

**LE POLITICHE DI ESTENSIONE
DELLA DURATA DEL BREVETTO
NEL SETTORE FARMACEUTICO**

ANNALISA BIAGI

DOTTORATO DI RICERCA IN ECONOMIA PUBBLICA

XXV CICLO

UNIVERSITÀ DI CATANIA

RINGRAZIAMENTI

Ringrazio prima di tutto il mio supervisore, il professor Roberto Cellini, per la competenza e la pazienza con cui ha guidato il mio lavoro di ricerca. Ringrazio poi il dipartimento di Economia dell'Università di Leicester per l'ospitalità e i commenti che ho ricevuto su una versione preliminare della tesi. Infine, la mia profonda gratitudine va a Vincenzo Denicolò, che mi ha suggerito un tema di ricerca appassionante, ha continuamente discusso con me lo sviluppo del mio lavoro, e mi ha aiutata a superare le inevitabili criticità incontrate. Questa tesi non avrebbe potuto essere scritta senza il suo aiuto.

INTRODUZIONE

L'industria farmaceutica è importante non solo per le sue dimensioni in termini strettamente economici, ma soprattutto per il contributo che può dare al miglioramento della durata e della qualità della nostra vita. Per gli studiosi della teoria dei brevetti, tuttavia, essa riveste un ruolo speciale per il fatto che è uno dei pochi settori (e forse il più importante) in cui i brevetti sono veramente essenziali per proteggere gli innovatori.¹ Questo dipende probabilmente dalla particolare chiarezza con cui i brevetti farmaceutici delimitano i confini del monopolio riservato all'inventore. Tale monopolio, infatti, riguarda semplicemente la formula chimica del principio attivo, il che rende relativamente semplice l'individuazione e la repressione delle violazioni – in una parola, l'*enforcement* del diritto. Inoltre nel settore farmaceutico la necessità della protezione brevettuale è evidente, a causa degli elevati costi fissi dell'introduzione di un nuovo farmaco² e della facilità dell'imitazione.

Tuttavia, la letteratura economica sui brevetti ha fino ad ora trascurato alcune importanti peculiarità della protezione brevettuale nel settore farmaceutico, che riguardano nello specifico la durata effettiva dei brevetti. Questa infatti non coincide con la durata nominale, attualmente pari a 20 anni. Da una parte, nel settore farmaceutico il tempo di sviluppo del

¹Questo è confermato da un'ampia evidenza empirica, basata sia su questionari somministrati ai responsabili della Ricerca e Sviluppo (*Research and Development*, R&D) di imprese manifatturiere, sia sulla stima del cosiddetto *patent premium*, ovvero l'extra-profitto garantito dalla protezione brevettuale rispetto al profitto che l'innovatore potrebbe ottenere con altri strumenti di protezione (come il segreto industriale). Si veda, ad esempio Cohen, Nelson and Walsh (2000), Arora, Ceccagnoli e Cohen (2008) e Sherer (2010). Secondo Bessen e Maurer (2008), oltre la metà del valore di tutti i brevetti a livello mondiale è costituito dal valore dei brevetti detenuti da un piccolo numero di grandi compagnie farmaceutiche.

²Secondo Di Masi et al. (2003) il costo dell'introduzione di un nuovo farmaco è dell'ordine delle centinaia di milioni di dollari (800 milioni per la precisione). Queste stime sono state contestate (il centro di ricerca che le ha prodotte è finanziato dall'industria farmaceutica), ma nessuno contesta che i costi fissi siano estremamente elevati.

nuovo prodotto (cioè il tempo che passa tra il deposito della domanda di brevetto e l'immissione in commercio del nuovo farmaco) è molto lungo per via della regolamentazione a cui è soggetta l'introduzione di nuovi farmaci. Dall'altra, e proprio per questo, una specifica legislazione estende il periodo di esclusiva di cui godono i titolari dei brevetti farmaceutici. Negli Stati Uniti, questa legislazione è stata introdotta nel 1984 dall'*Hatch-Waxman Act*, che ha reso più semplice l'immissione sul mercato dei farmaci generici ma al tempo stesso ha consentito al titolare del brevetto di recuperare una parte del tempo perso per ottenere l'approvazione del nuovo farmaco. Regole analoghe sono poi state introdotte in Europa e nella maggior parte dei paesi industrializzati.

Queste regole implicano che la politica della durata dei brevetti nel settore farmaceutico non è definibile come uno scalare, ma piuttosto come una funzione che lega la durata effettiva del brevetto al tempo di sviluppo del farmaco. La forma specifica di questa funzione varia nel tempo e da paese a paese, come discuteremo più ampiamente nel primo capitolo (ma si possono vedere, per farsi una prima idea, le figure da 1.2 a 1.5). Tale funzione può essere vista come un menù di opzioni che viene offerto all'inventore, le cui decisioni di investimento determinano in ultima analisi la scelta di un punto lungo questa funzione (cioè un particolare valore della durata effettiva del brevetto).

Questo suggerisce che la politica ottimale della durata dei brevetti nel settore farmaceutico possa essere studiata come un problema di *mechanism design*. Questo è l'approccio che seguiremo nei capitoli 4 e 5; esso costituisce l'elemento di maggiore originalità della tesi.

La letteratura economica sui brevetti farmaceutici, infatti, ha trattato il tempo di sviluppo come una variabile stocastica ma esogena. I lavori che hanno cercato di valutare gli effetti dell'*Hatch-Waxman Act* sulla durata effettiva dei brevetti farmaceutici, per esempio, hanno semplicemente stimato la durata media del tempo di sviluppo per determinare se nel complesso la nuova legislazione abbia comportato un aumento o una diminuzione della vita effettiva dei brevetti (Grabowski e Vernon, 1986 e 1987).

Indipendentemente da questo filone della letteratura, un'altra serie di lavori empirici ha però dimostrato chiaramente che il tempo di sviluppo non è una variabile esogena, ma dipende dalle decisioni di investimento delle compagnie farmaceutiche. Questi lavori hanno infatti evidenziato come i tempi di sviluppo dei farmaci di maggior valore siano mediamente inferiori alla media: questo non sarebbe possibile se il tempo di sviluppo fosse un parametro esogeno.³

Se quindi il tempo di sviluppo è una variabile endogena che di fatto viene scelta dall'innovatore, diventa allora possibile concepire la funzione che lega la durata effettiva del brevetto al tempo di sviluppo come, per l'appunto, un "meccanismo" e studiarne il disegno ottimale. Questa tesi è il primo lavoro che analizza la politica della durata dei brevetti da questo punto di vista.

Un approccio di *mechanism design* è stato adottato dalla letteratura che ha studiato i premi monetari come un meccanismo di incentivazione dell'innovazione alternativo (o complementare) rispetto alla proprietà intellettuale.⁴ Questa letteratura parte dal presupposto che i premi monetari sono uno strumento di incentivazione meno distortivo rispetto ai brevetti. Il problema, però, è che l'ammontare del premio dovrebbe essere correlato al valore dell'innovazione, il che è impossibile in situazioni di informazione asimmetrica, in cui il valore dell'innovazione è noto all'innovatore ma non al policy-maker. Offrendo diverse combinazioni di premi e brevetti, in cui brevetti di durata e/o ampiezza minore sono associati a premi monetari più grandi e viceversa, il policy-maker può indurre gli innovatori ad "auto-selezionarsi." Questa letteratura è sicuramente molto interessante, ma i meccanismi che propone sono per il momento puramente ipotetici e non trovano riscontro nella realtà. Come vedremo, anche la dipendenza funzionale tra durata effettiva del brevetto e tempo di sviluppo induce innovazioni di diverso valore ad "auto-selezionarsi," ma in questo caso si tratta di un meccanismo realmente impiegato in pratica (anche se fino ad ora non analizzato

³Il lavoro più significativo al riguardo è quello di Dranove e Meltzer (1994), che sarà discusso ampiamente nel capitolo 3.

⁴Si vedano, ad esempio, Wright (1983), Cornelli e Schankerman (1999), Scotchmer (1999), Shavell e van Ypersele (2001), Hopenhyan et al. (2006) e Chari et al. (2012).

dal punto di vista teorico).

Il risultato principale della nostra analisi è così riassumibile: quando il tempo di sviluppo è relativamente breve, non è ottimale estendere la durata del brevetto; una certa estensione è però ottimale quando il tempo di sviluppo è più lungo. Qualitativamente, questa è precisamente la forma della relazione attualmente adottata in pratica (con l'eccezione della legislazione italiana del 1991, che sarà discussa nel primo capitolo). Infatti, sia negli Stati Uniti che in Europa non vi è alcuna estensione della durata nominale del brevetto quando il tempo di sviluppo è inferiore a 5-6 anni, ma vi è un'estensione per tempi di sviluppo maggiori.

Superficialmente, si potrebbe immaginare che il legislatore abbia semplicemente voluto ridurre i costi amministrativi dell'adeguamento della durata del brevetto, considerando che se il tempo di sviluppo è breve l'estensione sarebbe comunque limitata. Tuttavia, questa spiegazione non regge ad un esame più attento. La procedura necessaria per estendere la durata del brevetto è semplice e non richiede scelte discrezionali, e in pratica le compagnie farmaceutiche richiedono l'estensione anche quando questa è limitata a pochi giorni. Ciò suggerisce che i costi amministrativi siano molto bassi, sia per le agenzie governative che per le imprese. La nostra analisi mostra che non è ottimale estendere la durata del brevetto quando il tempo di sviluppo è sufficientemente breve, anche a prescindere dai costi amministrativi.

Il resto della tesi è strutturato come segue. Il primo capitolo, di natura descrittiva, illustra in dettaglio le caratteristiche peculiari del processo innovativo nel settore farmaceutico cui abbiamo accennato sopra. Mostriamo come, per via della regolamentazione relativa all'immissione in commercio dei nuovi farmaci, la protezione brevettuale operi in modo diverso da ogni altro settore dell'economia, e descriviamo la legislazione che prevede la possibilità di estendere la durata del brevetto per compensare l'innovatore del tempo perso per ottenere l'approvazione all'immissione in commercio.

Il secondo capitolo propone un modello teorico del processo innovativo nel settore farmaceutico che cattura questi aspetti peculiari. Il modello genera diversi risultati di statica

comparata che sono in linea con le aspettative o hanno giustificazioni intuitive, il che ne conferma l'adeguatezza per lo studio delle politiche di estensione del brevetto. L'analisi permette di scomporre l'incentivo ad investire nella fase di sviluppo di un nuovo farmaco in due componenti: l'effetto di anticipazione dei profitti e l'effetto di durata effettiva del brevetto.

Il terzo capitolo cerca di identificare empiricamente l'effetto di durata effettiva del brevetto. Questo effetto opera quando l'estensione del termine del brevetto non compensa interamente l'innovatore per il tempo perso (il tempo di sviluppo), e genera un incentivo addizionale a investire per ridurre il tempo di sviluppo. Poichè questo effetto è cruciale nell'analisi normativa, è importante confermarne l'esistenza empiricamente.

Consideriamo tre possibili strategie per identificare empiricamente l'effetto di durata effettiva. La prima, che purtroppo ci è preclusa per l'indisponibilità dei dati relativi al periodo in questione, sfrutterebbe i cambiamenti legislativi creati dalle riforme introdotte nel tempo (come l'approvazione dell'*Hatch-Waxman Act*). La seconda è basata sul fatto che l'effetto di durata effettiva implica una non-linearità nella relazione tra durata della fase pre-clinica e incentivo ad investire nella fase clinica. Questa non-linearità è confermata dai dati. Infine, la terza strategia sfrutta in modo più completo la suddivisione del processo di sviluppo di un nuovo farmaco in diverse fasi (come la fase pre-clinica, clinica e di approvazione), condizionando l'investimento in ciascuna fase alla lunghezza della fase precedente. Anche questo approccio conferma l'esistenza di un effetto di durata effettiva.

I capitoli 4 e 5 rappresentano il nucleo della tesi e sviluppano un'analisi normativa delle politiche di estensione del brevetto nel settore farmaceutico. Nel capitolo 4 mostriamo come, in condizioni di completa informazione (cioè se il policy-maker potesse condizionare la durata del brevetto al valore dell'innovazione), la politica ottimale non prevede alcuna estensione. In pratica, però, non è possibile prevedere una politica diversa per ogni singola innovazione, vuoi per ragioni di semplicità amministrativa, vuoi perchè le caratteristiche specifiche dell'innovazione spesso non sono osservabili o verificabili. Quindi la stessa funzione deve potersi applicare ad innovazioni eterogenee: in particolare, ad innovazioni di diverso valore e diversa

profitabilità. Questo contesto di informazione incompleta viene studiato nel capitolo quinto. Qui mostriamo che vale comunque un risultato di *no distortion at the top*: per le innovazioni di valore maggiore (che avranno tempi di sviluppo più brevi) non deve essere prevista alcuna estensione. Per innovazioni di valore minore (e quindi tempi di sviluppo più lunghi), invece, è generalmente desiderabile distorcere la soluzione prevedendo una estensione della durata del brevetto.

Come si è detto, qualitativamente la soluzione ottimale con informazione incompleta corrisponde alle politiche adottate in pratica. Dal punto di vista metodologico, può essere interessante notare come il problema da noi studiato fornisce quello che forse è l'unico esempio concreto di come la proprietà di *no distortion at the top* possa essere osservata in una politica realmente adottata.⁵

⁵Nella teoria della tassazione ottimale, per esempio, l'implicazione del principio di *no distortion at the top* è che l'aliquota marginale sui redditi più alti dovrebbe essere nulla – una proprietà che chiaramente è violata dalle politiche adottate in pratica.

INDICE

	Page
RINGRAZIAMENTI	ii
INTRODUZIONE	iii
CAPITOLO	
1 ASPETTI ISTITUZIONALI	1
1.1 INTRODUZIONE	1
1.2 LA REGOLAMENTAZIONE DEI NUOVI FARMACI	2
1.3 RICERCA E SVILUPPO NEL SETTORE FARMACEUTICO	8
1.4 LA DURATA EFFETTIVA DEI BREVETTI FARMACEUTICI	11
1.5 CONCLUSIONI	21
2 IL MODELLO TEORICO	23
2.1 INTRODUZIONE	23
2.2 IPOTESI	24
2.3 LO STADIO DELLO SVILUPPO	31
2.4 LO STADIO DELLA RICERCA	37
2.5 STATICA COMPARATA	39
2.6 CONCLUSIONI	44
3 ANALISI EMPIRICA	47
3.1 INTRODUZIONE	47
3.2 VALORE DEI FARMACI E TEMPO DI SVILUPPO	48
3.3 L'EFFETTO DI DURATA EFFETTIVA	52

3.4	LA RACCOLTA DEI DATI	61
3.5	STATISTICHE DESCRITTIVE	65
3.6	STIME	67
3.7	CONCLUSIONI	69
4	ANALISI NORMATIVA: COMPLETA INFORMAZIONE	73
4.1	INTRODUZIONE	73
4.2	LA FUNZIONE DEL BENESSERE SOCIALE	76
4.3	CURVA DI DOMANDA RETTANGOLARE	80
4.4	INNOVAZIONI NON DRASTICHE	85
4.5	INNOVAZIONI DRASTICHE	89
4.6	LA DURATA OTTIMALE DEL BREVETTO	90
4.7	CONCLUSIONI	95
5	ANALISI NORMATIVA: INCOMPLETA INFORMAZIONE	97
5.1	INTRODUZIONE	97
5.2	INCOMPLETA INFORMAZIONE	100
5.3	IL CASO DI PARTECIPAZIONE DETERMINISTICA	104
5.4	IL CASO GENERALE	108
5.5	CONCLUSIONI	118
	CONCLUSIONI	121
	BIBLIOGRAFIA	124

CAPITOLO 1

ASPETTI ISTITUZIONALI

Riassunto. In questo capitolo, di natura descrittiva, vengono illustrate le caratteristiche peculiari del processo innovativo nel settore farmaceutico, dove, per via della regolamentazione relativa all'immissione in commercio dei nuovi farmaci, la protezione brevettuale opera in modo diverso da ogni altro settore dell'economia. La descrizione di tali peculiarità serve per giustificare le ipotesi dell'analisi teorica elaborata nel secondo capitolo e per delineare il contesto dell'analisi empirica sviluppata nel terzo capitolo.

1.1 INTRODUZIONE

Questo capitolo illustra alcune caratteristiche peculiari del processo innovativo nel settore farmaceutico, che implicano come la protezione brevettuale operi qui in modo diverso da ogni altro settore dell'economia. Descriveremo come la regolamentazione del processo di immissione in commercio dei nuovi farmaci si sia evoluta nel tempo, influenzando le caratteristiche dell'innovazione farmaceutica e motivando l'introduzione di una legislazione specifica che prevede la possibilità di estendere la durata nominale del brevetto. Questa legislazione ha generato nuove dimensioni della politica della vita dei brevetti farmaceutici, che fino ad ora sono state trascurate dalla letteratura economica. Questo capitolo è prevalentemente descrittivo, ma è essenziale per comprendere i problemi che saranno analizzati nel resto di questa tesi.

1.2 LA REGOLAMENTAZIONE DEI NUOVI FARMACI

La durata effettiva dei brevetti farmaceutici differisce sensibilmente da quella nominale – ed è questa la ragione per cui la legge prevede, in certe situazioni, la possibilità di estendere la durata nominale del brevetto. La ragione principale è che i nuovi farmaci non possono essere commercializzati prima dell'autorizzazione da parte delle autorità di regolamentazione.¹ Nel gergo legale, questa è chiamata “autorizzazione all'immissione in commercio.” Per cominciare illustriamo allora come sono cambiati nel tempo i requisiti necessari per ottenere l'autorizzazione all'immissione in commercio e le implicazioni che questi cambiamenti hanno avuto sul processo di sviluppo dei nuovi farmaci.

1.2.1 CENNI STORICI

La regolamentazione dell'introduzione sul mercato di nuovi farmaci si è evoluta nel tempo, seguendo tappe simili in tutti i principali Paesi industrializzati. L'impulso propulsivo ha però quasi sempre avuto origine negli Stati Uniti, e gli altri Paesi hanno seguito a ruota. Per questa ragione, il nostro breve *excursus* storico si concentra sugli Stati Uniti.

Per molto tempo non vi è stata alcuna legislazione specifica sui farmaci. Per tutto l'ottocento e la prima parte del novecento, per commercializzare un nuovo farmaco non era necessaria alcuna approvazione governativa: al contrario, era il governo ad avere l'onere della prova nel caso in cui avesse voluto ritirare un farmaco dal mercato. La prima legge federale sui farmaci degli Stati Uniti fu emanata nel 1906: *Pure Food and Drugs Act*, nota anche come *Wiley Act*, dal nome del proponente, H. Wiley, responsabile del *Bureau of Chemistry* istituito presso il Dipartimento dell'Agricoltura.² Questa legge prescriveva semplicemente che i farmaci, così come gli alimenti, dovessero essere esenti da adulterazioni e sofisticazioni.

¹Un farmaco – dal greco *pharmakon*, veleno e/o medicamento, rimedio, a seconda delle dosi o dei modi in cui è somministrato – è un bene che influenza notevolmente la durata e la qualità della vita delle persone. Questa è la ragione per cui i governi intervengono in diverse fasi dello sviluppo e della commercializzazione.

²Come vedremo, il *Bureau of Chemistry* può essere considerato l'antesignano della *Food and Drug Administration*.

Nulla di più veniva previsto sul fronte della sicurezza; in particolare, non veniva istituita nessuna procedura di controllo preventivo e/o di autorizzazione per la commercializzazione.

La prima proposta in tal senso venne formulata nel 1933 dalla *Food and Drug Administration* (FDA), un'agenzia governativa creata da una costola del *Bureau of Chemistry* nel 1927. La proposta era basata sulla considerazione di come la legge del 1906 fosse insufficiente a contrastare la distribuzione di medicinali insicuri, la cui assunzione spesso comportava gravi effetti collaterali. La proposta era corredata da una raccolta di casi illustranti le carenze della normativa, che venne definita dalla stampa, per i suoi contenuti drammatici, la “camera degli orrori.”³ Nonostante ciò, l'iter legislativo al Congresso era molto lento.

La situazione si sbloccò solo nel 1938, a seguito di un evento tragico: l'anno prima un farmaco, il Sulfanilamide, aveva causato la morte di 107 persone, tra cui molti bambini. Il Sulfanilamide era un medicinale per la cura delle infezioni della gola da streptococco, distribuito prima di allora in compresse. Nel 1937 cominciò a essere distribuito sotto forma di sciroppo e, come fu appurato in seguito, per la preparazione in forma liquida si utilizzò come solvente il dietilenglicole, una sostanza comunemente usata come antigelo per le automobili. Il cambiamento della modalità di somministrazione fu preceduto da test di laboratorio volti a migliorare il sapore dello sciroppo (al gusto di lampone), ma non a verificarne la tossicità.

Se lo sciroppo fosse stato previamente somministrato a cavie animali, ci si sarebbe subito accorti delle sue proprietà letali; anche una semplice disamina della letteratura scientifica dell'epoca avrebbe evidenziato che il dietilenglicole era tossico e potenzialmente letale. Ma a quel tempo, come si è detto, non erano previsti controlli preventivi sulla tossicità dei farmaci. La vendita di medicinali tossici poteva forse danneggiare la reputazione commerciale di un'impresa, ma non era illegale. Quindi nessuno dei controlli che si potevano fare fu fatto, e lo sciroppo fu prodotto e distribuito liberamente in molti stati americani. Le vittime dello ‘sciroppo al gusto di lampone’ patirono pene lancinanti, con lunghe agonie, in alcuni casi per

³Fu anche creata una mostra itinerante, per sensibilizzare l'opinione pubblica, con fotografie e illustrazioni. Questa costituì la base per la stesura del saggio “American Chamber of Horrors,” a cura di Ruth De Forest Lamb, funzionaria dell'*FDA* in prima linea nell'organizzare le azioni dei comitati di consumatori a favore della emanazione di una nuova legge per i farmaci.

più di 20 giorni; purtroppo nessun antidoto fu in grado di contrastare l'avvelenamento.

La madre di una delle vittime, una bambina di 6 anni, scrisse una commovente lettera al presidente degli Stati Uniti.⁴ Roosevelt, profondamente colpito, intervenne sul Congresso, e finalmente si giunse ad una rapida approvazione della proposta di legge. Il *Food, Drug, and Cosmetic Act* fu approvato negli Stati Uniti il 25 giugno del 1938, e stabilì che i produttori dovessero dimostrare la sicurezza dei farmaci prima di poterli commercializzare. Veniva così approvata la prima vera regolamentazione sull'introduzione di nuovi farmaci nel mercato, basata su un semplice, antico principio: *primum non nocere*; se un medicinale deve essere utilizzato sull'uomo, deve anzitutto non nuocere.

I requisiti di sicurezza previsti dal *Food, Drug, and Cosmetic Act* erano però ancora piuttosto blandi. Era inoltre previsto un meccanismo di silenzio-assenso, per cui l'approvazione si considerava acquisita se l'FDA non fosse intervenuta entro 180 giorni dalla domanda. Il risultato era un processo amministrativo molto veloce. Negli anni quaranta e cinquanta – l'epoca d'oro della farmacologia moderna, che vide l'introduzione di farmaci fondamentali come gli antibiotici e i cortisonici – spesso erano sufficienti pochi mesi per passare dalla scoperta di una *New Molecular Entity* (NME) alla commercializzazione del corrispondente farmaco. Da questo punto di vista, il settore farmaceutico non era allora molto diverso da tutti gli altri settori dell'economia. Non c'erano ragioni particolari per estendere la durata dei brevetti farmaceutici.

Le cose cambiarono nel 1962, anche in questo caso a seguito di un evento drammatico, verificatosi però in Europa: la tragedia del Talidomide. Nel 1959 questo farmaco, somminis-

⁴Eccone alcuni brani:

The first time I ever had occasion to call in a doctor for Joan and she was given Elixir of Sulfanilamide. All that is left to us is the caring for her little grave. Even the memory of her is mixed with sorrow for we can see her little body tossing to and fro and hear that little voice screaming with pain and it seems as though it would drive me insane ... It is my plea that you will take steps to prevent such sales of drugs that will take little lives and leave such suffering behind and such a bleak outlook on the future as I have tonight.

trato a donne incinte come antinausea e blando antidepressivo, aveva provocato la nascita di migliaia di bambini con gravi malformazioni agli arti.⁵ Negli Stati Uniti la distribuzione del Talidomide fu impedita grazie all'intervento di Frances Kelsey, una funzionaria dell'FDA che non autorizzò la commercializzazione del farmaco, non ritenendolo sufficientemente sicuro.⁶

Nonostante lo scampato pericolo, le richieste dell'opinione pubblica americana per una regolamentazione più severa aumentarono sensibilmente. A queste richieste il Congresso diede seguito appunto nel 1962 con un emendamento al *Food, Drug, and Cosmetic Act*, noto come il *Kefauver-Harris Amendment*. Il *Kefauver-Harris Amendment* introdusse un nuovo e più rigoroso sistema di regolamentazione, che in breve venne adottato in tutti i principali Paesi sviluppati e che sostanzialmente è quello tuttora in vigore.

1.2.2 IL KEFAUVER-HARRIS AMENDMENT

Il *Kefauver-Harris Amendment* stabilì che per ottenere l'autorizzazione all'immissione in commercio di un nuovo farmaco fosse necessario dimostrarne l'efficacia oltre che la sicurezza, e soprattutto introdusse standard di sicurezza molto più rigorosi che in passato. Di fatto, si creò così il sistema di autorizzazione all'immissione in commercio dei farmaci tuttora in vigore.

Oggi sicurezza ed efficacia dei nuovi farmaci devono essere dimostrate dalle compagnie farmaceutiche sulla base di test pre-clinici e clinici, definiti accuratamente dalle autorità di regolamentazione (l'FDA negli Stati Uniti e l'EMA in Europa).⁷ Le compagnie farmaceutiche devono documentare dettagliatamente anche ogni effetto collaterale emerso durante gli studi clinici.⁸ Le autorità di regolamentazione, una volta che i test sono stati eseguiti, devono

⁵Si stimano circa 8.000 casi. All'epoca, in molti Paesi europei non vi era alcuna regolamentazione dell'introduzione di nuovi farmaci sul mercato.

⁶E che per questo ricevette la 'Distinguished Federal Civilian Service Award' dal Presidente John Kennedy nel 1962.

⁷L'EMA (*European Medical Authority*, fino al 2003 nota come EMEA) è un ente creato nel 1985 dalla Comunità Europea. Esso si affiancò e in larga misura si sostituì alle diverse autorità nazionali preesistenti.

⁸Queste procedure si applicano anche quando emergono nuove indicazioni terapeutiche di farmaci già in precedenza commercializzati. Mentre in passato era possibile l'immissione in commercio

valutarne i risultati, senza restrizioni temporali di sorta, al fine di concedere, nel caso, la ratifica per la vendita.

La sperimentazione dei nuovi farmaci è suddivisa in due fasi, pre-clinica e clinica. La fase pre-clinica non coinvolge soggetti umani e quindi è regolamentata meno strettamente di quella clinica. Se la fase pre-clinica si conclude positivamente, la compagnia farmaceutica elabora una formulazione sperimentale del farmaco (ad esempio in pillole) adatta a condurre i test sugli esseri umani e presenta all'autorità di regolamentazione una richiesta di autorizzazione alla sperimentazione clinica (*Investigational New Drug application*, IND). Questa richiesta deve riportare i risultati dei test pre-clinici che dimostrino le potenzialità del nuovo farmaco. L'autorizzazione viene concessa se l'autorità di regolamentazione non interviene entro 30 giorni dalla richiesta.⁹

Se il nuovo farmaco supera anche la fase clinica, la compagnia farmaceutica presenta la domanda per ottenere la registrazione del farmaco e l'autorizzazione all'immissione in commercio (*New Drug Application*, NDA). La domanda deve essere corredata dai risultati di tutte le sperimentazioni condotte. L'autorità di regolamentazione, che come si è detto non ha scadenze predeterminate, mediamente impiega circa 18 mesi per valutare la domanda e dare eventualmente il via libera alla commercializzazione del nuovo farmaco. Tuttavia in alcuni casi (ad esempio, per i farmaci salva-vita) è prevista una procedura di approvazione abbreviata.

1.2.3 IL DIBATTITO SULLA REGOLAMENTAZIONE

Col *Kefauver-Harris Amendment*, lo sviluppo di un farmaco, a partire dalla scoperta di una nuova sostanza chimica, è diventato un processo molto più lungo, difficile e costoso che in passato. Secondo Di Masi et al. (2003), questo processo richiede mediamente circa 10 anni; semplicemente adottando un nuovo nome commerciale, oggi le nuove indicazioni terapeutiche sono a loro volta sottoposte alla procedura di approvazione e quindi richiedono nuovi, specifici test.

⁹Vi è quindi qui, ma non nella fase dell'approvazione finale, un meccanismo di silenzio-assenso.

stime più recenti sfiorano addirittura i 14 anni (Pammolli et al. 2010).¹⁰ Non vi è dubbio che tempi di sviluppo così lunghi, che si traducono in ritardi nella disponibilità di farmaci importanti, implicino costi sociali significativi.

Inoltre, l'introduzione della nuova e più incisiva regolamentazione è stata seguita da un vistoso calo nel numero di nuovi farmaci immessi sul mercato. Diversi autori hanno attribuito tale calo proprio alla regolamentazione, giudicata eccessiva (Wardel, 1973; Grabowski et al., 1978, Wiggins, 1981 e 1983). Ciò ha generato, negli anni settanta, un vivace dibattito sulla regolamentazione dei nuovi farmaci. In un contributo molto influente, Peltzman (1973) stima costi e benefici sociali del *Kefauver-Harris Amendment*, arrivando alla conclusione che i primi superano di molto i secondi. Le sue stime sono state però aspramente contestate (Mc Guire et al. 1975).

Altri autori hanno sostenuto la capacità del mercato di autoregolarsi. Secondo questo punto di vista, la minaccia dei danni per la responsabilità civile e della perdita di reputazione sarebbe sufficiente a creare incentivi adeguati a testare i nuovi farmaci, su base volontaria, prima di metterli in commercio. È stato anche sostenuto che la decisione di utilizzare un farmaco ancora in fase sperimentale debba essere lasciata al mercato – cioè ai singoli medici e pazienti. Le informazioni che si potrebbero ottenere in questo modo, secondo alcuni, sarebbero più complete rispetto a quelle ottenute con i test regolamentati. Vi sono diverse varianti di queste proposte: per esempio, è stato sostenuto che dovrebbe essere libera la distribuzione di farmaci ancora non approvati almeno ai malati terminali.¹¹

Il dibattito sulla regolamentazione dei nuovi farmaci si è oggi spento, non sia stato raggiunto un consenso sull'una o sull'altra tesi ma perché da tempo non sono state proposte argomentazioni veramente nuove. Tuttavia il problema è ancora largamente aperto, e sarebbe interessante riconsiderarlo alla luce degli sviluppi dell'economia comportamentale e dell'informazione incompleta. Questi temi non rientrano però nei confini di questa tesi.

¹⁰Le nostre stime, presentate nel Capitolo 3, sono più vicini a quelli di Di Masi et al. (2003) che a quelli di Pammolli et al. (2010).

¹¹Si veda la recente (2007) decisione della Corte d'Appello del District of Columbia nel caso *Abigail Alliance v. von Eschenbach*.

1.3 RICERCA E SVILUPPO NEL SETTORE FARMACEUTICO

Il sistema di regolamentazione dei nuovi farmaci implica che la commercializzazione possa avere inizio solo molto tempo dopo che la domanda di brevetto è stata depositata. Infatti, per ragioni che discuteremo meglio tra poco, la domanda per ottenere il brevetto solitamente viene presentata poco dopo la scoperta della nuova sostanza chimica (NME), quando ancora devono essere condotti gran parte dei test pre-clinici.

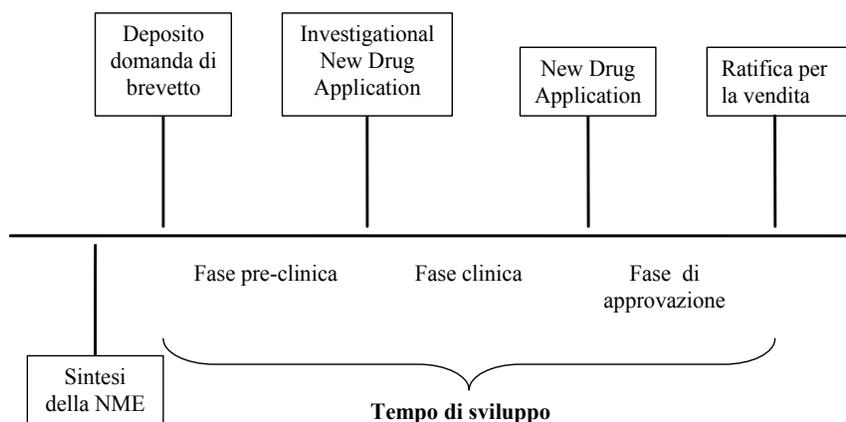


Figura 1.1: Il processo di sviluppo di un nuovo farmaco

Dal momento in cui viene concesso il brevetto (o, più precisamente, dal momento della presentazione della domanda) a quello dell'autorizzazione all'immissione in commercio passano molti anni. Come si è detto, secondo Di Masi et al. (2003), la durata di questa fase, che chiameremo tempo di sviluppo, è mediamente di circa 10 anni.¹² Pammolli et al. (2010), considerando un periodo più recente, trovano che il tempo medio di sviluppo sfiora i 14 anni.¹³ Questo ritardo è illustrato nella Figura 1.1, che rappresenta in forma stilizzata le varie fasi dello sviluppo di un farmaco.

¹²In realtà Di Masi et al. (2003) misurano il tempo di sviluppo a partire dal momento della sintesi della *NME*, non dal momento in cui viene presentata la domanda per il brevetto. Nella misura in cui la domanda viene presentata con un certo ritardo, il loro dato sovrastima il tempo di sviluppo come è definito in questa tesi.

¹³La loro definizione del tempo di sviluppo (“from patent filing to product launch”) coincide con la nostra.

Lo sviluppo di un nuovo farmaco non è solo molto lungo, ma anche molto costoso. La maggior parte delle ingenti spese necessarie per introdurre un nuovo farmaco vengono infatti sostenute proprio in questo stadio, e quindi dopo la scoperta e la brevettazione della *NME*. Inoltre, il tasso di insuccesso nello stadio dello sviluppo è, come vedremo, estremamente elevato.

Descriviamo ora più in dettaglio la “tecnologia” dello sviluppo di un nuovo farmaco. Tale sviluppo può essere suddiviso in quattro fasi: la fase pre-clinica e le tre fasi cliniche.¹⁴

I test pre-clinici, che mediamente durano circa 4 anni, consistono in una serie di sperimentazioni finalizzate alla valutazione dei possibili effetti tossici e sono condotte in un primo momento *in vitro* – cioè in provetta – e successivamente *in vivo* – cioè su animali da laboratorio. Anche se la fase pre-clinica è regolamentata meno strettamente di quella clinica, le sperimentazioni *in vivo* sono comunque soggette ad una normativa che ne limita l’uso allo stretto necessario.¹⁵ Già a partire dal 1959, per tutelare il diritto degli animali a non essere sottoposti a sofferenze immotivate è stata introdotta in tutti i laboratori la regola delle ‘3 R’: ridurre (diminuire al massimo il numero di cavie), raffinare (utilizzando antidolorifici ed anestetici per non prolungare la sofferenza dell’animale) e rimpiazzare (utilizzare quando possibile colture cellulari o altri metodi alternativi, come gli studi che usano il computer per simulazioni, chiamati studi *in silico*). Inoltre le sperimentazioni *in vivo* non sono permesse quando i test che devono essere effettuati siano già stati oggetto di sperimentazioni precedenti. Gli animali più usati sono il topo e il ratto; meno usati sono cani, gatti, criceti, conigli e primati – anche perchè comportano costi molti alti.¹⁶

¹⁴A queste si aggiunge una quarta fase, successiva alla commercializzazione del farmaco, che ai nostri scopi è però meno rilevante: si veda la nota 18.

¹⁵Per l’Italia, ad esempio, si veda il Decreto Legislativo n. 116 del 27/1/1992 (*Attuazione della direttiva CEE n. 609/86 in materia di protezione degli animali utilizzati a fini sperimentali o ad altri fini scientifici*) e la Legge n. 413 del 12/10/1993 (*Norme sull’obiezione di coscienza alla sperimentazione animale*) che regola l’obiezione di coscienza alla sperimentazione animale.

¹⁶Nel 2005 *Nature* ha pubblicato un articolo di critica al metodo di sperimentazione animale, in quanto non predittivo per la specie umana, basato su esempi di farmaci testati su animali rivelatisi poi nocivi per gli esseri umani (Abbott 2005). Nel frattempo stanno aprendosi nuovi scenari di indagine, come la farmacogenomica (che utilizza la biologia molecolare per analizzare le relazioni fra il patrimonio genetico di un individuo e la risposta ad un dato farmaco, aprendo la

Col tempo la regolamentazione dei test clinici da parte dell'FDA e dell'EMA è diventata sempre più minuziosa. Come abbiamo detto, la sperimentazione clinica consta di tre fasi. Nel corso della I fase il farmaco viene somministrato per la prima volta a esseri umani allo scopo di valutarne gli effetti indesiderati e la farmacocinetica, cioè l'assorbimento e il metabolismo della sostanza. Gli esperimenti vengono condotti su volontari sani, in numero da 20 a 80, sotto osservazione medica continua.¹⁷ I volontari devono essere sani per non confondere un eventuale effetto collaterale con il sintomo di una malattia preesistente.

Nella II fase il farmaco viene somministrato a poche centinaia di pazienti per stabilire la dose minima efficace sull'uomo, valutare la posologia ottimale e gli eventuali effetti collaterali. Il farmaco viene messo a confronto con un placebo oppure, nel caso esista, con una terapia standard.

Nella III fase il numero di pazienti reclutato viene sensibilmente aumentato, da varie centinaia fino ad alcune migliaia. Ciò consente di arricchire le informazioni inerenti l'efficacia e la tollerabilità del farmaco, perfezionando dosaggio e posologia. Spesso questa fase impegna ricercatori e medici per diversi anni.¹⁸ La regola di base che governa questa terza fase, del *double-blind testing*, è nota come *gold standard dell'FDA*. In questa tipologia di test, non è solo il paziente che non sa se gli è stato somministrato il farmaco o un placebo, ma anche lo stesso medico che somministra la cura.¹⁹

prospettiva per un uso personalizzato dei farmaci) e la bio-ingegneria (che mette a punto congegni elettronici che riproducono organi umani, come il polmone o il fegato, e possono essere usati per la sperimentazione).

¹⁷Principalmente vengono arruolati volontari di sesso maschile perchè in questo stadio della ricerca non è ancora esclusa la possibilità che la molecola possa avere effetti collaterali teratogeni, in grado di compromettere la salute dei feti.

¹⁸Vi è poi una IV fase, detta di sorveglianza post-marketing, che in questa sede non ci interessa, ma che citiamo per completezza. Essa riguarda i primi anni di vita del farmaco, dopo che è stata concessa l'autorizzazione all'immissione in commercio. In questa fase si valutano le reazioni avverse più rare, che non potevano emergere negli studi clinici. Queste informazioni servono per aggiornare il foglietto illustrativo ed eventualmente, nei casi più gravi, ritirare il farmaco dal commercio.

¹⁹Vi sono anche test in "triplo cieco," nel caso in cui anche chi si occupa della raccolta dei dati statistici non sappia se al paziente sia stato somministrato il farmaco o un placebo. D'altra parte, però vi sono casi in cui la regola aurea non viene fatta valere perché violerebbe il codice deontologico. Si pensi ad esempio ai malati di AIDS: per testare nuovi farmaci, oggi viene fatta una comparazione con farmaci già esistenti; in questi casi, infatti, la somministrazione di un placebo al

Abbiamo detto che il tasso di insuccesso nello stadio dello sviluppo di un nuovo farmaco è estremamente elevato. Di Masi et al. (2003) riportano un tasso di fallimento (detto anche *tasso d'attrito*) nei test clinici dell'80%. Le stime di Pammolli et al. (2010) danno un tasso d'attrito pari a circa il 70% nella fase pre-clinica, e valori comparabili a quelli di Di Masi et al. per la fase clinica. Si può quindi concludere che non più di una molecola brevettata su 20 diventi un farmaco commerciabile.²⁰

1.4 LA DURATA EFFETTIVA DEI BREVETTI FARMACEUTICI

Prima del *Kefauver-Harris Amendment*, la durata effettiva dei brevetti farmaceutici praticamente coincideva con la durata nominale. L'approvazione dei nuovi farmaci era piuttosto rapida, ma d'altra parte anche l'introduzione delle versioni generiche alla scadenza del brevetto era semplice e immediata. Invece, dopo il *Kefauver-Harris Amendment* e il conseguente aumento dei tempi di sviluppo, è naturale che durata effettiva e nominale dei brevetti possano divergere.

1.4.1 LA DURATA EFFETTIVA DOPO IL KEFAUVER-HARRIS AMENDMENT

Come abbiamo visto, il tempo di sviluppo dei nuovi farmaci aumentò di fatto considerevolmente dopo il *Kefauver-Harris Amendment*. Tuttavia, il *Kefauver-Harris Amendment* stabilì anche che i farmaci generici non potessero essere immessi sul mercato senza superare test simili a quelli previsti per i nuovi farmaci (al contrario di quanto avveniva precedentemente).

posto di un farmaco potrebbe essere letale.

²⁰Molti autori, basandosi su una stima fornita dalla stessa FDA, affermano che solo una nuova sostanza chimica su 5 o 10 mila molecole testate diventa un farmaco commerciabile. Tuttavia, questa stima include anche le molecole per cui non vi è prova di alcuna efficacia terapeutica, e che quindi non possono essere brevettate. Per le molecole brevettate, il tasso di insuccesso è in realtà molto minore, pur rimanendo molto alto.

poiché tali test non potevano iniziare prima della scadenza del brevetto sul nuovo farmaco,²¹ di fatto la durata nominale dei brevetti farmaceutici veniva in qualche modo ristabilita.

Tuttavia, il tempo di sviluppo dei farmaci generici non coincide necessariamente con quello del farmaco originario: anzi, a ben vedere, è indipendente da quest'ultimo. (Chiaramente, i tempi di sviluppo del farmaco originale e del farmaco generico sono correlati poiché entrambi dipendono da una serie di variabili esogene comuni. Ma se l'inventore riesce, in qualche modo, a ridurre il tempo di sviluppo del farmaco originario, questo non implica che il tempo di sviluppo del farmaco generico si riduca automaticamente.) Questo implica una relazione tra la durata effettiva del brevetto e il tempo di sviluppo del farmaco come quella illustrata nella Figura 1.2: si tratta semplicemente di una retta con una pendenza uguale a -1 . L'intercetta della retta è pari alla somma della durata nominale del brevetto e del tempo di sviluppo del farmaco generico. Se il tempo di sviluppo del farmaco originale coincide con quello del generico, la durata effettiva del brevetto è uguale a quella nominale. Ma i due tempi di sviluppo possono divergere, e ogni riduzione del tempo di sviluppo del farmaco originale si traduce in un aumento di ugual durata della vita effettiva del brevetto.

In pratica, prima del 1984 la situazione dell'innovatore era ancora migliore di quanto possa apparire dalla Figura 1.2, in quanto lo svolgimento dei test necessari per l'introduzione del farmaco generico richiedeva non solo molto tempo ma anche ingenti investimenti. Dato che i produttori di generici avrebbero dovuto comunque fronteggiare la concorrenza dell'innovatore, non sempre questi investimenti si prospettavano redditizi. Molti farmaci brevettati

²¹Secondo la giurisprudenza, infatti, la conduzione dei test non può godere della c.d. *experimental use defence*, costituendo sfruttamento dell'innovazione a fini commerciali che in quanto tale rappresenta un'infrazione del brevetto in assenza di un'autorizzazione del titolare. Questa interpretazione trova la sua classica formulazione nella decisione sul caso *Roche v. Bolar*. Roche, produttore del farmaco Dalmane (noto anche come Valium), fece causa a Bolar, un produttore di farmaci generici, che aveva cominciato, prima della scadenza brevettuale, a condurre i test per immettere sul mercato la versione generica del Valium. Bolar sostenne che l'uso del principio attivo brevettato non costituiva un'infrazione del brevetto, poiché era sperimentale. Invece il tribunale ribadì che l'uso che Bolar faceva del principio attivo di Dalmane nella conduzione dei test non era "for amusement, to satisfy idle curiosity, or for strictly philosophical inquiry," ma aveva chiari scopi commerciali. Come vedremo, proprio questa decisione indusse il Congresso a introdurre nel giro di pochi mesi, con l'*Hatch-Waxman Act* del 1984, un'esenzione esplicita nota come *Bolar exemption*.

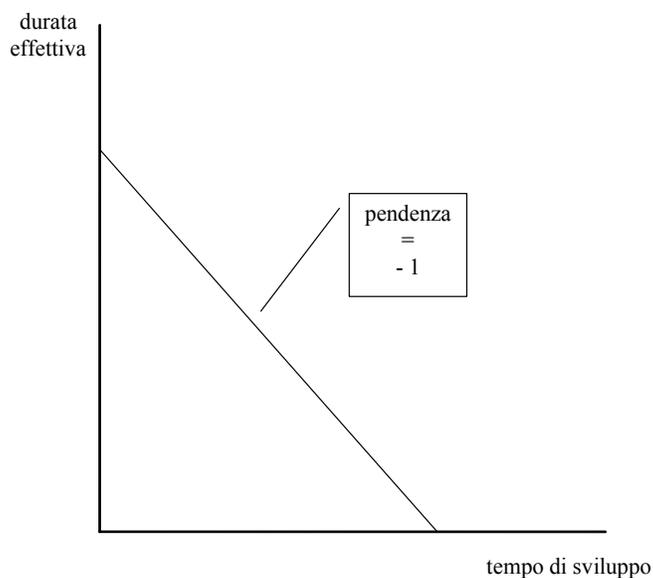


Figura 1.2: Durata effettiva e tempo di sviluppo dopo il *Kefauver-Harris Amendment*.

continuavano quindi a godere di una posizione di monopolio anche molto tempo dopo la scadenza del brevetto. È in questa situazione che intervenne l'*Hatch-Waxman Act*.

1.4.2 L'INCENTIVO A BREVETTARE RAPIDAMENTE

Prima di descrivere il sistema legislativo creato dall'*Hatch-Waxman Act* del 1984, è opportuno chiedersi perchè la domanda per ottenere il brevetto venga solitamente presentata subito dopo la scoperta di una NME.

Infatti, i problemi che sono al centro di questa tesi non sorgerebbero se il brevetto fosse richiesto solo nel momento in cui il farmaco viene autorizzato al commercio.

Come notano Dranove e Meltzer (1994), il fatto che i tempi di sviluppo siano così lunghi crea un incentivo a ritardare la richiesta del brevetto:

“it seems plausible that the long time to approval might provide an incentive for firms to delay patent application somewhat in order to postpone the date of

patent expiration” (p. 405).

Tuttavia, questi autori proseguono osservando:

“We raised this issue with several individuals within the pharmaceutical industry and found no one who was aware of such behavior” (*ibidem*).

In realtà, vi è una sia pur limitata giurisprudenza relativa a casi in cui la domanda per ottenere il brevetto è stata presentata dopo l’inizio dei test clinici – il che fa pensare che questa opzione sia ben presente alle compagnie farmaceutiche. Perché allora queste di solito preferiscono brevettare rapidamente, rinunciando apparentemente ad una interessante opportunità strategica? Una prima spiegazione è il rischio che l’innovazione venga duplicata indipendentemente da un secondo inventore.²² Poiché il secondo inventore può brevettare prima del primo, questi nella migliore delle ipotesi dovrebbe competere con un concorrente, e in assenza di “diritti di pre-uso” perderebbe addirittura la possibilità di commercializzare il nuovo farmaco.²³

Una seconda spiegazione è che il sistema di brevetti è disegnato in modo da indurre l’innovatore a brevettare rapidamente – una caratteristica che sta alla base della “prospect theory” di Kitch (1977). Vi sono diverse regole legali che spingono in questa direzione: in particolare, quella che nega la possibilità di brevettare una qualunque invenzione che viene usata in pubblico prima che sia presentata la domanda per ottenere il brevetto. (Tecnicamente, il fatto che l’innovazione sia già stata usata in pubblico prima del deposito della domanda per il brevetto fa venir meno il requisito della *novità*.) Secondo la giurisprudenza, la conduzione dei test clinici può essere considerata come un “uso pubblico” dell’innovazione,

²²C’è anche, e forse è ancora più importante, il rischio di spionaggio industriale. Ad esempio, un dipendente dell’impresa che ha scoperto la NME potrebbe trasferirsi presso un concorrente e trasmettergli l’informazione, o semplicemente vendergliela. In teoria questi comportamenti sono illegali e il soggetto entrato in possesso dell’informazione in questa maniera non potrebbe brevettare. In pratica, però, sarebbe difficile dimostrare che l’invenzione non sia stata riscoperta in modo indipendente.

²³Sul diritto a brevettare del secondo inventore e i diritti residuali del primo si veda Denicolò e Franzoni (2004) e Shapiro (2006). I diritti di pre-uso sono da tempo previsti nelle legislazioni europee, ma sono stati introdotti negli Stati Uniti solo nel 2011.

e in quanto tale ne preclude la successiva brevettabilità.²⁴ Si capisce allora perchè raramente le compagnie farmaceutiche rinviino la presentazione della domanda per ottenere il brevetto.

Naturalmente, sarebbe facile immaginare riforme legislative che consentano all'inventore di posticipare la domanda fino a quando il farmaco non sia stato autorizzato. Tuttavia, non è affatto scontato che il risultato – cioè una durata effettiva del brevetto farmaceutico indipendente dal tempo di sviluppo – sia socialmente desiderabile, come mostreremo nei capitoli 4 e 5.

1.4.3 L'HATCH-WAXMAN ACT

Lo scopo primario dell'*Hatch-Waxman Act* del 1984²⁵ era di favorire l'entrata sul mercato dei farmaci generici,²⁶ evitando la ripetizione di test molto costosi che nella maggior parte dei casi riproducevano informazioni già acquisite dall'innovatore.²⁷ Esso ha quindi previsto che per ottenere la ratifica per la vendita di un farmaco generico sia sufficiente dimostrarne

²⁴Si veda da ultimo la decisione del Southern District of New York nel caso *Dey v. Sunovion* (1 marzo 2012). In realtà, la giurisprudenza sul punto non è sempre univoca. Il discriminante sembra essere il fatto che l'innovazione possa essere considerata “perfezionata” prima dell'inizio dei test (nel qual caso la conduzione dei test precluderebbe la brevettabilità) o no. Per esempio, se il brevetto copre una nuova sostanza chimica, l'innovazione si deve considerare perfezionata nel momento in cui la nuova sostanza viene sintetizzata, e quindi la conduzione dei test rappresenta un “uso pubblico” dell'innovazione. Ma se il brevetto copre un nuovo involucro che consente al principio attivo di essere rilasciato solo in una certa parte dell'organismo, è ragionevole considerare che l'innovazione non sia ancora perfezionata fino a quando i test clinici non hanno consentito di dimostrare come avvenga il rilascio della sostanza. Nel caso *Bayer v. Barr Labs*, ad esempio, la conduzione dei test non fu considerata “uso pubblico” e il brevetto (la cui domanda era stata presentata dopo l'inizio dei test) fu ritenuto valido dal tribunale.

²⁵Public Law 98-417 (*Drug and Price Competition and Patent Term Restoration Act*). Per una analisi legale dell'*Hatch-Waxman Act* si veda, tra gli altri, Bagley (2011), Mossinghoff (1999), Ohly (2010) e Roin (2010).

²⁶Questa legge ebbe un lungo periodo di gestazione: infatti la prima proposta che contemplava l'esigenza di recuperare il tempo perso nel periodo di sviluppo, venne formulata da un gruppo di esperti di politica industriale nominato dal Presidente Carter nel 1978.

²⁷Non vanno inoltre dimenticati i problemi etici posti dalla ripetizione dei test clinici. Il trattamento di alcuni pazienti con un placebo in questi test può infatti essere giustificato quando l'efficacia e sicurezza di una terapia non sono ancora state dimostrate, ma diventa problematico quando lo sono.

l'equivalenza terapeutica con un farmaco già ratificato,²⁸ e ha introdotto la cosiddetta *Bolar exemption*, che consente di condurre i test del farmaco generico anche prima della scadenza del brevetto.²⁹ Queste riforme hanno contribuito in modo determinante alla crescita della quota di mercato dei farmaci generici, che è passata da meno del 30% prima del 1984 a oltre il 70% in anni recenti.

Facendo in modo che il monopolio dell'innovatore terminasse alla scadenza del brevetto, però, queste riforme implicavano che la durata effettiva del monopolio si riducesse considerevolmente. Nei termini della Figura 1.2, il tempo di sviluppo del farmaco generico viene di fatto azzerato, e quindi la retta si sposta verso il basso dato che l'intercetta diventa pari solamente alla durata nominale. In questo modo la durata effettiva può diventare molto breve, in alcuni casi addirittura nulla.

Per consentire all'innovatore di recuperare almeno una parte del tempo di sviluppo e garantire una durata minima dell'esclusiva, l'*Hatch-Waxman Act* ha allora previsto la possibilità di estendere il periodo di monopolio di cui gode l'innovatore. L'estensione è determinata da due meccanismi. Il primo è costituito dall'estensione vera e propria del termine del brevetto per un periodo pari ad una frazione del tempo di sviluppo. Il secondo opera attraverso la protezione dei dati ottenuti nei test pre-clinici e clinici (la cosiddetta *data exclusivity*), che impedisce di presentare una domanda di approvazione abbreviata (si veda la nota 28) prima di un certo lasso di tempo dall'approvazione del farmaco originale. Il produttore di farmaci generici che volesse commercializzare il farmaco prima della scadenza del periodo

²⁸Nel sistema di approvazione dei farmaci generici creato dall'*Hatch-Waxman Act*, il produttore del generico può presentare all'FDA una *Abbreviated New Drug Application* (ANDA) che richiede la sola prova dell'equivalenza terapeutica. (Naturalmente il produttore del generico ha sempre la possibilità di seguire la strada normale, che però è molto più lunga e costosa.) Per dimostrare l'equivalenza terapeutica occorre provare che il farmaco generico è bioequivalente a quello già approvato, cioè opera attraverso lo stesso meccanismo. Di fatto occorre dimostrare che la concentrazione nel sangue del principio attivo somministrato col farmaco generico è simile a quella del farmaco originario.

²⁹Il termine *Bolar exemption* deriva dalla sentenza che, come abbiamo visto nella nota 21, aveva ribadito che in assenza di una normativa *ad hoc* la conduzione dei test clinici non potesse godere della *experimental use defence*.

di *data exclusivity* dovrebbe quindi ripetere i test, proprio come avveniva prima dell'*Hatch-Waxman Act*. Dato che questo è molto costoso, i produttori di farmaci generici preferiscono aspettare che scada il periodo di *data exclusivity*. Questo meccanismo in pratica garantisce all'innovatore una durata minima del monopolio.

In termini quantitativi, negli Stati Uniti il brevetto può essere esteso di un periodo pari alla metà della durata della fase clinica più l'intera "fase di approvazione" (cioè il tempo impiegato dall'FDA per valutare i risultati dei test e concedere l'autorizzazione all'immissione in commercio). Vi sono però due tetti all'estensione: la durata effettiva del brevetto non può, per effetto dell'estensione, superare i 14 anni, e l'estensione stessa non può superare i 5 anni. La durata della *data exclusivity* è stata poi fissata a 5 anni, anche se recentemente è stata portata a 12 anni per i cosiddetti "farmaci biologici."³⁰

Il sistema che abbiamo appena descritto implica che la durata effettiva del brevetto, che indichiamo con il simbolo T , dipenda in maniera complicata dal tempo di sviluppo, τ , come mostrato nella Figura 1.3.³¹ Inizialmente, cioè per valori del tempo di sviluppo non superiori a 6 anni, il tetto dei 14 anni implica che non possa essere prevista alcuna estensione della durata nominale del brevetto. Quindi in questo tratto la curva è una retta con una pendenza pari a -1 . Quando il tempo di sviluppo supera i 6 anni, l'estensione diventa positiva. Essa però è ancora determinata dal tetto dei 14 anni sulla durata effettiva: la curva $T(\tau)$ quindi si appiattisce e diventa una retta orizzontale.

Quando il tempo di sviluppo aumenta ancora, sono possibili due casi. Se nessuno dei due tetti è operativo, l'estensione è pari a una certa frazione del tempo di sviluppo. Graficamente, questo corrisponde alla retta inclinata negativamente ma con una pendenza superiore a -1 della Figura 1.3 (il valore preciso della pendenza dipende dalla suddivisione del tempo di

³⁰Oltre alla *biologic data exclusivity*, esistono anche altre esclusive speciali, come la *pediatric exclusivity*, di 6 mesi di durata, che viene concessa quando un farmaco ottiene l'approvazione all'uso pediatrico, e la *orphan drug exclusivity*, di 7 anni, che si applica a farmaci per malattie rare.

³¹Si noti che la durata effettiva sarebbe costante se l'estensione della durata del brevetto coprisse l'intero tempo di sviluppo, oppure, come abbiamo discusso nella precedente sottosezione, se il termine della protezione brevettuale decorresse dal momento in cui viene concessa la ratifica per la vendita.

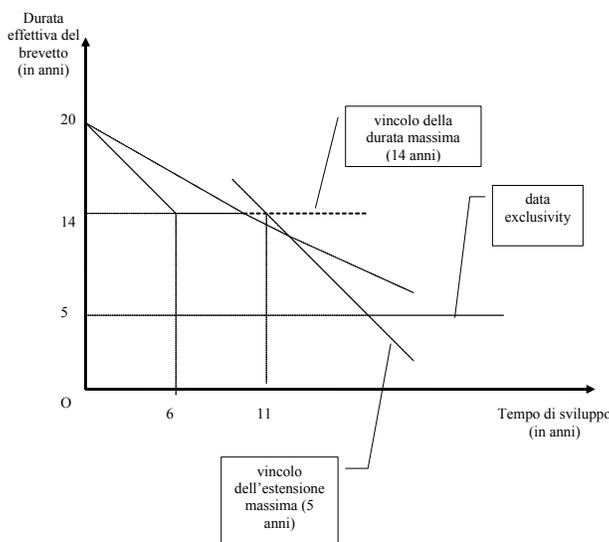


Figura 1.3: La durata effettiva del brevetto in funzione del tempo di sviluppo negli Stati Uniti

sviluppo tra fase pre-clinica, clinica e di approvazione). Se invece è operativo il tetto dei 5 anni all'estensione complessiva, subentra un tratto in cui la pendenza della curva $T(\tau)$ torna a essere -1 . Se anche inizialmente il tetto dei 5 anni non è operativo,³² è chiaro che lo deve diventare quando il tempo di sviluppo aumenta sufficientemente.

Quando infine il tempo di sviluppo è ancora maggiore, la durata effettiva del monopolio viene garantita esclusivamente dal meccanismo di *data exclusivity* e quindi T è di nuovo costante.

1.4.4 IL SISTEMA EUROPEO

Il sistema introdotto negli Stati Uniti con l'*Hatch-Waxman Act* è stato successivamente adottato in Europa, Giappone, Australia, Canada e quasi tutti gli altri Paesi industrializzati.

³²Supponiamo, per esempio, che il tempo di sviluppo complessivo sia di 11 anni, di cui 4 anni nella fase pre-clinica, 6 nella fase clinica e 1 anno per l'approvazione da parte dell'FDA. In questo caso nessuno dei due tetti è stringente, e l'estensione del termine del brevetto sarà di 4 anni per una durata effettiva di 13.

In particolare, in Europa ogni paese ha dapprima adottato una propria legislazione nazionale. Queste diverse legislazioni, pur nella sostanziale omogeneità dell'architettura di base, prevedevano termini e condizioni differenti. Per armonizzarle, l'Unione Europea ha emanato un proprio regolamento (Regolamento del Consiglio n. 1768 del 1992), che è stato poi incorporato nelle varie legislazioni nazionali. Recentemente la regolamentazione europea del 1992 è stata emendata dal Regolamento n. 469 del 2009.

Il sistema europeo è leggermente diverso da quella degli Stati Uniti. Dal punto di vista formale, in Europa non viene prevista un'estensione del termine del brevetto, ma sono stati creati i cosiddetti “Certificati di Protezione Complementare” (*Supplementary Protection Certificates*) che costituiscono un titolo giuridico distinto dal brevetto originario. Dal punto di vista sostanziale, la durata dell'estensione è diversa e più semplice da calcolare rispetto agli Stati Uniti. Per la precisione, essa è pari al periodo di tempo intercorso tra la data del deposito della domanda per ottenere il brevetto, aumentata di 5 anni, e la data dell'autorizzazione all'immissione in commercio.³³ Tuttavia, come negli Stati Uniti, l'estensione non può superare i 5 anni. Questo significa che la durata della protezione supplementare è pari al tempo di sviluppo eccedente i 5 anni, raggiungendo un massimo di 5 anni se il tempo di sviluppo è di 10 anni; poi, rimane costante. Nella maggior parte dei paesi europei, inoltre, la durata della *data exclusivity* è di 10 anni contro i 5 degli Stati Uniti.³⁴

La Figura 1.4 rappresenta in forma stilizzata la politica attualmente adottata in Europa. Come negli Stati Uniti, la curva di equazione $T(\tau)$ inizialmente ha una pendenza pari a -1 , non essendo prevista alcuna estensione della durata nominale del brevetto. La curva si appiattisce quando il tempo di sviluppo è superiore a 5 anni, per poi tornare a calare con

³³In altri termini, l'estensione è pari al tempo di sviluppo meno 5 anni. Ricordiamo che il tempo di sviluppo è qui inteso come il periodo dalla presentazione della domanda per ottenere il brevetto (che è il termine dai cui si comincia a calcolare la durata del brevetto) alla ratifica per la vendita. Naturalmente l'estensione è nulla se il tempo di sviluppo è inferiore a 5 anni.

³⁴In effetti, negli Stati Uniti sono state avanzate numerose proposte per armonizzare la durata della *data exclusivity* ai livelli europei (o addirittura estenderla a 12 o 14 anni). Al momento, però, l'unica estensione approvata riguarda i cosiddetti farmaci biologici (cioè i farmaci che sono sintetizzati da organismi viventi, come ad esempio l'eritropoietina), per i quali la *data exclusivity* è stata portata a 12 anni.

una pendenza pari a -1 quando τ supera i 10 anni. Infine, quando il tempo di sviluppo eccede i 15 anni, la durata effettiva viene garantita esclusivamente dal meccanismo di *data exclusivity* e quindi è di nuovo costante.

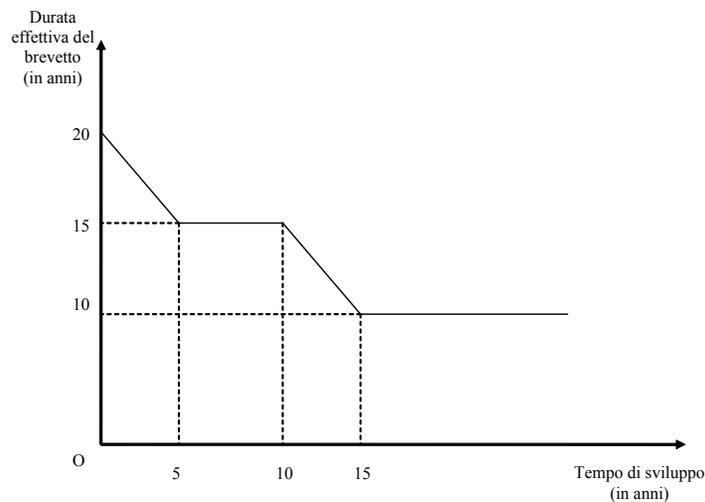


Figura 1.4: La durata effettiva del brevetto in funzione del tempo di sviluppo in Europa

1.4.5 IL SISTEMA ITALIANO

Dal 1992 il sistema descritto nella precedente sottosezione si applica anche in Italia. In precedenza, tuttavia, il Parlamento aveva approvato un sistema straordinariamente generoso (istituito con la Legge 349 approvata nel Novembre 1991), che è illustrato nella Figura 1.5. In breve, il sistema prevedeva il recupero dell'intero tempo di sviluppo fino ad un massimo di 18 anni.

L'estensione prevista era quindi sensibilmente maggiore rispetto a quella garantita dalle regole poi adottate a livello europeo. Si può pensare che poiché comunque la legislazione italiana è rimasta in vigore per meno di due anni, i suoi effetti devono essere stati limitati. Ma non è così. Senza alcuna vera giustificazione economica, infatti, il sistema è stato applicato retroattivamente. Nei pochi mesi in cui è rimasta in vigore, la Legge 349/91 ha quindi assicurato l'estensione della durata di oltre quattrocento brevetti, in una misura che non

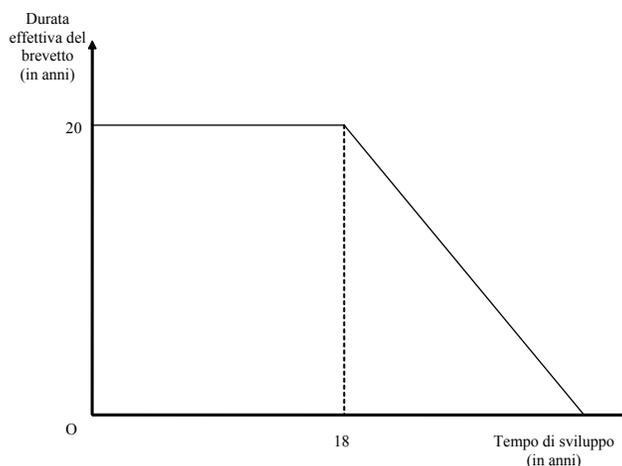


Figura 1.5: Il sistema dei Certificati di Protezione Complementare in Italia

ha alcun confronto né in Europa né negli Stati Uniti (dove per i farmaci già brevettati al momento della sua entrata in vigore l'*Hatch-Waxman Act* prevedeva un'estensione massima di soli due anni). Nel 2002 e poi nel 2006 il governo italiano ha gradualmente ridotto la durata dei *Certificati di Protezione Complementare* concessi nel 1991-92: ma ormai la maggior parte di questi era già scaduta.

Sarebbe interessante ricostruire le vicende che hanno condotto all'approvazione della Legge 349/91. Forse una lobby potente è semplicemente riuscita ad irretire un Parlamento distratto. Ma qualche dubbio è lecito, e solo gli storici potranno eventualmente fugarlo.

1.5 CONCLUSIONI

In questo capitolo abbiamo discusso le caratteristiche peculiari del processo innovativo nel settore farmaceutico, che implicano che la durata effettiva dei brevetti farmaceutici non coincida con quella nominale (attualmente pari a 20 anni). Abbiamo visto che il tempo di sviluppo (cioè il tempo che passa tra il deposito della domanda di brevetto e la commercializzazione del nuovo farmaco) è molto lungo per via della specifica regolamentazione a

cui è soggetta l'introduzione di nuovi farmaci, e che per compensare i titolari di brevetti farmaceutici una legislazione *ad hoc* estende il periodo di esclusiva di cui essi godono. Negli Stati Uniti, questa legislazione è stata introdotta nel 1984 dall'*Hatch-Waxman Act*; regole analoghe sono poi state emanate in Europa e nella maggior parte dei Paesi industrializzati. Per dare un'idea dell'importanza di queste politiche, basti dire che secondo Clift (2008) l'estensione della durata del brevetto negli Stati Uniti garantiva quasi il 40% dei ricavi delle vendite per i 40 farmaci più venduti nel 2006.

Da questo capitolo possiamo trarre due conclusioni principali, che costituiranno il punto di partenza della nostra analisi nei capitoli successivi. La prima è che l'invenzione di un nuovo farmaco non si conclude nel momento in cui viene concesso il brevetto: questo è solo il punto di partenza di una fase successiva, la fase dello sviluppo, che è almeno altrettanto importante di quella che la precede. La seconda conclusione è che la politica della durata dei brevetti farmaceutici non si può riassumere in uno scalare, ma è data da una funzione che lega la durata effettiva del brevetto al tempo di sviluppo del farmaco.

CAPITOLO 2

IL MODELLO TEORICO

Riassunto. Questo capitolo propone un modello teorico del processo innovativo nel settore farmaceutico che cattura gli aspetti istituzionali descritti nel capitolo introduttivo. Il modello permette di scomporre l'incentivo ad investire nella fase di sviluppo di un nuovo farmaco in due componenti: l'effetto di anticipazione dei profitti e l'effetto di durata effettiva del brevetto. L'analisi genera diversi risultati di statica comparata, alcuni dei quali verranno testati empiricamente nel capitolo 3, e serve da base per lo studio della politica ottimale che svilupperemo nei capitoli 4 e 5.

2.1 INTRODUZIONE

In questo capitolo sviluppiamo un modello teorico del processo innovativo nel settore farmaceutico che incorpora in forma stilizzata gli aspetti istituzionali discussi nel capitolo introduttivo. Il modello servirà come base per l'analisi empirica del capitolo 3 e per lo studio della politica ottimale nei capitoli 4 e 5.

A partire da Loury (1979), Lee e Wilde (1980) e Dasgupta e Stiglitz (1980), i modelli standard di *patent race* ipotizzano che il processo innovativo si concluda nel momento stesso in cui viene ottenuta l'innovazione, che viene poi protetta da un brevetto e commercializzata. Ma nel settore farmaceutico, come abbiamo visto, gran parte delle spese in R&D vengono in realtà sostenute dopo il conseguimento del brevetto, e servono per dimostrare l'efficacia e la

sicurezza della nuova sostanza chimica (*New Molecular Entity*, NME). Pertanto, il processo innovativo non può certo considerarsi concluso con l'ottenimento del brevetto. poiché nello sviluppo di una NME già brevettata l'innovatore gode, ovviamente, di una posizione di monopolio, la competizione tecnologica è però limitata alla fase della ricerca di una nuova sostanza.

È quindi essenziale distinguere due stadi della ricerca farmaceutica: lo *stadio della ricerca*, in cui diverse imprese competono per scoprire una NME dotata di effetti terapeutici, che viene brevettata dalla prima impresa che ha successo; e lo *stadio dello sviluppo*, in cui il detentore del brevetto investe nei test pre-clinici e clinici necessari per ottenere la ratifica per la vendita. Solo se la NME supera questi test, il titolare del brevetto può iniziare a commercializzare il nuovo farmaco.

2.2 IPOTESI

Il modello prevede dunque due stadi, denominati rispettivamente ricerca e sviluppo, caratterizzati da competizione (nella ricerca) e monopolio (nello sviluppo). A queste due fasi sono anche associate due diverse fonti di incertezza nel processo innovativo.

Nella fase della ricerca, la scoperta di una NME è ovviamente un evento incerto. Supponiamo che tale evento si realizzi con una probabilità θ che dipende dall'investimento aggregato effettuato dalle imprese che competono per il brevetto.

Nella fase dello sviluppo, vi è poi incertezza sulla possibilità che la NME si riveli efficace e sicura – un evento la cui probabilità è indicata da p . Per semplicità supponiamo che l'investimento nello stadio dello sviluppo non determini la probabilità p , che prendiamo come un parametro esogeno.¹ L'investimento in sviluppo determina invece quanto tempo è necessario per scoprire se la nuova sostanza chimica sia efficace e sicura. Come abbiamo visto nel primo capitolo, questo periodo di tempo, che chiamiamo tempo di sviluppo e indichiamo con τ , è di durata considerevole, mediamente superiore a 10 anni.

¹Come abbiamo visto nel Capitolo 1, valori realistici di p sono dell'ordine del 5% o meno.

Descriviamo ora più in dettaglio le nostre ipotesi sulla tecnologia di R&D nelle fasi della ricerca e dello sviluppo del farmaco.

2.2.1 LO STADIO DELLA RICERCA

Nello stadio della ricerca, diverse imprese gareggiano per scoprire e brevettare una NME. Assumiamo libertà di entrata, il che implica che i profitti attesi partecipando a questa gara debbano essere pari a zero.

I nostri risultati non dipendono dal modo particolare in cui viene rappresentata la “corsa al brevetto.” Possiamo quindi ricorrere a un modello in forma ridotta, che comprende come casi particolari diversi modelli di uso comune nella letteratura. Indichiamo con $\theta(R)$ la probabilità che una qualche impresa abbia successo, e assumiamo che tale probabilità dipenda dall’investimento aggregato in ricerca R . Nell’interpretazione letterale della forma ridotta, tutto l’investimento R è effettuato al tempo 0 e l’incertezza è risolta istantaneamente. In realtà, però, la nostra formulazione è formalmente equivalente ad un modello in cui il *timing* dell’innovazione è incerto e segue un processo di Poisson, come ipotizzato ad esempio da Loury (1979) e Lee e Wilde (1980).²

È naturale assumere che $\theta(0) = 0$ (non vi è possibilità di successo senza investimento in ricerca) e che la funzione $\theta(R)$ sia crescente. Ipotizziamo anche che $\theta(R)$ sia differenziabile e concava. L’ipotesi di concavità riflette l’idea di rendimenti decrescenti nella ricerca a livello di industria.³ L’elasticità della funzione $\theta(R)$,

$$\eta = \frac{R\theta'(R)}{\theta},$$

rappresenta l’incremento percentuale della probabilità di successo dovuto ad un aumento dell’1% dell’investimento in ricerca. Dato che $\theta(R)$ è concava e $\theta(0) = 0$, si dimostra facilmente che questa elasticità (nota nella letteratura come *elasticità dell’offerta di innovazioni*)

²Per la spiegazione dettagliata dell’equivalenza formale si veda Denicolò (2007).

³Infatti, la ricerca potrebbe richiedere input la cui offerta, almeno nel breve periodo, è fissa, come ad esempio ricercatori di talento o progetti di ricerca promettenti.

è necessariamente minore di uno.

2.2.2 LO STADIO DELLO SVILUPPO

Consideriamo ora lo stadio dello sviluppo. La differenza principale rispetto allo stadio della ricerca è che ora non vi è più competizione tecnologica: solo il detentore del brevetto può investire. Come già anticipato, assumiamo che la probabilità che la NME si riveli efficace e sicura, p , sia fissa. Tuttavia, le scelte di investimento in sviluppo del detentore del brevetto influenzano il tempo necessario per scoprire se la NME sia efficace e sicura. In altri termini, il tempo di sviluppo τ è determinato dall'investimento nella fase dello sviluppo, D .

Si potrebbe obiettare che il tempo di sviluppo è un parametro puramente tecnologico che non può essere influenzato dall'investimento dell'innovatore. In realtà, ci sono varie ragioni per cui il tempo di sviluppo può dipendere da tale investimento. Per esempio, i test pre-clinici e clinici possono essere condotti sequenzialmente o simultaneamente. Poniamo che una nuova molecola debba essere testata su 1000 pazienti, e che il trattamento di ogni paziente richieda un costo e un tempo dati. Il titolare del brevetto potrebbe testare i pazienti simultaneamente, minimizzando la durata complessiva dei test, oppure in sequenza, diciamo a gruppi di 100, iniziando da un gruppo e procedendo a quello successivo solo una volta completati i test sul primo gruppo. In quest'ultimo caso, il tempo necessario a completare con successo il test sarebbe 10 volte superiore, ma in media i costi sarebbero più bassi perchè il test potrebbe essere sospeso non appena diventasse sufficientemente probabile un esito negativo.⁴ Variando gradualmente la misura in cui i test vengono condotti sequenzialmente piuttosto che simultaneamente, il tempo di sviluppo può diminuire gradualmente al crescere dell'investimento.⁵

⁴Questo esempio suggerisce che, dopo tutto, l'investimento in sviluppo potrebbe influenzare anche la probabilità di approvazione p . Infatti, un farmaco potenzialmente valido potrebbe essere abbandonato prematuramente se il titolare del brevetto, non essendo disposto a sostenere il rischio di investimenti ingenti, interrompesse i test prima di raggiungere la certezza assoluta di un fallimento.

⁵Questo esempio non è puramente teorico. In pratica, le compagnie farmaceutiche devono sempre decidere in quanti centri (ospedali, cliniche universitarie ecc.) condurre la sperimentazione. Poichè

D'altra parte, che le scelte di investimento delle compagnie farmaceutiche abbiano un peso non trascurabile nella determinazione del tempo di sviluppo τ è confermato empiricamente. Un'ampia evidenza empirica mostra infatti come i tempi di sviluppo dei farmaci di maggior valore siano mediamente inferiori alla media. Ciò non sarebbe possibile se il tempo di sviluppo fosse un parametro esogeno.

Il lavoro più significativo al riguardo è quello di Dranove e Meltzer (1994).⁶ Analizzando un campione di 564 farmaci e considerando diverse misure dell'importanza del farmaco (dal volume delle vendite alle citazioni nella letteratura medica e nelle successive domande di brevetto, alla velocità della diffusione geografica), questi autori concludono che “more important drugs are developed and approved more rapidly than less important drugs” (p. 402); che “importance generally affects both phases [i.e., pre-clinical and clinical] of the drug development process” (p. 418); e che “the pharmaceutical manufacturers, rather than the FDA, are more responsible for relative acceleration of important drugs” (p. 421). Quantitativamente, Dranove e Meltzer (1994) stimano che un farmaco il cui valore eccede il valore medio di due volte la deviazione standard ha un tempo di sviluppo del 25% inferiore a quello medio.

Sulla base di considerazioni teoriche e dell'evidenza empirica, ipotizziamo dunque che il tempo di sviluppo τ dipenda dall'investimento in sviluppo, D , secondo una relazione decrescente $\tau(D)$. Supponiamo che la funzione $\tau(D)$ sia differenziabile e convessa, con $\lim_{D \rightarrow 0} \tau(D) = \infty$ e $\lim_{D \rightarrow \infty} \tau(D) = \bar{\tau}$, come illustrato nella Figura 2.1. In altre parole, se il detentore del brevetto non investe, l'incertezza non sarà mai risolta, mentre d'altra parte vi è una durata minima del tempo di sviluppo, $\bar{\tau}$, sotto la quale non si può scendere per quanto grande sia l'investimento.⁷

il tasso di arrivo di pazienti affetti da una certa patologia presso ogni centro di sperimentazione è un dato largamente esogeno, usare un numero più o meno grande di centri di sperimentazione equivale a condurre la ricerca in misura più o meno sequenziale.

⁶Rinviamo al capitolo 3 per una rassegna più ampia della letteratura su questo punto.

⁷Ove fosse necessario, assumeremo che $\tau(D)$ sia sufficientemente convessa da garantire che siano sempre soddisfatte le condizioni del secondo ordine nei problemi di massimizzazione considerati in seguito.

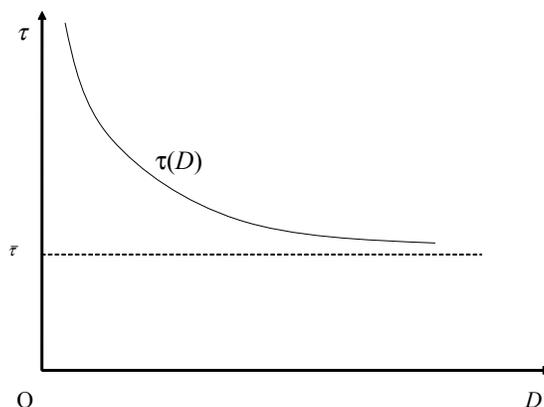


Figura 2.1: Il tempo di sviluppo τ in funzione dell'investimento nello stadio dello sviluppo D .

2.2.3 IL MERCATO DEL PRODOTTO

Per semplicità supponiamo che il mercato del prodotto sia stazionario, ma non sarebbe difficile estendere l'analisi al caso in cui domanda e costi variano nel tempo.

Anche se la curva di domanda e il costo marginale di produzione del nuovo farmaco sono costanti nel tempo, l'equilibrio nel mercato del prodotto cambia radicalmente quando scade il brevetto e vengono introdotte sul mercato le versioni generiche del nuovo farmaco. È un fatto ben documentato che la competizione dei farmaci generici erode in misura sensibile i profitti dell'innovatore.⁸ È altrettanto ben documentato che l'innovatore continua a godere di un certo potere di mercato anche dopo la scadenza del brevetto,⁹ ma per semplicità assumiamo che l'innovatore abbia un monopolio pieno sul nuovo farmaco fino a quando il brevetto è in vigore, e che la competizione dei farmaci generici annulli completamente i suoi profitti quando il brevetto scade. Sarebbe comunque sufficiente assumere, per i nostri risultati, che alla scadenza del brevetto i profitti dell'innovatore si riducano.

⁸Per esempio, nel caso del Naprosene, Mossinghoff (1999) riporta che dopo soli tre mesi dalla scadenza del brevetto, la quota di mercato del farmaco *brand* era già scesa al 25%.

⁹Si veda, ad esempio, Grabowski e Vernon (1992) e Reiffen e Ward (2005).

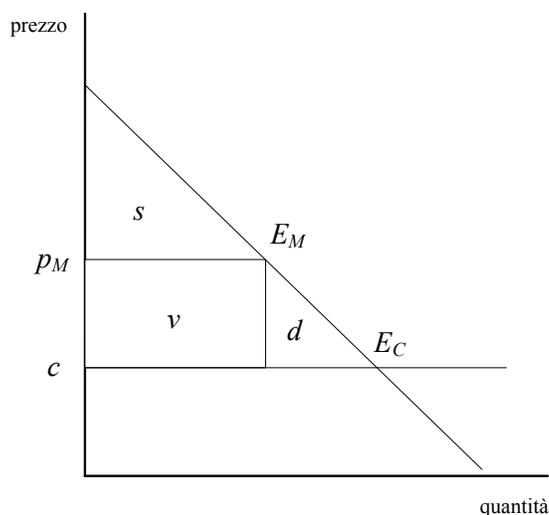


Figura 2.2: L'equilibrio nel mercato del prodotto.

L'equilibrio nel mercato del prodotto è illustrato nella Figura 2.2. Prima che il brevetto scada, si ha un equilibrio di monopolio nel punto E_M : l'innovatore fissa il prezzo di monopolio p_M e ottiene in ogni periodo un profitto v . Questo profitto rappresenta il valore privato dell'innovazione, in termini di flusso. I consumatori ottengono un surplus pari a s , mentre d rappresenta la perdita secca dovuta al fatto che il prezzo del farmaco è superiore al costo marginale.

Quando il brevetto scade, la concorrenza riduce il prezzo al livello del costo di produzione c , portando l'equilibrio nel punto E_C . I profitti e la perdita secca si annullano e tutto il beneficio sociale dell'innovazione è ottenuto dai consumatori, il cui surplus diventa uguale all'intera area compresa tra la curva di domanda e il costo marginale, che è pari alla somma delle aree s , v e d . Questa somma rappresenta il valore sociale dell'innovazione (sempre in termini di flusso), condizionatamente ad un suo uso efficiente.¹⁰

¹⁰Quando invece l'innovazione è monopolizzata dal titolare del brevetto, il valore sociale (cioè la somma del surplus dei consumatori e dei produttori) è pari solo a $s + v$.

I profitti futuri sono scontati a un tasso di sconto costante $r > 0$, che è anche il tasso sociale di sconto. Il profitto scontato, valutato al momento da cui comincia la commercializzazione del farmaco, cioè al tempo τ , è pari a

$$\int_0^T ve^{-rt} dt = \frac{(1 - e^{-rT})v}{r}$$

dove T è la durata effettiva del brevetto.¹¹

2.2.4 POLITICA DEI BREVETTI

Ipotizzando che la protezione brevettuale sia completa e impedisca ogni forma di concorrenza fino a che il brevetto è in vigore, la nostra analisi si concentrerà sulla durata effettiva del brevetto T . In particolare, supponiamo che la durata effettiva del brevetto dipenda dal tempo di sviluppo τ . Come abbiamo visto nel primo capitolo, il modo in cui T dipende da τ è determinato dalle regole legali di estensione della durata dei brevetti farmaceutici introdotte dall'Hatch-Waxman Act negli Stati Uniti e successivamente in Europa e negli altri Paesi industrializzati.

In questo capitolo consideriamo la funzione $T(\tau)$ come data, assumendo soltanto che soddisfi alcune condizioni di regolarità. In particolare, supponiamo: (i) che $T(\tau) : R_+ \rightarrow R_+$ sia continua, non crescente e differenziabile tranne che in un numero finito di punti; (ii) che $T(\bar{\tau})$ sia finito; e (iii) che $T'(\tau) \geq -1$, cioè che la scadenza del brevetto $\tau + T(\tau)$ possa essere rinviata, ma non possa essere anticipata, se lo sviluppo del farmaco richiede più tempo. Nei capitoli 4 e 5 analizzeremo la scelta ottimale della funzione $T(\tau)$ da parte di un policymaker benevolente che ha l'obiettivo di massimizzare il benessere sociale (ovvero la somma scontata del surplus atteso del consumatore e dei profitti).

¹¹Con l'espressione "durata effettiva del brevetto" indichiamo per brevità la durata del monopolio dell'innovatore, che, come abbiamo visto nel primo capitolo, dipende sia dalla durata del brevetto in senso stretto che dalle politiche di *data exclusivity*.

2.2.5 TIMING

Prima di procedere può essere utile riassumere il *timing* del modello, che è rappresentato nella Figura 2.3.

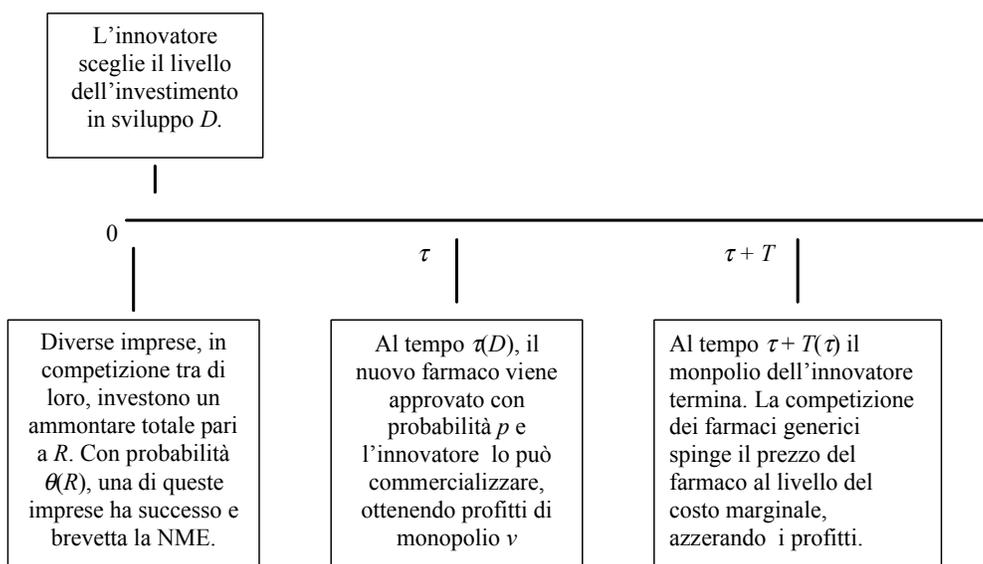


Figura 2.3: Il *timing* del modello.

2.3 LO STADIO DELLO SVILUPPO

Risolviamo ora il modello determinando i livelli di equilibrio dell'investimento in ricerca e in sviluppo per ogni data funzione $T(\tau)$ che soddisfa le nostre ipotesi di regolarità. Procediamo a ritroso, iniziando dallo stadio dello sviluppo.

Consideriamo un innovatore che ha appena scoperto e brevettato una NME. Il suo profitto atteso, scontato al tempo 0 e al lordo delle spese sostenute nella fase della ricerca, è:

$$\pi = p \int_{\tau}^{T+\tau} v e^{-rt} dt - D. \quad (2.1)$$

L'interpretazione della (2.1) è la seguente: con probabilità p , il farmaco viene approvato al tempo τ e il detentore del brevetto ottiene profitti v da τ fino al tempo $T + \tau$, quando

la protezione brevettuale scade. A questi profitti, che sono scontati al tasso r , vanno poi sottratte le spese per l'investimento in sviluppo, D .

Poichè $T = T(\tau)$ e $\tau = \tau(D)$, il profitto è in ultima analisi una funzione di D . L'innovatore sceglie la spesa in sviluppo D in modo da massimizzare $\pi(D)$. Le nostre ipotesi sulla funzione $\tau(D)$ implicano che π diventi negativo quando $D \rightarrow \infty$ e tenda a 0 quando $D \rightarrow 0$. Assumendo che $\pi(D)$ sia strettamente positivo per qualche valore di D (altrimenti nessuno mai investirebbe), la soluzione del problema di massimo è necessariamente interna. La condizione di primo ordine per un massimo è:

$$\frac{d\pi}{dD} = pv \{ [1 + T'(\tau)] e^{-r[T(\tau)+\tau]} - e^{-r\tau} \} \tau'(D) - 1 = 0. \quad (2.2)$$

La condizione (2.2) determina l'investimento ottimale nella fase dello sviluppo, D^* , e quindi implicitamente anche il tempo di sviluppo ottimale, $\tau^* = \tau(D^*)$. Entrambi dipendono dalla funzione $T(\tau)$.

2.3.1 EFFETTO DI FRONT LOADING E DI DURATA EFFETTIVA

Per capire meglio gli incentivi a investire nello stadio dello sviluppo, possiamo riscrivere la condizione (2.2) come segue:

$$pv \{ e^{-r\tau} - e^{-r[T(\tau)+\tau]} \} - pvT'(\tau)e^{-r[T(\tau)+\tau]} = -\frac{1}{\tau'(D)}. \quad (2.3)$$

L'espressione a sinistra del segno di uguale è il beneficio marginale di una riduzione del tempo di sviluppo, quella a destra è il costo marginale.

La condizione (2.3) dice che il beneficio marginale di una riduzione del tempo di sviluppo deve essere uguale al suo costo marginale. Il costo marginale di una riduzione del tempo di sviluppo è $-\frac{dD}{d\tau}$, che per la regola di derivazione delle funzioni inverse è uguale a $-\frac{1}{\tau'(D)}$. Il beneficio marginale è la somma di due termini. Il primo rappresenta l'effetto di anticipazione dei profitti (*front loading*). Dato che i profitti futuri sono scontati, l'innovatore ha incentivo

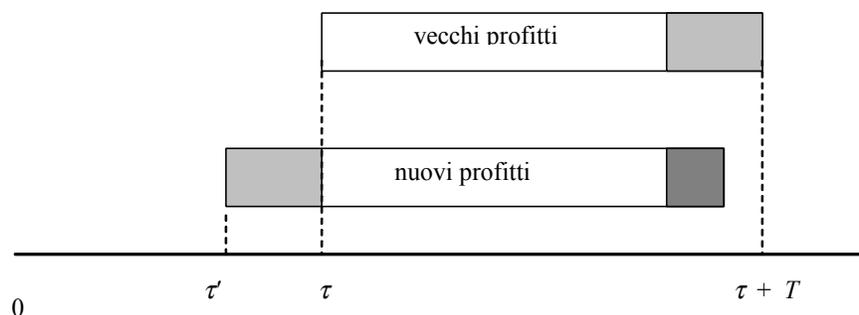


Figura 2.4: Effetti di *front loading* e di durata effettiva

a velocizzare lo sviluppo del nuovo farmaco per ottenere i suoi profitti al più presto. In particolare, se τ viene ridotto, poniamo, di un mese, l'innovatore guadagna un mese di profitti v alla data τ . Ma poiché il brevetto scade un mese prima (a parità di durata effettiva T), egli perde un mese di profitti v alla data $T + \tau$. Tuttavia, i profitti guadagnati valgono più di quelli perduti perché sono scontati per un periodo più breve. Il beneficio netto è così pari a $v [e^{-r\tau} - e^{-r(T+\tau)}]$. (Questo termine va poi moltiplicato per la probabilità che la NME sia efficace e sicura.)

Questo ragionamento presuppone che la durata del brevetto T sia costante. La seconda parte del beneficio marginale è invece legata al fatto che quando la funzione $T(\tau)$ è strettamente decrescente – cioè quando $T'(\tau) < 0$ – una riduzione del tempo di sviluppo aumenta la durata effettiva del brevetto. Ciò crea un ulteriore incentivo a ridurre τ , che è tanto più forte quanto più velocemente T cresce al calare di τ . A differenza dell'effetto di *front loading*, questo effetto di *durata effettiva* svanirebbe se T fosse indipendente da τ .

La Figura 2.4 rappresenta il meccanismo che genera i due effetti che abbiamo appena discusso. I due rettangoli in grigio chiaro, di uguale lunghezza ma sfasati nel tempo, rappresentano lo spostamento dei profitti quando il tempo di sviluppo si riduce da τ a τ' . Il rettangolino grigio scuro rappresenta il profitto aggiuntivo dovuto alla maggior durata

effettiva del brevetto.

2.3.2 ANALISI GRAFICA

Può essere utile una rappresentazione grafica del problema di ottimizzazione dell'inventore nello stadio dello sviluppo. Consideriamo il piano (τ, T) . Invertendo la funzione $\tau(D)$ per esprimere D in funzione di τ , la funzione di profitto può essere riscritta in funzione di τ e T come segue:

$$\pi(\tau, T) = p \int_{\tau}^{\tau+T} v e^{-rt} dt - D(\tau).$$

Consideriamo ora le curve di isoprofitto nel piano (τ, T) . L'inclinazione delle curve di isoprofitto è data da

$$\frac{dT}{d\tau} = -\frac{\frac{\partial \pi}{\partial \tau}}{\frac{\partial \pi}{\partial T}}.$$

La funzione $\pi(\tau, T)$ è sempre crescente in T , poiché

$$\frac{\partial \pi}{\partial T} = p v e^{-r(\tau+T)} > 0.$$

Abbiamo poi

$$\frac{\partial \pi}{\partial \tau} = p v [e^{-r(\tau+T)} - e^{-r\tau}] - D'(\tau).$$

Se la curva $\tau(D)$ è molto ripida per valori bassi di D e quasi piatta per valori alti, come rappresentato in Figura 2.1, questa derivata è positiva per valori bassi di τ e negativa per valori alti. Le curve di isoprofitto nel piano (τ, T) saranno allora prima decrescenti e poi crescenti. La derivata $\frac{\partial \pi}{\partial \tau}$ si annulla quando

$$\frac{\partial \pi}{\partial \tau} = p v [e^{-r(\tau+T)} - e^{-r\tau}] - D'(\tau) = 0,$$

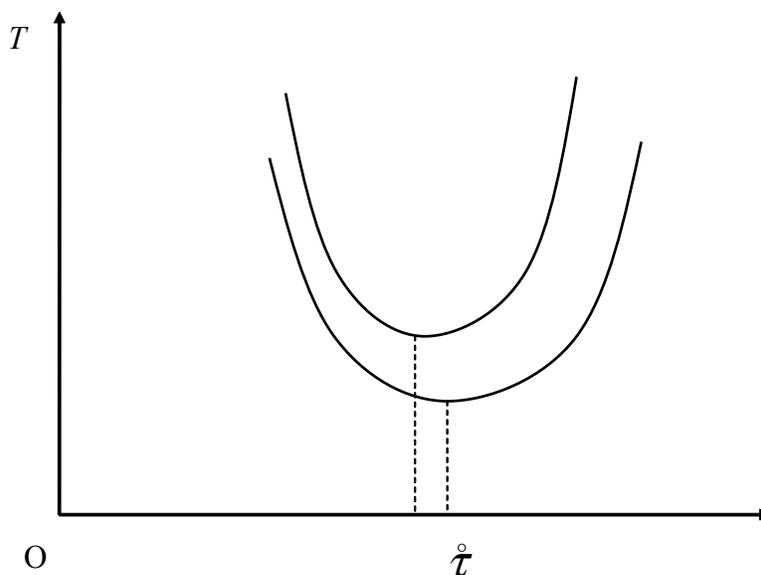


Figura 2.5: Curve di isoprofitto. Curve più alte hanno il punto di minimo spostato verso sinistra.

che corrisponde al punto di minimo della curva di iso-profitto. Indichiamo con $\hat{\tau}$ il valore di τ che soddisfa questa condizione. Si può facilmente verificare che curve di isoprofitto più alte avranno un minimo spostato a sinistra.¹²

L'equilibrio nello stadio dello sviluppo è illustrato graficamente nella Figura 2.6. La funzione $T(\tau)$ scelta dal policymaker può essere vista come un vincolo che deve essere soddisfatto nel processo di massimizzazione dei profitti. Graficamente, quindi, il problema dell'impresa è quello di raggiungere la curva di iso-profitto più alta dato il vincolo, poiché la curva $T(\tau)$ è non crescente, l'equilibrio sarà necessariamente nel tratto decrescente della curva di isoprofitto. Il tratto crescente delle curve di iso-profitto è quindi irrilevante, nel senso che l'equilibrio non si avrà mai in questo tratto: avremo sempre $\tau^* \leq \hat{\tau}$.

¹²Questo segue dal fatto che

$$\frac{\partial^2 \pi}{\partial T \partial \tau} = -r p v e^{-r(\tau+T)} < 0.$$

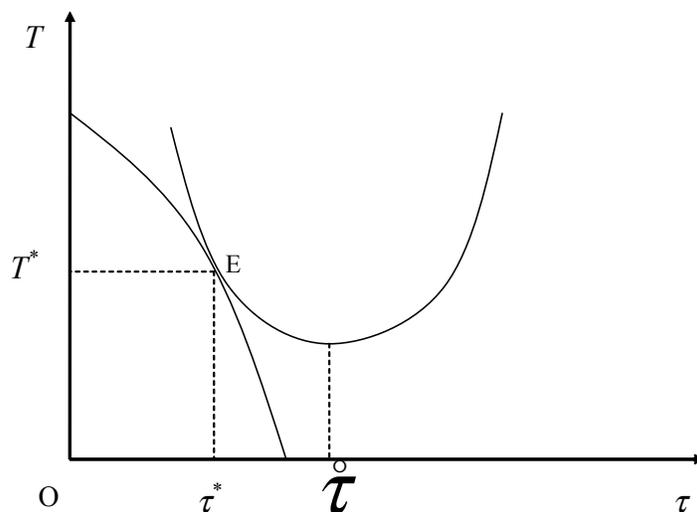


Figura 2.6: La scelta ottimale dell'innovatore nello stadio dello sviluppo.

Si noti che $\hat{\tau}$ è il valore del tempo di sviluppo che sarebbe ottimale se il vincolo $T(\tau)$ fosse una retta orizzontale, cioè se la durata effettiva del brevetto fosse indipendente dal tempo di sviluppo. In questo caso, l'effetto di durata effettiva svanirebbe e opererebbe solo l'effetto di *front loading* dei profitti. È allora intuitivo che quando opera anche l'effetto di durata effettiva, l'investimento in sviluppo sarà maggiore, e quindi il tempo di sviluppo si accorcerà: $\tau^* < \hat{\tau}$.

Si noti infine che la parte decrescente delle curve di isoprofitto (che è la sola rilevante) ha una forma simile a quella di una normale curva di indifferenza. Nel caso particolare in cui la funzione $T(\tau)$ è lineare, il vincolo è simile ad un vincolo di bilancio. L'intuizione grafica diventa quindi simile a quella di un problema standard di equilibrio del consumatore. Come vedremo, questa analogia può essere di aiuto per interpretare diversi risultati di statica comparata.

2.4 LO STADIO DELLA RICERCA

Passiamo ora allo stadio della ricerca. In questo stadio, diverse imprese competono per scoprire una NME. Le imprese sono in grado di prevedere l'equilibrio dello stadio dello sviluppo, per cui il premio per il vincitore della corsa per il brevetto è rappresentato dal profitto

$$\pi^* = p \int_{\tau^*}^{T(\tau^*)+\tau^*} v e^{-rt} dt - D^*.$$

Poichè vi è libertà di entrata nella corsa al brevetto, il profitto atteso nella fase della ricerca deve essere nullo. In equilibrio, l'investimento nella fase della ricerca R^* deve quindi soddisfare la seguente condizione:

$$R = \theta(R)\pi^*. \quad (2.4)$$

Questa equazione dice che con libertà di entrata, tutti i profitti attesi dall'innovazione sono investiti in ricerca. È immediato concludere che gli investimenti nello stadio di ricerca, R^* , e quindi la probabilità di successo, $\theta^* = \theta(R^*)$, siano una funzione crescente di π^* .

Proposizione 1. *L'investimento in ricerca R^* e la probabilità di successo θ^* sono funzioni crescenti del profitto atteso π^* .*

Dimostrazione. Differenziando implicitamente la (2.4) abbiamo

$$\frac{dR^*}{d\pi^*} = -\frac{\theta(R^*)}{\theta'(R^*)\pi^* - 1}.$$

Dato che $\theta(R)$ è concava e $\theta(0) = 0$, deve essere $\pi^* = \frac{R^*}{\theta(R^*)} < \frac{1}{\theta'(R^*)}$, il che implica che il denominatore è negativo. Ne segue che $\frac{dR^*}{d\pi^*} > 0$. ■

2.4.1 IL CASO DI ENTRATA BLOCCATA O RISTRETTA

L'ipotesi di libertà d'entrata, che implica che tutti i profitti attesi dall'innovazione siano investiti in ricerca, potrebbe essere considerata eccessivamente ottimista. Tuttavia, il risultato

che l'investimento in ricerca è una funzione crescente di π^* vale sotto ipotesi molto più generali.

Consideriamo per esempio il caso estremo di entrata bloccata con monopolio nella ricerca, in cui cioè, per qualche ragione, una sola impresa ha la possibilità di investire per scoprire una NME. Il profitto atteso del monopolista è $\theta(R)\pi^* - R$, e il monopolista sceglie R in modo da massimizzare tale profitto. La condizione di libertà d'entrata (2.4) deve quindi essere ora sostituita dalla condizione del primo ordine

$$\theta'(R)\pi^* - 1 = 0. \quad (2.5)$$

Dato che la funzione $\theta(R)$ è concava, è immediato concludere che, come nel caso di libertà d'entrata, R^* deve essere una funzione crescente di π^* .¹³

Si noti che la condizione (2.5) può essere riscritta come

$$R = \eta\theta(R)\pi^*.$$

dove η è l'elasticità della funzione $\theta(R)$, che come sappiamo è minore di uno. Quindi solo una frazione $\eta < 1$ dei profitti attesi viene investita in ricerca; la frazione rimanente $1 - \eta$ rappresenta il profitto atteso (positivo) del monopolista.

Questo suggerisce che i casi intermedi tra la libertà di entrata e il monopolio possano essere rappresentati immaginando che la frazione dei profitti attesi investiti nella ricerca sia compresa tra η e 1. Intuitivamente, tale frazione dovrebbe crescere al crescere dell'intensità della competizione nella ricerca. Anche se la verifica formale di questa congettura richiede un modello di investimento in ricerca esplicito (cioè un modello che specifichi la probabilità di successo di ogni singola impresa), e non una forma ridotta come quella utilizzata in questo lavoro, è intuitivo che un risultato simile alla Proposizione 1 debba valere anche per questi casi intermedi.

¹³E' sufficiente differenziare implicitamente la (2.5), ottenendo:

$$\frac{dR^*}{d\pi^*} = -\frac{\theta'(R)}{\theta''(R)\pi^*} > 0.$$

2.5 STATICA COMPARATA

Possiamo ora analizzare come i livelli di investimento in ricerca (R^*) e sviluppo (D^*) dipendano dalle variabili esogene del modello. Questa analisi di statica comparata non conduce a risultati sorprendenti: al contrario, i risultati sono del tutto in linea con le aspettative o comunque hanno semplici spiegazioni intuitive, il che può essere preso come una indicazione che il modello sia perlomeno sensato.

Le variabili esogene del modello sono: il valore dell'innovazione v , la probabilità che il nuovo farmaco sia approvato p , il tasso di interesse r e la durata effettiva del brevetto T , che è data in funzione di τ . È immediato verificare che le prime due variabili in realtà si possono combinare in un unico indice del valore atteso dell'innovazione, pv , lasciando così tre variabile esogene. Consideriamole una per volta.

2.5.1 VALORE DELL'INNOVAZIONE

Proposizione 2. *Un aumento del valore dell'innovazione, pv , aumenta sia l'investimento in ricerca R^* che quello in sviluppo D^* .*

Dimostrazione. La dimostrazione è immediata. Per quanto riguarda D^* , differenziando implicitamente la condizione del primo ordine (2.2) abbiamo:

$$\frac{\partial D^*}{\partial(pv)} = - \frac{\{[1 + T'(\tau)] e^{-r[T(\tau)+\tau]} - e^{-r\tau}\} \tau'(D)}{\frac{d^2\pi}{dD^2}} > 0$$

dove il segno positivo segue dai seguenti fatti: *i*) il denominatore è negativo per la condizione di secondo ordine; *ii*) $\tau'(D) > 0$; *iii*) il termine entro parentesi graffe è sicuramente negativo, potendosi riscrivere come $[e^{-r(T+\tau)} - e^{-r\tau}] + T'(\tau)e^{-r(T+\tau)} < 0$.

Per quanto riguarda R^* , è sufficiente osservare che per il teorema dell'inviluppo

$$\frac{d\pi^*}{d(pv)} = \frac{\partial\pi^*}{\partial(pv)} = \int_{\tau}^{T+\tau} e^{-rt} dt > 0.$$

La conclusione segue poi dalla Proposizione 1. ■

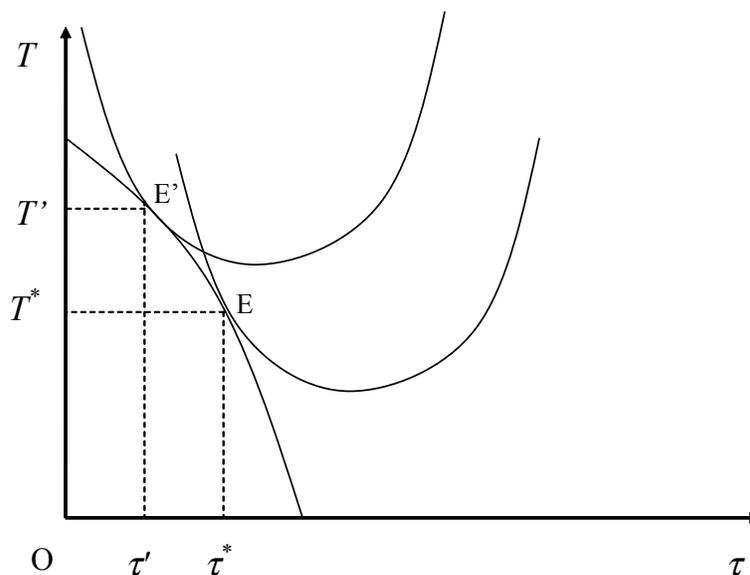


Figura 2.7: Effetti di un aumento del valore dell'innovazione sull'equilibrio nello stadio dello sviluppo.

Graficamente, un aumento di v e/o p fa spostare le curve di isoprofitto verso sinistra. Il conseguente spostamento dell'equilibrio nello stadio dello sviluppo è illustrato dalla Figura 2.7.

2.5.2 TASSO DI INTERESSE

Proposizione 3. *Un aumento del tasso di interesse r riduce l'investimento in ricerca R^* .*

Dimostrazione. Procedendo come nella dimostrazione della Proposizione 2, è sufficiente osservare che per il teorema dell'involuppo

$$\frac{d\pi^*}{dr} = \frac{\partial \pi^*}{\partial r} = pv \int_{\tau}^{T+\tau} (-t)e^{-rt} dt < 0.$$

La conclusione segue poi dalla Proposizione 1. ■

L'effetto di una variazione del tasso di interesse sull'investimento in sviluppo D^* è generalmente ambiguo, come si può verificare differenziando implicitamente la condizione del primo ordine (2.2):

$$\frac{\partial D^*}{\partial r} = - \frac{pv \{ \tau e^{-r\tau} - (\tau + T) [1 + T'(\tau)] e^{-r(\tau+T)} \} \tau'(D)}{\frac{d^2\pi}{dD^2}}.$$

Un aumento del tasso di interesse sicuramente indebolisce l'effetto di durata effettiva, ma l'effetto di *front loading* può essere rafforzato o indebolito. Per esempio, l'effetto di *front loading* è sicuramente rafforzato da un aumento di r se il valore iniziale di r è sufficientemente basso, dal momento che l'effetto di *front loading* è nullo quando $r = 0$.

2.5.3 DURATA EFFETTIVA DEL BREVETTO

Nel nostro modello (come nella realtà) la durata effettiva del brevetto non è un parametro ma una funzione del tempo di sviluppo. Per avere un'idea più precisa di come l'equilibrio di mercato dipenda dalla funzione $T(\tau)$, possiamo specificare una relazione lineare:

$$T = \Gamma - (1 - b)\tau \tag{2.6}$$

dove $\Gamma > 0$ e $0 \leq b \leq 1$. Il parametro Γ può essere interpretato come la durata nominale del brevetto, e il parametro b come la frazione del tempo di sviluppo τ che viene recuperata attraverso la protezione supplementare. Se $b = 1$ viene recuperato tutto il tempo di sviluppo e la durata effettiva coincide con quella nominale. Se invece $b = 0$, non vi è alcuna protezione supplementare.

Proposizione 4. *L'investimento in sviluppo D è una funzione crescente della durata nominale del brevetto Γ . L'investimento in sviluppo può essere sia crescente che decrescente in b , ma è sicuramente decrescente se b e/o il tasso di interesse sono sufficientemente piccoli.*

Dimostrazione. Differenziamo implicitamente la (2.3) dopo aver posto $T'(\tau) = -(1 - b)$. Consideriamo dapprima l'effetto di un cambiamento nella durata ufficiale del brevetto Γ .

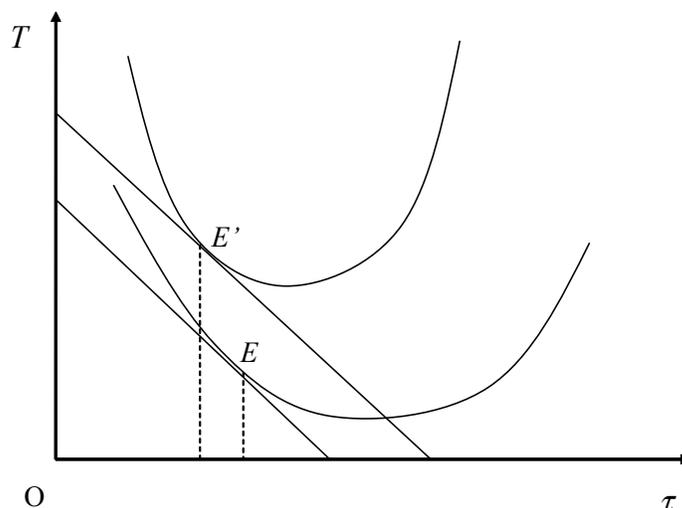


Figura 2.8: Effetti di un aumento di Γ sull'equilibrio nello stadio dello sviluppo.

Abbiamo:

$$\frac{\partial D^*}{\partial \Gamma} = \frac{e^{-r(T+\tau^*)} r p v b \tau'(D^*)}{\frac{d^2 \pi}{dD^2}} \geq 0$$

Si noti che la derivata è strettamente positiva a meno che $b = 0$.

Consideriamo poi un cambiamento del parametro b . Abbiamo:

$$\frac{\partial D^*}{\partial b} = - \frac{p v e^{-r(T+\tau^*)} (1 - r b \tau^*) \tau'(D^*)}{\frac{d^2 \pi}{dD^2}}.$$

Utilizzando ancora la condizione di secondo ordine, abbiamo che il segno di $\frac{\partial D^*}{\partial b}$ è opposto al segno di $1 - r b \tau^*$. In generale il segno di questa espressione è ambiguo, ma è certamente positivo (e quindi D^* è una funzione decrescente di b) se r e b sono sufficientemente piccoli.

■

Questi risultati sono illustrati nelle Figure 2.8 e 2.9. La Figura 2.8 mostra che un aumento di Γ è analogo ad un aumento del reddito nell'equilibrio del consumatore. La prima parte della Proposizione 4 stabilisce quindi che τ è un “bene inferiore,” nel senso che un aumento del “reddito” provoca una diminuzione del consumo del “bene τ .” La ragione intuitiva è che

l'effetto di anticipazione dei profitti è tanto più forte quanto maggiori sono i profitti. Un aumento di Γ aumenta i profitti, e quindi rafforza l'incentivo ad anticiparli: ciò porta ad un aumento degli investimenti e a una riduzione del tempo di sviluppo.

La Figura 2.9 mostra che un aumento di b è analogo ad una riduzione del prezzo di τ nell'equilibrio del consumatore. L'effetto è normalmente quello di far aumentare il consumo del "bene τ ." Tuttavia, non si può escludere la possibilità del paradosso di Giffen (cioè che τ diminuisca). Un aumento di b , infatti, ha due effetti contrastanti. Il primo è simile all'effetto di sostituzione: si riduce l'incentivo a investire nello stadio dello sviluppo per aumentare la durata effettiva del brevetto, perchè ogni data riduzione del tempo di sviluppo corrisponde ad un minore aumento della durata effettiva. Il secondo è simile all'effetto di reddito: *ceteris paribus*, un aumento di b aumenta la durata del brevetto di un ammontare uguale a τ^* , aumentando i profitti e rafforzando l'effetto di *front-loading*, il che porta ad una riduzione di τ . L'effetto totale è generalmente ambiguo, ma il primo effetto deve necessariamente prevalere quando il tasso di sconto è piccolo (e quindi l'effetto di *front-loading* diventa trascurabile) e anche b è piccolo (il che rafforza l'effetto di durata effettiva).

L'analogia con la teoria del consumatore suggerisce un risultato più generale, e cioè che se la pendenza della curva $T(\tau)$ nel punto di equilibrio (τ^*, T^*) viene ridotta, il valore ottimale di τ necessariamente diminuisce. Infatti, se la nuova curva $T(\tau)$ continua a passare per il punto (τ^*, T^*) ma la sua pendenza è inferiore, opera solo l'effetto di sostituzione che come abbiamo visto porta ad una diminuzione dell'investimento e quindi ad un aumento di τ .

Per completare l'analisi passiamo infine allo stadio della ricerca.

Proposizione 5. *L'investimento in ricerca R^* e la probabilità di successo θ^* sono funzioni crescenti della durata nominale del brevetto, Γ , e della frazione del tempo di sviluppo che viene recuperata, b .*

Dimostrazione. Il risultato segue dal teorema dell'involuppo e dalla Proposizione 1. Abbiamo

$$\frac{d\pi^*}{d\Gamma} = \frac{\partial\pi^*}{\partial\Gamma} = pve^{-r(T+\tau^*)} > 0$$

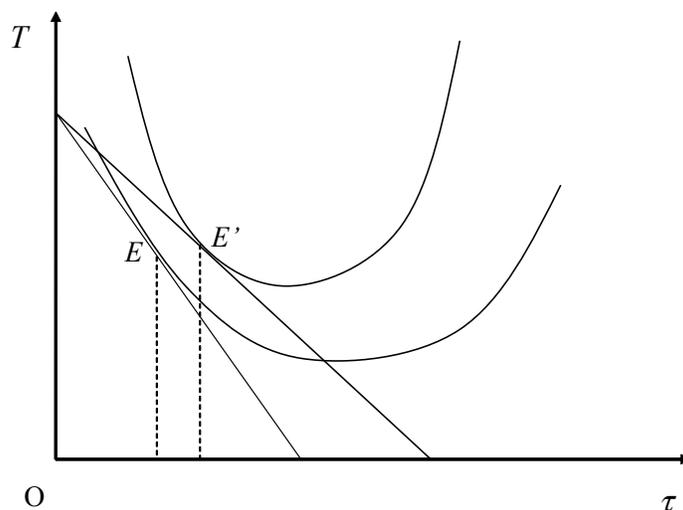


Figura 2.9: Effetti di un aumento di b sull'equilibrio nello stadio dello sviluppo.

e

$$\frac{d\pi^*}{db} = \frac{\partial\pi^*}{\partial b} = \tau^* p v e^{-r(T+\tau^*)} > 0. \quad \blacksquare$$

2.6 CONCLUSIONI

In questo capitolo abbiamo proposto un modello teorico del processo innovativo nel settore farmaceutico articolato in due stadi: lo stadio della ricerca e lo stadio dello sviluppo. I due stadi differiscono per due motivi. Primo, lo stadio dello sviluppo è necessariamente caratterizzato da monopolio, mentre nello stadio della ricerca diverse imprese possono competere per scoprire una NME. Secondo, nello stadio della ricerca l'investimento determina la probabilità di successo, mentre nello stadio dello sviluppo esso determina il tempo necessario per ottenere (con una probabilità data esogenamente) l'approvazione del farmaco.

Il modello genera una serie di risultati di statica comparata che hanno semplici spiegazioni intuitive. Esso inoltre consente di scomporre l'incentivo totale ad investire nello stadio dello sviluppo in due componenti distinte: l'effetto di *front loading* dei profitti e l'effetto di *durata effettiva*. L'effetto di *front loading* deriva dal fatto che l'innovatore ha incentivo a velocizzare lo sviluppo del nuovo farmaco per ottenere i suoi profitti al più presto, dato che i profitti futuri sono scontati. L'effetto di *durata effettiva* è invece legato al fatto che una riduzione del tempo di sviluppo aumenta la durata effettiva del brevetto. Questo effetto opera ogni qual volta l'estensione della durata nominale dei brevetti farmaceutici non compensa completamente l'innovatore del tempo impiegato per ottenere l'approvazione del farmaco. Come abbiamo visto nel primo capitolo, questa è precisamente la situazione attuale sia negli Stati Uniti che in Europa.

Il modello che abbiamo proposto in questo capitolo fornisce le basi sia per l'analisi empirica che svilupperemo nel prossimo capitolo, sia per l'analisi della politica ottimale nel capitolo 4.

CAPITOLO 3

ANALISI EMPIRICA

Riassunto. Il modello teorico presentato nel precedente capitolo ha diverse implicazioni che, almeno in linea di principio, possono essere testate empiricamente. In particolare, abbiamo identificato l'effetto di “durata effettiva,” che non era stato precedentemente considerato nella letteratura sugli effetti dei brevetti. In questo capitolo sviluppiamo un'analisi empirica dei tempi di sviluppo con l'obiettivo di identificare e quantificare l'effetto di durata effettiva.

3.1 INTRODUZIONE

La durata effettiva dei brevetti farmaceutici non coincide con la loro durata nominale (20 anni) perchè da un lato i tempi di sviluppo dei nuovi farmaci (cioè il tempo che passa tra l'ottenimento del brevetto e la commercializzazione) sono molto lunghi, dall'altro una specifica legislazione (introdotta inizialmente negli Stati Uniti dall'*Hatch-Waxman Act* del 1984 e successivamente adottata, con qualche variante, in Europa e nella maggior parte dei paesi industrializzati) estende il periodo di esclusiva di cui godono gli inventori. L'effetto di questa legislazione è che la durata effettiva dei brevetti farmaceutici dipende dal tempo di sviluppo. Nel precedente capitolo abbiamo visto come questa dipendenza funzionale possa fornire un ulteriore incentivo a investire risorse per accorciare i tempi di sviluppo: oltre al fatto che una commercializzazione più veloce implica in ogni caso che i profitti dell'innovatore siano anticipati e quindi siano scontati meno pesantemente (l'effetto di *front loading* dei

profitti), essa può anche tradursi in una maggiore durata effettiva del brevetto (l'effetto di *durata effettiva*).

Il nostro modello teorico implica diverse conclusioni che possono essere testate empiricamente. In primo luogo, implica una relazione negativa tra valore dell'innovazione e tempo di sviluppo. Questa relazione è stata confermata empiricamente da una letteratura che esamineremo nella prossima sezione. Ma il modello permette anche di scomporre l'incentivo ad investire nella fase dello sviluppo in un effetto di *front loading* e un effetto di durata effettiva. Questo secondo effetto, per quanto a nostra conoscenza, non è mai stato studiato empiricamente. La parte rimanente di questo capitolo propone una prima analisi empirica dei tempi di sviluppo con l'obiettivo di identificare, e se possibile quantificare, l'effetto di durata effettiva.

3.2 VALORE DEI FARMACI E TEMPO DI SVILUPPO

Una prima implicazione testabile del nostro modello teorico è che il tempo di sviluppo dipende negativamente dal valore dell'innovazione – cioè dal valore del nuovo farmaco. Questa conclusione discende immediatamente dal fatto che sia l'effetto di *front loading* dei profitti che l'effetto di durata effettiva sono più forti per innovazioni di valore più elevato (si veda l'equazione 2.3).

Come abbiamo già accennato nel secondo capitolo, la letteratura ha già testato e confermato empiricamente l'ipotesi che il tempo di sviluppo dei farmaci commercialmente più importanti sia più breve. Il lavoro fondamentale al riguardo è quello di Dranove e Meltzer (1994). Essi concludono che

more important drugs are developed and approved more rapidly than less important drugs (p. 402).

Più recentemente, anche se in modo meno raffinato, questa conclusione è stata confermata da Keyhani et al. (2006). In questa sezione discutiamo brevemente questi lavori.

L'analisi di Dranove e Meltzer (1994) è basata su due premesse. La prima è che le compagnie farmaceutiche siano in grado di identificare i farmaci più importanti in una fase precoce del processo di sviluppo. La seconda è che abbiano la possibilità di influenzare il tempo di sviluppo. A questo proposito, Dranove e Meltzer notano:

It is evident that firms have some ability to anticipate which drugs will be leading sellers. (p. 421)

Inoltre,

(...) firms already possess a tremendous ability to accelerate a given drug.
(p. 421)

I modi in cui le compagnie farmaceutiche possono velocizzare i tempi di sviluppo e approvazione sono molteplici: ad esempio, per rispondere alle richieste della FDA che i test debbano essere effettuati su un certo numero di volontari e/o pazienti, un'impresa potrebbe scegliere di lavorare in un piccolo numero di istituzioni per un periodo di tempo abbastanza lungo oppure lavorare con un più alto numero di istituzioni per un periodo di tempo più breve – spendendo però di più per raggiungere la misura del campione richiesta. Analogamente, un'impresa potrebbe decidere di compensare molto generosamente i volontari per i test per trovare volontari più rapidamente. Oppure potrebbe utilizzare per determinati progetti il proprio personale più qualificato ed ottenere il maggiore vantaggio dalle interazioni con l'FDA nel processo di sviluppo.

Sulla base di queste premesse, Dranove e Meltzer analizzano il periodo che va dal 1950 al 1986,¹ identificando un campione di 564 NME.² Per valutare l'importanza di ciascun farmaco, vengono considerate cinque diverse misure: il volume delle vendite, che però è disponibile solo per 75 farmaci del campione;³ l'estensione geografica dei mercati su cui il farmaco è stato

¹Per comodità d'analisi, questo lungo periodo viene suddiviso in 6 sottoperiodi di 6 anni ciascuno.

²Le NME approvate nel periodo di osservazione considerato furono in tutto 760. Gli autori però sono riusciti ad avere dati rilevanti solo per 564 NME.

³Vi sono in realtà anche dati più estesi, ma gli autori notano che, come purtroppo abbiamo potuto verificare personalmente, queste sono disponibili solo a costi proibitivi.

immesso; le citazioni ricevute nei testi di medicina; le citazioni nelle riviste di medicina; e infine le citazioni ricevute nelle domande di brevetto successive. L'analisi mette in relazione queste diverse misure di importanza (di mercato e scientifiche) con i tempi di sviluppo e di approvazione del farmaco.

Dranove e Meltzer definiscono il tempo di sviluppo come il periodo di tempo che va dalla data del deposito della domanda di brevetto alla data di approvazione della NDA.⁴ Anche se la loro analisi non distingue la fase pre-clinica da quella clinica, Dranove e Meltzer notano che circa la metà di tutte le spese in R&D venga sostenuta prima della IND, e quindi o nella fase della ricerca di una NME o nella fase pre-clinica. In realtà, il processo di sviluppo comincia a partire dalla scoperta della NME. Poiché non esistono fonti ufficiali dalle quali sia possibile desumere le date esatte di scoperta di una NME, Dranove e Meltzer utilizzano la data della domanda di brevetto internazionale per approssimare la data di scoperta di un farmaco.⁵ Questa non dovrebbe comunque essere molto diversa dalla data della scoperta: infatti è nell'interesse della compagnia farmaceutica depositare la domanda di brevetto subito dopo la scoperta della NME per proteggere i diritti brevettuali.⁶

Per determinare come l'importanza dei farmaci ne influenzi il tempo di sviluppo, Dranove e Meltzer stimano l'*hazard rate* del processo di approvazione, cioè la probabilità che un farmaco sia approvato, condizionatamente all'ipotesi che non lo sia già stato, in funzione del tempo passato dalla scoperta e delle misure di importanza descritte sopra. I risultati

⁴Questa definizione coincide con quella adottata in questo lavoro. Va però rilevato che altri autori usano definizioni diverse, il che crea una certa confusione nella letteratura. Ad esempio, Wiggins (1981), Hansen (1979) e Di Masi *et al.* (1991), definiscono il periodo di sviluppo come il periodo che va dalla data di sottomissione dell'*IND application* all'approvazione della NDA (nella nostra terminologia, questa è la durata della fase clinica). Altri, come Kaitin e al. (1991), come il periodo dalla data di sottomissione dell'NDA all'approvazione dell'NDA (che nella nostra terminologia è la fase di approvazione).

⁵Anche l'identificazione della data di deposito della domanda per il brevetto è problematica. Infatti, lo stesso farmaco è tipicamente protetto da molti brevetti: oltre al brevetto (o ai brevetti) sulla sostanza vera e propria, vi possono essere, per esempio, brevetti d'uso o brevetti di processo che non è facile identificare. Inoltre, Dranove e Meltzer non sono riusciti a trovare la data di deposito della domanda per il brevetto per 20 farmaci. In questi casi, la data della domanda è stata stimata sottraendo 2 anni alla data di concessione del brevetto.

⁶Si veda la discussione nella sezione 1.4.2.

indicano che i farmaci più importanti hanno avuto tempi di sviluppo più veloci dei farmaci meno importanti. La riduzione del tempo di sviluppo è dell'ordine del 25% per farmaci il cui valore sia di due deviazioni standard maggiore del valore medio. Risulta anche che la misura del valore del farmaco a cui il tempo di sviluppo è più sensibile è quella commerciale.

Gli autori si chiedono poi se l'accelerazione nel processo di sviluppo dei farmaci più importanti sia attribuibile alle scelte delle compagnie farmaceutiche o all'iniziativa della FDA. Infatti, anche il regolamentatore può avere un incentivo a velocizzare l'approvazione dei farmaci più importanti. In effetti, dalla metà degli anni settanta viene assegnata, ai farmaci ritenuti più importanti sulla base della novità della struttura chimica e della novità terapeutica, una procedura prioritaria nel processo di approvazione.⁷ Dranove e Melzter però concludono che il fattore determinante sono le scelte delle compagnie farmaceutiche (che nel nostro modello sono riassunte nel livello dell'investimento in sviluppo, D). Questo si desume anche dal fatto che sin dagli anni cinquanta, cioè prima del *Kefauver-Harris Amendment*, si è assistito ad un'accelerazione dei tempi per i farmaci più importanti.⁸

Oltre all'analisi articolata di Dranove e Meltzer, vanno ricordati anche gli studi dell'*Office of Technology Assessment* (OTA) e di Keyhani et al. (2006). Lo studio dell'OTA del 1989 evidenzia come farmaci commercialmente più importanti abbiano una durata effettiva del brevetto più lunga in media di 3,3 anni rispetto ai farmaci meno venduti.⁹ Keyhani et al. (2006) distinguono i farmaci in due categorie commerciali, a seconda che il volume annuale delle vendite sia maggiore o minore di 100 milioni di dollari. Questi autori non hanno dati sulla durata della fase pre-clinica, ma trovano che la durata mediana delle fasi clinica e di approvazione sia inferiore di 1,3 anni per i farmaci più importanti.

⁷Kaitin et al. (1991) trovano che effettivamente, con questa nuova procedura, i farmaci prioritari vengono approvati più velocemente. Dranove e Meltzer notano che i farmaci prioritari impiegano però più tempo per arrivare a sottomettere la domanda di NDA, forse perchè sono più difficili da testare o perchè le compagnie farmaceutiche perdono tempo per cercare di ottenere la priorità.

⁸Ciò naturalmente non significa che molto possa essere fatto ancora da parte dell'FDA per contribuire ad accelerare il processo di commercializzazione dei nuovi farmaci, ovviamente nel rispetto dei requisiti di efficacia e di sicurezza.

⁹Questo studio è citato in Dranove e Meltzer (1994), che riportano la sua principale conclusione, ma non siamo riusciti a reperirlo.

Possiamo quindi concludere che il tempo di sviluppo dipende negativamente dal valore del nuovo farmaco. Questa è la principale giustificazione dell'ipotesi, che è cruciale per la nostra analisi, che l'innovatore possa influenzare la durata del processo di approvazione attraverso le decisioni di investimento nella fase dello sviluppo. Una volta che si ammette questa possibilità, una correlazione negativa tra valore dei farmaci e il loro tempo di sviluppo sarà verosimilmente una proprietà condivisa da qualunque modello ragionevole dell'attività di ricerca nel settore farmaceutico.

3.3 L'EFFETTO DI DURATA EFFETTIVA

Una implicazione più specifica del nostro modello teorico è invece l'effetto di “durata effettiva,” che dipende dall'esistenza di una relazione tra durata effettiva del brevetto e tempo di sviluppo del farmaco. Questo effetto gioca un ruolo importante nell'analisi normativa dei capitoli 4 e 5, e pertanto ci sembra interessante identificarlo empiricamente. Per quanto a nostra conoscenza, nessuno fino ad ora ha cercato di farlo.

Per identificare l'effetto di “durata effettiva” si possono seguire diverse strategie. In questa sezione le analizziamo e ne discutiamo la fattibilità. Come si vedrà, la strategia che è forse più naturale ci è preclusa per l'indisponibilità dei dati, il che ci costringe a ricorrere ad approcci più indiretti.

3.3.1 RIFORME LEGISLATIVE

La prima e più naturale strategia empirica sfrutterebbe i cambiamenti nella politica della durata dei brevetti attuati negli anni ottanta e nei primi anni novanta del secolo scorso. Come abbiamo visto, l'approvazione dell'*Hatch-Waxman Act* e di provvedimenti analoghi in altri paesi in quel periodo ha di fatto ridotto la pendenza della curva $T(\tau)$ rispetto al regime pre-esistente, riducendo così l'intensità dell'effetto di durata effettiva. La predizione del modello è che questo dovrebbe aver allungato i tempi di sviluppo. Confrontando i tempi

di sviluppo prima e dopo la riforma dovrebbe quindi essere possibile identificare l'effetto di durata effettiva.

Questo approccio è comunque problematico per il fatto che nel tempo possono essere cambiate altre variabili che determinano i tempi di sviluppo. Questa difficoltà è accentuata dal fatto che le compagnie farmaceutiche sono imprese multinazionali i cui incentivi sono quindi determinati da una molteplicità di legislazioni nazionali, e che le riforme legislative non sono state introdotte simultaneamente in tutti i paesi, ma gradualmente nel tempo in un arco di diversi anni – un arco temporale sufficientemente lungo da far sì che nel frattempo possano essere cambiate molte altre determinanti del tempo di sviluppo.

Tuttavia, l'ostacolo principale all'implementazione di questa strategia non è tanto quello appena descritto, quanto, più semplicemente, l'indisponibilità dei dati. I dati necessari per seguire questo approccio in realtà esistono presso diversi centri di ricerca e società private, ma i centri di ricerca da noi interpellati non sono stati disponibili a condividere i dati, e le società private hanno richiesto pagamenti proibitivi. Abbiamo dunque dovuto procedere ad una raccolta diretta dei dati rilevanti. Nella prossima sezione descriveremo in dettaglio la procedura utilizzata, chiarendo la ragione per cui ci è stato impossibile ricostruire i dati per un passato troppo lontano.

3.3.2 INVESTIMENTO SEQUENZIALE

Dobbiamo quindi ricorrere ad un approccio più indiretto, ma che permette l'uso di dati relativi ad un periodo più recente. Per far questo, è necessario estendere l'analisi teorica del precedente capitolo in due direzioni per renderla più realistica. Per prima cosa, consideriamo il tempo di sviluppo come una variabile stocastica invece che deterministica. Questa ipotesi sembra naturale, dal momento che la durata delle varie fasi del processo di sviluppo dei nuovi farmaci è altamente variabile (si veda la sezione sulle statistiche descrittive per una illustrazione). Inoltre, ammettiamo la possibilità che le decisioni di investimento nella fase dello sviluppo non siano fatte una volta per tutte all'inizio di questa fase, ma possano essere

modificate nel tempo, col progredire dello sviluppo del farmaco. Come diventerà chiaro tra breve, queste due estensioni sono strettamente collegate.

Supponiamo quindi che per ogni farmaco vi sia una distribuzione di probabilità del tempo di sviluppo τ . L'investimento in sviluppo D modifica la distribuzione di probabilità: è ragionevole ipotizzare che al crescere dell'investimento D il tempo di sviluppo si riduca, nel senso della dominanza stocastica. Una semplice formulazione additiva che cattura questa ipotesi è

$$\tau = \tau(D) + \omega$$

dove il primo termine, $\tau(D)$, è deterministico e ha le stesse proprietà che abbiamo assunto nel precedente capitolo, e ω è una variabile stocastica con valor medio nullo e una data distribuzione di probabilità. L'interpretazione è che la funzione $\tau(D)$ rappresenta il tempo di sviluppo medio, ma il tempo di sviluppo effettivo può essere maggiore o minore di quello medio di un ammontare variabile ma indipendente dal comportamento dell'impresa.

Con questa formulazione, il profitto atteso scontato dell'innovatore all'inizio della fase dello sviluppo è:

$$\begin{aligned} \pi(D) &= E \left[p \int_{\tau}^{T(\tau)+\tau} v e^{-rt} dt - D \right] \\ &= pE \left[\int_{\tau}^{T(\tau)+\tau} v e^{-rt} dt \right] - D, \end{aligned}$$

dove $\tau = \tau(D) + \omega$. La condizione del primo ordine per un massimo diventa:

$$\frac{d\pi}{dD} = pv \{ E [e^{-r[T(\tau)+\tau]}] - E [e^{-r\tau}] + E [T'(\tau)e^{-r[T(\tau)+\tau]}] \} \tau'(D) - 1 = 0,$$

che può essere riscritta come segue:

$$pv \{ E [e^{-r\tau}] - E [e^{-r[T(\tau)+\tau]}] \} - pvE [T'(\tau)e^{-r[T(\tau)+\tau]}] = -\frac{1}{\tau'(D)}.$$

Queste espressioni mostrano come in presenza di incertezza le decisioni di investimento in sviluppo dipendano dall'effetto *medio* di anticipazione dei profitti (*front loading*) e dall'effetto *medio* della durata effettiva. L'effetto medio di *front loading* è la differenza tra la media dei fattori di sconto relativi alla data di commercializzazione e a quella di scadenza del brevetto, e l'effetto medio della durata effettiva è la *media* dei valori *scontati* della pendenza della curva $T(\tau)$. In questo contesto, l'effetto di durata effettiva non dipende solo dalla pendenza della curva $T(\tau)$ in un particolare punto, ma dalla sua pendenza media (nel senso appena precisato).

La seconda estensione del modello, come abbiamo detto, riguarda la dinamica dell'investimento nella fase dello sviluppo. Nel precedente capitolo abbiamo supposto per semplicità che l'investimento in sviluppo venga deciso una volta per tutte all'inizio della fase dello sviluppo. Ma in realtà la fase dello sviluppo è articolata in diverse sottofasi, illustrate nella Figura 1.1, che rappresenta la sequenza della fase pre-clinica (*pre-clinical phase*), che va dal momento in cui viene presentata la domanda per ottenere il brevetto a quello in cui viene presentata la IND (*Investigational New Drug application*),¹⁰ la fase clinica (*clinical phase*), che va dalla data della IND a quella della NDA (*New Drug Application*), e la fase dell'approvazione (*approval phase*), che va dalla data della NDA a quella dell'approvazione finale del farmaco. Il tempo di sviluppo è dato dalla somma di queste tre fasi. Così come il tempo di sviluppo complessivo è stocastico, vi è incertezza sulla lunghezza di ciascuna di queste fasi. Abbiamo anche visto che la fase clinica può essere a sua volta suddivisa in tre fasi (fase I, fase II e fase III). Il progresso da ciascuna fase a quella successiva non è automatico, e nella letteratura sono stati stimati le probabilità di fallimento (*attrition rates*) in ciascuna fase, e la durata di ciascuna delle fasi è incerta.

Nella realtà, le decisioni di investimento delle compagnie farmaceutiche non sono necessariamente predeterminate all'inizio del processo, ma possono essere modificate nel tempo,

¹⁰In realtà i test pre-clinici possono essere iniziati qualche tempo prima che venga presentata la domanda per il brevetto. Tuttavia, il sistema dei brevetti è congegnato in modo da incentivare gli inventori a presentare al più presto la domanda per ottenere il brevetto, per cui la nostra definizione della fase pre-clinica non introduce distorsioni significative.

col progredire dello sviluppo del farmaco da una fase all'altra. In altri termini, le decisioni di investimento in sviluppo non vengono prese al tempo $t = 0$, ma vengono prese sequenzialmente man mano che procede lo sviluppo del farmaco. In condizioni di certezza, il principio di coerenza temporale implica che se anche l'impresa potesse rideterminare i suoi progetti di investimento all'inizio di ciascuna fase, essa vorrebbe semplicemente riconfermare le scelte fatte all'inizio della fase dello sviluppo. Quindi in questo caso l'ipotesi che le decisioni di investimento vengono prese al tempo $t = 0$ non è veramente restrittiva. Ma in condizioni di incertezza le cose sono diverse. Al termine di ciascuna fase, l'impresa può rideterminare le proprie decisioni di investimento alla luce della risoluzione di ogni incertezza che si sia avuta nel frattempo.¹¹

In particolare, poichè la durata di ciascuna fase ha una componente stocastica, al termine di ciascuna fase si sarà risolta una parte dell'incertezza sulla durata. Quindi l'impresa potrà condizionare l'investimento nella fase successiva alla durata delle fasi precedenti. In sostanza, al termine di ciascuna fase l'impresa potrà ricalcolare l'effetto medio di *front loading* e l'effetto medio della durata effettiva in funzione della durata delle fasi precedenti. Questo la porterà a rideterminare l'investimento in sviluppo, influenzando la durata della fase corrente. Questo ragionamento ci porta a concludere che la durata media di ciascuna fase può dipendere dalla durata effettiva delle fasi precedenti.

Nel cercare di sfruttare questa possibile dipendenza per l'analisi empirica, siamo vincolati dal fatto che non abbiamo dati sulla durata delle diverse fasi cliniche, ma solo sulla durata della fase clinica nel suo complesso. E mentre abbiamo dati sulla durata della fase di approvazione, è ragionevole immaginare che la durata di questa fase non possa essere controllata direttamente dall'impresa. Questo ci induce a scomporre il processo di sviluppo in due sole fasi: la fase pre-clinica da una parte, e quella clinica e di approvazione dall'altra.¹² L'idea

¹¹Naturalmente l'impresa potrebbe, all'inizio del processo, determinare l'investimento in ciascuna fase successiva condizionatamente agli eventi che possono verificarsi nel frattempo. In altri termini, non stiamo dicendo che la strategia dell'impresa è temporalmente incoerente.

¹²Una ulteriore ragione per aggregare le fasi clinica e di approvazione è che le compagnie farmaceutiche spesso hanno una certa flessibilità nella scelta del momento in cui presentare l'NDA.

allora è che l'investimento nella fase clinica e di approvazione, e quindi la durata media di queste fasi, possa dipendere dalla durata della fase pre-clinica.

Vediamo ora perchè il fatto che la durata delle fasi clinica e di approvazione possa dipendere dalla durata della fase preclinica consenta di identificare l'effetto di durata effettiva. Abbiamo visto che in presenza di incertezza l'incentivo ad investire dipende dall'effetto medio di durata effettiva, $E [T'(\tau)e^{-r[T(\tau)+\tau]}$. Quando le scelte di investimento sono fatte sequenzialmente, il valore di $E [T'(\tau)e^{-r[T(\tau)+\tau]}$ rilevante per determinare l'incentivo ad investire nella fase clinica è quello condizionato alla durata realizzata della fase pre-clinica. Se la pendenza della curva $T(\tau)$ non è costante, l'effetto di durata effettiva dipende quindi dalla durata realizzata della fase pre-clinica.

L'effetto di *front loading*, invece, non dipende dalla durata della fase pre-clinica. È vero, infatti, che se scontiamo i profitti al tempo zero (l'inizio della fase dello sviluppo) una maggior durata della fase pre-clinica riduce l'effetto di *front loading*, ma è anche vero che si riduce nella stessa proporzione il costo marginale. In altri termini, sia costi marginali che benefici marginali sono scontati per un periodo più lungo. Forse più chiaramente, la stessa conclusione può essere ottenuta considerando che la decisione di investimento nella fase clinica è deciso all'inizio della fase clinica e che l'effetto di *front loading* valutato a quel punto è indipendente da tutto ciò che è successo in precedenza.

Possiamo quindi concludere che ogni genuina dipendenza funzionale della durata della fase clinica (e di approvazione) dalla durata della fase pre-clinica deve catturare l'effetto di durata effettiva. Si può quindi pensare di regredire la durata delle fasi clinica e di approvazione sulla durata della fase preclinica per individuare tale dipendenza.

Anticipando la presentazione della domanda, tuttavia, rischiano di allungare la fase di approvazione perchè potrebbero non essere ancora completati tutti i test previsti per la fase clinica. Di fatto la fase clinica sarebbe accorciata solo nominalmente. Viceversa, se la compagnia farmaceutica presenta l'NDA solo quando è sicura di aver completato tutti i test, la fase clinica risulterà più lunga ma la durata della fase di approvazione si accorcerà. Dunque la durata di ciascuna singola fase può essere soggetta a manipolazioni strategiche da parte dell'impresa, mentre la durata complessiva dipenderà genuinamente dall'investimento in sviluppo.

Purtroppo, le durate di queste due fasi sono verosimilmente correlate anche in maniera spuria. Sappiamo infatti che quanto maggiore è il valore dell'innovazione, tanto più l'impresa avrà incentivo a portarla al mercato rapidamente. In altri termini, per innovazioni di maggior valore le imprese avranno un maggior incentivo ad investire tanto nella fase pre-clinica che nella fase clinica. Questo si tradurrà in una minore durata di entrambe le fasi. Se il valore dell'innovazione è invece basso, ci aspettiamo una maggiore durata di entrambe le fasi. Questo fornisce un motivo per cui la durata della fase pre-clinica deve risultare *correlata positivamente* alla durata delle fasi clinica e di approvazione, anche in assenza di una genuina dipendenza funzionale di queste due variabili.

Se avessimo dati sul valore dei farmaci potremmo eliminare questa correlazione spuria. Ma poiché non li abbiamo, dobbiamo cercare di separare la correlazione spuria dalla dipendenza funzionale genuina in altri modi.

NON MONOTONICITÀ

Una prima possibilità è quella di sfruttare la non-linearità della dipendenza funzionale genuina. Per capire l'origine della non-linearità, occorre ricordare che l'estensione della durata del monopolio di cui gode l'innovatore è determinata da due meccanismi: una vera e propria estensione della durata del brevetto (*patent term extension* negli Stati Uniti; *supplementary protection* in Europa) e la cosiddetta *market exclusivity* (o *data exclusivity*), che impedisce di presentare una domanda di approvazione abbreviata (ANDA) prima di un certo lasso di tempo dall'approvazione originaria del farmaco.¹³ Le Figure 1.3 e 1.4 illustrano il modo in cui la durata effettiva del monopolio di cui gode l'innovatore, T , dipende dal tempo di sviluppo, τ , rispettivamente negli Stati Uniti e in Europa.

¹³Questo secondo istituto fornisce in realtà una protezione più forte del primo, essendo indipendente dalla validità del brevetto originario. In altri termini, se il brevetto originario viene contestato e dichiarato nullo questo annulla anche la sua estensione. Questa decisione, invece, non ha alcuna influenza sulla *market exclusivity*, che continua a valere indipendentemente dall'esistenza del brevetto.

L'effetto di durata effettiva dipende dalla pendenza della curva $T(\tau)$ ed è tanto più forte quanto maggiore, in valore assoluto, è tale pendenza. Come mostrano le Figure 1.3 e 1.4, la pendenza è però variabile: alcuni tratti della curva $T(\tau)$ sono piatti, altri hanno una pendenza pari a -1 , altri ancora hanno una pendenza intermedia tra questi valori. Aspetti comuni agli Stati Uniti e all'Europa sono che: (i) inizialmente la pendenza della curva è, in valore assoluto, massima (pari a -1); (ii) per valori intermedi del tempo di sviluppo, la curva si appiattisce; (iii) per valori più elevati del tempo di sviluppo, la curva torna a mostrare una pendenza (in valore assoluto) elevata; infine, (iv) per tempi di sviluppo estremamente lunghi la curva è di nuovo piatta (poichè il monopolio è di fatto garantito dalla *market exclusivity*).

Consideriamo ora un'impresa che ha completato la fase pre-clinica. Se il tempo impiegato per completare questa fase è molto breve, è probabile che l'impresa finirà per collocarsi nel tratto iniziale (con pendenza elevata) della curva $T(\tau)$. Qui l'effetto di durata effettiva è molto forte. Se il tempo impiegato nella fase pre-clinica è maggiore, sarà invece più probabile che l'impresa finisca per collocarsi nel tratto piatto della curva. In questo caso l'effetto di durata effettiva è più piccolo. Se però il tempo impiegato nella fase pre-clinica è ancora maggiore, sarà più probabile che il progetto si concluda nel secondo tratto ripido della curva $T(\tau)$, dove l'effetto di durata effettiva torna ad essere forte. Infine, l'effetto di durata effettiva sarà nuovamente ridotto quando l'impresa prevede tempi di sviluppo estremamente elevati.

Queste considerazioni implicano che la durata della fase pre-clinica abbia un effetto sull'incentivo dell'impresa ad investire per completare rapidamente la fase clinica (e la fase di approvazione), e che questo effetto sia non monotono: inizialmente (cioè quando la durata della fase pre-clinica è breve) una maggiore durata della fase pre-clinica sarà associata ad una maggiore durata di quella clinica, ma poi l'effetto si inverte e diventa negativo per durate maggiori della fase pre-clinica. Per durate ancora maggiori, l'effetto torna ad essere positivo. I casi in cui i tempi di sviluppo sono così lunghi sono però limitati. Escludendo questi casi, possiamo concludere che *la durata della fase pre-clinica ha un effetto non monotono, prima positivo e poi negativo, sulla durata delle fasi clinica e di approvazione.*

La correlazione spuria è invece sempre positiva. Possiamo quindi pensare di stimare una relazione quadratica tra la durata delle fasi clinica e di approvazione, y , e la durata della fase pre-clinica x :

$$y_i = \alpha + \beta_1 x_i + \beta_2 x_i^2 + \varepsilon_i \quad (3.1)$$

dove ε_i è l'errore stocastico. Le nostre ipotesi sono che $\beta_1 > 0$ e $\beta_2 < 0$. L'ipotesi che $\beta_1 > 0$ cattura sia la correlazione spuria, sia il fatto che inizialmente una fase pre-clinica più lunga tende ad associarsi ad una fase clinica anch'essa più lunga. L'ipotesi che $\beta_2 < 0$ cattura invece la non-monotonicità dell'effetto di durata effettiva.

L'EFFETTO MEDIO DI DURATA EFFETTIVA

Un approccio più diretto consiste nell'usare come variabile esplicativa non la durata della fase pre-clinica, ma la pendenza attesa della curva $T(\tau)$ calcolata alla fine della fase pre-clinica. Questa può essere calcolata utilizzando la distribuzione di probabilità della durata delle fasi clinica e di approvazione. In assenza di un effetto di durata effettiva, questa distribuzione di probabilità non dovrebbe dipendere dalla durata della fase pre-clinica. Usando questa distribuzione, possiamo quindi calcolare la distribuzione di probabilità condizionata del tempo di sviluppo, e quindi il valore medio della pendenza della curva $T(\tau)$,

$$z(x) = E [T'(\tau)e^{-r[T(\tau)+\tau]} | x] .$$

L'equazione da stimare a questo punto diventa

$$y_i = \alpha + \beta z(x_i) + \varepsilon_i \quad (3.2)$$

dove ε_i è ancora l'errore stocastico. La nostra ipotesi in questo caso è che $\beta_1 > 0$. Una curva più piatta (cioè uno z più alto) corrisponde ad un effetto di durata effettiva più piccolo e quindi ad una maggiore durata della fase pre-clinica e di approvazione.

Dopo aver descritto la raccolta dei dati e presentato una serie di statistiche descrittive, procederemo a stimare le equazioni (3.1) e (3.2).

3.4 LA RACCOLTA DEI DATI

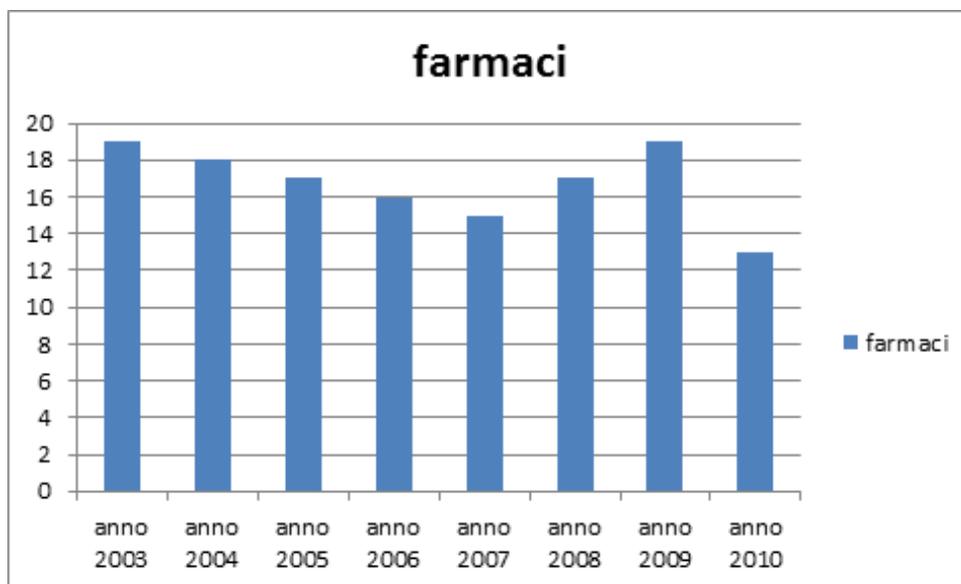
Abbiamo raccolto dati sulla durata delle varie fasi dello sviluppo del farmaco per tutti i nuovi principi attivi (NME) approvati negli Stati Uniti dal gennaio 2003 al dicembre 2010, per un totale di 8 anni. Ritardi amministrativi implicano che dalle nostre fonti non sia possibile ricavare i dati rilevanti sulla durata delle varie fasi per i farmaci approvati più recentemente.

Anche se la *patent term extension* si applica anche a nuove composizioni o nuovi metodi di somministrazione, abbiamo escluso questi farmaci dalla nostra analisi sia perchè sono meno innovativi, sia perchè la protezione brevettuale è meno forte di quella garantita da un brevetto che copre una nuova sostanza chimica.

Per i farmaci sintetizzati chimicamente, l'identificazione delle NME si è basata sulla classificazione operata dalla stessa FDA. Per i farmaci biologici, invece, la FDA non distingue tra NME e nuove composizioni o nuovi metodi di somministrazione. Inoltre la durata della *data exclusivity* per i farmaci biologici è molto maggiore che per i farmaci sintetizzati chimicamente, e quindi gli effetti che ci interessano sono diversi. Abbiamo quindi deciso di escludere i farmaci biologici dal nostro *data set*.

La Figura 3.1 mostra il numero dei nuovi farmaci approvati anno per anno. Recentemente vi è stato un vivace dibattito su una presunta tendenza ad una riduzione del tasso di innovazione nel settore farmaceutico. Direttamente o indirettamente le case farmaceutiche hanno sostenuto la tesi di una riduzione del tasso di innovazione, vista come premessa per giustificare la necessità di una protezione più forte. Ma la Figura 3.1 mostra che non vi è alcun evidente trend negativo nel numero dei nuovi farmaci approvati (a parte il calo del 2010).

L'informazione sui farmaci approvati è tratta dal sito della FDA. Oltre alla distinzione tra farmaci sintetizzati chimicamente e farmaci biologici, la FDA riporta l'eventuale classificazione di un farmaco come "orfano" (*orphan drug*). Questi farmaci, che servono per curare malattie rare, ricevono un trattamento speciale sia in termini di semplicità e rapidità del processo di approvazione, sia perchè godono di un periodo di esclusiva più lungo degli altri.



Oltre che dal sito della FDA, abbiamo tratto le nostre informazioni dal *Federal Register* e dall'Ufficio dei Brevetti (USPTO). Per ottenere la *patent term extension*, il titolare del brevetto deve presentare domanda all'USPTO entro un certo termine. Prima di essere presa in considerazione dall'USPTO che concede l'estensione, la domanda passa però al vaglio dell'FDA, che verifica le date della IND, della NDA e dell'approvazione riportate nella domanda. Al termine di questo processo, i risultati del controllo sono pubblicati sul *Federal Register* (la "Gazzetta Ufficiale" degli Stati Uniti). Fortunatamente, per gli anni più recenti il *Federal Register* è disponibile *on line* (questa è la ragione per cui la nostra raccolta dei dati non si è potuta spingere fino al periodo delle riforme legislative degli anni ottanta, periodo per il quale il *Federal Register* è disponibile solo in forma cartacea). Ricercando sul *Federal Register* possiamo quindi ottenere la durata ufficiale della fase clinica e della fase di approvazione. I dati pubblicati sul *Federal Register* nel giugno 2012 (quando è stata svolta la nostra raccolta dei dati) sono, per l'appunto, quelli relativi ai farmaci approvati nel dicembre 2010.

Il *Federal Register* riporta anche l'estensione richiesta dal titolare del brevetto nella domanda originaria. Spesso però vi sono piccole discrepanze tra le date dell'*IND application* e

della NDA riportate nella domanda e quelle ufficiali fornite dalla FDA. In alcuni casi, poi, la richiesta di estensione non è conforme alla norma. Per queste ragioni l'estensione concessa dall'USPTO può essere diversa (minore ma anche maggiore) da quella richiesta. L'USPTO mantiene un elenco dei brevetti estesi in base all'*Hatch-Waxman Act*: i dati più recenti sono relativi a farmaci approvati tra il 2007 e il 2008. Per i farmaci approvati più recentemente, abbiamo tuttavia proceduto a correggere l'estensione richiesta quando questa, sulla base delle informazioni disponibili, appare incongrua.

La legge stabilisce che l'inventore possa ottenere l'estensione di un solo brevetto per ogni farmaco approvato. In realtà un farmaco può essere protetto da molti brevetti, che possono riguardare formulazioni alternative del principio attivo, il metodo di somministrazione o il processo di produzione del farmaco. In un solo caso l'USPTO ha concesso un'estensione a due brevetti anziché uno. Questo però ha fatto sì che nelle domande più recenti i titolari dei brevetti abbiano presentato richieste di estensione di diversi brevetti (due o tre). Per le domande che non sono ancora state approvate (e per le quali quindi è ancora incerto quale brevetto godrà dell'estensione)¹⁴, abbiamo considerato il brevetto più vecchio. Questo normalmente è il brevetto che copre il principio attivo, e che quindi fornisce la protezione più forte.

Una volta identificato il brevetto che copre il farmaco, il sito dell'USPTO fornisce la data di approvazione del brevetto e la data in cui la domanda è stata presentata. Come abbiamo detto, ai nostri fini la durata della fase pre-clinica è il tempo che passa tra la presentazione della domanda per il brevetto e la presentazione della IND. A questo proposito vanno però fatte due precisazioni.

La prima è che la data della presentazione della domanda non coincide sempre con la data da cui comincia a decorrere il termine dei 20 anni. Infatti, se l'USPTO impiega più di tre anni per concedere il brevetto, vi può essere un *patent term adjustment* che compensa l'innovatore del tempo perso nella fase di approvazione. Si possono quindi distinguere due

¹⁴Questo presumibilmente viene concordato tra il titolare dei brevetti e l'USPTO.

date: la data in cui è stata presentata la domanda, e la data da cui comincia a decorrere il termine.

La seconda precisazione è che non sempre la domanda per il brevetto viene presentata per la prima volta negli Stati Uniti. Quando l'inventore presenta la prima domanda in un altro paese, ha poi un anno di tempo per presentare domanda negli Stati Uniti. Solitamente questo anno di grazia è sfruttato interamente, perchè l'inventore può comunque godere della priorità stabilita dalla data di presentazione nell'altro paese e quindi ha un incentivo a ritardare più che può la presentazione della domanda per godere di una maggior durata effettiva.

Questi fatti implicano che la fase pre-clinica possa essere definita in due modi diversi: *i*) dal momento in cui è stata depositata la prima domanda di brevetto (*filing date*) alla *IND application*; oppure *ii*) dal momento in cui comincia a decorrere il termine dei 20 anni (*starting date*) alla data della *IND application*. Usando queste due diverse definizioni si ottengono valori diversi della durata della fase pre-clinica leggermente diversi. La prima definizione è quella che sembra più appropriata per misurare il tempo effettivamente impiegato nella fase pre-clinica (anzi forse costituisce ancora una sottostima della durata reale di questa fase, perchè i test possono cominciare anche prima del deposito della domanda di brevetto). La seconda definizione è però quella più appropriata per la determinazione della durata effettiva del brevetto, ed è quindi ad essa che faremo riferimento nel resto di questo lavoro.

Dobbiamo anche osservare che i ritardi nell'approvazione dei brevetti non sono sempre imputabili all'USPTO. Spesso l'innovatore abbandona la domanda originaria e ne riformula una nuova alla luce delle conoscenze che ha acquisito nel frattempo. Questo consente di definire le rivendicazioni (*claims*) in modo più preciso e quindi di ottenere una protezione legale più adeguata. La riformulazione può essere ripetuta diverse volte, col risultato che alcuni brevetti vengono concessi molti anni dopo la domanda originaria. Il rovescio della medaglia è che può capitare che l'approvazione della domanda per il brevetto arrivi dopo la presentazione della *IND application*. In questo caso, l'estensione prevista dall'*Hatch-Waxman Act* non è più pari a metà della durata della fase clinica più quella della fase di approvazione, ma solo

alla metà del tempo che passa tra l'approvazione del brevetto e la NDA. Questa situazione non è infrequente, e quindi in linea di principio dovrebbe essere tenuta in considerazione nell'analisi empirica e in particolare nella costruzione della variabile che cattura l'effetto medio di durata effettiva.

3.5 STATISTICHE DESCRITTIVE

Cominciamo presentando alcune semplici statistiche descrittive. La Tabella 3.1 presenta diverse informazioni che smentiscono tesi che vengono spesso date per scontate nel dibattito corrente.

Una prima osservazione è che *il tempo di sviluppo medio è di circa 11 anni e mezzo*, dunque in linea con i dati riportati da Di Masi et al. (2003),¹⁵ ma sensibilmente inferiore a quelli di Pammolli et al. (2010). Questi ultimi autori riportano un valore medio del tempo di sviluppo di circa 14 anni, che è sensibilmente diverso da quello da noi trovato. In parte questa differenza si spiega per il fatto che questi autori prendono come data iniziale della fase di sviluppo quella della presentazione della domanda per il brevetto, mentre noi consideriamo il momento da cui comincia a decorrere il termine dei 20 anni. Questo però spiega solo una parte (poco più di un anno) della differenza tra i nostri risultati e i loro. Poichè i dati usati da Pammolli e i suoi collaboratori non sono disponibili pubblicamente, è impossibile identificare le ragioni della differenza residua, che è di circa 2 anni.

La stima del tempo di sviluppo medio è rilevante perchè da questa dipende la stima della durata media del brevetto. Se il tempo di sviluppo medio fosse veramente pari a 14 anni, dato che l'estensione massima possibile è di 5 anni ma mediamente è molto minore (meno di 3), la durata media del brevetto dovrebbe essere sensibilmente inferiore a 10 anni. Invece, come mostra la Tabella 3.2, *in media*, nel nostro campione di farmaci, *la durata effettiva del*

¹⁵In realtà Di Masi et al. (2003) misurano il tempo di sviluppo a partire dal momento della sintesi della *NME*, mentre noi lo misuriamo dal momento in cui comincia a decorrere il termine dei 20 anni. Quindi i nostri dati suggeriscono che i tempi di sviluppo siano aumentati negli ultimi anni, ma non tanto quanto suggeriscono Pammolli et al. (2010).

*brevetto è di quasi 12 anni.*¹⁶ Di nuovo, questo valore è perfettamente in linea con quello trovato da Di Masi et al. (2003). Possiamo quindi concludere che nell'ultimo decennio non vi è stato alcun deterioramento della durata effettiva della protezione dei nuovi farmaci, contrariamente a quanto viene spesso sostenuto.

Una seconda osservazione immediata è l'estrema variabilità della durata delle varie fasi del processo di sviluppo e del tempo complessivo di sviluppo. Questo varia da meno di un anno e mezzo ad oltre 23 anni. I casi in cui il tempo di sviluppo complessivo è così lungo sono però limitati; in questi casi, la protezione di cui gode l'innovatore è determinata interamente dal meccanismo della *data exclusivity*. *La fase pre-clinica dura in media 3 anni e mezzo*; se però fosse calcolata a partire dalla presentazione della domanda per il brevetto, la durata media supererebbe i 4 anni e mezzo. Vi è un certo numero di farmaci per cui la durata della fase pre-clinica è nulla. Questo in parte è dovuto alla nostra particolare definizione (che comunque non è arbitraria, ma al contrario è quella appropriata ai fini della nostra analisi), in parte al fatto che in alcuni casi le compagnie farmaceutiche possono cercare di rinviare nel tempo il momento in cui presentano la domanda di brevetto – anche se, come abbiamo visto nel primo capitolo, questa strategia è rischiosa. *La fase clinica dura mediamente più di 6 anni*, di nuovo con una estrema variabilità: si va da meno di 8 mesi a oltre 15 anni. La fase di approvazione varia da 2 mesi a 14 anni, con una durata media di circa 1 anno e mezzo. Bisogna dire che il caso di una fase di approvazione lunga 14 anni è veramente un caso estremo, in cui di fatto la compagnia farmaceutica ha dovuto ripetere tutti i test clinici perchè gli studi iniziali furono giudicati totalmente inattendibili. La variabilità della somma delle fasi clinica e di approvazione è minore di quella che ci si dovrebbe attendere se la durata di queste fasi fosse indipendente: vi è invece evidentemente una certa correlazione negativa, che conferma che le compagnie farmaceutiche possono in una

¹⁶Si noti che la durata effettiva media non è data dalla formula:

$$20 - \text{tempo di sviluppo} + \text{estensione},$$

perchè per i farmaci con un tempo di sviluppo molto lungo la durata effettiva del monopolio è uguale alla durata della *data exclusivity*.

certa misura anticipare o posticipare strategicamente il momento in cui presentano l'NDA.

3.6 STIME

La prima strategia che seguiamo per identificare l'effetto di durata effettiva è quella di stimare una relazione non-lineare tra la durata delle fasi clinica e di approvazione (la nostra variabile *Clinapp*) e la durata della fase pre-clinica (*Preclinica*). La relazione stimata è dunque:

$$clinapp = \alpha + \beta_1 Preclinica + \beta_2 Preclinica + \varepsilon \quad (3.3)$$

La stima è stata effettuata con il metodo dei Minimi Quadrati Ordinari (OLS), usando il programma *Stata*. I risultati sono riportati nella Tabella 3.2. I parametri hanno il segno atteso, e il coefficiente β_2 è statisticamente significativo. Come ci si poteva aspettare, il coefficiente di determinazione è piuttosto basso (10%) dal momento che vi sono molte variabili omesse che presumibilmente concorrono a determinare il tempo di sviluppo.

Per esempio, abbiamo classificato i nuovi farmaci in 7 diverse categorie terapeutiche (come antibiotici, anti-tumorali ecc.). Abbiamo introdotto 6 *dummies* per catturare un possibile effetto di queste categorie terapeutiche, ma la stima non risulta migliorata.

La seconda strategia consiste nel costruire una variabile che rappresenta l'effetto di durata effettiva atteso per ciascuna innovazione. Come abbiamo visto, questo dipende dalla durata della fase pre-clinica e dalla distribuzione di probabilità della durata delle fasi clinica e di approvazione. Per prima cosa abbiamo quindi specificato una distribuzione di probabilità appropriata per la durata delle fasi clinica e di approvazione. La Figura 3.2 riporta la distribuzione empirica. La forma di questa distribuzione suggerisce che possa essere approssimata da una distribuzione Gamma. Abbiamo quindi stimato i parametri della distribuzione Gamma partendo dalla distribuzione empirica. Il parametro k è stato stimato come $k = 8.23$, mentre il parametro θ è risultato $\theta = 345$. Sommando la distribuzione di probabilità stimata alla durata realizzata della fase pre-clinica otteniamo la distribuzione di probabilità del tempo di sviluppo, condizionata alla durata realizzata della fase pre-clinica.



Abbiamo quindi supposto, per semplicità, che la pendenza della curva $T(\tau)$ possa assumere soltanto tre valori: -1 per τ compreso tra 0 e 6 anni, 0 per τ compreso tra 6 e 11 anni, e di nuovo -1 per τ maggiore di 11 anni. Questa ipotesi comporta due approssimazioni: *i*) nella realtà, in certi casi (quando i due tetti discussi nel Capitolo 1 non sono stringenti) la pendenza della curva è maggiore di -1 ma comunque negativa, mentre noi supponiamo che possa essere solo -1 o 0 ; *ii*) nella realtà, per valori molto elevati di τ la curva torna ad avere una pendenza nulla. Per ridurre gli errori dovuti a questa seconda approssimazione, abbiamo escluso i farmaci con una durata della fase pre-clinica particolarmente lunga, rimanendo con $n = 128$ osservazioni. La prima approssimazione, invece, non sembra particolarmente problematica data la variabilità ipotizzata della fase clinica e di approvazione.

Procedendo in questo modo, la pendenza media della curva $T(\tau)$, condizionata dalla durata della fase preclinica, diventa semplicemente uguale a 1 meno la probabilità che il tempo di sviluppo, sempre condizionato alla durata della fase pre-clinica, sia compreso tra 6 e 11 anni. Questa probabilità può essere calcolata facilmente, per ciascun farmaco; in questo modo possiamo quindi costruire una variabile ausiliaria che misura la pendenza attesa

della curva $T(\tau)$, condizionata alla durata della fase pre-clinica. Chiamiamo questa variabile *Ependenza*.

Il nostro modello predice che la durata delle fasi clinica e di approvazione dipenda negativamente da *Ependenza*. Come sappiamo, l'analisi teorica suggerisce anche che vi possa essere una correlazione spuria tra la durata delle fasi clinica e di approvazione e la durata della fase pre-clinica. Abbiamo quindi stimato la relazione

$$clinapp = \alpha + \beta_1 Preclinica + \beta_2 Ependenza + \varepsilon \quad (3.4)$$

La stima è stata effettuata con il metodo dei Minimi Quadrati Ordinari (OLS), usando il programma *Stata*. I risultati sono riportati nella Tabella 3.3. Complessivamente la stima risulta migliorata rispetto alla stima della (3.3); il coefficiente di determinazione è quasi raddoppiato. Il coefficiente β_2 ha il segno atteso e è statisticamente significativo al livello dell'1%. Anche il coefficiente β_1 è ora significativo (anche se solo al livello del 5%), ma ha il segno opposto a quello atteso.

3.7 CONCLUSIONI

In questo capitolo abbiamo esaminato due implicazioni del nostro modello teorico che possono essere testate empiricamente. La prima è che vi sia una relazione negativa tra valore del farmaco e tempo di sviluppo. Questa relazione è già stata confermata in diversi studi, che abbiamo discusso nella sezione 2. La seconda, che invece è peculiare al nostro modello, riguarda l'effetto di durata effettiva. Abbiamo discusso come questo effetto possa essere identificato empiricamente in generale, e come i dati che abbiamo raccolto ci consentano di seguire due strade: *i*) regredire la durata della fase clinica su quella della fase preclinica, ricercando un effetto non-lineare negativo; *ii*) costruire una misura dell'intensità attesa dell'effetto di durata effettiva, condizionata alla durata della fase pre-clinica.

Entrambe le strade portano a stime del coefficiente che cattura l'effetto di durata effettiva che sono statisticamente significative e hanno il segno atteso. Nel complesso, i risultati

di questa analisi si possono considerare incoraggianti. Tuttavia, vi è sicuramente spazio per migliorare le stime, sia per quanto riguarda la dimensione del campione di farmaci considerato, sia per quanto riguarda le tecniche di stima (che per ora sono limitate a OLS).

Tabella 3.1: Statistiche descrittive

Variabili	Media	Deviazione standard	Minimo	Massimo
Fase pre-clinica	1346.6	1239.7	0	5965
Fase clinica	2319.8	1145.4	235	5919
Fase di approvazione	522.0	631.0	68	5590
Clinica + approvazione	2841.8	1224.1	494	6768
Tempo di sviluppo	4188.3	1556.8	595	8436
Durata effettiva brevetto	4358.7	1346.5	1826	6808
Estensione durata brevetto	1042.8	672.2	0	1826
Osservazioni	134	134	134	134

Nota: Tutte le durate sono espresse in giorni

Tabella 3.2: Stima della relazione non-lineare

Variabile	<i>Clinapp</i>	
<i>Preclinica</i>	0.24967 (0.23107)	$F(2,131) = 6.33$
<i>Preclinica</i> ²	-0.000116** (0.0000512)	$R^2 = 0.1023$
Osservazioni	134	

Tabella 3.3: Stima con pendenza attesa

Variabile	<i>Clinapp</i>	
<i>Preclinica</i>	-0.21112** (0.0821)	$F(2,125) = 12.64$
<i>Ependenza</i>	-1772.5*** (416.8)	$R^2 = 0.1861$
Osservazioni	128	

Nota: Variabile dipendente: *Clinapp*. Entro parentesi è riportato lo *standard error*. La significatività dei parametri è indicata dalla seguente notazione: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.01$.

CAPITOLO 4

ANALISI NORMATIVA: COMPLETA INFORMAZIONE

Riassunto. Abbiamo visto che il settore farmaceutico è caratterizzato da politiche di estensione della durata dei brevetti che possono essere espresse tramite una funzione, $T(\tau)$, che lega la durata effettiva del brevetto T al tempo di sviluppo τ . Nell'ultima parte della tesi analizziamo la scelta ottimale della funzione $T(\tau)$ da parte di un policy-maker benevolente che si propone come obiettivo la massimizzazione del benessere sociale. In questo capitolo cominciamo l'analisi considerando il caso di completa informazione.

4.1 INTRODUZIONE

Nel settore farmaceutico, il meccanismo di protezione brevettuale opera in modo del tutto particolare, a causa della dettagliata regolamentazione cui è soggetta l'introduzione sul mercato di nuovi farmaci. L'immissione in commercio di un nuovo farmaco richiede una speciale autorizzazione da parte delle autorità preposte e avviene parecchio tempo dopo la concessione del brevetto sulla nuova sostanza chimica (*New Molecular Entity*, NME). La durata effettiva dei brevetti farmaceutici dunque dipende dal tempo necessario per ottenere l'approvazione (cioè per dimostrare che la NME sia efficace e sicura), che è chiamato "tempo di sviluppo" del nuovo farmaco.

La durata effettiva dei brevetti farmaceutici non è però semplicemente uguale alla differenza tra la durata nominale (pari attualmente a 20 anni) e il tempo di sviluppo. Esiste

infatti un complesso sistema di regole legali che permette di prolungare la durata del brevetto, recuperando almeno una parte del tempo di sviluppo. Queste regole sono state introdotte per la prima volta negli Stati Uniti con l'*Hatch-Waxman Act* del 1984, e successivamente sono state adottate, con qualche variante, in Europa, Giappone, Australia, Canada e quasi tutti gli altri Paesi industrializzati. Come abbiamo visto nel primo capitolo, l'estensione della durata del monopolio di cui gode l'innovatore è determinata da due meccanismi. Il primo è un vero e proprio prolungamento della durata del brevetto. Il secondo è la *data exclusivity*, che, rendendo proibitivamente costosa per un certo periodo l'introduzione sul mercato dei farmaci generici, di fatto garantisce una certa durata minima del monopolio per l'inventore.

Nei capitoli precedenti abbiamo analizzato come questi meccanismi influenzino gli incentivi ad investire nella fase di ricerca e in quella dello sviluppo di un nuovo farmaco. In questo capitolo e in quello successivo ne studieremo invece gli aspetti normativi. È giusto che la durata effettiva dei brevetti farmaceutici dipenda dal tempo di sviluppo? Non sarebbe preferibile recuperare completamente il tempo perso, assicurando una durata effettiva costante? O, all'estremo opposto, escludere qualsiasi estensione della durata del brevetto? Qual è, insomma, la relazione ottimale, dal punto di vista sociale, tra la durata effettiva del brevetto e il tempo di sviluppo del nuovo farmaco?

L'analisi di questi problemi non serve solo a soddisfare una legittima curiosità teorica, ma anche a gettar luce su importanti scelte di politica economica. Infatti, le regole di estensione della durata dei brevetti farmaceutici variano nel tempo e da paese a paese, e sono costantemente oggetto di dibattito tra i responsabili di politica economica.¹ In questo lavoro la nostra analisi rimarrà astratta e puramente qualitativa. Tuttavia, come discuteremo nel capitolo conclusivo, in un futuro l'analisi potrebbe essere arricchita di contenuto empirico e condurre ad indicazioni di politica economica più concrete. In ogni caso, la calibrazione

¹Per esempio, vi è attualmente una differenza significativa tra la durata della *data exclusivity* in Europa e negli Stati Uniti, dove sono state presentate diverse proposte per prolungare la *data exclusivity* in modo da armonizzarne la durata con quella europea, o addirittura spingerla oltre. Per una rassegna del dibattito di politica dei brevetti farmaceutici negli Stati Uniti si veda Roin (2009).

empirica deve necessariamente essere preceduta da un'analisi delle proprietà qualitative della soluzione ottimale; ed è questo l'obiettivo che ci proponiamo qui.

Per quanto a nostra conoscenza, l'analisi sviluppata in questo capitolo e in quello successivo è del tutto nuova. Fino ad ora, l'analisi normativa dei brevetti ha quasi sempre considerato la durata del brevetto come uno scalare. Alcuni lavori hanno studiato la combinazione ottimale tra durata e ampiezza,² o la possibilità di condizionare la durata del brevetto ad altre variabili economiche,³ ma nessuno ha studiato come la durata effettiva del brevetto dovrebbe dipendere dal tempo di sviluppo del farmaco – cioè le politiche ottimali di estensione della durata del brevetto.

Mancava, in realtà, un modello teorico di riferimento che incorporasse le peculiarità essenziali del processo innovativo nel settore farmaceutico e quindi consentisse un'analisi formale del problema. Il modello che abbiamo sviluppato nel secondo capitolo colma questa lacuna nella letteratura. Come abbiamo visto, esso ci consente di capire come le politiche di estensione della durata dei brevetti, espresse da una funzione $T(\tau)$ che lega la durata effettiva T al tempo di sviluppo τ , influenzino gli incentivi ad investire nella fase della ricerca e in quella dello sviluppo di un nuovo farmaco. Ora si tratta di analizzare la scelta ottimale della funzione $T(\tau)$ da parte di un policy-maker benevolente, che si propone come obiettivo la massimizzazione del benessere sociale.

L'idea centrale della nostra analisi è che la funzione $T(\tau)$ rappresenti di fatto un menù di possibili opzioni per l'innovatore. Questi, in pratica, può scegliere una durata effettiva del brevetto più lunga velocizzando il processo di approvazione del farmaco. Al contrario, ad uno sviluppo del farmaco meno costoso ma più lento sarà associata una durata effettiva minore (assumendo che la funzione $T(\tau)$ sia decrescente). Naturalmente le scelte dell'innovatore dipenderanno dal valore potenziale del nuovo farmaco. Quando l'innovazione è molto

²Si veda Gilbert e Shapiro (1990) per il contributo pionieristico in questo campo, e Denicolò (1996) per una sintesi.

³Si vedano, ad esempio, Wright (1983), Cornelli e Schankerman (1999), Scotchmer (1999), Shavell e van Ypersele (2001), Hopenhyan et al. (2006) e Chari et al. (2012).

profittevole e quindi l'incentivo ad anticipare i profitti è elevato, l'innovatore sceglierà investimenti maggiori nella fase dello sviluppo, e quindi tempi di sviluppo più rapidi, garantendosi una durata effettiva del monopolio più lunga. Quando invece l'innovazione è meno profittevole, sarà ottimale investire meno nella fase di sviluppo, optando di fatto per una durata effettiva del brevetto più breve. Il problema del policymaker può quindi essere visto come un problema di *mechanism design*.

In questo capitolo consideriamo un contesto di completa informazione, in cui la funzione $T(\tau)$ si applica ad una singola innovazione, le cui caratteristiche sono note al policymaker. In questo caso, il nostro risultato principale è che la soluzione ottimale non prevede alcuna estensione della durata del brevetto – cioè nessuna protezione supplementare o esclusività dei dati. Analiticamente, la soluzione ottimale è $T(\tau) = \Gamma - \tau$, dove la variabile Γ può essere interpretata come la durata nominale del brevetto e in generale può dipendere dalle caratteristiche dell'innovazione.⁴

La ragione per cui non è ottimale fornire alcuna protezione supplementare è che questa diluisce l'incentivo ad investire nella fase dello sviluppo, quando l'innovatore è già protetto dal brevetto e quindi non corre più il rischio di essere anticipato dai concorrenti. Per usare la terminologia introdotta nel Capitolo 2, estendere la durata del brevetto riduce l'effetto di durata effettiva e quindi accentua il problema del sotto-investimento nella fase dello sviluppo.

4.2 LA FUNZIONE DEL BENESSERE SOCIALE

Il modello di riferimento è quello sviluppato nel Capitolo 2, a cui rimandiamo per la spiegazione delle ipotesi e delle condizioni di equilibrio.

Supponiamo che il policymaker scelga la funzione $T(\tau)$ in modo da massimizzare il benessere sociale, tenendo conto del comportamento del settore privato (che abbiamo già analizzato nel Capitolo 2). Il policymaker può scegliere la funzione $T(\tau)$ liberamente, a

⁴Naturalmente, con completa informazione la funzione $T(\tau)$ è determinata solo localmente, cioè in un intorno del punto che sarà effettivamente scelto dall'innovatore. Il risultato è quindi, più precisamente, che la pendenza della funzione nel punto rilevante è uguale a -1 .

parte alcune naturali restrizioni che discutiamo subito. Primo, supponiamo che la funzione $T(\tau) : R_+ \rightarrow R_+$ debba essere continua e differenziabile tranne che eventualmente in un numero finito di punti. Questa è una ipotesi di regolarità che è comune nella letteratura di teoria degli incentivi. Secondo, ipotizziamo che quando la funzione $T(\tau)$ è differenziabile, debba essere $-1 \leq T'(\tau) \leq 0$. Questo significa che la funzione non è crescente e che tempi di sviluppo più lunghi non possono essere associati a date di scadenza del brevetto più vicine nel tempo: al massimo, la durata effettiva si riduce nella stessa misura in cui aumenta il tempo di sviluppo. Si noti che le funzioni $T(\tau)$ attualmente adottate in pratica, illustrate per esempio nelle Figure 1.3 e 1.4, soddisfano entrambe le condizioni. Discuteremo la condizione $-1 \leq T'(\tau)$ più ampiamente in seguito, quando considereremo situazioni nelle quali il vincolo diventa stringente.

Utilizzeremo la nozione di benessere sociale che è comune nelle analisi di equilibrio parziale, ovvero il valore atteso della somma scontata del surplus dei consumatori e dei produttori. Ricordando la decomposizione del valore sociale dell'innovazione, in termini di flusso, nelle tre parti v , s e d illustrate nella Figura 2.2, possiamo scrivere il benessere sociale atteso, in generale, come:

$$W = \theta(R) \left\{ p \left[\int_{\tau}^{\tau+T} (v + s)e^{-rt} dt + \int_{\tau+T}^{\infty} (v + s + d)e^{-rt} dt \right] - D \right\} - R. \quad (4.1)$$

Il significato di questa espressione è il seguente. Con probabilità $1 - \theta(R)$, la ricerca non ha successo e la NME non viene scoperta. In questo caso la società non ottiene alcun beneficio, ma sostiene comunque i costi di ricerca R . Con probabilità $\theta(R)$, la NME viene scoperta e si passa allo stadio dello sviluppo. In questo stadio, l'innovatore investe D e scopre al tempo τ se il nuovo farmaco può essere approvato o no. Se viene approvato (cosa che avviene con probabilità p), l'innovazione è protetta dal brevetto per un periodo pari a T , mentre dal tempo $\tau + T$ in poi diventa liberamente disponibile (cioè le versioni generiche del farmaco possono essere commercializzate). Quando l'innovazione è coperta da brevetto, il benessere sociale è pari al profitto v più il surplus dei consumatori s ; vi è una perdita secca d dovuta

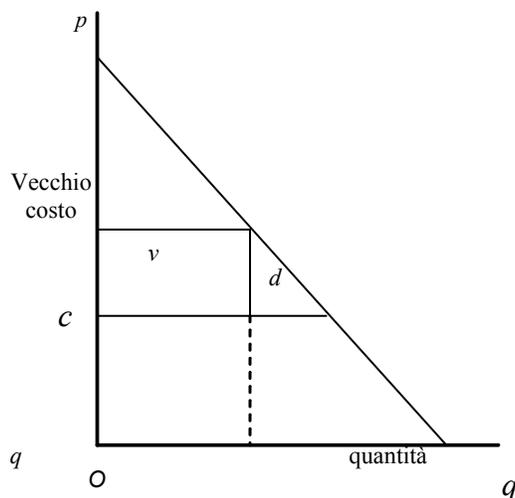


Figura 4.1: Il caso di un'innovazione non drastica

all'imposizione di un prezzo superiore al costo marginale. Quando il brevetto scade, profitto e perdita secca vengono azzerati e il surplus dei consumatori diventa pari a $v + s + d$.⁵

Ci sono due casi particolari in cui l'espressione per il benessere sociale si semplifica, e che pertanto meritano una considerazione particolare. Il primo si ha quando l'innovazione è non-dradastica. Questo caso è illustrato nella Figura 4.1. La situazione rappresentata nella figura è quella di un'innovazione che riduce il costo di produzione di un farmaco, portandolo a c .⁶ Si suppone che il prezzo di monopolio (che non è riportato nella figura) sia maggiore

⁵Dovrebbe essere chiaro a questo punto che, dopo che il brevetto è scaduto, v e d sono solo componenti ideali del surplus dei consumatori e non rappresentano più, rispettivamente, il profitto e la perdita secca. Per comodità di notazione, però, non introduciamo un nuovo simbolo ma indichiamo il surplus dei consumatori alla scadenza del brevetto con $v + s + d$.

⁶Anche se nulla impedisce di applicare l'estensione della durata del brevetto a innovazioni di processo, è molto più frequente che essa si applichi a innovazioni di prodotto. (La ragione è che la legge prevede che per ogni farmaco si possa estendere un solo brevetto, anche se il farmaco è protetto da diversi brevetti, e normalmente il brevetto che copre la formula chimica del principio attivo è quello più importante e più facile da far valere.) Tuttavia, anche un'innovazione di prodotto può essere concettualizzata secondo lo schema della Figura 4.1. Si può infatti pensare alla curva di domanda come alla domanda di un certo "effetto terapeutico," e a un nuovo e più efficace farmaco come ad un modo per produrre lo stesso "effetto terapeutico" ad un costo inferiore. In effetti, è frequente il caso in cui NME siano usate per curare malattie per cui già esiste una cura, rispetto

del vecchio costo. Se prima dell'innovazione il mercato era perfettamente competitivo, o se il nuovo innovatore e il precedente *incumbent* competono alla Bertrand, l'innovatore praticherà un prezzo limite, appena inferiore al vecchio costo. Ciò significa che finché l'innovazione è coperta dal brevetto, i consumatori di fatto non ottengono alcun beneficio rispetto alla situazione iniziale, per cui si ha $s = 0$.⁷ L'innovatore ottiene un flusso di profitti pari a v , mentre d rappresenta la perdita secca dovuta al fatto che il prezzo eccede il costo marginale. Quando invece il brevetto scade, il prezzo scende al livello del nuovo costo marginale c e la società ottiene l'intero valore sociale dell'innovazione, $v + d$, nella forma di surplus dei consumatori.

Dunque il caso di innovazione non drastica è simile a quello di un'innovazione drastica (come la scoperta di un farmaco radicalmente innovativo), con la sola differenza che $s = 0$. In questo caso quindi il benessere sociale diventa

$$W = \theta(R) \left\{ p \left[\int_{\tau}^{\tau+T} v e^{-rt} dt + \int_{\tau+T}^{\infty} (v + d) e^{-rt} dt \right] - D \right\} - R. \quad (4.2)$$

Un caso ancora più particolare è quello di una curva di domanda rettangolare, come in Green e Scotchmer (1995). Questo caso è illustrato nella Figura 4.2. In questo caso abbiamo $s = d = 0$, e v rappresenta sia il valore sociale che quello privato dell'innovazione (di nuovo in termini di flusso). Il caso $s = d = 0$ si può verificare anche con una curva di domanda "regolare," se l'innovatore è in grado di praticare una discriminazione perfetta di prezzo. Quale che sia la ragione per cui s e d possono annullarsi, se lo fanno il benessere sociale si riduce a

$$W = \theta(R) \left[p \int_{\tau}^{\infty} v e^{-rt} dt - D \right] - R. \quad (4.3)$$

alla quale la NME rappresenta un miglioramento marginale. Si pensi, ad esempio, alle generazioni successive di farmaci anti-retrovirali per la cura dell'AIDS.

⁷Si tenga presente che s indica non il livello del surplus dei consumatori ma l'aumento creato dall'innovazione. Quando l'innovazione crea un nuovo prodotto, queste due variabili coincidono ma esse differiscono quando l'innovazione migliora una tecnologia preesistente.

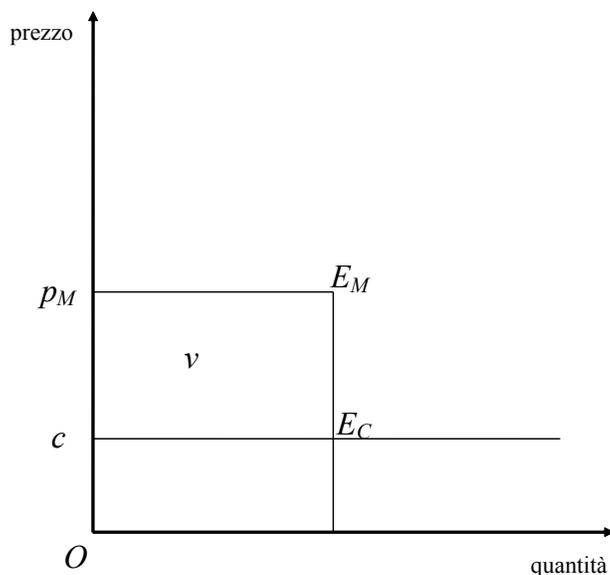


Figura 4.2: L'equilibrio nel mercato del prodotto con una curva di domanda rettangolare.

Nell'analisi che segue faremo spesso riferimento a questi due casi particolari, che permettono di semplificare l'analisi o di dare un'interpretazione più chiara dei risultati.

Come abbiamo detto, in questo capitolo consideriamo un contesto di completa informazione. In altri termini supponiamo che la funzione $T(\tau)$ si applichi ad una singola innovazione, le cui caratteristiche sono note al policy-maker. Si suppone cioè che il policy-maker conosca perfettamente sia le funzioni $\theta(R)$ e $\tau(D)$ che i parametri p , v , s , d e r .

4.3 CURVA DI DOMANDA RETTANGOLARE

Per chiarire meglio la logica che sta alla base dei nostri risultati, cominciamo dal caso di una curva di domanda rettangolare, in cui $s = d = 0$. Questa ipotesi semplifica notevolmente l'analisi. Infatti, la determinazione della funzione $T(\tau)$ presuppone un'analisi della durata ottimale del brevetto – a cui si aggiunge la complicazione che tale durata non è necessariamente costante. L'analisi della durata ottimale del brevetto è incentrata sul *trade-off*

di Nordhaus tra incentivi ad investire in R&D ed inefficienza statica (Nordhaus, 1969). L'ipotesi di curva di domanda rettangolare elimina completamente tale trade-off, perchè implica che il monopolio non provochi alcuna inefficienza statica. Questo è molto utile sul piano espositivo, perchè consente di concentrare l'analisi sugli aspetti di maggiore novità e interesse della nostra analisi, che sono legati alle politiche di estensione del brevetto e al ruolo che possono giocare per generare gli obiettivi appropriati ad investire nei due stadi dell'innovazione farmaceutica, ricerca e sviluppo.⁸

4.3.1 IL TRADE-OFF TRA RICERCA E SVILUPPO

Si potrebbe pensare che se si pone $d = 0$, astraendo completamente dal costo sociale del monopolio in termini di inefficienza statica, la politica ottimale debba prevedere una durata del brevetto infinita. Ma nel nostro modello non è così. La ragione è che la nostra ipotesi di libertà di entrata nello stadio della ricerca implica una tendenza al sovra-investimento in ricerca, per via dell'effetto *winner-take-all*.⁹ In particolare, se la durata del brevetto fosse infinita, il valore privato dell'innovazione coinciderebbe con il valore sociale e l'effetto di *winner-take-all* genererebbe sicuramente sovra-investimento in ricerca.¹⁰ Per ovviare a questo problema, la durata del brevetto deve essere finita, in modo che il valore privato dell'innovazione sia inferiore al valore sociale, contrastando così la tendenza al sovra-investimento.

Tuttavia, se da un lato ridurre il valore privato al di sotto del valore sociale elimina il sovra-investimento nello stadio della ricerca, dall'altro lato crea un problema di sotto-investimento nella fase dello sviluppo. Anche se l'ipotesi di curva di domanda rettangolare

⁸Allo stesso modo, Green e Scotchmer (1995) fanno l'ipotesi di curva di domanda rettangolare nella loro analisi pionieristica dell'innovazione sequenziale al fine di concentrare l'attenzione sul trade-off tra la protezione offerta al primo e al secondo innovatore.

⁹Per una discussione dell'effetto *winner-take-all* si veda ad esempio Denicolò e Franzoni (2010) e la letteratura ivi citata.

¹⁰Se vi fosse invece monopolio nello stadio della ricerca, con una curva di domanda rettangolare sarebbe sufficiente fissare una durata infinita del brevetto per generare gli incentivi appropriati ad investire sia nella fase della ricerca che in quella dello sviluppo. Questa politica consentirebbe di riprodurre esattamente la soluzione di first-best. Con monopolio nello stadio della ricerca, l'analisi di benessere diventa non banale solo se si suppone che $d > 0$.

elimina il trade-off di Nordhaus, vi è quindi comunque un trade-off tra la necessità di garantire gli incentivi appropriati ad investire nello stadio della ricerca e nello stadio dello sviluppo. Per usare la terminologia della teoria della politica economica, un solo strumento (la durata del brevetto) non consente, in generale, di raggiungere due obiettivi indipendenti (il livello appropriato di investimento in ricerca R e in sviluppo D).

Qui entra però in gioco la possibilità di legare la durata effettiva del brevetto al tempo di sviluppo. Come abbiamo visto nel Capitolo 2, questa dipendenza crea un effetto di durata effettiva che incentiva l'investimento nello stadio dello sviluppo, fornendo di fatto al policymaker un secondo strumento. Si tratta di vedere se i due strumenti (cioè il valore della funzione $T(\tau)$ e la sua pendenza $T'(\tau)$) sono sufficienti a risolvere il problema del sovra-investimento nello stadio della ricerca e del sotto-investimento nello stadio dello sviluppo.

Con completa informazione e nel caso $s = d = 0$, la risposta a questo interrogativo è positiva: è cioè possibile utilizzare gli strumenti a disposizione del policymaker, ovvero la durata nominale del brevetto e la dipendenza della durata effettiva dal tempo di sviluppo, per riprodurre esattamente la soluzione di first best. Chiaramente, una funzione $T(\tau)$ che riproduce la soluzione di first best è necessariamente quella ottimale.¹¹

4.3.2 LA SOLUZIONE DI FIRST BEST

La soluzione di first best è quella che si ottiene immaginando che il policymaker possa controllare direttamente la scelta degli investimenti in ricerca e sviluppo, R e D , e che queste variabili siano determinate in modo da massimizzare W .

Indichiamo con \hat{R} e \hat{D} la soluzione di first best, e, per semplicità di notazione, poniamo $\hat{\theta} = \theta(\hat{R})$ e $\hat{\tau} = \tau(\hat{D})$. Con alcuni semplici passaggi algebrici, le condizioni del primo ordine

¹¹Di fatto con completa informazione la funzione $T(\tau)$ risulta determinata solo in un intorno del punto $\hat{\tau}$ (che come vedremo rappresenta il tempo di sviluppo di first best). Una volta scelti opportunamente $T(\hat{\tau})$ e $T'(\hat{\tau})$, la specificazione del resto della funzione $T(\tau)$ è cioè irrilevante. Come vedremo, una certa indeterminatezza permane anche con informazione incompleta nel caso di un numero finito di “tipi,” ma scompare con un continuo di tipi.

per un massimo del benessere sociale (4.3) possono essere riscritte come segue:¹²

$$pv e^{-r\hat{\tau}} = -\frac{1}{\tau'(\hat{D})}. \quad (4.4)$$

e

$$pv \int_{\hat{\tau}}^{\infty} e^{-rt} dt - \hat{D} = \frac{1}{\theta'(\hat{R})}. \quad (4.5)$$

La condizione (4.4) richiede che il beneficio sociale marginale di una riduzione del tempo di sviluppo (il primo membro della (4.4)) sia uguale al suo costo marginale (il secondo membro). Il beneficio sociale marginale è dato semplicemente dal valore sociale dell'innovazione scontato dal tempo τ (il momento in cui l'innovazione diventa disponibile) a zero. La condizione (4.5) richiede invece che il costo marginale di un aumento della probabilità di successo (il secondo membro della (4.5)) sia uguale al corrispettivo rendimento marginale, che è il valore sociale di una NME. Questo è dato dal primo membro della (4.5).

La soluzione del sistema di equazioni (4.4) e (4.5) può essere ottenuta recursivamente. La condizione (4.4) determina direttamente \hat{D} . Intuitivamente, la scelta di D deve massimizzare il termine entro parentesi quadre nella (4.3), che è il valore sociale di una NME. Come per il corrispondente problema privato analizzato nel Capitolo 2, date le nostre ipotesi sulla funzione $\tau(D)$ la soluzione è necessariamente interna. Posto che $D = \hat{D}$ (e quindi $\tau = \hat{\tau}$), il valore sociale della NME a quel punto è dato, e quindi la (4.5) determina il valore ottimale dell'investimento nella fase della ricerca, \hat{R} .

4.3.3 IMPLEMENTAZIONE

Consideriamo ora un'economia di mercato in cui le decisioni di investimento in R&D sono prese da imprese private che massimizzano il proprio profitto; il policymaker può quindi controllare tali investimenti solo indirettamente, attraverso la politica dei brevetti $T(\tau)$. Mostriamo che nel caso di una curva di domanda rettangolare la soluzione di first best possa

¹²Nel caso generale, la soluzione di first best è data da formule simili a queste, con $(v + s + d)$ al posto di v .

essere riprodotta nell'equilibrio di mercato attraverso la scelta di una opportuna funzione $T(\tau)$.

Per prima cosa (e questo è il punto cruciale di questa discussione), dimostriamo che ponendo $T(\tau) = \Gamma - \tau$ si garantisce che $D^* = \hat{D}$. Come abbiamo visto sopra, il parametro Γ è interpretabile come la durata nominale del brevetto, mentre porre (nella notazione introdotta nel Capitolo 2) $b = 0$ significa che non vi è alcun recupero del tempo di sviluppo. La dimostrazione dell'asserto è immediata confrontando la condizione del primo ordine che determina l'investimento in sviluppo nell'equilibrio di mercato (cioè l'equazione (2.2) – o, equivalentemente, la (2.3) – del Capitolo 2) con la (4.4). Quando $T(\tau) = \Gamma - \tau$ e quindi $T'(\tau) = -1$, la condizione (2.3) diventa

$$pve^{-r\tau} = -\frac{1}{\tau'(D)}$$

e quindi coincide con la condizione (4.4).

La ragione intuitiva per cui vale questo risultato è che il beneficio sociale marginale di una riduzione del tempo di sviluppo è semplicemente il valore sociale dell'innovazione scontato dal tempo τ a zero. Per esempio, se il tempo di sviluppo viene ridotto di un ammontare Δt , la società può godere dell'innovazione per un periodo addizionale pari a Δt al tempo τ . Il maggior beneficio sociale è allora pari a $v\Delta t$, scontato dal tempo τ al tempo zero. Se non vi è alcuna estensione della durata del brevetto, il beneficio privato marginale di una riduzione del tempo di sviluppo è anch'esso uguale al valore dell'innovazione scontato dal tempo τ , e quindi coincide col beneficio sociale marginale. Pertanto l'incentivo privato ad investire nello stadio dello sviluppo viene perfettamente allineato all'incentivo sociale.

Per chiarire ancora meglio, consideriamo il caso estremo opposto, in cui l'estensione della durata del brevetto è uguale al tempo di sviluppo. In questo caso, il beneficio privato marginale di ridurre il tempo di sviluppo sarebbe dato dall'effetto di anticipazione dei profitti. Come sappiamo, questo è pari alla differenza tra il valore dell'innovazione scontato dal tempo τ e quello scontato dal tempo $T + \tau$. Quindi l'incentivo privato ad investire nello stadio dello

sviluppo sarebbe inferiore all'incentivo sociale. La stessa conclusione continuerebbe a valere anche se vi fosse una estensione parziale della durata del brevetto.

Avendo garantito che $D^* = \hat{D}$, rimane da dimostrare che si può avere anche $R^* = \hat{R}$. La dimostrazione di quest'ultima parte del risultato è molto semplice. Intuitivamente, dal momento che con una durata infinita del brevetto si ha necessariamente sovra-investimento nello stadio della ricerca (a causa dell'effetto di *winner-take-all*) è sufficiente ridurre in misura adeguata la durata nominale del brevetto per generare il livello di investimento in ricerca socialmente desiderabile.¹³

Possiamo quindi concludere che con una curva di domanda rettangolare la politica ottimale è $T(\tau) = \Gamma - \tau$ per un opportuno valore della durata nominale del brevetto Γ .

4.4 INNOVAZIONI NON DRASTICHE

Dimostriamo ora che questa conclusione vale più in generale. Ci concentreremo dapprima sul caso di innovazioni non drastiche (in cui $s = 0$ ma $d > 0$); poi mostreremo intuitivamente che il risultato vale anche con innovazioni drastiche.

Quando la perdita secca d è positiva, il monopolio implica un'inefficienza statica e quindi il nostro modello presenta il classico trade-off di Nordhaus. Questo significa che la soluzione di first best non può più essere riprodotta in un'economia di mercato, neanche in condizioni

¹³Formalmente, consideriamo la condizione di libertà di entrata (la (5) del Capitolo 2). Sfruttando il fatto che $D^* = \hat{D}$ e $T(\tau) = \Gamma - \tau$, questa può essere riscritta come segue:

$$p \int_{\hat{\tau}}^{\Gamma} v e^{-rt} dt - \hat{D} = \frac{R^*}{\theta(R^*)}.$$

Questa condizione determina R^* in funzione di Γ . Il primo membro di questa espressione è crescente in Γ mentre il secondo membro è decrescente in R^* , quindi R^* sarà una funzione continua e crescente di Γ . Si osservi poi che il primo membro si annulla per un valore di Γ sufficientemente basso, mentre per $\Gamma = \infty$ la condizione diventa

$$\frac{1}{\theta'(\hat{R})} = \frac{R^*}{\theta(R^*)},$$

che per la concavità della $\theta(R)$ implica $R^* > \hat{R}$. Per il teorema del valor medio esiste quindi un valore finito di Γ tale che $R^* = \hat{R}$.

di completa informazione.¹⁴ In particolare, per limitare l'inefficienza statica del monopolio, la durata effettiva del brevetto dovrà essere ridotta in misura tale da creare sotto-investimento nella fase della ricerca. Oltre al trade-off di Nordhaus, rimane però operante anche il trade-off, che abbiamo messo in luce nella precedente sotto-sezione, tra la necessità di garantire il livello appropriato di investimenti nella fase della ricerca da un lato e in quella dello sviluppo dall'altro. La politica ottimale ha un compito complicato, perchè deve bilanciare in modo appropriato questi diversi trade-off.

Dimostriamo per prima cosa che di nuovo la soluzione ottimale non prevede alcuna estensione della durata del brevetto, e quindi è data ancora dalla formula $T(\tau) = \Gamma - \tau$.

Proposizione 1. *La politica dei brevetti ottimale con completa informazione non comporta alcuna estensione della durata del brevetto. In altre parole, $\tilde{T}(\tau) = \Gamma - \tau$ e quindi $\tilde{T}'(\tau) = -1$*

Dimostrazione. Cominciamo calcolando il valore del benessere sociale in corrispondenza dell'equilibrio di mercato. Utilizzando la condizione di libertà di entrata (2.4) del Capitolo 2, che implica che i profitti attesi sono nulli, è facile vedere che il benessere sociale atteso diventa

$$W^* = \theta^* p \int_{T(\tau^*) + \tau^*}^{\infty} (v + d)e^{-rt} dt. \quad (4.6)$$

Intuitivamente, il benessere sociale è uguale al valore scontato del surplus del consumatore, poichè i profitti attesi sono pari a zero per via della condizione di libertà di entrata nella corsa al brevetto. In altre parole, il valore dell'innovazione dal tempo τ^* fino alla scadenza del brevetto copre esattamente i costi totali di ricerca e sviluppo, il cui valore scontato atteso

¹⁴Dal punto di vista terminologico, chiamiamo *soluzione di first best* quella che si ottiene immaginando che il policymaker possa controllare direttamente la scelta degli investimenti in ricerca e sviluppo, R e D ; *soluzione ottimale con completa informazione* quella che si ottiene immaginando che il policymaker scelga la funzione $T(\tau)$ conoscendo perfettamente le caratteristiche dell'innovazione; e *soluzione ottimale con incompleta informazione* (o soluzione di second best, anche se a rigore si dovrebbe parlare di third best) quella in cui il policymaker sceglie la funzione $T(\tau)$ conoscendo solo la distribuzione di probabilità delle caratteristiche dell'innovazione. Per ridurre il rischio di confusione, indichiamo con \hat{x} il valore di una generica variabile x nella soluzione di first best, con \tilde{x} la soluzione ottimale con informazione completa e con \check{x} la soluzione ottimale con informazione incompleta.

è $R^* + \theta(R^*)D^*$. Ciò che rimane per la società è il valore sociale dell'innovazione, $v + d$, da $T(\tau^*) + \tau^*$ in poi.

Si noti che il benessere sociale dipende dalla durata del brevetto $T(\tau)$ non solo direttamente ma anche indirettamente, dato che come sappiamo la funzione $T(\tau)$ determina R^* e D^* , e quindi θ^* e τ^* . Un esame della (4.6) rivela che una condizione necessaria per un massimo è che la scelta della funzione $T(\tau)$ massimizzi la probabilità di successo $\theta^* = \theta(R^*)$ fra tutte le funzioni $T(\tau)$ per cui il valore di $T(\tau^*) + \tau^*$ è costante. In altre parole, la soluzione deve risolvere il problema

$$\begin{aligned} & \max \theta^* \\ & \text{sotto il vincolo} \quad T(\tau^*) + \tau^* = \Upsilon \quad (\text{costante}). \end{aligned} \quad (4.7)$$

La variabile Υ può essere interpretata come un indice del livello assoluto di protezione brevettuale.

La soluzione di questo problema, $T(\tau, \Upsilon)$, determina il valore massimo di θ^* in funzione di Υ . Indichiamo tale funzione con la notazione $\tilde{\theta}^*(\Upsilon)$. Data la funzione $\tilde{\theta}^*(\Upsilon)$, il benessere sociale (6) può essere espresso in funzione di Υ :

$$W^*(\Upsilon) = p\tilde{\theta}^*(\Upsilon) \int_{\Upsilon}^{\infty} (v + d)e^{-rt} dt.$$

Il problema del policymaker a questo punto si riduce a trovare il valore di Υ che massimizza $W^*(\Upsilon)$.

Questo ragionamento mostra che il problema di trovare la funzione ottimale $T(\tau)$ si può scomporre in due stadi: trovare la funzione $T(\tau)$ ottimale per ogni dato valore di Υ , cioè $T(\tau, \Upsilon)$, e poi trovare il valore ottimale di Υ .

Per dimostrare il nostro risultato è sufficiente considerare il primo stadio. Dato che come sappiamo $\frac{d\theta^*}{d\pi^*} > 0$ (si veda la Proposizione 1 del Capitolo 2), massimizzare θ^* equivale a massimizzare

$$\pi^* = p \int_{\tau^*}^{\Upsilon} v e^{-rt} dt - D(\tau^*).$$

La condizione di primo ordine per un massimo è

$$pve^{-r\tau^*} = -\frac{1}{\tau'(D)}. \quad (4.8)$$

Indichiamo con $\tilde{\tau}$ la soluzione di questa equazione. Si noti che $\tilde{\tau} < \hat{\tau}$, essendo $\hat{\tau}$ ora dato da una condizione simile alla (4.4) ma con $(v + d)$ al posto di v . (Questo significa, come sopra anticipato, che la soluzione ottimale con completa informazione implica sotto-investimento nella fase dello sviluppo rispetto alla soluzione di first best.)

Sappiamo che in una economia di mercato τ^* deve soddisfare la condizione (2.3). Combinando tale condizione con la (4.8) otteniamo

$$pv [1 + T'(\tau)] e^{-r[T(\tau)+\tau]} = 0,$$

la cui soluzione è evidentemente

$$T'(\tau) = -1.$$

Questo dimostra che non è ottimale prevedere alcuna estensione della durata del brevetto.

■

La Proposizione 1 può essere illustrata graficamente nel modo seguente. Per un dato valore di π^* , l'incentivo ad investire nella fase della ricerca, e quindi la probabilità di ottenere l'innovazione, θ^* , sono dati. Il problema sociale, che è quello di massimizzare la (4.6), si riduce quindi alla minimizzazione dell'estremo inferiore dell'integrale, $T + \tau$. Graficamente, il problema è allora quello di minimizzare $T + \tau$ sotto il vincolo che all'innovatore sia garantito un livello pre-determinato del profitto. La Figura 4.3 mostra come la soluzione di questo problema richiede che la pendenza della curva di iso-profitto, nel punto che sarà scelto dall'innovatore, deve essere pari a -1 . Ma allora anche la pendenza della curva $T(\tau)$ in quel punto deve essere pari a -1 .

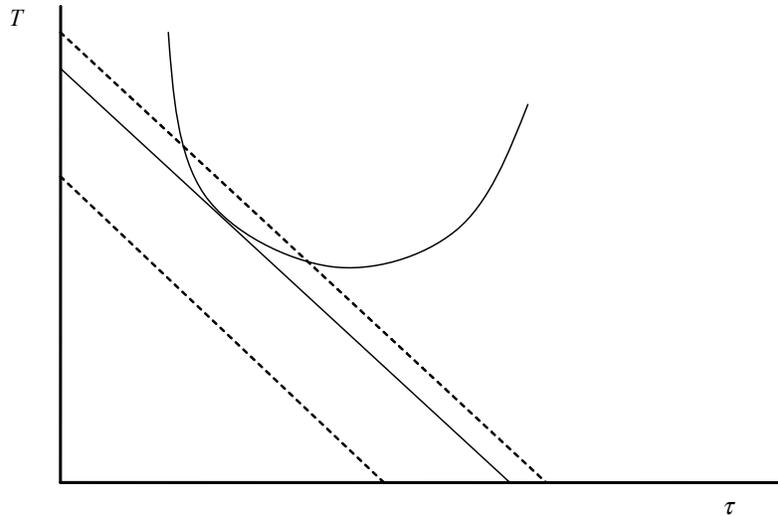


Figura 4.3: Illustrazione grafica della Proposizione 1

4.5 INNOVAZIONI DRASTICHE

La Proposizione 1 continua a valere anche con innovazioni drastiche. Per capire il perché, si noti che tenendo conto della condizione di libertà di entrata, in generale il benessere sociale (4.1) può essere riscritto come:

$$W^* = \theta^* p \left[\int_{\tau^*}^{T(\tau^*)+\tau^*} s e^{-rt} dt + \int_{T(\tau^*)+\tau^*}^{\infty} (v + s + d) e^{-rt} dt \right]. \quad (4.9)$$

Rispetto al caso $s = 0$ abbiamo un termine in più che è una funzione crescente di τ^* . Questo termine addizionale riflette il fatto che ora, anche al netto dei costi di ricerca e di sviluppo, la società ottiene un beneficio netto positivo s anche prima che il brevetto scada.

Quando $s = 0$, il pianificatore sociale (come abbiamo visto nella dimostrazione della Proposizione 1) di fatto sceglie la funzione $T(\tau)$ semplicemente in modo da massimizzare $\pi^* = p \int_{\tau^*}^{\tau} v e^{-rt} dt - D(\tau^*)$ a parità di $T(\tau^*) + \tau^*$, o, equivalentemente, in modo da minimizzare $T + \tau$ per un dato π^* . Quando $s > 0$, a causa del termine addizionale che abbiamo appena

discusso, il pianificatore sociale ha un incentivo addizionale a ridurre τ . Idealmente, quindi, l'effetto di durata effettiva dovrebbe essere ancora più forte che nel caso $s = 0$. Dovrebbe cioè essere $T'(\tau) < -1$.

Se questo non è possibile, per via del vincolo $T'(\tau) \geq -1$, la politica ottimale resta $T'(\tau) = -1$, ovvero $T(\tau) = \Gamma - \tau$. Il vincolo $T'(\tau) \geq -1$ può essere giustificato sulla base della considerazione che i brevetti sono diritti di proprietà, e in quanto tali non possono subire limitazioni per il fatto di non essere esercitati. Dunque se l'innovatore perde tempo e non sviluppa l'innovazione, al massimo la scadenza del brevetto può non venire prolungata, ma non può essere accorciata.

A dire il vero questa giustificazione del vincolo $T'(\tau) \geq -1$ non è inattaccabile. Già attualmente è prevista una limitazione di diritti di proprietà intellettuale che non vengono esercitati, come nel caso della licenza obbligatoria di un brevetto. Più in generale, il policy-maker potrebbe proteggere l'innovatore con strumenti legali che non siano soggetti a regole generali che possono sembrare, dal punto di vista economico, arbitrarie. Se non si vuole imporre il vincolo $T'(\tau) \geq -1$, rimane comunque vero che la soluzione ottimale con completa informazione non prevede alcuna estensione della durata nominale del brevetto – in effetti, l'estensione ottimale sarebbe in questo caso negativa.

4.6 LA DURATA OTTIMALE DEL BREVETTO

Avendo dimostrato che la politica ottimale con completa informazione è $\tilde{T}(\tau) = \Gamma - \tau$, consideriamo ora la scelta ottimale della durata nominale del brevetto, Γ . Abbiamo visto come la caratterizzazione della durata ottimale sia semplice nel caso di una curva di domanda rettangolare. Il problema, nel caso generale, è però più complicato. Cominciamo ancora l'analisi dal caso di innovazioni non drastiche ($s = 0$); il caso di innovazioni drastiche sarà trattato brevemente alla fine di questa sezione.

Per ragioni che diventeranno chiare tra breve, non conviene ragionare direttamente in termini della durata nominale Γ , ma piuttosto in termini di una variabile ad essa collegata. Più precisamente, osserviamo che se il tempo di sviluppo è fissato al valore $\hat{\tau}$ che si ottiene quando $T'(\tau) = -1$, se la durata nominale del brevetto fosse infinita l'innovatore conseguirebbe profitti pari a $p \int_{\hat{\tau}}^{\infty} v e^{-rt} dt - D(\hat{\tau})$. Con una durata nominale finita Γ , invece, i profitti sono soltanto $p \int_{\hat{\tau}}^{\Gamma} v e^{-rt} dt - \tilde{D}$ (dove $\tilde{D} = D(\hat{\tau})$). Una variabile collegata a Γ e che può essere vista come una misura alternativa dell'intensità della protezione brevettuale è il rapporto tra profitti effettivi e profitti massimi teorici (Denicolò, 2007)

$$\Phi = \frac{p \int_{\hat{\tau}}^{\Gamma} v e^{-rt} dt - \tilde{D}}{p \int_{\hat{\tau}}^{\infty} v e^{-rt} dt - \tilde{D}}.$$

Tale rapporto varia tra 0 (quando $\Gamma \leq \hat{\tau}$ e quindi la durata effettiva è nulla) e 1 (quando $\Gamma = \infty$ e quindi anche la durata effettiva è infinita).

Utilizzando la variabile Φ , la caratterizzazione della durata nominale ottimale è molto semplice.

Proposizione 2. *La politica ottimale dei brevetti con completa informazione prevede una durata nominale tale per cui il rapporto tra i profitti uguaglia l'elasticità dell'offerta di innovazioni:*

$$\tilde{\Phi} = \eta, \tag{4.10}$$

dove $\eta = \frac{\theta'(\hat{R})\hat{R}}{\hat{\theta}}$.

Dimostrazione. Quando $T(\tau) = \Gamma - \tau$, il benessere sociale (4.6) diventa

$$W^* = \theta^* p \int_{\Gamma}^{\infty} (v + d) e^{-rt} dt, \tag{4.11}$$

dove θ^* è dato dalla condizione di libertà di entrata (2.4). Dato che con $T(\tau) = \Gamma - \tau$ deve necessariamente essere $\tau = \hat{\tau}$, la condizione di libertà d'entrata può essere riscritta come:

$$\frac{R(\theta^*)}{\theta^*} = p \int_{\hat{\tau}}^{\Gamma} v e^{-rt} dt - \tilde{D}.$$

Quindi θ^* è in ultima analisi una funzione di Γ . La derivata di questa funzione si può calcolare differenziando implicitamente la precedente espressione, ottenendo:

$$\frac{d\theta^*}{d\Gamma} = \frac{r\theta^* \int_{\hat{\tau}}^{\infty} v e^{-rt} dt}{\frac{dR(\theta^*)}{d\theta} - \left[p \int_{\hat{\tau}}^{\Gamma} v e^{-rt} dt - \tilde{D} \right]}.$$

La condizione del primo ordine per la massimizzazione del benessere sociale, $\frac{dW}{d\Gamma} = 0$, si riduce facilmente a

$$\frac{d\theta^*}{d\Gamma} - r\theta^* = 0.$$

Sostituendo in questa formula il valore di $\frac{d\theta^*}{d\Gamma}$ calcolato sopra otteniamo

$$\frac{dR(\theta^*)}{d\theta} = R'(\theta^*) = p \int_{\hat{\tau}}^{\infty} v e^{-rt} dt - \tilde{D}.$$

Utilizzando nuovamente la condizione di libertà d'entrata (5) del Capitolo 2 abbiamo infine

$$R'(\theta^*) \frac{\theta^*}{R(\theta^*)} = \frac{p \int_{\hat{\tau}}^{\infty} v e^{-rt} dt - \tilde{D}}{\Gamma \int_{\hat{\tau}}^{\infty} v e^{-rt} dt - \tilde{D}},$$

da cui segue immediatamente la (4.10). ■

Si noti che con una curva di domanda rettangolare la (4.10) implementa esattamente la soluzione di first best. Per dimostrarlo, confermando così il risultato già ottenuto in una precedente sezione, è sufficiente sostituire la (4.10) nella condizione di libertà di entrata (2.4)

e osservare che la condizione così ottenuta coincide con la condizione (4.5), che determina il valore di first best di θ .

Per l'analisi del caso di incompleta informazione, che sarà sviluppata nel prossimo capitolo, è importante capire anche come varia la durata nominale ottimale del brevetto, che è determinata dalla regola (4.10), al variare del valore dell'innovazione. La risposta a questo interrogativo è fornita dal seguente risultato.

Proposizione 3. *Se l'elasticità dell'offerta di innovazioni $\frac{R\theta'(R)}{\theta}$ è costante o decrescente in R , nella soluzione ottimale con completa informazione la durata nominale del brevetto Γ è una funzione decrescente del valore dell'innovazione v .*

Dimostrazione. Supponiamo per cominciare che l'elasticità dell'offerta di innovazioni $\eta = \frac{R\theta'(R)}{\theta}$ sia costante. Differenziando implicitamente la condizione (4.10) si ottiene

$$\frac{d\Gamma}{dv} = -\frac{\frac{\partial\Phi}{\partial v}}{\frac{\partial\Phi}{\partial\Gamma}}.$$

Chiaramente,

$$\frac{\partial\Phi}{\partial\Gamma} = \frac{pve^{-r\Gamma}}{p\int_{\tilde{\tau}}^{\infty} ve^{-rt} dt - \tilde{D}} > 0.$$

Per determinare il segno di $\frac{d\Gamma}{dv}$ dobbiamo quindi trovare il segno di $\frac{\partial\Phi}{\partial v}$. Abbiamo

$$\begin{aligned} \frac{\partial\Phi}{\partial v} \propto & \left[p\int_{\tilde{\tau}}^{\Gamma} e^{-rt} dt - pve^{-r\tilde{\tau}}\tau'(\tilde{D})\frac{d\tilde{D}}{dv} - \frac{d\tilde{D}}{dv} \right] \left[p\int_{\tilde{\tau}}^{\infty} ve^{-rt} dt - \tilde{D} \right] + \\ & - \left[p\int_{\tilde{\tau}}^{\infty} e^{-rt} dt - pve^{-r\tilde{\tau}}\tau'(\tilde{D})\frac{d\tilde{D}}{dv} - \frac{d\tilde{D}}{dv} \right] \left[p\int_{\tilde{\tau}}^{\Gamma} ve^{-rt} dt - \tilde{D} \right] \end{aligned}$$

dove il simbolo \propto significa “ha lo stesso segno di.” Ma la condizione (4.8) implica che

$$-pve^{-r\tilde{\tau}}\tau'(\tilde{D})\frac{d\tilde{D}}{dv} - \frac{d\tilde{D}}{dv} = 0,$$

quindi abbiamo

$$\frac{\partial H}{\partial v} \propto p\int_{\tilde{\tau}}^{\Gamma} e^{-rt} dt \left(p\int_{\tilde{\tau}}^{\infty} ve^{-rt} dt - \tilde{D} \right) - p\int_{\tilde{\tau}}^{\infty} e^{-rt} dt \left(p\int_{\tilde{\tau}}^{\Gamma} ve^{-rt} dt - \tilde{D} \right),$$

da cui si ottiene immediatamente

$$\frac{\partial H}{\partial v} \propto \tilde{D}p \int_{\Gamma}^{\infty} v e^{-rt} dt > 0.$$

Ne segue che se l'elasticità η è costante abbiamo $\frac{d\Gamma}{dv} < 0$.

Se invece l'elasticità non è costante, la differenziazione implicita porta a

$$\frac{d\Gamma}{dv} = -\frac{\frac{\partial H}{\partial v} - \frac{\partial \eta}{\partial v}}{\frac{\partial H}{\partial \Gamma}}.$$

Chiaramente, R è una funzione crescente di v . Ne segue che se l'elasticità η è decrescente in R , rimane vero *a fortiori* che la durata nominale ottimale decresce al crescere del valore dell'innovazione. ■

L'intuizione è che se la durata nominale del brevetto fosse costante, il rapporto tra i profitti Φ sarebbe una funzione crescente del valore dell'innovazione. Ma tale rapporto deve essere costante se l'elasticità di offerta di innovazioni è costante. Perchè il rapporto tra i profitti rimanga costante, la durata nominale del brevetto deve diminuire al crescere di v .

Questo ragionamento continua a valere se l'elasticità di offerta di innovazioni è una funzione decrescente dell'investimento in ricerca. Il risultato potrebbe però essere invertito se l'elasticità dell'offerta di innovazioni crescesse abbastanza velocemente al crescere dell'investimento in ricerca. In questo caso la durata nominale del brevetto potrebbe essere costante o una funzione crescente di v . Come vedremo nel prossimo capitolo, la politica ottimale con completa informazione dipende precisamente da quale di questi casi si realizzi.

Abbiamo quindi fornito una caratterizzazione completa della politica ottimale con completa informazione per il caso $s = 0$. Senza ripetere in dettaglio tutti i passaggi analitici, discutiamo brevemente come cambia l'analisi nel caso di innovazioni drastiche, cioè quando $s > 0$. In questo caso, la “regola dell'elasticità” (4.10) va modificata semplicemente come segue:

$$\tilde{\Phi} = \frac{v + s + d}{v + d} \eta. \quad (4.12)$$

La dimostrazione è del tutto analoga a quella della Proposizione 2 ed è pertanto omessa.

La Proposizione 3 non cambia di una virgola se si suppone che quando varia v anche s e d varino nella stessa proporzione, come appare ragionevole e come infatti supporremo nella prossima sezione. (Se invece s fosse indipendente da v , il risultato che la durata nominale ottimale del brevetto è una funzione decrescente di v sarebbe addirittura rafforzato.)

Possiamo quindi concludere che tutte le principali conclusioni che abbiamo ottenuto per il caso di innovazioni non drastiche si estendono a quello di innovazioni drastiche. Confortati da questa osservazione, continueremo a concentrarci sul caso $s = 0$ anche nell'analisi della politica ottimale con informazione incompleta.

4.7 CONCLUSIONI

In questo capitolo abbiamo analizzato la scelta della funzione $T(\tau)$ in condizioni di completa informazione. Abbiamo visto che la politica ottimale non prevede alcuna estensione della durata del brevetto, e abbiamo caratterizzato la durata ottimale grazie ad una semplice regola, la “regola dell’elasticità,” che dice che il rapporto tra i profitti (una misura indiretta della durata del brevetto) deve essere uguale all’elasticità dell’offerta di innovazioni. Questi risultati sono stati dimostrati per il caso di innovazioni non-dramatiche, ma con semplici adattamenti si applicano anche al caso di innovazioni drastiche.

CAPITOLO 5

ANALISI NORMATIVA: INCOMPLETA INFORMAZIONE

Riassunto. Abbiamo visto che se il policymaker può determinare una diversa funzione $T(\tau)$ per ogni specifica innovazione, non deve essere prevista alcuna estensione della durata del brevetto. In questo capitolo studiamo la politica ottimale quando una stessa funzione $T(\tau)$ si applica ad innovazioni eterogenee: in particolare, ad innovazioni di diverso valore e diversa profittabilità. In questo contesto più realistico, il nostro risultato principale è che non sia ottimale estendere la durata del brevetto quando il tempo di sviluppo è relativamente breve, mentre una certa estensione è ottimale quando il tempo di sviluppo è più lungo.

5.1 INTRODUZIONE

Abbiamo visto che se il policymaker può specificare una funzione $T(\tau)$ diversa per ogni diversa innovazione, cioè con completa informazione, non deve essere prevista alcuna estensione della durata del brevetto. In pratica, però, non è possibile prevedere una politica diversa per ogni singola innovazione, vuoi per ragioni di semplicità amministrativa, vuoi perchè le caratteristiche specifiche dell'innovazione spesso non sono osservabili o verificabili. Quindi una stessa funzione $T(\tau)$ deve potersi applicare ad innovazioni eterogenee: in particolare, ad innovazioni di diverso valore e diversa profittabilità.

In questo contesto più realistico, la determinazione della funzione $T(\tau)$ può essere vista come un problema di *screening* ottimale. La teoria degli incentivi ci insegna che la soluzione

ottimale richiede che al soggetto informato (nel nostro caso, l'innovatore) venga proposto un menù di scelte, determinato in modo tale da indurre un opportuno processo di auto-selezione. La teoria ci insegna anche che la politica ottimale con completa informazione non è generalmente implementabile in condizioni di asimmetria informativa. Nel nostro contesto, se il policymaker continuasse a porre $T(\tau) = \Gamma - \tau$, come sarebbe ottimale fare in condizioni di completa informazione, la durata effettiva del brevetto verrebbe a dipendere dal valore dell'innovazione in un modo che non è necessariamente ottimale. La soluzione che sarebbe ottimale con completa informazione deve quindi essere distorta per tener conto dei vincoli di compatibilità con gli incentivi.

In questo contesto, il nostro risultato principale è che non sia ottimale estendere la durata del brevetto quando il tempo di sviluppo è relativamente breve, mentre una certa estensione è ottimale quando il tempo di sviluppo è più lungo. Sotto questo aspetto, la politica adottata attualmente sembra, almeno nelle sue linee di fondo, corretta. Come abbiamo visto nel primo capitolo, infatti, sia negli Stati Uniti che in Europa non vi è alcuna estensione della durata nominale del brevetto quando il tempo di sviluppo è inferiore a 5-6 anni, ma vi è un'estensione per tempi di sviluppo maggiori.

Una giustificazione superficiale della politica attuale potrebbe immaginare che il legislatore abbia semplicemente voluto ridurre i costi amministrativi del meccanismo, considerando che se il tempo di sviluppo è breve l'estensione sarebbe comunque limitata. Tuttavia, questa spiegazione non regge ad un esame più attento. La procedura necessaria per estendere la durata del brevetto è semplice e non richiede scelte discrezionali, e in pratica le compagnie farmaceutiche richiedono l'estensione anche quando questa è limitata a pochi giorni. Ciò suggerisce che i costi amministrativi, sia per le autorità pubbliche che per le imprese private, siano molto bassi. La nostra analisi mostra invece che è ottimale non estendere la durata del brevetto se il tempo di sviluppo è sufficientemente breve, a prescindere dai costi amministrativi.

Le politiche di estensione della durata del brevetto rappresentano precisamente la distorsione che è ottimale introdurre nelle situazioni di informazione incompleta. Nessuna distorsione è però necessaria per i “tipi alti” (cioè, nel nostro caso, per le innovazioni di valore più elevato). Questo ben noto principio di *no-distortion-at-the-top* discende dal fatto che i tipi bassi non hanno comunque incentivo a riprodurre le scelte ottimali per i tipi alti. Nel nostro caso, questo principio implica che per le innovazioni di valore più elevato, e quindi con tempi di sviluppo più brevi, la politica ottimale con informazione incompleta coincida con quella con informazione completa, che come abbiamo visto non prevede alcuna estensione della durata nominale.

In conclusione, la nostra analisi fornisce una spiegazione sia per il tratto iniziale della curva di equazione $T(\tau)$, in cui sia negli Stati Uniti che in Europa la pendenza è uguale a -1 , sia per il tratto successivo, in cui la curva tende ad appiattirsi. L’ottimalità del tratto iniziale si spiega in base al principio di *no-distortion-at-the-top*.¹ L’ottimalità del tratto successivo, in cui la curva $T(\tau)$ ha una pendenza maggiore di -1 (cioè vi è una qualche estensione della durata nominale del brevetto), si spiega in base alla necessità di distorcere la soluzione di completa informazione per tener conto dei vincoli di compatibilità con gli incentivi. Al livello di astrazione a cui si colloca l’analisi, non è possibile dire se il modo preciso in cui viene prevista un’estensione della durata nominale sia ottimale. E, a dire il vero, la forma particolare delle politiche adottate in pratica, come quelle rappresentate nelle Figure 1.3 e 1.4, in cui tratti orizzontali si alternano a tratti con pendenza pari a -1 , appare piuttosto bizzarra e difficile da giustificare nel dettaglio. Una idea più precisa sulla forma ottimale della funzione $T(\tau)$ può essere ottenuta solo calibrando empiricamente il modello – un esercizio che però esula dai limiti di questo lavoro.

¹Questo risultato ha forse un interesse più generale. Per quanto a nostra conoscenza, infatti, il principio di *no-distortion-at-the-top* è stato generalmente considerato come un principio teorico privo di implicazioni pratiche rilevanti. (Questo è sicuramente vero per la teoria della tassazione ottimale del reddito.) Nel nostro contesto, invece, non solo le implicazioni sono chiare ma sono anche evidentemente soddisfatte dalle politiche adottate in pratica!

5.2 INCOMPLETA INFORMAZIONE

In linea di principio, le innovazioni cui si applica una stessa funzione $T(\tau)$ possono differire l'una dall'altra da molti punti di vista: le NME possono essere più o meno difficili da scoprire, possono risultare efficaci e sicure con diverse probabilità, possono avere tempi di sviluppo più o meno lunghi e più o meno costosi da accorciare, possono essere più o meno profittevoli, ecc. È quindi evidente che in generale l'eterogeneità tra innovazioni è multi-dimensionale.

Purtroppo però la teoria del *multi-dimensional screening* è generalmente complicata e intrattabile; fino ad ora, nella letteratura sono state fornite soluzioni solo per alcuni casi particolari del tutto speciali. Per ragioni di semplicità analitica, e seguendo in ciò la vasta maggioranza della letteratura sulla teoria degli incentivi, assumeremo quindi che l'eterogeneità sia unidimensionale.

Ci concentreremo, in particolare, sulle differenze nel valore dell'innovazione. Nel nostro modello, tale valore è determinato dal prodotto tra due parametri: la probabilità che la NME sia efficace e sicura, p , e la profittabilità dell'innovazione, v . Se si assume, come sembra naturale, che s e d siano proporzionali a v (in altri termini, che $s = \sigma v$ e $d = \delta v$), il prodotto pv determina univocamente sia gli incentivi privati ad investire in R&D che il benessere sociale. Assumeremo pertanto che il prodotto pv possa variare da un'innovazione all'altra, rappresentando il fattore di eterogeneità nella nostra analisi. Anche se in effetti è solo il prodotto pv che conta, per comodità di notazione supporremo che p sia costante e solo v vari.

Nel resto di questo capitolo analizzeremo allora la scelta ottimale della funzione che lega la durata effettiva del brevetto al tempo di sviluppo in un contesto di asimmetria informativa. Nel momento in cui scelgono quanto investire in R&D, gli innovatori conoscono il valore esatto di v , mentre quando sceglie la funzione $T(\tau)$ il policymaker conosce soltanto la distribuzione di probabilità di v . Questa ipotesi riflette il fatto che la funzione $T(\tau)$ è fissata per un periodo di tempo considerevole, nell'arco del quale si applica a innovazioni molto diverse tra cui il policymaker non può fare distinzioni.

Applicando il principio di rivelazione, possiamo immaginare che il policymaker scelga la durata effettiva del brevetto T e la durata del tempo di sviluppo τ per ciascun tipo di innovazione v sotto opportuni vincoli di compatibilità con gli incentivi.

5.2.1 LA SORTING CONDITION

Cominciamo osservando che anche con informazione incompleta il policymaker ha la possibilità di discriminare tra diversi tipi di innovazioni. La ragione è che il nostro modello soddisfa una opportuna *sorting condition*, il che implica che il policymaker possa implementare funzioni $T(v)$ e $\tau(v)$ strettamente monotone, se queste sono ottimali, inducendo tipi diversi a scegliere diverse combinazioni dal “menù” $T(\tau)$.

Per dimostrare la *sorting condition*, notiamo prima di tutto che l’innovatore sceglie dal “menù” $T(\tau)$ nel momento in cui decide l’investimento in sviluppo, D , e quindi quando ha già innovato e conseguito il brevetto su una NME. A quel punto, invertendo la funzione $\tau(D)$ il profitto del titolare del brevetto può essere scritto in funzione di T , τ e v , come segue:

$$\pi(T, \tau, v) = p \int_{\tau}^{T+\tau} v e^{-rt} dt - D(\tau).$$

Consideriamo le curve di isoprofitto nel piano (τ, T) , lungo le quali per ogni dato v il profitto π è costante. Vale il seguente risultato:

Lemma 1. (*Sorting condition*) La pendenza delle curve di isoprofitto $\frac{dT}{d\tau}$ è decrescente in v .

Dimostrazione. Differenziando implicitamente l’equazione $\pi(T, \tau, v) = k$, per k costante, si ottiene:

$$\frac{dT}{d\tau} \Big|_{\pi=\text{cost}} = \frac{(1 - e^{-rT}) + \frac{D'(\tau)}{vp} e^{r\tau}}{e^{-rT}}$$

da cui il lemma discende immediatamente. ■

Sappiamo dal Capitolo 2 che l’equilibrio si ha sicuramente in un tratto decrescente della curva di isoprofitto. Una pendenza maggiore significa quindi, graficamente, che la curva è più piatta.

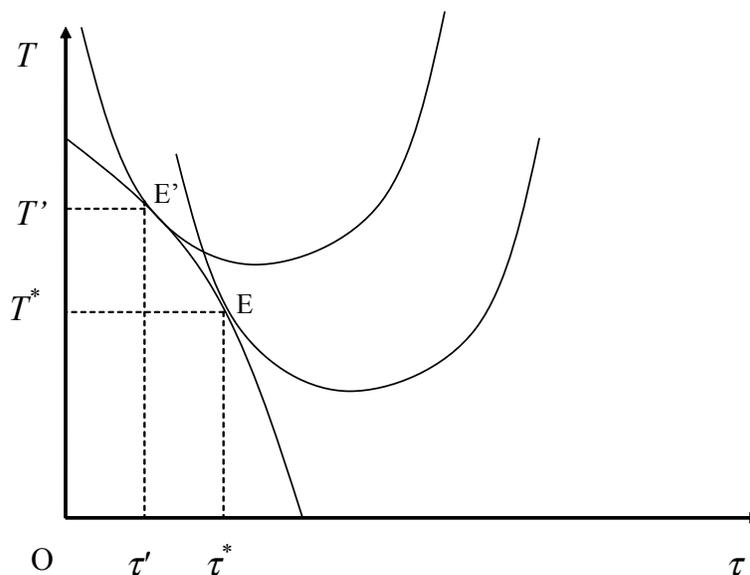


Figura 5.1: Effetti di un aumento del valore dell'innovazione sull'equilibrio nello stadio dello sviluppo.

Intuitivamente, il valore di un prolungamento del brevetto è tanto maggiore quanto maggiore è il valore dell'innovazione. Quindi al crescere del valore dell'innovazione il titolare del brevetto è disposto ad accettare un maggiore aumento dell'investimento in sviluppo per ottenere un dato aumento della durata effettiva del brevetto, come mostra la figura 5.1.

5.2.2 IL PROGRAMMA OTTIMALE

Per procedere, consideriamo per semplicità il caso di due tipi: il valore dell'innovazione può essere alto, per la precisione pari a v_H , con probabilità q , oppure basso, cioè pari a $v_L < v_H$, con probabilità $1 - q$.

Con due tipi di innovazioni, il policymaker di fatto propone due sole opzioni. Indichiamo con (T_H, τ_H) la combinazione di durata effettiva e tempo di sviluppo che in equilibrio sarà scelta dal titolare del brevetto su una innovazione di valore elevato (v_H , il “tipo alto”) e

con (T_L, τ_L) quella che sarà scelta da chi ha ottenuto un'innovazione di valore relativamente minore (v_L , il “tipo basso”).²

Per ciascun tipo di innovazione, nell'equilibrio di mercato il benessere sociale è dato da un'espressione simile alla (4.6). Il benessere sociale atteso è quindi

$$W = q\theta(R_H)p \int_{T_H+\tau_H}^{\infty} (1+\delta)v_H e^{-rt} dt + (1-q)\theta(R_L)p \int_{T_L+\tau_L}^{\infty} (1+\delta)v_L e^{-rt} dt. \quad (5.1)$$

Questa funzione deve essere massimizzata rispettando i vincoli di “compatibilità con gli incentivi:”

$$\pi_L = p \int_{\tau_L}^{T_L+\tau_L} v_L e^{-rt} dt - D(\tau_L) \geq p \int_{\tau_H}^{T_H+\tau_H} v_L e^{-rt} dt - D(\tau_H) \quad (5.2)$$

$$\pi_H = p \int_{\tau_H}^{T_H+\tau_H} v_H e^{-rt} dt - D(\tau_H) \geq p \int_{\tau_L}^{T_L+\tau_L} v_H e^{-rt} dt - D(\tau_L), \quad (5.3)$$

e tenendo conto che gli investimenti in ricerca R_H e R_L sono dati dalle condizioni di libertà di entrata

$$\frac{\theta(R_H)}{R_H} = \pi_H \quad (5.4)$$

e

$$\frac{\theta(R_L)}{R_L} = \pi_L. \quad (5.5)$$

I vincoli di compatibilità con gli incentivi assicurano che il tipo alto scelga effettivamente la combinazione (T_H, τ_H) e il tipo basso la combinazione (T_L, τ_L) . Si noti che la condizione di libertà di entrata deve valere separatamente per i due tipi perchè le imprese conoscono il vero valore dell'innovazione quando decidono il livello dell'investimento nella fase della ricerca.

Il programma di ottimizzazione che abbiamo appena descritto presenta due particolarità rispetto ai problemi simili tipicamente considerati nella letteratura. In primo luogo, non sono

²Questo significa che in realtà la curva $T(\tau)$ risulta determinata, per quanto riguarda il suo livello e la sua pendenza, solo localmente, in un intorno dei due punti (T_H, τ_H) e (T_L, τ_L) . (Nel caso di completa informazione, come si ricorderà, la curva è determinata solo in un intorno di un punto.) Aumentando il numero di tipi possibili, l'indeterminatezza però si riduce e scompare completamente nel caso di un continuo di tipi.

possibili trasferimenti monetari tra principale e agente, e quindi indirettamente tra diversi tipi di agenti. La seconda particolarità è che il problema di ottimizzazione sociale non è soggetto a veri e propri vincoli di partecipazione. Tuttavia, le probabilità di successo θ_H e θ_L possono essere interpretate come il “grado di partecipazione” di ciascun tipo. Questo assimila il nostro problema ai problemi di screening con *random participation*: si veda al riguardo Rochet e Stole (2001).

La soluzione del programma ottimale non è quindi un’applicazione meccanica di risultati tradizionali. Cominciamo però da un caso particolare in cui ricompaiono i vincoli di partecipazione nella loro forma standard.

5.3 IL CASO DI PARTECIPAZIONE DETERMINISTICA

Supponiamo che la funzione $\theta(R)$ abbia la forma a “gradino” illustrata nella Figura 5.2: la probabilità di successo è nulla se l’investimento è inferiore a un certo valore critico \bar{R} ed è pari a 1 se l’investimento è maggiore o uguale a tale valore. Con questa ipotesi, nella fase della ricerca non vi è alcuna incertezza: per ottenere l’innovazione è sufficiente investire un certo ammontare minimo.

Questo implica che l’innovazione non sarà ottenuta se $\pi^* < \bar{R}$, e sarà ottenuta certamente se $\pi^* \geq \bar{R}$. Per rendere il problema interessante, supponiamo che sia ottimale fare in modo che vengano ottenute anche le innovazioni di tipo basso: questo richiede che v_L non sia troppo più piccolo di v_H e che q non sia troppo grande. Nel caso queste condizioni non valgano, il problema diventa banale in quanto si riduce nella sostanza ad un problema di informazione completa: il policymaker di fatto può ragionare come se tutte le innovazioni fossero di tipo alto.³

³Naturalmente questa politica ha un costo, che è quello di rinunciare alle innovazioni di valore più basso. Tuttavia, se v_L è piccolo e/o q è grande, questo costo è trascurabile.

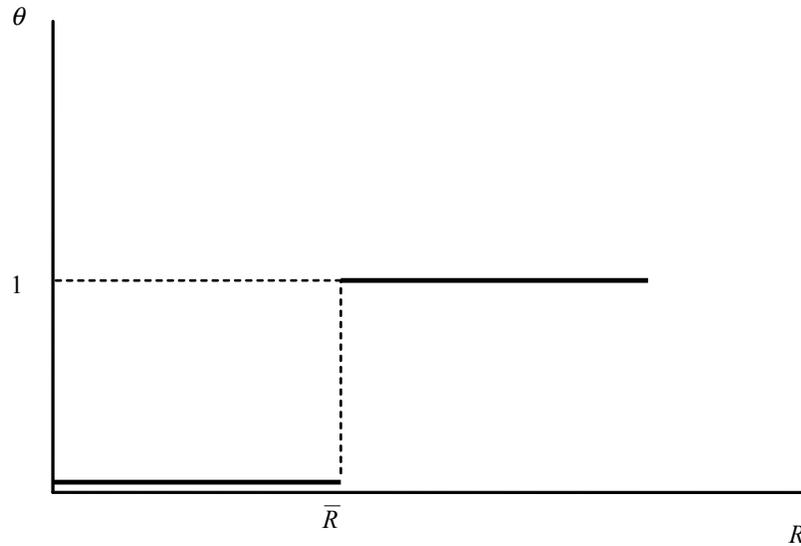


Figura 5.2: Una funzione $\theta(R)$ a gradini

Le condizioni (5.4) e (5.5) devono quindi essere sostituite dai seguenti vincoli di partecipazione:

$$\pi_H \geq \bar{R} \quad (5.6)$$

$$\pi_L \geq \bar{R}. \quad (5.7)$$

È subito evidente che il vincolo (5.6) non sarà mai binding, ma più in generale è facile dimostrare il seguente risultato:

Lemma 2. *Nel problema di massimizzazione del benessere sociale, i vincoli stringenti sono il vincolo di partecipazione del tipo basso (5.3) e il vincolo di compatibilità con gli incentivi del tipo alto (5.2).*

Dimostrazione. La dimostrazione è del tutto standard ed è quindi omessa. ■

Dato che $\theta(R_H) = \theta(R_L) = 1$, il benessere sociale atteso diventa

$$W = qp \int_{T_H + \tau_H}^{\infty} (1 + \delta)v_H e^{-rt} dt + (1 - q)p \int_{T_L + \tau_L}^{\infty} (1 + \delta)v_L e^{-rt} dt. \quad (5.8)$$

Sappiamo inoltre che il vincolo di partecipazione del tipo basso deve essere stringente

$$p \int_{\tau_L}^{T_L+\tau_L} v_L e^{-rt} dt - D(\tau_L) = \bar{R}, \quad (5.9)$$

e così pure il vincolo di compatibilità con gli incentivi del tipo alto:

$$\int_{\tau_H}^{T_H+\tau_H} v_H e^{-rt} dt - D(\tau_H) = \int_{\tau_L}^{T_L+\tau_L} v_H e^{-rt} dt - D(\tau_L). \quad (5.10)$$

Vale il seguente risultato:

Proposizione 1. *La soluzione ottimale con incompleta informazione non prevede alcuna estensione della durata del brevetto per le innovazioni di maggior valore e una certa estensione per le innovazioni di valore più basso. In altri termini, $\check{\tau}_H = \tilde{\tau}_H$ e $\check{\tau}_L > \tilde{\tau}_L$.*

Dimostrazione. Questo risultato segue una logica ben nota nella letteratura di *screening* ottimale e quindi la dimostrazione sarà soltanto abbozzata. Un esame della funzione obiettivo (5.8) rivela che la scelta ottimale di T_H e τ_H deve ancora minimizzare la somma $T_H + \tau_H$ sotto il vincolo che il profitto π_H sia uguale ad un valore pre-determinato, che in questo caso è $\int_{\tau_L}^{T_L+\tau_L} v_H e^{-rt} dt - D(\tau_L)$. Proprio come in condizioni di informazione completa, l'implicazione è che la soluzione per il tipo alto deve essere in un punto in cui la pendenza della curva di iso-profitto è uguale a -1 , e quindi anche l'inclinazione della curva $T(\tau)$ deve essere uguale a -1 . Per quanto riguarda il tipo basso, invece, il problema non è semplicemente quello di minimizzare $T_L + \tau_L$ sotto il vincolo che il profitto π_L sia uguale a \bar{R} , perchè T_L e τ_L influenzano anche l'altro vincolo del problema sociale (cioè il vincolo di compatibilità con gli incentivi del tipo alto, (5.10)). In particolare, un aumento di τ_L rende più stringente il vincolo di compatibilità con gli incentivi del tipo alto e quindi è socialmente costoso. Ne segue che il valore ottimale di τ_L sarà più basso che con completa informazione. Dunque la pendenza della curva di isoprofitto sarà maggiore di -1 nel punto ottimale per il tipo basso, e conseguentemente anche la pendenza della curva $T(\tau)$ sarà maggiore di -1 in quel punto.

■

La logica di questo risultato è illustrata nella Figura 5.3. Ragionando per assurdo, supponiamo che la soluzione ottimale con completa informazione per il tipo basso sia ancora ottimale con informazione incompleta. Questa soluzione prevede che innovazioni di valore v_L generino un profitto esattamente pari a \bar{R} , il che fissa la curva di isoprofitto per il tipo basso (quella rappresentata nella Figura 5.3), e che la pendenza della curva di isoprofitto sia pari a -1 . Queste condizioni identificano il punto L, che corrisponde alla combinazione (T_L, τ_L) che sarebbe ottimale con informazione completa. (Le rette tratteggiate nella Figura 5.3 hanno una pendenza uguale a -1 .) Dato che innovazioni di valore v_H possono comunque accedere alla combinazione (T_L, τ_L) , esse genereranno un profitto superiore a \bar{R} . La sovra-remunerazione per le innovazioni di valore elevato è costosa dal punto di vista sociale, perchè con libertà di entrata tutto l'extra-profitto viene investito in ricerca, anche se l'investimento in eccesso di \bar{R} è completamente inutile dal punto di vista sociale. Data la curva di isoprofitto che è assicurata al tipo alto, la scelta ottimale sarebbe di nuovo quella che minimizza la somma $T_H + \tau_H$, cioè il punto H nella Figura 5.3.

È ora facile vedere come, in realtà, il punto L non sia ottimale con informazione incompleta. La soluzione deve comunque collocarsi sulla curva di indifferenza del tipo basso rappresentata nella figura, per via del vincolo di partecipazione. Ma se ora spostiamo leggermente il punto L verso destra, avremo una perdita di benessere sociale sulle innovazioni di valore basso che è del secondo ordine, dato che ci stiamo leggermente allontanando dal punto che sarebbe ottimale con completa informazione. Questo però ci consente di ridurre la sovra-remunerazione per innovazioni di valore elevato, con un effetto positivo sul benessere sociale che è del primo ordine, e che quindi più che compensa l'effetto negativo. La politica ottimale deve quindi prevedere un valore di τ_L più alto che con completa informazione – e questo come sappiamo richiede che vi sia una qualche estensione della durata del brevetto.

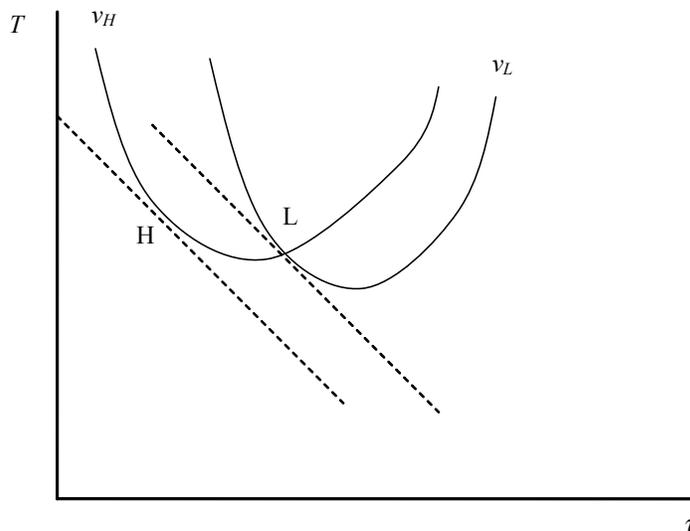


Figura 5.3: Illustrazione grafica della Proposizione 1

5.4 IL CASO GENERALE

Quando la funzione $\theta(R)$ è “liscia,” l’analisi diventa più complicata perchè, come abbiamo accennato sopra, il problema sociale diventa in effetti un problema di *optimal screening with random participation constraints*.

Cominciamo dimostrando che la soluzione deve sicuramente soddisfare la proprietà di *no-distortion-at-the-top*.

Proposizione 2. *La soluzione ottimale con incompleta informazione soddisfa la proprietà di no-distortion-at-the-top. In altri termini, $\check{\tau}_H = \tilde{\tau}_H$, o, equivalentemente, $\check{T}'(\tilde{\tau}_H) = -1$.*

Dimostrazione. Il modo più semplice per dimostrare questo risultato è attraverso un’analisi grafica (si veda la Figura 5.4). Ipotizziamo per assurdo che la pendenza della curva $T(\tau)$ nel punto H scelto dal tipo alto sia maggiore di -1 , come nel caso della retta continua nella Figura 5.4. (La curva $T(\tau)$ è supposta lineare solo per semplicità grafica; l’argomentazione che segue vale in generale.) Si consideri un segmento (il segmento tratteggiato nella Figura

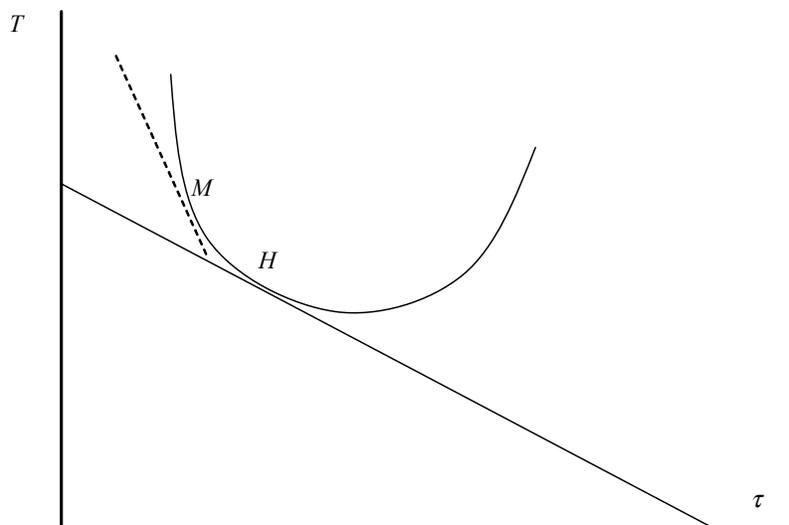


Figura 5.4: La proprietà di *no distortion at the top*

5.4) con pendenza pari a -1 e tangente nel punto M alla stessa curva di isoprofitto del tipo alto che passa per il punto H . Chiaramente, il valore di $T + \tau$ lungo il segmento tratteggiato, e quindi in particolare nel punto M , è più basso che nel punto H . Ora si consideri una curva $T(\tau)$ formata dall'involucro interno della curva originaria e il segmento può costruito, e si supponga questa nuova curva sostituisca la curva $T(\tau)$ originaria. Chiaramente, il tipo basso non cambia le sue scelte. Questo implica che il secondo termine della funzione del benessere sociale non cambia. Il tipo alto, invece, si sposta da H a M . Poiché il profitto del tipo alto non cambia, anche il suo grado di partecipazione θ_H rimane invariato. Ma siccome $T_H + \tau_H$ si è ridotto, il primo termine della funzione del benessere sociale è necessariamente aumentato. Questo significa che una curva $T(\tau)$ con una pendenza maggiore di -1 nel punto di tangenza con la curva di isoprofitto del tipo alto non può essere ottimale. In tale punto, la pendenza deve necessariamente essere pari a -1 . ■

Intuitivamente, la proprietà di *no-distortion-at-the-top* discende dal fatto che i tipi bassi non hanno incentivo a imitare le scelte ottimali per i tipi alti. In altri termini, il vincolo di compatibilità con gli incentivi per il tipo basso – la condizione (5.2) – non è mai stringente nella soluzione ottimale. Questo risultato ha un suo valore in sé, per cui vale la pena di enunciarlo e dimostrarlo formalmente:

Lemma 3. *Nella soluzione ottimale con informazione incompleta, il vincolo di compatibilità con gli incentivi per il tipo basso non è mai stringente.*

Dimostrazione. Di nuovo ricorriamo ad un'analisi grafica. Nella Figura 5.5, il punto H rappresenta la combinazione $(\check{T}_H, \check{\tau}_H)$ ottimale per il tipo alto e il punto L la combinazione $(\check{T}_L, \check{\tau}_L)$ ottimale per il tipo basso. Se, ragionando per assurdo, il vincolo di compatibilità con gli incentivi per il tipo basso fosse stringente, i punti H e L dovrebbero giacere sulla stessa curva di isoprofitto per il tipo basso, come è rappresentato nella Figura 4.5. perché però H sia la scelta ottimale per il tipo alto e L quella ottimale per il tipo basso, la curva $T(\tau)$ dovrebbe avere un angolo nel punto H . E poichè per la Proposizione 2 la pendenza della curva di isoprofitto del tipo alto nel punto H deve essere uguale a -1 , a destra di tale punto la curva dovrebbe avere pendenza minore di -1 . Questo è però impossibile, dimostrando che il vincolo di compatibilità con gli incentivi per il tipo basso non può essere stringente. ■

Nel nostro caso, il principio di *no-distortion-at-the-top* implica che per le innovazioni di valore più elevato, e quindi con tempi di sviluppo più brevi, non sia prevista alcuna estensione della durata nominale del brevetto. Come abbiamo visto nel capitolo introduttivo, di fatto sia negli Stati Uniti che in Europa non vi è alcuna estensione della durata nominale del brevetto quando il tempo di sviluppo è inferiore a 5-6 anni. Da questo punto di vista, la politica adottata attualmente sembra, sul piano qualitativo, corretta.

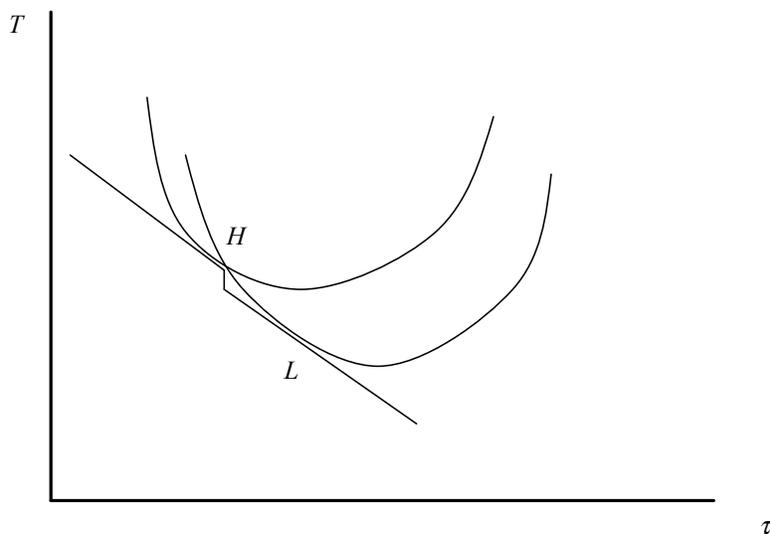


Figura 5.5: Il vincolo di compatibilità con gli incentivi per il tipo basso non è stringente

5.4.1 CARATTERIZZAZIONE COMPLETA DELLA SOLUZIONE

La caratterizzazione completa della soluzione richiede di distinguere tra diversi casi, in funzione della relazione tra la durata nominale ottimale con completa informazione per le innovazioni di valore alto e di valore basso, $\tilde{\Gamma}_L$ e $\tilde{\Gamma}_H$. Questi casi sono:

1. $\tilde{\Gamma}_L < \tilde{\Gamma}_H$;
2. $\tilde{\Gamma}_L = \tilde{\Gamma}_H$;
3. $\tilde{\Gamma}_L$ poco più grande di $\tilde{\Gamma}_H$ (il significato preciso dell'espressione "poco più grande" sarà chiarito tra un momento);
4. $\tilde{\Gamma}_L$ molto più grande di $\tilde{\Gamma}_H$.

Come sappiamo dalla Proposizione 3 del precedente capitolo, i primi due casi si possono presentare solo se l'elasticità dell'offerta di innovazioni cresce abbastanza velocemente al crescere dell'investimento in ricerca. Anche se in linea di principio non possiamo escluderli, essi sembrano quindi meno probabili degli ultimi due.

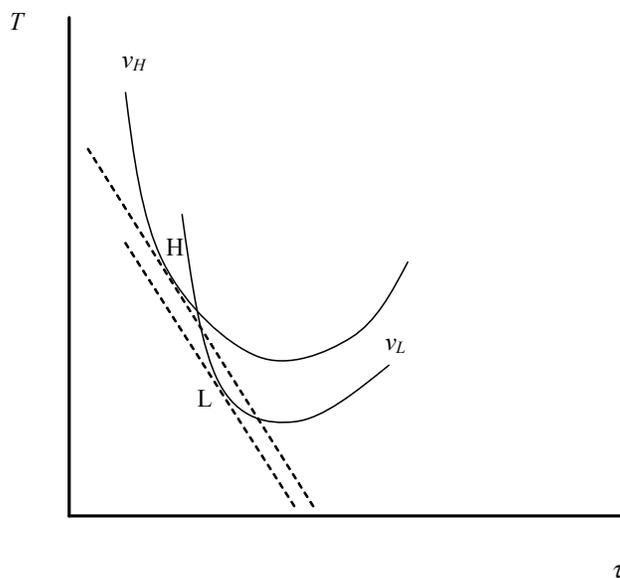


Figura 5.6: Il caso $\tilde{\Gamma}_L < \tilde{\Gamma}_H$

IL CASO $\tilde{\Gamma}_L < \tilde{\Gamma}_H$

In questo caso, la soluzione ottimale con completa informazione è quella rappresentata nella Figura 5.6. Chiaramente, questa soluzione non può essere implementata con informazione incompleta, o perché (se $\tilde{\Gamma}_L$ è molto più piccolo di $\tilde{\Gamma}_H$) non rispetta i vincoli di compatibilità con gli incentivi, o perché (se anche tali vincoli fossero rispettati perché $\tilde{\Gamma}_L$ è di poco più piccolo di $\tilde{\Gamma}_H$ – come nel caso rappresentato nella Figura 5.6) l’implementazione richiederebbe una curva $T(\tau)$ che in qualche tratto deve avere una pendenza minore di -1 . In questo caso, la soluzione ottimale con incompleta informazione sarà quindi $T(\tau) = \check{\Gamma} - \tau$, con $\tilde{\Gamma}_L < \check{\Gamma} < \tilde{\Gamma}_H$. La scelta del tempo di sviluppo τ è non distorta sia al *top* che al *bottom*, ma la durata nominale è distorta rispetto alla situazione di completa informazione.

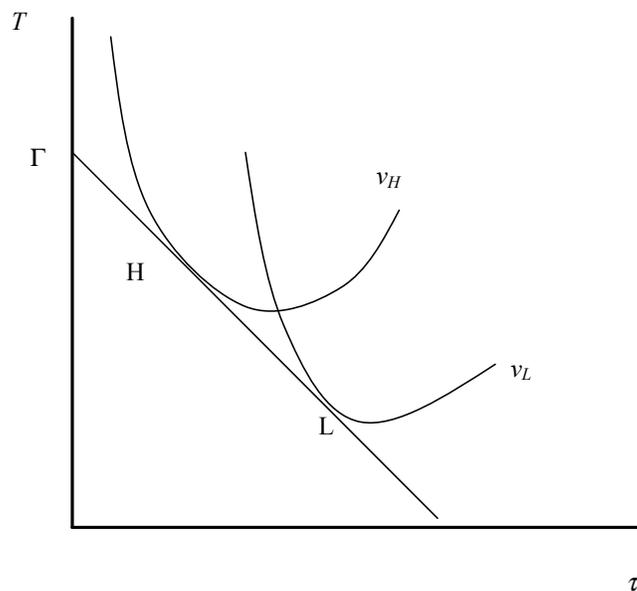


Figura 5.7: Il caso $\tilde{\Gamma}_L = \tilde{\Gamma}_H$

IL CASO $\tilde{\Gamma}_L = \tilde{\Gamma}_H$

Questo è il caso più semplice, perchè la politica ottimale con completa informazione è implementabile anche con informazione incompleta ponendo semplicemente $T(\tau) = \check{\Gamma} - \tau$ con $\check{\Gamma} = \tilde{\Gamma}_L = \tilde{\Gamma}_H$. Abbiamo quindi $(\check{T}_H, \check{\tau}_H) = (\tilde{T}_H, \tilde{\tau}_H)$ e $(\check{T}_L, \check{\tau}_L) = (\tilde{T}_L, \tilde{\tau}_L)$. La soluzione ottimale in questo caso è rappresentata nella Figura 5.7.

Veniamo ora ai due casi che sono verosimilmente più rilevanti.

IL CASO IN CUI $\tilde{\Gamma}_L$ È POCO PIÙ GRANDE DI $\tilde{\Gamma}_H$

Anche in questo caso la politica ottimale con completa informazione è implementabile con informazione incompleta. Ora tuttavia la funzione $T(\tau)$ deve comprendere due tratti di equazione $T(\tau) = \Gamma - \tau$ ma con un diverso valore di Γ : deve essere $\Gamma = \tilde{\Gamma}_H$ nel tratto iniziale della curva e $\Gamma = \tilde{\Gamma}_L$ nel tratto finale. Questi due tratti possono essere collegati da un tratto

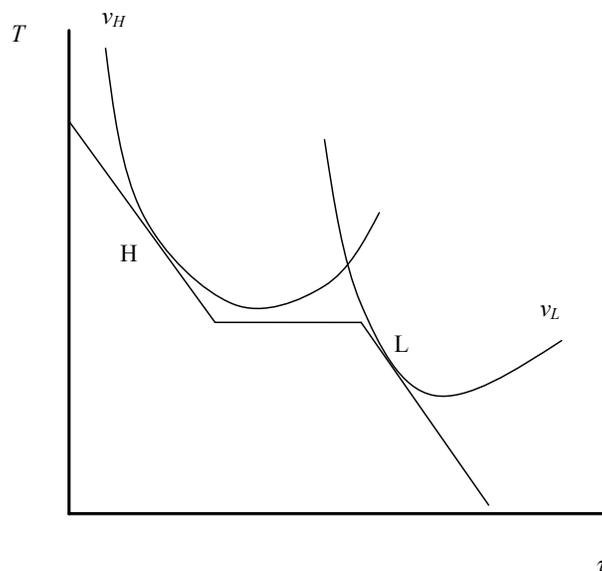


Figura 5.8: Il caso in cui $\tilde{\Gamma}_L$ è poco più grande di $\tilde{\Gamma}_H$

orizzontale (o comunque con una pendenza maggiore di -1), come illustrato nella Figura 5.8.

In questo caso la soluzione prevede sia *no-distortion-at-the-top* che *no-distortion-at-the-bottom*. La curva $T(\tau)$ ha quindi pendenza uguale a -1 sia nel tratto iniziale che nel tratto finale. Vi deve tuttavia essere un tratto intermedio più piatto, che corrisponde ad una certa estensione della durata nominale del brevetto.⁴

La condizione necessaria perché la soluzione possa avere questa forma è che la soluzione ottimale con informazione incompleta rispetti i vincoli di compatibilità con gli incentivi, e in particolare che il punto \tilde{L} giaccia al di sotto della curva di isoprofitto del tipo alto che passa per il punto \tilde{H} . Questo richiede che $\tilde{\Gamma}_L$ sia solo di poco più grande di $\tilde{\Gamma}_H$. Se questa

⁴Se il tratto intermedio fosse orizzontale, la soluzione sarebbe sorprendentemente simile alle politiche di estensione adottate in pratica, e in particolare in Europa. Va però detto che questa forma particolare è un artefatto dell'ipotesi che vi siano solo due tipi. In ogni caso, sarebbe difficile giustificare il tratto orizzontale finale delle curve rappresentate nelle Figure 1.3 e 1.4, corrispondente al meccanismo di *data exclusivity*.

condizione non è verificata, ci troviamo nel quarto e ultimo caso da considerare.

IL CASO IN CUI $\tilde{\Gamma}_L$ È MOLTO PIÙ GRANDE DI $\tilde{\Gamma}_H$

Quando $\tilde{\Gamma}_L$ è molto più grande di $\tilde{\Gamma}_H$, la soluzione ottimale con completa informazione non rispetta i vincoli di compatibilità con gli incentivi. Questo caso è rappresentato nella Figura 5.9.⁵

Il vincolo di compatibilità con gli incentivi che viene violato dalla soluzione con completa informazione è quello del tipo alto. Questo significa che un innovatore che ha sviluppato un'innovazione di valore elevato preferirebbe collocarsi nel punto \tilde{L} piuttosto che nel punto \tilde{H} . Ne segue che nella soluzione con informazione incompleta il vincolo di compatibilità con gli incentivi del tipo alto deve essere stringente. Ciò implica, come è ben noto, una distorsione delle scelte che si applicano al tipo basso.

Questa osservazione conduce al risultato principale di questa sezione (e dell'intero capitolo):

Proposizione 3. *Se l'elasticità dell'offerta di innovazioni $\frac{R\theta'(R)}{\theta}$ è costante o decrescente in R , la soluzione ottimale con incompleta informazione non prevede alcun prolungamento della durata del brevetto per valori bassi del tempo di sviluppo τ , mentre prevede un prolungamento per valori del tempo di sviluppo più elevati.*

Dimostrazione. In base alla Proposizione 3 del Capitolo 4, se l'elasticità dell'offerta di innovazioni $\frac{R\theta'(R)}{\theta}$ è costante o decrescente in R , deve essere $\tilde{\Gamma}_L > \tilde{\Gamma}_H$. Il caso in cui $\tilde{\Gamma}_L$ è di poco più grande di $\tilde{\Gamma}_H$, cosicché il vincolo di compatibilità con gli incentivi del tipo alto non è stringente, è già stato considerato nella precedente sottosezione. Rimane quindi da considerare il caso in cui $\tilde{\Gamma}_L$ è sufficientemente più grande di $\tilde{\Gamma}_H$ da far sì che il vincolo di compatibilità con gli incentivi del tipo alto sia stringente.

⁵Questo caso è forse quello più rilevante, anche perché al crescere del numero dei tipi diventa sempre meno probabile che la soluzione ottimale con completa informazione rispetti tutti i vincoli di compatibilità con gli incentivi.

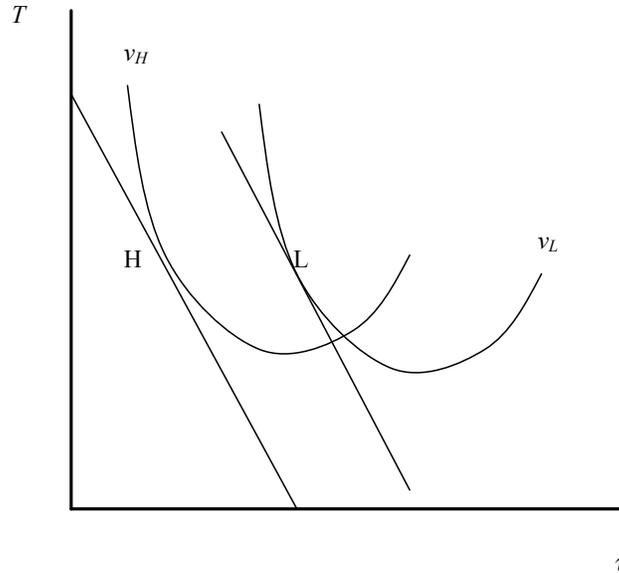


Figura 5.9: Il caso in cui $\tilde{\Gamma}_L$ è molto più grande di $\tilde{\Gamma}_H$

Se il vincolo del tipo alto è stringente, deve essere

$$\pi_H = p \int_{\tau_L}^{T_L + \tau_L} v_H e^{-rt} dt - D(\tau_L).$$

Quindi π_H , e di conseguenza anche R_H , non dipende più da T_H e τ_H , ma piuttosto da T_L e τ_L .⁶

⁶Naturalmente T_H e τ_H devono rispettare il vincolo che

$$\pi_H = p \int_{\tau_H}^{T_H + \tau_H} v_H e^{-rt} dt - D(\tau_H)$$

sia costante e pari a $p \int_{\tau_L}^{T_L + \tau_L} v_H e^{-rt} dt - D(\tau_L)$. E' facile verificare analiticamente che questo implica la proprietà di *no-distortion-at-the-top*. Infatti, essendo θ_H indipendente da T_H e τ_H , la scelta di T_H e τ_H deve semplicemente massimizzare

$$\int_{T_H + \tau_H}^{\infty} (1 + \delta) v_H e^{-rt} dt$$

Consideriamo quindi la scelta di T_L e τ_L . Ricordiamo che il benessere sociale è

$$W = q\theta(R_H)p \int_{T_H+\tau_H}^{\infty} (1+\delta)v_H e^{-rt} dt + (1-q)\theta(R_L)p \int_{T_L+\tau_L}^{\infty} (1+\delta)v_L e^{-rt} dt, \quad (5.11)$$

dove gli investimenti in ricerca R_H e R_L sono dati dalle condizioni di libertà di entrata

$$\frac{\theta(R_H)}{R_H} = \pi_H$$

e

$$\frac{\theta(R_L)}{R_L} = \pi_L.$$

Procedendo in modo simile al caso di completa informazione, scomponiamo il problema di massimizzazione sociale in due stadi: nel primo stadio si determina il valore ottimale di $T_L + \tau_L$, che indichiamo con Υ_L ; nel secondo la scelta ottimale di τ_L a parità di $T_L + \tau_L$, cioè per $T_L + \tau_L = \Upsilon_L$.

Consideriamo il secondo stadio. Come sappiamo (si veda la dimostrazione della Proposizione 2) se R_H fosse indipendente da τ_L , il problema si ridurrebbe alla massimizzazione di R_L , cioè di π_L , e quindi coinciderebbe con il problema di completa informazione. Ma siccome ora anche R_H dipende da τ_L , la scelta ottimale di τ_L deve massimizzare non più π_L , ma una media ponderata di π_L e π_H , cioè

$$\psi\pi_L + (1-\psi)\pi_H. \quad (5.12)$$

Volendo, il modo in cui il peso ψ dipende dalle variabili e dalle funzioni esogene del modello può essere reso esplicito, ma ai fini della dimostrazione l'unica cosa che conta è che $0 < \psi < 1$.

La condizione del primo ordine per la massimizzazione della (5.11) a parità di $T_L + \tau_L$ è:

$$\psi [pv_L e^{-r\tau_L} + D'(\tau_L)] + (1-\psi) [pv_H e^{-r\tau_L} + D'(\tau_L)] = 0. \quad (5.13)$$

sotto il vincolo che

$$p \int_{\tau_H}^{T_H+\tau_H} v_H e^{-rt} dt - D(\tau_H) = \text{costante}.$$

Ma questo problema è duale al problema (4.7), e quindi la soluzione deve prevedere $\tau_H = \tilde{\tau}_H$.

La soluzione di questa equazione, $\check{\tau}_L$, è necessariamente minore di $\tilde{\tau}_L$. Infatti, $\tilde{\tau}_L$ è il valore di τ_L che annulla il primo termine entro parentesi quadre, mentre il secondo termine è annullato da $\tilde{\tau}_H < \tilde{\tau}_L$. Quindi, quando $\tau_L = \tilde{\tau}_L$ il secondo termine entro parentesi quadre è negativo (cioè il costo marginale di una riduzione del tempo di sviluppo eccede il relativo beneficio marginale per l'innovatore). Ne segue che

$$pv_L e^{-r\check{\tau}_L} + D'(\check{\tau}_L) > 0.$$

Sappiamo che in una economia di mercato τ_L deve soddisfare la condizione (2.3), che si può riscrivere come

$$pv_L e^{-r\tau} + D'(\tau_L) - pv [1 + T'(\tau)] e^{-r[T(\tau)+\tau]} = 0$$

Siccome quando $\tau_L = \check{\tau}_L$ la somma dei primi due termini è positiva, ne segue che

$$pv [1 + T'(\tau)] e^{-r[T(\tau)+\tau]} > 0,$$

che evidentemente implica

$$T'(\tau) > -1.$$

Questo dimostra che è ottimale prevedere un'estensione della durata del brevetto quando il valore dell'innovazione è basso e quindi il tempo di sviluppo è relativamente alto. ■

Abbiamo così determinato le distorsioni nella scelta del tempo di sviluppo implicate dall'asimmetria informativa. Naturalmente anche la durata effettiva del brevetto risulterà conseguentemente distorta, anche se in generale è più difficile determinare il segno di questa distorsione.

5.5 CONCLUSIONI

In questo capitolo abbiamo studiato la scelta ottimale delle regole che determinano la relazione tra la durata effettiva del brevetto e il tempo di sviluppo. Il nostro risultato principale

è che non sia ottimale estendere la durata del brevetto quando il tempo di sviluppo è relativamente breve, mentre una certa estensione è ottimale quando il tempo di sviluppo è più lungo.

Da questo punto di vista, la politica adottata attualmente sembrerebbe, almeno sul piano qualitativo, corretta. Come abbiamo visto nel capitolo introduttivo, infatti, sia negli Stati Uniti che in Europa non vi è alcuna estensione della durata nominale del brevetto quando il tempo di sviluppo è inferiore a 5-6 anni, mentre è prevista un'estensione per tempi di sviluppo maggiori.

Quella sviluppata in questo capitolo e nel precedente è la prima analisi economica normativa della relazione tra la durata effettiva del brevetto e il tempo di sviluppo nel settore farmaceutico. Essa ha messo in luce effetti economici e proprietà qualitative della politica ottimale che gettano luce su un importante dibattito di politica economica. L'analisi potrebbe essere estesa in diverse direzioni, che saranno discusse nel capitolo conclusivo.

CONCLUSIONI

In questo lavoro abbiamo studiato le politiche che prevedono la possibilità di estendere la durata dei brevetti farmaceutici per recuperare almeno una parte del tempo perso per ottenere l'autorizzazione all'immissione in commercio. Queste politiche di fatto definiscono la durata effettiva del brevetto T non come uno scalare, ma come una funzione del tempo di sviluppo del farmaco, $T(\tau)$. Abbiamo quindi analizzato come la forma di questa funzione influenzi le decisioni di investimento in ricerca e sviluppo delle compagnie farmaceutiche, e come la funzione stessa dovrebbe essere disegnata da un policy-maker benevolente che volesse massimizzare il benessere sociale.

Il risultato principale della nostra analisi è così riassumibile: quando il tempo di sviluppo è relativamente breve, non è ottimale estendere la durata del brevetto, mentre una certa estensione è ottimale quando il tempo di sviluppo è più lungo. Qualitativamente, questo è precisamente il modello attualmente adottato sia negli Stati Uniti che in Europa: infatti, quando il tempo di sviluppo è inferiore a 5-6 anni non vi è alcuna estensione della durata nominale del brevetto, ma è prevista un'estensione per tempi di sviluppo maggiori.

Questo risultato segue dalla formulazione del problema sociale come un problema di *screening* ottimale. La funzione $T(\tau)$ rappresenta di fatto un menù di possibili opzioni per l'innovatore, che può scegliere una durata effettiva del brevetto maggiore, velocizzando il processo di approvazione del farmaco, o minore, con uno sviluppo del farmaco più lento ma meno costoso. Naturalmente, quando l'innovazione è molto profittevole e quindi l'incentivo ad anticipare i profitti è elevato, l'innovatore sceglierà investimenti maggiori nella fase dello

sviluppo, e quindi tempi di sviluppo più rapidi, garantendosi una durata effettiva del monopolio più lunga. Quando invece l'innovazione è meno profittevole, sarà ottimale investire meno nella fase di sviluppo, optando di fatto per una durata effettiva del brevetto più breve.

In questa prospettiva, il risultato che non sia ottimale prevedere alcuna estensione quando il tempo di sviluppo è breve discende dal fatto che non è ottimale prevedere alcuna estensione in condizioni di completa informazione, e dal principio di *no distortion at the top*. Come è ben noto, questo principio è molto robusto e generalmente deve valere in ogni problema di *screening* ottimale in cui i vincoli di compatibilità con gli incentivi operino nella direzione standard. Il fatto che la politica ottimale con completa informazione non preveda alcuna estensione del brevetto discende invece dalla necessità di non diluire gli incentivi ad investire nella fase dello sviluppo, quando non vi è più competizione tecnologica e quindi viene meno l'effetto incentivante del meccanismo del *winner-take-all*. In realtà, per il caso di innovazioni drastiche sarebbe addirittura desiderabile punire gli innovatori che non sviluppano il farmaco rapidamente: ma se questo non è possibile, la pendenza della curva $T(\tau)$ dovrà essere uguale a -1 non solo in un punto, ma per tutto un intervallo iniziale di valori del tempo di sviluppo.

La nostra analisi si presta a diverse estensioni, che in questa sede ci sono precluse dalla mancanza di tempo e di dati. Per esempio, la strategia empirica più efficace per identificare l'effetto di durata effettiva richiede dati sul periodo precedente l'*Hatch-Waxman Act*. Questi dati sono difficili da ricostruire o proibitivamente costosi da acquistare.

Probabilmente l'estensione più interessante dell'analisi normativa riguarda la possibilità di calibrare empiricamente la funzione $T(\tau)$ ottimale. Per arrivare a una calibrazione empirica della funzione $T(\tau)$ sono necessari diversi passaggi. Per cominciare, la nostra analisi si è limitata al caso di due tipi di innovazioni. Come abbiamo osservato, questo permette di identificare solo due punti della funzione $T(\tau)$ e la forma della funzione in un intorno di ciascuno dei due punti (dato che l'analisi determina la pendenza della curva). Per identificare pienamente la funzione $T(\tau)$ occorre estendere l'analisi al caso di un continuo di tipi.

Una volta fatto questo, il passo successivo sarebbe quello di stimare empiricamente gli ingredienti del calcolo della funzione ottimale. Questi sono:

1. *La funzione $\theta(R)$* . Come primo approccio, sembra ragionevole scegliere una specificazione iso-elastica della funzione e stimare l'elasticità dell'offerta di innovazioni. Una vasta letteratura empirica ha cercato di stimare tale elasticità (per una rassegna, si veda Scotchmer, 2004 e Denicolò, 2007). Ci sono stime anche specifiche per il settore farmaceutico (Acemoglu e Linn, 2004). Un approccio alternativo è invece quello di fissare questa elasticità al livello che risulta necessario perchè la durata nominale attuale del brevetto (20 anni) sia invero quella ottimale.
2. *La distribuzione di probabilità del valore delle innovazioni*. Esistono stime empiriche della distribuzione di probabilità del valore dei brevetti, ma per quanto a nostra conoscenza non vi sono stime specifiche per il settore farmaceutico.
3. *La funzione $\tau(D)$* . Se si conosce la distribuzione di probabilità del valore delle innovazioni, la funzione $\tau(D)$ potrebbe essere calibrata in modo da generare una distribuzione di probabilità del tempo di sviluppo simile a quella osservata (e stimata nel capitolo 3).

In questo modo dovrebbe essere possibile ottenere una caratterizzazione non solo qualitativa, e non solo locale, della funzione $T(\tau)$.

Un'altra estensione riguarda il caso di eterogeneità multi-dimensionale. Come abbiamo osservato nel capitolo sesto, le innovazioni possono differire da molti punti di vista, e quindi è evidente che in generale l'eterogeneità è multi-dimensionale. Anche se il problema del *multi-dimensional screening* è ancora largamente irrisolto sul piano teorico, nel nostro caso qualche progresso è forse possibile in virtù del fatto che lo spazio delle allocazioni è comunque unidimensionale (cioè T dipende solo da τ , per quando grande sia il numero di variabili che possono differire da un'innovazione all'altra).

BIBLIOGRAFIA

Abbot, A. (2005). Animal testing: More than a cosmetic change. *Nature*, vol. 438, 144-146.

Acemoglu, D. e J. Lin (2004). Market size in innovation: Theory and evidence from the pharmaceutical industry. *Quarterly Journal of Economics*, vol. 119, 1049-1090.

Arora, A., Ceccagnoli M. e W. Cohen (2008). R&D and the patent premium. *International Journal of Industrial Organization*. vol.26, 1153-1179.

Bagley, M. (2011). Patent term restoration and non-patent exclusivity in the United States. Virginia Public Law and Legal Theory Research Paper No. 2011-25.

Bessen, J. e M. J. Meurer (2008). *Patent failure: How judges, bureaucrats, and lawyers put innovators at risk*. Princeton: Princeton University Press.

Bhat, V. (2005). Patent term extension strategies in the pharmaceutical industry. *Pharmaceuticals Policy and Law*, vol. 6, 109-122.

Chari, V.V., Golosov M. e A. Tsyvinski (2012). Prizes and patents: using market signals to provide incentives for innovations. *Journal of Economic Theory*, vol. 147(2), 781-801.

Clift, C. (2008). The value of patent term extension to the pharmaceutical industry in the US. *Journal of Generic Medicines*, vol. 5, 201-208.

Cohen, W., Nelson R. e J. Walsh (2000). Protecting their intellectual assets: appropriability conditions and why U.S. manufacturing firms patent (or not). NBER Working Paper No. 7552.

Cornelli, F. e M. Schankerman (1999). Patents renewals and R&D incentives. *Rand Journal of Economics*. vol. 30(2), 197-213.

- Dasgupta, P. e J. Stiglitz (1980). Industrial structure and the nature of innovative activity. *Economic Journal*, vol. 90, 266-293.
- Denicolò, V. (1996). Patent races and optimal patent breadth and length. *Journal of Industrial Economics*, vol. 44, 249-265.
- Denicolò, V. (2007). Do patents over-compensate innovators? *Economic Policy*, vol. 52, 679-729.
- Denicolò, V. e L.A. Franzoni (2004). Patent, secrets and the first invention defense. *Journal of Economics and Management Strategy*, vol. 13, 537-538.
- Denicolò, V. e L.A. Franzoni (2010). On the winner-take-all principle in innovation races. *Journal of the European Economic Association*, vol. 8, 1133-1158.
- Di Masi, J., R. Hansen e H. Grabowski (2003). The price of innovation: new estimates of drug development costs. *Journal of Health Economics*, vol. 22, 151-185.
- Dranove D. e D. Meltzer (1994). Do important drugs reach the market sooner?. *Rand Journal of Economics*, vol. 25, 402-423.
- Frantz, S. e A. Smith (2003). New drug approvals for 2002. *Nature Reviews Drug Discovery*, vol. 2, 95-96.
- Gieringer, D. (1985). The safety and efficacy of new drug approval. *Cato Journal*, vol. 5, 177-201.
- Gilbert, R. e C. Shapiro (1990). Optimal patent breadth and length. *Rand Journal of Economics*, vol. 21(1), 106-112.
- Goldman, D. e D. Lakdawalla (2012). Intellectual property, information technology, biomedical research, and marketing of patented products. In *Handbook of Health Economics*, a cura di M. V. Pauly, T. G McGuire e P. Pita Barros, vol. 2, 825-872.
- Grabowski H. e M. Kyle (2007). Generic competition and market exclusivity periods in pharmaceuticals. *Managerial and Decision Economics*, vol. 27, 1-12.

Grabowski H. e J. Vernon (1992). Brand loyalty, entry and price competition in pharmaceuticals after the 1984 Drug Act. *Journal of Law and Economics*, vol. 35, 331-350.

Grabowski H., J. Vernon (1987). Pioneers, imitators, and generics. A simulation model of Schumpeterian competition. *Quarterly Journal of Economics*, vol. 102(3), 491-525.

Grabowski H., J. Vernon (1986). Longer patents for lower imitation barriers: the 1984 Drug Act. *American Economic Review*, vol. 76(2), 195-198.

Grabowski H., J. Vernon e L. G. Thomas (1978). Estimating the effects of regulation on innovation: an international comparative analysis of the pharmaceutical industry. *Journal of Law and Economics*, vol. 21, 133-163.

Grabowski H. e Y. R. Wang (2006). The quantity and quality of worldwide new drug introductions, 1982-2003. *Health Affairs*, vol. 25, 452-459.

Green, J. e S. Scotchmer (1995). On the division of profit with sequential innovation. *Rand Journal of Economics*, vol. 26, 20-33.

Hopenhayn, H., Mitchell M. e G. Llobet (2006). Rewarding sequential innovations: prizes, patents and buyouts. *Journal of Political Economy*, vol. 104(6), 1041-1068.

Hughes, B. (2008). 2007 FDA drug approvals. *Nature Review Drug Discovery*, vol. 7, 107-109.

Hughes, B. (2009). 2008 FDA drug approvals. *Nature Review Drug Discovery*, vol. 8, 93-96.

Jensen, E. (1987). Research expenditure and the discovery of new drugs. *Journal of Industrial Economics*, vol. 36(1), 83-95.

Kaitin, K. e E. Ealy (2000). The new drug approvals of 1996, 1997, and 1998: drug development trends in the user fee era. *Drug Information Journal*, vol. 34, 1-14.

Kaitin, K. e M. Manocchia (1997). The new drugs approvals of 1993, 1994, and 1995: trends in drug development. *American Journal of Therapeutics*, vol. 4, 46-54.

Keyhani, S., M. Diener -West e N. Powe (2006). Are development times for pharmaceuticals increasing or decreasing? *Health Affairs*, vol. 25, 461-468.

Kitch, E.W. (1977). The nature and function of the patent system. *Journal of Law and Economics*, vol. 20, 265-290.

Lee, T. e L.L. Wilde (1980). Market structure and innovation: a reformulation. *Quarterly Journal of Economics*, vol. 94, 429-436.

Loury, G.C. (1979). Market structure and innovation. *Quarterly Journal of Economics*, vol. 93, 395-410.

Mc Guire, T., Nelson R. e T. Spavins (1975). An evaluation of consumer protection legislation: the 1962 drug amendments. *Journal of Political Economy*, vol. 83(3), 655-661.

Mossinghoff, G. (1999). Overview of the Hatch-Waxman Act and its impact on the drug development process. *Food and Drug Law Journal*, vol. 54, 187-194.

National Institute for Health Care Management Foundation (2002). Changing patterns of pharmaceutical innovation.

Nordhaus, W. (1969). *Invention growth and welfare*. Cambridge, Mass.: MIT Press.

Ohly, D.C. (2010). The Hatch-Waxman Act: Prescriptions for innovative and inexpensive medicines. Mimeo.

Pammolli, F., Magazzini, L. e M. Riccaboni (2011). The productivity crisis in pharmaceuticals R&D. *Nature Reviews*, vol. 10, 428-438.

Peltzman, S. (1973). An evaluation of consumer protection legislation: the 1962 drug amendments. *Journal of Political Economy*, vol. 81(5), 1049-1091.

Reiffen D. e M. Ward (2005). Generic drug industry dynamics. *Review of Economics and Statistics*, vol. 87, 37-49.

Roin B. (2009). Drug patent length. Mimeo, Harvard Law School.

Scherer, F.M. (2010). *Pharmaceutical innovation*. In *Handbook in Economics*, vol. 1, 539-574.

Scotchmer, S. (1999). On the optimality of patent renewal system. *Rand Journal of Economics*, vol. 30, 181-196.

Scotchmer, S. (2004). *Innovation and Incentives*. Cambridge, Mass., MIT Press.

Shapiro, C. (2006). Prior user rights. *American Economic Review*, vol. 96(2), 92-96.

Shavell, S. e T. Van Ypersele (2001). Rewards versus intellectual property rights. *The Journal of Law and Economics*, vol. 44(2), 525-547.

Temin, P. (1980). Taking your medicine-drug regulation in the United States. Harvard University Press.

Thomas, G. (1990). Regulation and firm size. *Rand Journal of Economics*, vol. 21(4), 497-517.

Wardell, W M. (1973). Introduction of new therapeutic drugs in the United States and Great Britain: an international comparison. *Clin. Pharmacol Ther.*, vol. 14, 773-790.

Wardell, W M. (1978). The drug lag revisited: comparison by therapeutic area of patterns of drugs marketed in the United States and Great Britain from 1972 through 1976. *Clin. Pharmacol Ther.*, vol. 24, 499-524.

Wiggins, S. e R. Maness (1983). Price competition in pharmaceuticals: the case of anti-infectives. *Economic Enquiry*, vol. 42, 247-263.

Wiggins, S. (1981). Product quality regulation and new drug introductions: some new evidence from the 1970s. *Review of Economics and Statistics*, vol. 63(4), 615-619.

Wiggins, S. (1983). The impact of regulation on pharmaceutical research expenditures: a dynamic approach. *Economic Enquiry*, vol. 21, 115-128.

Wright, B.D. (1983). The economics of invention incentives: patents, prizes, and research contracts. *American Economic Review*, vol. 73, 691-707.