complessiva ma indipendenti dalla localizzazione.

Nel caso concreto, il modello che presiede ad esempio alle decisioni di investimento di un'impresa dipenderebbe dal settore e/o dalla classe dimensionale di appartenenza e sarebbe lo stesso in qualsiasi regione. Nota (per esempio da dati censuari) la struttura per settore e dimensione dell'universo della regione R, diviene possibile stimare l'aggregato regionale degli investimenti partendo dalla stima campionaria del valor medio (nazionale) degli investimenti delle imprese appartenenti a ciascuna cella dimensione/settore, ed aggregando tali valori medi stimati con pesi basati sulla struttura di R, ossia con la numerosità di ciascuna cella nell'universo regionale.

Poiché il campione nazionale è più grande di quello regionale, lo stimatore sintetico ha varianza minore di quello basato sul solo sottocampione regionale. Se l'ipotesi sottostante è esattamente vera, lo stimatore è anche corretto. Se non è vera (il che è più plausibile), lo stimatore è invece distorto, e il valore della distorsione dipende in segno e modulo dalle specificità regionali di comportamento cella per cella. Questa distorsione, per ipotesi, non è nota a priori e non si può stimare nel campione se non con alta varianza, più esattamente con una varianza tale da annullare il vantaggio conseguito con l'impiego dello stimatore sintetico.

Una tale circolarità logica impedisce di valutare a priori i termini del trade-off tra media e varianza che la scelta tra stimatori classici e sintetici implica. L'unico modo di avere informazioni su di esso è quello di sperimentare "in laboratorio" differenti stimatori con campionamenti ripetuti su un piccolo universo noto. I risultati così ottenuti "in vitro" per un determinato contesto non possono essere automaticamente estesi, dal punto di vista statistico, a contesti differenti "in vivo" e valgono solo come esempio; ma se ci sono ragionevoli motivi per ritenere che il collegamento logico tra il caso di laboratorio e il caso reale sia stretto, allora l'esempio è comunque rilevante.

L'esperimento compiuto ai fini del presente lavoro è consistito nel prendere come universo di riferimento il campione B.I. rilevato nelle indagini sul 1987, 1988 e 1989, escludendo tuttavia (i) le imprese con dati mancanti per una o più delle variabili considerate, (ii) le imprese con 500 addetti o più, dato che il loro peso elevato avrebbe falsato i risultati mentre il loro numero esiguo non avrebbe consentito un'analisi statistica disaggregata.

L'universo 'nazionale' è quindi risultato composto, in ciascun anno, da 600 imprese circa e da esso sono stati estratti, per ciascun anno, 50 campioni casuali senza ripetizione ciascuno di dimensione pari a 250 unità.

Le variabili considerate per la stima sono otto, e precisamente:

- investimenti nell'anno precedente;
- investimenti nell'anno corrente previsti l'anno precedente;
- investimenti nell'anno corrente;
- investimenti previsti per l'anno successivo;
- occupazione alla fine dell'anno corrente;
- occupazione prevista per l'anno successivo;
- fatturato nell'anno precedente;
- fatturato nell'anno corrente.

Per ciascuna variabile e per ciascuna regione ([4]) si è valutata la performance ex post (su 50 iterazioni) di undici stimatori, di cui tre classici e sei sintetici, descritti brevemente più sotto e definiti in dettaglio nell'Appendice metodologica.

Sebbene la procedura di estrazione del campione non preveda una stratificazione, due dei tre stimatori classici e (necessariamente) tutti gli stimatori sintetici sono ottenuti previa post-stratificazione ([5]); le stime sono cioè medie ponderate in cui alle imprese appartenenti a ciascuna cella viene attribuito un peso pari al rapporto tra la quota della cella nell'universo (in termini di numero di imprese o di occupati) e la quota della stessa cella nel campione (o sotto-campione territoriale). Le celle sono definite in base alla dimensione e al settore e sono in tutto otto ([6]).

2.3 Stimatori classici

Gli stimatori regionali "classici" (cioè basati solo sull'informazione contenuta in ciascun sottocampione regionale) utilizzati nella simulazione sono i seguenti:

1. Stimatore semplice, non ponderato (media campionaria semplice)
2. Stimatore post-stratificato, ponderato con la numerosità delle imprese
3. Stimatore post-stratificato, ponderato con il totale degli addetti.

Lo stimatore 3, che assume come perfettamente nota la variabile di ponderazione ([7]), si fonda sull'ipotesi che la suddivisione delle imprese in un numero molto limitato di classi dimensionali non sia sufficiente per tener conto adeguatamente dell'effetto di scala. Gli stimatori 2 e 3 presuppongono che il sotto-campione regionale non abbia celle vuote (eccetto quelle che sono vuote anche nell'universo). Questa condizione non è quasi mai verificata per le regioni più piccole a causa dell'insufficiente numerosità delle imprese. Perciò le stime 2 e 3 sono state computate soltanto per un gruppo di 5 "grandi" regioni (Piemonte-Val d'Aosta, Lombardia, Veneto, Emilia-Romagna, Toscana). Tuttavia anche per le "grandi" regioni alcune iterazioni hanno dovuto essere scartate per la presenza di celle campionarie vuote; il numero di iterazioni per gli stimatori 2 e 3 è quindi generalmente inferiore a 50.

2.4 Stimatori sintetici

Gli stimatori sintetici fanno uso di informazione esterna al sottocampione regionale (qualche volta chiamata "informazione super-regionale"). Poiché l'ipotesi sottostante, cioè che il comportamento delle imprese appartenenti a una determinata cella non vari tra una regione e l'altra, è presumibilmente tanto meno vera quanto più ampia è l'area super-regionale di riferimento, è ragionevole attendersi che esista un trade-off fra distorsione e varianza anche all'interno della classe degli stimatori sintetici, e precisamente tra quegli stimatori che fanno uso dell'intero campione nazionale e quelli che fanno uso di un sottoinsieme di esso, composto di un certo numero di regioni "simili". Per questo motivo si sono sperimentati stimatori di entrambi i tipi. Inoltre le aree intermedie, oltre quella nazionale, sono state definite in due modi alternativi: con riferimento alle tradizionali aree Istat ([8]) (quindi con una definizione a priori delle aree: Nord-Ovest, Nord-Est, Centro, Sud e Isole) e con una procedura di clustering automatico ([9]). Infine, come per gli stimatori classici post-stratificati, si sono usati due diversi criteri di ponderazione. Quindi gli stimatori sintetici, post-stratificati, sono in tutto 6, e precisamente:

4. campione = area Istat, ponderazione = numerosità
5. campione = area Istat, ponderazione = addetti
6. campione = cluster, ponderazione = numerosità

7. campione = cluster, ponderazione = addetti
8. campione = Italia, ponderazione = numerosità
9. campione = Italia, ponderazione = addetti

Come gli stimatori 2 e 3 tutti gli stimatori sintetici presuppongono l'assenza di celle vuote, ma a livello di campione nazionale o di area intermedia. I casi di scarto sono quindi risultati pochi e le stime da 4 a 9 sono state calcolate per tutte le regioni. Le proprietà degli stimatori sono discusse nell'Appendice metodologica.

In sintesi, ci si attende che gli stimatori classici siano corretti e che la distorsione cresca passando dagli stimatori basati su aree intermedie a quelli basati sul campione nazionale; ci si attende d'altra parte che la variabilità delle stime diminuisca passando dallo stimatore semplice a quelli classici post-stratificati, da questi agli stimatori sintetici intermedi, e infine da questi a quelli basati sul campione nazionale.

3. Risultati

I risultati delle simulazioni sono riassunti nelle tavole 1 e 2, separatamente per il gruppo delle "grandi" e quello delle "piccole" regioni (medie dei risultati regionali).

Per tutti gli stimatori calcolati è riportato il valore, in percentuale del valore effettivo da stimare, del bias assoluto, della deviazione standard, del root mean square error e della t di Student per l'ipotesi di non distorsione dello stimatore.

Per semplicità si indicherà d'ora in avanti lo stimatore n con Xn. Le Figg.2a-2c e 3a-3c mostrano graficamente, con riferimento a investimenti, occupazione e fatturato nel 1988, i termini del trade-off tra distorsione e deviazione standard nei 9 stimatori considerati. I risultati sono per certi aspetti diversi per le "grandi" regioni (con un numero medio di 37 imprese nei 50 campioni, v. Tav. 3) rispetto alle "piccole" regioni (6 imprese campionate in media). Ciò ha suggerito di calcolare le statistiche riassuntive separatamente per i due gruppi e la trattazione dei risultati sarà quindi separata.

3.1 "Grandi" regioni

Si considerino i dati relativi alle cinque "grandi" regioni (Tav.1). Gli stimatori classici presentano, nelle tre rilevazioni, scostamenti dalla media relativamente contenuti (ad esempio nel caso di X1 e in riferimento al 1988 il bias va dall' 1,2% dell'occupazione al 2,3% degli investimenti, per X2 dal 3,7% del fatturato al 7,2% degli investimenti, valori entro gli usuali livelli di significatività) ma hanno una variabilità elevata.

In particolare X1 (lo stimatore non ponderato) ha una deviazione standard dell'ordine del 15-20% del valore vero (inferiore per l'occupazione), e X2 (lo stimatore ponderato con la numerosità delle imprese) di quasi il 7,5% per l'occupazione e del 20% circa per gli investimenti (livelli intermedi per il fatturato). Lo stimatore X3 (ponderato con l'occupazione totale) dà risultati lievemente migliori rispetto ai primi due.

Gli stimatori sintetici, sempre nelle "grandi" regioni, hanno, come atteso, una distorsione più elevata rispetto a quelli classici: nettamente più elevata per quanto riguarda gli investimenti e il fatturato (fino al 19% per gli investimenti), non altrettanto per quanto riguarda l'occupazione (in alcuni casi la distorsione media è persino inferiore a quella degli stimatori classici, sebbene sempre significativa). La variabilità invece è alquanto inferiore rispetto agli stimatori classici (sotto il 5% per l'occupazione, 7-10% per il fatturato e gli investimenti correnti e previsti). In maggior misura che nel caso degli stimatori classici, la ponderazione con gli addetti migliora la distorsione e la deviazione standard.

Per tutte le tre rilevazioni, le macro aree 'istituzionali' (aree Istat) utilizzate dagli stimatori 4 e 5 risultano la migliore scelta in termini di distorsione mentre in termini di deviazione standard risulta preferibile l'unica macro-area 'Italia'. I cluster automatici risultano dominati sotto entrambi gli aspetti, fatta eccezione per le variabili relative all'anno (1987) rispetto a cui sono stati formati i clusters.

E' evidente che nel nostro esperimento (i) le stime dell'occupazione sono complessivamente migliori delle altre, (ii) anche i termini del trade-off tra distorsione e variabilità sono migliori per l'occupazione che per il fatturato e, soprattutto, per gli investimenti, al punto che alcuni stimatori sintetici dominano (minor valore del bias e minor varianza) gli stimatori classici ([10]).

Questi risultati, tuttavia, sono in parte spurii; nella definizione delle "celle" per la post-stratificazione entra (anche se in termini molto approssimativi e con riferimento all'anno precedente) la classe dimensionale, definita appunto in termini di occupati; è ovvio che le stime dell'occupazione in queste condizioni sono più accurate. Quindi, in questo senso, il risultato dell'esperimento non può automaticamente essere esteso a una realtà in cui la post-stratificazione debba essere fatta ricorrendo a dati molto arretrati.

Naturalmente, la scelta dello stimatore "migliore" dipende dalla funzione di preferenza dell'analista. Il più comune criterio di decisione è l'errore quadratico medio (MSE o la sua radice quadrata RMSE), che naturalmente non è altro che una funzione di perdita quadratica dove distorsione e deviazione standard entrano con ugual peso.

La Tav. 1 mostra una situazione variegata, su cui è interessante soffermarsi brevemente.

1. Per l'occupazione tutti gli stimatori sintetici

sono preferibili a quelli classici in termini di RMSE; in particolare gli stimatori ponderati con gli addetti (quelli dispari, a parte X1) sono preferibili a quelli ponderati con le numerosità. Tra gli stimatori sintetici nessuno emerge nettamente, ma l'RMSE minimo (1% circa) si raggiunge con lo stimatore che fa uso dell'intero campione nazionale.

2. Anche per il fatturato si ricava una qualche preminenza di quest'ultimo stimatore (e di X5 su posizioni molto vicine), benchè in riferimento alla rilevazione del 1989 la forte distorsione degli stimatori sintetici renda preferibili le stime classiche.

3. Gli investimenti sono caratterizzati da un RMSE superiore rispetto al fatturato; non è facile generalizzare sulla performance delle varie classi di stimatori. Anche qui, tuttavia, la ponderazione con gli addetti e le aree Istat sembrano complessivamente le scelte migliori.

3.2 "Piccole" regioni

Per le regioni rimanenti (Tav. 2) distorsione e variabilità degli stimatori si collocano su livelli nettamente peggiori. Da un punto di vista qualitativo, invece, la situazione è in un certo senso simile, perché distorsione e variabilità cambiano da uno stimatore all'altro nel modo previsto: più bassa (non significativa) la distorsione dello stimatore classico X1 (gli stimatori post-stratificati non si sono potuti calcolare per l'eccessiva probabilità di celle vuote) rispetto a quella degli stimatori sintetici e assai più elevata la variabilità delle sue stime. In particolare, gli stimatori sintetici almeno per quanto riguarda investimenti e fatturato presentano distorsioni tra il 20 e il 40% circa mentre mostrano deviazioni standard del 5-25% (oltre 30% lo stimatore classico). Tra gli stimatori sintetici si segnalano decisamente quelli (X8, X9) che fanno riferimento alla macro-area 'Italia' i quali hanno una distorsione simile agli altri ma una variabilità anche 3 volte inferiore. Come già nel caso delle "grandi" regioni, la ponderazione con gli addetti prevale, in generale, su quella con il numero di imprese. Anche in questo caso le statistiche relative alla variabile occupazione sono migliori, specialmente per quanto riguarda gli stimatori sintetici; ma come si è già avvertito in questo risultato c'è un elemento spurio dovuto al meccanismo di formazione delle celle.

In termini di RMSE il trade-off distorsione/varianza viene risolto a favore degli stimatori sintetici (ad esempio, per gli investimenti realizzati nel 1988 si ha un RMSE minimo del 38% (X5 e X9) contro il 50% dello stimatore classico; per gli investimenti previsti 22% di X9 contro 47% circa di X1; per l'occupazione 3% contro 25% circa).

4. Un approccio generale alla scelta dello stimatore ottimale

4.1 Funzioni generalizzate di perdita

Il RMSE è solo uno degli infiniti possibili criteri di decisione. Per generalizzare e disporre di una valutazione alternativa, e in un certo senso più accurata, dei termini del trade-off, si consideri la funzione di perdita

[1] $$L(|\beta_i|,\sigma_i,\alpha,h) = \alpha\sigma_i^h + (1-\alpha)|\beta_i|^h$$

dove $0 \leq \alpha \leq 1$, e l'indice i identifica lo stimatore X_i caratterizzato dalla distorsione $\beta_i=E(X_i)-\mu$ e dalla deviazione standard σ_i. Se X_0 è uno stimatore assunto come benchmark e X_i è uno stimatore alternativo tale che $\sigma_i < \sigma_0$ e $|\beta_i|>|\beta_0|$, X_i sarà preferibile a X_0 per ogni funzione del tipo [1] (dato h), in cui $\alpha > \alpha_i^*(h)$, dove $\alpha_i^*(h)$ è dato da: $\alpha_i^*(h)/(1-\alpha_i^*(h)) \equiv \tau_i(h) = [|\beta_i|^h - |\beta_0|^h]/[\sigma_0^h - \sigma_i^h]$ cioè $\alpha_i^*(h) = \tau_i(h)/(1+\tau_i(h))$

La statistica $\tau_i(h)$ rappresenta in modo sintetico il cutoff-point della funzione di perdita [1] dato h e può essere usata come misura dei termini del trade-off tra correttezza ed efficienza implicito nella scelta di uno stimatore alternativo rispetto a X_0. In particolare, $\tau_i(1)$ è semplicemente il rapporto tra incremento del bias e diminuzione della deviazione standard conseguenti al passaggio dallo stimatore di riferimento allo stimatore alternativo X_i. Quanto più basso è τ_i, tanto migliori sono i termini del trade-off.

Per le "grandi" regioni, scartato X1, che è (quasi) sistematicamente dominato, si è usato come stimatore classico di riferimento X2, e si è calcolato $r_i(1)$ e $\alpha_i^*(1)$ per ciascuno stimatore sintetico. Per le "piccole" regioni, necessariamente, si è assunto X1 come stimatore di riferimento.

I risultati, esposti nelle Tavv. 4 e 5, confermano che gli stimatori sintetici possono essere preferiti a quelli classici sulla base di funzioni di perdita ragionevoli.

In particolare, tra gli stimatori sintetici, X8 e X9 (area totale Italia) risultano in generale quelli preferibili per il più ampio sottoinsieme di funzioni di perdita lineari.

Infatti, escludendo l'occupazione per la quale, come già notato, non c'è trade-off e la preferenza per gli stimatori sintetici è pressoché assoluta, gli stimatori X8 e/o X9 sono preferibili a X2 per funzioni di perdita che attribuiscano alla deviazione standard, (i) nel caso del fatturato un peso pari o superiore a circa 2/3 di quello del bias (rilevazioni 1987 e 1988; circa 2 volte per il 1989), (ii) nel caso degli investimenti un peso circa pari o superiore a quello del bias (rilevazioni 1987 e 1989; circa 1/4 per il 1988).

Per le "piccole" regioni il cutoff point della funzione di preferenza di primo grado (h=1) si trova su valori analoghi per il fatturato (tra 1/2 e 2/3) e inferiori per gli investimenti (tra 1/2 e 1).

4.2 La frontiera degli stimatori efficienti

In forma sintetica, l'approccio usualmente seguito in un esercizio di simulazione di stimatori alternativi come quello descritto nel paragrafo 2, è quello di generare un meccanismo casuale di campionamento e di calcolare "in vitro" distorsione β (in valore assoluto) e deviazione standard σ di ciascun stimatore considerato. Una funzione di perdita, diciamo $L(|\beta|,\sigma)$, consente di attribuire un valore alla performance degli stimatori, fornendone un ordinamento dal 'migliore' al 'peggiore'.

In realtà, almeno in linea di principio, un approccio più generale può essere delineato.

Infatti, sebbene, a quanto ci consta, non sia stato mai osservato neppure a livello teorico, il problema della scelta dello stimatore ottimale si può porre in stretta analogia con il problema della scelta del miglior portafoglio di investimenti, argomento centrale della teoria finanziaria ([11]).

In ambito statistico, il problema può essere formulato come segue: si vuole individuare lo stimatore ottimale, $\hat{X}_*$, del parametro μ caratteristico della distribuzione della variabile X nell'universo da cui si campiona, dato un vettore di n stimatori di base $\hat{X}=(\hat{X}_1,....,\hat{X}_n)'$ e data una funzione di perdita generalizzata $L(|\beta|,\sigma)$, ove $|\beta|$ è la distorsione dello stimatore in valore assoluto e σ è la sua deviazione standard:

$\beta_i \equiv E(\hat{X}_i)-\mu, \quad \sigma^2_i \equiv E[(\hat{X}_i-E(\hat{X}_i))^2].$

Per definizione, $L_1>0$, $L_2>0$, L_j essendo la derivata parziale di L rispetto al j-esimo argomento.

Si assuma che l'insieme degli stimatori ammissibili sia dato dalle combinazioni lineari degli stimatori di base $\{\hat{Z}=w'\hat{X},\ w\in R^n,\ w'1=1\}$ ([12]).

Si avrà:

$E(\hat{Z})=w'E(\hat{X}), \quad \beta=E(\hat{Z})-\mu = w'E(\hat{X})-\mu=w'(E(\hat{X})-\mu 1)+\mu(w'1-1)=w'b$

$\sigma^2=w'\Sigma w$

ove Σ è la matrice delle varianze-covarianze di $\hat{X}$ e b il vettore delle distorsioni.

Il principio fondamentale della 'diversificazione' finanziaria si applica anche nella scelta dello stimatore, risultando efficace ogni qualvolta le correlazioni tra gli stimatori-base, presumibilmente positive, non siano perfette.

In analogia con la teoria delle scelte di portafoglio (cfr.Ingersoll (1987) ch.4) si possono definire di 'minima varianza' gli stimatori (combinazioni di stimatori-base) con la minor varianza per ogni livello di distorsione β.

Essi saranno identificati dalla soluzione del problema di minimo vincolato:

$\text{Min}_w \ \frac{1}{2}w'\Sigma w$ dati $w'b = \beta$ e $w'1=1$ ([13])

ove per comodità si è moltiplicata per ½ la varianza da minimizzare. Formando il lagrangiano

$\pounds \equiv \frac{1}{2}w'\Sigma w - \delta(w'b-\beta)-\phi(w'1-1)$ e derivando si ha

$\Sigma w-\delta b-\phi 1=0$ da cui

[2] $w=\delta\Sigma^{-1}b+\phi\Sigma^{-1}1$

e quindi usando i vincoli si ottiene

[3] $\delta=(\beta A-B)/D, \quad \phi=(C-\beta B)/D$

ove $D\equiv AC-B^2$ $A\equiv 1'\Sigma^{-1}1$ $B\equiv 1'\Sigma^{-1}b$ $C\equiv b'\Sigma^{-1}b$

$A,C,D>0$.

Sostituendo il vettore w così ottenuto nella forma quadratica della varianza si ottiene la frontiera degli stimatori di minima varianza, data da una parabola nello spazio σ^2,β, (iperbole in σ,β)

[4] $\sigma^2= w'\Sigma w = (A\beta^2-2B\beta+C)/D$

Si noti che lo stimatore di minima varianza globale si può ricavare ponendo a zero il moltiplicatore di Lagrange sul vincolo di distorsione, $\delta=0$, da cui $\beta=B/A$ (distorsione positiva o negativa) e $\sigma^2=1/A$ $(=\phi)$ (minima varianza globale), che identificano il vettore di pesi $w_v\equiv\Sigma^{-1}1/A$. All'opposto, lo stimatore con distorsione nulla ha varianza C/D $(=\phi)$ ed è identificato dal vettore $w_B=-B\Sigma^{-1}b/D+C\Sigma^{-1}1/D$.

La frontiera degli stimatori efficienti è l'arco tra questi due estremi (v. Fig.4), le altre combinazioni essendo tutte dominate per qualunque forma funzionale di L.

Ad esempio, il criterio RMSE equivale alla funzione di perdita

$L(|\beta|,\sigma)= \beta^2+\sigma^2$ che determina curve di indifferenza nello spazio (β,σ) costituite dalle semicirconferenze con centro in (0,0) e raggio crescente col livello di perdita. Posto $\beta^2+\sigma^2=L°$, usando [4] e imponendo un'unica soluzione per la distorsione β, si ottiene $L°=(C+1)/(A+D)$ e la distorsione $\beta=B/(A+D)$ che sostituita nei moltiplicatori [3] dà il vettore ottimale dei pesi $w=-B\Sigma^{-1}b/(A+D)+(C+1)\Sigma^{-1}1/(A+D)$.

Analogamente si procede considerando la funzione di perdita [1] per h=1:

$L(|\beta|,\sigma)= \alpha\sigma+(1-\alpha)|\beta|$

che dà luogo a curve di indifferenza del tipo di quelle rappresentate nella Fig.4 con coefficiente angolare (in valore assoluto) pari a $1/\tau$.

Da un punto di vista operativo, si noti che se $E(\hat{Z})$ ha lo stesso ordine di grandezza di μ (ad esempio nella stima di un valor medio), allora tali saranno anche β e σ mentre C e il vettore dei pesi w sono di ordine 0 (numeri puri) rispetto a μ, cioè indipendenti dall'unità di misura.

Chiaramente, se nell'approccio tradizionale la validità dei risultati dell'esperimento dipende dalla corrispondenza tra relazioni in simulazione e relazioni nel caso concreto: ad esempio

$$\beta_0{}^2+\sigma_0{}^2 < \beta_j{}^2+\sigma_j{}^2 \quad \Longleftrightarrow \quad \beta_0{}^{2*}+\sigma_0{}^{2*} < \beta_j{}^{2*}+\sigma_j{}^{2*}$$

(ove X_0 è lo stimatore 'migliore' nell'esperimento e i parametri asteriscati si riferiscono agli stimatori nella situazione concreta), nell'approccio appena delineato non si può richiedere che si verifichi la corrispondenza:

$$(w_0'b)^2+w_0'\Sigma w_0 < (w_s'b)^2+w_s'\Sigma w_s \iff (w_0'b^*)^2+w_0'\Sigma^* w_0 < (w_s'b^*)^2+w_s'\Sigma^* w_s$$

ove w_s è un qualunque altro vettore di pesi, dato che si avrà quasi sicuramente $w_0 \neq w_0^*$ (vettore ottimale nel caso concreto).

Tuttavia, per il successo dell'approccio proposto è sufficiente che si abbia:

$$(w_0'b^*)^2+w_0'\Sigma^* w_0 < \beta_0^{2*}+\sigma_0^{2*}$$

una condizione che può realizzarsi in numerose circostanze ([14]) in funzione degli effettivi parametri degli stimatori nella realtà indagata.

Per la variabile investimenti (a titolo esemplificativo per il 1989 e per la regione Emilia Romagna) la simulazione ha consentito di calcolare la frontiera disegnata nella Fig.5. Come atteso, gli stimatori 8 e 9 sono quasi sulla frontiera di minima varianza (dove è rappresentato anche lo stimatore ottimale in termini di RMSE w_0), ma presentano una maggiore distorsione rispetto agli stimatori 4 e 5 e agli stimatori classici.

5. Conclusioni

Il campione Banca d'Italia delle imprese manifatturiere non è sufficientemente numeroso per fornire stime regionali sufficientemente attendibili degli investimenti e di altre variabili, neppure nel caso delle regioni più grandi, a causa della notevole variabilità delle stime campionarie "classiche". L'uso di stimatori "sintetici", che sfruttano l'informazione relativa ad aree super-regionali, può costituire una valida alternativa a una costosa estensione del campione, purché per ottenere una riduzione della varianza delle stime si sia disposti a pagare un prezzo in termini di distorsione.

Nella prospettiva di costruire ed utilizzare simili stimatori, si è compiuta una verifica preliminare, sia analitica sia numerica, delle caratteristiche e delle proprietà degli stimatori stessi, che ha fornito elementi di valutazione dei termini del trade-off tra media e varianza nel confronto fra stimatori classici e sintetici. Un esperimento di tipo Montecarlo, condotto tutto all'interno dei campioni relativi alle rilevazioni '87,'88 e '89 ed esteso alle principali variabili quantitative dell'indagine (investimenti, occupazione e fatturato) e a numerose varianti di stimatori classici e sintetici, ha permesso di stabilire quanto segue.

a) Per le regioni più grandi i termini del trade-off tra distorsione e variabilità sono tali che gli stimatori sintetici, o alcune varianti di essi, sono preferibili a quelli classici sulla base di funzioni di perdita che attribuiscano alla variabilità (deviazione standard delle stime) un peso pari o superiore a quello della distorsione nel caso della variabile investimenti e a circa 2/3 o più nel caso della variabile fatturato. Per la variabile occupazione gli stimatori sintetici dominano quelli classici per qualunque funzione lineare o quadratica di preferenza distorsione/varianza, sebbene tale risultato sia in parte un effetto spurio della post-stratificazione settore/addetti.

b) Per le regioni di minori dimensioni gli stimatori sintetici risultano preferibili a quelli classici per pesi relativi sulla variabilità (rispetto alla distorsione) analoghi a quelli delle "grandi" regioni per il fatturato e inferiori per gli investimenti (tra 1/2 e 1). Rapporti significativamente più bassi si ottengono anche in questo caso per la variabile occupazione.

c) Il trade-off distorsione/varianza si manifesta su livelli della funzione di perdita, lineare o quadratica, molto diversi nei due casi delle "grandi" e delle "piccole" regioni, nel senso che anche usando stimatori sintetici di vario tipo queste ultime restano caratterizzate nell'esperimento da livelli di distorsione e/o varianza anche doppi o più rispetto alle "grandi" regioni e presumibilmente inaccettabili a fini analitici (as es. RMSE degli investimenti '88 intorno al 40% del valore vero contro il 15% circa per le "grandi" regioni). Va comunque tenuto presente che anche per le "piccole" regioni gli stimatori sintetici risultano migliori dello stimatore classico 'media campionaria semplice' per ampie classi di funzioni di perdita.

d) Tra i vari tipi di stimatori sintetici non emerge una gerarchia univoca (ossia rapporti di dominanza generalizzati) tra quelli basati sull'intero campione nazionale, da un lato, e quelli basati su aggregazioni intermedie dall'altro; tuttavia sia per le "grandi" sia per le "piccole" regioni e per ampie classi di funzioni di perdita (RMSE e lineari) i risultati dell'esperimento segnalano gli stimatori basati sulla macro-area Italia (X8 e X9) e, per le sole "grandi" regioni, gli stimatori basati sulle aree Istat (X4 e X5), su posizioni vicine ai precedenti.

e) Le procedure automatiche di clustering in base alla variabile investimenti e al rapporto investimenti/occupati non forniscono in generale (fatta eccezione per le variabili '87 usate nel clustering) risultati migliori dell'aggregazione "standard" per aree Istat; quest'ultima risulta quindi da preferirsi nelle applicazioni (a fortiori per le ragioni esposte nella nota 9).

I risultati conseguiti incoraggiano, a parere di chi scrive, a sviluppare la metodologia con riferimento, almeno in via preliminare, alle regioni maggiori. La pubblicazione di serie Istat relative agli universi regionali di riferimento e alle variabili esaminate dovrebbe consentire in tempi brevi la costruzione di stime regionali di investimenti e fatturato e la verifica, su dati effettivi, della bontà degli stimatori sintetici proposti. In prospettiva, l'applicazione congiunta degli stimatori sintetici (per il problema di undercoverage) e delle tecniche di imputazione (per il problema dei missing data: v. Trimarchi (1990)) dovrebbe consentire non solo miglioramenti delle stime sulle regioni maggiori ma anche stime sufficientemente affidabili per le regioni più piccole.

APPENDICE METODOLOGICA

MEDIA DISTORSIONE E VARIANZA DEGLI STIMATORI REGIONALI CLASSICI E SINTETICI POST-STRATIFICATI

Si consideri la variabile X presso una popolazione finita $P=\{X^{\sim}_p,\ p=1(1)N\}$, ove il simbolismo $p=1(1)N$ indica la successione $p=1,2,\ldots,N-1,N$, suddivisa in un certo numero di 'regioni' (sub-popolazioni o dominii: v. Cochran (1977) p.34) di numerosità ${}_RN$, ove il sub-prefisso R indica una generica regione. Si è interessati alla stima del valor medio ${}_RX^{\approx}\equiv{}_RX^{\sim}/{}_RN$ della variabile $X^{\sim}$ presso la sub-popolazione regionale, sulla base di un campionamento della popolazione P senza reimmissione, post-stratificato, di dimensione n.

Si noti che la dimensione del sub-campione regionale ${}_Rn$ e quella degli strati (celle) ${}_Rn_h$, per $h=1(1)H$, sono variabili casuali.

In particolare, dato un sub-campione regionale $\{{}_RX_j,\ j=1(1){}_Rn\}$, post-stratificato in $\{{}_RX_{hk},\ k=1(1){}_Rn_h\ h=1(1)H\}$, si è assunto per ogni strato ${}_Rn_h>0$ e ${}_RN_h$ noto senza errore e si sono considerati i seguenti 5 stimatori ([15]):

1. <u>Stimatore semplice</u>

$${}_R\hat{X}_1 = 1/{}_Rn\ \Sigma_j\ {}_RX_j$$

2. <u>Stimatore post-stratificato ponderato con la numerosità</u>

$${}_R\hat{X}_2 = 1/{}_RN\ \Sigma_h\ \Sigma_k\ {}_RX_{hk}\ {}_RN_h/{}_Rn_h$$

3. <u>Stimatore post-stratificato ponderato con la variab. Y</u>

${}_R\hat{X}_3 = 1/{}_RN\ \Sigma_h\ \Sigma_k\ {}_RX_{hk}\ {}_RY^{\sim}{}_h/{}_RY_h$ ove

${}_R Y_h = \Sigma_k \; {}_R Y_{hk}$ mentre ${}_R Y^{\sim}{}_h$ è l'effettivo valore della variabile regionale ${}_R Y^{\sim}$ nello strato h.

4. <u>Stimatore super-regionale post-stratificato ponderato con la numerosità</u>

${}_R\hat{X}_4 = 1/{}_R N \; \Sigma_h \; \Sigma_k \; {}_A X_{hk} \; {}_R N_h / {}_A n_h$ $\qquad k=1(1)_A n_h$ numerosità delle osservazioni nello strato h della macro area (circoscrizione) A.

5. <u>Stimatore super-regionale post-stratificato ponderato con la variabile Y</u>

${}_R\hat{X}_5 = 1/{}_R N \; \Sigma_h \; \Sigma_k \; {}_A X_{hk} \; {}_R Y^{\sim}{}_h / {}_A Y_h$

Utilizzando le proprietà del valor medio condizionato (v. Mood, Graybill e Boes(1974) p.158) si possono verificare per ciascuno stimatore i seguenti risultati:

1. $E({}_R\hat{X}_1) = E(E({}_R\hat{X}_1 | {}_R n)) = {}_R X^{\sim} / {}_R N \equiv {}_R X^{\approx}$ cioè lo stimatore semplice è corretto;

$$Var({}_R\hat{X}_1) = E(Var({}_R\hat{X}_1 | {}_R n)) + Var(E({}_R\hat{X}_1 | {}_R n)) =$$
$$= {}_R\underline{\sigma}^2 E(1/{}_R n) - {}_R\underline{\sigma}^2 / {}_R N$$

ove ${}_R\underline{\sigma}^2 \equiv 1/({}_R N-1) \Sigma_p \; ({}_R X^{\sim}{}_p - {}_R X^{\approx})^2$ è la varianza 'corretta' della sub-popolazione, ${}_R X^{\approx} \equiv 1/{}_R N \; \Sigma_p \; {}_R X^{\sim}{}_p$ è la media della sub-popolazione e ${}_R n / {}_R N$ è la frazione (aleatoria) di campionamento.

Una stima approssimativamente corretta della varianza dello stimatore 1 si può ottenere osservando che ([16]) ([17]) in approssimazione quadratica il valor medio di $1/{}_R n$ è, per ${}_R W \equiv {}_R N/N$,

$E(1/{}_R n) \approx 1/(n\ {}_R W) + (1-{}_R W)/(n^2\ {}_R W^2)(N-n)/(N-1)$

e una stima corretta di ${}_R\underline{\sigma}^2$ è data da

${}_R S^2 \equiv 1/({}_R n-1)\ \Sigma_j\ ({}_R Xj-{}_R X^{\char126})^2$, ove

${}_R X^{\char126} \equiv {}_R\hat{X}_1 = \Sigma_j\ {}_R X_j/{}_R n$ è lo stimatore semplice, media campionaria.

Si tenga presente che una stima corretta della varianza dello stimatore implica, per la diseguaglianza di Jensen, una distorsione per difetto dello scarto quadratico medio.

2. $E({}_R\hat{X}_2) = E(E({}_R\hat{X}_2 | {}_R n_h)) = E({}_R X^{\char126}/{}_R N) = {}_R X^{\approx}$

corretto (essendo ${}_R n_h > 0$)

$$\begin{aligned}Var({}_R\hat{X}_2) &= E(Var({}_R\hat{X}_2 | {}_R n_h)) + Var(E({}_R\hat{X}_2 | {}_R n_h)\\ &= E(\Sigma_h\ {}_R W_h^2/{}_R n_h\ {}_R\underline{\sigma}^2{}_h(1-{}_R n_h/{}_R N_h))\\ &= \Sigma_h\ {}_R W_h^2\ E(1/{}_R n_h)\ {}_R\underline{\sigma}^2{}_h - \Sigma_h\ {}_R W_h^2/{}_R N_h\ {}_R\underline{\sigma}^2{}_h =\end{aligned}$$

essendo ${}_R W_h \equiv {}_R N_h/{}_R N$ e ([18])

$$\begin{aligned}E(1/{}_R n_h) &= E(E(1/{}_R n_h | {}_R n)) \approx E(1/{}_R n)\ ({}_R N_h-1)/(({}_R N-1){}_R W_h^2) +\\ &+ E(1/{}_R n^2)\ ({}_R N-{}_R N_h)/(({}_R N-1){}_R W_h^2)\end{aligned}$$

ove $E(1/{}_R n)$ è approssimabile come sopra mentre in approssimazione di ordine n^{-2} si può porre $E(1/{}_R n^2) \approx 1/E^2({}_R n) = 1/(n^2\ {}_R W^2)$.

Si noti che la varianza di ${}_R\hat{X}_2$ è funzione lineare della varianza <u>entro</u> gli strati, ${}_R\underline{\sigma}^2{}_h$ ([19]) e si riduce al crescere della frazione (media) di campionamento.

3. Usando l'approssimazione in serie di Taylor al II ordine si ottiene ([20])

$E({}_R\hat{X}_3) \approx {}_R\tilde{X}/{}_RN + E({}_R\triangle_3)$ ove

$${}_R\triangle_3 \equiv \Sigma_h \; {}_RW_h \; (1/{}_Rn_h - 1/{}_RN_h) \; 1/{}_R\tilde{\tilde{Y}}_h ({}_R\tilde{\tilde{X}}_h/{}_R\tilde{\tilde{Y}}_h \; {}_R\underline{\sigma}_{hYY} - {}_R\underline{\sigma}_{hXY})$$

e il suo valor medio è rispetto alla distribuzione di ${}_Rn_h$. Si noti che il fattore $({}_R\tilde{\tilde{X}}_h/{}_R\tilde{\tilde{Y}}_h \; {}_R\underline{\sigma}_{hYY} - {}_R\underline{\sigma}_{hXY})$ ([21]) si può scrivere anche ${}_R\tilde{\tilde{X}}_h/{}_R\tilde{\tilde{Y}}_h \; {}_R\underline{\sigma}_{hYY}(1 - {}_R\beta_{hXY} \; {}_R\tilde{\tilde{Y}}_h/{}_R\tilde{\tilde{X}}_h)$ con ${}_R\beta_{hXY}$ coefficiente di regressione di ${}_RX_h$ su ${}_RY_h$. Pertanto la distorsione si annullerà quando ${}_R\beta_{hXY}$ eguaglia il reciproco del rapporto tra le medie (cioè quando la retta di regressione dei rapporti con le rispettive medie passa per l'origine con coefficiente unitario).

La varianza dello stimatore è

$$\begin{aligned} Var({}_R\hat{X}_3) &= E(Var({}_R\hat{X}_3 | {}_Rn_h)) + Var(E({}_R\hat{X}_3 | {}_Rn_h) \\ &\approx \Sigma_h \; {}_RW_h^2 \; {}_R\tilde{\tilde{X}}_h^2 \; E(1/{}_Rn_h - 1/{}_RN_h) \\ &\qquad [{}_R\underline{\sigma}_{hXX}/{}_R\tilde{\tilde{X}}_h^2 + {}_R\underline{\sigma}_{hYY}/{}_R\tilde{\tilde{Y}}_h^2 - 2{}_R\underline{\sigma}_{XY}/({}_R\tilde{\tilde{X}}_h \; {}_R\tilde{\tilde{Y}}_h)] \\ &\quad + Var({}_R\triangle_3) \end{aligned}$$

e risulta anch'essa in funzione della bontà della relazione lineare tra le variabili rapportate alla media.

4. $E({}_R\hat{X}_4) = {}_R\tilde{X}/{}_RN + {}_R\triangle_4$ ove

$${}_R\triangle_4 \equiv \Sigma_h \; {}_RW_h({}_A\tilde{\tilde{X}}_h - {}_R\tilde{\tilde{X}}_h)$$

$$Var({}_R\hat{X}_4) = \Sigma_h \; {}_RW^2_h \; {}_A\underline{\sigma}^2_h \; E(1/{}_An_h) - \Sigma_h \; {}_RW^2_h \; {}_A\underline{\sigma}^2_h/{}_AN_h$$

Si noti che la distorsione dello stimatore non dipende dalle numerosità campionarie di cella e si riduce al ridursi delle differenze tra le medie di cella della macro e della micro area.

Confrontando gli stimatori 2 e 4 si nota che lo stimatore super-regionale introduce un trade-off tra distorsione e varianza: se la macro area ha varianza di cella non superiore a quella regionale (condizione di ottimalità nella scelta dell'area) lo stimatore super-regionale ${}_R\hat{X}_4$ è distorto ma ha varianza minore di ${}_R\hat{X}_2$.

5. $E({}_R\hat{X}_5) \approx {}_RX^{\sim}/{}_RN + {}_R\triangle_{51} + E({}_R\triangle_{52})$ ove

$${}_R\triangle_{51} = \Sigma_h \; ({}_A\overset{\approx}{X}_h/{}_A\overset{\approx}{Y}_h - {}_R\overset{\approx}{X}_h/{}_R\overset{\approx}{Y}_h)\,{}_RW_h \; {}_R\overset{\approx}{Y}_h$$

$${}_R\triangle_{52} = \Sigma_h \; {}_RW_h \; {}_R\overset{\approx}{Y}_h/{}_R\overset{\approx}{Y}_h^{\,2} \; (1/{}_An_h - 1/{}_AN_h) \; [{}_A\overset{\approx}{X}_h/{}_A\overset{\approx}{Y}_h \; {}_A\underline{\sigma}_{hYY} - {}_A\underline{\sigma}_{hXY}]$$

Si noti che la prima componente di distorsione dipende dalla media 'relativa' di cella della variabile X nelle due aree (effetto macro-area come per lo stimatore 4) mentre la seconda componente dipende dalla correlazione tra X e Y (effetto ponderazione mediante la variabile aleatoria Y come per lo stimatore 3).

La varianza è approssimabile con

$$\begin{aligned} Var({}_R\hat{X}_5) &= E(Var({}_R\hat{X}_5|{}_Rn_h)) + Var(E({}_R\hat{X}_5|{}_Rn_h) \\ &\approx \Sigma_h \; {}_RW_h^{\,2} \; {}_R\overset{\approx}{Y}_h^{\,2} \; {}_A\overset{\approx}{X}_h^{\,2}/{}_A\overset{\approx}{Y}_h^{\,2} \; E(1/{}_An_h - 1/{}_AN_h) \\ &\qquad [{}_A\underline{\sigma}_{hXX}/{}_A\overset{\approx}{X}_h^{\,2} + {}_A\underline{\sigma}_{hYY}/{}_A\overset{\approx}{Y}_h^{\,2} - 2\,{}_A\underline{\sigma}_{XY}/({}_A\overset{\approx}{X}_h \; {}_A\overset{\approx}{Y}_h)] \\ &\quad + Var({}_R\triangle_{52}) \end{aligned}$$

Si noti che per A=R si ottengono, come caso particolare, le formule relative allo stimatore ${}_R\hat{X}_3$.

NOTE

1. Si consideri il caso di campionamento semplice in blocco da una popolazione di numerosità N. La varianza dello stimatore $\hat{X}$ della media della variabile Y è $Var(\hat{X}) = \sigma^2/n(N-n)/(N-1) \equiv V$ ove $\sigma^2 \equiv Var(Y)$ varianza di Y nella popolazione e n è la numerosità campionaria. La numerosità campionaria n° richiesta per ottenere una certa varianza V° dello stimatore $\hat{X}$ (e quindi una certa 'qualità' delle stime) risulta quindi $n^\circ = 1/(V^\circ/\sigma^2(N-1)/N + 1/N)$ che sale a $\sigma^2/V^\circ (>1)$ al crescere di N. Se i costi di rilevazione crescono linearmente con la numerosità campionaria n (a fortiori nel caso di concavità nella funzione di costo) mentre le risorse crescono in misura (più che) proporzionale alla numerosità dell'universo N, un'indagine su un'area più ampia comporterà, ceteris paribus, un costo complessivo proporzionalmente minore e la domanda di informazione statistica locale non troverà, a livello locale, risorse sufficienti (v. Fig.1). In tal senso si potrebbe dire che la domanda di informazione su aree ridotte non è sempre una 'domanda effettiva' in senso smithiano.

2. Significative eccezioni sono costituite ad esempio da "Congiuntura industriale in Emilia Romagna" a cura di Unioncamere, Federcasse e Federindustria regionali e da "Congiuntura economica regionale" a cura di Unioncamere e Regione Lombardia che forniscono trimestralmente stime quantitative accurate sulla base di campioni statistici di imprese regionali.

3. V. Banca d'Italia (1990), nota alla Tav.B7 in Appendice (p.226) e Barca, Cannari e Di Benedetto (1987).

4. Ai fini dell'esercizio sono state definite 17 regioni poichè alcune regioni amministrative (Piemonte e Val d'Aosta, Abruzzo e Molise, Basilicata e Calabria) sono state accorpate per mancanza di un numero sufficiente di osservazioni.

5. Sulle proprietà della post-stratificazione si vedano Jagers, Odén e Trulsson (1985) e Jagers (1986), nonché gli stimatori proposti da Fuller (1966).

6. Quattro settori per due classi dimensionali. I settori sono: industrie di base (metallurgiche, minerali non metalliferi, chimiche e affini: classi Istat da 21 a 26); meccaniche (inclusi mezzi di trasporto: classi da 31 a 37); moda (tessili, abbigliamento, pelli e cuoio, calzature: classi 43-44-45); altre (alimentari,legno e

mobilio, carta e stampa, gomma e plastica, etc.: classi 41-42 e da 46 a 49). Le classi dimensionali sono: da 50 a 199 addetti; da 200 a 499 addetti. Richiedendosi, ai fini della post-stratificazione, celle campionarie non vuote, per certe regioni il numero dei campioni è risultato in realtà inferiore a 50 (v. più oltre).

7. Per la stima delle variabili occupazionali la ponderazione è stata fatta con gli addetti dell'anno precedente.

8. Nei casi concreti le aree Istat potrebbero presentare un grado di omogeneità interna superiore a quello presente nell' universo fittizio adottato, con effetti di annullamento o riduzione della distorsione degli stimatori corrispondenti (X4 e X5), che dipende dai differenziali di cella area-regione (v. Appendice Metodologica).

9. La procedura di clustering adottata raggruppa progressivamente le regioni 'più vicine' fino alla formazione di un numero prefissato -nel nostro caso, 4 come le aree Istat- di clusters. La 'distanza' è definita dal quadrato della norma euclidea, $d(X_i,X_j) \equiv \|X_i - X_j\|^2$ ove X_i è un vettore di variabili relative all'i-esimo elemento (nel nostro caso investimenti e rapporto investimenti/occupazione nel 1987). La distanza tra cluster è assunta come (metodo 'average linkage') $D_{hk} \equiv \Sigma_i \Sigma_j \; d({}_hX_i, {}_kX_j)/(N_h \; N_k)$ essendo ${}_hX_i$ il livello di X presso l'elemento i del cluster h e N_h il numero di elementi del cluster h. La classificazione ottenuta è stata sottoposta a test con l'analisi discriminante, assumendo una distribuzione multinormale entro le classi. Inutile aggiungere che gli specifici clusters ottenuti non hanno validità al di là dell'esercizio di simulazione, dipendendo dalle caratteristiche delle regioni nell'universo fittizio costituito dal campione BI.

10. Tuttavia si noti che la minor varianza rende significativi gli scarti dalla media e quindi distorti gli stimatori sintetici.

11. Si veda il classico lavoro di Markowitz (1959). Un approccio analogo è stato utilizzato da Bates e Granger (1969) per la combinazione di previsioni alternative.

12. Il vincolo di somma a 1 dei pesi non è necessario ed è introdotto per completezza. La soluzione per il caso di somma non vincolata si ottiene ponendo a zero il corrispondente moltiplicatore di Lagrange. Tuttavia il

vincolo consente di escludere lo stimatore degenere $w \equiv 0$ e, nel caso di n stimatori corretti, mantiene la proprietà di non distorsione.

13. Nel caso di n stimatori corretti non perfettamente correlati la soluzione di $\min \frac{1}{2} w'\Sigma w$ dato $w'1=1$ è $w=\Sigma^{-1}1/(1'\Sigma^{-1}1)$ con varianza $1/1'\Sigma^{-1}1$, inferiore alla minima varianza tra gli n stimatori di base.

14. Si consideri, come semplice esempio, il caso di due stimatori 0,1 quest'ultimo corretto con varianza effettiva σ_1^{2*}. Sia $\beta_0^{2*} + \sigma_0^{2*} < \sigma_1^{2*}$.
Per la preferibilità della combinazione espressa dal vettore $w=(w_0,w_1)'$ rispetto allo stimatore-base 0 deve aversi:
$(w'b^*)^2 + w'\Sigma^* w < \beta_0^{2*} + \sigma_0^{2*}$ che da', per $w_0<1$,
$w_0 > (\sigma_1^{2*}-(\beta_0^{2*}+\sigma_0^{2*}))/(\beta_0^{2*}+var(X0-X1))$.

15. Fuller(1966) ha costruito stimatori corretti anche nel caso che per qualche strato si abbia ${}_Rn_h=0$.
Nel caso di conoscenza imprecisa delle numerosità della sub-popolazione si può tener presente che ${}_Rn/{}_RN$ è una stima corretta di n/N e quindi quest'ultima frazione di campionamento può essere utilizzata come stima della prima. Gli errori nella stima della dimensione degli strati possono tuttavia più che compensare i guadagni di precisione della stratificazione (v. Cochran(1977) p.117) con effetti quindi sull'efficienza relativa di disegni campionari alternativi (v.Thomsen, Tesfu e Binder(1986)).

16. La variabile aleatoria ${}_Rn$ ha distribuzione ipergeometrica con media e varianza date da
$E({}_Rn)=n\ {}_RW$, $Var({}_Rn)=n\ {}_RW(1-{}_RW)(N-n)/(N-1)$.
Approssimando nel continuo in serie di Taylor al II ordine:
$E(1/{}_Rn)\approx 1/(n\ {}_RW)+(1-{}_RW)/(n^2\ {}_RW^2)(N-n)/(N-1)$.
Approssimazioni più accurate e le formule esatte si trovano in Stephan(1945) e Kabe(1976). In particolare il troncamento ${}_Rn>0$ richiede un fattore di normalizzazione della distribuzione di probabilità pari a $1-prob({}_Rn=0)$.

17. Nel caso di (sub-)popolazioni finite e campionamento casuale senza reimmissione si ha
$E({}_RX_j) = \Sigma_{p=1(1)RN}\ {}_RX^{\sim}{}_p/{}_RN \equiv {}_RX^{\sim}/{}_RN \equiv {}_RX^{\approx}$

$E({}_RX_{hk}) = \Sigma_{t=1(1)RNh}\ {}_RX^{\sim}{}_{ht}/{}_RN_h \equiv {}_RX^{\sim}{}_h/{}_RN_h \equiv {}_RX^{\approx}{}_h$

$$\mathrm{Cov}({}_RX_{hk},{}_RX_{js}) = E[({}_RX_{hk}-E({}_RX_{hk}))({}_RX_{js}-E({}_RX_{js}))] =$$

$$= E[({}_RX_{hk}-E({}_RX_{hk}))\ E({}_RX_{js}-E({}_RX_{js})\,|\,{}_RX_{hk}-E({}_RX_{hk}))]$$

$$= \begin{cases} 0 & \text{se } h\#j \\ E[({}_RX_{hk}-E({}_RX_{hk})) \\ \quad ((\Sigma_{t\,R}X^{\sim}{}_{ht})-{}_RX_{hk}-E({}_RX_{hk})({}_RN_h-1))/({}_RN_h-1)] = \\ -E[({}_RX_{hk}-E({}_RX_{hk}))^2]/({}_RN_h-1) = -{}_R\sigma^2{}_h/({}_RN_h-1) & \text{se } h=j,\ k\#s \\ E[({}_RX_{hk}-E({}_RX_{hk}))^2] \equiv {}_R\sigma^2{}_h \equiv {}_R\sigma_{hXX} & \text{se } h=j, k=s \end{cases}$$

In base a tali risultati si può dimostrare che lo stimatore ${}_RS_{hXY}$ è (incondizionatamente) corretto per ${}_R\underline{\sigma}_{hXY}$ ove

$${}_RS_{hXY} \equiv 1/({}_Rn_h-1)\Sigma_k({}_RX_{hk}-{}_RX^{\cdot h})({}_RY_{hk}-{}_RY^{\cdot h}) \quad \text{e}$$

$${}_R\underline{\sigma}_{hXY} \equiv 1/({}_RN_h-1)\Sigma_k({}_RX^{\sim}{}_{hk}-E({}_RX_{hk}))({}_RY_{hk}-E({}_RY_{hk})) =$$

$$= {}_R\sigma_{hXY}\ {}_RN_h/({}_RN_h-1).$$

Si noti che per ${}_RN_h \uparrow \infty$ si ottiene il risultato relativo al caso di popolazioni infinite (o di campionamento con reimmissione).

18. La variabile aleatoria condizionata ${}_Rn_h|{}_Rn$ ha distribuzione ipergeometrica con media e varianza date da $E({}_Rn_h|{}_Rn)={}_Rn\ {}_RW_h$, $\mathrm{Var}({}_Rn_h|{}_Rn)={}_Rn\ {}_RW_h(1-{}_RW_h)({}_RN-{}_Rn)/({}_RN-1)$. Pertanto, usando l'approssimazione in serie di Taylor al II ordine si ottiene
$E(1/{}_Rn_h|{}_Rn)=1/({}_Rn\ {}_RW_h)+(1-{}_RW_h)/({}_Rn^2\ {}_RW_h{}^2)({}_RN-{}_Rn)/({}_RN-1)$

19. Vale la scomposizione della varianza 'corretta'
${}_R\underline{\sigma}^2 = \Sigma_h\ ({}_RN_h-1)/({}_RN-1){}_R\underline{\sigma}^2{}_h + \Sigma_h\ {}_RN_h/({}_RN-1)({}_RX^{\approx}{}_h-{}_RX^{\approx})^2$

20. Si consideri una funzione di 2 v.a. $g(X,Y)$. Espandendo g in serie di Taylor intorno ai valori medi X_0,Y_0 e trascurando i termini di grado superiore al secondo si ottengono le approssimazioni per media e varianza:
$E(g(X,Y)) \approx g(X_0,Y_0) + \frac{1}{2}g_{XX}\ \sigma_{XX} + \frac{1}{2}g_{YY}\ \sigma_{YY} + g_{XY}\ \sigma_{XY}$

$$\mathrm{Var}(g(X,Y)) \approx g_X{}^2\sigma_{XX}+g_Y{}^2\sigma_{YY}+2g_X\ g_Y\ \sigma_{XY}$$

$$-(\tfrac{1}{4}g_{XX}{}^2\sigma_{XX}{}^2+\tfrac{1}{4}g_{YY}{}^2\sigma_{YY}+g_{XY}{}^2\sigma_{XY}{}^2+\tfrac{1}{2}g_{XX}\ g_{YY}\ \sigma_{XX}\ \sigma_{YY}$$

$$+g_{XX}\ g_{XY}\ \sigma_{XX}\ \sigma_{XY}+g_{XY}\ g_{YY}\ \sigma_{XY}\ \sigma_{YY})$$

Nel caso in esame $g(X,Y)=X/Y$ per cui

$$E(X/Y) \approx X_0/Y_0 + [X_0/Y_0^3\sigma_{YY}-\sigma_{XY}/Y_0^2]$$

$$\begin{aligned}Var(X/Y) \approx\ & (X_0/Y_0)^2(\sigma_{XX}/X_0^2+\sigma_{YY}/Y_0^2-2\sigma_{XY}/X_0Y_0) \\ & - [X_0/Y_0^3\sigma_{YY}-1/Y_0^2\sigma_{XY}]^2\end{aligned}$$

ove i termini in parentesi quadre sono caratteristici dell'approssimazione quadratica e quindi non appaiono in Hansen,Hurwitz e Madow (1953, vol.I, ch.4, §18) che si limitano all'approssimazione lineare. Qui si usa l'approssimazione quadratica per il valor medio e quella lineare per la varianza (cfr. Mood, Graybill e Boes (1974) p.181).

21. Negli stimatori ponderati con la variabile Y, anche se $Y^{\sim}_h$ e $Y^{\approx}_h$ sono assunti noti le varianze $\underline{\sigma}_{hYY}$ sono generalmente da stimare sulla base dei dati campionari.

FIG. 1

MUMEROSITA' CAMPIONARIA (COSTI) E RISORSE AL CRESCERE DELL'UNIVERSO

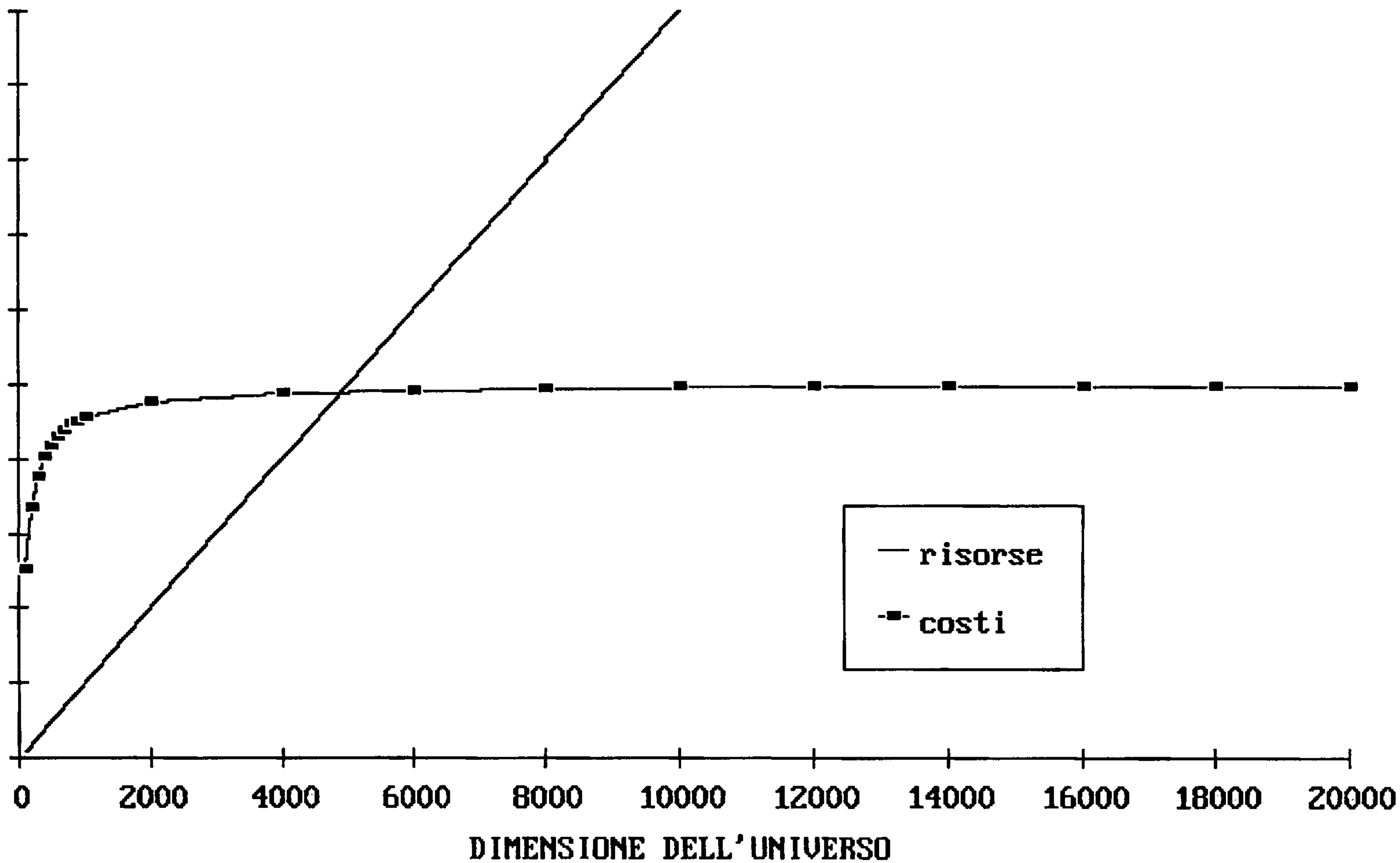


FIG. 2a

BIAS (IN V.A.) E DEVIAZIONE STANDARD DEGLI INVESTIMENTI (GRANDI REGIONI)

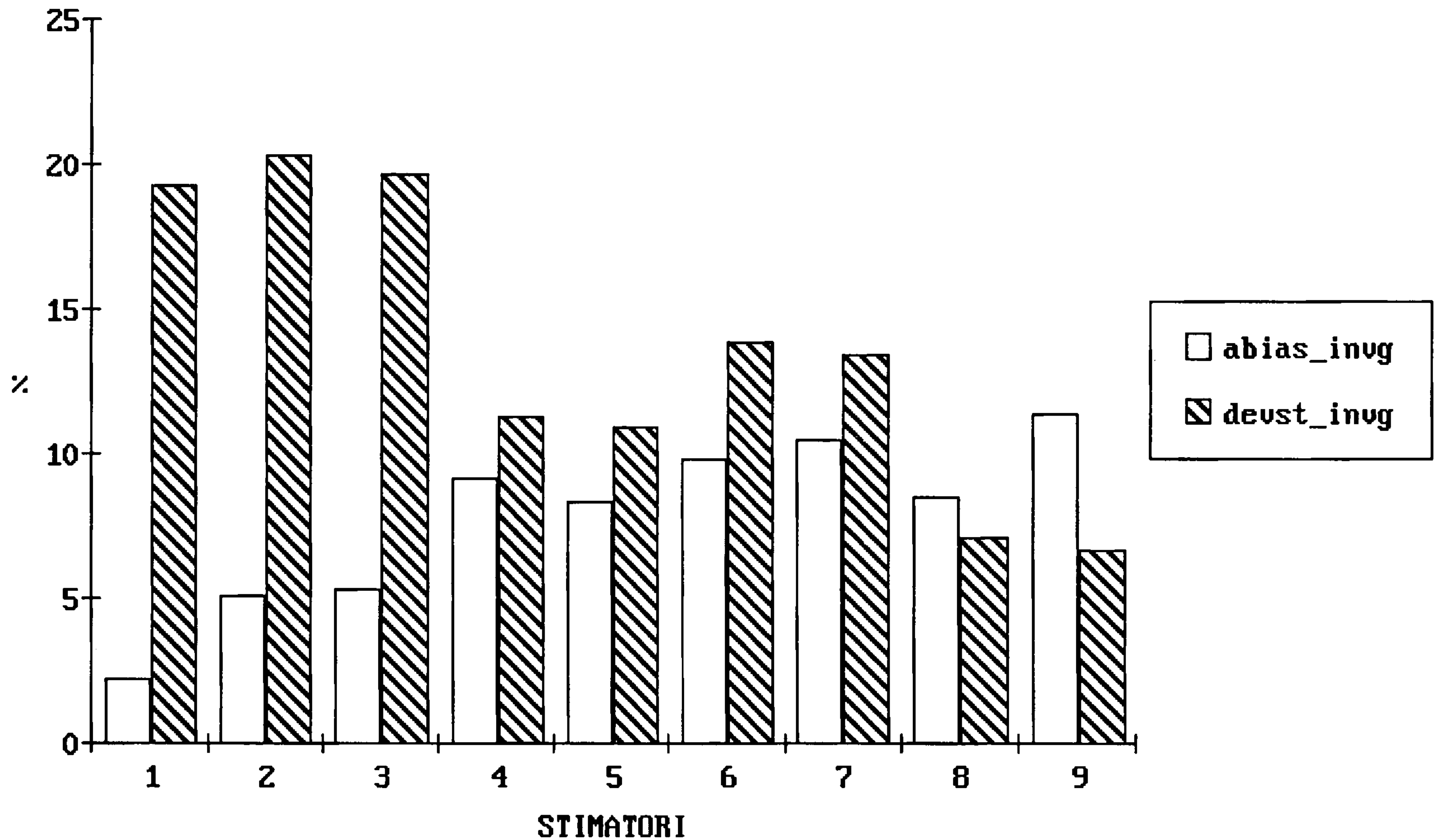


FIG. 2b

BIAS (IN V.A.) E DEVIAZIONE STANDARD DELL'OCCUPAZIONE (GRANDI REGIONI)

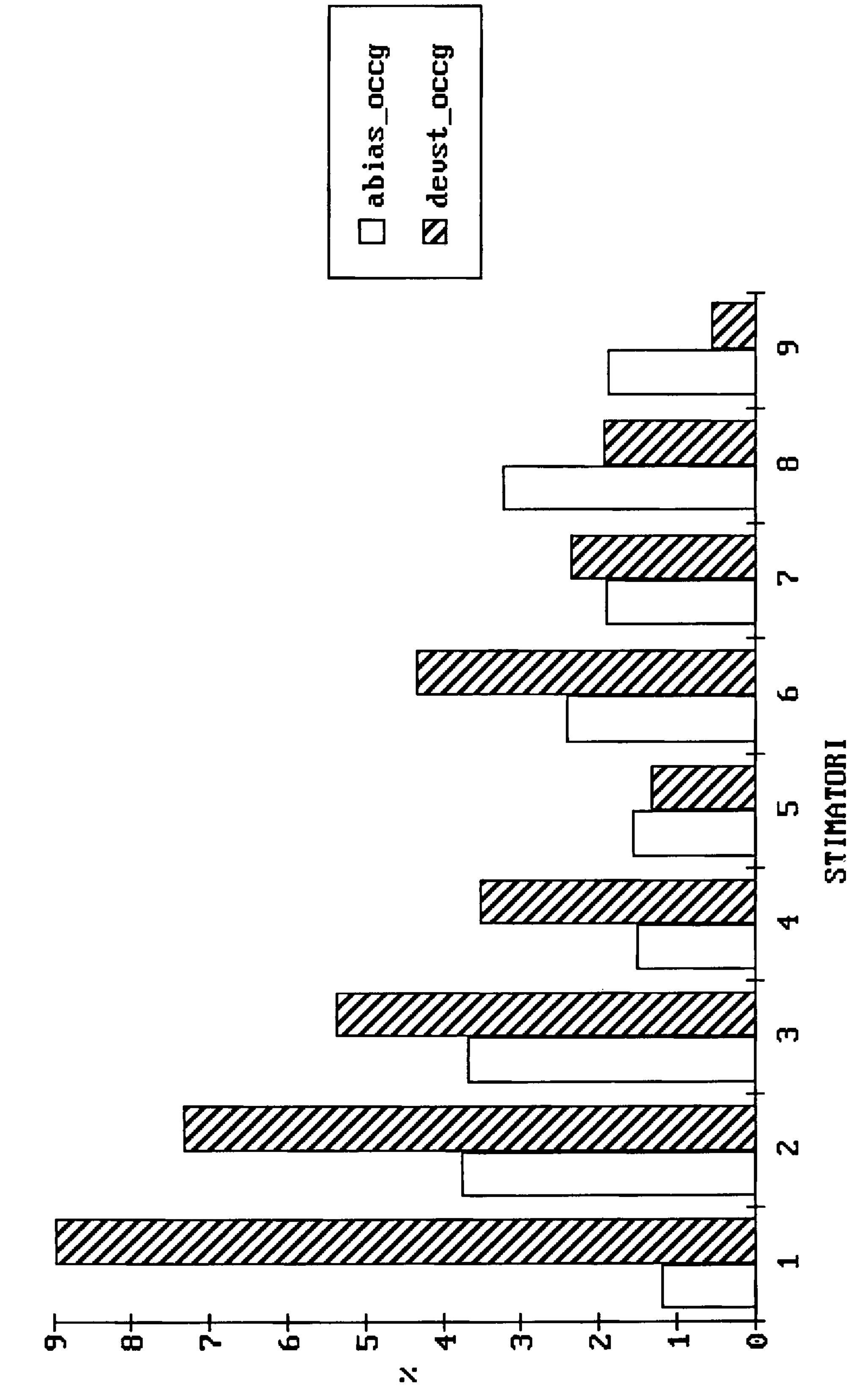


FIG. 2c

BIAS (IN V.A.) E DEVIAZIONE STANDARD DEL FATTURATO (GRANDI REGIONI)

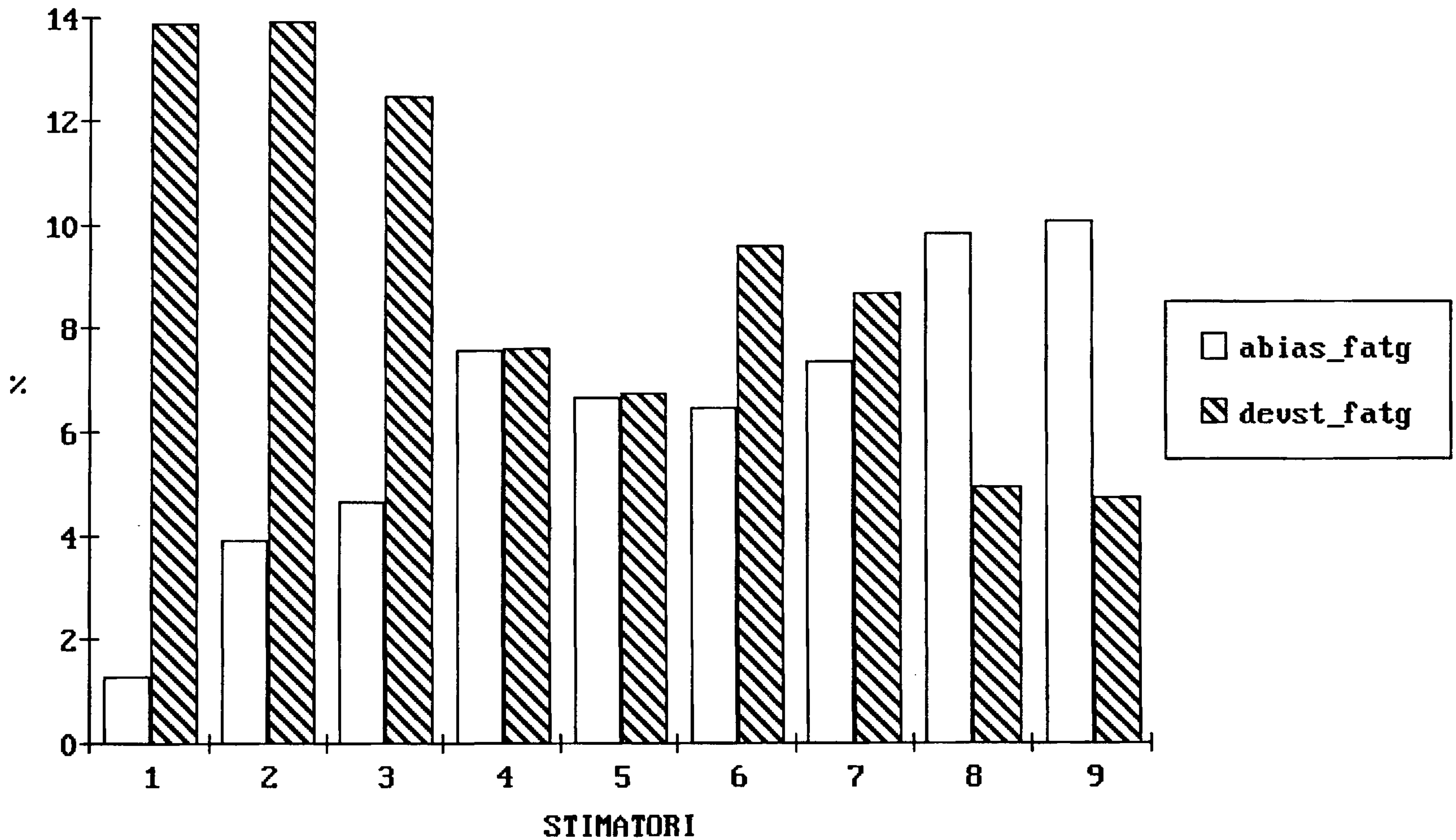


FIG. 3a

BIAS (IN V.A.) E DEVIAZIONE STANDARD DEGLI INVESTIMENTI (PICCOLE REGIONI)

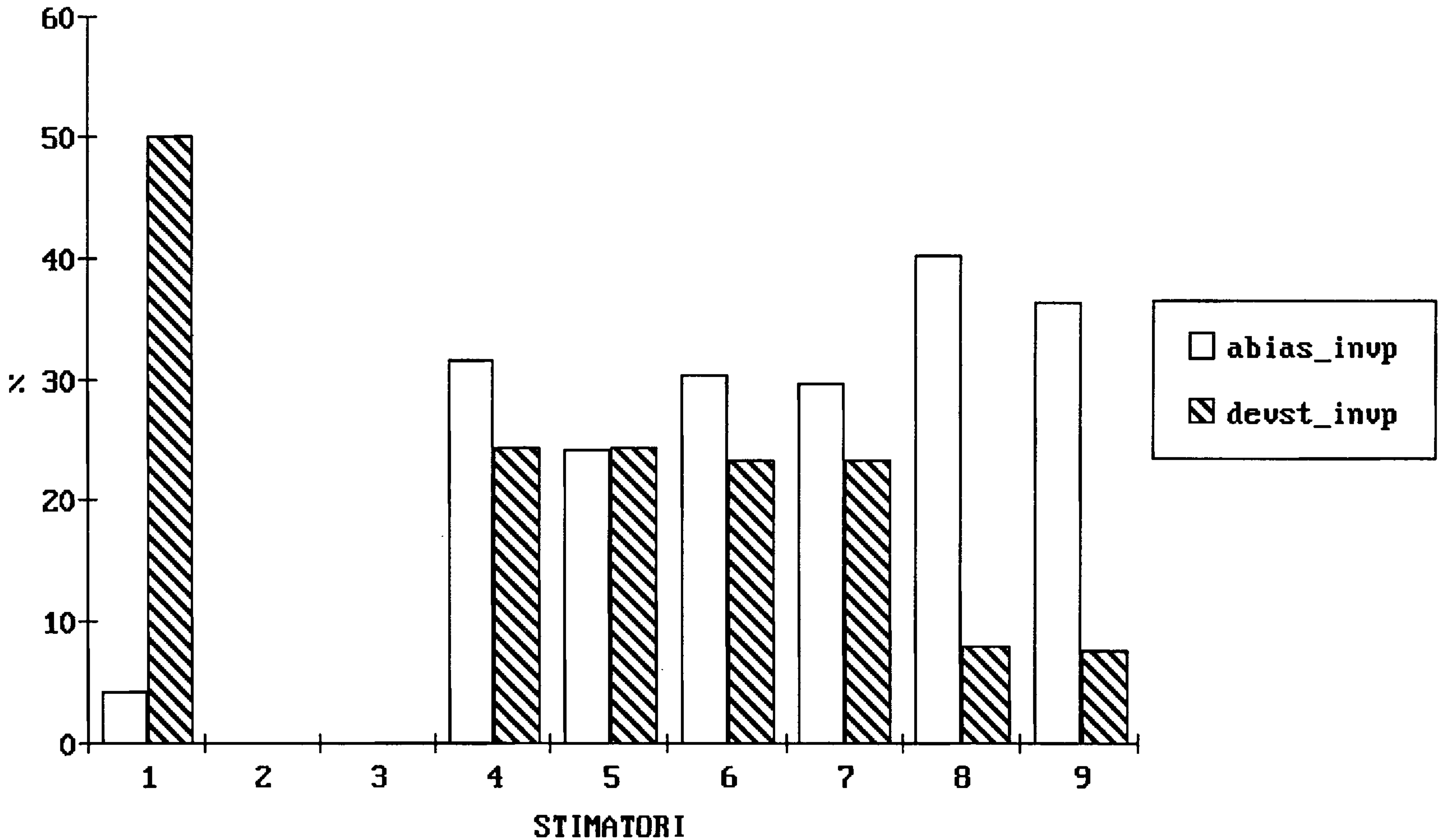


FIG. 3b

BIAS (IN V.A.) E DEVIAZIONE STANDARD DELL'OCCUPAZIONE (PICCOLE REGIONI)

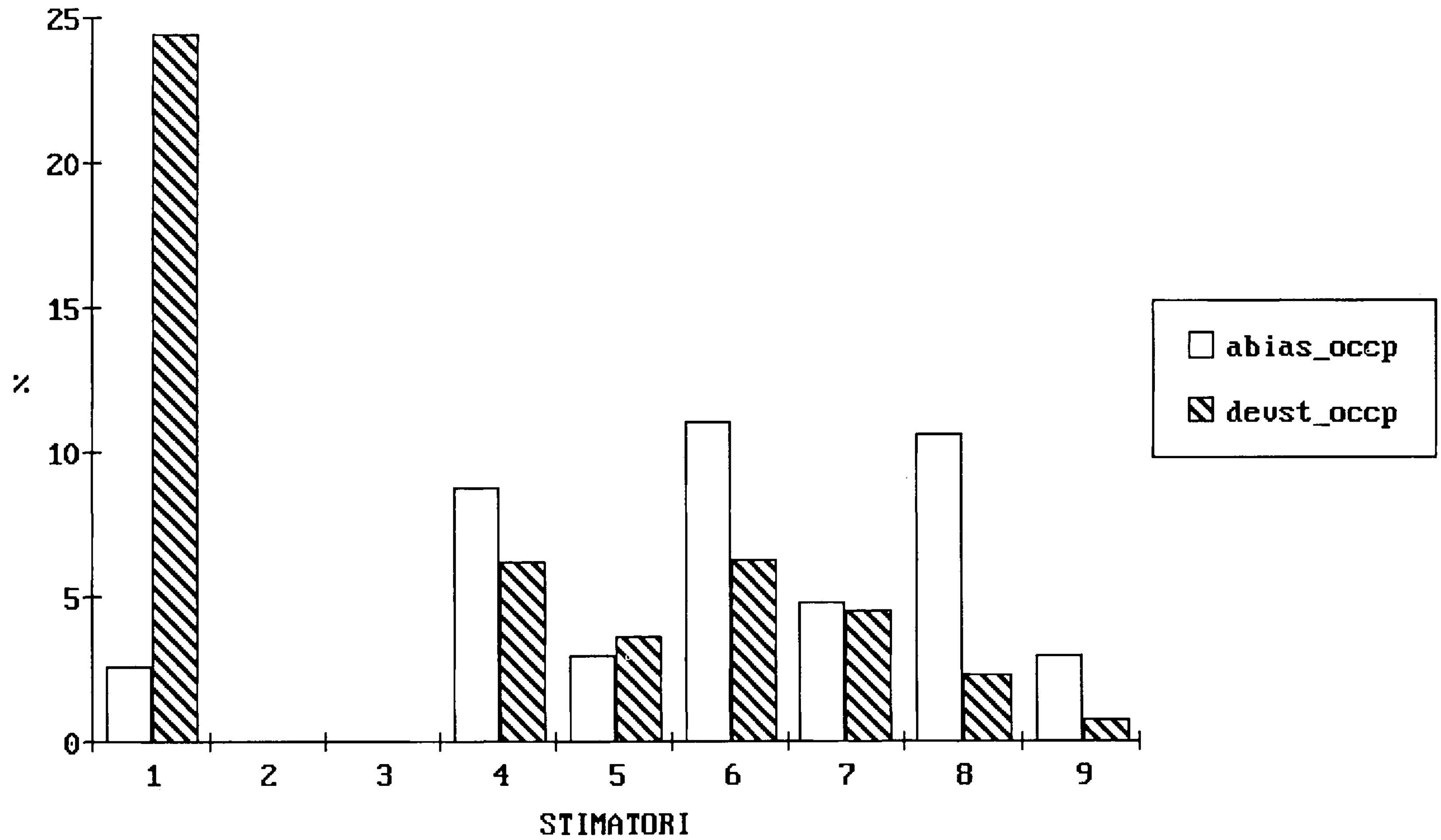


FIG. 3c

BIAS (IN V.A.) E DEVIAZIONE STANDARD DEL FATTURATO (PICCOLE REGIONI)

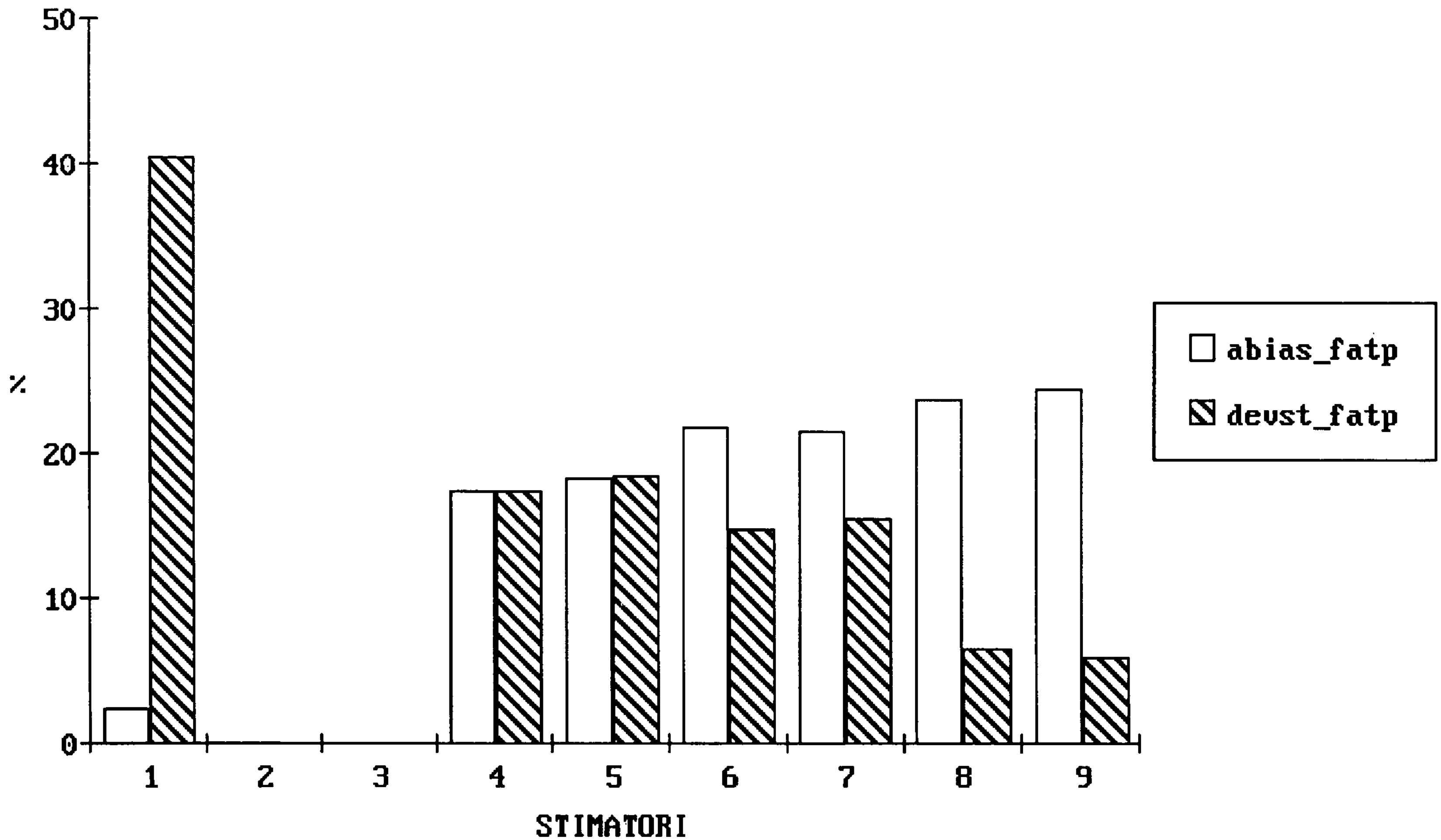


FIG. 4

FRONTIERA DEGLI STIMATORI EFFICIENTI E CURVE DI INDIFFERENZA

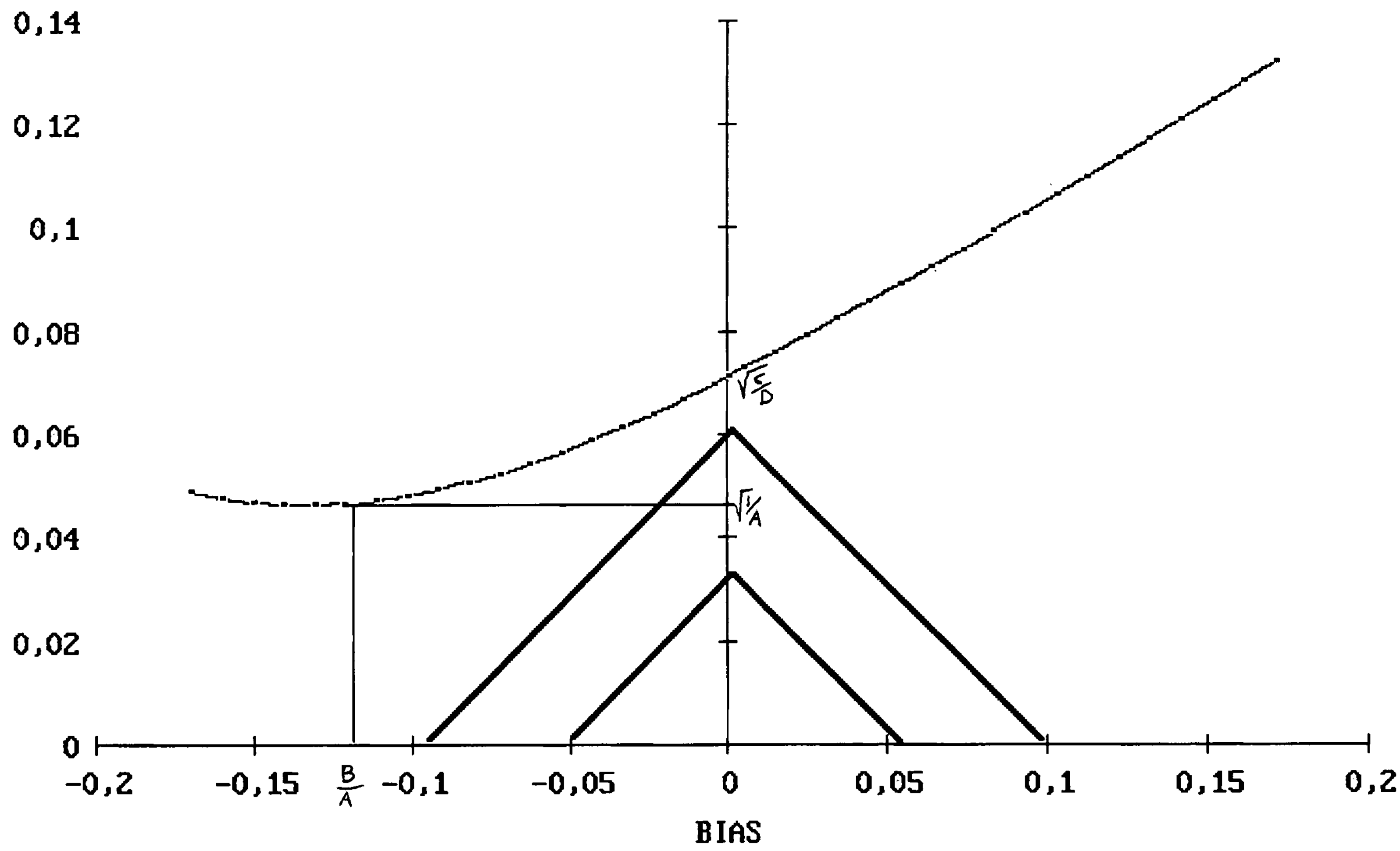
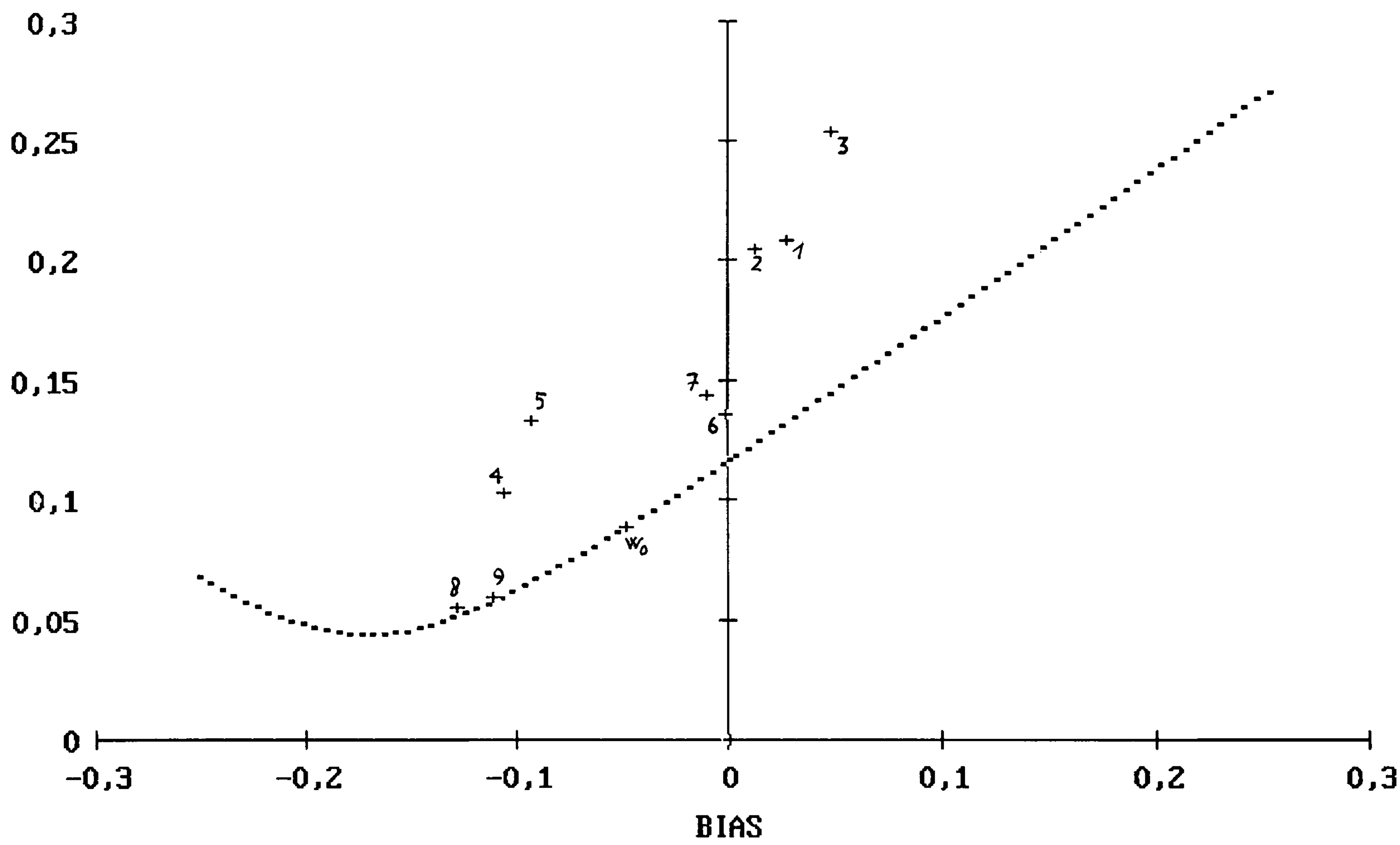


FIG. 5

FRONTIERA DEGLI STIMATORI EFFICIENTI DEGLI INVESTIMENTI '89



Tav. (1)

TIPREG=GRA

STAT	STIMAT	INV86 MEAN	INVP87 MEAN	INV87 MEAN	INVP88 MEAN	OCC87 MEAN	OCCP88 MEAN	FAT86 MEAN	FAT87 MEAN
ABIAS	1	2.01	3.26	3.93	2.31	0.92	0.86	1.55	1.44
	2	3.59	9.41	10.03	3.64	3.27	3.30	2.62	2.51
	3	3.29	10.24	10.60	3.28	3.10	3.03	2.71	2.86
	4	8.59	24.22	24.49	6.36	1.13	1.29	9.47	9.99
	5	9.01	26.69	26.34	5.33	0.65	0.56	9.68	10.52
	6	11.42	15.91	12.24	16.22	3.99	4.08	5.28	6.16
	7	12.09	16.23	9.54	14.45	0.29	0.44	5.99	6.52
	8	16.24	24.50	21.59	13.27	4.45	4.12	9.33	9.77
	9	17.13	25.79	21.31	12.65	0.76	0.88	10.93	11.69
DEVST	1	20.80	28.56	30.20	18.90	9.82	9.63	17.45	16.30
	2	22.23	22.10	23.50	19.31	7.56	7.37	17.28	16.43
	3	20.09	20.36	21.35	17.95	5.08	5.12	15.93	14.32
	4	11.97	26.77	26.99	12.37	3.29	3.28	10.95	10.46
	5	11.97	29.04	28.70	12.46	1.28	1.30	10.47	9.59
	6	14.70	15.04	13.72	15.78	3.83	3.74	11.12	10.36
	7	14.80	15.48	12.86	15.34	1.79	1.52	10.32	9.05
	8	8.04	12.77	12.73	6.26	1.79	1.83	4.97	4.72
	9	7.82	12.87	12.78	6.51	0.58	0.53	4.78	4.51
RMSE	1	20.95	28.84	30.57	19.05	9.88	9.69	17.61	16.46
	2	22.77	25.31	27.03	20.04	8.47	8.36	17.77	16.91
	3	20.78	23.84	25.23	18.91	6.17	6.18	16.47	14.87
	4	15.84	38.78	39.23	14.33	3.52	3.71	15.28	15.29
	5	15.86	42.45	41.88	13.84	1.49	1.51	14.82	14.84

TIPREG=GRA

STAT	STIMAT	INV86 MEAN	INVP87 MEAN	INV87 MEAN	INVP88 MEAN	OCC87 MEAN	OCCP88 MEAN	FAT86 MEAN	FAT87 MEAN
RMSE	6	19.58	22.62	18.70	23.10	6.06	5.98	12.65	12.17
	7	19.30	23.29	17.45	21.28	1.82	1.59	12.31	11.71
	8	19.92	29.00	27.23	15.13	4.86	4.70	11.44	11.80
	9	20.00	30.44	27.43	14.29	1.00	1.09	12.15	12.86
TSTUD	1	0.64	0.76	1.09	0.82	0.67	0.64	0.83	0.84
	2	0.94	2.63	2.75	1.30	2.28	2.38	1.11	1.13
	3	0.92	3.12	3.17	1.49	3.08	2.99	1.40	1.57
	4	5.06	8.59	9.19	3.84	2.66	3.35	5.18	5.55
	5	5.02	9.51	10.08	3.14	3.93	4.56	5.95	6.65
	6	4.54	6.69	5.66	7.26	9.48	9.60	3.80	4.45
	7	5.24	6.40	4.26	6.99	1.85	2.23	3.66	4.34
	8	12.96	24.99	22.71	14.32	17.64	15.78	13.39	14.56
	9	14.67	26.14	23.46	13.57	9.20	12.96	16.09	18.36

Tav. (1) (cont.)

TIPREG=GRA

STAT	STIMAT	INV87 MEAN	INVP88 MEAN	INV88 MEAN	INVP89 MEAN	OCC88 MEAN	OCCP89 MEAN	FAT87 MEAN	FAT88 MEAN
ABIAS	1	1.98	1.98	2.27	2.05	1.20	1.21	1.58	1.31
	2	7.24	4.78	5.12	5.14	3.78	3.78	3.67	3.93
	3	7.02	5.09	5.32	5.38	3.71	4.06	4.34	4.67
	4	17.65	7.21	9.14	12.65	1.53	2.39	6.59	7.60
	5	15.70	5.09	8.35	12.41	1.58	1.48	6.32	6.71
	6	10.96	10.31	9.83	14.74	2.44	3.26	4.05	6.48
	7	9.70	10.23	10.49	12.59	1.92	1.54	6.58	7.40
	8	15.97	9.24	8.54	17.33	3.26	3.58	7.55	9.87
	9	18.67	8.97	11.34	15.85	1.91	1.50	8.68	10.12
DEVST	1	31.43	17.30	19.28	17.93	9.00	8.95	14.12	13.90
	2	28.02	18.63	20.36	17.98	7.36	7.49	13.99	13.95
	3	27.99	16.96	19.66	17.91	5.39	5.26	12.65	12.52
	4	20.21	9.95	11.32	10.22	3.53	3.64	7.93	7.63
	5	20.07	9.87	10.97	10.37	1.34	1.23	7.19	6.77
	6	20.05	12.54	13.90	13.60	4.36	4.60	9.09	9.61
	7	20.40	12.10	13.41	13.33	2.37	2.12	8.28	8.71
	8	10.82	6.47	7.10	6.82	1.96	2.04	4.72	4.97
	9	11.08	6.26	6.68	6.60	0.58	0.48	4.57	4.77
RMSE	1	31.54	17.47	19.47	18.10	9.12	9.08	14.23	13.99
	2	29.19	19.29	21.09	18.75	8.53	8.68	14.59	14.67
	3	29.08	17.89	20.49	18.93	6.76	6.86	13.52	13.51
	4	28.27	12.64	14.65	17.95	3.90	4.58	10.64	11.24
	5	26.86	11.70	13.99	18.07	2.27	2.05	9.94	10.08

TIPREG=GRA

STAT	STIMAT	INV87 MEAN	INVP88 MEAN	INV88 MEAN	INVP89 MEAN	OCC88 MEAN	OCCP89 MEAN	FAT87 MEAN	FAT88 MEAN
RMSE	6	23.09	17.00	17.92	21.07	5.56	6.13	10.15	11.98
	7	22.95	16.26	17.51	20.15	3.16	2.74	11.45	12.43
	8	21.74	11.69	12.04	19.68	3.84	4.14	9.09	11.16
	9	23.36	11.24	13.53	18.52	2.12	1.68	10.27	11.53
TSTUD	1	0.57	0.93	0.96	0.93	1.12	1.15	0.87	0.75
	2	1.49	1.66	1.52	1.85	2.85	2.72	1.53	1.67
	3	1.84	2.00	1.75	2.06	3.52	4.01	2.07	2.24
	4	8.22	5.35	5.49	8.49	2.97	4.24	5.64	7.02
	5	6.62	4.02	5.19	8.35	7.09	7.21	6.31	7.27
	6	3.61	5.52	4.76	7.36	5.77	7.02	3.45	5.31
	7	3.03	6.63	5.75	5.43	6.01	4.65	6.81	7.20
	8	20.68	9.97	8.05	16.17	11.79	12.45	11.01	13.67
	9	22.71	10.11	11.38	15.04	20.75	20.18	13.05	14.49

Tav. (1) (cont.)

TIPREG=GRA

STAT	STIMAT	INV88 MEAN	INVP89 MEAN	INV89 MEAN	INVP90 MEAN	OCC89 MEAN	OCCP90 MEAN	FAT88 MEAN	FAT89 MEAN
ABIAS	1	2.23	2.96	2.27	3.13	0.88	0.91	0.72	0.80
	2	3.38	3.89	3.42	3.26	3.46	3.44	3.10	2.99
	3	4.15	4.62	4.72	4.54	3.49	3.44	3.57	3.36
	4	11.04	11.59	13.64	20.96	2.54	2.08	20.91	20.18
	5	9.12	10.19	11.85	18.76	0.57	0.89	17.47	16.72
	6	16.27	18.82	19.15	24.12	5.12	4.64	19.01	17.47
	7	11.28	14.02	14.12	18.07	1.03	0.96	14.14	13.76
	8	15.59	17.86	23.73	26.55	4.96	4.54	22.47	21.43
	9	12.71	14.09	18.49	21.40	0.97	0.76	19.05	18.65
DEVST	1	21.04	23.20	20.84	21.03	8.77	8.68	14.39	13.38
	2	19.35	21.89	20.52	20.80	6.91	6.93	13.78	12.97
	3	18.14	20.41	21.29	18.37	5.10	4.97	12.57	11.93
	4	11.68	14.69	12.53	18.07	3.37	3.42	11.33	10.20
	5	10.87	13.77	12.62	16.63	1.61	1.39	9.75	8.50
	6	14.50	15.19	13.92	17.85	3.83	3.78	11.48	10.12
	7	12.95	14.29	13.44	15.28	1.72	1.58	10.09	8.82
	8	7.67	8.51	7.54	9.84	1.60	1.59	6.15	5.41
	9	7.02	8.21	7.03	8.85	0.54	0.41	5.91	5.05
RMSE	1	21.19	23.39	20.99	21.31	8.85	8.76	14.42	13.42
	2	19.82	22.48	21.02	21.28	7.88	7.90	14.52	13.75
	3	18.89	21.39	22.25	19.18	6.28	6.14	13.46	12.78
	4	16.55	19.75	19.36	31.09	4.61	4.24	25.28	23.78
	5	14.48	17.82	17.97	27.89	1.74	1.75	21.31	19.70

TIPREG=GRA

STAT	STIMAT	INV88 MEAN	INVP89 MEAN	INV89 MEAN	INVP90 MEAN	OCC89 MEAN	OCCP90 MEAN	FAT88 MEAN	FAT89 MEAN
RMSE	6	22.34	26.37	26.27	32.79	7.23	6.76	23.49	21.82
	7	17.73	21.99	21.90	26.16	2.08	1.91	19.11	17.95
	8	18.14	20.57	25.07	29.96	5.36	4.94	24.04	22.80
	9	14.87	17.08	20.20	24.78	1.15	0.93	21.10	20.02
TSTUD	1	0.67	0.87	0.82	0.93	0.78	0.78	0.39	0.44
	2	1.26	1.10	1.09	1.26	2.97	2.92	1.74	1.73
	3	1.61	1.40	1.58	1.82	3.89	3.93	1.96	1.85
	4	7.08	5.74	8.00	8.11	4.82	4.02	13.38	13.99
	5	5.93	5.12	6.42	8.34	2.54	3.52	13.05	14.27
	6	8.00	9.31	10.86	9.92	11.67	10.75	9.52	9.83
	7	5.90	6.84	8.57	8.50	5.06	4.97	8.46	9.31
	8	13.99	13.67	20.05	15.76	21.27	19.66	25.85	27.65
	9	12.89	11.59	16.35	14.40	12.64	13.30	22.28	25.30

Tav. (2)

TIPREG=PIC

STAT	STIMAT	INV86 MEAN	INVP87 MEAN	INV87 MEAN	INVP88 MEAN	OCC87 MEAN	OCCP88 MEAN	FAT86 MEAN	FAT87 MEAN
ABIAS	1	3.47	4.81	3.51	4.20	2.27	2.19	2.41	2.13
	2	.	.	.	.	.	.	.	.
	3	.	.	.	.	.	.	.	.
	4	36.35	37.81	28.38	27.84	7.96	8.55	19.61	12.62
	5	35.85	37.46	26.66	24.01	2.36	2.67	25.24	15.64
	6	21.95	22.63	18.61	22.13	9.98	10.25	17.47	17.13
	7	17.66	31.75	23.83	21.00	4.92	4.65	24.18	16.27
	8	41.23	34.41	26.24	35.13	9.39	9.91	21.41	17.37
	9	41.43	37.49	28.43	32.91	2.05	2.05	26.48	17.14
DEVST	1	47.36	48.43	39.43	46.97	23.79	23.25	34.82	34.70
	2	.	.	.	.	.	.	.	.
	3	.	.	.	.	.	.	.	.
	4	22.89	26.69	26.55	19.14	6.14	6.08	15.11	14.12
	5	24.47	28.28	27.06	19.92	3.28	3.27	15.31	14.26
	6	16.16	36.58	39.56	16.74	6.85	6.87	19.53	16.95
	7	18.83	44.09	44.97	16.77	4.59	4.42	21.58	17.25
	8	8.58	13.21	14.37	6.56	2.12	2.17	7.33	6.49
	9	8.54	13.82	14.45	6.79	0.71	0.62	6.75	5.91
RMSE	1	47.59	48.78	39.72	47.28	23.97	23.42	34.95	34.82
	2	.	.	.	.	.	.	.	.
	3	.	.	.	.	.	.	.	.
	4	44.91	48.77	41.34	36.39	10.86	11.18	27.60	20.35
	5	45.31	50.23	40.56	34.14	4.43	4.49	31.69	21.87

TIPREG=PIC

STAT	STIMAT	INV86 MEAN	INVP87 MEAN	INV87 MEAN	INVP88 MEAN	OCC87 MEAN	OCCP88 MEAN	FAT86 MEAN	FAT87 MEAN
RMSE	6	29.14	46.92	45.14	32.15	13.06	13.24	28.50	26.67
	7	29.40	57.98	52.16	31.20	6.90	6.65	35.01	26.91
	8	43.52	38.45	33.73	36.33	9.74	10.34	23.47	19.80
	9	43.16	41.00	34.79	34.54	2.25	2.19	28.25	19.24
TSTUD	1	0.52	0.69	0.75	0.62	0.76	0.76	0.45	0.42
	2	.	.	.	.	.	.	.	.
	3	.	.	.	.	.	.	.	.
	4	12.34	16.68	8.63	17.30	11.56	12.37	8.16	7.01
	5	10.81	14.69	7.45	12.97	9.16	8.55	9.33	7.43
	6	13.61	9.91	3.98	15.34	12.24	12.59	7.59	8.61
	7	12.47	12.60	4.91	14.78	10.17	9.46	10.37	10.11
	8	37.76	19.87	13.38	52.49	29.98	31.15	19.68	19.42
	9	39.69	20.07	14.53	43.75	21.01	23.83	24.36	21.18

Tav. (2) (cont.)

TIPREG=PIC

STAT	STIMAT	INV87	INVP88	INV88	INVP89	OCC88	OCCP89	FAT87	FAT88
		MEAN	MEAN	MEAN	MEAN	MEAN	MEAN	MEAN	MEAN
ABIAS	1	3.91	5.35	4.29	3.16	2.62	2.69	2.56	2.44
	2	.	.	.	.	.	.	.	.
	3	.	.	.	.	.	.	.	.
	4	41.92	26.09	31.68	27.57	8.79	8.78	19.14	17.52
	5	38.10	25.03	24.31	23.62	2.98	2.69	22.28	18.35
	6	32.58	26.56	30.52	27.46	11.09	11.53	22.59	21.80
	7	33.98	29.06	29.79	21.12	4.81	4.92	21.81	21.55
	8	41.98	34.96	40.34	27.32	10.63	10.59	23.85	23.80
	9	34.69	34.37	36.45	20.17	3.00	2.52	25.38	24.45
DEVST	1	45.05	48.93	50.20	47.09	24.44	24.68	38.17	40.51
	2	.	.	.	.	.	.	.	.
	3	.	.	.	.	.	.	.	.
	4	32.82	19.92	24.44	18.87	6.23	6.35	18.44	17.49
	5	37.97	21.29	24.52	18.64	3.65	3.13	20.29	18.56
	6	45.63	17.51	23.45	15.88	6.29	6.46	14.30	14.82
	7	50.96	17.47	23.42	16.72	4.54	4.23	15.22	15.49
	8	16.47	6.53	8.12	6.65	2.33	2.39	6.34	6.63
	9	15.34	6.36	7.66	6.38	0.78	0.59	5.89	6.04
RMSE	1	45.40	49.69	50.46	47.23	24.63	24.88	38.31	40.62
	2	.	.	.	.	.	.	.	.
	3	.	.	.	.	.	.	.	.
	4	55.95	36.85	42.55	35.42	11.24	11.29	28.52	27.29
	5	57.53	36.26	37.62	31.88	5.21	4.32	31.79	27.86

TIPREG=PIC

STAT	STIMAT	INV87	INVP88	INV88	INVP89	OCC88	OCCP89	FAT87	FAT88
		MEAN	MEAN	MEAN	MEAN	MEAN	MEAN	MEAN	MEAN
RMSE	6	58.26	34.24	41.21	33.39	13.96	14.12	29.12	29.43
	7	62.56	36.00	40.60	28.59	7.07	7.04	30.10	29.74
	8	47.09	36.11	41.83	28.90	11.21	11.03	25.27	25.03
	9	40.32	35.37	38.00	22.24	3.22	2.65	26.41	25.49
TSTUD	1	0.76	0.89	0.60	0.52	0.81	0.84	0.46	0.42
	2	.	.	.	.	.	.	.	.
	3	.	.	.	.	.	.	.	.
	4	10.78	18.13	14.73	12.86	11.56	11.13	10.34	9.70
	5	7.68	15.02	10.50	10.77	8.32	8.46	9.27	8.22
	6	7.39	15.71	12.53	16.30	15.93	16.07	15.47	13.74
	7	6.55	20.48	14.31	11.84	11.09	14.56	18.56	17.08
	8	16.13	54.06	45.66	38.51	31.93	30.61	28.88	25.21
	9	13.29	51.45	39.28	28.32	28.40	27.69	35.88	29.55

Tav. (2) (cont.)

TIPREG=PIC

STAT	STIMAT	INV88 MEAN	INVP89 MEAN	INV89 MEAN	INVP90 MEAN	OCC89 MEAN	OCCP90 MEAN	FAT88 MEAN	FAT89 MEAN
ABIAS	1	5.29	4.03	3.56	4.89	3.41	3.40	3.18	3.49
	2	.	.	.	.	.	.	.	.
	3	.	.	.	.	.	.	.	.
	4	32.22	32.13	29.10	31.21	9.89	9.42	17.96	17.29
	5	24.94	28.26	22.45	26.26	2.41	1.72	20.99	20.86
	6	29.51	30.33	29.09	28.87	12.20	11.23	19.54	19.98
	7	26.96	25.78	21.83	27.27	5.09	4.29	22.14	21.10
	8	52.78	29.30	29.80	32.03	12.78	11.91	20.50	19.44
	9	45.40	22.16	19.13	26.11	3.19	1.84	21.92	21.57
DEVST	1	45.78	45.38	44.89	49.55	22.03	22.02	36.67	34.43
	2	.	.	.	.	.	.	.	.
	3	.	.	.	.	.	.	.	.
	4	27.27	23.18	23.20	24.51	5.72	5.73	16.74	15.69
	5	25.12	20.80	21.44	24.62	2.79	2.39	18.68	16.86
	6	20.00	18.92	16.45	24.25	6.41	6.45	17.09	15.80
	7	18.87	17.58	16.91	24.00	4.26	3.86	17.17	15.48
	8	10.81	8.44	8.02	10.19	2.04	2.03	6.99	6.45
	9	10.00	7.72	7.33	9.27	0.71	0.48	6.82	6.16
RMSE	1	46.24	45.66	45.13	49.89	22.39	22.38	36.88	34.69
	2	.	.	.	.	.	.	.	.
	3	.	.	.	.	.	.	.	.
	4	44.51	41.47	38.48	42.58	11.86	11.44	26.30	25.23
	5	36.74	35.87	32.08	38.73	4.01	3.23	29.21	28.12

TIPREG=PIC

STAT	STIMAT	INV88 MEAN	INVP89 MEAN	INV89 MEAN	INVP90 MEAN	OCC89 MEAN	OCCP90 MEAN	FAT88 MEAN	FAT89 MEAN
RMSE	6	38.09	38.97	35.68	41.66	14.39	13.71	29.57	29.17
	7	34.76	33.25	30.22	39.72	7.00	6.04	31.68	29.79
	8	54.64	31.69	31.21	35.71	12.99	12.22	22.79	21.61
	9	47.67	24.96	21.43	29.46	3.32	1.91	23.84	23.15
TSTUD	1	0.94	0.68	0.57	0.73	1.12	1.11	0.65	0.75
	2	.	.	.	.	.	.	.	.
	3	.	.	.	.	.	.	.	.
	4	10.96	9.38	8.76	8.77	13.96	13.38	9.47	7.93
	5	8.26	9.09	7.17	7.63	7.78	9.13	9.47	9.74
	6	14.60	17.72	19.32	14.22	19.97	18.19	11.51	11.74
	7	14.15	16.17	15.88	14.23	14.53	12.46	16.51	15.86
	8	37.21	31.57	30.65	23.80	45.09	42.18	23.38	20.43
	9	32.47	27.86	23.93	22.55	33.48	26.95	24.58	25.39

Tav. (4)

TIPREG=GRA

	TAU_INV	ALFA_INV	TAU_OCC	ALFA_OCC	TAU_FAT	ALFA_FAT
	1987	1987	1987	1987	1987	1987
STIMAT						
4	-4.15	1.00	-0.50	0.00	1.25	0.56
5	-3.14	1.00	-0.42	0.00	1.17	0.54
6	0.23	0.18	0.19	0.16	0.60	0.38
7	-0.05	0.00	-0.52	0.00	0.54	0.35
8	1.07	0.52	0.20	0.17	0.62	0.38
9	1.05	0.51	-0.36	0.00	0.77	0.44

TIPREG=GRA

	TAU_INV	ALFA_INV	TAU_OCC	ALFA_OCC	TAU_FAT	ALFA_FAT
	1989	1989	1989	1989	1989	1989
STIMAT						
4	1.28	0.56	-0.26	0.00	6.20	0.86
5	1.07	0.52	-0.54	0.00	3.07	0.75
6	2.38	0.70	0.54	0.35	5.09	0.84
7	1.51	0.60	-0.47	0.00	2.59	0.72
8	1.56	0.61	0.28	0.22	2.44	0.71
9	1.12	0.53	-0.39	0.00	1.98	0.66

TIPREG=GRA

	TAU_INV	ALFA_INV	TAU_OCC	ALFA_OCC	TAU_FAT	ALFA_FAT
	1988	1988	1988	1988	1988	1988
STIMAT						
4	0.44	0.31	-0.59	0.00	0.58	0.37
5	0.34	0.26	-0.36	0.00	0.39	0.28
6	0.73	0.42	-0.45	0.00	0.59	0.37
7	0.77	0.44	-0.37	0.00	0.66	0.40
8	0.26	0.21	-0.09	0.00	0.66	0.40
9	0.45	0.31	-0.27	0.00	0.67	0.40

Tav. (5)

TIPREG=PIC

STIMAT	TAU_INV 1987	ALFA_INV 1987	TAU_OCC 1987	ALFA_OCC 1987	TAU_FAT 1987	ALFA_FAT 1987
4	1.93	0.66	0.32	0.24	0.51	0.34
5	1.87	0.65	0.00	0.00	0.66	0.40
6	-115.58	1.00	0.45	0.31	0.85	0.46
7	-3.67	1.00	0.14	0.12	0.81	0.45
8	0.91	0.48	0.33	0.25	0.54	0.35
9	1.00	0.50	-0.01	0.00	0.52	0.34

TIPREG=PIC

STIMAT	TAU_INV 1988	ALFA_INV 1988	TAU_OCC 1988	ALFA_OCC 1988	TAU_FAT 1988	ALFA_FAT 1988
4	1.06	0.52	0.34	0.25	0.66	0.40
5	0.78	0.44	0.02	0.02	0.72	0.42
6	0.98	0.50	0.47	0.32	0.75	0.43
7	0.95	0.49	0.11	0.10	0.76	0.43
8	0.86	0.46	0.36	0.27	0.63	0.39
9	0.76	0.43	0.02	0.02	0.64	0.39

TIPREG=PIC

STIMAT	TAU_INV 1989	ALFA_INV 1989	TAU_OCC 1989	ALFA_OCC 1989	TAU_FAT 1989	ALFA_FAT 1989
4	1.18	0.54	0.40	0.28	0.74	0.42
5	0.81	0.45	-0.05	0.00	0.99	0.50
6	0.90	0.47	0.56	0.36	0.89	0.47
7	0.65	0.39	0.09	0.09	0.93	0.48
8	0.71	0.42	0.47	0.32	0.57	0.36
9	0.41	0.29	-0.01	0.00	0.64	0.39

BIBLIOGRAFIA

BANCA D'ITALIA (1990), Assemblea Generale ordinaria dei partecipanti. Relazione del Governatore sull'esercizio 1989, Roma.

BARCA F., CANNARI L. e DI BENEDETTO C. (1987), Il disegno campionario nell'indagine sugli investimenti delle imprese industriali, Roma, Banca d'Italia (dattiloscritto).

BATES J.M. e GRANGER C.W. (1969), The combination of forecasts, "Operational Research Quarterly", 4, pp.451-468

COCHRAN W.G. (1977), Sampling Techniques, 3rd ed., New York, J.Wiley

FULLER W.A. (1966), Estimation Employing Post Strata, "Journal of The American Statistical Society", 61, pp.1172-1183

GORI E. E MARCHETTI G. (1988), Problemi e metodi di analisi di dati statistici per piccole aree, in: "Informazione e analisi statistica per aree subregionali", Società Italiana di Statistica, Convegno 1987, Perugia, Galeno Editrice.

HANSEN M.H., HURWITZ W.N. E MADOW W.G. (1953), Sample Survey Methods and Theory. Vol.I Methods and Applications, Vol.II Theory, New York, J.Wiley

INGERSOLL J.E. (1987), Theory of Financial Decision Making, Totowa (N.J.), Rowman & Littlefield

JAGERS P. (1986), Post-Stratification against Bias in Sampling, "International Statistical Review", 54, 2, pp.159-167

JAGERS P., ODÉN A. E TRULSSON L. (1985), Post-Stratification and Ratio Estimation: Usages of Auxiliarry Information in Survey Sampling and Opinion Polls, "International Statistical Review", 53, 3, pp.221-238

KABE D.G. (1976), Inverse Moments of Discrete Distributions, "The Canadian Journal of Statistics", Section D, 4, 1, pp.133-141

MARKOWITZ H. (1959), Portfolio Selection: Efficient Diversification of Investments, New York, J.Wiley

MOOD A.M.,GRAYBILL F.A. E BOES D.C. (1974), Introduction to the Theory of Statistics, 3rd ed., London, McGraw-Hill

STEINBERG J. (ed.) (1979), Synthetic Estimation for Small Areas, National Institute on Drug Abuse Research Monograph 24, Washington, U.S. Government Print. Off.

STEPHAN F.F. (1945), The Expected Value and Variance of the Reciprocal and Other Negative Powers of a Positive Bernoullinan Variate, "Annals of Mathematical Statistics", 16, pp.50-61

THOMSEN I., TESFU D. E BINDER D.A. (1986), Estimation of Design Effects and Intraclass Correlations when using Outdated Measuers of Size, "International Statistical Review", 54, 3, pp.343-349

TRIMARCHI F. (1990), L'imputazione dei dati mancanti nelle indagini campionarie: un'applicazione delle tecniche di regressione, Banca d'Italia, "Temi di discussione", n.143, novembre.

ELENCO DEI PIÙ RECENTI TEMI DI DISCUSSIONE (*)

n. 125 — *On the estimation of stochastic differential equations: the continuous-time maximum-likelihood approach,* di R. CESARI (settembre 1989).

n. 126 — *La misurazione dell'efficienza nei modelli di "frontiera"*, di M. GRESTI (settembre 1989).

n. 127 — *Do intergenerational transfers offset capital market imperfections? Evidence from a cross-section of Italian households,* di L. GUISO - T. JAPPELLI (settembre 1989).

n. 128 — *La struttura dei rendimenti per scadenza secondo il modello di Cox, Ingersoll e Ross: una verifica empirica,* di E. BARONE - D. CUOCO - E. ZAUTZIK (ottobre 1989).

n. 129 — *Il controllo delle variabili monetarie e creditizie: un'analisi con il modello monetario della Banca d'Italia,* di I. ANGELONI - A. CIVIDINI (novembre 1989).

n. 130 — *L'attività in titoli delle aziende di credito: un'analisi di portafoglio,* di G. FERRI - C. MONTICELLI (dicembre 1989).

n. 131 — *Are asymmetric exchange controls effective?* di F. PAPADIA - S. ROSSI (gennaio 1990).

n. 132 — *Misurazione dell'offerta di lavoro e tasso di disoccupazione,* di P. SESTITO (marzo 1990).

n. 133 — *Progressing towards European Monetary Unification: Selected Issues and Proposals,* di L. BINI SMAGHI (aprile 1990).

n. 134 — *Il valore informativo delle variabili finanziarie: un'analisi con il modello econometrico trimestrale della Banca d'Italia,* di I. ANGELONI e A. CIVIDINI (aprile 1990).

n. 135 — *A Model for Contingent Claims Pricing on EMS Exchange Rates,* di A. ROMA (maggio 1990).

n. 136 — *Le attività finanziarie delle famiglie italiane,* di L. CANNARI - G. D'ALESSIO - G. RAIMONDI - A. I. RINALDI (luglio 1990).

n. 137 — *Sistema pensionistico e distribuzione dei redditi,* di L. CANNARI - D. FRANCO (luglio 1990).

n. 138 — *Time Consistency and Subgame Perfection: the Difference between Promises and Threats,* di L. GUISO - D. TERLIZZESE (luglio 1990).

n. 139 — *Test di integrazione e analisi di cointegrazione: una rassegna della letteratura e un'applicazione,* di G. BODO - G. PARIGI - G. URGA (luglio 1990).

n. 140 — *The Experience with Economic Policy Coordination: the Tripolar and the European Dimensions,* di G. GOMEL - F. SACCOMANNI - S. VONA (luglio 1990).

n. 141 — *The Short-Term Behavior of Interest Rates: Did the Founding of the Fed Really Matter?,* di P. ANGELINI (ottobre 1990).

n. 142 — *Evoluzione e performance dei fondi comuni mobiliari italiani,* di F. PANETTA - E. ZAUTZIK (ottobre 1990).

n. 143 — *L'imputazione dei dati mancanti nelle indagini campionarie: un'applicazione delle Tecniche di regressione,* di F. TRIMARCHI (dicembre 1990).

n. 144 — *On the Measurement of Intra-Industry Trade: Some Further Thoughts,* di S. VONA † (dicembre 1990).

n. 145 — *Exchange Rate Variability and Trade: Why is it so Difficult to Find Any Empirical Relationship?,* di L. BINI SMAGHI (dicembre 1990).

n. 146 — *La scelta del meccanismo di collocamento dei titoli di Stato: analisi teorica e valutazione dell'esperienza italiana,* di L. BUTTIGLIONE - A. PRATI (gennaio 1991).

n. 147 — *Diversification and Performance,* di M. BIANCO (gennaio 1991).

n. 148 — *Exchange Rate and Pricing Strategies in a Model of International Duopoly,* di P. CASELLI (gennaio 1991).

n. 149 — *Concorrenza e redditività nell'industria bancaria: un confronto internazionale,* di V. CONTI (febbraio 1991).

n. 150 — *Economie di scala e di diversificazione nel sistema bancario italiano,* di C. CONIGLIANI - R. DE BONIS - G. MOTTA - G. PARIGI (febbraio 1991).

n. 151 — *Politiche di offerta e riallocazione del credito bancario negli anni ottanta,* di C. GIANNINI - L. PAPI - A. PRATI (febbraio 1991).

(*) I «Temi» possono essere richiesti a:
Banca d'Italia - Servizio Studi - Divisione Biblioteca e Pubblicazioni - Via Nazionale, 91 - 00184 Roma.

BANCA D'ITALIA - CENTRO STAMPA