

Traccia delle lezioni per il corso di Economia delle risorse umane*

Andrea Ichino

(Corso di Laurea in Economia e Finanza, Università di Bologna)

30 aprile, 2008

Abstract

L'obiettivo di questo corso è analizzare il mercato del lavoro e i soggetti che in esso operano, rendendo gli studenti capaci di seguire in modo non ideologico il dibattito politico-sociale e la letteratura teorica e applicata che studia questo mercato.

Il lavoro ha un'importanza fondamentale nella vita degli esseri umani e influisce in modo determinante sulle loro decisioni economiche e sul loro benessere. Nel mercato del lavoro gli individui offrono energie, talento e tempo disponibile per produrre beni e servizi utili alla società, conseguendone sia uno status sociale sia un reddito che, nella maggior parte dei casi, costituisce la principale fonte di sostentamento per loro stessi e per le loro famiglie. La rilevanza delle decisioni che interessano l'attività lavorativa non è limitata al tempo dedicato al lavoro retribuito, ma coinvolge l'intero ciclo di vita di un individuo: dalle scelte di istruzione e formazione professionale, alle decisioni di pensionamento, all'utilizzo del tempo libero. Per lo svolgimento di un'attività lavorativa e per l'accesso a determinate professioni sempre più necessario disporre di un buon livello di alfabetizzazione e di competenze specifiche al tipo di lavoro svolto, da acquisire attraverso l'istruzione scolastica e la formazione professionale continua. Inoltre, l'allungamento della vita attesa degli individui e la disponibilità di un reddito pensionistico rendono l'abbandono della vita lavorativa più una scelta economica che una necessità biologica. Dal lato delle imprese, il lavoro e il "capitale umano" in esso incorporato rappresentano risorse strategiche per il successo o il fallimento di un progetto di investimento. Nell'impresa moderna non basta utilizzare una tecnologia adeguata: è anche necessario attirare gli individui più capaci, incentivarli ad impegnarsi nel lavoro e impedire che abbandonino l'impresa per passare alla concorrenza.

Non dovrebbe sorprendere pertanto che il funzionamento del mercato del lavoro risulti essenziale per comprendere le caratteristiche e la performance dei sistemi sociali e, come tale, rivesta un ruolo centrale nell'analisi economica. In fondo il reddito generato dai servizi di lavoro rappresenta, tuttora, la quota più rilevante del reddito nazionale, mentre la patologia economica probabilmente più discussa e studiata dagli scienziati sociali, la disoccupazione, trova origine e si sviluppa proprio nel mercato del lavoro.

Tuttavia, proprio perché il rapporto di impiego coinvolge in modo così profondo la persona, il mercato del lavoro si configura come un mercato molto particolare e alquanto diverso dagli altri mercati in cui vengono scambiati beni e servizi. È un mercato che spesso non opera secondo le leggi concorrenziali della domanda e dell'offerta e nel quale le imperfezioni informative, le peculiarità sociali e le istituzioni giocano un ruolo rilevante. Qui più che altrove disturbano le situazioni di disparità di potere contrattuale, di ineguaglianza e di bisogno anche se generate da libere scelte degli individui. E, forse per questo, qui più che altrove è sentito (ancorché da alcuni aspramente osteggiato) il bisogno di un intervento pubblico che corregga, dove possibile, le imperfezioni del mercato, anche a costo di eventuali inefficienze.

*Andrea Ichino, Dipartimento di Scienze Economiche, Università di Bologna, Piazza Scaravilli 2, Bologna, e-mail: andrea.ichino@unibo.it <http://www2.dse.unibo.it/ichino/>

Contents

1	Introduzione	3
2	Come si forma l'offerta di lavoro	7
2.1	Il ruolo della famiglia di origine	8
2.1.1	Mobilità intergenerazionale: i fatti	9
2.1.2	Una teoria per riflettere su questi fatti	27
2.2	Perchè andiamo a scuola?	36
2.2.1	Un modello della quantità di istruzione acquisita da un individuo	37
2.2.2	Un primo confronto tra modello e dati	43
2.2.3	Come introdurre eterogeneità nel modello base	46
2.2.4	Un ripasso di alcune nozioni e strumenti di econometria	57
2.2.5	Il problema della causalità	65
2.2.6	Regressione multivariata e Matching	74
2.2.7	Il ricorso a esperimenti "naturali"	80
2.3	Quanto lavoro viene offerto sul mercato e a quale salario	85
2.3.1	Un modello dell'offerta di lavoro individuale	86
2.3.2	Effetto dei trasferimenti sulla partecipazione al lavoro	88
2.3.3	Offerta di lavoro, tassazione del reddito e benessere sociale	89
3	La domanda e l'equilibrio in un mercato del lavoro concorrenziale	98
3.1	La domanda di forza lavoro omogenea	99
3.2	L'equilibrio in un mercato del lavoro concorrenziale	104
3.3	Disoccupazione e fluttuazioni cicliche	108
3.4	Differenziali salariali in un mercato concorrenziale	110
3.5	Equilibrio concorrenziale e mercati del lavoro reali	115
4	Istituzioni che interferiscono con il funzionamento del mercato concorrenziale	116
4.1	Il sindacato	117
4.1.1	Modelli di contrattazione sindacale	119
4.1.2	Sindacato e rappresentanza: la contrapposizione tra <i>insiders</i> e <i>outsiders</i>	129

4.1.3	Il sindacato come strumento per correggere imperfezioni del mercato	139
4.2	I regimi di protezione dell'impiego (<i>Employment Protection Legislation</i>)	147
4.2.1	Il modello base: neutralità della EPL	150
4.2.2	EPL in presenza di salari rigidi	153
4.2.3	EPL, efficienza produttiva ed efficienza distributiva	157
5	Dentro l'impresa	160
5.1	Il problema della selezione dei lavoratori	162
5.1.1	L'equilibrio con selezione avversa	163
5.1.2	Il modello di segnalazione	171
5.1.3	Il modello di <i>screening</i>	177
5.2	Il problema della incentivazione dei lavoratori	184
5.2.1	Il problema della contrattazione tra un "principale" ed un "agente"	186
5.2.2	Performance relativa e tornei	194
5.2.3	Incentivazione collettiva e "free riding"	204
5.2.4	Salari di efficienza e incentivazione basata sul rischio di licenziamento	214

1 Introduzione

Il lavoro ha un'importanza fondamentale nella vita degli esseri umani e influisce in modo determinante sulle loro decisioni economiche e sul loro benessere.

Il funzionamento del mercato del lavoro è essenziale per comprendere le caratteristiche e la performance dei sistemi sociali e, come tale, riveste un ruolo centrale nell'analisi economica.

- Il reddito generato dai servizi di lavoro rappresenta, tuttora, la quota più rilevante del reddito nazionale.
- La “patologia” economica probabilmente più discussa e studiata dagli scienziati sociali, la disoccupazione, trova origine e si sviluppa proprio nel mercato del lavoro.

Tuttavia, proprio perchè il rapporto di impiego coinvolge in modo così profondo la persona, il mercato del lavoro si configura come un mercato molto particolare e alquanto diverso dagli altri mercati in cui vengono scambiati beni e servizi.

Cosa ci proponiamo di studiare

Organizzeremo l'analisi secondo queste tappe:

1. L'offerta: come i lavoratori acquisiscono le dotazioni con cui si presentano sul mercato, cosa e quanto decidono di offrire.
2. La domanda: quali e quante risorse umane sono domandate dalle imprese.
3. Come si incontrano domanda e offerta a seconda che:
 - i mercati siano competitivi o meno;
 - la forza lavoro sia omogenea o eterogenea;
 - esista incertezza;
 - le informazioni siano incomplete e/o asimmetriche;
 - esistano istituzioni che interferiscono con il funzionamento del mercato
4. Come e con quali risultati i lavoratori interagiscono con l'imprenditore e gli altri fattori della produzione all'interno di una impresa.
5. Come e perchè i lavoratori e le imprese interrompono il rapporto di lavoro.

Tre tematiche aggiuntive di grande attualità

Sono tematiche che interessano trasversalmente molti degli argomenti trattati nella parte principale del corso.

- Flussi migratori.
- Discriminazione di genere.
- “Globalizzazione” .

Quali strumenti utilizza l'economista del lavoro

- Analisi empirica:
 - Econometria descrittiva;
 - Econometria inferenziale;
 - Econometria per l'identificazione di relazioni causali.
- Analisi delle istituzioni:
 - Conoscenza delle istituzioni giuridiche culturali e sociali;
 - Analisi delle determinanti e del disegno di queste istituzioni;
 - Analisi dell'impatto di queste istituzioni sui mercati.
- Analisi teorica:
 - Analisi normativa;
 - Analisi positiva;
 - Modelli: “la carta geografica” e “il lampione”.

2 Come si forma l'offerta di lavoro

Le caratteristiche con cui un potenziale lavoratore si presenta sul mercato dipendono da:

- La famiglia.
- L'ambiente.
- La scuola.

Date le caratteristiche acquisite prima di entrare nel mercato, il lavoratore decide:

- se cercare un lavoro;
- quanto lavoro offrire;
- dove e come offrire un lavoro;
- quale salario richiedere per ogni unità di lavoro offerto.

2.1 Il ruolo della famiglia di origine

La famiglia d'origine determina le caratteristiche di un potenziale lavoratore in molti modi:

- dotazione genetica;
- dotazione ambientale e culturale;
- dotazione economica;
- guida nelle scelte scolastiche;
- educazione alle regole della convivenza sociale.

Che evidenza empirica abbiamo sulla rilevanza della famiglia d'origine?

Possiamo distinguere empiricamente i canali attraverso cui la famiglia d'origine esercita i suoi effetti?

E' efficiente e/o equo che la famiglia d'origine abbia un ruolo?

2.1.1 Mobilità intergenerazionale: i fatti

Considerate la seguente equazione

$$Y_t = \alpha + \beta Y_{t-1} + \epsilon_t \quad (1)$$

- Y_t è il reddito dei figli e Y_{t-1} è il reddito dei padri;
- ϵ_t è un shock casuale ai redditi dei figli;

Il parametro β è uno degli indicatori più utilizzati per misurare:

- il grado di mobilità intergenerazionale e l'esistenza di pari opportunità: ossia se la situazione economica dei genitori determina quella dei figli;
- la tendenza della diseguaglianza tra i redditi delle dinastie a ridursi nel tempo;
- la tendenza dei redditi di una dinastia a convergere (regredire) verso la media.

Con il metodo dei Minimi Quadrati Ordinari possiamo stimare β nella popolazione sostituendo i momenti del campione nella definizione:

$$\beta = \frac{\text{Cov}(Y_t, Y_{t-1})}{\text{Var}(Y_{t-1})}$$

Interpretazione del parametro β nella equazione 1

Se $\beta = 1$ allora:

$$Y_t = \alpha + Y_{t-1} + \epsilon_t$$

- il reddito Y_t è una “passeggiata casuale” (*random walk with drift*);
- $E(Y_t) = \alpha + Y_{t-1}$ e quindi il reddito dei padri è il miglior predittore del reddito dei figli;
- assenza di mobilità intergenerazionale e di pari opportunità.

Se $\beta = 0$ allora:

$$Y_t = \alpha + \epsilon_t$$

- il reddito dei figli si differenzia dalla media α solo per effetto di una variabile casuale indipendente;
- in particolare il reddito dei figli è indipendente da quello dei padri;
- perfetta mobilità intergenerazionale e esistenza di pari opportunità.

Dato $\Delta Y_t = Y_t - Y_{t-1}$ possiamo riscrivere la **1** come:

$$\Delta Y_t = \alpha - (1 - \beta)Y_{t-1} + \epsilon_t$$

Se $0 < \beta < 1$ allora $1 - \beta > 0$ e:

- il reddito di una dinastia cresce tra la generazione $t - 1$ e t in modo inversamente proporzionale al reddito della generazione $t - 1$;
- il reddito dei figli di padri poveri crescerà di più di quello dei figli di padri ricchi, rispetto al reddito dei padri;
- col passare del tempo le dinastie inizialmente povere o ricche cambiano sistematicamente posizione nella gerarchia dei redditi.

Se $\beta > 1$ allora $1 - \beta < 0$ e:

- il reddito di una dinastia cresce tra la generazione $t - 1$ e t in modo direttamente proporzionale al reddito della generazione $t - 1$;
- il reddito dei figli di padri poveri crescerà meno di quello dei figli di padri ricchi, rispetto al reddito dei padri;
- a meno di shocks casuali ϵ_t l'ordinamento delle dinastie nella gerarchia dei redditi non cambia nel tempo.

Stime parametro β nella equazione 1

- [Becker et al. \(1986\)](#) “Human Capital and the Rise and Fall of Families” .
 - Stima una bassa correlazione intergenerazionale e conclude che la persistenza osservata è essenzialmente di natura genetica, ineliminabile ed efficiente.
- [Solon, Gary \(1992\)](#), “Intergenerational Income Mobility in the United States”
 - Usando stime più precise del reddito permanente la persistenza intergenerazionale è molto maggiore di quanto stimato da Becker et al (1986).
- [Blanden et al.\(1992\)](#) “Intergenerational Mobility in Europe and North-America” .
 - Confronti internazionali: maggiore mobilità nei paesi scandinavi.
- [Mocetti, 2006](#) “Intergenerational Mobility in Italy”
 - Stime per l'Italia dove si osserva poca mobilità sociale.

Nature vs. Nurture: i canali dell'effetto familiare

Numerosi studi hanno cercato di stimare empiricamente l'importanza di questi due canali usando dati su gemelli, fratelli naturali e fratelli adottati

- Behrman e Taubman (1977), usano dati su gemelli:
 - 44% della varianza nel grado di istruzione conseguito dipende da *Nature*.
- Behrman e Taubman (1989), estendono il campione a fratelli e parenti:
 - 80% del successo scolastico dipende da *Nature*
- [Plug e Vijverberg \(2003\)](#), usano dati su fratelli naturali e adottati
 - 55-75% del successo scolastico dipende da *Nature*

Questi studi richiedono ipotesi piuttosto “forti”.

Alcune domande poste da questi fatti

- Che relazione c'è tra mobilità intergenerazionale e pari opportunità?
- Che cosa vuol dire assicurare nel concreto una situazione di pari opportunità?
- Esiste un *trade off* tra pari opportunità ed efficienza?
- Quali sono i criteri che possono aiutarci a definire un grado ottimale di mobilità sociale?
- Che relazione c'è tra mobilità sociale e allocazione dei “talenti individuali” nelle occupazioni?
- Può (deve?) lo stato influenzare la mobilità sociale attraverso il sistema scolastico o fiscale?
- Che relazione c'è tra mobilità sociale e sistema scolastico pubblico o privato?

Il puzzle studiato da [Checchi et al. \(1999\)](#)

Italia	Stati Uniti
sistema scolastico pubblico	sistema scolastico privato
meno diseguaglianza di reddito	più diseguaglianza di reddito
meno mobilità sociale	più mobilità sociale

Due possibili spiegazioni:

- Mercati del lavoro e dei prodotti non competitivi in Italia;
- Minori incentivi ad accumulare capitale umano in un sistema scolastico statale.

I dati

Table 1: Age distribution for both generations, Italy US

Paese	Father/son	Av.age	St.Dev.	Min. age	Max. age
Italy N = 1666	Father	47	7	31	83
	Son	44	11	25	65
United States N = 1050	Father	47	7	27	74
	Son	33	5	25	59

- *Italia*

figli intervistati nel 1985; padri a 14 anni (ricordati dai figli).

- *US*

figli intervistati nel 1990; padri intervistati nel 1974.

- *Misura dello status economico:*

Reddito permanente di un individuo in una data occupazione, calcolato come il reddito mediano dei lavoratori in quella stessa occupazione nella generazione dei figli.

Diseguaglianza dei redditi in Italia e negli USA

Table 2: Inequality measures for Italy and the US

Measure	Italy	US	Italy	US
	Father	Father	Son	Son
90-10 percentile differential	140.6	164.3	131.5	150.3
relative mean deviation	12.2	14.6	13.2	14.3
coefficient of variation	33.8	37.5	34.8	36.0
standard deviation of logs	30.0	35.6	31.3	34.9
Gini coefficient	16.8	20.2	17.9	19.6
Atkinson ($\epsilon = 2$)	8.7	11.8	9.3	11.4
Theil entropy	5.0	6.6	5.5	6.1

La diseguaglianza è maggiore negli Stati Uniti:

- indipendentemente dalla misura utilizzata;
- in entrambe le generazioni.

Per il significato degli indicatori di disuguaglianza della Tabella 2, vedi il [Capitolo 18](#) del manuale Brucchi Luchino (BL nelle pagine che seguono).

Classi di reddito per studiare la mobilità intergenerazionale

Table 3: Italy: distribution across quartiles or income classes

Quartiles			Income classes		
Quartile 1	Minimum	100	100	Minimum	Class 1
	Median	144	135	Median	
	Maximum	156	144	Maximum	
Quartile 2	Minimum	156	150	Minimum	Class 2
	Median	164	164	Median	
	Maximum	186	216	Maximum	
Quartile 3	Minimum	186	219	Minimum	Class 3
	Median	219	234	Median	
	Maximum	234	318	Maximum	
Quartile 4	Minimum	234	331	Minimum	Class 4
	Median	267	369	Median	
	Maximum	474	474	Maximum	

Poichè la distribuzione dei redditi non è simmetrica, l'uso dei quartili rende il quarto gruppo troppo ampio in Italia.

Table 4: USA: distribution across quartiles or income classes

Quartiles			Income classes		
Quartile 1	Minimum	100	100	Minimum	Class 1
	Median	139	130	Median	
	Maximum	165	139	Maximum	
Quartile 2	Minimum	165	148	Minimum	Class 2
	Median	187	174	Median	
	Maximum	217	215	Maximum	
Quartile 3	Minimum	217	215	Minimum	Class 3
	Median	252	261	Median	
	Maximum	261	314	Maximum	
Quartile 4	Minimum	261	322	Minimum	Class 4
	Median	330	337	Median	
	Maximum	463	463	Maximum	

Anche negli USA poichè la distribuzione dei redditi non è simmetrica, l'uso dei quartili rende il quarto gruppo troppo ampio.

Matrici di transizione tra classi di reddito

Table 5: Italy: interclass transition probabilities

	Son C1	Son C2	Son C3	Son C4	Abs.freq.
Father C1	21.8	50.4	22.3	5.4	367
Father C2	12.0	55.9	25.8	6.3	884
Father C3	5.9	27.0	51.6	15.5	341
Father C4	4.0	16.2	32.4	47.3	74
Abs.freq.	209	783	510	164	1666

Table 6: USA: interclass transition probabilities

	Son C1	Son C2	Son C3	Son C4	Abs.freq.
Father C1	25.9	36.4	31.4	6.3	239
Father C2	22.5	37.7	29.7	10.1	337
Father C3	9.3	31.0	41.7	18.0	355
Father C4	4.2	15.1	42.0	38.7	119
Abs.freq.	176	342	373	159	1050

Table 7: Scalar indicators of mobility for interclass transition matrices

	Italy	USA	Eq. opp.
$ML = 1 - \lambda_2 $	0.55	0.65	1
$MT = \frac{k - tr(P)}{k - 1}$	0.74	0.85	1
$MD = 1 - det(P) ^{(1/(k-1))}$	0.79	0.90	1
$MB = \sum_i \sum_j f_{ij} i - j $	0.62	0.80	-
$MA = \sum_i \sum_j f_{ij} W_i - W_j $	22.44	27.55	-

Gli Stati Uniti sono più vicini dell'Italia ad una situazione di pari opportunità e perfetta mobilità sociale.

Una matrice di transizione implica l'esistenza di pari opportunità se tutte le righe sono uguali, ossia se le dinastie hanno la stessa probabilità di arrivare in ogni classe di arrivo (figli), indipendentemente dalla classe di partenza (padri).

Per un approfondimento sugli indicatori di mobilità della Tabella 7, vedi il [Checchi et al. \(1995\)](#).

Table 8: Determinants of the probability that a son is in income class 3 or 4

	ITALY			USA		
	model 1	model 2	model 3	model 1	model 2	model 3
Father: class 3 or 4	0.37 (.03)	0.35 (.03)		0.22 (.03)	0.19 (.03)	
Father: college degree	0.18 (.09)	0.02 (.09)		0.19 (.04)	0.05 (.05)	
Son: college degree		0.31 (.05)	0.39 (.05)		0.47 (.03)	0.50 (.03)
Father's age	-0.001 (.002)	-0.001 (.002)		0.009 (.002)	0.004 (.003)	
Son's age		-0.003 (.001)	-0.03 (.001)		0.005 (.004)	0.007 (.003)
observed prob.	.427	.427	.427	.508	.508	.508
predicted prob.	.427	.428	.428	.511	.532	.530
Pseudo R2	.08	.10	.04	.07	.19	.16
log-likelihood	-939	-918	-984	-665	-578	-597
sample size	1505	1505	1505	1037	1037	1037

Effetto della scolarità e del background familiare sul successo economico dei figli (Tabella 8)

La variabile dipendente prende valore 1 se il figlio è nelle due classi più alte.

Il modello 1 dice che in entrambi i paesi il reddito paterno è più rilevante dell'istruzione paterna nel determinare la classe di reddito del figlio.

Il modello 2 dice che:

- in entrambi i paesi l'effetto dell'istruzione paterna passa essenzialmente attraverso l'istruzione del figlio;
- In Italia ... conta più scegliersi una “buona” famiglia che conseguire la laurea; in USA è vero il contrario.

Il confronto tra modello 2 e modello 3 indica che il background familiare è più rilevante in Italia che non negli USA.

Mobilità nei livelli di istruzione

Table 9: Actual marginal and limiting distributions for education in Italy and USA

	Italy no coll.	Italy coll.	Italy no HS	Italy HS+	USA no coll.	USA coll.
Father	0.97	0.03	0.92	0.08	0.84	0.16
Son	0.91	0.09	0.71	0.29	0.73	0.27
Limit	0.83	0.17	0.30	0.70	0.65	0.35

Il livello di istruzione dei figli è aumentato notevolmente rispetto a quello dei padri, soprattutto in Italia.

Ma in Italia non tutte le classi sociali hanno partecipato in modo simile a questo aumento assoluto di capitale umano.

Table 10: Italy: transition probabilities from “no college” to “college”

	Son E1	Son E2	Abs.freq.
Father E1	92.9	7.1	1462
Father E2	34.9	65.1	43
Abs.freq.	1374	131	1505

Note: each cell contains the row-to-column transition probability. E1 = no college degree; E2 = completed college degree.

Table 11: USA: transition probabilities from “no college” to “college”

	Son E1	Son E2	Abs.freq.
Father E1	79.2	20.8	870
Father E2	38.9	61.1	167
Abs.freq.	754	283	1037

Note: each cell contains the row-to-column transition probability. E1 = no college degree; E2 = completed college degree.

Table 12: Italy: transition probabilities from “less than highschool” to “highschool or +”

	Son E1	Son E2	Abs.freq.
Father E1	75.9	24.1	1389
Father E2	10.3	89.7	116
Abs.freq.	1066	439	1505

Note: each cell contains the row-to-column transition probability. E1 = less than highschool; E2 = completed highschool or more.

Table 13: Scalar indicators of mobility for educational transition matrices

	Italy E2 = coll.	USA E2 = coll.	Italy E2 = HS or +	Eq. opp.
$OR = \frac{p_{12}/p_{11}}{p_{22}/p_{21}}$	24.6	6.0	27.3	1
$MT = \frac{k - tr(P)}{k - 1}$	0.42	0.60	0.34	1
$MB = \sum_i \sum_j f_{ij} i - j $	0.12	0.27	0.14	-

La probabilità di laurearsi se il padre non è laureato è molto maggiore negli Usa che non in Italia.

Il “rischio relativo” (*odds ratio*) di laurearsi aumenta di

- 6 volte in USA
- 25 volte Italia

avendo il padre laureato piuttosto che non laureato.

Anche definendo il diploma di scuola superiore come livello alto di istruzione Il “rischio relativo” (*odds ratio*) di avere questo diploma aumenta di 27 volte in Italia se il padre ha il diploma.

2.1.2 Una teoria per riflettere su questi fatti

Un sistema scolastico “pubblico” eguaglia l’offerta di istruzione disponibile alle dinastie rendendola indipendente dal reddito familiare e dai talenti individuali:

- favorisce una maggiore mobilità sociale.

Un sistema scolastico “privato” permette alle dinastie di differenziare gli investimenti in istruzione in modo meritocratico a seconda del talento dei figli:

- favorisce una maggiore mobilità sociale.

Il secondo effetto prevale se un buon abbinamento tra talenti individuali e qualità dell’istruzione è il fattore più rilevante per l’accumulazione di capitale umano:

- sistema scolastico pubblico per l’istruzione primaria e privato per la terziaria?

L’uniformità del sistema scolastico pubblico toglie ai poveri con talento uno strumento per competere con i ricchi.

Il sistema dei “Vouchers” può consentire di separare l’obiettivo distributivo dall’obiettivo di efficienza.

Gli ingredienti principali di un possibile modello per razionalizzare questi fatti

Ogni persona fa parte di una dinastia e vive per due periodi:

- ha un capitale umano base pari a 1;
- ha un talento naturale che può essere “Low” o “High”.

Il talento non serve per produrre ma solo per accumulare capitale umano.

Ogni persona guadagna un salario pari al suo capitale umano.

Ogni generazione inferisce dalle generazioni precedenti la probabilità di avere un talento alto o basso.

L'accumulazione di capitale umano a scuola dipende da:

- “self confidence”: quanto uno crede nel proprio talento;
- talento effettivo;
- impegno;
- capitale umano (ossia reddito) dei genitori;
- qualità del sistema scolastico (diversa nel sistema pubblico e privato).

Tempi

Nel primo periodo t la persona:

- conosce la storia scolastica familiare;
- ipotizza una probabilità (*belief*) di avere un talento alto;
- prende coscienza della qualità dell'istruzione disponibile;
- decide se andare a scuola e se sì, con quanto impegno.

Nel secondo periodo $t + 1$ la persona:

- conosce il suo talento effettivo solo se ha deciso di andare a scuola;
- aumenta il capitale umano se ha scelto di andare a scuola e ha talento alto;
- produce;
- decide che livello di istruzione lasciare alla generazione successiva (in modo diverso nel sistema pubblico e privato);
- consuma quel che rimane;
- ha un figlio e il ciclo ricomincia.

Trasmissione del talento e accumulazione di capitale umano

	$a_{t+1} = L$	$a_{t+1} = H$
$a_t = L$	$1 - \alpha$	α
$a_t = H$	α	$1 - \alpha$

Nota: $a =$ talento e $\alpha \in (0, 1/2)$

Se la persona decide di andare a scuola il capitale umano accumulato è:

$$h_{t+1} = \begin{cases} 1 & \text{if } a_{t+1} = L; \\ \theta (1 - n_t)^\beta e_t^\gamma h_t^\delta & \text{if } a_{t+1} = H; \end{cases}$$

- n_t tempo libero (ossia tempo non dedicato alla istruzione),
- e_t qualità dell'istruzione scolastica,
- h_t capitale umano (e quindi reddito) del padre.
- $h_{t+1} > 1$ implica che il talento è alto.

“Self-confidence” da una generazione alla successiva

Data una probabilità soggettiva ν^{t-1} che un padre abbia un talento alto, che probabilità soggettiva può avere il figlio riguardo al suo talento?

$$\nu^t = \nu^{t-1}(1 - \alpha) + (1 - \nu^{t-1})\alpha = \alpha + (1 - 2\alpha)\nu^{t-1}$$

Dopo un tentativo fallito di scolarizzazione di una generazione, la successiva ha una probabilità α di avere un talento elevato (la probabilità minima possibile).

Con il passare delle generazioni la “self-confidence” aumenta anche se la dinastia non prova mai ad andare a scuola, semplicemente perchè i “difetti” degli avi si “diluiscono” nel tempo.

Quando la “self-confidence” è aumentata a sufficienza, la dinastia riprova a salire la scala sociale andando a scuola.

La mobilità sociale è maggiore dove è necessario un livello inferiore di “self-confidence” per investire in istruzione.

Preferenze

$$U(n_t, c_{t+1}, \nu_{t+1}, e_{t+1}) = \log n_t + \log c_{t+1} + \nu_{t+1} \log e_{t+1}$$

- n_t : tempo libero,
- c_{t+1} : consumo,
- ν_{t+1} : probabilità che la generazione successiva abbia talento alto
- e_{t+1} : qualità dell'istruzione trasferita alla generazione successiva.

e_{t+1} dipende dal sistema scolastico privato o pubblico .

Qualità dell'istruzione privata o pubblica

Sistema privato

- e_{t+1} è deciso dal padre in base al proprio reddito.
- La decisione è presa in funzione del livello di “self-confidence”.
- deve soddisfare il vincolo: $c_{t+1} + e_{t+1} \leq h_{t+1}$.

Sistema pubblico

- Ogni generazione è tassata per finanziare l'istruzione della successiva.
- L'aliquota τ_{t+1} è votata a maggioranza dalla popolazione.
- $e_{t+1} = \tau_{t+1}H_{t+1}$ l'istruzione è uguale per tutti i figli della generazione successiva, indipendentemente dalla probabilità che abbiano talento alto o basso.
- deve soddisfare il vincolo: $c_{t+1} \leq h_{t+1}(1 - \tau_{t+1})$.

Il modello si risolve confrontando le scelte di una ipotetica dinastia nel sistema pubblico e nel sistema privato.

Il ciclo di ascesa e caduta delle dinastie

Consideriamo la storia di una dinastia dopo una caduta. La generazione successiva ha il minimo di “self-confidence”

Di generazione in generazione la “self-confidence” aumenta fino a che la dinastia riprova a investire in istruzione e a risalire la scala sociale.

Il sistema scolastico pubblico:

- migliora la qualità dell’istruzione per i poveri, rendendo necessaria una minore “self-confidence” per investire;
- ma riduce l’incentivo a investire perchè tassa il reddito dell’investimento.

In un sistema scolastico privato:

- è più costoso per i poveri ottenere una istruzione di qualità elevata
- ma i genitori possono decidere in modo più efficiente la qualità di istruzione adatta ai figli
- e i rendimenti dell’investimento non sono tassati.

Conclusioni dal modello

La considerazione dei “talenti individuali” arricchisce l’analisi della mobilità sociale

La mobilità sociale è efficiente quando consente una perfetta corrispondenza tra talenti ed occupazioni.

Un sistema scolastico pubblico:

- riduce i costi di un investimento in istruzione per le famiglie povere,
- ma può anche ridurre l’incentivo ad effettuare un tale investimento

Il sistema ideale è un sistema che distingue il problema distributivo dal problema della combinazione efficiente tra talenti e qualità dell’istruzione.

2.2 Perché andiamo a scuola?

Il modello sopra esaminato suggerisce che:

- l'acquisizione di capitale umano a scuola è un fattore fondamentale di potenziale riequilibrio delle differenze familiari;
- la decisione di andare scuola può essere interpretata come un investimento al fine di aumentare la nostra capacità di produrre reddito;
- la decisione se investire o no dipende dai costi e dai benefici attesi.

In aggiunta, l'istruzione può anche essere considerata un bene di consumo

In questa parte del corso ci proponiamo di studiare

- che cosa determina le scelte scolastiche di una persona;
- che conseguenze ha l'istruzione sulla carriera lavorativa di una persona.

Iniziamo analizzando alcuni fatti stilizzati relativi all'istruzione nel [Capitolo 2](#) di BL.

2.2.1 Un modello della quantità di istruzione acquisita da un individuo

Consideriamo adesso una delle pietre miliari dell'analisi economica del mercato del lavoro: il modello di accumulazione del capitale umano di Becker (1975), nella esposizione di Card (1995a, 1999) contenuta nel [Capitolo 20](#) BL (per una esposizione alternativa vedi il [Capitolo 2](#)).

Becker propone una spiegazione razionale delle scelte di istruzione degli individui, con precise implicazioni empiriche che gli economisti hanno cercato di verificare.

Il modello ipotizza che un anno di istruzione abbia un effetto “causale” (non necessariamente uguale per tutti) sui redditi che un lavoratore conseguirà.

Le implicazioni di politica economica di questa teoria hanno una rilevanza fondamentale. Riflettete ad esempio sui paesi in via di sviluppo o sulla decisione di prolungare l'obbligo scolastico.

L'identificazione e la misurazione empirica di questo effetto causale non è facile, come vedremo, e richiede di definire che cosa sia un effetto “causale”: un problema su cui scienziati e filosofi si sono interrogati fin dall'antichità.

La funzione generatrice del reddito

L'assunzione centrale del modello è che andare a scuola sia un modo per accumulare capitale umano e che un maggiore capitale umano si traduca in maggiori redditi nel mercato del lavoro secondo questa funzione:

$$Y = Y(S) \quad (2)$$

dove:

- S è il numero di anni di scuola;
- $Y(S)$ è il reddito generato dal capitale umano accumulato in S anni di scuola;
- la funzione generatrice del reddito è crescente e concava

$$Y' > 0 \quad \text{e} \quad Y'' < 0$$

Esercizio: qual'è il [grafico](#) di questa funzione? (Vedi Figura 20.1 BL)

La funzione obiettivo degli individui

Gli individui scelgono il numero ottimale di anni di scuola S in modo da massimizzare il valore scontato attuale atteso dei loro redditi futuri:

$$V(S, Y) = \int_S^{\infty} Y(S)e^{-rt} dt = \frac{Y(S)e^{-rS}}{r}, \quad (3)$$

dove $Y(S)$ è il reddito potenziale in funzione di S e r è il tasso di sconto.

Passando ai logaritmi possiamo riscrivere questo valore scontato atteso come

$$\tilde{U}(S, Y) = \log(V(S, Y)) = \log(Y) - rS - \log(r) \quad (4)$$

- andare a scuola è utile a un individuo perchè produce reddito
- ma è costoso perchè costringe l'individuo a rinunciare al reddito che conseguirebbe se lavorasse invece di andare a scuola. (“costo opportunità”)

(Figura 2.5 BL: esempio di andamento del reddito a due livelli di S)

È questa una visione troppo restrittiva (e ragionieristica...) della decisione di andare a scuola?

Una specificazione più generale della funzione obiettivo

È possibile generalizzare la funzione obiettivo assumendo:

$$U(S, Y) = \log(Y) - h(S) \quad (5)$$

dove assumiamo che $h(s)$ sia crescente e convessa o al massimo lineare

$$h' > 0 \quad \text{e} \quad h'' \geq 0$$

$h(s)$ racchiude in modo implicito l'aggregazione delle diverse componenti dei costi e dei benefici dell'istruzione:

- il reddito da lavoro perso negli anni passati a scuola (costo opportunità);
- le tasse scolastiche;
- il costo psicologico di star seduti nel banco ad ascoltare una lezione;
- il beneficio intrinseco dell'imparare qualcosa (scuola come bene di consumo);
- costo del denaro necessario per finanziare l'istruzione, ragionevolmente maggiore per gli individui più poveri.

Forma delle curve di indifferenza in questo problema di ottimo

Possiamo fare diverse ipotesi sulla concavità della funzione $h(s)$ da cui discendono forme diverse delle curve di indifferenza:

- Se $h(s)$ è lineare ritorniamo al caso dell'equazione 4 (vedi la [Figura 20.1 BL](#));
- Se $h(s)$ è convessa le curve di indifferenza sono convesse (vedi la [Figura 2.1 BL](#)).

In entrambi i casi notate che le curve di indifferenza sono crescenti nel piano S, Y . Qual'è il significato economico di questa ipotesi?

Più in generale cosa cambia nelle curve di indifferenza e nel vincolo di questo problema di ottimo rispetto al caso classico del problema del consumatore nel corso base di microeconomia?

Soluzione del problema di ottimo

Il problema di ottimo che l'individuo deve risolvere è quindi:

$$\text{Max } U(Y, S) = \log(Y) - h(S) \quad \text{sotto il vincolo} \quad Y = Y(S) \quad (6)$$

Gli anni ottimali di scuola sono dati dalla condizione del primo ordine:

$$\frac{Y'(S)}{Y(S)} = h'(S) \quad (7)$$

- $\frac{Y'(S)}{Y(S)} =$
 - è il rendimento marginale di un anno di scuola, ovvero
 - il tasso marginale a cui è possibile trasformare un anno di scuola in reddito;
- $h'(S) =$
 - è il costo marginale di un anno di scuola, ovvero
 - il tasso marginale a cui si è disposti a sostituire reddito e anni di scuola.

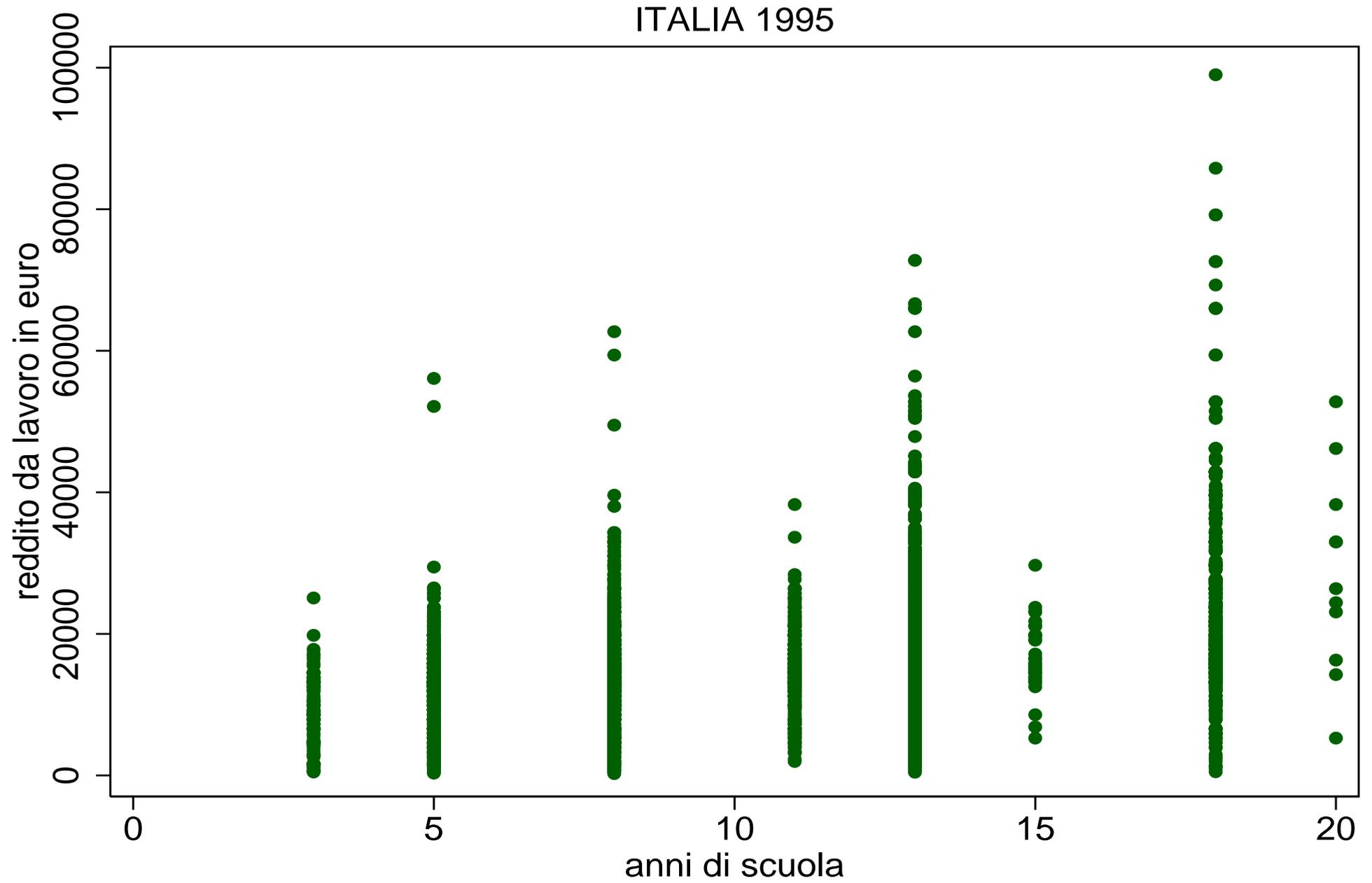
Vedi le Figure [2.1 BL](#) o [20.1 BL](#).

2.2.2 Un primo confronto tra modello e dati

Per confrontare il modello con l'evidenza empirica abbiamo bisogno di:

- Una banca dati:
 - utilizzeremo un estratto dell' "Indagine sui Bilanci delle Famiglie Italiane" (SHIW), raccolta dalla Banca d'Italia per gli anni 1991, 1993, 1995;
 - la documentazione e le indagini per anni precedenti e successivi sono scaricabili dal sito della [Banca d'Italia](#);
 - Indagini analoghe sono disponibili per numerosi altri paesi.
- Un software per analizzare i dati:
 - i software disponibili sono numerosi;
 - useremo Stata che è installato nell'aula informatica della facoltà;
 - sul sito (<http://www.stata.com/>) potete trovare documentazione, tutorial e corsi on line;
 - [qui](#) potete trovare un utile stata tutorial preparato da Daniela Vuri per gli studenti dell'Università di Firenze.

Figure 1: Anni di scuola e reddito da lavoro



Il modello base non è compatibile con i dati

Il modello base così come sopra descritto non prevede eterogeneità tra individui e quindi genera una sola combinazione ottimale di anni di scuola S e redditi da lavoro Y

La Figura 1 indica invece scelte eterogenee di istruzione e redditi diversi a parità di anni di scuola.

Entriamo in Stata per conoscere i dati e vedere come la figura è stata costruita:

- i comandi di stata utilizzati;
- cosa succede se li eseguiamo (nota: .do file e .dta file devono essere nella stessa cartella).

Il confronto con l'evidenza empirica suggerisce che dobbiamo introdurre qualche elemento di eterogeneità nel modello se vogliamo che possa dire qualche cosa di interessante.

2.2.3 Come introdurre eterogeneità nel modello base

Card (1995a, 1999) assume che esista eterogeneità tra individui

- nei rendimenti marginali dell'istruzione:

$$\left[\frac{Y'(S)}{Y(S)} \right]_i = \beta_i(S) = b_i - k_b S \quad (8)$$

dove b_i misura, ad esempio, differenze nell'abilità individuale;

- nei costi marginali dell'istruzione.

$$[h'(S)]_i = \delta_i(S) = r_i + k_r S \quad (9)$$

dove r_i misura, ad esempio, differenze di costi di accesso al credito per l'istruzione.

Interpretazione dell'eterogeneità dei rendimenti marginali

Stiamo assumendo che il rendimento marginale del tempo passato a scuola

$$\left[\frac{Y'(S)}{Y(S)} \right]_i = \beta_i(S) = b_i - k_b S$$

- sia una funzione lineare decrescente degli anni di scuola;
- abbia inclinazione k_b uguale per tutti gli individui;
- abbia un intercetta b_i diversa da individuo a individuo;
- individui con intercetta b_i più elevata sono più “abili” nel senso che da ogni unità marginale di tempo passata a scuola traggono un rendimento maggiore.

Notate che questa è una ipotesi sulla “derivate” della funzione generatrice del reddito. Integrando $\beta_i(S)$ nella [9](#) si ottiene che:

$$[Y(S)]_i = a e^{(b_i S - (\frac{k_b}{2} S^2))} \quad (10)$$

Esercizio: disegnate la funzione generatrice del reddito e la funzione del rendimento marginale degli anni di scuola per due individui con diversa abilità.

È questa una definizione ragionevole di abilità?

Notate che questa assunzione implica una caratterizzazione particolare di cosa voglia dire “abilità”:

- l’abilità è una dote che aumenta la capacità di far fruttare il tempo passato a scuola

Con curve di indifferenza omotetiche standard, questa ipotesi assicura che individui *più* abili scelgano un numero maggiore di anni di scuola.

Avremmo potuto definire l’“abilità” in modo diverso:

- l’abilità è una dote che consente ad una persona di guadagnare di più indipendentemente dal tempo passato a scuola

In questo secondo caso, con curve d’indifferenza omotetiche standard, accadrebbe che siano gli individui *meno* abili ad andare a scuola per un numero maggiore di anni.

Esercizio: disegna i grafici delle scelte ottime al variare dell’abilità a seconda delle due diverse ipotesi sul significato di abilità.

Interpretazione dell'eterogeneità dei costi marginali

Stiamo assumendo che il costo marginale del tempo passato a scuola

$$[h'(S)]_i = \delta_i(S) = r_i + k_r S$$

- sia una funzione lineare crescente degli anni di scuola;
- abbia inclinazione k_r uguale per tutti gli individui;
- abbia un intercetta r_i diversa da individuo a individuo;
- gli individui con intercetta r_i più elevata sono individui per i quali ogni unità marginale di tempo passata a scuola ha un costo maggiore.

Esercizio: disegnate le curve di indifferenza corrispondenti all'ipotesi che abbiamo fatto su $\delta_i(S)$ a seconda che

1. $k_r = 0$ and $r_i \neq r_j$ for $i \neq j$
2. $k_r > 0$ and $r_i \neq r_j$ for $i \neq j$

Per rendere più concreta l'analisi interpretiamo r_i come un indicatore di povertà: gli individui più poveri hanno vincoli di liquidità maggiori e quindi costi marginali maggiori di finanziamento dell'istruzione.

Scelte ottimali di istruzione con eterogeneità

Sostituendo il rendimento marginale 8 e il costo marginale 9 nella condizione del primo ordine 7, possiamo ricavare il numero ottimale di anni di scuola scelto da ciascun individuo:

$$S_i^* = \frac{(b_i - r_i)}{k_b + k_r} \quad (11)$$

Il modello adesso genera risultati compatibili con la Figura 1.

Notate che:

- le scelte ottimali di istruzione variano tra individui perchè gli individui hanno costi e benefici marginali diversi da ogni anno aggiuntivo di istruzione;
- abbiamo ipotizzato che i benefici marginali dipendano dall'abilità e i costi marginali dipendendo dalla ricchezza (della famiglia d'origine);
- per data ricchezza, gli individui più abili scelgono più anni di scuola;
- per data abilità gli individui più ricchi scelgono più anni di scuola.

Una versione semplificata ma istruttiva del modello

Consideriamo una versione semplificata del modello in cui assumiamo (vedi esempio 1 a pagina 49) curve di indifferenza lineari con intercette diverse da individuo a individuo

$$k_r = 0 \quad \text{and} \quad r_i \neq r_j \quad \text{for} \quad i \neq j$$

Indicando con y , il logaritmo del reddito, il modello assume questa forma:

$$\text{Max } U_i(y, S) = y - r_i S \quad (12)$$

$$\text{subject to } y = b_i S - \frac{k_b}{2} S^2$$

$$\beta_i(S) = b_i - k_b S. \quad (13)$$

$$S_i^* = \frac{(b_i - r_i)}{k_b} \quad (14)$$

$$\beta_i^* = b_i - k_b S_i^* = r_i. \quad (15)$$

Nel seguito omettiamo per semplicità * nel denotare le scelte ottimali.

Un esempio con 4 tipi di soggetti

Assumiamo che ci siano due valori per ogni parametro di eterogeneità:

$$b_H > b_L$$

$$r_H > r_L$$

e quindi 4 possibili combinazioni differenti di abilità e ricchezza:

$$g = \{LH, HH, LL, HL\}.$$

Ogni gruppo di soggetti $g = \{i, j\}$ sceglie un numero diverso di anni di istruzione:

$$S_g \equiv S_{i,j} = \frac{(b_i - r_j)}{k_b}, \quad (16)$$

Vedi la Figura [20.2 BL](#)

Il numero di anni di scuola cresce con gli indicatori di abilità b_i e di ricchezza $-r_j$.

Il rendimento di un anno di istruzione nei quattro gruppi

Il rendimento di un anno di istruzione è diverso per i poveri ed i ricchi, ma, a parità di ricchezza, è uguale per i due livelli di abilità data la linearità delle curve di indifferenza (Vedi la [Figura 20.2](#) BL):

$$\begin{aligned}\beta_{LH} &= \beta_{HH} = r_H \\ \beta_{LL} &= \beta_{HL} = r_L.\end{aligned}\tag{17}$$

In particolare notate che essendo $r_H > r_L$, i poveri hanno un rendimento marginale maggiore dei ricchi indipendentemente dall'abilità.

Infatti la concavità della funzione generatrice del reddito implica che il rendimento marginale del tempo passato a scuola sia decrescente.

Quindi chi va a scuola più a lungo (i ricchi) ha un rendimento marginale inferiore.

Esercizio: come cambierebbe il rendimento marginale di un anno di istruzione in funzione dell'abilità a parità di ricchezza se le curve di indifferenza fossero convesse e omotetiche?

Il rendimento medio di un anno di istruzione nella popolazione

Per definire il rendimento medio nella popolazione ipotizziamo una distribuzione di frequenza della popolazione nei quattro gruppi.

$$\{P_{LL}, P_{LH}, P_{HL}, P_{HH}\}$$

Data questa distribuzione il rendimento medio è

$$\bar{\beta} = (P_{LH} + P_{HH})r_H + (P_{LL} + P_{HL})r_L = \bar{r}, \quad (18)$$

Ossia è la media aritmetica del rendimento nei quattro gruppi

Esercizio: come possiamo scrivere il rendimento medio in funzione di r_H e r_L se la distribuzione è uniforme (ossia: $P_g = P = 0,25 \forall g$)?

La controversa correlazione tra abilità e ricchezza

Ipotizziamo che:

- l'abilità sia parzialmente ereditata dai genitori;
- i genitori più abili abbiano in media più anni di scuola e un reddito maggiore;
- genitori con redditi maggiori abbiano tassi di sconto r più bassi perchè
 - hanno vincoli inferiori di accesso al credito,
 - traggono maggiore utilità dall'istruzione come bene di consumo.

Date queste ipotesi (che non sono le uniche possibili):

1. come è distribuita la popolazione nei quattro gruppi?
2. sono più alte le frequenze P_{HH} e P_{LL} oppure le frequenze P_{HL} e P_{LH} ?
3. che correlazione deve esistere nei dati tra anni di scuola e reddito?
4. e come deve essere la correlazione tra abilità e povertà?

Esercizio: Prova a ragionare su ipotesi alternative.

Un confronto più approfondito tra modello e dati

Abbiamo esteso il modello ottenendo un set più ricco di implicazioni empiriche.

Ci proponiamo ora di affiancare l'analisi teorica con quella empirica al fine di:

1. verificare se il modello è una buona rappresentazione del “modello vero di generazione dei dati osservati” (*Data Generating Process* (DGP));
2. stimare
 - la distribuzione delle variabili osservabili nella popolazione;
 - la correlazione tra le variabili nella popolazione;
3. risalire, se possibile ai parametri strutturali del modello;
4. identificare relazioni “causali” tra le variabili che “possiamo manipolare” e le “variabili di risultato” a cui siamo interessati.

Dobbiamo però fare un passo indietro e ripassare alcuni strumenti per l'analisi empirica a disposizione dell'economista. A questo fine faremo riferimento al [Capitolo 19](#).

2.2.4 Un ripasso di alcune nozioni e strumenti di econometria

Considerate la seguente relazione lineare:

$$W_i = \alpha + \beta S_i + \epsilon_i \quad (19)$$

- S_i = anni di istruzione;
- W_i = logaritmo del reddito da lavoro (perchè il logaritmo?);
- α = intercetta della retta di regressione;
- β = coefficiente angolare della retta di regressione;
- ϵ = ciò che come ricercatori non osserviamo e non conosciamo, ma che può essere noto ai soggetti studiati.

Questa relazione è una approssimazione empirica della relazione teorica tra anni di istruzione e reddito implicita nel nostro modello (DGP).

Impariamo qualcosa dalla stima di questa relazione?

I problemi che dobbiamo risolvere

1. come stimare i parametri α e β ;
2. come valutare la bontà delle stime ottenute;
3. come testare delle ipotesi sui parametri α e β ;
4. stabilire se esiste una corrispondenza, ed eventualmente quale, tra le stime $\hat{\alpha}$ e $\hat{\beta}$ e i parametri del modello teorico (DGP) che ipotizziamo abbia generato i dati;
5. stabilire se la relazione stimata può essere interpretata come una relazione “causale” ed eventualmente per quali soggetti nella popolazione.

L'ultimo problema richiede una definizione di “causalità”.

La distinzione tra

- il problema della identificabilità;
- il problema dell'inferenza.

Come stimare i parametri α e β

Il concetto di stimatore

Il concetto di stima.

Lo stimatore dei Minimi Quadrati Ordinari (OLS).

$$\hat{\beta} = \frac{\widehat{Cov}(W_i, S_i)}{\widehat{Var}(S_i)} \quad (20)$$

$$\hat{\alpha} = \widehat{E}(W_i) - \hat{\beta}\widehat{E}(S_i) \quad (21)$$

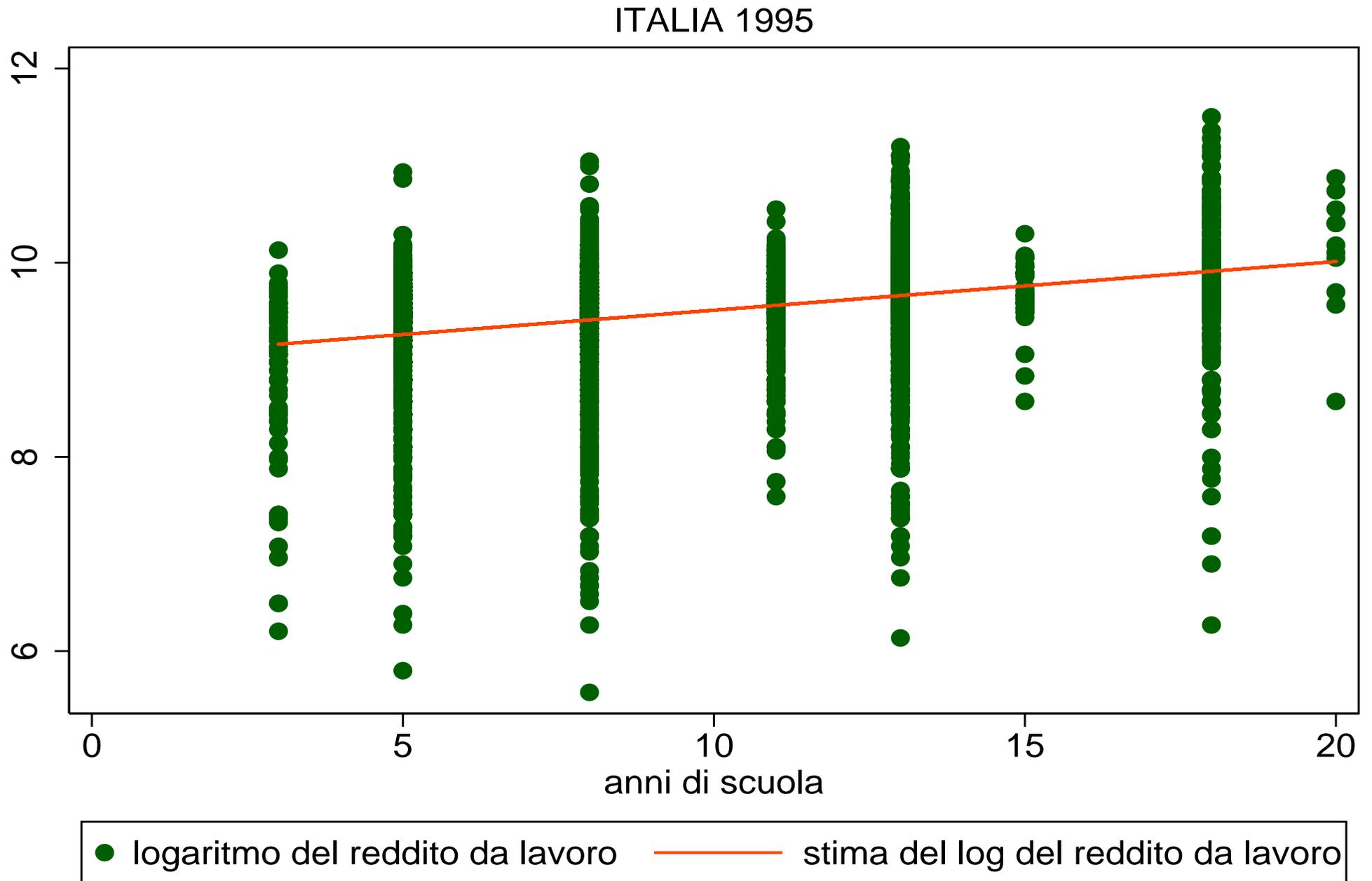
Errore campionario, varianza, errore standard di uno stimatore.

Come stimare la varianza di ϵ e la varianza dello stimatore OLS

Come [stimare i parametri \$\alpha\$ e \$\beta\$](#) con Stata usando i dati della [Figura 1](#).

$$\hat{\alpha} = 9.01 \quad \hat{\beta} = 0.05$$

Figure 2: Anni di scuola e reddito da lavoro osservato e stimato



Come valutare la bontà delle stime ottenute

Alcune proprietà desiderabili di uno stimatore $\hat{\beta}$ per un parametro β :

- non distorsione: in media $\hat{\beta}$ è uguale a β ;
- efficienza: la varianza di $\hat{\beta}$ è “piccola”;
- consistenza: al crescere del campione, $\hat{\beta}$ diventa “indistinguibile” da β .

Quali ipotesi sono necessarie per queste proprietà?

1. Gli errori devono avere media zero

$$E(\epsilon_i) = 0$$

2. Gli errori devono avere la stessa varianza (*omoschedasticità*).

$$\text{Var}(\epsilon_i) = E(\epsilon_i^2) = \sigma^2$$

3. Gli errori di due osservazioni diverse non devono essere correlati

$$\text{Cov}(\epsilon_i, \epsilon_j) = 0 \quad \text{for } i \neq j$$

4. Gli errori e il regressore non devono essere correlati

$$\text{Cov}(\epsilon_i, S_j) = 0 \quad \text{for all } i \text{ and } j$$

Come testare ipotesi sui parametri α e β

La legge dei grandi numeri e la normalità dello stimatore OLS

Come calcolare la probabilità che la stima assuma valori compresi in un dato intervallo.

Intervalli di confidenza

Test di ipotesi sul valore del parametro β .

- ipotesi nulla e ipotesi alternativa;
- livello di significatività.
- la statistica *t-di Student*;
- teoria classica del test delle ipotesi;
- errori di prima e seconda specie e potenza di un test.

Che relazione esiste tra $\hat{\alpha}$ e $\hat{\beta}$ e i parametri del DGP

Nella Figura 2 abbiamo stimato una relazione lineare crescente tra anni di scuola e reddito da lavoro con un coefficiente angolare (derivata)

$$\hat{\beta} = \frac{\widehat{Cov}(W_i, S_i)}{\widehat{Var}(S_i)} = 0.05$$

Possiamo da questa stima risalire ai parametri del modello (DGP) che ipotizziamo abbia generato i dati?

Considerate la controversa relazione tra abilità e ricchezza discussa a pag. 55.

Se il modello teorico proposto è quello che ha generato i dati, una covarianza positiva tra anni di istruzione e reddito implica che:

- P_{HH} e P_{LL} potrebbero essere uguali o inferiori a P_{HL} e P_{LH} ;
- la correlazione tra b_i e r_i potrebbe essere nulla o negativa;
- la correazione tra abilità e ricchezza potrebbe essere nulla o positiva.

La Figura 20.4 BL descrive con dati simulati il caso ipotetico in cui $\hat{\beta} < 0$, dal quale trarremmo conclusioni opposte.

Possiamo interpretare $\hat{\beta}$ come la stima di un effetto causale?

Abbiamo visto che possiamo imparare qualcosa di rilevante riguardo al modello dal segno e dal valore di

$$\hat{\beta} = \frac{\widehat{Cov}(W_i, S_i)}{\widehat{Var}(S_i)} = 0.05$$

$\hat{\beta}$ è una misura della correlazione tra S_i e W_i , e, dato il DGP, il fatto che questa correlazione sia positiva ci dice qualcosa sulle relazioni strutturali del modello.

Non dice molto di più però. Direbbe di più se potessimo essere certi che le ipotesi di pagina 61 fossero verificate (in particolare la 1, 3 e 4).

Se quelle ipotesi non sono vere la stima OLS di $\hat{\beta}$ che abbiamo ottenuto:

- *non identifica* alcun parametro strutturale del modello;
- *non identifica* l'effetto causale degli anni di istruzione sul reddito.

Cerchiamo di capire perchè aprendo una parentesi sul problema della causalità (vedi il [Capitolo 20](#) BL).

2.2.5 Il problema della causalità

Considerate queste domande:

- Il fumo causa il cancro e le malattie polmonari?
- L'aspirina riduce il rischio di infarto?
- Un anno in più di istruzione aumenta i redditi di una persona
- Un lavoro temporaneo è un trampolino verso un lavoro permanente?
- I vincoli ai licenziamenti hanno un effetto sulla disoccupazione?

Le risposte a queste domande (e a molte altre che influenzano la nostra vita di tutti i giorni) richiedono di sapere identificare e misurare le relazioni causali: un problema antico in filosofia e statistica.

Abbiamo bisogno di un modello per definire la causalità e per misurarla: analizziamo quello proposto nel [Capitolo 20](#) BL

Un modello per definire la causalità

Considerate una popolazione di soggetti; per ogni soggetto osserviamo una variabile D e una variabile Y .

Osserviamo che D e Y sono correlati. Possiamo concludere che questa *correlazione* implica una relazione *causale*?

In generale no perchè:

- la relazione può essere affetta da fattori di confondimento;
- la direzione di causalità non è immediatamente ovvia.

Ci proponiamo di capire in quale senso e sotto quali ipotesi possiamo concludere dall'evidenza che D *causa* Y .

È utile pensare a questo problema utilizzando la terminologia dell'analisi medica sperimentale

- i è un indice per i soggetti della popolazione studiata.
- D_i denota lo stato di esposizione al *trattamento*:
 $D_i = 1$ se il soggetto i è stato esposto al trattamento;
 $D_i = 0$ se il soggetto i non è stato esposto al trattamento.
- $Y_i(D_i)$ indica il risultato potenziale a seconda del trattamento:
 $Y_i(1)$ è il risultato in caso di trattamento;
 $Y_i(0)$ è il risultato in caso di non trattamento;

Il risultato osservato per ciascun soggetto può essere scritto come:

$$Y_i = D_i Y_i(1) + (1 - D_i) Y_i(0) \quad (22)$$

Questo approccio alla causalità richiede di pensare in termini di eventi “controfattuali” .

Il problema fondamentale dell'inferenza causale

Definition 1. Effetto Causale

Per un soggetto i , il trattamento D_i ha un effetto causale sul risultato Y_i se l'evento $D_i = 1$ invece dell'evento $D_i = 0$ implica $Y_i = Y_i(1)$ invece di $Y_i = Y_i(0)$. In questo caso l'effetto di D_i su Y_i è

$$\Delta_i = Y_i(1) - Y_i(0)$$

L'identificazione e la misurazione è evidentemente impossibile dal punto di vista logico per il motivo seguente.

Proposition 1. Il Problema Fondamentale dell'Inferenza Causale.

È impossibile osservare per uno stesso soggetto i i valore $D_i = 1$ e $D_i = 0$ così come i valore $Y_i(1)$ e $Y_i(0)$; quindi, è impossibile osservare e misurare l'effetto di D su Y per il soggetto i (Holland, 1986).

Un altro modo per esprimere questo problema è dire che non possiamo inferire l'effetto del trattamento perchè non abbiamo l'evidenza *controfattuale* di quello che sarebbe accaduto in assenza di trattamento.

La soluzione suggerita dall'analisi statistica

L'analisi statistica risolve il problema cercando di identificare l'effetto causale medio nella popolazione o in sottogruppi particolarmente rilevanti.

L'effetto medio del trattamento nella popolazione (ATE):

$$\begin{aligned} E\{\Delta_i\} &= E\{Y_i(1) - Y_i(0)\} \\ &= E\{Y_i(1)\} - E\{Y_i(0)\} \end{aligned} \quad (23)$$

L'effetto medio del trattamento sui trattati(ATT):

$$\begin{aligned} E\{\Delta_i \mid D_i = 1\} &= E\{Y_i(1) - Y_i(0) \mid D_i = 1\} \\ &= E\{Y_i(1) \mid D_i = 1\} - E\{Y_i(0) \mid D_i = 1\} \end{aligned} \quad (24)$$

Sotto quali ipotesi questi effetti sono interessanti per un economista o per un medico?

Questo approccio costituisce un passo avanti per la soluzione del Problema Fondamentale dell'Inferenza Causale?

La semplice comparazione tra trattati e non trattati

In generale il confronto tra il risultato dei trattati e dei non trattati produce una stima distorta del ATT:

$$\begin{aligned} E\{Y_i \mid D_i = 1\} - E\{Y_i \mid D_i = 0\} & \quad (25) \\ &= E\{Y_i(1) \mid D_i = 1\} - E\{Y_i(0) \mid D_i = 1\} \\ & \quad + E\{Y_i(0) \mid D_i = 1\} - E\{Y_i(0) \mid D_i = 0\} \\ &= \tau + E\{Y_i(0) \mid D_i = 1\} - E\{Y_i(0) \mid D_i = 0\} \end{aligned}$$

dove $\tau = E\{\Delta_i \mid D_i = 1\}$ è ATT.

La differenza tra il lato sinistro (che possiamo stimare) e τ è il cosiddetto *sample selection bias* (errore di autoselezione del campione) dato dalla differenza tra il risultato dei trattati e dei non trattati (ossi i “controlli”) nella situazione di non trattamento (ad esempio alla *baseline* prima del trattamento).

Questo tipo di confronto non ci dice nulla quando i trattati e i non trattati sono diversi già prima del trattamento, cosa che normalmente accade quando il trattamento è scelto dai soggetti, o comunque ha luogo in modo non casuale.

Esperimenti causali controllati

Considera due campioni casuali C e T estratti dalla popolazione Poichè per costruzione questi campioni sono statisticamente identici alla intera popolazione possiamo scrivere:

$$E\{Y_i(0)|i \in C\} = E\{Y_i(0)|i \in T\} = E\{Y_i(0)\} \quad (26)$$

e

$$E\{Y_i(1)|i \in C\} = E\{Y_i(1)|i \in T\} = E\{Y_i(1)\}. \quad (27)$$

Sostituendo 26 e 27 in 23 è immediato ottenere:

$$\begin{aligned} E\{\Delta_i\} &\equiv E\{Y_i(1)\} - E\{Y_i(0)\} \\ &= E\{Y_i(1)|i \in T\} - E\{Y_i(0)|i \in C\}. \end{aligned} \quad (28)$$

La randomizzazione risolve il Problema Fondamentale dell'Inferenza Causale perchè utilizza i *controlli* C come immagine di quello che sarebbe accaduto ai *trattati* T nella situazione controfattuale di non trattamento, e viceversa.

Un esperimento controllato per il rendimento dell'istruzione

Consideriamo il modello teorico studiato nella sezione 2.2.1 e supponiamo di estrarre dalla popolazione generata dal DGP i due campioni casuali C e T .

Supponiamo inoltre di offrire agli individui del campione T una borsa di studio che li induca a stare a scuola più a lungo. Questo implica per loro un riduzione del costo marginale r_j (Vedi la Figura 20.3 BL).

Per semplificare l'analisi senza perdita di generalità, assumiamo che la borsa di studio sia strutturata in modo che ciascun trattato vada a scuola per un stesso tempo aggiuntivo $\Delta S = 1$ (ad esempio un anno, o il triennio universitario base):

$$\Delta S_g = \Delta S = 1 \quad \forall g. \quad (29)$$

Grazie alla *randomizzazione* i trattati sono una immagine controfattuale dei controlli e viceversa: quindi adattando l'equazione 28 otteniamo:

$$E(y_i|i \in T) - E(y_i|i \in C) = [(P_{LH} + P_{HH})r_H + (P_{LL} + P_{HL})r_L]\Delta S = \bar{r} = \bar{\beta} \quad (30)$$

Le conseguenze della impossibilità di effettuare esperimenti

Lalonde (1986) da una descrizione efficace e provocatoria degli errori che un ricercatore può fare utilizzando dati “osservati” invece che dati ottenuti attraverso esperimenti controllati.

D'altro canto, la sperimentazione controllata non è sempre possibile a causa di:

- preoccupazioni di tipo etico;
- difficoltà tecniche;
- validità esterna e replicabilità.

Cosa può fare il ricercatore quando un esperimento casuale controllato non è possibile?

- tecniche di regressione multivariata e *matching* tra trattati e controlli;
- tecniche basate su esperimenti non controllati (“naturali”).

2.2.6 Regressione multivariata e Matching

Consideriamo di nuovo la regressione 19: $W_i = \alpha + \beta S_i + \epsilon_i$: Notate che:

$$\begin{aligned}\hat{\beta} &= \frac{\widehat{Cov}(W_i, S_i)}{\widehat{Var}(S_i)} \\ &= \frac{\widehat{Cov}(\alpha + \beta S_i + \epsilon_i, S_i)}{\widehat{Var}(S_i)} \\ &= \beta + \frac{\widehat{Cov}(\epsilon_i, S_i)}{\widehat{Var}(S_i)}\end{aligned}\tag{31}$$

In un esperimento randomizzato controllato $\widehat{Cov}(\epsilon_i, S_i) \xrightarrow{p} Cov(\epsilon_i, S_i) = 0$, perchè il trattamento S_i è per costruzione non correlato con alcuna caratteristica dei soggetti. In questo caso si ha che

$$\hat{\beta} \xrightarrow{p} \beta\tag{32}$$

E grazie alla randomizzazione la stima con il metodo dei minimi quadrati identifica l'effetto causale del trattamento.

Per convincervene, considerate il caso semplice in cui S_i sia una variabile *dummy* (indicatore binario):

- $S_i = 1$ se il soggetto i ha la laurea (o un titolo più elevato);
- $S_i = 0$ se il soggetto i ha un titolo inferiore alla laurea.

Notate che:

$$\begin{aligned} E(W_i | S_i = 1) &= \alpha + \beta \\ E(W_i | S_i = 0) &= \alpha \end{aligned} \tag{33}$$

poichè $E(\epsilon_i | S_i) = 0$ grazie alla randomizzazione e quindi

$$E(W_i | S_i = 1) - E(W_i | S_i = 0) = \beta = \frac{\text{Cov}(W_i, S_i)}{\text{Var}(S_i)} \tag{34}$$

Se i soggetti provengono da un esperimento controllato in cui chi prosegue gli studi fino alla laurea è scelto a caso, questa formula corrisponde (a parte la notazione) alla equazione 30.

Notate inoltre che il coefficiente di una regressione su una variabile binaria è pari alla differenza tra le medie della variabile dipendente nei due sottocampioni definiti dalla variabile binaria (ciò è vero sempre anche senza randomizzazione).

Il metodo dei minimi quadrati con dati non sperimentali

Consideriamo il significato della termine ϵ_i nella equazione 19:

$$W_i = \alpha + \beta S_i + \epsilon_i$$

ϵ_i denota tutti gli altri fattori che determinano W_i , diversi da S_i .

In un esperimento randomizzato per costruzioni questi fattori sono in media uguali in tutti i soggetti indipendentemente dal loro S_i .

Facendo riferimento al [Capitolo 19](#) BL, consideriamo uno di questi fattori: gli anni di esperienza lavorativa che denotiamo con X_i . Possiamo riscrivere la 19 nella forma canonica della cosiddetta Equazione di *Mincer*

$$W_i = \alpha + \beta S_i + \lambda X_i + \nu_i \quad (35)$$

dove ν_i denota le altre determinanti del reddito.

L'accumulazione di capitale umano sul posto di lavoro (*on the job training*) suggerisce che λ debba essere positivo.

Supponiamo di non avere dati sperimentali e che al crescere del livello di istruzione i soggetti abbiano un numero minore di anni di esperienza lavorativa, ad esempio perchè più giovani:

$$E(X_i|S_i) = \phi + \rho S_i \quad (36)$$

Nota $\rho \neq 0$ perchè i dati non sono sperimentali. Immaginiamo di non conoscere gli anni di esperienza lavorativa e quindi di poter stimare solo l'equazione 19:

$$W_i = \alpha + \beta S_i + \epsilon_i$$

In questo caso

$$\begin{aligned} E(W_i|S_i) &= \alpha + \beta S_i + \lambda E(X_i|S_i) + E(\nu_i|S_i) \\ &= (\alpha + \lambda\phi) + (\beta + \lambda\rho)S_i \end{aligned} \quad (37)$$

E quindi la stima è distorta

$$\hat{\beta} \xrightarrow{p} (\beta + \lambda\rho) \quad (38)$$

In particolare se $\lambda > 0$ e $\rho < 0$ l'errore è negativo perchè l'effetto positivo vero della scuola è ridotto dall'effetto negativo del possedere una minore esperienza lavorativa.

Il significato delle variabili di controllo in una regressione

In assenza di un esperimento, se potessimo aggiungere ad una regressione tutti i possibili “fattori di confondimento” riusciremmo a identificare l’effetto causale che ci interessa.

Riprendiamo i dati della SHIW e il file [lezione1.do](#) per verificare come cambia la stima dell’effetto di un anno di istruzione controllando per

- esperienza lavorativa
- sesso
- area geografica
- numero di figli

Controllare per queste variabili è equivalente a confrontare due campioni di soggetti con uguale esperienza, sesso, area geografica e numero di figli (*matching*), ma con scolarità diversa, e vedere chi guadagna di più.

Notate che il sesso è (a meno di casi estremi) non modificabile: è rilevante questa osservazione per l'interpretazione delle stime?

Se ci interessa l'effetto globale della scolarità incluso anche dell'effetto che passa per le scelte lavorative, di migrazione e di fertilità, allora dobbiamo controllare solo per il sesso

Idealmente vorremmo controllare per tutte le caratteristiche precedenti al trattamento, ossia confrontare persone identiche al momento del trattamento (*matching*).

Tuttavia è molto improbabile che il ricercatore abbia dati sufficienti a controllare per tutti i casi di confondimento.

Cosa può fare allora il ricercatore?

2.2.7 Il ricorso a esperimenti “naturali”

Consideriamo di nuovo l'equazione base 19, ipotizzando che l'unico fattore di confondimento sia l'abilità genetica A_i che non possiamo osservare.

$$\begin{aligned}W_i &= \alpha + \beta S_i + \epsilon \\ &= \alpha + \beta S_i + \gamma A_i + u_i\end{aligned}\tag{39}$$

È ragionevole ipotizzare che l'abilità genetica A_i non sia completamente misurabile e di non poter confrontare l'effetto della scuola a parità di abilità.

Riusciamo ugualmente ad identificare e stimare β se esiste una variabile Z_i (detta “strumentale”) tale che

- $\text{Cov}(Z_i, S_i) \neq 0$
- $\text{Cov}(Z_i, \epsilon_i) = 0$

Ossia,

- abbia un effetto causale sugli anni di istruzione,
- non sia correlata con l'errore della regressione (ossia con l'abilità) e abbia quindi un effetto sul reddito soltanto attraverso gli anni di istruzione.

Se questa variabile esiste, lo stimatore

$$\begin{aligned}\hat{\beta}_{IV} &= \frac{\widehat{Cov}(W_i, Z_i)}{\widehat{Cov}(S_i, Z_i)} \\ &= \frac{\widehat{Cov}(\alpha + \beta S_i + \epsilon_i, Z_i)}{\widehat{Cov}(S_i, Z_i)} = \beta + \frac{\widehat{Cov}(\epsilon_i, Z_i)}{\widehat{Cov}(S_i, Z_i)} \xrightarrow{p} \beta\end{aligned}\quad (40)$$

è consistente. Concettualmente, una variabile Z_i con queste caratteristiche produce in modo “incontrollato” gli stessi effetti di un esperimento “controllato”:

- assegna in modo casuale i soggetti alle diverse dosi di trattamento;
- consente di comparare i risultati di soggetti che hanno dosi di trattamento diverse ma sono “uguali” rispetto al resto;

Esempi (vedi anche la dibattito riassunto in questo [articolo](#)):

- la *VietNam War Draft* per l'effetto del servizio militare in USA;
- il trimestre di nascita per l'effetto della scolarità;
- la distanza dall'università più vicina per l'effetto della scolarità.

Una stima IV per il nostro modello teorico

Torniamo ancora al modello teorico studiato nella sezione [2.2.1](#) e consideriamo uno strumento Z_i tale che:

- $Z_i = 1$ se il soggetto i vive vicino ad una università.
- $Z_i = 0$ se il soggetto i vive lontano da una università.

Chi vive vicino ad una università a costi inferiori di accesso all'istruzione e quindi sceglie più anni di istruzione a parità di altre caratteristiche:

$$E(S_i|Z_i = 1) > E(S_i|Z_i = 0).$$

Dobbiamo anche assumere che vivere vicino ad una università

- non offra accesso a salari diversi per data scelta universitaria
- non sia correlato con l'abilità genetica.

Sotto queste ipotesi (con $\Delta S = 1$) abbiamo la stessa situazione dell'esperimento controllato di pagina [72](#) e rappresentato graficamente nella [Figura 20.3](#) BL.

$$\hat{\beta}_{IV} = E(y_i|Z_i = 1) - E(y_i|Z_i = 0) = [(P_{LH} + P_{HH})r_H + (P_{LL} + P_{HL})r_L]\Delta S = \bar{r} \quad (41)$$

Cosa accade se i soggetti non “obbediscono” al trattamento?

È possibile che parte dei soggetti non obbedisca all'assegnazione ricevuta per il trattamento. Ad esempio:

- per i soggetti ricchi, i costi di accesso alla scuola sono irrilevanti, e quindi vivere vicino o lontano non influenza la scelta universitaria
- per i soggetti poveri, invece, la vicinanza è rilevante

In questo caso l'aumento di scolarità indotto dallo strumento Z nei 4 gruppi di popolazione g è

- $\Delta S_{g|Z} = 0$ per i gruppi $g = HL$ e $g = LL$
- $\Delta S_{g|Z} > 0$ per i gruppi $g = HH$ e $g = LH$

In questo caso l'esperimento (e lo stimatore IV) consentono di stimare soltanto l'effetto per la parte della popolazione che risponde alla assegnazione

$$\hat{\beta}_{IV} \xrightarrow{p} r_H \quad (42)$$

come illustrato [Figura 20.5](#) e nella [Tabella 20.1](#) BL.

Il problema fondamentale della stima IV

La stima IV risolve il problema della identificazione di un effetto causale in assenza di un esperimento controllato.

Richiede però un esperimento “naturale” convincente, ossia un situazione in cui

- il trattamento sia influenzato da una variabile esogena
- questa variabile esogena non deve avere effetti diretti sul risultato e sulle variabili contenute nel termini di errore.
- quindi questa variabile deve svolgere lo stesso ruolo dell'assegnazione casuale al trattamento in un esperimento controllato

Non è facile trovare queste variabili ma se non siamo convinti di averle trovate, è meglio astenersi dal fare affermazioni di tipo causale utilizzando questo metodo di stima.

2.3 Quanto lavoro viene offerto sul mercato e a quale salario

Nelle sezioni precedenti abbiamo analizzato come si determinano le caratteristiche della forza lavoro *potenziale* di un sistema economico.

Ci chiediamo adesso (sulla rilevanza di queste domande vedi ad es. [Giavazzi 2007](#)):

- se e a quali condizioni questa forza lavoro *potenziale* decide di offrirsi effettivamente sul mercato;
- posto che un individuo decida di lavorare, quanto lavoro offrirà a seconda del salario di mercato.

Un chiarimento sui termini “offerta” e “domanda” di lavoro in economia

Definizioni e statistiche descrittive ([grafici](#) e [tabelle](#)) riguardo a:

- popolazione attiva e popolazione non in età lavorativa;
- persone “nella forza lavoro” e “non nella forza lavoro” ;
- occupati, disoccupati;
- tassi di partecipazione, disoccupazione e occupazione.

2.3.1 Un modello dell'offerta di lavoro individuale

Definiamo:

- T = quantità domandata di tempo libero;
- C = quantità domandata di beni e servizi (consumo) ;
- P = prezzo di C , che possiamo normalizzare a 1;
- $U(C, T)$ = funzione di utilità dell'individuo;
- T_{\max} = tempo totale disponibile per lavoro e tempo libero;
- $L = T_{\max} - T$ quantità di lavoro offerta;
- W = salario per unità di tempo;
- X = altri redditi (ad esempio reddito del partner, da capitale o trasferimenti);
- W^* = salario di riserva.

Il problema di ottimo del consumatore

L'individuo risolve il seguente problema per scegliere la quantità ottimale di consumo e tempo libero (quindi anche di lavoro):

$$\max_{C,T} U(C, T) \quad (43)$$

sotto il vincolo

$$PC + WT \leq WT_{\max} + X \quad (44)$$

Notate che il salario è il “prezzo” del tempo libero, oltre ad essere una determinante del reddito disponibile.

Seguiamo la soluzione grafica nel [Capitolo 3](#) del BL:

- la scelta ottimale ([Figura 3.5](#));
- la determinazione del salario di riserva ([Figura 3.5](#));
- l'effetto di reddito e l'effetto di sostituzione ([Figure 3.6 - 3.8](#));
- la curva di offerta di lavoro ([Figura 3.9](#));
- la possibilità di una curva di lavoro inclinata negativamente ([Figure 3.7 e 3.9](#)).

2.3.2 Effetto dei trasferimenti sulla partecipazione al lavoro

Come varia l'offerta di lavoro individuale al variare dei redditi non da lavoro?

La [Figura 3.10](#) BL mostra gli effetti di un sussidio ai disoccupati (*welfare program*), che cessa di essere erogato se la persona trova un lavoro.

- Un sussidio di questo tipo può disincentivare la partecipazione al lavoro.

Esercizio: Come disegnereste, in questo modello, un sistema di trasferimenti che non disincentivi la partecipazione al lavoro?

Comparete l'effetto di un sussidio come quello descritto nella [Figura 3.10](#), con aumenti del reddito non da lavoro prodotti da

- X = reddito da capitale;
- X = reddito degli altri familiari.

Perché gli effetti sono differenti?

2.3.3 Offerta di lavoro, tassazione del reddito e benessere sociale

Per finanziare i servizi offerti ai cittadini, lo stato deve imporre delle tasse, che in parte gravano sui redditi da lavoro.

La tassazione dei redditi ha un effetto distorsivo sull'offerta di lavoro che produce una perdita di benessere sociale.

I modelli di offerta di lavoro vengono utilizzati dagli economisti per disegnare i sistemi di imposizione fiscale in modo da minimizzare le distorsioni e la perdita di benessere sociale.

Discutiamo una proposta di [Alesina e Ichino](#) per la quale vedi anche [l'articolo sul Sole24ore, 27 marzo 2007](#).

Secondo questa proposta, lo stato dovrebbe ridurre la tassa sul reddito delle donne e aumentare quella degli uomini.

Per valutare questa proposta dobbiamo prima capire perchè una tassa sul reddito può essere distorsiva.

Perchè una tassa sul reddito può essere distorsiva

Immaginiamo una popolazione di lavoratori maschi (m) e femmine (f) che ricevano un salario pari alla loro produttività marginale ipotizzata uguale per tutti e pari a w .

Ipotizziamo che l'offerta di lavoro delle donne L_f (e analogamente quella degli uomini) sia una funzione crescente del salario al netto delle tasse, con elasticità costante:

$$L_f = s[(1 - t)w]^{\sigma_f} \quad (45)$$

dove:

- t è l'aliquota della tassa sul reddito;
- $\sigma_f = \frac{\partial L_f}{\partial Z} \frac{Z}{L_f} < 1$ è l'elasticità costante e $Z = (1 - t)w$.
- s è un fattore di scala

Esercizio: disegnate la domanda e l'offerta di lavoro in un mercato con queste caratteristiche.

La tassa ottimale e la perdita di benessere sociale

Partendo da una situazione in cui $t = 0$, l'imposizione di una tassa $t > 0$ riduce l'offerta di lavoro e produce una perdita di benessere sociale pari a:

$$C_f = w(L_n - L_f) - \int_{L_f}^{L_n} \left(\frac{1}{s}L\right)^{\frac{1}{\sigma_f}} dL \quad (46)$$

Esercizio: rappresentate graficamente il “triangolo della perdita sociale” indotto dalla tassa sul reddito (vedi [figura](#)).

La derivata di C_f rispetto a t è:

$$\frac{dC_f}{dt} = -wt \frac{dL_f}{dt} = s\sigma_f w^{\sigma_f+1} (1-t)^{\sigma_f-1} t > 0 \quad (47)$$

Ma quel che conta ai fini della proposta è che la perdita sociale cresce al crescere della elasticità:

$$\frac{dC_f}{dt d\sigma_f} = \frac{1}{1-t} sw [w(1-t)]^{\sigma_f} \{1 + \log[(1-t)w]\} > 0$$

(essendo ragionevole assumere che $w(1-t) > 1$).

Distorsione fiscale ed elasticità dell'offerta di donne e uomini

Le stime esistenti indicano che l'elasticità dell'offerta di lavoro femminile è molto maggiore di quella maschile per ragioni non ancora completamente chiarite:

- motivi biologici
- motivi culturali
- discriminazione

Quali che siano i motivi, se l'offerta di lavoro femminile è più elastica di quella maschile ($\sigma_f > \sigma_m$), una stessa aliquota fiscale t_0 produce effetti più distorsivi nel mercato delle donne che in quello degli uomini:

$$L_{f0} < L_{m0} < L_n. \quad (48)$$

$$C_{f0} > C_{m0} > 0. \quad (49)$$

La distribuzione del reddito, più concentrata verso il basso per le donne, è un secondo motivo per cui una tassazione non differenziata genera maggiori distorsioni per le donne (per questo argomento più complesso vedi [l'articolo](#)).

Tassa ottimale ed elasticità dell'offerta di donne e uomini

Sia $T_0 = t_0(wL_{f0} + wL_{m0})$ il gettito fiscale con una singola aliquota t_0 .

Per avere un sistema fiscale ottimale a parità di gettito il governo dovrebbe imporre tasse differenziate t_f e t_m tali da minimizzare la somma delle perdite sociali nei due mercati:

$$C = C_f + C_m \quad \text{s.v.} \quad t_f w L_f + t_m w L_m \geq T_0 \quad (50)$$

Vedi [figura](#).

La soluzione di questo problema suggerisce che le tasse ottimali debbano soddisfare questa relazione:

$$t_f = \frac{\theta}{\sigma_f + \theta} < t_0 < \frac{\theta}{\sigma_m + \theta} = t_m \quad (51)$$

dove $\theta = \frac{\lambda}{1+\lambda}$ e λ è il moltiplicatore di Lagrange del problema di ottimo.

L'[articolo](#) dimostra che la proposta rimane valida anche considerando le elasticità incrociate in un modello unitario familiare di offerta (§ 4 del [capitolo 3](#) BL).

Table 14: Optimal gender based taxation in Italy, Norway and USA

		Italy	Norway	USA
<i>Assuming zero cross elasticities:</i>				
Female elasticity ^a	σ_f	0.66	0.52	0.40
Male elasticity ^a	σ_m	0.12	0.39	0.10
Maximum optimal gender tax ratio ^b	$\max\left(\frac{t_f}{t_m}\right)$	0.67	0.91	0.79
Optimal revenues-constant tax rate for females ^c	t_f	0.27	0.49	0.38
Optimal revenues-constant tax rate for males ^c	t_m	0.68	0.56	0.71
<i>Assuming non-zero cross-elasticities;</i>				
Female cross-elasticity ^a	γ_{fm}	-0.49	-0.42	-0.19
Male cross-elasticity ^a	γ_{mf}	-0.12	-0.23	0.03
Gender labor force participation ratio ^d	$\frac{l_f}{l_m}$	0.64	0.92	0.84
Optimal revenues-constant tax rate for females ^e	\bar{t}_f	0.35	0.57	0.35
Optimal revenues-constant tax rate for males ^e	\bar{t}_m	0.77	0.70	0.91

Figure 3: Hazard rates of the earnings distribution in Italy, Norway and USA

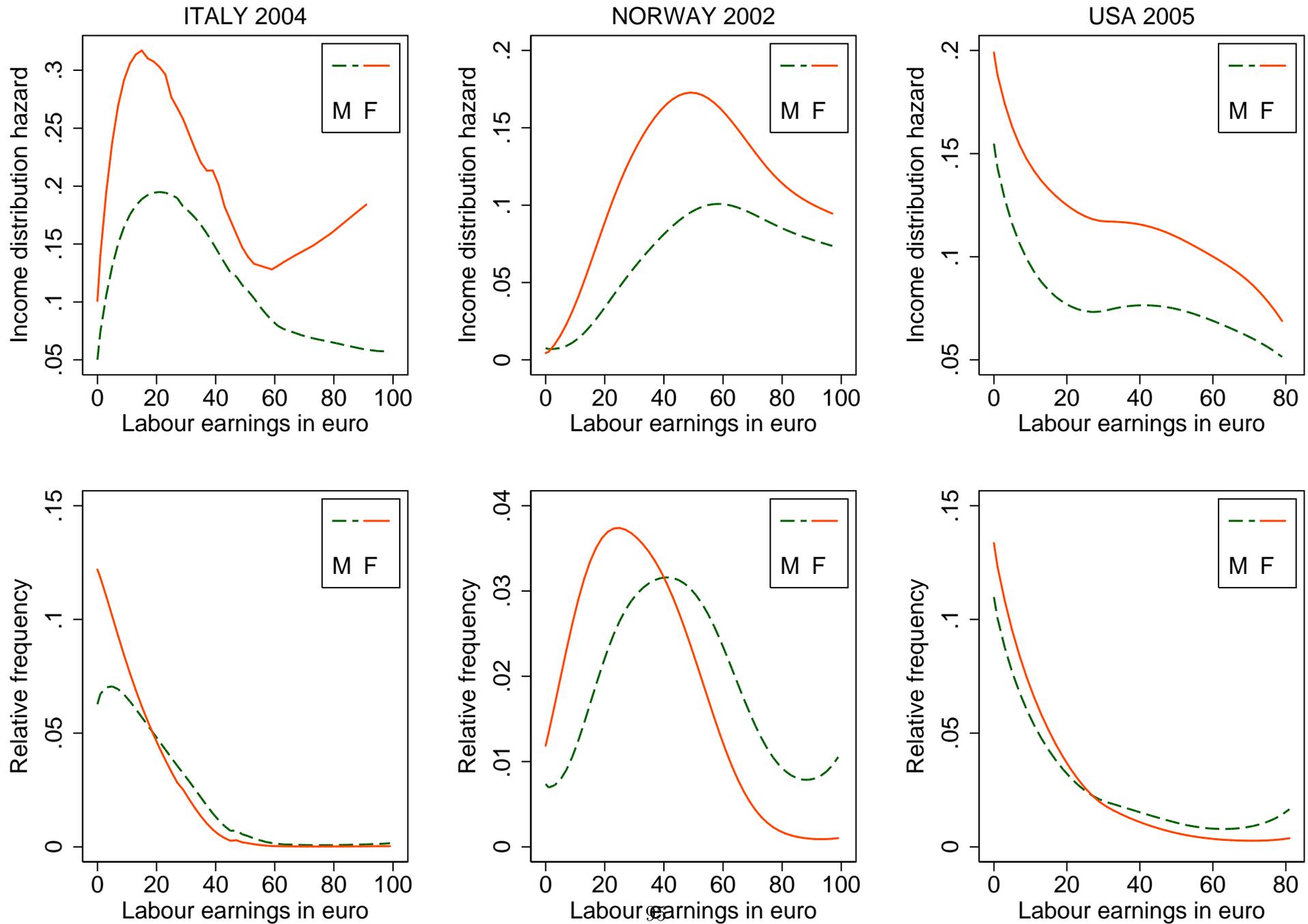


Table 15: Weighted averages of the hazard rates of the earnings distribution in Italy, Norway and USA

	Italy	Norway	USA
Females	.25	.12	.16
Males	.17	.07	.11

Effetti collaterali della tassazione differenziata per genere

Questa proposta consente al governo di ridurre la pressione fiscale media (e quindi la distorsione) a parità di gettito fiscale: di fatto riduce il “cuneo fiscale”.

La tassazione differenziata produce anche altre effetti collaterali desiderabili:

- aumenta l'occupazione femminile, un obiettivo auspicato da molti: in Italia nel 2006 lavorava il 46.3% delle donne mentre l'Agenda di Lisbona indica l'obiettivo del 60% nel 2010;
- aumenta l'equità del sistema perchè compensa le donne per i costi biologici e sociali che ostacolano il loro lavoro;
- combatte la discriminazione in modo meno distorsivo che altre misure quali “quote rosa”, “daddy month”, “affirmative action” etc.;
- riequilibra il potere contrattuale tra i coniugi e il tempo da essi speso in casa, in particolare con i figli;
- aumenta il reddito familiare al netto delle tasse, consentendo di finanziare la cura dei figli in asilo o con “babysitters”;
- consente compensazioni “paretiane” all'interno della famiglia.

3 La domanda e l'equilibrio in un mercato del lavoro concorrenziale

Ci chiediamo adesso come si determina la domanda di lavoro di una azienda nel caso in cui:

- il lavoro sia l'unico fattore della produzione e sia omogeneo (oppure esistano altri tipi di forza lavoro e altri fattori della produzione, ma siano fissi nel breve periodo);
- il mercato del lavoro sia perfettamente concorrenziale.

Seguiamo i Capitoli 4 e 5 BL.

Combinando insieme la domanda e l'offerta di lavoro ci proponiamo di studiare le caratteristiche di un equilibrio competitivo nel mercato del lavoro.

Ci chiederemo se e per quali motivi questo equilibrio sia vicino o lontano dalla realtà, e qualora risulti lontano, perchè valga ugualmente la pena di studiarlo.

3.1 La domanda di forza lavoro omogenea

Definiamo:

- Y = quantità prodotta dall'impresa;
- L = quantità di lavoro domandata;
- \bar{K} = Quantità di altri fattori della produzione, fissi nel breve periodo;
- $Y = f(L, \bar{K})$ è la funzione di produzione (vedi la figura 4.1 BL);
- $f_L(L, \bar{K})$ = produttività marginale del lavoro (vedi la 4.1 BL);
- $\frac{Y}{L} = \frac{f(L, \bar{K})}{L}$ = produttività media del lavoro (vedi la figura 4.1 BL);
- Π = profitto;
- W = salario di una unità di forza lavoro;
- P = prezzo di una unità di prodotto.

Il problema di ottimo dell'impresa

L'impresa sceglie la quantità di lavoro che massimizza il profitto, ossia risolve il seguente problema

$$\max_L \Pi = PY - WL - R\bar{K} \quad \text{sotto il vincolo} \quad Y = f(L, \bar{K}) \quad (52)$$

La condizione del primo ordine per la massimizzazione dell'obiettivo richiede che:

$$f_L(L, \bar{K}) = \frac{W}{P} \quad (53)$$

ossia che si assuma forza lavoro fino al punto in cui la produttività marginale dell'ultima unità di lavoro eguaglia il salario reale.

Notate che questa condizione identifica il punto di tangenza tra la funzione di produzione e la curva di iso-profitto

$$Y = \frac{W}{P}L + \frac{\bar{\Pi}}{P} + \frac{R\bar{K}}{P} \quad (54)$$

Vedi la [Figura 4.2](#) BL.

La curva della domanda di lavoro al variare del salario

Se nella [Figura 4.2](#) facciamo variare il salario reale e quindi la pendenza della curva di iso-profitto, otteniamo una sequenza di punti di ottimo che descrivono la quantità ottimale di lavoro in funzione del salario reale

Formalmente questo esercizio grafico corrisponde all'inversione della condizione del primo ordine che consente di ottenere la funzione di domanda di lavoro:

$$L = f_L^{-1}\left(\frac{W}{P}, \bar{K}\right) = L\left(\frac{W}{P}, \bar{K}\right) \quad (55)$$

Vedi la [Figura 4.2](#) BL.

Notate che l'inclinazione della domanda di lavoro è

$$\frac{dL}{d\frac{W}{P}} = f_{L\frac{W}{P}}^{-1}\left(\frac{W}{P}, \bar{K}\right) = L_{\frac{W}{P}}\left(\frac{W}{P}, \bar{K}\right) < 0 \quad (56)$$

dove l'ineguaglianza finale vale se la produttività marginale del lavoro è decrescente: la domanda di lavoro diminuisce se il salario reale cresce.

Alcune precisazioni

1. Come sarebbe la domanda di lavoro se la produttività marginale fosse costante? Notate che questa è l'ipotesi che giustifica la domanda di lavoro ipotizzata per discutere gli effetti della tassazione nella Sezione [2.3.3](#).
2. Questa funzione di domanda di lavoro è stata derivata sotto l'ipotesi che:
 - esista concorrenza perfetta nel mercato del prodotto e quindi l'impresa prenda il prezzo P per dato;
 - esista concorrenza perfetta nel mercato del lavoro e quindi l'impresa prenda il salario nominale W per dato.

Gran parte del programma rimanente del corso è finalizzato a studiare cosa accade quando rimuoviamo l'ipotesi di concorrenza perfetta nel mercato del lavoro.

Qui vediamo brevemente cosa accade se rimuoviamo l'ipotesi di concorrenza perfetta nel mercato del prodotto.

Domanda di lavoro con mercato del prodotto non competitivo

La condizione del primo ordine [53](#) può anche essere espressa come uguaglianza tra costo marginale e ricavo marginale:

$$P f_L(L, \bar{K}) = W \quad (57)$$

Se la domanda per il bene prodotto dall'impresa (Y) è decrescente rispetto al prezzo, ossia il mercato del prodotto non è competitivo, il prezzo P non può essere considerato costante dall'impresa in quanto $P = P(Y)$ con $\frac{dP}{dY} < 0$.

L'uguaglianza tra ricavi e costi marginali che massimizza il profitto diventa

$$P \left(1 - \frac{1}{|\eta_{YP}|} \right) f_L(L, \bar{K}) = W \quad (58)$$

dove $\eta_{YP} = \frac{dY}{dP} \frac{P}{Y} < 0$. Vedi la [Figura 4.3](#). Esercizio: come si ricava la condizione [58](#)?

Notate che in questo caso la domanda di lavoro è decrescente anche se la produttività marginale del lavoro è costante. Perché?

3.2 L'equilibrio in un mercato del lavoro concorrenziale

L'incontro tra la domanda e l'offerta di lavoro determina il livello di occupazione e il salario di una economia.

La [Figura 5.1](#) BL descrive l'andamento di queste variabili in alcuni paesi Europei e negli USA. L'[articolo](#) di Francesco Giavazzi sul Corriere della Sera del 16 marzo 2007 compara la capacità di creare posti di lavoro in Italia e Spagna.

In questa parte del corso ci chiediamo se e in quale misura l'equilibrio descritto sotto le ipotesi di concorrenza perfetta nel mercato del lavoro può spiegare gli andamenti osservati di variabili quali:

- occupazione
- salario reale
- disoccupazione
- differenziale salariali tra lavoratori

Le ipotesi per un mercato del lavoro concorrenziale

- il numero di imprese e lavoratori è elevato e quindi gli agenti sono *price takers*;
- tutti gli agenti hanno informazioni complete e perfette sulle opportunità offerte dal mercato;
- lavoratori e imprese sono omogenei e nessuno ha una posizione dominante;
- non esistono barriere alla mobilità e all'entrata di agenti nel mercato
- non esistono vincoli istituzionali all'aggiustamento dei prezzi e delle quantità.

Sotto queste condizioni l'equilibrio concorrenziale assicura che

- valga la *legge dell'unico prezzo*, ossia lo “stesso bene” non possa avere due prezzi diversi;
- non esistano possibilità di arbitraggio non sfruttate dagli agenti;
- valga una situazione di *ottimo paretiano*, ossia non sia possibile un'allocazione delle risorse tale per cui un agente stia meglio senza che nessun'altro stia peggio.

Analisi formale dell'equilibrio concorrenziale

Il comportamento del lavoratore è descritto dalla soluzione del problema:

$$\max_{C,T} U(C, T) \quad \text{s. v.} \quad PC = W(T_{max} - T) + X = WL + X \quad (59)$$

Vedi la [Figura 5.2](#)

Il comportamento dell'impresa è descritto dalla soluzione del problema:

$$\max_L \Pi = PY - WL - R\bar{K} \quad \text{s. v.} \quad Y = f(L, \bar{K}) \quad (60)$$

Vedi la [Figura 5.3](#).

I due problemi possono essere analizzati congiuntamente, come nella [Figura 5.4](#), per descrivere studiare le caratteristiche dell'equilibrio:

$$L^s \left(\frac{W^e}{P}, Z \right) = L^d \left(\frac{W^e}{P}, K \right) \quad (61)$$

dove Z e K sono gli insiemi di altre variabili che influenzano la domanda L^d e l'offerta L^s . (Quali sono queste variabili?)

Caratteristiche dell'equilibrio e statica comparata

Il punto di A nella [Figura 5.4](#) definisce i valori delle variabili rilevanti in equilibrio:

- il salario reale $\frac{W}{P}$, corrispondente al coefficiente angolare della retta di iso-profitto che coincide con il vincolo di bilancio,
- la quantità di forza lavoro occupata,
- la quantità di beni prodotti dalle imprese.

Il sistema è in equilibrio nel punto A perchè se viene perturbato, ha al suo interno forze che ripristinano l'equilibrio: vedi le [Figure 5.5](#) e [5.6](#).

Notate in particolare che nell'equilibrio concorrenziale:

- non esiste disoccupazione “involontaria” nel senso che le unità di forza lavoro disoccupate sono unità che preferiscono non lavorare al salario prevalente;
- l'equilibrio concorrenziale garantisce il raggiungimento di una situazione di *ottimo paretiano*.

3.3 Disoccupazione e fluttuazioni cicliche

In un equilibrio di concorrenza perfetta, la disoccupazione è solo “volontaria” e il suo livello è determinato unicamente:

- dalle preferenze dei lavoratori riguardo a quanto tempo dedicare al lavoro per produrre beni e a quanto tempo dedicare al riposo (disoccupazione);
- da shocks tecnologici che influenzano la produttività del lavoro e quindi la vantaggiosità di lavorare piuttosto che dedicarsi al tempo libero.

La [Figura 5.9](#) BL spiega la cosiddetta “Teoria reale delle fluttuazioni cicliche” (*Real Business Cycle Theory*).

In un contesto dinamico multiperiodale, un shock tecnologico positivo rende vantaggioso “lavorare” di più oggi e “riposare” di più domani. Viceversa per uno shock negativo.

Questo genera le fluttuazioni cicliche, che sono comunque “fisiologiche”, non “patologiche”, e quindi non richiedono interventi correttivi di politica economica.

Il dibattito sulla disoccupazione volontaria e involontaria

Il dibattito macroeconomico sulla natura “volontaria” o “involontaria” della disoccupazione ha origini antiche nella letteratura economica.

Fu in particolare Keynes, che osservava gli effetti della Grande Depressione del 1929, a suggerire che possano esistere anche situazioni di disoccupazione “involontaria” che richiedono correttivi di politica economica

L'ipotesi Keynesiana era centrata sull'idea che la domanda aggregata fosse insufficiente, e che quindi fosse necessario incrementare la spesa pubblica per stimolare la domanda e l'occupazione

Il dibattito recente ha però messo in luce che il disequilibrio Keynesiano richiede rigidità nella determinazione dei prezzi e dei salari o altre imperfezioni che impediscano l'aggiustamento (vedi le Figure [5.5](#) e [5.6](#) BL senza aggiustamento).

Gran parte delle prossime lezioni sarà dedicata proprio allo studio di queste imperfezioni che impediscono l'esplicarsi delle forze positive della concorrenza.

3.4 Differenziali salariali in un mercato concorrenziale

L'equilibrio concorrenziale che abbiamo descritto è anche in grado di spiegare l'esistenza di differenziali salariali, rimuovendo l'ipotesi che la forza lavoro sia di un unico tipo omogeneo.

Abbiamo infatti visto nella Sezione [2.2](#) che l'eterogeneità delle scelte di istruzione genera lavoratori che si presentano sul mercato con titoli di studio diversi e, presumibilmente, differenti abilità.

Al tempo stesso le imprese e i posti di lavoro da esse creati non sono omogenei.

Tuttavia, nella misura in cui queste disomogeneità siano note e osservabili, l'equilibrio concorrenziale può essere esteso alla considerazione di queste differenze nelle caratteristiche dei lavoratori e dei posti di lavoro.

Un punto fermo dell'equilibrio concorrenziale, anche in questa versione estesa, rimane comunque il fatto che ciascun lavoratore deve ricevere un salario pari alla sua produttività marginale.

Eterogeneità dei lavoratori e equilibrio concorrenziale

Considerate nuovamente la [Figura 5.4](#) BL. Possiamo introdurre eterogeneità nel modello rappresentato in questa figura in modo parzialmente analogo a quanto fatto nella [Sezione 2.2](#) per le scelte di istruzione.

Ad esempio, assumiamo che esistano due funzione di produzione diverse, una per laureati e una per non laureati.

- Come le disegnereste per rappresentare il fatto che la produttività marginale dei laureati è maggiore a parità di unità di lavoro?
- Considerate due mappe di curve di indifferenza lineari per i due tipi di lavoratori: Saranno più inclinate quelle dei laureati o quelle dei non laureati?
- Da cosa deduciamo che il salario reale è uguale alla produttività marginale per ciascun tipo di lavoratore?
- Che differenze ci sono nei tassi di occupazione e nei salari reali a seconda del livello di istruzione?
- Che cosa accade se uno shock aumenta il differenziale di produttività tra laureati e non laureati?

Il dibattito sui differenziali salariali in USA e nel mondo

Durante gli anni '80 i differenziali salariali tra lavoratori sono aumentati enormemente in USA dopo essersi compressi negli anni '70. Vedi ad esempio le Figure I e II in [Katz and Murphy \(1992\)](#).

Il fenomeno è meno evidente in Europa (vedi ad esempio Freeman e Katz, 1994).

[Katz and Murphy \(1992\)](#) mostrano come un semplicissimo modello concorrenziale competitivo possa spiegare questi fenomeni in termini di domanda e offerta relativa di lavoratori *skilled* e *unskilled*.

Questi autori stimano i trends dell'occupazione relativa e del salario relativo in vari sottoperiodi (vedi la loro Figura III) concludendo che:

- i trends degli anni '70 sono compatibili solo con aumenti della offerta relativa di lavoratori *skilled*;
- i trends degli anni '80 sono compatibili solo con aumenti della domanda relativa di lavoratori *skilled*.

Sapreste ricostruire il loro ragionamento modificando opportunamente la [Figura 5.8](#)?

Il dibattito sui differenziali salariali in USA e nel mondo

Questi fenomeni hanno suscitato un dibattito acceso su tre possibili spiegazioni:

- il progresso tecnico e in particolare la rivoluzione informatica
- il commercio internazionale e la concorrenza dei paesi in via di sviluppo
- l'evoluzione delle istituzioni e delle preferenze per l'uguaglianza nei vari paesi

Il dibattito è parzialmente ancora aperto, ma l'opinione prevalente è che:

- l'espansione dell'accesso all'istruzione negli anni '60 abbia generato un eccesso offerta di lavoratori *skilled* negli anni '70;
- la rivoluzione dei computers sia la causa dell'aumento della domanda relativa di lavoratori *skilled* negli anni '80;
- La stessa rivoluzione sta oggi causando una “polarizzazione” del mercato del lavoro con una contrazione della classe sociale intermedia ([Autor et al. \(2006\)](#)).

La spiegazione basata sul ruolo del commercio internazionale è discussa nel [Capitolo 15](#) BL per chi fosse interessato.

I differenziali compensativi per posti di lavoro eterogenei

Considerate un lavoratore che:

- debba scegliere tra due posti di lavoro A e B nei quali è ugualmente produttivo;
- ma A è in città dove la casa costa molto e B è in campagna dove la casa costa poco.

L'intuizione suggerisce che il lavoratore voglia un salario maggiore nel posto A piuttosto che nel posto B .

Anche questo tipo di differenziale salariale che, notate, ha luogo tra lavoratori identici, può essere spiegato da un modello competitivo.

L'origine del differenziale sta nel fatto che il posto di lavoro *meno gradevole* deve pagare un salario più alto *a parità di caratteristiche del lavoratore*.

Come usereste la [Figura 5.4](#) per razionalizzare questa situazione?

3.5 Equilibrio concorrenziale e mercati del lavoro reali

Per continuare la nostra analisi è utile riflettere sulla contrapposizione tra due posizioni antagoniste nel dibattito economico:

1. quella di chi ritiene che la realtà sia vicina ad un equilibrio concorrenziale, e che quindi, essendo questo equilibrio ottimale, sia meglio lasciare che il mercato operi senza interferenze (gli economisti *di destra*);
2. quella di chi ritiene invece che la realtà sia lontana da un equilibrio concorrenziale per numerose ragioni, e che quindi sia necessario intervenire per correggere le imperfezioni del mercato (gli economisti *di sinistra*).

Le implicazioni di queste due posizioni sono radicalmente diverse, ma la scelta tra esse non dovrebbe essere *ideologica*. Dovrebbe dipendere dalle risposte a:

- Esistono imperfezioni tali per cui la realtà è sensibilmente diversa da quella descritta da un equilibrio concorrenziale?
- Se queste imperfezioni esistono, che cosa le determina e possono essere corrette o eliminate?

Le lezioni che seguono si propongono di rispondere a queste domande.

4 Istituzioni che interferiscono con il funzionamento del mercato concorrenziale

In questa parte del corso studieremo due istituzioni che interferiscono con il funzionamento di un mercato del lavoro concorrenziale:

- il sindacato;
- i regimi di protezione dell'impiego (*Employment Protection legislation*).

Inizieremo analizzando i meccanismi di funzionamento di queste istituzioni e le conseguenze per l'equilibrio del mercato.

Ci chiederemo poi i motivi dell'esistenza di queste istituzioni:

- nascono per correggere una imperfezione del mercato e quindi configurano un "second best efficient equilibrium" ?
- rispondono all'esigenza di posizionare la società nella posizione ottimale lungo il *trade-off* uguaglianza ed efficienza, posizione diversa da quella indotta dall'equilibrio concorrenziale?
- difendono interessi corporativi di una parte degli agenti con un danno per la collettività nel suo complesso?

4.1 Il sindacato

I sindacati sono associazione di lavoratori che possono avere due tipi di obiettivi:

1. proteggere gli interessi e migliorare il benessere dei propri iscritti ;
2. proteggere gli interessi e migliorare il benessere della forza lavoro nel suo complesso.

Storicamente i sindacati Europei continentali, soprattutto affiliati al movimento socialista e cattolico, si rifanno (almeno a parole) al secondo obiettivo, mentre i sindacati di origine anglosassone sono più esplicitamente corporativi.

Il [Capitolo 8](#) BL contiene:

- una descrizione del contesto istituzionale in cui operano i sindacati;
- dati statistici sull'andamento dei tassi di sindacalizzazione (vedi la [Figura 8.1](#));
- un modello per analizzare il comportamento e le conseguenze della presenza di un sindacato in un mercato del lavoro.

Sindacato e concorrenza perfetta nel mercato del prodotto

Ricordate dall'esame di microeconomia che in un mercato del prodotto perfettamente concorrenziale le imprese hanno *extra* profitti pari a zero.

Questo perchè se in quel mercato si potessero fare profitti superiori a quelli *normali* entrerebbero altre imprese, la produzione aumenterebbe, diminuirebbe il prezzo e gli *extra* profitti si annullerebbero.

E analogamente, se i costi di una o più imprese aumentassero facendo scendere i profitti sotto il livello normale, l'impresa uscirebbe dal mercato.

Ne consegue che i sindacati avrebbero pochi margini d'azione in imprese che vendano prodotti in mercati perfettamente concorrenziali:

- se ottengono miglioramenti salariali o ambientali che comportino costi aggiuntivi per le imprese, le costringono ad uscire dal mercato;
- se non lo fanno perdono iscritti.

Riservandoci di tornare su questa problema quando ci chiederemo perchè esistono i sindacati, ora assumiamo concorrenza imperfetta nel mercato del prodotto.

4.1.1 Modelli di contrattazione sindacale

In un mercato del lavoro sindacalizzato i livelli di salario e occupazione non sono determinati dalle forze di mercato ma da qualche forma di contrattazione tra imprese e lavoratori.

La contrattazione è un processo complicato di interazione tra due parti che intendono trovare un accordo su come produrre un bene e come suddividere i vantaggi generati dalla produzione.

Gli economisti hanno proposto numerosi modelli matematici per descrivere il processo di contrattazione e soprattutto i possibili equilibri che da esso possono emergere.

Seguendo il [Capitolo 8](#) BL utilizziamo il modello di contrattazione di Nash.

Per utilizzare questo modello dobbiamo prima definire le funzioni obiettivo che guidano il sindacato e l'impresa durante la contrattazione.

La funzione obiettivo del sindacato

Se il sindacato rappresenta N lavoratori il suo obiettivo è:

$$V(W, L) = \frac{L}{N}U(W) + \left(1 - \frac{L}{N}\right)U(\bar{W}) \quad (62)$$

- $U(.)$ è l'utilità che i lavoratori derivano dal reddito;
- W è il salario (che riassume l'insieme delle condizioni di lavoro);
- L è il numero di occupati;
- \bar{W} è il reddito alternativo dei membri del sindacato se non sono occupati.

Il sindacato massimizza il reddito atteso dei suoi membri. La definizione di chi sono i membri sarà essenziale per valutare la desiderabilità sociale e l'efficienza degli effetti del sindacato.

Le curve di indifferenza nel piano $\{L, W\}$ sono definite nel modo usuale per un generico livello $V(W, L) = v$, risolvendo W in funzione di L .

Esercizio: che forma avranno in generale queste curve di indifferenza? E nel caso in cui $U(.)$ sia lineare? (Vedi la [Figura 8.2 BL](#))

La funzione obiettivo dell'impresa

L'impresa ha come obiettivo la massimizzazione del profitto come nel caso di mercato del lavoro concorrenziale, assumendo per semplicità che il lavoro sia l'unico fattore di produzione.

$$\Pi(W, L) = R(L) - WL \quad (63)$$

- $R(L)$ è il ricavo ottenuto con L lavoratori;
- WL è il costo del lavoro

Le curve di isoprofitto nel piano $\{L, W\}$ sono definite nel modo usuale per un generico livello $\Pi(W, L) = \pi$, risolvendo W in funzione di L .

Esercizio: che forma avranno in generale le curve di isoprofitto? Che cosa rappresenta il luogo dei punti di massimi di queste curve? (Vedi la [Figura 8.3](#) BL)

Modello di contrattazione di Nash

L'esito della contrattazione è la soluzione del seguente problema di ottimo

$$\max B = V(W, L)^\gamma \Pi(W, L)^{(1-\gamma)} \quad (64)$$

dove γ misura il potere contrattuale del sindacato e $1 - \gamma$ quello dell'impresa.

Una contrattazione che abbia come risultato la soluzione di questo problema, soddisfa alcuni criteri di razionalità, efficienza e desiderabilità.

Tassonomia dei modelli di contrattazione:

1. Se la funzione B nella 64 è massimizzata soltanto rispetto a W , abbiamo i modelli in cui l'impresa mantiene il “diritto a gestire l'occupazione” (“right to manage”);
 - “sindacato monopolista” se $\gamma = 1$;
 - “right to manage classico” se $0 < \gamma < 1$;
 - equilibrio concorrenziale con sindacato privo di potere se $\gamma = 0$.
2. Se la funzione B nella 64 è massimizzata rispetto a W e L abbiamo i modelli di contrattazione efficiente.

Il sindacato monopolista

Supponiamo che:

- l'impresa mantenga il diritto di stabilire il livello di occupazione L ;
- il sindacato abbia tutto il potere contrattuale ($\gamma = 1$) ai fini della scelta del solo salario W .

Il problema di ottimo diventa:

$$\max_W V(W, L) \quad \text{s. v.} \quad L = L^d(W) = R'^{-1}(W) \quad (65)$$

dove $L = L^d(W) = R'^{-1}(W)$ è la domanda di lavoro ossia il luogo delle combinazioni di W e L che garantisce la massimizzazione del profitto, ossia $R'(L) - W = 0$.

In questo caso il sindacato sceglie liberamente il salario mentre l'impresa sceglie l'occupazione lungo la domanda di lavoro, dato il salario scelto dal sindacato.

L'equilibrio è rappresentato nella [Figura 8.2](#) e corrisponde al punto di tangenza tra la domanda di lavoro e la curva di indifferenza del sindacato.

Il modello “right to manage” classico

Supponiamo che:

- l'impresa mantenga il diritto di stabilire il livello di occupazione L ;
- nessuna delle due parti abbia un potere contrattuale completo.

In questo caso il risultato della contrattazione è descritto da:

$$\max_W B = V(W, L)^\gamma \Pi(W, L)^{(1-\gamma)} \quad \text{s. v.} \quad L = L^d(W) = R'^{-1}(W) \quad (66)$$

dove $L = L^d(W) = R'^{-1}(W)$ è sempre la domanda di lavoro.

Il sindacato non può imporre il salario ottimale dal suo punto di vista, e la contrattazione determina un salario inferiore.

L'impresa continua a scegliere l'occupazione lungo la domanda di lavoro, dato il salario emerso dalla contrattazione

L'equilibrio è rappresentato nella figura [Figura 8.2](#).

Equilibrio concorrenziale con sindacato privo di potere

Supponiamo che:

- l'impresa mantenga il diritto di stabilire il livello di occupazione L ;
- l'impresa abbia un potere contrattuale completo.

In questo caso il risultato della contrattazione è descritto da:

$$\max_W B = \Pi(W, L) \quad \text{s. v.} \quad L = L^d(W) = R'^{-1}(W) \quad (67)$$

dove $L = L^d(W) = R'^{-1}(W)$ è sempre la domanda di lavoro.

Questo è un caso limite in cui il risultato della contrattazione coincide con l'equilibrio concorrenziale.

Impresa e sindacato contrattano sul salario ma l'impresa può imporre il salario minimo pari a $W = \bar{W}$.

Dato $W = \bar{W}$, l'impresa sceglie il livello di occupazione L che quindi coincide con quello concorrenziale

L'equilibrio è rappresentato nella Figura [Figura 8.2](#).

Equilibrio con contrattazione efficiente

Consideriamo la [Figura 8.3](#), ad esempio nel punto corrispondente all'equilibrio del “sindacato monopolista”.

Da questo punto le due parti possono aumentare l'occupazione e diminuire il salario raggiungendo *entrambe* livelli più alti delle loro rispettive funzioni obiettivo.

Infatti il luogo delle combinazioni “Pareto efficienti” di W ed L è dato dai punti di tangenza tra curva di indifferenza e curva di isoprofitto, detto “curva dei contratti efficienti”

Lungo la “curva dei contratti” non è possibile trovare combinazioni di W e L tali per cui entrambe le parti stiano meglio o almeno una stia meglio, e l'altra non stia peggio.

Il punto che prevale lungo la curva dipende dal potere contrattuale delle parti.

Esercizio: scrivete l'equazione che definisce la curva dei contratti.

In che senso si ha efficienza lungo la curva dei contratti?

L'equilibrio perfettamente concorrenziale è uno dei punti della curva dei contratti, ma ve ne sono altri che non coincidono con questo equilibrio e quindi non possono essere efficienti nello stesso senso dell'equilibrio perfettamente concorrenziale.

Il motivo è che la curva dei contratti è efficiente dal punto di vista delle *sole due parti* coinvolte nella contrattazione, ossia sindacato e impresa.

La curva dei contratti potrebbe non essere efficiente dal punto di vista:

- dei consumatori;
- di altri lavoratori non rappresentati dal sindacato

Le combinazioni di W e L lungo la curva dei contratti

- assicurano l'efficienza paretiana tra impresa e sindacato;
- non assicurano che la situazione dei consumatori e degli eventuali lavoratori non rappresentati sia parimenti ottimale.

Sindacato e consumatori

Supponete che

- il sindacato rappresenti tutti i lavoratori (ma veramente tutti a livello globale);
- le imprese abbiano potere di monopolio nel mercato del prodotto.

In questa situazione il sindacato svolge il ruolo di trasferire parte delle rendite (quasi) monopolistiche delle imprese ai lavoratori, e può farlo efficientemente lungo la curva dei contratti.

Però i consumatori (ossia ... i lavoratori quando non lavorano) potrebbero stare meglio nell'equilibrio perfettamente concorrenziale:

- il salario nominale sarebbe al livello minimo \bar{W} ;
- ma anche i prezzi sarebbero inferiori.

Se la competizione tra imprese riducesse le rendite monopolistiche, l'esito della contrattazione dovrebbe muoversi verso l'equilibrio perfettamente concorrenziale, senza sindacato.

4.1.2 **Sindacato e rappresentanza: la contrapposizione tra *insiders* e *outsiders***

Studiamo adesso che cosa accade quando il sindacato non rappresenta la totalità dei lavoratori.

Inizieremo studiando una economia in cui esistano due settori (o due imprese), di cui uno solo sindacalizzato; le conseguenze sono:

- emergenza di un differenziale salariale tra lavoratori sindacalizzati e non;
- i lavoratori non sindacalizzati stanno peggio che in assenza del sindacato nell'altro settore;
- si determina una perdita di benessere sociale complessivo.

Esamineremo poi che evidenza empirica esiste sull'esistenza di un differenziale salariale tra lavoratori sindacalizzati e non.

Infine esamineremo come si modifica l'equilibrio in una economia sindacalizzata al variare delle regole che caratterizzano la rappresentanza sindacale.

Sindacato in una economia a due settori

Consideriamo la [Figura 8.7](#) BL:

- gli elementi costitutivi della rappresentazione grafica dei due settori;
- l'equilibrio generale in assenza di sindacato (Punto A);
- l'equilibrio del sindacato monopolista nel settore S (Punto B)
- l'equilibrio nel settore competitivo (Punto C)
- variazioni dell'occupazione nei due settori indotte dal sindacato;
- variazioni del salario nei due settori indotte dal sindacato;
- la differenza nel valore della produttività marginale dei due settori;
- la perdita di benessere sociale (il triangolo ABC);
- implicazioni distributive: il conflitto tra lavoratori sindacalizzati e non;
- sindacati universalisti e sindacati corporativi;
- sindacato, tutela dei lavoratori nei paesi in via di sviluppo e globalizzazione.

L'effetto del sindacato sui salari

Abbiamo visto che se in un sistema economico solo una parte dei lavoratori è sindacalizzata, dovremmo osservare un premio salariale a vantaggio dei membri del sindacato

L'evidenza descrittiva sembrerebbe confermare questo risultato: [Tabella 8.2](#) BL

Sappiamo però, dall'analisi del problema della causalità, che le stime della [Tabella 8.2](#) potrebbero non cogliere un effetto causale.

Se consideriamo l'essere membri del sindacato come un "trattamento", esiste un problema di auto-selezione; il premio salariale dei lavoratori sindacalizzati:

- potrebbe riflettere il fatto che sono lavoratori "diversi dagli altri" (magari perchè operanti in aziende diverse);
- oppure potrebbe riflettere un vero effetto causale del trattamento.

Lo studio di [Di Nardo e Lee \(2004\)](#), basato su un "quasi-esperimento", suggerisce che se il premio esiste (come effetto causale), non sembra ridurre la performance delle imprese e la loro sopravvivenza.

“Insiders”, “outsiders” ed equilibrio macroeconomico

Il modello di Blanchard e Summers (1986) presentato nel [Capitolo 9](#) illustra come gli effetti macroeconomici della sindacalizzazione differiscono a seconda delle regole di appartenenza al sindacato.

Consideriamo un sistema economico in cui la domanda di mercato per l'impresa j è (in logaritmi):

$$y^j = -g(p^j - p) + (m - p) \quad (68)$$

(Nota bene: ricerca questa specificazione nel manuale di macroeconomia).

- y^j è la quantità domandata per l'impresa j ;
- p^j è il prezzo di offerta per l'impresa j ;
- p è il prezzo di mercato;
- $m - p$ è la domanda aggregata, definita come funzione dell'offerta di moneta;
- g è una costante

La domanda dell'impresa j dipende negativamente dalla differenza tra il suo prezzo e il prezzo di mercato, e positivamente dalla domanda aggregata.

La funzione di produzione e la domanda di lavoro

Ogni impresa j ha rendimenti di scala costanti e quindi:

$$y^j = l^j \quad (69)$$

dove per semplicità si assume che ogni lavoratore produca una unità di prodotto

La massimizzazione del profitto richiede che il prezzo di una unità di prodotto equagli il costo marginale, ossia:

$$p^j = w^j - e \quad (70)$$

dove e è una variabile casuale che rappresenta uno shock tecnologico.

Esercizio: identificate qualche esempio di shock tecnologico che possa essere descritto dalla variabile casuale e (nota il segno $-$); scrivete il profitto dell'impresa.

Sostituendo nella [68](#) otteniamo la domanda di lavoro dell'impresa j :

$$l^j = -g(w^j - e - p) + (m - p) \quad (71)$$

che dipende negativamente dal salario reale e positivamente dalla domanda aggregata.

La funzione obiettivo del sindacato

Supponiamo che esistano l^* lavoratori di cui l_I^j siano gli “insiders” associati all’impresa j .

Possiamo specificare in modo generale l’obiettivo occupazionale del sindacato con la seguente espressione:

$$E(l^j) = al_I^j + (1 - a)l^* \quad (72)$$

- se $a = 0$ il sindacato rappresenta la totalità dei lavoratori;
- se $a = 1$ il sindacato rappresenta solo gli “insiders”.

Il sindacato (monopolista) sceglie il salario che garantisce in media un livello di occupazione pari all’obiettivo [72](#), e quindi l’occupazione di equilibrio è data da:

$$al_I^j + (1 - a)l^* = -g[w^j - E(e) - E(p)] + [E(m) - E(p)] \quad (73)$$

Notate che tutte le variabili casuali compaiono in valore atteso, ipotizzando che la contrattazione avvenga sulla base di ciò che le parti si aspettano accada in media.

L'equilibrio macroeconomico: il salario

In equilibrio generale con aspettative razionali deve accadere che

$$E(p^j) = E(p) = w^j - E(e) = w - E(e) \quad (74)$$

perchè tutte le imprese sono idendiche e il salario è unico per tutti i lavoratori.

Sostituendo la 74 nella 73, otteniamo il salario di equilibrio:

$$w = E(m) + E(e) - al_I^j - (1 - a)l^* \quad (75)$$

ossia il salario dipende

- positivamente dalla quantità nominale di moneta, al crescere della quale, *coeteris paribus*, si genera inflazione;
- positivamente dagli shocks tecnologici;
- negativamente dall'obiettivo occupazionale che il sindacato vuole conseguire.

L'equilibrio macroeconomico: l'occupazione

Poichè tutte le imprese sono uguali, e ipotizzando $g = 1$ per semplicità, l'occupazione in una generica azienda può essere scritta come:

$$l - l^* = a(l_I - l^*) + [m - E(m)] + [e - E(e)] \quad (76)$$

Questa è anche l'occupazione totale, a meno di un fattore di scala.

La [76](#) dice che la differenza tra occupati e forza lavoro totale cresce con:

- la differenza tra insiders e forza lavoro totale;
- gli shocks inattesi alla domanda aggregata;
- gli shocks inattesi alla produttività.

Il termine $a(l_I - l^*)$ manca nella specificazione classica di un modello macroeconomico: è pari a zero se:

- $a = 0$, ossia se il sindacato si impegna a rappresentare gli interessi di tutti i lavoratori;
- $l_I = l^*$, ossia se gli “insiders” coincidono con l'intera forza lavoro.

Cosa accade se il sindacato rappresenta solo gli occupati

Supponiamo che

- gli “insiders” siano gli occupati del periodo precedente:

$$l_{I,t}^j = l_{t-1}^j \quad (77)$$

- il sindacato rappresenti solo gli “insiders” e quindi $a = 1$.

In questo caso il salario di equilibrio è:

$$w_t = E(m) + E(e) - l_{t-1} \quad (78)$$

e dipende negativamente dal livello di occupazione del periodo precedente.

Il livello di occupazione di equilibrio è

$$l_t = l_{t-1} + [m - E(m)] + [e - E(e)] \quad (79)$$

e quindi è una “passeggiata casuale” (*random walk*).

“Insiders”, “Outsiders” e mercato del lavoro Europeo

Sotto le ipotesi della pagina precedente:

- l'occupazione del periodo corrente predice l'occupazione futura;
- se uno shock negativo (domanda o produttività) riduce l'occupazione, non vi sono meccanismi automatici che facciano ri-crescere l'occupazione;
- questo perchè se l'occupazione scende, il sindacato impone un salario appena sufficiente a occupare gli “insiders”;
- gli outsiders non hanno alcuna influenza sull'equilibrio del mercato del lavoro.

Questo modello è stato usato per spiegare l'“isteresi” del mercato del lavoro europeo durante gli anni '70-'80; in quel periodo:

- gli shock petroliferi avevano ridotto la produttività e provocato un aumento della disoccupazione;
- l'occupazione non tornava a crescere dopo l'esaurimento dello shock.

Esercizio: riflettete sul dibattito Italiano in tema di mercato del lavoro, alla luce di questo modello.

4.1.3 Il sindacato come strumento per correggere imperfezioni del mercato

Nelle sezioni precedenti abbiamo visto modelli che descrivono il sindacato in una luce negativa, come interferenza con l'equilibrio concorrenziale.

Secondo questi modelli il sindacato determina:

- non solo una perdita di efficienza rispetto all'equilibrio concorrenziale,
- ma anche una perdita di equità nella misura in cui il sindacato protegga gli “insiders” senza tener conto delle esigenze degli “outsiders”.

Esistono però anche modelli che descrivono il sindacato in una luce più positiva, come strumento che riduce le iniquità prodotte dal libero mercato:

- Il modello di monopsonio;
- Il modello di Solow - Weibull (Solow, 1994);
- I meccanismi di “exit” e “voice” (Hirshman, 1970; Freeman and Medoff, 1984).

Il modello di monopsonio

La parola monopsonio indica un mercato in cui esista un solo compratore.

Immaginate una città relativamente isolata in cui esista un'unica impresa.

Se gli abitanti vogliono lavorare con un salario che garantisca la sopravvivenza, possono farlo solo in questa impresa, oppure devono emigrare con costi proibitivi.

L'offerta (inversa) di lavoro degli abitanti è data dalla funzione

$$W^S = W(L) \quad \text{con} \quad W' > 0 \quad (80)$$

ossia l'offerta di lavoro è positivamente inclinata.

L'impresa produce utilizzando il solo fattore lavoro, con una funzione di ricavo

$$R = R(L) \quad \text{con} \quad R' > 0 \quad (81)$$

Il costo di produzione e'

$$C(L) = LW(L) \quad (82)$$

Il problema di ottimo dell'impresa monopsonistica

L'impresa massimizza il profitto definito come:

$$\Pi = R(L) - C(L) = R(L) - LW(L) \quad (83)$$

La condizione del primo ordine richiede che il salario soddisfi la seguente equazione:

$$R'(L) = W(L) - LW' = W(L) \left(1 - \frac{1}{\eta_{LW}} \right) \quad (84)$$

dove $\eta_{LW} = \frac{dL}{dW} \frac{W}{L}$ è l'elasticità dell'offerta di lavoro.

Questo equilibrio è rappresentato nella [Figura 4.1.3.A](#):

- L'impresa assume lavoratori fino a che il ricavo marginale è uguale al costo marginale (Punto *A*)
- I lavoratori assunti, non avendo alternative, sono disposti a lavorare per il salario W , corrispondente al Punto *M*.
- L'equilibrio concorrenziale corrisponde invece al Punto *C* in cui il ricavo marginale è uguale al salario.

L'inefficienza dell'equilibrio di monopsonio

In un mercato del lavoro monopsonistico:

- l'occupazione è inferiore a quella concorrenziale;
- il salario è inferiore a quello concorrenziale.

I lavoratori non hanno alternative e quindi sono costretti ad accettare le condizioni dell'impresa.

Ne consegue una perdita di benessere pari all'area gialla (*deadweight loss*) nella [Figura 4.1.3.B](#).

Questa perdita di benessere è analoga a quella che ha luogo in un mercato monopolistico, ossia in cui esiste un solo venditore.

Se in un mercato monopsonistico potessero entrare altre imprese, vi sarebbe competizione per assumere i lavoratori; occupazione e salari crescerebbero fino al livello concorrenziale.

È ancora attuale questo modello?

Il modello di monopsonio si adatta bene a descrivere il mercato del lavoro agli albori della rivoluzione industriale e del sistema capitalistico

In quel contesto, i lavoratori riuniti in sindacato possono imporre all'impresa salari più elevati e maggiore occupazione, aumentando l'efficienza del sistema.

Il monopsonio è stato anche utilizzato per giustificare l'introduzione di un Salario Minimo Garantito, come ad esempio in Francia. Se in concorrenza lo SMG riduce l'efficienza, in monopsonio ottiene l'effetto opposto ([Figura 4.1.3.C](#)).

Al crescere della mobilità globale del lavoro e del capitale, oggi giorno diventa più difficile descrivere il mercato del lavoro come un monopsonio.

Il modello è stato comunque riproposto in chiave moderna da Manning (2003) il quale sostiene che i lavoratori fronteggiano una situazione di “monopsonio informativo”, perchè sono meno informati delle imprese sulle opportunità di mercato.

Il modello di Solow - Weibull

Questo modello sostiene che non c'è contrapposizione di interessi tra “insiders” e “outsiders” perchè se i secondi facessero la concorrenza ai primi, alla fine starebbero tutti peggio.

Per gli “outsiders” è preferibile attendere “in coda” il momento di diventare “insiders”, piuttosto che scatenare una “race to the bottom” offrendosi alle imprese con un salario inferiore.

Un requisito fondamentale perchè questo accada è che la “coda” per diventare “insider” non sia troppo lunga.

Sotto questa luce, il sindacato non è altro che un “patto di non concorrenza” tra i lavoratori (un “cartello”) che si auto sostiene in equilibrio, fino a che nessuno a interesse a deviare.

Il sindacato quindi di fatto protegge anche gli “outsiders” (ma comunque non i consumatori perchè i salari più alti implicano prezzi più alti).

Gli elementi costitutivi del modello di Solow - Weibull

- Esistono due livelli di salario: alto e basso.
- Se il salario è alto solo gli insiders hanno lavoro
- Se è basso anche gli outsiders trovano lavoro, ma il salario è basso per tutti.
- Quando un insider va in pensione lascia il posto per un outsider.

L'outsider ha due possibilità:

1. Offrirsi alle imprese per un salario basso e lavorare subito.
2. Attendere in coda per diventare outsider e da lì in avanti guadagnare un salario alto

Per dato tasso di sconto, la strategia 2 è dominante se il periodo da outsider è corto mentre quello da insider è lungo

Esercizio: è realistico questo modello? Pensate all'opposizione dei giovani francesi al CPE.

I meccanismi di “exit” e “voice”

Hirshman (1970), mette in luce due possibili volti della relazione tra lavoratori e aziende:

- “exit” : ossia uscita, conflitto;
- “voice” : ossia dialogo, risoluzione pacifica dei contrasti;

L’idea di Hirshman è che il sindacato possa organizzare le richieste dei lavoratori rendendo più efficiente e meno frammentata la contrattazione.

Il sindacato quindi di fatto aumenta l’efficienza perchè riduce il conflitto, e riduce i costi di contrattazione per l’impresa, che deve contrattare con una sola controparte invece che con ciascun lavoratore.

Questa idea è stata ripresa da Freeman e Medoff (1984): Paradossalmente, non sono le imprese che hanno rendite ad attirare il sindacato. È il sindacato a rendere le imprese più efficienti e quindi a favorire la formazione di rendite.

Esercizio: è convincente questo modello oggi in Italia, pensando ad esempio al caso dell’Alfa Romeo di Arese?

4.2 I regimi di protezione dell'impiego (*Employment Protection Legislation*)

In un mercato perfettamente concorrenziale, le imprese dovrebbero essere libere di interrompere il rapporto di lavoro se la sua continuazione non è vantaggiosa.

La [Tabella 10.1](#) BL, mostra che, i paesi dell'OCSE differiscono notevolmente tra loro riguardo alla disciplina della risoluzione del rapporto di lavoro per volontà dell'impresa:

- paesi rigidi: nei quali le imprese sono fortemente vincolate; ad esempio i paesi dell'Europa continentale non-scandinava;
- paesi flessibili: nei quali le imprese sono libere di licenziare i lavoratori quando vogliono: ad esempio i paesi anglosassoni.

Ci proponiamo di studiare le ragioni per cui la EPL viene introdotta e le sue conseguenze in termini di

- efficienza
- equità.

Seguiamo il [Capitolo 10](#) BL. Vedi anche Ichino (1997).

In cosa consiste la EPL

I vincoli alla libertà di licenziamento delle imprese possono essere suddivisi in due categorie:

- trasferimenti dall'impresa al lavoratore (*severance payments*)
- vincoli amministrativi o pagamenti a terzi (e.g. avvocati nell'eventualità di un giudizio).

La differenza tra le due categorie è importante per la valutazione dell'efficienza della EPL.

Esempio: caratteristiche principali della disciplina Italiana:

- Art. 18 dello Statuto dei Lavoratori: giustificato motivo soggettivo e oggettivo; giusta causa.
- Il processo del lavoro e la tutela del lavoratore in giudizio.
- Il diritto alla reintegrazione.
- Quanto costa all'azienda un licenziamento? e quanto al lavoratore?

Gli argomenti a favore della EPL

Considerate nuovamente il mercato monopsonistico di cui abbiamo parlato, con un'unica impresa e una moltitudine di lavoratori che hanno quella impresa come unica possibilità di lavoro.

In questa situazione (tipica del capitalismo selvaggio delle origini) i lavoratori (l'esercito dei disoccupati di riserva, secondo Marx) sono la parte debole del mercato.

La piena libertà di licenziamento appare iniqua e inefficiente:

- i lavoratori non sono fazzoletti “kleenex” da usare e gettare al bisogno.

Per questo motivo il sindacato ha da sempre combattuto per l'imposizione di vincoli ai licenziamenti.

Ma in un mercato del lavoro competitivo e moderno, che conseguenze ha la EPL?

4.2.1 Il modello base: neutralità della EPL

Immaginiamo un'economia perfettamente concorrenziale, in cui il governo imponga alle imprese di pagare un "buono uscita" T ai lavoratori in caso di licenziamento.

Assumiamo che:

- i salari siano perfettamente flessibili;
- i lavoratori siano neutrali rispetto al rischio (ossia siano interessati solo al salario medio indipendentemente dalla sua variabilità).

In assenza di buono uscita, l'impresa prevede di assumere un lavoratore

- per τ periodi,
- ad un salario pari W^* per periodo

Poichè tutti i mercati sono concorrenziali il reddito totale del lavoratore $W^*\tau$ è pari alla sua produttività durante l'intera durata del contratto:

- un reddito maggiore metterebbe l'impresa fuori mercato;
- un reddito inferiore indurrebbe il lavoratore a cercare un altro lavoro.

La reazione del mercato all'imposizione del buono uscita

In questo caso,

- se il salario è perfettamente flessibile e
- il costo totale del lavoro non può cambiare perchè altrimenti il rapporto di lavoro cesserebbe,

le parti si accorderanno per ridurre il salario in ogni periodo in un modo tale per cui alla fine il lavoratore percepisca esattamente il valore della sua produttività.

Un modo per farlo è descritto nella [Figura 10.1](#) BL. Questa soluzione si applica al caso in cui il licenziamento non avviene prima di τ^* , e può invece avvenire ad una data successiva τ incerta.

Esercizio: come potreste modificare il contratto se τ fosse una data certa e scegliendo $\tau^* = \tau$?

Notate che un altro modo per leggere questo risultato è che se i lavoratori rinunciassero al “buono uscita” avrebbero un salario più alto in ogni periodo.

Perchè allora la EPL crea tanto dibattito se non ha effetti?

Se i mercati sono veramente concorrenziali tendono ad annullare gli effetti di interferenze legislative (cosa di cui tipicamente i giuristi non si rendono conto).

Perchè la EPL abbia effetti reali, è necessario che qualcosa d'altro impedisca ai mercati di aggiustarsi in modo da annullarne l'effetto.

Se ad esempio il costo del licenziamento ha due componenti:

- un trasferimento T al lavoratore;
- un trasferimento R a terzi (stato, giudici) oppure un costo in termini di tempo di giudizio, incertezza, etc.;

allora le parti possono annullare gli effetti della componente T ma non quelli della componente R .

Anche senza componente R , se la EPL si accompagna a rigidità salariale verso il basso, ad esempio per la presenza del sindacato o di patti impliciti di non concorrenza tra lavoratori (Solow - Weibull), il mercato non può aggiustarsi.

4.2.2 EPL in presenza di salari rigidi

Consideriamo due paesi e ipotizziamo per semplicità due situazioni estreme:

- il paese F , flessibile, in cui il licenziamento è sempre possibile senza costi;
- il paese R , rigido, in cui il licenziamento è impossibile (ossia ha costi infiniti).

La funzione di produzione è identica in ciascun paese

$$Y = A_i \log(L) \quad (85)$$

ed esiste una sola impresa per paese.

Il termini A_i con $i \in \{l, h\}$ rappresenta gli shocks di produttività che possono colpire i due paesi, con $A_h > A_l$.

La probabilità di uno shock positivo A_h è p mentre quella di uno shock negativo A_l è $1 - p$.

In entrambi i paesi il salario è fisso e pari a W .

L'equilibrio nel paese flessibile

Nel paese flessibile, in ogni periodo l'impresa osserva lo shock di produttività e poi decide quanti lavoratori assumere risolvendo il problema di ottimo:

$$\max_L \Pi^F = A_i \log(L) - WL \quad (86)$$

dove il prezzo di vendita del prodotto è stato normalizzato a 1. La soluzione è:

$$L_i^F = \frac{A_i}{W} \quad (87)$$

che descrive una domanda di lavoro negativamente inclinata che si sposta a destra o a sinistra a seconda degli shocks (vedi la [Figura 10.2](#) BL).

Quindi passando da un periodo negativo a uno positivo l'occupazione del paese flessibile aumenta di

$$\Delta L^F = \frac{A_h - A_l}{W} \quad (88)$$

e viceversa nel caso opposto.

Il livello di occupazione medio di lungo periodo è:

$$\bar{L}^F = \frac{pA_h + (1-p)A_l}{W} \quad (89)$$

L'equilibrio nel paese rigido

Nel paese rigido l'impresa non può aggiustare l'occupazione a seconda degli shocks: non assume nelle espansioni e non può licenziare nelle recessioni.

Il meglio che l'impresa può fare è scegliere un livello fisso di occupazione che massimizzi in media, ossia in valore atteso, i profitti.

Quindi l'impresa risolve il problema

$$\max_L E(\Pi^R) = [pA_h + (1 - p)A_l] \log L - WL \quad (90)$$

Notate che W non è una variabile casuale.

La soluzione a questo problema è:

$$L^R = \frac{[pA_h + (1 - p)A_l]}{W} \quad (91)$$

e nel paese rigido l'occupazione non registra alcuna fluttuazione ciclica, mentre la variabilità degli shocks si riflette tutta sui profitti.

Il confronto tra i due paesi

Il confronto suggerisce tre importanti conseguenze della EPL (vedi [Figura 10.2](#)).

1. Nel lungo periodo l'occupazione media è uguale nei due paesi.
 - Notate che il dibattito giornalistico si focalizza tipicamente sull'effetto della EPL sui livelli di disoccupazione, mentre la teoria suggerisce che l'effetto diretto sui livelli è irrilevante.
2. La EPL riduce la variabilità dell'occupazione e aumenta la variabilità dei profitti (vedi le [Figure 10.3 e 10.4](#)).
 - Quindi assicura i lavoratori contro il rischio occupazionale ma se le imprese sono avverse al rischio, potrebbero emigrare verso paesi flessibili.
3. La EPL rende l'economia meno efficiente (e per questa via indiretta potrebbe ridurre anche il tasso di occupazione).
 - L'economia flessibile è più efficiente perchè occupa un numero maggiore di lavoratori quando lo shock è positivo e quindi essi sono più produttivi; viceversa nel caso di shock negativo.

4.2.3 EPL, efficienza produttiva ed efficienza distributiva

La EPL riduce la capacità produttiva di un sistema:

- non solo perchè impedisce l'aggiustamento efficiente dell'occupazione al variare degli shocks,
- ma anche perchè riduce l'impegno esercitato dai lavoratori (vedi, ad esempio, [Ichino e Riphahn, 2005](#)).

D'altro canto:

- se il mercato è ancora simile a quello monopsonistico del primo capitalismo,
- o comunque se i lavoratori sono avversi al rischio e non hanno modo di assicurarsi contro il rischio di disoccupazione o di incapacità dei managers, perchè i mercati assicurativi e dei capitali sono imperfetti,

la società potrebbe preferire una perdita di efficienza produttiva al fine di una maggiore efficienza distributiva, a vantaggio dei lavoratori.

Perchè questo valga, però, una condizione necessaria è che tutti i lavoratori abbiano accesso alla protezione.

EPL ed equita: insiders, outsiders e consumatori

Nei sistemi economici avanzati, invece, la EPL tende a proteggere un gruppo relativamente ristretto di “insiders” mentre gli “outsiders” lavorano senza protezioni in un mercato competitivo:

- In Italia sono completamente protetti solo i dipendenti pubblici e quelli della grande impresa privata;
- gli altri lavoratori (la maggior parte) non gode di alcuna protezione o di protezioni inferiori

Posto che la concorrenza internazionale non consente l'estensione a tutti delle protezioni, è equo che la EPL protegga solo gli insiders?

- si, se valgono ad esempio le ipotesi del modello Solow - Weibull;
- no, se la EPL finisce per essere una protezione delle rendite parassitarie degli insiders a danno degli outsiders.

Anche nel caso in cui la EPL protegga gli outsiders, bisognerebbe anche tenere conto dell'interesse dei consumatori: pensate ad esempio all'efficienza dei servizi pubblici in Italia.

In mezzo al guado: i rischi di una riforma della EPL

Chi difende la EPL ricorre all'argomento che nei casi in cui una flessibilizzazione è stata introdotta, non si sono visti effetti positivi (vedi le [Figure 10.5 e 10.6](#)).

Questo argomento non è convincente per due motivi.

In primo luogo perchè abbiamo visto che non ci sono ragioni per attendersi un effetto immediato e diretto sui livelli di occupazione e disoccupazione.

- Gli effetti sono sull'efficienza (e sull'equità) del sistema e richiedono tempo.

In secondo luogo, se il governo è incerto e le imprese non percepiscono un segnale chiaro e definitivo di flessibilizzazione, l'economia potrebbe fermarsi “in mezzo al guado” con i difetti di entrambi i sistemi (vedi Bertola e Ichino 1995):

- le imprese licenziano i lavoratori in eccesso approfittando della flessibilizzazione;
- non assumono nuovi lavoratori per paura di non poterli licenziare in seguito se il governo cambia idea.

5 Dentro l'impresa

Fino ad ora abbiamo studiato:

- come e con quali caratteristiche si determina l'offerta di lavoro;
- come e con quali caratteristiche si determina la domanda di lavoro;
- l'equilibrio tra domanda e offerta se i mercati sono competitivi;
- come si modifica l'equilibrio competitivo in presenza di istituzioni che interferiscono con il suo funzionamento;
- che ragioni hanno di esistere queste istituzioni.

Ora utilizziamo lo “zoom” per spostarci dal mercato all'impresa, e studiare cosa accade quando impresa e lavoratore si incontrano.

Diventa essenziale in questa fase abbandonare l'ipotesi che i lavoratori siano omogenei e assumere che essi possano differire tra loro in termini di:

- produttività innata, che indicheremo con θ .
- impegno (*effort*) nel lavoro che indicheremo con e .

I problemi dell'impresa e del lavoratore

L'impresa, che assumiamo sia interessata a massimizzare il profitto, deve risolvere due problemi principali:

- come selezionare i lavoratori migliori se la produttività innata non è osservabile;
- come far in modo che i lavoratori, una volta assunti, si impegnino al massimo nel lavoro, internalizzando gli obiettivi dell'impresa.

Assumiamo invece che il lavoratore sia interessato a massimizzare la sua utilità che dipende:

- positivamente dal salario;
- negativamente dalla variabilità del salario, ossia il lavoratore è avverso al rischio;
- negativamente dall'impegno e , che è costoso per il lavoratore.

Esiste quindi un problema di potenziale contrapposizione tra gli obiettivi dell'impresa e del lavoratore.

5.1 Il problema della selezione dei lavoratori

Questo problema è interessante quando le caratteristiche dei lavoratori, in particolare la loro produttività innata, non sono perfettamente osservabili.

Inizieremo studiando questa situazione al fine di mostrare che

- quando la produttività dei lavoratori potenziali non è osservabile, il mercato potrebbe collassare (equilibri con *adverse selection*);
- esistono istituzioni del mercato del lavoro finalizzate ad impedire all'impresa di osservare la produttività individuale;
- queste istituzioni possono essere giustificate se nella società esiste una forte preferenza per l'uguaglianza ma rischiano di portare il mercato al collasso.

Esamineremo poi che cosa possono fare i lavoratori e le imprese quando la produttività non è osservabile:

- i lavoratori più abili cercheranno di segnalarsi (*equilibri di signalling*);
- le imprese cercheranno strumenti per distinguere e identificare i lavoratori più abili (*equilibri di screening*).

5.1.1 L'equilibrio con selezione avversa

Esiste un insieme continuo di lavoratori nell'intervallo $[0, 1]$, che differiscono in termini di output prodotto $\theta \in [\underline{\theta}, \bar{\theta}]$ con una distribuzione $F(\theta)$.

Il salario di riserva è $r(\theta)$ con $r' \geq 0$ e i lavoratori accettano offerte $w \geq r(\theta)$.

L'impresa produce con rendimenti di scala costanti usando solo lavoro.

L'impresa è *price taker*, il prezzo è normalizzato a 1 e non esistono barriere all'entrata (quindi zero extra profitti).

Come termine di riferimento, se θ fosse osservabile l'equilibrio sarebbe

$$w_o^*(\theta) = \theta \quad (92)$$

e i lavoratori che accettano questa offerta sarebbero definiti da:

$$\Theta_o^* = \{\theta : r(\theta) \leq w_o^*(\theta)\} \quad (93)$$

Questo è un equilibrio competitivo e quindi Pareto efficiente, ma genera ex post una situazione di disuguaglianza che la società potrebbe considerare non desiderabile se valutata prima che ciascun individuo conosca il suo θ .

Equilibrio quando la produttività non è osservata dall'impresa

Se l'impresa non può osservare θ , l'equilibrio è descritto da un unico livello di salario w_u^* e da un insieme di lavoratori Θ_u^* che accetta l'offerta, tali che:

$$w_u^* = E(\theta \mid \theta \in \Theta_u^*) \quad (94)$$

$$\Theta_u^* = \{\theta : r(\theta) \leq w_u^*\} \quad (95)$$

L'equazione 94 assicura che il salario offerto dall'impresa sia uguale alla produttività media dei lavoratori che accetteranno l'offerta in equilibrio.

L'equazione 95 definisce l'insieme di lavoratori che accetta l'offerta w_u^* .

L'impresa anticipa correttamente la produttività di chi accetterà l'offerta.

Questo equilibrio potrebbe essere preferito a quello precedente in cui θ era osservabile, se esiste una forte preferenza sociale per l'uguaglianza.

Notate che in linea di principio potrebbe non esserci contrapposizione tra uguaglianza ed efficienza nel confronto tra i due equilibri.

Ragioni della non osservabilità di θ

Ci sono ostacoli “tecnologici” che non consentono all’impresa di osservare perfettamente la produttività individuale prima di assumere un lavoratore.

Ma oltre agli ostacoli tecnologici, esistono anche ostacoli istituzionali, tipicamente eretti dal sindacato.

Infatti se θ è artificialmente reso *non osservabile*, il salario di equilibrio non può dipendere da θ e questo è un tipico obiettivo egualitario del sindacato.

Questo obiettivo può essere giustificato utilizzando la teoria della giustizia del filosofo Rawls, secondo cui la funzione di benessere sociale dovrebbe coincidere con il benessere degli “ultimi” nella società,

$$R = \text{Min} \{w(\theta)\} \quad (96)$$

Questo perchè, sotto “il velo dell’ignoranza ex ante” su quale sarà la nostra produttività, vogliamo tutelarci contro il rischio di essere “ultimi”.

Notate: θ è una caratteristica innata, non una variabile controllabile come e .

Istituzioni che impediscono di osservare θ

- Liste numeriche per le assunzioni.
- Restrizioni sulla possibilità di acquisire informazioni sui chi fa domanda per un posto di lavoro.
- Restrizioni sulla possibilità di monitorare i lavoratori durante il periodo di prova.
- Restrizioni sulla possibilità di usare contratti temporanei prolungati come periodo di prova.

Esistono poi istituzioni che impediscono di utilizzare informazioni su θ anche se disponibili:

- Sistemi retributivi non basati sulla produttività.
- Promozioni automatiche basate sulla anzianità e non sul merito.
- Liste di licenziamento basate su anzianità (*last in first out*).
- In generale, istituzioni non meritocratiche.

Il rischio di selezione avversa e collasso del mercato

Questa situazione è stata studiata dal Premio Nobel George Akerlof in un famoso articolo del 1973, in cui utilizzava come esempi:

- il mercato delle macchine usate;
- il mercato delle assicurazioni.

Con riferimento al mercato del lavoro supponiamo che:

- $r(\theta) < \theta \quad \forall \theta \in [\underline{\theta}, \bar{\theta}]$, ossia tutti i lavoratori siano felici di lavorare se potessero ricevere un salario pari a θ ;
- $r(\theta)$ sia strettamente crescente in θ , ossia i lavoratori più produttivi hanno alternative migliori e quindi hanno salari di riserva più alti.

Sotto queste ipotesi, la qualità media dei lavoratori interessati al posto offerto aumenta con il salario offerto, e viceversa.

Questo genera la possibilità di “selezione avversa”: se l’impresa abbassa il salario per ridurre i costi, attira lavoratori peggiori e quindi meno produttivi, e questo la costringe ad abbassare ulteriormente il salario, con un circolo vizioso che può determinare il collasso del mercato.

Analisi formale dell'equilibrio con selezione avversa

Questo equilibrio è caratterizzato da:

$$w_u^* = E(\theta \mid r(\theta) \leq w_u^*) \quad (97)$$

$$\Theta_u^* = \{\theta : r(\theta) \leq w_u^*\} \quad (98)$$

Per analizzarlo utilizziamo le figure seguenti:

Figura 13.B.1:

- Se il mercato non collassa otteniamo un equilibrio egualitario che aumenta il benessere Rawlsiano ma è ex post inefficiente perchè persone che dovrebbero lavorare, non lavorano.

Figura 13.B.2:

- tuttavia il mercato collassa se si arriva ad un livello di salario tale per cui la qualità dei lavoratori potenziali è troppo bassa per convincere l'impresa ad assumere qualcuno.

La “lama di rasoio” su cui operano le istituzioni Rawlsiane

Supponiamo di poter caratterizzare l'equilibrio del primo capitalismo come un equilibrio competitivo in cui θ è perfettamente osservabile. Questo equilibrio è Pareto efficiente ma genera ineguaglianza ex post.

In questa situazione si genera una pressione sociale a favore di istituzioni Rawlsiane. Tuttavia, impedendo l'osservazione di θ , queste istituzioni:

- massimizzano il benessere del lavoratore meno produttivo;
- generano un equilibrio che è ex ante Pareto efficiente “sotto il velo Rawlsiano di ignoranza”,
- con il costo di generare una situazione che potrebbe rivelarsi Pareto inefficiente ex post.

Ma il problema più grave è che impedendo l'osservazione di θ (ossia la meritocrazia) il mercato può collassare e tutti finiscono per stare peggio.

Come evitare il rischio di selezione avversa?

Indipendentemente dal fatto che abbia origini istituzionali o tecnologiche, l'impossibilità di osservare θ danneggia i lavoratori più produttivi e le imprese.

Questo può determinare le seguenti reazioni delle parti danneggiate:

- I lavoratori più produttivi (ossia la parte “informata” del mercato) possono intraprendere azioni finalizzate a *segnalare* la loro maggiore abilità.
- Le imprese (ossia la parte “non informata” del mercato) possono offrire
 - contratti con caratteristiche tali per cui solo i lavoratori più abili trovano vantaggioso offrirsi;
 - menù di contratti per cui i lavoratori trovino vantaggioso distinguersi volontariamente l'uno dall'altro sulla base del loro θ ;
- Le imprese (ossia la parte “non informata” del mercato) possono anche assumere i lavoratori e usare il periodo di prova per fare “screening”.

Notate che i modelli di segnalazione e *screening* hanno infinite applicazioni anche in campi diversi dal mercato del lavoro e in particolare nei mercati finanziari.

5.1.2 Il modello di segnalazione

Supponiamo ora che la produttività assuma solo due valori:

θ_H = produttività elevata (con probabilità λ);

θ_L = produttività bassa (con probabilità $1 - \lambda$).

Una versione di questo modello è descritta con una notazione diversa nella [Sezione 3](#) del Capitolo 2 BL.

La produttività θ non può essere osservata direttamente dall'impresa.

L'impresa vorrebbe poter distinguere i lavoratori per poterli pagare in proporzione alla loro produttività.

Il salario di riserva è zero per entrambi i lavoratori:

$$r(\theta_L) = r(\theta_H) = 0 < \theta_L < \theta_H.$$

L'equilibrio se la segnalazione non è possibile

Sappiamo dalla Sezione 5.1.1 che un equilibrio possibile in questo caso è:

$$w_u^* = E(\theta) \quad (99)$$

$$\Theta_u^* = \{\theta_L, \theta_H\} \quad (100)$$

Tutti lavorano e il salario è pari alla produttività media.

Notate che anche se θ non è osservabile, questo equilibrio:

- è Pareto efficiente perchè $r(\theta) < \theta$ per tutti i θ ;
- la piena occupazione è efficiente perchè nessuno preferisce il salario di riserva;
- tuttavia i lavoratori più produttivi sono pagati meno della loro produttività;
- l'impresa perde sui lavoratori θ_L e guadagna su quelli θ_H ;
- è una situazione egualitaria secondo il principio di Rawls;
- tuttavia il mercato può collassare se r diventa maggiore di $E(\theta)$.

Cosa accade se un segnale per θ diventa possibile

Supponiamo che i lavoratori possano andare a scuola per un numero di anni pari a s a loro scelta.

Per quanto paradossale, è interessante ipotizzare che a scuola i lavoratori non imparino nulla di utile per modificare la loro produttività che rimane quindi pari a quella innata.

Supponiamo però che per i lavoratori più abili sia più facile conseguire il diploma.

Sotto queste ipotesi è possibile che i lavoratori θ_H vadano a scuola solo per segnalare che sono diversi dagli altri.

Ovviamente il segnale non sarebbe informativo se fosse conveniente lanciarlo anche per i lavoratori θ_L .

Il segnale è informativo, solo se i lavoratori θ_H hanno un incentivo a inviarlo mentre i θ_L preferiscono non inviarlo.

Ipotesi sul costo di lanciare il segnale

- Indicando con $C(s, \theta)$ il costo di andare a scuola:

$C(0, \theta) = 0$: il costo è zero per chi non va a scuola.

$C_s(s, \theta) > 0$: il costo aumenta con gli anni di scuola per tutti i θ .

$C_{ss}(s, \theta) > 0$: il costo è convesso negli anni di istruzione per tutti i θ .

$C_\theta(s, \theta) < 0$: il costo diminuisce al crescere di θ_H .

$C_{s\theta}(s, \theta) < 0$ il costo marginale di un anno di scuola è inferiore per i lavoratori θ_H .

- Quindi, in particolare, aumentare s è più costoso per i lavoratori θ_L che per i lavoratori θ_H .

- Per ogni lavoratore, l'utilità di accettare un offerta w dell'impresa è

$$U(w, s | \theta) = w - C(s, \theta) \quad (101)$$

- Curve di indifferenza che soddisfano queste ipotesi sono rappresentate nella [Figura 2.4](#) BL per i due tipi di lavoratori.

Equilibri *separating* e *pooling*

Supponete che l'impresa annunci che pagherà un salario pari a $w = \beta s$ dove β ha un valore corrispondente alla retta rappresentata nella [Figura 2.4](#) (nella notazione della figura il salario è I).

È evidente che i lavoratori più abili θ_H sceglieranno un livello di istruzione maggiore di quello dei θ_L

In questo caso si ha quindi un equilibrio *separating*, in cui i lavoratori si “separano” per livello di abilità, e in particolare i più produttivi vanno a scuola più a lungo per lanciare il segnale.

Esercizio: disegna nella [Figura 2.4](#) una una funzione dei salari $w = \gamma s$ tale per cui si determini un equilibrio *pooling* in cui tutti i lavoratori scelgono lo stesso livello di scolarità e nessuno segnala.

Esercizio: se l'impresa annunciassse due livelli salariali (e quali?) potrebbe fare profitti maggiori che nella [Figura 2.4](#)?

Equilibri di separazione ed efficienza

È importante osservare che l'equilibrio di segnalazione non è un *first best* perché inviare il segnale è costoso per i lavoratori più produttivi.

Il *first best* continua ad essere solo la situazione in cui θ sia perfettamente osservabile.

Posto che θ non sia osservabile, però, l'equilibrio di separazione è *second best efficient*.

Questo modello, studiato dal Premio Nobel Mark Spence nel 1973, ha infinite applicazioni in numerosi contesti.

Notate che se questo modello è vero:

- andare a scuola non aumenta il capitale umano, come ipotizzato nella Sezione [2.2](#);
- andare a scuola è solo un segnale della produttività innata.

Attualmente gli economisti non sono ancora riusciti a trovare un modo convincente per testare quale di questi due modelli sia vicino alla realtà.

5.1.3 Il modello di *screening*

Se la “segnalazione” da parte dei lavoratori più produttivi non è possibile, l’impresa può proporre un contratto di lavoro tale per cui solo i lavoratori θ_H trovino ottimale fare domanda.

Le caratteristiche di questo meccanismo di *screening* sono:

- i lavoratori vengono assunti per un primo periodo di prova alla fine del quale devono superare un test;
- solo se superano il test vengono assunti per un secondo periodo (ossia a tempo indeterminato);
- il salario è più alto nel secondo periodo rispetto al primo, anche se la produttività del lavoratore rimane costante in entrambi i periodi.

Vogliamo dimostrare che sotto certe condizioni, un contratto di questo tipo è attraente solo per i lavoratori θ_H .

In questo modo, quindi, l’impresa può identificare e assumere solo i lavoratori migliori.

Le ipotesi

Consideriamo di nuovo i due tipi di lavoratori θ_H e θ_L della Sezione 5.1.2 ma assumiamo che i salari di riserva possano essere diversi:

$$r(\theta_L) = W_L < r(\theta_H) = W_H \quad (102)$$

Il contratto di lavoro dura due periodi:

- W_1 è il salario offerto dall'impresa nel primo periodo;
- W_2 è il salario offerto dall'impresa nel secondo periodo;
- alla fine del periodo 1 i lavoratori devono superare un test tale per cui:
 - i lavoratori θ_H hanno probabilità P_H di passare;
 - i lavoratori θ_L hanno probabilità $P_L < P_H$ di passare.
- Assumiamo inoltre, per completezza, che la reputazione impedisce all'impresa di rinnegare il contratto e non pagare W_2 nel secondo periodo.

L'impresa vuole stabilire quale profilo salariale nei due periodi è necessario per convincere solo i lavoratori θ_H a fare domanda.

Caratteristiche della soluzione per questo problema

Un salario sufficientemente alto nel secondo periodo induce i lavoratori θ_H a fare domanda.

È però necessario abbassare sufficientemente il salario nel primo periodo per disincentivare i lavoratori θ_L a fare domanda.

Questo perchè i lavoratori più abili sono disposti a guadagnare poco nel primo periodo dato che hanno una probabilità alta di passare il test e guadagnare tanto nel secondo periodo.

I lavoratori meno abili invece, non avendo molte chance di passare il test, subirebbero solo l'aspetto meno attraente del contratto.

Esercizio: Che esempi potete trovare di contratti di questo tipo nei mercati del lavoro reali?

Il vincolo di partecipazione per i due tipi di lavoratori

I lavoratori θ_H fanno domanda se:

$$W_1 + P_H W_2 + (1 - P_H)W_H \geq 2W_H \quad (103)$$

I lavoratori θ_L fanno domanda se:

$$W_1 + P_L W_2 + (1 - P_L)W_L < 2W_L \quad (104)$$

Notate che c'è un unico set di salari che soddisfa entrambi i vincoli [103](#) e [104](#) con il segno di uguale dato P_H e P_L .

Il differenziale salariale che non basta per fare screening

Risolvendo rispetto a W_2 i vincoli 103 e 104 con il segno di uguale otteniamo:

$$\bar{W}_2 = \frac{(1 + P_H)W_H - (1 + P_L)W_L}{P_H - P_L} \quad (105)$$

Risolvendo rispetto a W_1 i vincoli 103 e 104 con il segno di uguale otteniamo:

$$\bar{W}_1 = \frac{-P_L(1 + P_H)W_H + P_H(1 + P_L)W_L}{P_H - P_L} \quad (106)$$

Il differenziale tra questi due salari per i due periodi:

$$\bar{W}_2 - \bar{W}_1 = \frac{(1 + P_H)(1 + P_L)(W_H - W_L)}{P_H - P_L} = X \quad (107)$$

non basta a fare *screening* perchè entrambi i lavoratori sono indifferenti tra fare domanda o meno.

Per avere *screening* è necessario un differenziale salariale maggiore tra i due periodi.

Un profilo salariale che consente lo screening

Considerate invece questa relazione

$$W_1^* = 2W_H - P_H W_2^* - (1 - P_H)W_H \quad (108)$$

Se W_1 e W_2 soddisfano questa relazione (che è la 103) i lavoratori θ_H sono indifferenti e assumiamo che in caso di indifferenza facciano domanda.

Sostituendo la 108 in 104, otteniamo il differenziale che sarebbe necessario per disincentivare i lavoratori θ_L :

$$W_2^* - W_1^* > \frac{(1 + P_H)(1 + P_L)(W_H - W_L)}{P_H - P_L} = \bar{W}_2 - \bar{W}_1 = X \quad (109)$$

Ossia iniziando da un differenziale pari a X , che rende i due lavoratori indifferenti, se l'impresa aumenta W_2 e riduce W_1 in modo da non violare il vincolo 108, i lavoratori θ_H rimangono indifferenti.

I lavoratori θ_L , invece, non sono più interessati perchè perderebbero tanto nel primo periodo per un guadagno troppo incerto nel secondo.

Perchè è necessario sia il profilo salariale che il test

- Risolvendo la 103 con $W_1 = W_2 = \tilde{W}$ mostra che un salario costante convincerebbe i lavoratori θ_H solo se $\tilde{W} = W_H$.
- Ma l'equazione 103 mostra che ovviamente un salario costante $\tilde{W} = W_H$ convincerebbe i lavoratori θ_L a fare domanda indipendentemente dalla probabilità di passare il test.
- E con un salario inferiore a W_H , solo i lavoratori θ_L (o nessuno) farebbero domanda.
- Analogamente, con probabilità identiche di passare il test o farebbero domanda tutti e due i lavoratori, o solo i peggiori o nessuno.
- È proprio la combinazione ottimale di differenze salariali e differenze di probabilità di passare il test che rende convenienti ai soli lavoratori θ_H pagare “un pegno” nel primo periodo per avere un vantaggio nel secondo.

5.2 Il problema della incentivazione dei lavoratori

Fino ad ora abbiamo considerato il problema di una impresa che deve assumere lavoratori con caratteristiche diverse, non osservabili da parte dell'impresa.

L'impresa vorrebbe assumere i lavoratori per lei "migliori" rispetto a qualche criterio, ma l'asimmetria informativa complica le cose:

- il mercato può collassare per via della selezione avversa;
- non è detto che si verifichino le condizioni che consentono:
 - ai lavoratori più produttivi di segnalarsi;
 - all'impresa di identificare e distinguere i migliori dai peggiori.

Passiamo ora allo stadio successivo del rapporto di lavoro, in cui l'impresa ha assunto dei lavoratori i quali, indipendentemente dalla loro produttività,

- possono lavorare con diversi livelli di impegno e (*effort*);
- e esercitare un maggiore impegno è costoso per loro.

Dove sta il problema? Il *moral hazard*

Supponete che l'impresa osservi il prodotto Y del lavoratore dato da

$$Y = e + \varepsilon \quad (110)$$

dove ε è una componente casuale non osservabile: ossia il prodotto dipende:

- dall'impegno esercitato del lavoratore, che *l'impresa non vede*;
- da altri fattori casuali che nè il lavoratore nè l'impresa possono controllare.

L'impresa potrebbe pagare il lavoratore “a cottimo” ossia in proporzione al prodotto Y , ma questi potrebbe non impegnarsi perchè anche se lui esercitasse molto impegno uno shock negativo potrebbe annullare tutto il suo guadagno.

In altri termini l'impresa fronteggia il problema di conciliare

- l'esigenza di incentivare il lavoratore (ossia di evitare il *moral hazard*) e
- l'esigenza di assicurare il lavoratore contro shocks non controllabili.

Questo è noto come il “problema del contratto di agenzia” o “problema del principale e dell'agente” e ha infinite applicazioni non solo in economia del lavoro.

5.2.1 Il problema della contrattazione tra un “principale” ed un “agente”

L'agente sceglie quanto impegno e esercitare a fronte del quale sopporta un costo $c(e)$, crescente e convesso ($c' > 0$ e $c'' > 0$).

Il prodotto Y che interessa al principale è dato da:

$$Y = e + \varepsilon \quad (111)$$

e il principale osserva solo il totale non le due componenti.

La componente di errore ε è distribuita normalmente con media 0 e varianza σ^2 .

Il principale offre all'agente un contratto che prevede una relazione lineare tra retribuzione e prodotto

$$W(Y) = s + bY \quad (112)$$

Ci chiediamo che valori debbano assumere i parametri a e b per assicurare una conciliazione ottimale delle esigenze di incentivazione e assicurazione.

Caratteristiche delle possibili soluzioni

Per dare il massimo degli incentivi al lavoratore, il principale deve:

- farsi pagare dall'agente un ammontare fisso e lasciare il resto del prodotto (non necessariamente positivo) all'agente;
- nei termini della equazione 112 questa soluzione implica $W = s + Y$ con $s < 0$ e $b = 1$;
- Esercizio: che esempi potete trovare di soluzioni di questo tipo nella realtà?

Per dare il massimo della assicurazione al lavoratore, il principale deve:

- pagare all'agente un ammontare fisso e tenersi il resto del prodotto (non necessariamente positivo);
- nei termini della equazione 112 questa soluzione implica $W = s$ con $s > 0$ e $b = 0$;
- Esercizio: che esempi potete trovare di soluzioni di questo tipo nella realtà?.

È ragionevole ipotizzare che l'ottimo (notate: un *second best*, perchè?) sia collocato in qualche posizione intermedia tra questi estremi.

Passi per la soluzione del problema

La soluzione di questo tipo di modelli di “teoria dei contratti” viene solitamente decomposta nei seguenti passi.

Il principale massimizza il profitto dati due vincoli relativi al comportamento dell’agente:

- il *vincolo di partecipazione*: il contratto deve avere caratteristiche tali per cui l’agente abbia voglia di firmarlo volontariamente;
- il *vincolo di incentivazione*: il principale anticipa che l’agente reagirà in modo per lui ottimale agli incentivi ricevuti.

La soluzione procede “in senso logicamente inverso” (rispetto a come la proposta contrattuale verrebbe naturalmente fatta nella realtà):

1. per dati parametri s e b del contratto si risolve il problema dell’agente;
2. si derivano i vincoli di partecipazione e incentivazione dell’agente;
3. si risolve il problema del principale, trovando i parametri s e b che massimizzano il profitto dati i vincoli.

Il problema dell'agente

In questo corso descrivo i lineamenti della soluzione senza entrare nei dettagli che richiedono strumenti matematici relativamente più complessi di quelli a vostra conoscenza.

Una versione analoga di questo modello è presentata anche nel [Capitolo 13](#) BL.

L'utilità dell'agente è

$$U(W, e) = -\exp\{-r(W - c(e))\} \quad (113)$$

dove $r \geq 0$ è il coefficiente di avversione assoluta al rischio.

$U_0 = -1$ è l'utilità di riserva sotto cui l'agente non vuole scendere, e che quindi determina la sua disponibilità a partecipare

La normalizzazione a -1 è irrilevante. Cattura semplicemente le alternative a disposizione del lavoratore, prese come standard di riferimento.

I vincoli di incentivazione e partecipazione

Sostituendo 112 e 111 nella funzione di utilità l'agente risolve il seguente problema:

$$\max_e E(U(W, e)) = \int_{\varepsilon} -\exp\{-r(s + b(e + \varepsilon) - c(e))\} dF(\varepsilon) \quad (114)$$

che ha come soluzione un effort ottimale $e^*(b)$ che soddisfa la condizione del primo ordine.

$$b = c'(e) \quad (115)$$

Non importa che siate in grado di seguire i passaggi matematici, ma provate a capire come interpretare queste relazioni:

La 115 è il *vincolo di incentivazione* che indica al principale come l'agente reagirà al parametro b : l'agente sceglierà un ammontare di impegno e tale per cui il costo marginale dell'impegno eguagli il suo beneficio marginale pari a b .

Il *vincolo di partecipazione* è invece:

$$E(U) \geq U_0 = -1 \quad (116)$$

Il problema del principale

Il principale sceglie i parametri del contratto salariale in modo da massimizzare il profitto

$$\Pi = Y - s - bY = (1 - b)(e + \varepsilon) - s \quad (117)$$

ma tenendo conto del fatto che:

- l'agente sceglie l'impegno in modo per lui ottimale dati gli incentivi;
- se l'agente non consegue un'utilità superiore a quella di riserva non partecipa.

Poichè $E(\varepsilon) = 0$, il problema si riduce a:

$$\max_{b,s} E(\Pi) = (1 - b)e - s \quad (118)$$

sotto i vincoli

$$e = e^*(b) \quad (119)$$

$$E(U) \geq U_0 = -1 \quad (120)$$

Chi fosse interessato alla soluzione matematica può trovarla nelle *Lecture notes* di *Personnel Economics* scaricabili dal [mio sito](#) alle pagine 89-91.

Il contratto ottimale

I parametri ottimali del contratto $W(Y) = s + bY$ sono:

$$s^* = -b^*e^* + c(e^*) - \frac{1}{2}r(b^*)^2\sigma^2 \quad (121)$$

$$b^* = \frac{1}{1 + r\sigma^2c''} \leq 1 \quad (122)$$

Il parametro su cui è interessante riflettere è b :

- Se l'agente è neutrale rispetto al rischio ($r = 0$), allora $b^* = 1$:
 - ossia il principale “affitta” l'impresa per un canone fisso $F = s_{b^*=1}^*$;
 - l'agente si tiene tutto il prodotto che rimane dopo il pagamento del canone;
 - l'incentivazione è massima.
- Se l'agente è avverso al rischio ($r > 0$), allora $b^* < 1$ e diminuisce al crescere:
 - della varianza degli shocks σ^2 ;
 - della convessità c'' della funzione di costo, ossia al crescere del costo *marginale* dell'impegno, in funzione dell'impegno stesso.
 - l'agente è parzialmente assicurato e l'incentivazione è meno forte.

Perchè nella realtà osserviamo molti contratti con $b^* \approx 0$?

- L'avversione al rischio dei lavoratori potrebbe essere molto elevata.
- Economisti e psicologici hanno mostrato che l'incentivazione monetaria a volte ottiene effetti opposti a quelli previsti.
- Il prodotto individuale è misurabile solo in modo imperfetto.
- È più facile misurare l'input piuttosto che l'output.
- Esistono altri meccanismi impliciti di relazione tra lavoratori, principale e mercati che stimolano l'impegno.
- In presenza di più agenti il principale ottiene risultati migliori basando gli incentivi sulla performance relativa, non assoluta.
- Se i lavoratori operano in gruppo gli incentivi non possono essere individuali.
- Anche un salario indipendente dal prodotto può incentivare i lavoratori se esiste un rischio positivo di licenziamento.

Questi problemi sono l'oggetto di una delle aree di ricerca più attive nell'economia del lavoro moderna: la *Personnel Economics*. Ne esaminiamo alcuni.

5.2.2 Performance relativa e tornei

Supponete che esistano numerosi agenti e che gli shocks casuali alla produttività abbiano due componenti:

- una componente individuale specifica per ciascun agente;
- una componente aggregata che influenza tutti gli agenti insieme.

Pensate ad esempio al gestore di un fondo di investimento: i risultati della sua gestione dipendono, oltre che dal suo impegno, da:

- shocks all'intero mercato, e che quindi colpiscono tutti i fondi;
- shocks specifici del suo portafoglio.

Il principale può offrire all'agente un contratto in cui la retribuzione dipenda dalla differenza tra la performance del fondo e quella del mercato.

In questo modo può aumentare l'effetto incentivante del contratto senza ridurre l'assicurazione offerta all'agente, la cui retribuzione, non è affetta da shocks che influenzano l'intero mercato.

Contratti di questo tipo sono frequenti per gestori di fondi e managers di aziende.

Il ricorso alla produttività relativa all'interno dell'aziende

Dentro l'impresa, l'incentivazione dei lavoratori sulla base della performance relativa viene tipicamente utilizzata nella forma di "tornei" per le promozioni.

L'uso della parola "torneo", proposto per descrivere queste situazioni da Lazear and Rosen (1981), non è casuale

Una situazione in cui:

- più lavoratori competono per essere promossi ad un livello superiore e
- un solo lavoratore può essere promosso,

ha le stesse caratteristiche di un torneo sportivo e produce effetti incentivanti analoghi.

Secondo Lazear e Rosen la retribuzione basata sulla performance è poco frequente nel mercato del lavoro, perchè i tornei per le promozioni sono il vero modo con cui le imprese incentivano i lavoratori.

Un modello matematico di un torneo

Considerate una impresa con due lavoratori j e k che producono rispettivamente,

$$Y_j = e_j + \varepsilon_j \quad (123)$$

$$Y_k = e_k + \varepsilon_k \quad (124)$$

dove

- e_j e e_k sono i livelli di impegno dei due agenti;
- ε_j e ε_k sono variabili casuali a media zero che descrivono la fortuna dei due agenti;
- il prodotto è osservabile e può essere oggetto di contratto.

I lavoratori hanno un salario iniziale W_0 e hanno un costo dell'impegno $C(e_j)$.

Il principale offre un salario $W_1 > W_0$ all'agente che riesce a produrre di più nel primo periodo.

In molti casi una promozione comporta un cambiamento del tipo di lavoro svolto. Qui astraiano da questo aspetto, e consideriamo soltanto il fatto che il lavoratore preferisce W_1 piuttosto che W_0 .

Che decisioni devono prendere i lavoratori

L'agente j massimizza rispetto a e_j il seguente obiettivo

$$\max_{e_j} W_0(1 - P^j(e_j, e_k)) + W_1 P^j(e_j, e_k) - C(e_j) \quad (125)$$

where $P^j(e_j, e_k)$ è la probabilità endogena di vincere il premio.

La condizione del primo ordine è

$$(W_1 - W_0)P_{e_j}^j - C'(e_j) = 0 \quad (126)$$

dove $P_{e_j}^j$ è una derivata parziale, e la condizione per l'agente k è analoga.

La probabilità di vincere il premio per j è data da

$$\begin{aligned} P^j(e_j, e_k) &= \text{Prob}(e_j + \varepsilon_j > e_k + \varepsilon_k) = \text{Prob}(e_j - e_k > \varepsilon_k - \varepsilon_j) \\ &= G(e_j - e_k) \end{aligned} \quad (127)$$

dove G è la distribuzione cumulata di frequenza della differenza tra gli shocks $\varepsilon_k - \varepsilon_j$, e la condizione per l'ottimo è

$$(W_1 - W_0)g(e_j - e_k) - C'(e_j) = 0 \quad (128)$$

L'impegno esercitato dai due lavoratori

I due lavoratori sono identici e risolvono lo stesso problema nello stesso modo. Quindi:

$$e_j^* = e_k^* \quad (129)$$

e la condizione del primo ordine si semplifica in

$$(W_1 - W_0)g(0) = C'(e_j^*) \quad (130)$$

che può essere risolta per e_j^* (e in modo simile per e_k^*).

Questo risultato ha due implicazioni interessanti:

- maggiore è il premio salariale, maggiore è l'impegno;
- maggiore è il parametro $g(0)$ maggiore è l'impegno.

Notate che $g(0)$ descrive quanto importante è la fortuna nel determinare il vincitore

Il ruolo della fortuna

Cosa accade se la fortuna è *poco* importante.

- La fortuna è poco importante quando i due shocks sono molto simili. Nel caso estremo sono idenetici: $\varepsilon_j = \varepsilon_k$.
- In questo caso la distribuzione G è degenere in 0 e $g(0)$ va all'infinito.
- Sotto questa ipotesi l'impegno determina chi vince e quindi è redditizio impegnarsi.

Cosa succede invece se la fortuna è *molto* importante.

- La fortuna è importante nel determinare il vincitore quando la distribuzione G della differenza $\varepsilon_k - \varepsilon_j$ ha molta massa nelle code.
- Questo significa che con probabilità elevata un giocatore è stato fortunato proprio quando l'altro è stato sfortunato.
- Sotto queste ipotesi la distribuzione G è piatta e $g(0)$ ha un valore basso.
- Quindi non è redditizio impegnarsi perchè la fortuna conta troppo.

Il problema del principale

Anticipando il comportamento degli agenti, il principale vuole scegliere i due livelli di salario W_1 and W_0 in modo da massimizzare il profitto per lavoratore:

$$\text{Max } \Pi = e - \frac{W_0 + W_1}{2} \quad (131)$$

dato il vincolo di incentivazione

$$(W_1 - W_0)g(0) = C'(e) \quad (132)$$

e il vincolo di partecipazione

$$\frac{W_0 + W_1}{2} \geq C(e) \quad (133)$$

Notate che il lato sinistro del vincolo di partecipazione è il salario medio dato che la probabilità di vittoria si pari a $G(0) = .5$.

Stiamo anche assumendo che gli agenti sono neutrali rispetto al rischio e hanno un'utilità

$$U = E(W) - C \quad (134)$$

La soluzione del problema del principale

Sostituendo il vincolo di partecipazione nell'obiettivo da massimizzare:

$$\max_{W_0, W_1} \Pi = e - C(e) \quad (135)$$

e le condizioni del primo ordine sono:

$$[1 - C'(e)] \frac{\partial(e)}{\partial W_0} = 0 \quad (136)$$

e

$$[1 - C'(e)] \frac{\partial(e)}{\partial W_1} = 0 \quad (137)$$

Queste condizioni sono soddisfatte quando $C'(e) = 1$. Usando il vincolo di incentivazione questa condizione richiede che il premio salariale sia pari a

$$(W_1 - W_0) = \frac{1}{g(0)} \quad (138)$$

Commenti

- $C'(e) = 1$ implica che il torneo induce gli agenti a esercitare un impegno che eguaglia il costo marginale per il lavoratore al beneficio marginale dell'azienda, e questo è socialmente ottimale.
- Dal punto di vista dei lavoratori questa è una situazione analoga a quella del “Dilemma del prigioniero” in teoria dei giochi.
- Questo ammontare di impegno deve anche soddisfare il vincolo di partecipazione.
- La dimensione del premio deve essere inversamente proporzionale all'importanza della fortuna misurata da $g(0)$.
 - Quando la fortuna non è importante, $g(0)$ è elevato e un piccolo premio è sufficiente per indurre un impegno elevato.
 - Quando la fortuna è molto importante, $g(0)$ è basso e un premio grande è necessario per indurre gli agenti ad esercitare impegno.

Performance relativa e mercati reali

Il modello dei tornei spiega perchè possano avere senso le “retribuzioni d’oro” dei managers o degli sportivi.

Queste retribuzioni non sono pensate per incentivare chi le riceve ma soprattutto per incentivare i lavoratori che stanno sotto di loro nella gerarchia e che ambiscono ad una promozione.

I tornei sono quindi uno strumento potente per pensare alla struttura delle retribuzioni.

Presentano però il rischio di disincentivare la cooperazione tra i dipendenti e incentivare il sabotaggio dei compagni di lavoro.

Ecco perchè la letteratura aziendale sulla gestione delle risorse umane enfatizza la necessità di mantenere “l’armonia interna all’azienda, lo spirito di gruppo ...”

5.2.3 Incentivazione collettiva e “free riding”

L'incentivazione diventa ancora più problematica se

- gli agenti operano in gruppo (*team working*);
- il prodotto complessivo è osservabile e contrattabile,
- ma l'output specifico di ciascun individuo non è osservabile,
- e l'effort specifico di ciascun individuo non è osservabile.

In questa situazione c'è il rischio che gli agenti si comportino come *free riders* (vedi Holmstrom, 1982), riducendo l'impegno individuale anche in assenza di incertezza.

Disegnare una incentivazione individuale appropriata è impossibile in questo contesto, ma il problema può essere risolto attraverso:

- l'emergere di una gerarchia all'interno del gruppo;
- il ripetersi a lungo termine della interazione tra i membri del gruppo.

Un modello semplice di free riding

Considerate

- due lavoratori $i \in \{1, 2\}$;
- che contribuiscono ad un progetto congiunto con un impegno $e_i \in [0, \infty)$;
- sopportano un costo dell'impegno $c(e)$, con $c' > 0$ e $c'' > 0$;
- hanno una funzione di utilità

$$U_i = w_i - c(e_i) \quad (139)$$

dove w_i è il compenso per la partecipazione al progetto.

Il prodotto congiunto è osservabile senza incertezza ed è dato da:

$$Y = f(e_1 + e_2) \quad (140)$$

con $f' > 0$ e $f'' < 0$.

L'impegno individuale non è osservabile.

Il “first best”

Il “first best” è dato dalla soluzione di questo problema:

$$\max_{e_1, e_2} f(e_1 + e_2) - c(e_1) - c(e_2) \quad (141)$$

La condizione del primo ordine è:

$$f' = \frac{dc}{de_1} = \frac{dc}{de_2} \quad (142)$$

Ad esempio, se

- $C = \frac{e^2}{2}$
- $f(e_1 + e_2) = e_1 + e_2$

l'impegno dei due agenti che consentirebbe il “first best” è:

$$e_1^* = e_2^* = 1 \quad (143)$$

Il prodotto se gli agenti si “dividono la torta” a metà

Supponiamo che gli agenti decidano di dividere il prodotto a metà con un compenso individuale pari a:

$$w_i = \frac{Y}{2} \quad (144)$$

Questo è ad esempio quello che accadrebbe in una società cooperativa.

Data questa regola di divisione, ogni agente sceglie l'impegno ottimale e_i massimizzando:

$$\max_{e_i} \frac{f(e_i + e_j)}{2} - c(e_i) \quad (145)$$

e la condizione del primo ordine diventa

$$\frac{f'}{2} = \frac{dc}{de_1} = \frac{dc}{de_2} \quad (146)$$

Indicando con \bar{e} il livello di impegno scelto dagli agenti in questo caso, è immediato verificare che

$$\bar{e}_i < e_i^* \quad (147)$$

Il problema del “free riding”

Considerando di nuovo il caso specifico prima descritto

- $f(e_1 + e_2) = e_1 + e_2$
- $C = \frac{e^2}{2}$

si ha che

$$\bar{e}_i = \frac{1}{2} < e_i^* = 1 \quad (148)$$

inoltre il prodotto totale

$$\bar{Y} = 1 < Y^* = 2 \quad (149)$$

è inferiore al first best ma è interamente distribuito tra gli agenti data la regola di divisione:

$$w_i = \frac{1}{2} \quad (150)$$

Con due agenti il prodotto totale si riduce a metà. Con N membri, il fattore di riduzione sarebbe $\frac{1}{N}$.

“Assunzione di un principale” per risolvere il prolema

Holmstrom (1982) suggerisce l'idea che le gerarchie emergano in una comunità primitiva cooperativa proprio per risolvere il problema del free riding.

La soluzione consiste nell'assumere un principale:

- che si impegni a infliggere penalità o premi a seconda del comportamento degli agenti;
- che sia disposto a farsi carico di quanto rimane (positivo o negativo) del prodotto dopo la distribuzione agli agenti, se la regola di divisione prevede che il prodotto non sia interamente distribuito.

Proprio il fatto che i membri del gruppo accettino di non distribuire tutto, riduce la propensione al free riding.

Un semplice esempio

Supponiamo di nuovo che:

- $f(e_1 + e_2) = e_1 + e_2$
- $C = \frac{e^2}{2}$
- $U_i = w_i - C(e_i)$

e che il principale si impegni a far osservare questa regola di divisione:

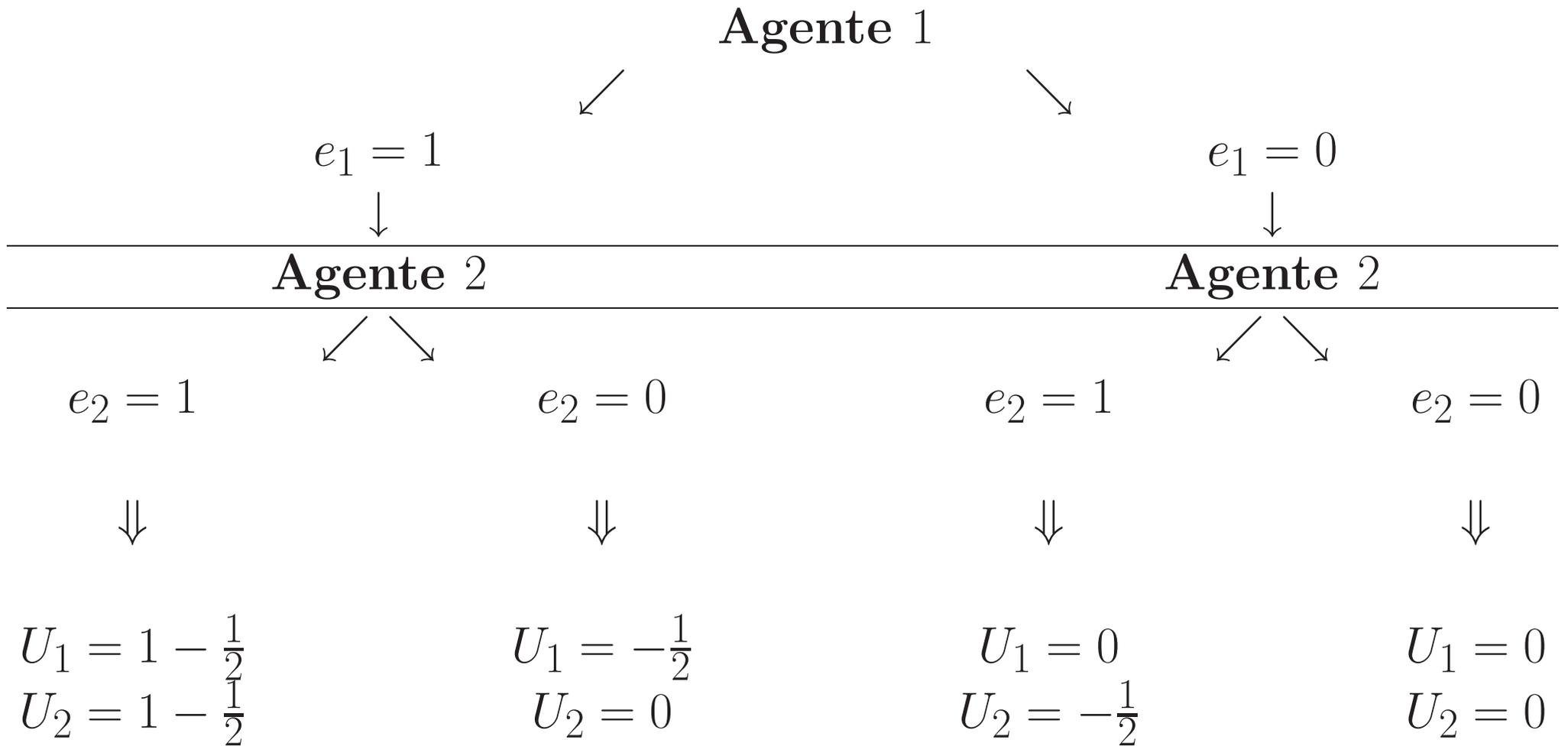
$$w_i = \begin{cases} \frac{Y}{2} & \text{if } Y \geq Y^* = e_1^* + e_2^* = 2 \\ 0 & \text{if } Y < Y^* = e_1^* + e_2^* = 2 \end{cases} \quad (151)$$

Notate che con questa regola di divisione e costi convessi di impegno:

- Ogni impegno $e > 1$ è dominato da un impegno $e = 1$;
- Ogni impegno $0 < e < 1$ è dominato da un impegno $e = 0$.

Considerate il “gioco” di interazione tra i due membri del gruppo data questa regola di divisione.

Il gioco in presenza di un principale



Equilibrio del gioco

Ci sono due equilibri di Nash in questo gioco:

- $e_1 = 1$ e $e_2 = 1$;
- $e_1 = 0$ e $e_2 = 0$.

Ma l'equilibrio cooperativo corrispondente al first best ora è possibile.

Quando prevale l'equilibrio cooperativo il principale non trattiene nulla.

Ma nell'altro caso, in cui gli agenti fanno free riding, il principale minaccia di trattenersi parte del prodotto.

Quindi l'equilibrio di free riding (in cui per altro nulla verrebbe prodotto) non ha ragione di verificarsi.

Interazioni ripetute

L'equilibrio cooperativo può anche essere indotto endogenamente dal ripetersi "all'infinito" delle interazioni degli agenti, con questa regola di divisione del prodotto:

$$w_{it} = \begin{cases} \frac{Y_t}{2} & \text{se } Y_{t-1} \geq Y^* \\ 0 \text{ per sempre} & \text{se } Y_{t-1} < Y^* \end{cases} \quad (152)$$

Con questo schema, se gli agenti non sono troppo impazienti, il beneficio immediato di fare free riding è inferiore del costo di lungo periodo derivante dalla punizione dei compagni

Il ragionamento è analogo a quello che abbiamo usato per spiegare il risultato di Solow- Weibull nella Sezione. [4.1.3](#)

Questo argomento è stato utilizzato per spiegare perchè i sistemi retributivi con partecipazione ai profitti possono incentivare la produttività.

5.2.4 Salari di efficienza e incentivazione basata sul rischio di licenziamento

Fino ad ora abbiamo visto che, a seconda del contesto:

- la retribuzione basata sulla performance individuale;
- la valutazione della performance relativa;
- l'emergenza endogena di una gerarchia;

possono risolvere il problema della incentivazione.

Un'altra soluzione è offerta dalla combinazione

- della “carota”: i lavoratori sono pagati più del loro salario alternativo e
- del “bastone”: i lavoratori rischiano il licenziamento se non si impegnano.

Questa strategia è possibile solo se l'impegno degli agenti è osservabile almeno in modo imperfetto attraverso l'osservazione del prodotto o attraverso qualche forma di monitoraggio.

Il salario definito da questa strategia viene chiamato salario di efficienza (vedi il [Capitolo 14](#) BL).

La letteratura sui salari di efficienza

Con il termine “salari di efficienza” si indica una vasta letteratura che ha come denominatore comune l’individuazione meccanismi per i quali un salario reale maggiore di quello competitivo consente alle imprese di aumentare i profitti.

- Economia dello sviluppo: vedi le citazioni in Katz (1986).
- Minaccia di sindacalizzazione: Dickens (1986).
- Selezione avversa: Guash and Weiss (1981); è la situazione studiata nella sezione [Section 5.1.1](#).
- Disoccupazione come strumento per il mantenimento della disciplina: Shapiro and Stiglitz (1984).

Gran parte di questa letteratura è stata utilizzata per giustificare l’ipotesi Keynesiana di rigidità del salario e quindi la possibilità di disoccupazione “involontaria” di equilibrio (vedi il [Capitolo 14](#) BL.)

Qui invece ci focalizziamo sul salario di efficienza come strumento per incentivare la forza lavoro indipendentemente dalle sue conseguenze macroeconomiche (vedi Gibbons and Waldman, 1999).

Un modello di salario di efficienza

Consideriamo una impresa neutrale rispetto al rischio che offre ad un lavoratore anche lui neutrale rispetto al rischio un salario w in ogni periodo.

Se il lavoratore rifiuta l'offerta la sua alternativa è un salario w_a .

Se il lavoratore accetta può scegliere tra due comportamenti:

- un livello di impegno alto e_H con un costo c ;
- un livello di impegno basso e_L con un costo pari a zero.

Il prodotto Y può essere alto H o basso L e quindi $H > L$.

- se $e = e_H$ allora $Y = H$;
- se $e = e_L$ allora:
 - $Y = H$ con probabilità p
 - $Y = L$ con probabilità $1 - p$

Quindi l'osservazione del prodotto rivela solo imperfettamente quanto impegno è stato esercitato dall'agente.

L'equilibrio se il rapporto di lavoro dura un solo periodo

Assumiamo inoltre:

$$H - c > w_a > pH + (1 - p)L \quad (153)$$

ossia che è socialmente efficiente per il lavoratore essere occupato in questa impresa ed esercitare un livello di impegno alto.

Tuttavia se il rapporto di lavoro dura un solo periodo il risultato efficiente non può essere raggiunto perchè:

- se l'impresa si impegna a pagare ex post il lavoratore non ha motivo di credere che l'impresa mantenga la promessa.
- se l'impresa paga ex ante non ha motivo di credere che il lavoratore si impegnerà.
- Quindi in questo caso l'impresa offre

$$w_{os} = pH + (1 - p)L < w_a \quad (154)$$

e il lavoratore rifiuta perchè il salario alternativo è più alto.

L'equilibrio se il rapporto di lavoro dura più periodi

Supponiamo che r sia il tasso di sconto comune a impresa e agente.

Consideriamo questa offerta dell'impresa:

$$w_t = \begin{cases} w^* > w_a & \text{if } w_{t-1} = w^* \text{ and } Y_{t-1} = H \\ 0 \text{ per sempre} & \text{altrimenti} \end{cases} \quad (155)$$

Notate che questa offerta implica tornare all'equilibrio monoperiodale nel caso in cui una delle due parti non mantenga il suo impegno.

Se il lavoratore non si impegna

- risparmia il costo c per un periodo;
- ma perde il valore atteso scontato della differenza tra salario di efficienza e salario alternativo per tutti i periodi successivi:

$$(1 - p) \frac{w^* - w_a - c}{r} \quad (156)$$

Il salario di efficienza w^*

Possiamo quindi determinare il salario di equilibrio w^* pari al salario minimo sufficiente a far scegliere al lavoratore un livello di impegno alto:

$$w^* = w_a + c + \frac{rc}{1-p} \quad (157)$$

Il punto chiave è che il salario di efficienza paga non solo il costo opportunità del lavoratore $w_a + c$ ma anche una rendita $\frac{rc}{1-p}$.

Il rischio di perdere questa rendita in caso di licenziamento induce il lavoratore ad impegnarsi.

Il salario di efficienza W^* è maggiore del salario competitivo nell'azienda $w_a + c$, e quindi se tutte le imprese si comportano nello stesso modo, può generare disoccupazione involontaria (Shapiro-Stiglitz (1984)).

Notate tuttavia che entrambe le parti, dal loro punto di vista privato, preferiscono il salario di efficienza e il salario può quindi rimanere stabilmente lontano dall'equilibrio competitivo, generando disoccupazione involontaria.

Bibliografia

- Akerlof, George (1973), “The market for lemons: Quality uncertainty and the market mechanism” *Quarterly Journal of Economics*, 89:488-500.
- Alesina Alberto and Andrea Ichino (2007) “Gender Based Taxation” <http://www2.dse.unibo.it/ichino>
- Alesina Alberto and Andrea Ichino (2007) “Meno tasse per lei, più tasse per lui”, *Il Sole24ore* 27 marzo 2007.
- Autor D., Katz L. and M. Kearney (2006) “The Polarization of the U.S. Labour Market” NBER Working Paper No. 11986.
- Becker, Gary S. (1975), “Human Capital and the Personal Distribution of Income: An Analytical Approach” (W.S. Woytinsky Lecture), in: Gary S. Becker, *Human Capital*, 2nd edition. New York: Columbia University Press.
- Becker, Gary S. and Nigel Tomes (1986) “Human Capital and the Rise and Fall of Families” *Journal of Labor Economics*, Vol. 4, No. 3, Part 2, pp. S1-S39.
- Behrman, Jere R., Taubman, Paul, and Wales, Terence J. (1977), Controlling for and Measuring the Effects of Genetics and Family Environment in Equations for Schooling and Labor Market Success. In *Kinometrics: Determinants of Socioeconomic Success within and between Families*, edited by Paul Taubman. Amsterdam: North-Holland.
- Behrman, Jere R., Taubman, Paul, (1989) Is Schooling Mostly in the Genes? Nature-Nurture Decomposition Using Data on Relatives. *J.P.E.* 97 (December): 1425-46.
- Bertola G. e A. Ichino (1995) Crossing the River: A Comparative Perspective on Italian Employment Dynamics, *Economic Policy: A European Forum*, October, (21), 359-415.
- Blanchard, Olivier J. and Summers Larry (1986) “Hysteresis and the European Unemployment Problem”, NBER Working Paper, n. 1950.
- Blanden, Jo, Paul Gregg and Stephen Machin (2005) “Intergenerational mobility in Europe and North America”, *Center for Economic Performance*
- Bjorklund, Anders, Eriksson, Tor, Jantti, Markus, Raaum, Oddbjorn and Osterbacka, Eva, (2002) “Brother Correlations in Earnings in Denmark, Finland, Norway and Sweden Compared to the United States”, *Journal of Population Economics*, Vol. 15 (4), pp. 757-72.
- Card, David (1995a), “Earnings, Schooling, and Ability Revisited”, *Research in Labor Economics* 14, 23–48.
- Card, David (1995b), “Using Geographic Variation in College Proximity to Estimate the Returns to Schooling”, in: L.N. Christofides, E.K. Grant, and R. Swidinsky (eds.), *Aspects of Labour Market Behaviour: Essays in Honour of John Vanderkamp*. Toronto: University of Toronto Press, 201–222,.
- Card, David (1999), “The Causal Effect of Education on Earnings”, in: Orley Ashenfelter and David Card (eds.). in: *Handbook of Labor Economics* Vol. 3, North-Holland, Amsterdam.

- Checchi, Daniele , Andrea Ichino and Aldo Rustichini (1999) “More equal but less mobile? Education financing and intergenerational mobility in Italy and in the US”, *Journal of Public Economics*, 74, pp 351-393.
- Checchi, Daniele , Andrea Ichino and Aldo Rustichini (1995) “Immobili perchè eguali. Mobilit  occupazionale e scolastica tra generazioni in Italia e Stati Uniti”, in Galli G. *La mobilit  sociale in Italia*, Roma, Edizioni Confindustria, .
- Dickens, W.T. (1986), “Wages, employment and the threat of collective action by workers”, mimeo, Berkeley University.
- Erickson, C. and A. Ichino (1994) “Wage Differentials in Italy. Market Forces, Institutions and Inflation”, in: Richard Freeman and Larry Katz (eds.), *Differences and Changes in the Wage Structure*, Chicago and London: University of Chicago Press, 265-305.
- Freeman, R. and Medoff J. (1984), “What do unions do?”, New York, Basic Books.
- Freeman, R. and Katz L. (eds.) (1994), *Differences and Changes in the Wage Structure*, Chicago and London: University of Chicago Press.
- Giavazzi, Francesco (2007) “Il caso spagnolo e i dilemmi italiani: come si creano i posti di lavoro” *Corriere della Sera*, 16 marzo 2007.
- Gibbons, Robert and Michael Waldman (1999), “Careers in Organizations: Theory and Evidence”, in Orey Ashenfelter and David Card, “Handbook of Labor Economics”, Volume 3, ch. 36, Amsterdam, Elsevier Science.
- Guash, Luis and Andrew Weiss (1981), “Self-selection in the labor market” *American Economic Review*, June, 275-284.
- Hassler, John, Jos V. Rodriguez and Morayand Joseph Zeira (2007), “Inequality and Mobility”, mimeo, Hebrew University
- Hirschman, Albert (1970) “Exit, Voice and Loyalty”, Cambridge, Harvard University Press.
- Holmstrom, Bengt (1979), “Moral hazard and observability” *Bell Journal of Economics*, 9: 74-91.
- Holmstrom, Bengt (1982), “Moral hazard in teams” *Bell Journal of Economics*, 13: 324-340.
- Holmstrom, Bengt (1999), “Managerial Incentive Problems: A Dynamic Perspective” *Review of Economic Studies*, 66: 169-182.
- Katz, Lawrence F. (1986), “Efficiency Wage Theories: A partial Evaluation” in the “NBER Macro Annual.
- Katz, L. , and K. Murphy (1992) “Changes in relative wages , 1963-1987: Supply and demand factors”, *Quarterly Journal of Economics*, February.
- Ichino A. (1997) La disciplina limitativa dei licenziamenti. Effetti e giustificazioni nella letteratura economica recente, *Politica Economica* , Dicembre, 13(3), 375-408.
- Ichino A. e R. Riphahn (2005) The Effect of Employment Protection on Worker Effort. A comparison of absenteeism during and after probation, *Journal of the European Economic Association*, March, 3(1), 120-43.

- Lalonde, Robert (1986), Evaluating the econometric evaluations of training programs with experimental data, *The American Economic Review*, Vol. 76, No. 4.
- Lazear, Edward P. and Sherwin Rosen (1981), "Rank-Order Tournaments as Optimum Labor Contract", *Journal of Political Economy* 89(5), 841-864.
- Lazear, Edward P. (1995), "Personnel Economics", Cambridge, MIT press.
- Manning, A. (2003), "Monopsony in motion: imperfect competition in labour markets", Princeton, Princeton University Press.
- Mocetti, S. (2007), "Intergenerational income mobility in Italy", mimeo, Università di Bologna
- Holland, Paul W. (1986), "Statistics and Causal Inference", *Journal of the American Statistical Association* 81, 945-970.
- Plomin, Robert, and Petrill, Stephen A. (1997) Genetics and Intelligence: Whats New?, *Intelligence* 24 (January/February): 5377.
- Rawls, CHECK (19??), "A theory of justice", ????
- Shapiro, Carl and Joseph Stiglitz (1984), "Equilibrium Unemployment as a Discipline device", *American Economic Review*, 74, 433-444.
- Solon, Gary (1992), "Intergenerational Income Mobility in the United States" *The American Economic Review*, Vol. 82, No. 3., pp. 393-408.
- Solow, Robert (1994), "Il mercato del lavoro come istituzione sociale", Bologna, Il Mulino.
- Spence, M (1973) "Job market signalling" *Quarterly Journal of Economics*, 87, 355-374.
- Zimmerman, David J. (1992) "Regression Toward Mediocrity in Economic Stature", *The American Economic Review*, Vol. 82, No. 3., pp. 409-429.