

Università degli Studi di Salerno
DIPARTIMENTO DI SCIENZE ECONOMICHE

Gianluigi Coppola

CELPE – Centro di Economia del Lavoro e di Politica Economica

**DISOCCUPAZIONE E MERCATO DEL LAVORO:
UN'ANALISI SU DATI PROVINCIALI***

Ricerca condotta nell'ambito del progetto strategico CNR su
"Disoccupazione e bassi livelli di attività in Italia" (contratto n. 95.04485-ST7.4).

WORKING PAPER 3.77
febbraio 1999

Indice

1. Premessa.....	5
2. Il mercato del lavoro.....	6
3. La struttura provinciale dell'economia italiana: i risultati di un'Analisi Multivariata.....	24
4. Conclusioni.....	31
5. Appendice.....	32
6. Bibliografia.....	34

1. Premessa

Gli squilibri territoriali propri dell'economia italiana vengono quasi sempre sintetizzati nella rappresentazione dualistica Nord-Sud. Tale ottica, se da un lato ha un indubbio vantaggio di sintesi delle diversità presenti nel paese e ne costituisce una rappresentazione più che attuale, dall'altro può non cogliere appieno le differenze esistenti tra le diverse realtà economiche e le dinamiche che ciascuna di esse ha presentato nel corso degli anni. Ciò vale in modo particolare per valutare le diverse *performance* territoriali di quelle variabili economiche che appartengono al mercato del lavoro. Si rende pertanto necessario, al fine di ottenere una più completa raffigurazione della realtà italiana e di come essa sia articolata, un'analisi che vada oltre il dualismo Nord-Sud e che individui aree territoriali omogenee.

In questo lavoro ci si prefigge tale obiettivo e per il suo raggiungimento si descrivono e si individuano le omogeneità e le divergenze delle realtà economiche italiane sulla base di indicatori del mercato del lavoro, del settore della produzione e delle famiglie.

Lo studio è diviso in due parti. Nella prima parte si descrive la struttura del mercato del lavoro e la sua evoluzione negli ultimi quindici anni. Nella seconda parte si procede all'individuazione di elementi di convergenza e di divergenza delle realtà economiche analizzando unitamente agli indici del mercato del lavoro già visti nella prima parte, indici del settore della produzione e delle famiglie. In questa fase vengono applicati metodi di statistica multivariata quali l'analisi delle componenti principali e l'analisi di *cluster*.

Occorre precisare che l'analisi non ha un preciso modello teorico di riferimento. Tuttavia essa è propedeutica alla costruzione ed alla stima di un modello diretto a spiegare i differenziali territoriali [Decressin J. Fatàs A, 1995; Blanchard O. Katz L, 1992] poichè il fine ultimo dello studio, si ripete, è quello di gruppi di province omogenee sulle quali effettuare il confronto delle variabili economiche.

I dati usati per l'analisi sono principalmente dati disaggregati a livello provinciale e si riferiscono al periodo storico 1981-1995. Essi appartengono a due fonti statistiche diverse. Per quanto riguarda il mercato del lavoro sono state utilizzate le serie storiche relative alle rilevazioni delle forze di lavoro dell'ISTAT mentre come indicatori del settore delle famiglie della produzione ci si è avvalsi dei dati provinciali dell'Istituto Guglielmo Tagliacarne sul reddito e consumo disponibile e sul reddito prodotto. Si precisa che per gli anni in cui i dati Tagliacarne non sono disponibili sono state effettuate delle loro stime.

2. Il Mercato del lavoro¹

Le peculiarità del mercato del lavoro in Italia sono essenzialmente due: l'elevato tasso di disoccupazione ed il divario territoriale esistente a livello di *performance* ottenute nelle diverse aree geografiche del Paese. La disoccupazione non è affatto distribuita in modo omogeneo tanto che può definirsi a tutti gli effetti l'elemento preminente dell'economia meridionale. Secondo dati ISTAT relativi all'anno 1995 in Italia vi sono circa tre milioni di disoccupati pari a circa al 12% delle forze di lavoro. Tale valore non è di molto superiore alla media degli ultimi quindici anni (11%). Tuttavia ciò che ha caratterizzato gli anni '80 e la prima metà degli anni '90 è stato l'accentuarsi del divario esistente tra l'Italia Settentrionale ed il

1 L'analisi che segue si basa essenzialmente sulle statistiche delle rilevazioni sulle forze di lavoro. Si avverte che nel 1992 sono cambiati i metodi di rilevazione e, pertanto, i dati relativi agli anni che vanno dal 1993 in poi non sono immediatamente confrontabili con gli anni precedenti. In questo lavoro non si sono apportate rettifiche alle serie storiche, al contrario di quanto è stato fatto in altri lavori (Amendola, Caroleo, Coppola, 1997). Per questo motivo si potranno riscontrare minime differenze nei valori con gli studi citati.

Mezzogiorno. Infatti se nel Nord Italia il tasso di disoccupazione è stato nel quindicennio in media pari al 6,8%, nell'Italia Meridionale lo stesso indice ha registrato una media del 17,5% sino a raggiungere il 21,1% nel 1988. L'acuirsi del divario può essere visto in modo immediato confrontando le seguenti cartine dell'Italia che si riferiscono rispettivamente agli anni 1960 1981 e 1995 (grafici nn. 2.1, 2.2, 2.3).

Fig. 2.1 – Tasso di disoccupazione. Anno 1960

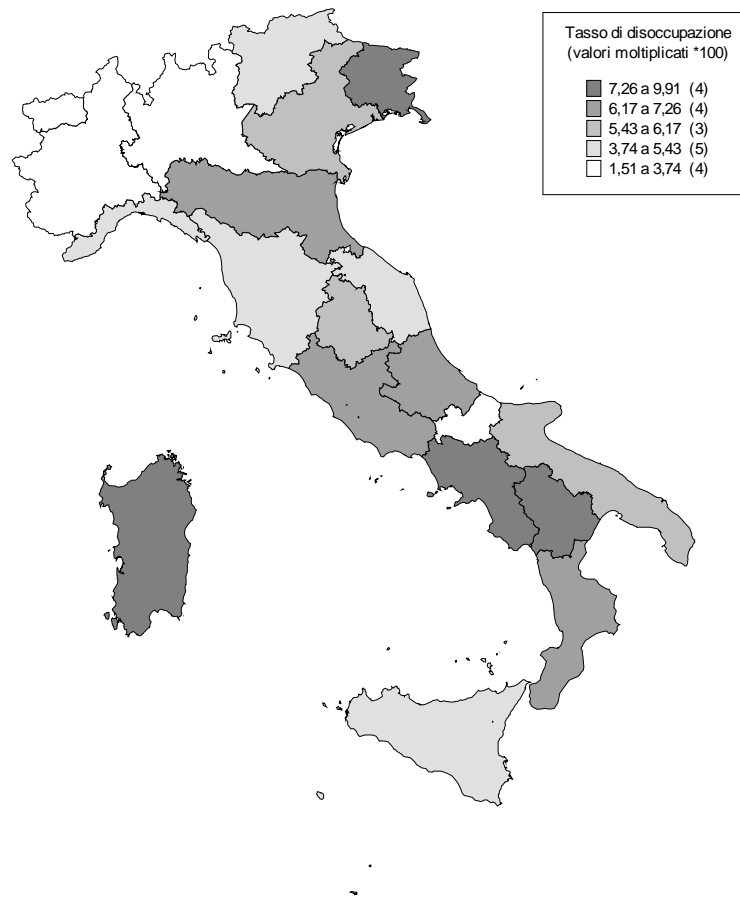


Fig. 2.2 – Tasso di disoccupazione. Anno 1981

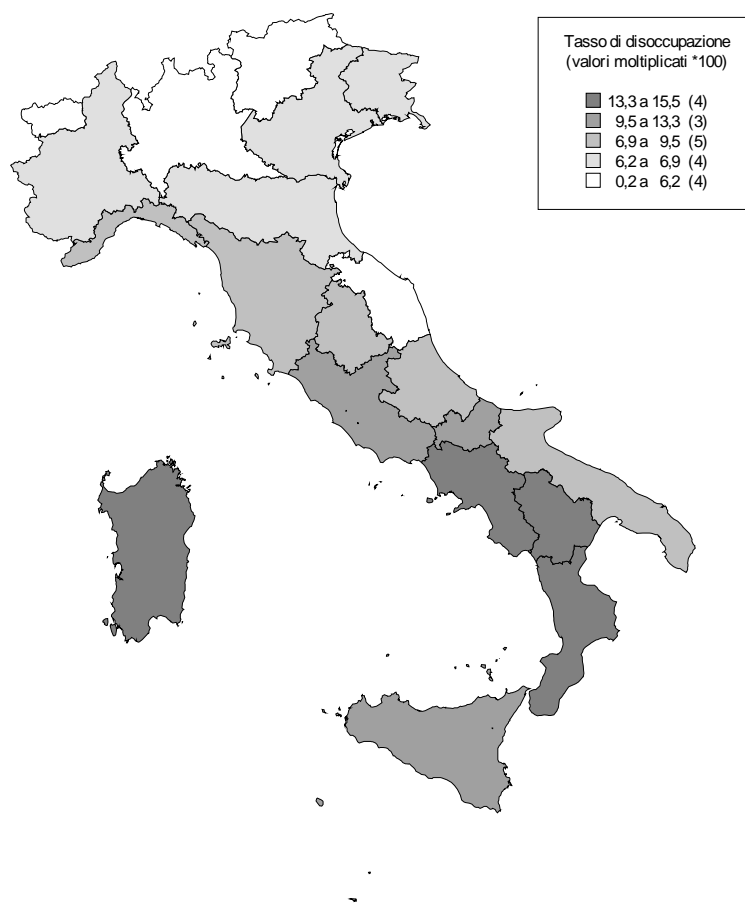
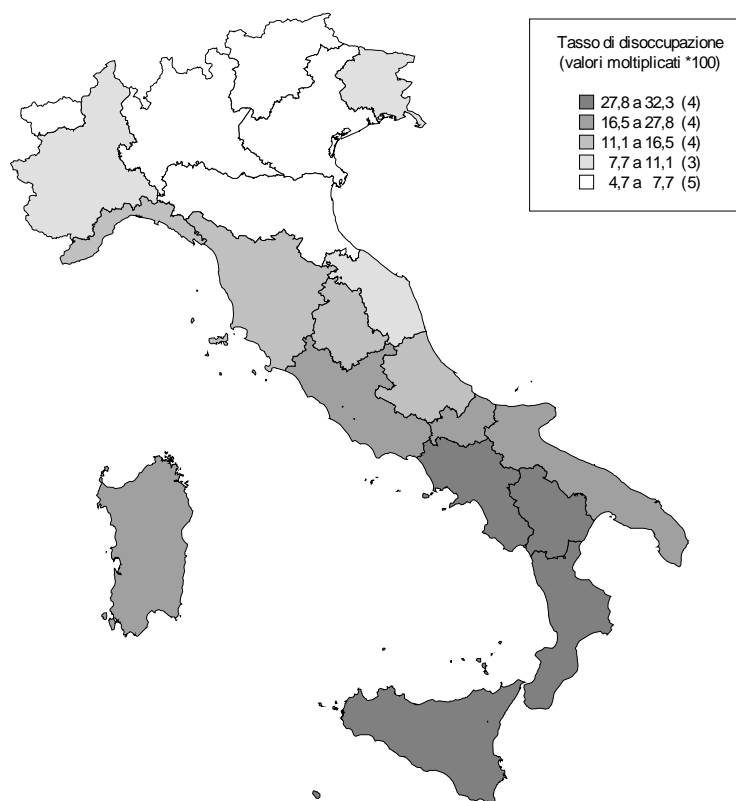


Fig. 2.3 – Tasso di disoccupazione. Anno 1995



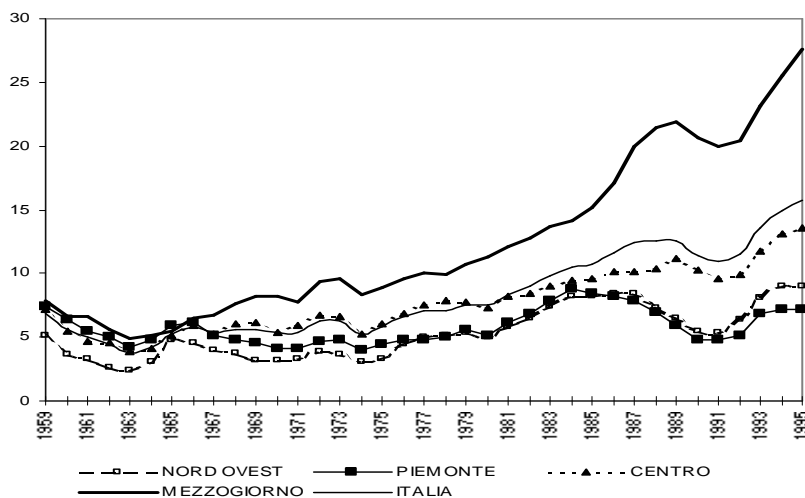
Nel 1960 il tasso di disoccupazione in Italia era del 5,5% e la distanza tra il Meridione (6,65%) ed il Nord Italia (5,3%) era meno di un punto e mezzo percentuale. Inoltre si nota che il Nord-Est aveva un tasso di disoccupazione (6,4%) superiore alla media nazionale mentre in alcune regioni meridionali, come la Sicilia (5,4%) ed il Molise (2%), si avevano tassi di disoccupazione inferiori alla stessa media. Nel 1960 la disoccupazione era presente sia in alcune regioni del Nord-Est (Emilia Romagna e Friuli) ed era minima in alcune regioni del Sud (Molise e Sicilia).

Nel 1981, ed in modo molto più evidente nel 1995, la disoccu-

pazione è diventata l'elemento caratterizzante dell'economia meridionale. In particolare nel 1995 le regioni del Nord-Est hanno registrato un tasso di disoccupazione (5,9%) che è la metà di quello nazionale ed inferiore di un punto e mezzo percentuale a quello dell'Italia nord-occidentale (7,4%) mentre nelle regioni meridionali la percentuale delle forze lavoro che sono disoccupate (21%) è quasi il doppio della media nazionale ed addirittura il triplo dell'Italia Settentrionale (7%) fino ad arrivare al 25% in Campania.

Nel grafico 2.4 è rappresentata la dinamica temporale della disoccupazione per ognuna delle quattro macroaree. E' ben evidente la notevole differenza che esiste tra la serie storica del tasso di disoccupazione meridionale e quelle dell'Italia nord-occidentale e dell'Italia nord-orientale. La prima evidenzia un forte trend crescente mentre le altre sembrano avere una più accentuata natura ciclica.

Fig. 2.4 – Tasso di disoccupazione. Macroregioni d'Italia (1960-1995)



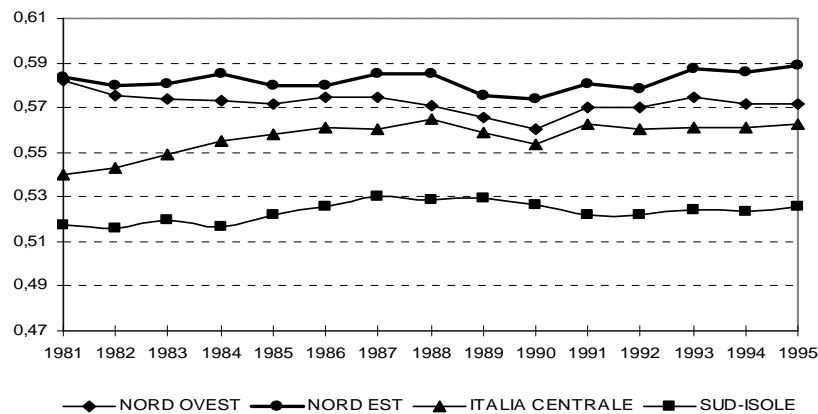
Analizzando i dati a livello di singola provincia il quadro complessivo non cambia. Nel grafico che segue sono riportati i valori minimi medi e massimi registrati per provincia nel periodo 1981-1995.

Come si può vedere i valori medi ed i rispettivi campi di variazione aumentano quando ci si sposta dalle province settentrionali a quelle meridionali. In particolare valori medi elevati si sono registrati in molte province siciliane. Caltanissetta ha avuto la media più alta del quindicennio (26%) con un valore massimo del 34% seguita da Enna (24%) e da Messina (20%). Anche in Calabria ed in Basilicata le medie provinciali sono state molto alte. In Campania Napoli, Caserta e Salerno il tasso di disoccupazione è stato molto consistente con una media rispettivamente del 21% 23% e 28%, anche se per le altre due province della regione, Avellino e Benevento, si sono rilevati dei tassi di disoccupazione inferiori alla media meridionale. Anche in Puglia il quadro provinciale è più variegato rispetto alle regioni Sicilia e Calabria. Infatti Bari ha avuto nel periodo di riferimento un tasso di disoccupazione del 11% Brindisi del 13% mentre Taranto ha avuto il 17% delle forze di lavoro disoccupate. Il Nord, pur registrando tassi di disoccupazione di molto inferiori a quelli meridionali, offre un quadro non del tutto omogeneo del fenomeno. Infatti, se molte province presentano tassi di disoccupazione che non si discostano molto dalla media settentrionale (7%), altre come Torino Rovigo, La Spezia, Genova e Ferrara presentano dei valori abbastanza elevati.

Per un'analisi più completa del fenomeno si è ritenuto utile analizzare in modo più approfondito la dinamica degli indici dell'offerta e della domanda di lavoro poichè la disoccupazione può essere considerata come la risultante del divario esistente tra queste due grandezze.

Come proxy dell'offerta di lavoro si è adottato il tasso di attività calcolato come il rapporto tra la consistenza delle forze di lavoro e la popolazione in età da lavoro. La media italiana del tasso di attività nel quindicennio 1980 -1995 è stata pressoché stabile. Disaggregando tale indice per aree geografiche non si sono registrate notevoli differenze. Il tasso di attività dell'area Nord-Est è stato costante, è diminuito leggermente quello del Nord Ovest, mentre il tasso di attività dell'Italia meridionale è stato sempre il più basso d'Italia.

Fig. 2.6 – Tasso di attività. Macroregioni d'Italia 1981-1995



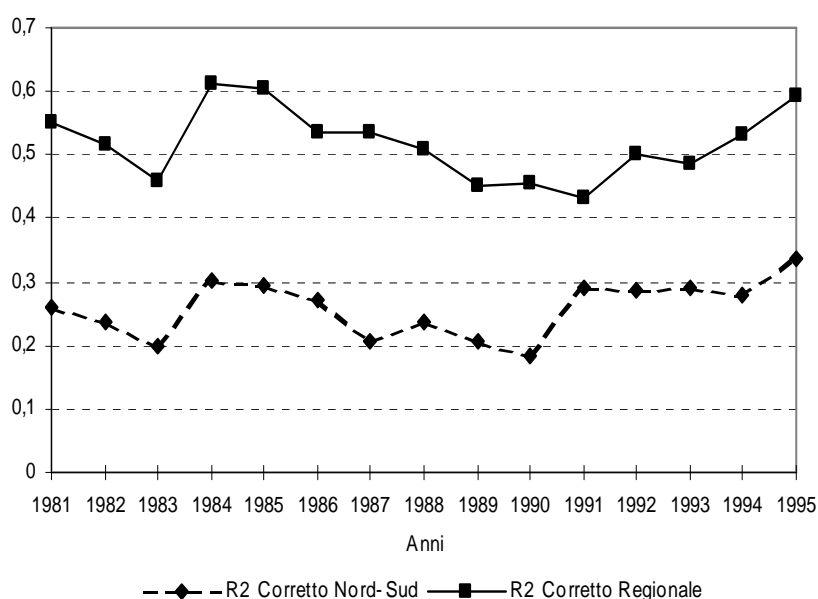
I differenziali dei tassi di attività tra le regioni italiane e tra il Nord ed il Sud del paese sono rimasti invariati nel corso degli anni '80 ed hanno registrato una crescita solo nell'ultimo quinquennio. Essi sono stati calcolati per ogni provincia italiana e per ogni anno scomponendo, la devianza totale della serie in due grandezze: la devianza tra le classi, denominata devianza *between*, e la devianza interna alle classi, detta anche devianza *within*, dove per classi si intendono gruppi predefiniti di province².

2 La scomposizione della devianza totale è stata calcolata considerando come classi sia le due macroregioni Centro-Nord e Meridione sia le singole regioni amministrative. Il peso percentuale della devianza *between* sulla devianza totale misura quanto del differenziale del tasso di attività tra le provincie è imputabile all'appartenenza delle stesse alle singole classi. Lo stesso risultato si ottiene stimando l' R^2 di una regressione lineare nella quale la variabile dipendente è la serie dei tassi di attività provinciali ed i regressori sono delle dicotomiche che assumono valore pari a 0 o ad 1 a seconda se la provincia appartiene o meno ad una determinata classe. In particolare nella classificazione Nord-Sud delle provincie italiane si avrà nella regressione una sola variabile dicotomica che assumerà valore se la provincia è situata al Nord e valore 1 se è situata al Sud mentre se le provincie vengono raggruppate in base alla loro regione amministrativa di appartenenza si avranno, tante dicotomiche, una per regione, che avranno valore 0 o 1 a seconda se la provincia appartiene o meno a quella regione.

I due indici, il primo relativo alla ripartizione Nord-Sud e l'altro relativo alla classificazione regionale, non sono tra loro immediatamente confrontabili, perché le due regressioni hanno un diverso numero di regressori e quindi di gradi di libertà. Pertanto si sono confrontati non gli R^2 delle regressioni

Si rileva che il differenziale dei tassi di attività tra le province italiane è imputabile solo per il 30% circa all'appartenenza delle stesse alle due macroaree del paese (Nord e Sud) e per il 55% alla propria regione³.

**Fig. 2.7 – Serie R2 corretto Nord-Sud ed R2 corretto regionale.
Tasso di attività – (1981-1995)**



Diversi sono i risultati ottenuti calcolando il tasso di attività maschile e quello femminile. Il primo è diminuito in tutte le aree del paese mentre il secondo ha registrato una continua crescita. In

bensi gli R^2 corretti per i rispettivi gradi di libertà (\bar{R}^2). Il calcolo è stato effettuato per ogni anno della serie dei tassi di attività 1981-1995 ottenendo in questo modo due serie storiche di \bar{R}^2 ognuna relativa alla classificazione adottata che vengono sinteticamente indicati come \bar{R}^2 Nord-Sud ed \bar{R}^2 regionale.

3 Medie del periodo 1981-1995.

particolare i tassi di partecipazione maschile delle singole macroaree presentano una forte omogeneità soprattutto se confrontati con altri indici che verranno visti da qui a poco. Inoltre il divario regionale del tasso di attività femminile è risultato in media più elevato rispetto a quello maschile.

Fig. 2.8 – Tasso di attività – Maschi (1981-1995)

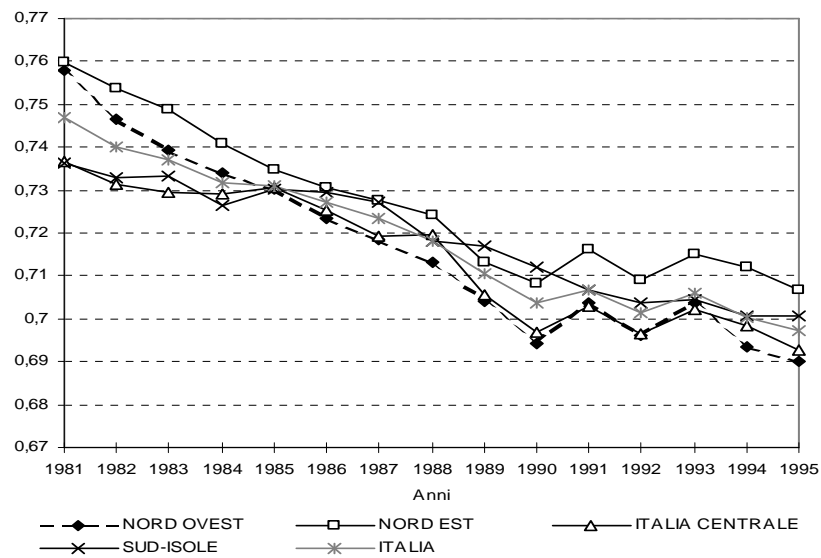
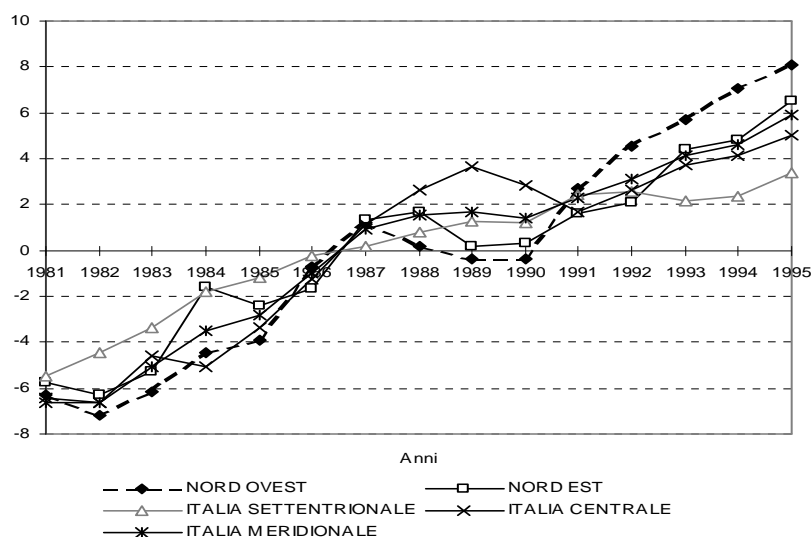


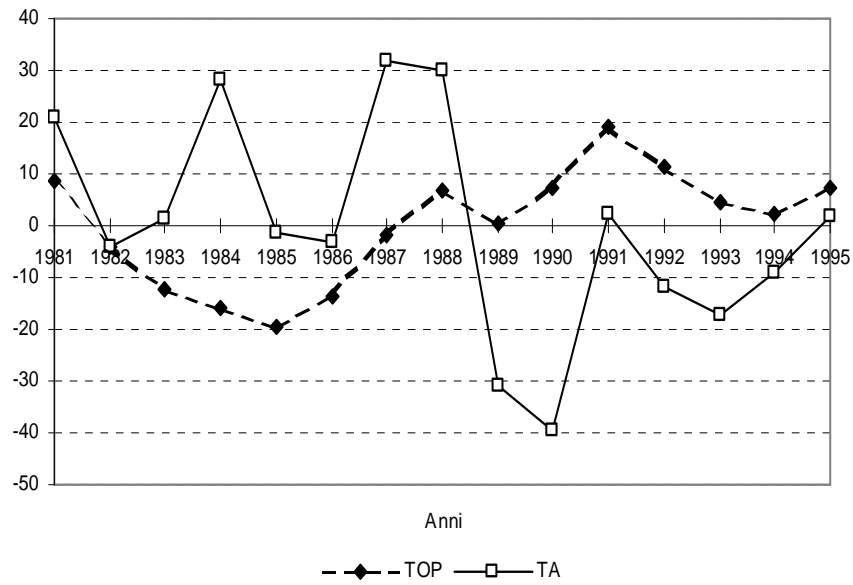
Fig. 2.9 – Tasso di attività – Femmine (1981-1995)



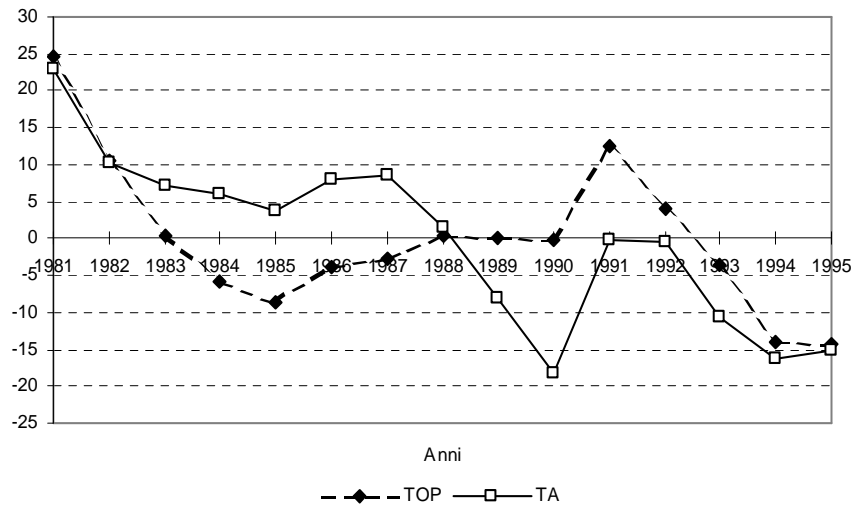
Più articolata è stata la dinamica del tasso di occupazione (TOP) assunto in questo contesto come proxy della domanda di lavoro. L'indice dato dal rapporto tra le persone occupate e la popolazione in età da lavoro ha subito un forte calo nel Meridione mentre nelle altre aree ha registrato una crescita negativa meno accentuata ed una ciclicità più marcata. Anche per questo indice è importante la distinzione per sesso. Infatti il tasso di occupazione femminile, a differenza di quello maschile è cresciuto in tutte le aree del paese ad esclusione del Meridione dove si è mantenuto essenzialmente stabile. Tutto questo si è risolto in un sostanziale ampliarsi dei divari territoriali delle opportunità di occupazione.

Tali divari risultano essere più marcati ed evidenti se si considerano unitamente le serie dei tassi di attività e dei tassi di occupazione. Nei grafici seguenti sono riportate le serie standardizzate rispetto ai relativi valori medi, dei due tassi per ciascuna delle quattro macroaree in cui è stata divisa l'Italia.

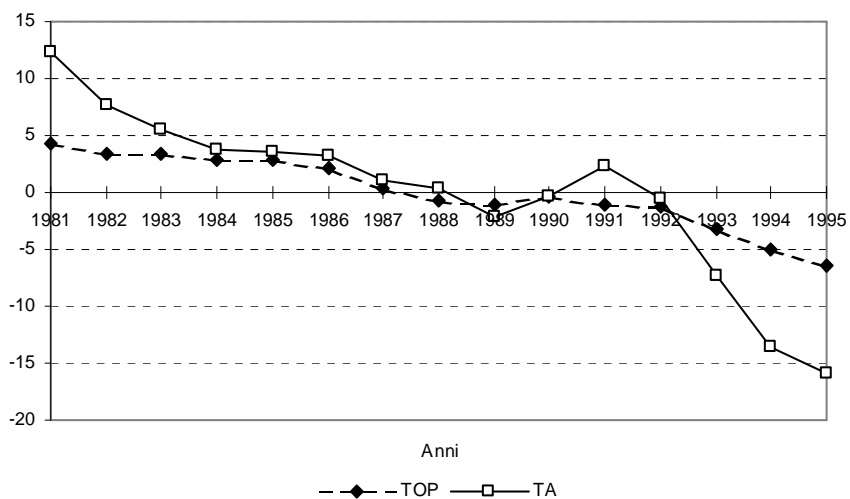
**Fig. 2.10 – Tasso di attività e tasso di occupazione.
Serie standardizzate – Nord-Ovest (1981-1995)**



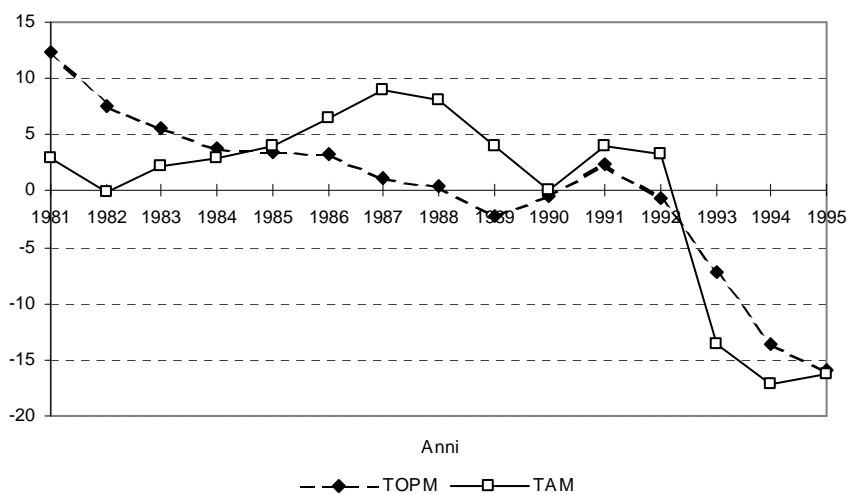
**Fig. 2.11 – Tasso di attività e tasso di occupazione.
Serie standardizzate – Nord-Est (1981-1995)**



**Fig. 2.12 – Tasso di attività e tasso di occupazione.
Serie standardizzate – Centro (1981-1995)**



**Fig. 2.13 – Tasso di attività e tasso di occupazione.
Serie standardizzate – Sud-Isole (1981-1995)**

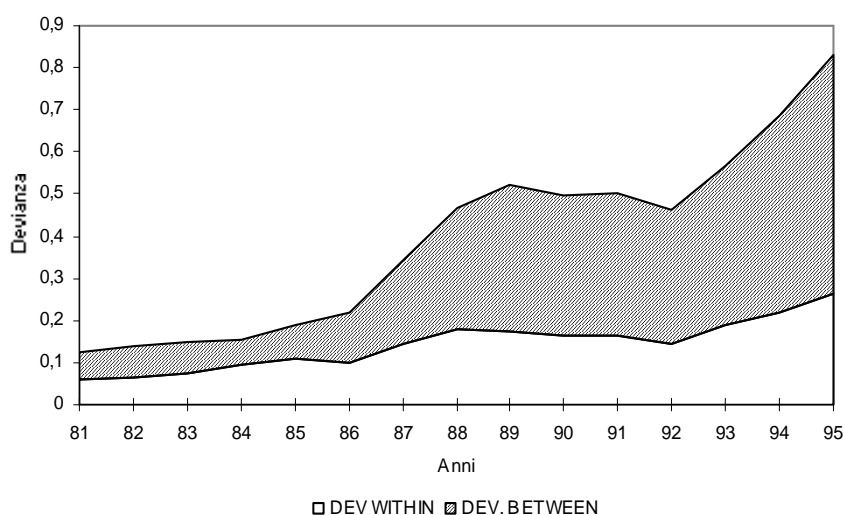


Come si può vedere ogni area presenta dinamiche diverse dei due tassi. Nel Nord-Est il tasso di occupazione è diminuito sino al 1985 e poi è cresciuto negli anni successivi mentre il tasso di attività ha avuto un andamento ciclico. In complesso vi è stato un aumento delle opportunità di lavoro a fronte di un'offerta stabile. Nel Meridione, invece, vi è stata una costante caduta del TOP ed una crescita (almeno fino al 1992) del tasso di attività che ha comportato una crescita della disoccupazione.

La combinazione degli andamenti territoriali differenziati dei due indici, del tasso di attività e del tasso di occupazione, ha comportato una crescita del divario territoriale del tasso di disoccupazione. Tale differenziale esistente tra le province italiane, e la sua dinamica nel tempo può essere studiato in modo alternativo e complementare usando due diversi approcci.

Il primo metodo, già adottato nell'analisi dei tassi di attività, consiste nello scomporre la varianza totale nelle due sue componenti *between* e *within*. Il seguente grafico riporta l'andamento della devianza così scomposta dove le classi sono le due macroregioni.

Fig. 2.14 - Scomposizione della devianza totale in devianza within e between. Tasso di disoccupazione – dati provinciali (1981-1995)



Come si può vedere nel corso degli anni ottanta è aumentata notevolmente la variabilità del tasso di disoccupazione tra le province italiane ed allo stesso tempo il divario tra Centro-Nord e Sud. Il differenziale esistente tra le province è in gran parte spiegato il dualismo Nord Sud e ciò conferma, anche in modo più forte, che la disoccupazione rimane una peculiarità dell'economia meridionale.

Con il secondo approccio si è calcolato il tasso di crescita espresso in logaritmi del rapporto percentuale della disoccupazione della provincia *iesima* sul totale nazionale e si è scomposto tale rapporto in tre fattori che sono il tasso di crescita del rapporto percentuale dei tassi di disoccupazione, dei tassi di attività e delle popolazione. In formula si ha⁴:

$$\Delta \log \left(\frac{U_{it}}{U_t} \right) = \Delta \log \left(\frac{TDIS_{it}}{TDIS_t} \right) + \Delta \log \left(\frac{TA_{it}}{TA_t} \right) + \Delta \log \left(\frac{POP_{it}}{POP_t} \right)$$

L'ultima identità quantifica quanto della crescita del divario in termini di disoccupazione di ogni singola provincia nei confronti della media nazionale, è dovuta alla crescita dei differenziali dei tassi di disoccupazione, dei tassi di attività, e della percentuale della popolazione della provincia. Il calcolo dei tassi di crescita è stato elaborato per tutte le province italiane e per i quindici anni oggetto di quest'analisi (1981-1995) e sono state successivamente raggruppate nelle quattro aree (Nord-Ovest, Nord-Est, Centro e Meridione) che vengono riportate nei seguenti grafici:

4 Si rimanda in Appendice per una dimostrazione analitica.

Fig. 2.15 – Scomposizione della variazione della disoccupazione relativa. Nord-Ovest (1981-1995)

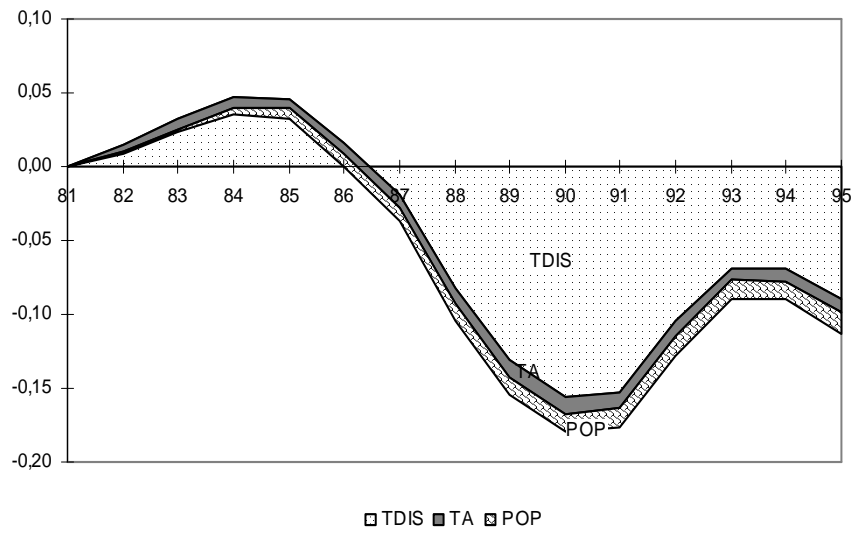


Fig. 2.16 – Scomposizione della variazione della disoccupazione relativa. Nord-Est (1981-1995)

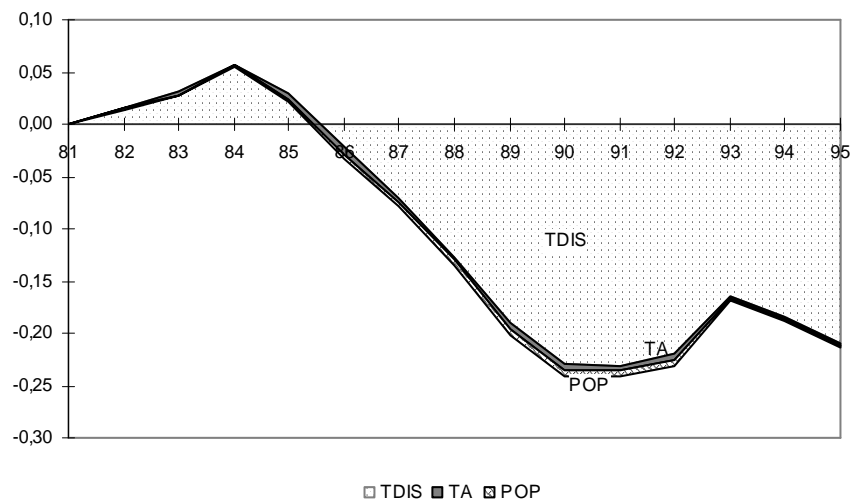


Fig. 2.17 – Scomposizione della variazione della disoccupazione relativa. Centro (1981-1995)

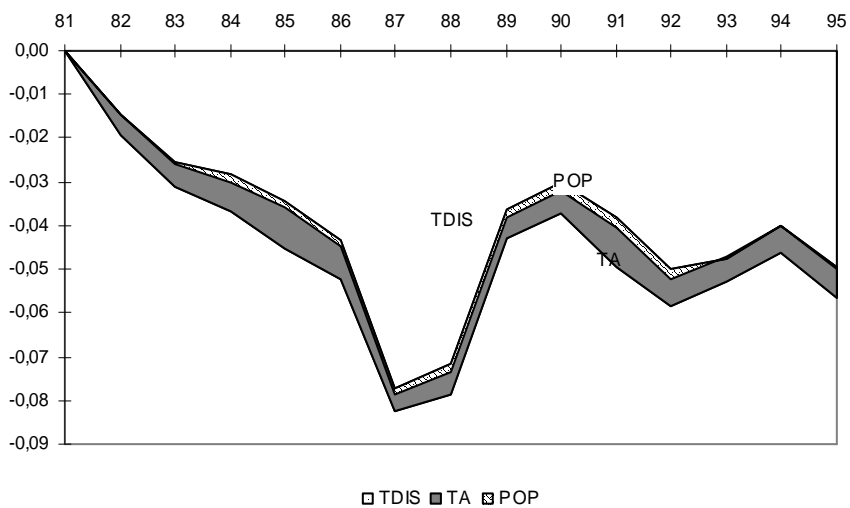
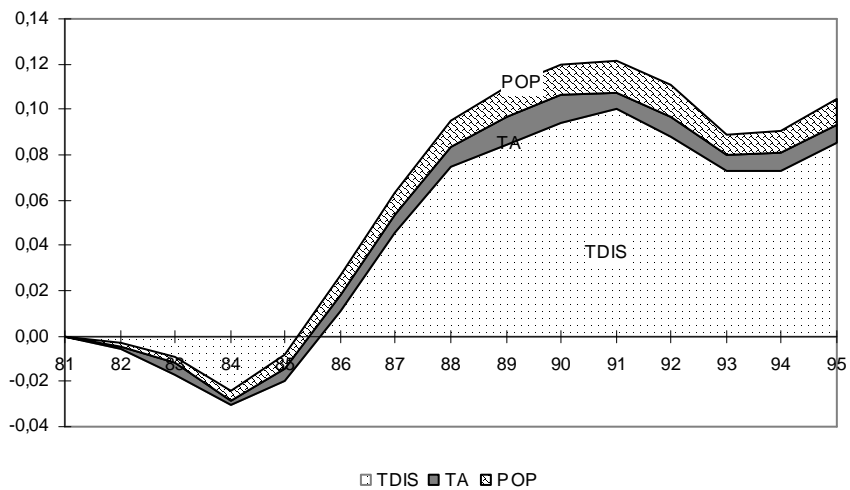


Fig. 2.18 – Scomposizione della variazione della disoccupazione relativa. Sud-Isole (1981-1995)



I grafici riassumono, in modo diverso, ciò che è stato già evidenziato in precedenza e come si può ben vedere le differenze tra le quattro aree sono molto nette. Il Mezzogiorno ha registrato, a differenza delle altre aree del paese una forte crescita del differenziale di disoccupazione dovuta ad in gran parte ad una caduta della domanda di lavoro ed in minima parte ad una crescita del differenziale dei tassi di attività e della popolazione residente mentre il Nord ha diminuito il proprio differenziale di disoccupazione grazie quasi esclusivamente alla crescita della domanda di lavoro⁵.

Dalla analisi svolta emerge che i divari tra le aree del paese sono aumentati nel corso del tempo. Vi è stata una crescita dei differenziali già esistenti delle *performance* del mercato del lavoro sia tra Nord e Sud in misura così consistente tale da aggravare il radicato dualismo presente nel paese, sia tra le singole regioni, a testimonianza dell'esistenza di una dinamica interna anche nelle due macroaree. Se è lecito mutuare un'espressione propria della geologia si può asserire che esiste una sorta di *effetto deriva* in virtù del quale le economie regionali del nostro Paese stanno assumendo delle proprie peculiarità che portano ad accrescere le distanze tra esse esistenti.

3. La struttura provinciale dell'economia italiana: i risultati di un'Analisi Multivariata

Sino adesso si è ragionato quasi esclusivamente in termini di confronti tra aree del paese anche se l'analisi è stata elaborata basandosi su dati provinciali. Ciò che si intende fare in questo paragrafo è completare lo studio già sviluppato con un'analisi compiuta ad un livello territoriale più disaggregato al fine di individuare le similitudini e le distanze che esistono tra le province italiane. Per raggiungere tale obiettivo si sono considerati, unitamente agli indici del mercato del lavoro, anche indici relativi al settore delle famiglie (reddito disponibile e consumi procapite) ed a quello della produzione (reddito prodotto) al fine di avere un insieme più

⁵ Una diminuzione del tasso disoccupazione può essere assimilata ad una proxy della crescita dell'occupazione e quindi ad un aumento della domanda di lavoro. [Decressin J. Fatàs A., 1995, p. 1633].

ampio di indicatori delle economie locali. Essendo il numero delle variabili aumentato ci si è avvalsi delle tecniche proprie dell'analisi multivariata, in particolare dell'analisi delle componenti principale e dell'analisi di *cluster*, per ottenere un quadro di sintesi della struttura dell'economia. Le due tecniche sono state utilizzate in modo complementare tra loro e si basano sui dati relativi agli anni 1984 e 1993⁶ per i quali è stato effettuato un confronto.

Gli indici impiegati nell'analisi multivariata sono riportati nella tabella 3.1.

6 Il confronto è stato effettuato sul primo e sull'ultimo anno per i quali i dati sono disponibili.

Tabella 3.1 - Indicatori utilizzati nell'analisi multivariata

Sigla	Indice	Descrizione
	Indicatori Mercato del Lavoro	
TA	Tasso di attività (1)	Forze di lavoro/Popolazione in età da lavoro
TAP	Tasso di attività (2)	Forze di lavoro/Popolazione
TOP	Tasso di occupazione	Occupati/Forze di lavoro
TDIS	Tasso di disoccupazione	Persone in cerca di lavoro/Forze di lavoro
TDIA	Tasso di disoccupazione implicita	(Persone in cerca di lavoro + disposte a lavorare)/(Forze di lavoro + Non Forze di lavoro in età da lavoro)
TNEL	Percentuale di persone non in età da lavoro	Popolazione non in età da lavoro/Popolazione
TAGR	Percentuale occupati in agricoltura	Occupati in agricoltura/Totale occupati
TIND	Percentuale occupati nell'industria	Occupati nell'industria/Totale occupati
TSER	Percentuale occupati nei servizi	Occupati nei servizi/Totale occupati
DIPTOT	Percentuali occupati dipendenti	Occupati dipendenti/Occupati
TAGRD	Percentuale dipendenti in agricoltura	Occupati dipendenti in agricoltura/Occupati dipendenti
TINDD	Percentuale dipendenti nell'industria	Occupati dipendenti in agricoltura/Occupati dipendenti
TSERD	Percentuale dipendenti nei servizi	Occupati dipendenti in agricoltura/Occupati dipendenti
AFT	Grado di femminilizzazione dell'offerta lavoro	Forza lavoro femminile/Forza lavoro
	Indicatori Settore delle famiglie	
RPC	Reddito disponibile pro-capite	Reddito disponibile/Popolazione
CPC	Consumo pro-capite	Consumo/Popolazione
	Indicatori Settore delle imprese	
VAPOP	Valore aggiunto prodotto per abitante	Valore aggiunto al netto dei servizi bancari/popolazione
VAG	Valore aggiunto per occupato in agricoltura	Valore aggiunto prodotto in agricoltura/occupati in agricoltura
VIND	Valore aggiunto per occupato nell'industria	Valore aggiunto dell'industria/occupati nell'industria
VSER	Valore aggiunto servizi per occupato nel settore dei servizi	Valore aggiunto dei servizi/occupati nel settore dei servizi
PVSER	Percentuale valore aggiunto servizi	Valore aggiunto servizi/Totale Valore aggiunto (al lordo dei servizi bancari)
PSNV	Percentuale valore aggiunto servizi non destinabili alla vendita	Valore aggiunto servizi non destinabili alla vendita/Totale Valore aggiunto (al lordo dei servizi bancari)

Obiettivo del metodo delle componenti principali⁷ è quello di sintetizzare le variabili inserite nel modello in un numero limitato di *componenti principali* conservandone il massimo della loro variabilità al fine di perdere il minimo possibile dell'informazione statistica da esse espressa. Con questa tecnica, si raggiungono due scopi: la riduzione del numero delle variabili e l'individuazione di variabili latenti che possono essere interpretati come indicatori di sintesi della struttura economica delle economie locali. Ogni *componente principale* è una variabile latente combinazione lineare di tutte le variabili. E' quindi possibile individuare la correlazione che ogni variabile ha con ciascuna componente estratta e, *va da sé*, che maggiore è tale correlazione o *peso* della variabile, più importante è il contributo dato dalla variabile stessa nella determinazione della singola componente. L'ordine con il quale vengono estratte le componenti è importante poiché la variabilità spiegata da ogni componente è direttamente ad esso correlata. Tutte le componenti principali sono tra loro ortogonali. Ogni osservazione, in questo caso le province italiane, ha un proprio punteggio dato dai relativi valori stimati delle componenti principali ed è quindi possibile classificare le osservazioni stesse in base ai punteggi ottenuti.

Come è stato appena detto l'analisi delle componenti principali è stata applicata per i dati relativi all'anno 1984 e 1993. Nella prima fase sono state estratte le componenti principali⁸. Nella tabella 3.2 sono riportate per le prime cinque componenti estratte il grado di correlazione esistente tra ogni variabile e le componenti stesse per l'anno 1984 e per l'anno 1993 e la percentuale di varianza spiegata rispettivamente per ciascuno dei due anni.

7 Per una trattazione formale si rimanda a Rizzi, 1989.

8 Si è effettuata la rotazione degli assi con il metodo del Quartimax Raw.

**Tabella 3.2 - Risultati dell'analisi delle componenti principali
metodo di rotazione:Quartimax Raw - anni 1984,1993**

	anno 1984					anno 1993				
	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
TA	0,77	0,26	-0,39	-0,36	-0,09	0,87	-0,18	-0,27	0,11	-0,25
TAP	0,88	0,19	-0,24	-0,31	-0,06	0,93	-0,13	-0,16	0,07	-0,20
TOP	0,85	0,22	-0,38	-0,14	-0,14	0,94	-0,19	-0,21	0,03	-0,11
TDIS	-0,81	-0,06	0,29	-0,32	0,16	-0,92	0,16	0,09	0,09	-0,11
TDIA	-0,63	0,03	-0,05	-0,61	0,08	-0,78	0,06	0,07	0,18	-0,29
TNEL	-0,82	0,08	-0,24	0,07	-0,04	-0,80	-0,04	-0,17	0,04	-0,01
TAGR	-0,58	0,09	-0,78	-0,03	-0,07	-0,62	-0,08	-0,69	0,07	0,02
TIND	0,65	0,64	0,38	0,05	-0,07	0,62	-0,71	0,22	0,02	0,04
TSER	-0,18	-0,88	0,37	-0,01	0,16	-0,21	0,91	0,29	-0,05	-0,06
TAGRD	-0,73	0,08	-0,28	0,04	0,00	-0,73	-0,07	-0,33	0,03	0,17
TINDD	0,66	0,67	0,21	0,06	-0,10	0,74	-0,61	0,20	0,04	0,03
TSERD	-0,28	-0,91	-0,05	-0,11	0,12	-0,50	0,81	-0,05	-0,07	-0,14
RPC	0,92	-0,10	0,23	0,07	0,12	0,95	0,06	0,14	-0,03	0,12
CPC	0,87	-0,17	0,25	0,09	0,14	0,88	0,19	0,17	-0,02	0,16
PVSER	0,10	-0,80	0,11	0,34	-0,23	-0,09	0,78	-0,04	0,30	0,39
PSNV	-0,78	-0,32	0,00	-0,25	-0,13	-0,83	0,23	-0,09	0,15	-0,17
VAG	0,35	-0,08	0,19	0,12	0,66	0,27	-0,04	-0,01	-0,89	0,03
VIND	0,27	-0,26	0,09	-0,05	0,79	0,52	0,33	0,47	-0,29	0,01
VSER	0,65	0,08	0,17	0,42	0,24	0,48	0,01	0,18	-0,06	0,75
VAPOP	0,93	0,01	0,12	0,08	0,29	0,94	0,01	0,18	-0,19	0,16
AFT	0,73	0,05	-0,18	-0,53	-0,12	0,82	0,05	-0,09	0,28	-0,20
DIPTOT	0,06	-0,06	0,86	0,04	0,06	0,37	-0,02	0,77	0,11	0,19
Expl.Var	9,87	3,50	2,50	1,42	1,44	11,39	3,29	1,88	1,19	1,16
Perc. Tot.	0,45	0,16	0,11	0,06	0,07	0,52	0,15	0,09	0,05	0,05
cumulata	0,45	0,61	0,72	0,79	0,85	0,52	0,67	0,75	0,81	0,86

Relativamente all'anno 1984 la prima componente spiega da sola il 45% dell'intera varianza e la seconda il 16% per un totale complessivo del 61% circa mentre per il 1993 gli stessi valori sono pari al 52% per la prima componente, 15% per la seconda per un totale di varianza spiegata pari a 67%. La prima componente principale è correlata con alcuni indici del mercato del lavoro. In particolare è positivamente correlata con il tasso di attività (TA +0,77), con il tasso di occupazione (TOP, +0,85), con l'indice di femminilizzazione della forza lavoro (AFT, +0,73) e con la percentuale degli occupati nell'industria (TIND, +0,65) mentre è negativamente correlata con il tasso di disoccupazione (TDIS, -0,81) e con la percentuale di persone non in età da lavoro (TNEL, -0,82). Inoltre è

positivamente correlata anche con il valore aggiunto prodotto procapite (VAPOP, +0,93) con il reddito procapite (RPC, +0,92) e con il consumo procapite (CPC, +0,87), ed è negativamente correlata con quota di valore aggiunto dei servizi non destinabili alla vendita (PSNV, -0,78). La prima componente principale può essere in sintesi interpretata come il grado di *performance* del mercato del lavoro e, più in generale, come il livello dello sviluppo del reddito e della produzione. La seconda componente, essendo positivamente correlata con la percentuale degli occupati nell'industria (TIND, -0,78) e negativamente correlata con la percentuale degli occupati nei servizi (TSER, -0,88) ed anche con la quota di valore aggiunto prodotto dal settore terziario (PVSER, -0,80), è una variabile latente che esprime il grado di terziarizzazione della struttura economica delle province.

Nel 1993 i segni dei coefficienti di correlazione tra le variabili e le componenti principali rimangono invariati e pertanto le prime due componenti principali estratte conservano lo stesso significato interpretativo. Sono aumentati gli indici di correlazione della prima componente con le variabili. In particolare ha assunto maggior peso il tasso di disoccupazione (TDIS, -0,92), il tasso di disoccupazione implicito (TDIA) che è aumentato da -0,63 del 1984 a -0,78 e la percentuale della forze lavoro femminile (AFT, +0.82).

Nelle Figure nn. 3.3 e 3.4 sono rappresentati graficamente i risultati ottenuti. L'asse orizzontale rappresenta la prima componente principale e quello verticale la seconda componente principale. La posizione di ciascuna variabile all'interno del grafico è funzione del grado di correlazione esistente tra le variabili stesse e le singole componenti principali. Nell'area sinistra del grafico sono situati gli indici che hanno una forte correlazione negativa con la prima componente principale, vale a dire il tasso di disoccupazione (TDIS), il tasso di disoccupazione implicito (TDIA), la percentuale di persone non in età da lavoro (TNEL) e la quota di servizi non vendibili (PSNV), mentre a destra si trovano le variabili con essa fortemente correlate, come il tasso di attività (TA), il tasso di occupazione (TOP), il consumo procapite (CPC) ed il reddito procapite (RPC). Nella zona bassa del grafico sono disposte le variabili relative alla quota di occupati nei servizi (TSER), la percentuale degli occupati dipendenti nei servizi (TIND), e la percentuale del valore aggiunto dei servizi non destinabili alla vendita (PVSER). Andando

da sinistra verso destra, quindi, migliorano le *performances* degli indici del mercato del lavoro, del reddito e del consumo pro-capite e scorrendo il grafico dall'alto verso il basso aumenta il grado di terziarizzazione delle economie delle province.

I grafici nn 3.5 e 3.6 sono simili a quelli precedenti soltanto che al posto delle variabili vi sono le osservazioni, ossia le province italiane. La posizione occupata dalla singola provincia è funzione dei punteggi della prima e della seconda componente principale relativi alla provincia stessa. Esprimendo la prima componente principale il livello di sviluppo delle economie provinciali e la seconda il grado di terziarizzazione della struttura economica, la dislocazione delle province all'interno del grafico risulta essere strettamente collegata alla struttura dell'economia delle singole province.

Nella zona sinistra del grafico sono situate le province le cui economie sono caratterizzate da alti tassi di disoccupazione, da un'alta quota di valore aggiunto prodotta imputabile alla Pubblica Amministrazione e da un peso maggiore del settore agricolo. In quest'area si trovano quasi la totalità delle province meridionali. Al contrario a destra si trovano le province con un'alta percentuale di occupati, con forte partecipazione femminile al lavoro, con elevato reddito, consumo e valore aggiunto prodotto procapite. Le province dell'Italia centro-settentrionale sono situate in questa area. Nella parte inferiore del grafico sono situate province con una più alta percentuale di occupati nel settore della pubblica amministrazione (Roma e Trieste) e in quella superiore del grafico vi sono le province del Nord che hanno in media un settore industriale più sviluppato.

Come si è detto prima, il 1993 presenta una struttura economica delle province analoga a quella del 1984 ma è evidente uno squilibrio maggiore tra le due aree del paese, essendo cresciuta la distanza tra Nord e Sud. Gli indicatori più significativi del mercato del lavoro, come il tasso di occupazione, di disoccupazione e di partecipazione femminile, hanno acquistato un peso, dato dai coefficienti di correlazione, maggiore nella determinazione delle componenti principali. La figura n. 3.6, che mostra la posizione delle variabili rispetto alle componenti principali relative all'anno 1993, evidenzia una maggiore distanza tra le variabili stesse. Questo fenomeno è certamente imputabile alla crescita nel Paese del divario economico. Anche dal grafico delle province emerge in modo

speculare lo stesso fenomeno. In particolare si ha che il gruppo delle province meridionali sito sempre nell'area sinistra del grafico è più compatto ed anche più distante dal gruppo del Nord. Anche il gruppo delle province settentrionali sembra aver una maggiore dispersione nel suo interno. Pertanto dall'analisi delle componenti principali si ha conferma dei risultati ottenuti nella prima parte di questo lavoro.

Per avere una visione più netta dei gruppi di province ci si è avvalsi dell'analisi di Cluster implementata sullo stesso set di dati. Se con il metodo delle componenti principali si è individuata la struttura interna alle variabili ed in via mediata gruppi omogenei di province, con quest'ultima tecnica multivariata si sono aggregate le province in via diretta sulla base delle similarità e delle distanze tra loro esistenti. Essa è stata sviluppata in modo complementare alla prima poichè si sono individuate in modo più marcato i gruppi di province dato che si è considerato l'intera variabilità espressa dall'insieme delle variabili. In particolare si è adottato un algoritmo gerarchico⁹ che permette di aggregare prima gruppi di province che presentano distanze minime e successivamente gruppi con elementi sempre meno omogenei che comprendono in essi i gruppi formati negli stadi precedenti. Il risultato finale non è un'unica partizione delle 95 province ma più partizioni che vanno da quelle più piccole fino a quelle più grandi. L'insieme delle partizioni viene rappresentato graficamente attraverso il *dendrogramma* o *diagramma ad albero*.

Nei grafici 3.7 e 3.8 sono riportati i *dendrogrammi* ottenuti dall'analisi di cluster per gli anni 1984 e 1993. Nel dendrogramma, sono indicati sull'asse delle ordinate i livelli di distanza a cui vengono raggruppati gli elementi, mentre sull'asse delle ascisse sono indicate le singole unità. Ogni ramo del dendrogramma corrisponde ad un gruppo di province mentre la linea orizzontale di congiunzione di due o più rami individua la distanza che separa i gruppi.

Anche dal confronto dei due grafici emerge che la divergenza tra Nord e Sud è maggiore poichè nel 1993 è aumentata la distan-

⁹ Esistono diversi metodi di classificazione gerarchica dei gruppi. In questa analisi si è adottato il metodo di Ward e la distanza considerata tra gli elementi è la distanza Euclidea.

za alla quale si uniscono i due gruppi di province. Anche in questa sede è evidenziato il sostanziale accorpamento delle province meridionali mentre è più evidente la dinamica sviluppatasi tra le province del Centro-Nord. Tutto questo conferma ciò che in parte è emerso nel corso dell'analisi descrittiva osservando gli andamenti dell' \bar{R}^2 . In quel contesto si è osservato il crescere del divario tra le due macroaree del Paese tra le singole regioni. E' possibile ora asserire che la dinamica regionale interessa maggiormente le regioni del Nord.

4. Conclusioni

L'analisi svolta, nei suoi diversi momenti, conferma l'esistenza e l'acuirsi del dualismo Nord-Sud e rileva una crescita differenziata delle regioni soprattutto di quelle settentrionali. Come è stato detto in premessa e come si è potuto notare nel corso della lettura del presente lavoro ci si è astenuti da indicare le cause degli squilibri territoriali poichè lo spirito con il quale si è svolta l'indagine è stato quello di definire una semplice radiografia delle economie regionali. In letteratura non mancano di certo studi che spiegano l'esistenza del dualismo italiano. Ciò che si è voluto qui mettere in evidenza è la presenza della differente dinamica esistente tra il Nord ed il Sud dell'Italia che porta ad allargare lo storico divario piuttosto che ridurne la consistenza. E' d'uopo ricordare che i dati impiegati sono disaggregati a livello provinciale e quindi presenze di microaree interne alle province che presentano una forte vivacità soprattutto nel Mezzogiorno possono non emergere dall'analisi dei dati. L'analisi svolta vuole essere propedeutica per la costruzione di un modello diretto a spiegare le dinamiche regionali presenti in Italia.

5. Appendice

La disoccupazione della provincia *i*-esima al tempo *t*, U_{it} può essere data dal prodotto dei tre fattori:

$$U_{it} = \frac{U_{it}}{FL_{it}} \frac{FL_{it}}{POP_{it}} POP_{it}$$

dove FL sono le forze di lavoro $\frac{U_{it}}{FL_{it}}$ è il tasso di disoccupazione (TDIS), $\frac{FL_{it}}{POP_{it}}$ è il tasso di attività (TA) e POP_{it} è la popolazione della provincia *i*-esima al tempo *t*.

Definita U_t la disoccupazione italiana al tempo *t*, la percentuale della disoccupazione della provincia *i*-esima è data dal rapporto $\frac{U_{it}}{U_t}$ ed è pari al prodotto dei rapporti dei tassi di disoccupazione dei tassi di attività e della popolazione della provincia *i*-esima rispetto all'Italia. In formula si ha:

$$\frac{U_{it}}{U_t} = \frac{U_{it}}{FL_{it}} * \frac{FL_{it}}{POP_{it}} * \frac{POP_{it}}{FL_t} * \frac{POP_t}{POP_t}$$

oppure

$$\frac{U_{it}}{U_t} = \frac{TDIS_{it}}{TDIS_t} \frac{TA_{it}}{TA_t} \frac{POP_{it}}{POP_t}$$

Calcolando i logaritmi di ambo i membri e successivamente la loro differenza rispetto a due periodi successivi si ha:

$$\Delta \log \left(\frac{U_{it}}{U_t} \right) = \Delta \log \left(\frac{TDIS_{it}}{TDIS_t} \right) + \Delta \log \left(\frac{TA_{it}}{TA_t} \right) + \Delta \log \left(\frac{POP_{it}}{POP_t} \right)$$

Il primo membro dell'ultima equazione è una proxy del tasso di crescita della percentuale della disoccupazione della provincia *i*-esima sul totale Italia ed è pari alla somma dei tassi di crescita dei rapporti dei rispettivi tassi di disoccupazione, dei tassi di attività e delle popolazioni.

6. Bibliografia

- AA-VV, 1994, *Statistica*, manuale, Statsoft.
- AMENDOLA A. (A CURA DI), 1995, *Disoccupazione: Analisi Macroeconomica e Mercato del Lavoro*, Edizioni Scientifiche Italiane.
- AMENDOLA A., CAROLEO F.E., COPPOLA G., *Differenziali territoriali nel mercato del lavoro e sviluppo in Italia*, CELPE Università degli Studi di Salerno, Discussion Paper n. 36.
- AMENDOLA A., SCATTAGLIA M., 1992, *Disoccupazione e tassi di attività nel Mezzogiorno*, CELPE Università degli Studi di Salerno, Discussion Paper n. 5.
- BLANCHARD O., KATZ L., 1992, *Regional evolution*, Brookings Papers on Economic Activity 1, 1-75.
- CAROLEO F.E., 1989, *Le cause economiche dei differenziali regionali del tasso di disoccupazione*, Working Paper n. 3.3, Università degli Studi di Salerno.
- DECRESSIN J., FATÀS A., 1995, *Regional labor market dynamics in Europe*, European Economic Review, 39, 1627-1655.
- FABBRIS LUIGI, 1983, *Analisi esplorativa dei dati multidimensionali*, Cleup editore.
- ISTAT, *Indagine sulle forze di lavoro anni 1981-1995*.
- ISTAT, 1996, *Rapporto sull'Italia*, Il Mulino.
- ISTITUTO TAGLIACARNE, *Il reddito prodotto in Italia: un'analisi su dati provinciali 1980-1991*.
- ISTITUTO TAGLIACARNE, *Il reddito disponibile, consumo e risparmio analisi su dati provinciali 1980-1991*.
- PASCALE FILOMENA, 1995, *Metodi di analisi multivariata. Un'applicazione ai fenomeni socio-economici*, Tesi di Laurea in Statistica, Università degli studi di Salerno.
- RIZZI ALFREDO, 1989, *Analisi dei dati*, La nuova Italia scientifica.