

IL FINANZIAMENTO DELLA SANITÀ IN ITALIA: FINANZA AUTONOMA O DERIVATA?

AGNESE SACCHI

JEL Classification: H51, H75, H77

Keywords: Sistema sanitario nazionale – Federalismo fiscale

*Il finanziamento della sanità in Italia: finanza autonoma o derivata?**

Agnese Sacchi

Università degli Studi di Roma Tre, *E-mail*: asacchi@uniroma3.it

Abstract

Recent reforms of the National Health Service (NHS) financing system in Italy (D.Lgs. 446/97 and D.Lgs. 56/2000) have potentially given more responsibility for health care to Regions, increasing their power over this function. Starting from this normative process, the aim of the paper is to analyse the Italian health spending in the Ordinary Statute Regions (OSR), by proposing a model to verify empirically whether the financing mechanisms are actually defined at local level or if Regions act as agents of the central government.

The empirical results do not support the evolution of fiscal federalism in health sector, highlighting in fact that health financing decisions are mostly centralized in Italy, even if the institutional context has recently moved toward decentralization of health financing sources, extending the autonomy of sub-national governments.

JEL Classification: H51, H75, H77

1. Introduzione

L'avvio del federalismo fiscale e il riordino del Sistema Sanitario Nazionale (SSN) costituiscono due innovazioni potenzialmente in grado di modificare in modo sostanziale l'assetto della finanza regionale. La sanità, infatti, può essere considerata a pieno titolo il settore che ha operato da banco di prova del federalismo fiscale, in quanto essa ha attuato cambiamenti che vanno nella direzione del trasferimento delle responsabilità a livelli di governo sub-centrali e della progressiva espansione del ruolo delle regioni.

Partendo da queste considerazioni, l'obiettivo del lavoro consiste nell'analizzare il caso della spesa sanitaria italiana – nelle Regioni a Statuto Ordinario (RSO)¹ – proponendo un semplice modello teorico e un metodo empirico per valutare quanto i meccanismi di erogazione e di prelievo relativi alla funzione sanità, formalmente in capo alle regioni, siano effettivamente stabiliti a livelli locale.

(*) Ringrazio Paolo Liberati per i preziosi commenti e suggerimenti al lavoro. Inoltre, ringrazio Fabio Fiorillo e tutti i partecipanti al 13° convegno dell'Associazione Italiana di Economia Sanitaria (AIES), Matera, ottobre 2008, per le utili osservazioni. Mi assumo la piena responsabilità delle proposizioni e delle opinioni espresse nell'articolo e degli errori rimanenti.

¹ L'analisi empirica è stata condotta prendendo in esame solo le Regioni a Statuto Ordinario dato il diverso regime di finanziamento della spesa sanitaria nelle Regioni a Statuto Speciale (RSS). Infatti, l'attribuzione delle competenze e il sistema di finanziamento in tali Regioni sono determinati da statuti regionali – a differenza delle RSO, in cui valgono le leggi statali nel rispetto dei vincoli costituzionali – anche sensibilmente diversi tra regioni. In dettaglio, il modello di finanziamento delle RSS si basa essenzialmente sul sistema delle compartecipazioni al gettito dei tributi erariali, cioè sul versamento a tali regioni del gettito di molti tributi statali incassati entro i confini regionali.

La necessità di determinare empiricamente quanto questa funzione sia accentrata o decentrata è sembrato un tema rilevante, specie osservando le recenti riforme istituzionali che hanno riguardato l'Italia. Infatti, il nuovo meccanismo di finanziamento delle RSO (D.Lgs. 446/97 e D.Lgs. 56/2000) costituisce il primo vero sistema di finanziamento di cui si è dotata la sanità². La nuova normativa ha previsto che il finanziamento del SSN sia quasi interamente affidato alle Regioni³ e basato sulla finanza regionale complessiva, senza più vincoli specifici di destinazione su uno o più tributi (Muraro, 2002).

Tuttavia, se dal 2001 sono venuti meno tali vincoli formali, di fatto, restano i vincoli sostanziali tali per cui il gettito dei principali tributi regionali (Irap e Addizionale Irpef) viene destinato principalmente al finanziamento della spesa sanitaria regionale (Liberati, 2001). In aggiunta, con l'introduzione della compartecipazione all'Iva contestualmente all'abolizione dei trasferimenti dallo Stato alle Regioni - di cui il 99,2% riferiti alla sanità (abolizione del FSN) – permane il meccanismo di trasferimento di risorse dal centro alle unità locali⁴. Questo aspetto è strettamente legato al fatto che la garanzia ai cittadini di “livelli essenziali ed uniformi di assistenza” (LEA)⁵ è ancora un obiettivo fondamentale del governo centrale, il quale è interessato ad assicurare che su tutto il territorio nazionale siano forniti livelli minimi di servizio sanitario. La questione da porsi è se sia possibile mantenere degli standard nazionali, quando la sanità rientri tra le competenze proprie delle Regioni, le quali potrebbero non avere un'adeguata capacità di risorse per garantire tali livelli minimi, sui quali continuerebbe dunque ad incidere la componente di finanziamento nazionale, presumibilmente nella forma di trasferimenti specifici o vincolati. In altri termini, la presenza dei livelli essenziali potrebbe non essere totalmente compatibile con un sistema di finanziamento decentrato della spesa sanitaria. In tal senso, si è ritenuto opportuno analizzare se il cambiamento prodotto dal decentramento fiscale sul settore sanitario – almeno formalmente - si sia effettivamente realizzato.

Ipotizzando un modello teorico che sia in grado di rappresentare le soluzioni verso cui dovrebbe tendere un sistema di *finanza autonoma* o di *finanza derivata*, caratterizzati da una diversa gestione della funzione sanità (in termini di responsabilità di provvedere a tale funzione), ci si è chiesti se le soluzioni effettivamente osservate nel caso italiano siano più compatibili con l'uno o l'altro degli schemi teorici. In particolare, se prevalga un modello in cui le decisioni di copertura del fabbisogno sanitario sono sostanzialmente indipendenti (*finanza autonoma*) o uno in cui le decisioni sono

² La legge 833/78 istituì il SSN in sostituzione del sistema mutualistico a rimborso per le singole categorie professionali. Fino a quel momento, non era previsto alcun sistema di finanziamento della sanità pubblica.

³ Ad eccezione di una quota gestita a livello centrale per finalità o programmi particolari (accordi internazionali, ricerca).

⁴ La compartecipazione all'Iva è, infatti, il tributo che compensa circa il 94% dei finanziamenti aboliti, fornendo alle regioni entrate pari a 53.127 miliardi di lire.

⁵ Il compito dei LEA è, infatti, quello di realizzare il dettato dell'art. 32 della Costituzione sul diritto di ogni cittadino, in qualunque parte del territorio egli viva, di avere un'adeguata tutela della salute.

centralizzate (si pensi ai livelli essenziali – *finanza derivata*). Da qui, l'esigenza di verificare empiricamente lo scostamento tra le soluzioni ottimali fornite dai modelli proposti e il contesto italiano. In definitiva, il punto che si vuole mettere in evidenza è che nonostante le riforme istituzionali abbiano potenzialmente aumentato l'autonomia delle Regioni - per alcune funzioni essenziali come la sanità (D.Lgs. 56/2000) - i risultati empirici potrebbero non confermare tale evoluzione normativa. Detto altrimenti, si potrebbe verificare che la finanza pubblica italiana – in riferimento alla funzione sanità - si comporti ancora in maniera compatibile con uno schema di *finanza derivata*, benché di recente si siano fatti passi avanti verso un sistema di *finanza autonoma*.

2. Il modello teorico

Considerando un'economia caratterizzata dalla presenza di almeno due livelli di governo (Goodspeed, 2000)⁶, si intende costruire un modello che sia in grado di rappresentare le soluzioni ottimali per i livelli di spesa e tassazione, relativamente alla funzione sanità, in due scenari alternativi:

- a) Stato e regioni finanziano contemporaneamente ed in modo autonomo la spesa sanitaria per le rispettive comunità, ognuno secondo le proprie competenze (*finanza autonoma*);
- b) un *decisore centrale* unico si occupa in via esclusiva della funzione sanitaria, prendendo decisioni per tutti i livelli di governo (*finanza derivata*).

L'ipotesi di base è che entrambi i livelli di governo (centrale e locale) utilizzino un'imposta proporzionale sul reddito⁷ per finanziare la spesa sanitaria, erogata sotto forma di beni e servizi assegnati su base pro capite, ai rispettivi residenti (Arzaghi e Henderson, 2005).

⁶ Una struttura federale di governo implica almeno due livelli di governo: uno più alto (es. governo centrale) e un set di unità inferiori (es. regioni). Molti paesi democratici hanno un governo nazionale e una gamma di governi locali sottostanti. Senza perdere generalità nei risultati, si assume l'esistenza di due enti locali (ad esempio, due regioni) e di un governo centrale. Le due regioni, assunte di uguale popolazione, si differenziano per il reddito totale prodotto – dato dalla somma dei redditi individuali – e quindi per il reddito pro capite. Per maggiori dettagli sul modello teorico, si permetta il rinvio a Sacchi (2008).

⁷ Dal punto di vista del governo locale, il finanziamento della sanità con imposta proporzionale sul reddito trova un riscontro reale nell'addizionale regionale Irpef che viene applicata nella maggior parte delle RSO con aliquota proporzionale tra 0,9% e 1,4%. Anche l'Irap prevede un'aliquota costante del 3,9% (con l'eccezione di alcuni settori agevolati). Benché questo tributo gravi formalmente sulle imprese, o meglio sui soggetti economici che esercitano un'attività produttiva, si può ragionevolmente ipotizzare una traslazione, almeno parziale, dell'imposta sui consumatori, che va quindi ad incidere sul loro reddito personale. Sulla base di questa osservazione, è sembrato sensato ipotizzare un prelievo proporzionale sul reddito individuale a livello locale. Dal lato del governo centrale, l'applicazione dell'imposta proporzionale per finanziare la sanità non trova una corrispondenza diretta con la situazione reale. In effetti, non esiste un tributo proprio del governo centrale destinato a sostenere tale spesa. Tuttavia, la principale fonte di finanziamento per la sanità di "competenza" dello Stato è rappresentata dalla compartecipazione all'Iva, che si basa su una struttura sostanzialmente proporzionale applicata alla base imponibile nazionale. Essendo l'Iva un'imposta generale sugli scambi di beni e servizi la cui incidenza si manifesta sui consumi, ed essendo i consumi correlati con il reddito degli individui, è plausibile assumere una certa equivalenza tra questa imposta e l'imposta proporzionale sul reddito. Accettando dunque questa interpretazione, nel modello teorico si propone un prelievo proporzionale sul reddito anche da parte del governo centrale.

In aggiunta, ogni *policy maker* è benevolente e, come tale, massimizza una funzione obiettivo espressione del benessere collettivo dei propri residenti, dato il vincolo di bilancio in pareggio. Si assume inoltre l'assenza di relazioni tra i vari livelli di governo e tra governi di pari livello⁸. Nel definire gli strumenti per attuare la funzione sanità, l'operatore pubblico tiene conto delle preferenze della rispettiva collettività per tale funzione pubblica. In particolare, si assume che a fronte dell'eterogeneità delle preferenze individuali, l'operatore pubblico – sia esso centrale o locale – consideri una preferenza media per la spesa sanitaria. Tale parametro è legato alla distribuzione del reddito dentro l'ente: esso è definito dal rapporto tra il reddito medio e il reddito massimo e, come tale, fornisce informazioni sul grado di disuguaglianza. L'idea è che maggiore è la distanza tra reddito medio e massimo, maggiore è la sperequazione del reddito, maggiore sarà anche il peso delle politiche sanitarie sull'utilità aggregata. In altri termini, si assume che l'importanza di tale spesa sul benessere degli individui sia tanto maggiore quanto più elevata è la disuguaglianza del reddito, data la sua natura di spesa redistributiva *in kind*⁹.

Il quadro appena descritto rappresenta il punto di partenza dell'analisi. Nel prossimo paragrafo vengono descritti gli scenari istituzionali di *finanza autonoma e derivata*.

2.1 La funzione sanità nel caso di *finanza autonoma*

Il primo scenario prevede che Stato e regioni abbiano, entrambi, la responsabilità di provvedere all'erogazione e al finanziamento della spesa sanitaria, in modo autonomo e simultaneo. Nel farlo, i governi utilizzano gli strumenti a loro disposizione, massimizzando il benessere delle rispettive collettività: ogni *policy maker* prende come base del problema una funzione che rappresenti una misura di sintesi dell'utilità dei suoi membri. In dettaglio, si assume che ogni governo locale

⁸ Il fatto che non siano previsti legami verticali tra Stato e regioni esclude la possibilità che il governo centrale intervenga nel bilancio delle autorità locali per ripianare eventuali disavanzi, evitando anche una serie di problemi legati ai *soft budget constraints*, ai *bailouts* e ai *commons pool* (Pisauro, 2001; 2003). Tali fenomeni, infatti, si potrebbero verificare con il decentramento delle responsabilità fiscali se, per gli enti locali, permanga la possibilità di essere finanziati e “coperti” in ultima istanza dalle risorse del governo centrale. Va detto che la presenza di *policy makers* benevolenti, interessati dunque alla massimizzazione del benessere sociale piuttosto che a quella di una funzione obiettivo propria, garantisce che anche nel caso in cui i governi locali ricevano trasferimenti dall'autorità centrale, questi li spendano in modo efficiente. In definitiva, l'assenza di relazioni tra livelli di governo si traduce nell'assenza di trasferimenti verticali dallo Stato alle unità locali, assumendo una perfetta corrispondenza tra la dimensione delle entrate e quella delle spese a tutti i livelli. In realtà, anche la direzione opposta del legame è teoricamente possibile, ma rappresenta il caso meno frequente. Osservando anche la recente riforma istituzionale italiana (si vedano i decreti sull'attuazione del federalismo fiscale che prevedono l'istituzione di nuovi tributi locali - D.Lgs. 446/97 - e la contemporanea abolizione dei trasferimenti erariali dallo Stato alle Regioni - D.Lgs. 56/2000), questa ipotesi trova un effettivo riscontro nella realtà. Si nota, infatti, una crescente autonomia delle Regioni, caratterizzata da un minor legame effettivo – in termini di finanziamento della spesa locale – tra il centro e gli enti sub-centrali. I dati (Buglione, 2005) suggeriscono che la capacità di autofinanziamento – misurata dalla percentuale delle entrate proprie sulle entrate correnti – delle RSO è passata dal 1996 al 2000 da circa il 20% a oltre il 50%. Similmente, il livello di decentramento del gettito tributario si è quasi raddoppiato.

⁹ L'evidenza empirica mostra che un comportamento sempre più diffuso da parte dei governi consiste nell'assegnare maggiori risorse alla spesa redistributiva in beni e servizi - *in kind* - rispetto a quelle destinate per i trasferimenti monetari - *in cash*. In Italia, ad esempio, nel 2001 la spesa sanitaria è pari a circa il 6,3% del PIL; la spesa per l'istruzione è circa il 4,9% del PIL; mentre le spese per le politiche a sostegno della famiglia e del mercato del lavoro (comprehensive dei sussidi alla disoccupazione) sono rispettivamente pari all'1% e all'1,1% del PIL (OECD, 2004).

massimizzi l'utilità media dei suoi residenti, mentre il governo centrale consideri una funzione obiettivo che è la somma pesata delle funzioni obiettivo dei governi locali (Tresch, 2002) ovvero la somma pesata delle utilità medie considerate dai *policy maker* locali.

I processi di massimizzazione dei due livelli di governo vengono analizzati distintamente, benché essi facciano parte dello stesso scenario istituzionale.

2.1.1 I governi locali

Il problema di massimizzazione si presenta in modo analogo per i due governi locali. La funzione obiettivo della regione i -esima ($i = 1,2$) è data dalla seguente espressione:

$$W_i = \alpha \bar{y}_i (1 - t - t_i) + \bar{\gamma}_i (\ln g + \ln g_i) \quad (1)$$

$$\text{con } \bar{\gamma}_i = \left(1 - \frac{\bar{y}_i}{y_{\max,i}} \right) \quad (1a)$$

L'equazione (1) rappresenta l'utilità media dei cittadini nella regione¹⁰. Oltre alle variabili fiscali locali (t_i e g_i) e centrali (t e g), nella funzione obiettivo della regione è presente il parametro di preferenza media per la spesa sanitaria ($\bar{\gamma}_i$)¹¹. Rispetto a tale parametro, l'idea è che la preferenza per la spesa sia definita sul complesso della funzione sanità, e non sulle componenti di tale funzione (es. ospedali e farmacie)¹². In tal senso, l'importanza che tale spesa ha sul benessere degli individui non dipende dal fatto che a fornire il servizio sia lo Stato o la regione, ma che entrambi forniscano

¹⁰ Il reddito medio locale (\bar{y}_i) rappresenta la misura di sintesi di tutti i redditi. Il reddito di riferimento è quello medio – e non il reddito mediano - in quanto i governi sono benevolenti e non interessati alla loro rielezione (in tal caso, sarebbe stato più opportuno considerare le preferenze dell'elettore mediano). Inoltre, essi non conoscono le preferenze puntuali ma solo la forma della distribuzione. Anche per questo motivo, si calcola la media.

¹¹ Il parametro di preferenza media assume valori tra zero e uno: $0 \leq \bar{\gamma}_i < 1$. Se tutti i redditi fossero uguali, allora anche il reddito medio coinciderebbe con il reddito massimo e $\bar{\gamma}_i$ assumerebbe il suo valore minimo. D'altra parte, se il reddito massimo fosse molto distante dalla media, la distribuzione sarebbe maggiormente sperequata rispetto al caso precedente. In tal caso, il parametro $\bar{\gamma}_i$ si avvicina al suo valore massimo. L'estremo superiore, infatti, è escluso in quanto $\bar{\gamma}_i = 1$ implica che il reddito medio sia zero e questo è possibile solo se tutti i redditi sono uguali a zero, in assenza di redditi negativi, ovviamente. In tal caso, anche il reddito massimo sarebbe pari a zero e il rapporto tra reddito medio e reddito massimo sarebbe indeterminato.

¹² In media, non si preferisce avere più ospedali o più farmacie, ma si ha preferenza affinché ospedali e farmacie esercitino il loro ruolo di complementi nello svolgimento della funzione sanità. Inoltre, dal momento che non si conoscono le preferenze individuali, attribuire lo stesso peso – in termini di funzione di utilità aggregata - per beni diversi (ospedali e farmacie) è arbitrario tanto quanto assumere pesi diversi per tali beni. Di certo, un solo parametro rende il modello più parsimonioso.

beni essenziali – e non intercambiabili - di tale funzione¹³. In sede di massimizzazione della funzione obiettivo, il *policy maker* locale è tenuto al rispetto del pareggio di bilancio¹⁴:

$$t_i Y_i = g_i N_i \quad (2)$$

Dalla massimizzazione vincolata, si ottengono le scelte ottime del governo locale:

$$\bullet \quad g_i^* = \frac{\bar{\gamma}_i}{\alpha} \quad (3)$$

$$\bullet \quad t_i^* = \frac{\bar{\gamma}_i}{\alpha} \left(\frac{1}{\bar{y}_i} \right) \quad (4)$$

La quota di spesa sanitaria pro capite erogata dalla regione dipende direttamente dal parametro che esprime la disuguaglianza del reddito e inversamente dal parametro di preferenza per il reddito privato. Similmente, l'aliquota ottimale è funzione crescente del parametro $\bar{\gamma}_i$; decrescente del parametro α e del reddito medio locale (\bar{y}_i) ¹⁵.

A fronte di una crescente disparità tra i redditi, l'ammontare di spesa erogato sarà più alto; per finanziarlo, occorre agire sul gettito. Il governo locale alzerà l'aliquota, a parità di base imponibile; mentre se aumenta la base imponibile, *ceteris paribus*, può permettersi di fissare un'aliquota più bassa.

2.1.2 Il governo centrale

Il problema di massimizzazione del governo centrale è analogo a quello dei governi locali. In particolare, la funzione obiettivo del governo centrale è data dalla somma pesata - dove i pesi sono dati dal rapporto tra la popolazione locale e quella nazionale: N_i/N ¹⁶ - delle utilità medie dei

¹³ In tal senso, si assume l'esistenza di un livello minimo di spesa (si pensi, ai livelli minimi essenziali stabiliti per le prestazioni sanitarie) che deve essere fornito da entrambe le autorità di governo (centrale e locale). In definitiva, si assume: $g \geq 1$ e $g_i \geq 1$. A tale proposito, si può osservare che anche storicamente si è spesso verificata una sovrapposizione di competenze tra diversi livelli di governo che ha portato ad avere che una stessa funzione sia svolta dal governo centrale e dalle unità sub-centrali.

¹⁴ Si assume che tutte le risorse prelevate dalla collettività ($t_i Y_i$) siano destinate al finanziamento della spesa pubblica a favore della popolazione locale ($g_i N_i$).

¹⁵ In particolare, per la spesa si ha: $\frac{\partial g_i^*}{\partial \gamma_i} = \frac{1}{\alpha} > 0$ e $\frac{\partial g_i^*}{\partial \alpha} = -\frac{\bar{\gamma}_i}{\alpha^2} < 0$; per l'aliquota: $\frac{\partial t_i^*}{\partial \gamma_i} > 0$; $\frac{\partial t_i^*}{\partial \alpha} < 0$; $\frac{\partial t_i^*}{\partial \bar{y}_i} < 0$.

¹⁶ Data, per ipotesi, la stessa dimensione della popolazione locale ($N_1 = N_2$), la somma pesata delle utilità medie locali, si riduce alla media di tali funzioni di utilità.

residenti delle due regioni ovvero dalla somma pesata delle funzioni obiettivo dei governi locali (Tresch, 2002):

$$W = \sum_{i=1}^2 \frac{N_i}{N} [\alpha \bar{y}_i (1 - t - t_i) + \bar{\gamma}_i (\ln g + \ln g_i)] \quad (5)$$

Il mix ottimale di spesa e aliquota¹⁷ si ottiene nel rispetto del vincolo di pareggio del bilancio:

$$tY = gN \quad (6)$$

I valori ottimali per la spesa centrale (g^*) e per l'aliquota unica (t^*) sono dati dalle seguenti espressioni:

$$\bullet \quad g^* = \frac{\bar{\beta}}{\alpha} \quad (7)$$

$$\bullet \quad t^* = \left(\frac{\bar{\beta}}{\alpha} \right) \frac{1}{y} \quad (8)$$

$$\text{con} \quad \bar{\beta} = \sum_{i=1}^2 \left(\frac{N_i}{N} \right) \bar{\gamma}_i \quad (9)$$

Il parametro $\bar{\beta}$ è definito dalla (9), come somma pesata delle disuguaglianze locali ($\bar{\gamma}_1$ e $\bar{\gamma}_2$)¹⁸. Esso tiene conto delle caratteristiche della distribuzione del reddito in entrambe le regioni – e del grado di disuguaglianza ad esse connesso - e, come tale, fornisce indicazioni sulla distribuzione del reddito nel complesso. Dunque, $\bar{\beta}$ esprime una misura di sintesi delle diverse preferenze per la spesa sanitaria erogata dal governo centrale nelle due regioni.

Sulla base di queste considerazioni, il governo centrale sceglie un ammontare di spesa – equazione (7) - che è funzione crescente della disuguaglianza del reddito presente nelle due regioni;

¹⁷ Il governo centrale fissa un'aliquota unica – proporzionale sul reddito – e indifferenziata su tutto il territorio per evitare disparità geografiche tra individui e distorsioni dovute ad una differente tassazione, secondo il principio del “trattamento uguale degli uguali” che risiedono nello stesso territorio. Infatti, il “trattamento differenziato degli uguali” è ammesso solo a livello locale ovvero tra individui che appartengono a diverse giurisdizioni e come tali sono soggetti alla rispettiva politica locale, che può differire da quella promossa dalle altre regioni.

¹⁸ Essendo i pesi dati dal rapporto tra la popolazione regionale (uguale nelle due regioni) e quella complessiva, la somma pesata dei $\bar{\gamma}_i$ si riduce alla media di tali parametri. In altri termini, $\bar{\beta}$ è dato dalla media dei $\bar{\gamma}_i$ regionali.

un'aliquota – equazione (8) - che dipende positivamente dal parametro $\bar{\beta}$ e negativamente dal reddito medio complessivo (\bar{y}), *ceteris paribus*¹⁹.

2.2 La funzione sanità nel caso di *finanza derivata*

Il secondo scenario prevede che ci sia un unico decisore centrale che si occupi del benessere di tutta la collettività. Senza ricorrere al decentramento fiscale, esso utilizza indirettamente il meccanismo del trasferimento di risorse alle unità di governo sottostanti. Queste diventano come delle “scatole vuote”, ovvero prive dei poteri di prelievo e di spesa, e si limitano a trasferire, a loro volta, risorse sotto forma di beni e servizi ai rispettivi cittadini (*finanza derivata*).

L'idea di base è che il decisore definisca la spesa sanitaria a livello centrale e per le due regioni, assegnando in modo opportuno le risorse per la popolazione nazionale (g) e quelle per la popolazione delle due regioni (g_1 e g_2), applicando un'aliquota unica (t_d) alla base imponibile nazionale²⁰. Esso massimizza la sua funzione obiettivo, espressione dell'utilità aggregata degli individui, nel rispetto del vincolo di pareggio di bilancio:

$$W_d = \sum_{i=1}^2 \frac{N_i}{N} \left[\alpha \bar{y}_i (1 - t_d) + \bar{\gamma}_i (\ln g + \ln g_i) \right] \quad (10)$$

$$t_d Y = gN + \sum_{i=1}^2 g_i N_i \quad (11)$$

La soluzione del problema del decisore centrale è caratterizzata dalle seguenti equazioni:

$$\bullet \quad g^* = \frac{\bar{\beta}}{\alpha} \quad (12)$$

$$\bullet \quad g_i^* = \frac{\bar{\gamma}_i}{\alpha} \quad (13)$$

$$\bullet \quad t_d^* = 2 \frac{\bar{\beta}}{\alpha} \left(\frac{1}{\bar{y}} \right) \quad (14)$$

¹⁹ Infatti, si ha: $\frac{\partial g^*}{\partial \beta} = \frac{1}{\alpha} > 0$; $\frac{\partial t^*}{\partial \beta} = \left(\frac{1}{\alpha} \right) \frac{1}{\bar{y}} > 0$ e $\frac{\partial t^*}{\partial \bar{y}} = - \left(\frac{\bar{\beta}}{\alpha} \right) \frac{1}{(\bar{y})^2} < 0$.

²⁰ Si assume che il decisore centrale fissi un'aliquota proporzionale indifferenziata sul territorio, per evitare disparità regionali. Essa sarà presumibilmente diversa dall'aliquota centrale, da quella locale - ottenute nel caso di *finanza autonoma* - e dalla somma di queste ultime.

Le espressioni che definiscono la spesa sanitaria pro capite a livello centrale e locale, sono analoghe a quelle ottenute nel caso in cui entrambi i livelli di governo si occupino di tale funzione in modo autonomo²¹ – equazioni (3) e (7).

L'uguaglianza delle soluzioni conferma che indipendentemente da chi gestisce la funzione sanità, le scelte ottime per la spesa sono le stesse, in quanto il decisore pur pensando alla collettività nel suo complesso riesce a fissare il set di beni e servizi che Stato e regioni avrebbero scelto autonomamente per le rispettive popolazioni.

L'aliquota unica è invece il doppio di quella che il governo centrale fisserebbe in presenza di tassazione locale (*finanza autonoma*)²². Dato il differente vincolo di bilancio, l'aliquota applicata dal decisore non coincide con quella scelta dal governo centrale ($t_d^* \neq t^*$), né con quella scelta dal governo regionale ($t_d^* \neq t_i^*$) a seguito dei loro processi di massimizzazione autonoma²³.

Prima di analizzare empiricamente quale tra i due scenari proposti sia maggiormente rappresentativo per il caso italiano, si fornisce un quadro del processo normativo che ha riguardato il finanziamento della sanità in Italia.

3. Il decentramento fiscale e la sanità in Italia: aspetti istituzionali

Il decentramento fiscale in Italia ha prodotto consistenti modifiche nel bilancio delle RSO: a partire dal 2001, è aumentata la dimensione dell'autofinanziamento basato sui tributi propri (Irap e addizionale regionale Irpef) e devoluti (compartecipazione all'Iva e all'accisa sulla benzina) e si è contemporaneamente ridotto il peso dei trasferimenti erariali (Messina, 2001). In particolare per l'anno 2001, l'ammontare totale dei trasferimenti soppressi è pari a 56.656 miliardi di lire, di cui ben il 99,2% riferiti alla sanità (Dirindin, 2001). Di qui, è evidente il ruolo del settore sanitario nella realizzazione del federalismo fiscale.

D'altra parte, dal lato delle nuove entrate, i cambiamenti si concentrano principalmente in una voce: la compartecipazione all'Iva. Essa è, infatti, il tributo che compensa circa il 94% dei trasferimenti aboliti, fornendo alle regioni entrate pari a 53.127 miliardi di lire. Dall'osservazione di tali dati, si

²¹ Anche i rapporti tra i parametri sono gli stessi: gli strumenti fiscali sono tutti funzioni crescenti del rispettivo parametro di disuguaglianza e decrescenti del parametro di preferenza per il reddito (o consumo) privato. Questo si nota dai segni delle derivate parziali: $\frac{\partial g^*}{\partial \beta} > 0$, $\frac{\partial g_i^*}{\partial \gamma_i} > 0$ e $\frac{\partial t_d^*}{\partial \beta} > 0$. Mentre: $\frac{\partial g^*}{\partial \alpha} < 0$, $\frac{\partial g_i^*}{\partial \alpha} < 0$ e $\frac{\partial t_d^*}{\partial \alpha} < 0$.

²² Infatti, risulta: $t_d^* = 2t^*$, poiché $t^* = \frac{\beta}{\alpha} \left(\frac{1}{y} \right)$.

²³ In realtà, essa non coincide nemmeno con la somma delle due aliquote ($t_d^* \neq t^* + t_i^*$). In questo caso, infatti, l'aliquota unica deve garantire contestualmente la copertura di tutte le spese sanitarie, intendendo quelle che il governo centrale e le due regioni erogherebbero separatamente in un regime di *finanza autonoma*. In ogni caso, si riscontra una certa simmetria tra l'espressione che identifica t_d^* e le equazioni che definiscono le aliquote ottimali dei precedenti contesti (t^* e t_i^*).

potrebbe quasi affermare che con la riforma introdotta nel 2001, si sia attuata la sostituzione del Fondo Sanitario Nazionale con la compartecipazione all'Iva.

Prima di tale riforma, il SSN era finanziato sulla base del fabbisogno nazionale corrispondente all'erogazione dei livelli essenziali di tutela sanitaria e la ripartizione di tale importo tra le regioni veniva determinata sulla base della cosiddetta "quota capitaria ponderata" (Muraro, 2002) commisurata ai bisogni della popolazione (indicatori relativi a sesso, età, densità della popolazione ecc.). Il fabbisogno regionale così calcolato era coperto, in parte, dal 90% del gettito locale dell'Irap e, per la parte residua, da un apposito trasferimento erariale a valere sul FSN.

Con il nuovo regime, invece, si realizza la completa abolizione dei trasferimenti erariali integrativi alle RSO per la spesa sanitaria (con alcune eccezioni) che vengono sostituiti con: l'aumento dell'aliquota dell'addizionale regionale Irpef (di 0,4 punti percentuali²⁴); l'aumento della compartecipazione all'accisa sulla benzina (da 242 a 250 lire al litro); la compartecipazione regionale al gettito Iva (nella misura del 27,5%²⁵).

In realtà, tra le fonti di finanziamento della spesa sanitaria, rimane l'Irap, istituita (D.Lgs. 446/97) per sostituire i contributi sanitari, precedentemente regionalizzati. Da trasferimenti vincolati e prelievi "di scopo" - i contributi sanitari - si è passati a risorse proprie locali e "non di scopo", benché vincolati per legge al finanziamento della sanità (Liberati, 2001). Infatti, il meccanismo di imputazione dell'Irap prevede che il relativo gettito (al 90%) rimanga vincolato al finanziamento della spesa sanitaria, accompagnato dal gettito (per intero) dell'addizionale regionale Irpef.

In sintesi, per ciascuna Regione il fabbisogno per la sanità è coperto attraverso la fiscalità generale e, in particolare, con imposte dirette (Irap e addizionale Irpef) ed indirette (compartecipazione all'Iva e all'accisa sulla benzina) e residuali trasferimenti erariali.

Alla luce di questi fatti e del crescente peso che la spesa per la sanità ha sul bilancio pubblico nazionale e regionale, si è effettuato un esercizio di verifica empirica del modello teorico, per verificare se il cambiamento normativo prodotto dal decentramento fiscale sul finanziamento del settore sanitario si sia effettivamente realizzato.

4. Una verifica empirica per l'Italia

Dopo aver derivato le soluzioni ottime del modello teorico (spesa sanitaria pro capite e aliquota proporzionale) per entrambi gli scenari istituzionali, si procede al calcolo dei valori numerici di tali variabili utilizzando un campione di dati reali (microdati di Banca d'Italia, "Indagine sui bilanci

²⁴ L'aliquota base del tributo è stata elevata dallo 0,5 allo 0,9 per cento; quella massima dall'1 all'1,4 per cento. Corrispondentemente, è stata ridotta di 40 punti base l'aliquota erariale (Messina, 2001).

²⁵ Questa è la misura indicata dal D.Lgs. 56/2000. Essa è stata rideterminata sulla base dei dati consuntivi del 1999 e dal 2001 innalzata al 38,55%.

delle famiglie italiane”, 2002²⁶). Questo è dovuto al fatto che le grandezze teoriche dipendono da parametri legati alla distribuzione del reddito regionale e nazionale.

Il passo successivo consiste nel confrontare tali valori teorici con le misure effettive di politica sanitaria, per le singole Regioni e per lo Stato. In tal senso, si sono utilizzati i dati di fonte Istat (2002) e del Ministero della Salute (2002) per le misure della spesa sanitaria regionale e nazionale e per le relative fonti di finanziamento (Irap, addizionale regionale Irpef, compartecipazione all’Iva). Valutando il quadrato degli scarti tra le due misure, si calcola l’errore che si commette approssimando le grandezze effettive con quelle teoriche degli scenari istituzionali di *finanza autonoma e derivata*²⁷.

Il confronto più interessante riguarda il lato della tassazione²⁸ ovvero tra le aliquote applicate dagli operatori pubblici nei due contesti e le aliquote effettive legate al finanziamento della spesa sanitaria. Tecnicamente, il valore delle aliquote teoriche è ottenuto dal modello opportunamente calibrato sulla realtà italiana; il valore delle aliquote effettive come rapporto tra gettiti e basi imponibili²⁹. Lo scarto minore tra tassazione teorica ed effettiva fornisce l’errore più piccolo, che dovrebbe indicare a quale dei due scenari proposti si avvicina maggiormente il finanziamento della sanità in Italia. In altri termini, se l’errore che si commette approssimando le aliquote effettive con quelle teoriche nel caso di *finanza autonoma* è inferiore a quello ottenuto effettuando la stessa approssimazione nel caso di *finanza derivata*, allora è accettabile l’interpretazione secondo cui in Italia le politiche pubbliche per la sanità (lato entrate) sono condotte in modo autonomo da Stato e Regioni, ognuno secondo le proprie competenze.

D’altra parte, se l’errore calcolato nell’ipotesi di *finanza derivata* è inferiore a quello nel caso di *finanza autonoma*, si può affermare che il modello teorico più vicino alla realtà italiana è quello caratterizzato dalla presenza di un solo operatore pubblico che, a fronte di un prelievo uniforme su tutto il territorio³⁰, definisce la spesa sanitaria per tutti i livelli governo (decisioni centralizzate).

4.1 Il campione di dati

L’analisi viene condotta utilizzando l’indagine della Banca d’Italia sui bilanci delle famiglie italiane, che contiene le informazioni raccolte presso le famiglie italiane nell’indagine campionaria

²⁶ Per la verifica empirica, si è scelto di utilizzare il 2002 come anno di riferimento per due principali motivi: esso rappresenta il primo anno di attuazione della riforma e, al contempo, il più recente come disponibilità di dati per il campione di Banca d’Italia. A tale proposito, si veda anche la nota 32.

²⁷ Tale scostamento non viene considerato in modo a sé stante, ma nell’ottica di comparare due diversi regimi istituzionali.

²⁸ Questo è dovuto al fatto che le soluzioni del modello per la spesa sanitaria pro capite restano immutate, anche considerando regimi istituzionali diversi. Si confrontino a tale proposito le equazioni (3) e (7) per la *finanza autonoma*, con la (12) e la (13) per la *finanza derivata*.

²⁹ In altri termini, si ricorre alla definizione di aliquota implicita.

³⁰ La differenza principale tra i due regimi istituzionali risiede nel fatto che il prelievo fiscale messo in atto dal decisore centrale non coincide con quello definito dai *policy makers* locale e centrale. In altri termini, l’aliquota ottima nel caso di *finanza derivata* è diversa da quella fissata dal governo centrale e dai governi locali nel caso di *finanza autonoma*.

per l'anno 2002³¹. In dettaglio, per costruire la distribuzione del reddito su base regionale e nazionale – necessaria per il calcolo dei parametri di disuguaglianza $\bar{\gamma}_i$ e $\bar{\beta}$ - si è preso in considerazione il reddito individuale (come somma di altre componenti: reddito da lavoro dipendente, da capitale, da pensioni, ecc.) lordo³². Per una corretta elaborazione dei microdati di Banca d'Italia, è necessario procedere alla pesatura dei redditi. Questo perché la rappresentatività del campione di Banca d'Italia è abbastanza buona a livello nazionale, ma meno a livello regionale. In tal senso, si rendono necessari aggiustamenti in termini di pesatura del campione per consentire l'allineamento alle principali distribuzioni marginali socio-demografiche di fonte Istat, migliorando così il grado di rappresentatività del dataset utilizzato.

4.2 Il problema dei pesi

In generale, lo scopo della calibrazione dei pesi, da utilizzare con i dati sulla distribuzione del reddito delle famiglie, è quello di aumentare la precisione e di ridurre le differenze tra la distribuzione del campione e quella della popolazione reale con l'aiuto di variabili ausiliarie³³ (Faiella e Gambacorta, 2007). In altri termini, partendo dai pesi originari contenuti nel dataset di Banca d'Italia ($w_{h,i}$), è possibile ottenere, con opportune procedure di parametrizzazione, un nuovo set di pesi finali ($\hat{w}_{h,i}$) per rimediare a fenomeni di sottorappresentazione del campione. La correzione effettuata – benché molto parziale – permette di ottenere una buona rappresentatività regionale del campione, almeno per quanto concerne la numerosità della popolazione³⁴. Tecnicamente, si è applicata la seguente formula per ottenere i nuovi pesi per regione:

$$\hat{w}_{h,i} = \hat{N}_i \cdot \left(\frac{w_{h,i}}{\sum_{h \in i} w_{h,i}} \right) \quad (15)$$

³¹ L'indagine ha riguardato 8.011 famiglie estratte dalle liste anagrafiche di 344 comuni, composte di 21.148 individui.

³² Esso è stato ottenuto con opportune procedure a partire dal reddito disponibile netto fornito da Banca d'Italia. A tale proposito, si ringraziano Maria Grazia Pazienza e Chiara Rapallini (Università di Firenze, Dipartimento di Studi sullo Stato) per avere messo a disposizione dell'autrice il dataset per l'anno 2002 contenente la variabile reddito individuale disponibile (e tutte le sue componenti) *lordo*, ricostruito dalle medesime con elaborazioni, a partire dai microdati sul reddito individuale disponibile *netto* diffuso da Banca d'Italia.

³³ In questo contesto, come variabile ausiliaria si è utilizzato il dato della numerosità reale della popolazione.

³⁴ In realtà, il problema (della numerosità) della popolazione, principalmente a livello regionale, è solo una delle possibili variabili che causano la non corretta rappresentazione del campione. Infatti, potrebbero verificarsi problemi analoghi per il reddito (se, ad esempio, gli intervistati sono tutti soggetti senza reddito, minorenni, ecc.), per la composizione familiare (se, ad esempio, i soggetti intervistati appartengono tutti a famiglie con 5 figli, ecc.) e per altri aspetti relativi all'indagine. In tal senso, la correzione dei pesi sulla base della reale numerosità della popolazione è una correzione solo parziale ovvero per una sola dimensione, che permette di avere una buona rappresentatività del campione relativamente a tale aspetto. Per le altre variabili (tra cui il reddito) si assume che esse siano, invece, ben rappresentate.

\hat{N}_i è la dimensione effettiva della popolazione regionale (dato Istat); $w_{h,i}$ è la variabile dei pesi originari indicizzata per regione (ogni sua componente indica un individuo di quella regione); il denominatore della formula indica la somma dei pesi originari per regione. In altri termini, il rapporto indicato tra parentesi rappresenta il peso relativo dell'individuo all'interno della regione. Si veda l'Appendice (*Tavola A.1*) per i valori della popolazione - regionale e nazionale - calcolati a partire dal campione di Banca d'Italia con la tecnica del peso "regionalizzato" – espressione (15).

4.3 Il reddito individuale

Al fine di ricostruire la distribuzione del reddito a livello regionale e nazionale, si utilizza la variabile che descrive il reddito disponibile lordo individuale (Y), dato dalla somma delle seguenti componenti: reddito da lavoro dipendente (YL), pensioni e trasferimenti (YT), reddito da lavoro autonomo (YM) e reddito da capitale (YC), secondo la classificazione di Banca d'Italia. Tale grandezza è omnicomprensiva in quanto permette di considerare tutte le possibili fonti di reddito per l'individuo³⁵. Dopo aver costruito il nuovo set di coefficienti di ponderazione ($\hat{w}_{h,i}$), è stato possibile pesare correttamente i redditi individuali. La somma degli stessi, indicizzata per regione, consente di calcolare il reddito totale disponibile lordo all'interno di ogni regione e per l'Italia nel complesso. Avendo effettuato la correzione solo sulla popolazione, è possibile che le altre variabili del campione – reddito - non trovino un perfetto allineamento con i dati reali. Si è valutata, dunque, la corrispondenza tra il reddito campionario, per regione e in aggregato, e il reddito reale.

La *Tavola A.2* (in Appendice) mostra il reddito di Banca d'Italia (dopo le tecniche di riproporzionamento della popolazione) e quello rilevato dall'Istat, per il 2002. La prima colonna (A) contiene i redditi individuali lordi pesati con $\hat{w}_{h,i}$ e poi sommati, per regione e per l'Italia; la colonna (B) descrive il reddito disponibile³⁶ lordo per il totale delle famiglie all'interno di ogni regione; la colonna (C) riporta i valori del Prodotto Interno Lordo (PIL) regionale e per l'Italia. Non sorprende notare che il reddito BdI sottostima parzialmente il reddito disponibile lordo Istat³⁷: dalle ultime due colonne si nota che il primo copre, in media, intorno all'80% del reddito disponibile lordo totale, e circa il 55% del PIL. Alla luce di queste discrepanze tra le misure di reddito, si è

³⁵ Dal campione originario, sono stati eliminati gli individui con età inferiore ai 18 anni dal momento che, essendo interessati alla distribuzione del reddito individuale, si presume che non detengano un reddito proprio. Inoltre, sono stati esclusi gli individui residenti nelle Regioni a Statuto Speciale. In definitiva, il numero totale di osservazioni per la variabile Y nell'anno 2002 è pari a 14.853. Il totale degli individui viene poi ripartito nelle varie regioni, attraverso la variabile IREG che fornisce il codice Istat della regione di appartenenza.

³⁶ La dicitura "reddito disponibile lordo totale" si riferisce alla voce "reddito disponibile lordo", totale famiglie, Conto della distribuzione secondaria del reddito (Impieghi). Per reddito disponibile si intende "l'ammontare di risorse correnti degli operatori per gli impieghi finali (consumo e risparmio)", secondo la definizione Istat.

³⁷ Questo potrebbe essere dovuto sostanzialmente a due ragioni: la prima è che l'indagine di Banca d'Italia non riesce a stimare esattamente tutte le componenti di reddito, nel senso che alcune di queste sono sottorappresentate, come è il caso del reddito da lavoro autonomo; la seconda è che, come evidenziato più volte, la correzione apportata è parziale, in quanto è solo sulla popolazione. In tal senso, le altre variabili del campione potrebbero risentire di questa correzione e risultare sottorappresentate.

ritenuto opportuno procedere ad ulteriori elaborazioni per testare la robustezza dei risultati raggiunti con i dati campionari di Banca d'Italia (paragrafo 8.3).

5. Metodologia e calcolo dei parametri

Il punto di partenza dell'analisi è calcolare i parametri del modello che determinano a loro volta le soluzioni ottime (spesa pro capite e aliquota) dei due scenari. In particolare, ci si riferisce a $\bar{\beta}$ e $\bar{\gamma}_i$. Essi vengono calcolati come valori puntuali a partire dalle distribuzioni del reddito regionale e nazionale utilizzando i microdati di Banca d'Italia.

Dal modello, si ha che il parametro $\bar{\gamma}_i$ è funzione della distanza tra il reddito medio e il reddito massimo locale – espressione (1a) – mentre $\bar{\beta}$ è dato dalla somma pesata dei $\bar{\gamma}_i$ regionali – equazione (9). Ai fini della misurazione empirica di $\bar{\beta}$, si estendono i risultati del modello considerando 15 Regioni, invece che 2, dato l'elevato grado di simmetria riscontrato per i $\bar{\gamma}_i$ regionali. In tal senso, il parametro $\bar{\beta}$ - disuguaglianza complessiva - viene calcolato come somma ponderata di tutti gli indici di disuguaglianza regionali ($\bar{\gamma}_i$). Riguardo ai pesi, si considera il rapporto tra la popolazione locale e quella nazionale - $\frac{N_i}{N}$ - come nel modello, “ignorando” però l'ipotesi di uguale popolazione nelle due regioni per considerare, invece, popolazioni regionali differenti sulla base dei dati del campione (*Tavola A.1*). Dunque, il calcolo di $\bar{\beta}$ avviene a partire dai $\bar{\gamma}_i$ di tutte le regioni, secondo la seguente espressione:

$$\bar{\beta} = \sum_{i=1}^{15} \left(\frac{N_i}{N} \right) \bar{\gamma}_i \quad (16)$$

5.1 I parametri regionali

Il primo passo dell'analisi consiste nel ricostruire la distribuzione del reddito all'interno di ogni regione e per l'Italia. I redditi degli individui di una data regione – adeguatamente pesati - danno luogo alla distribuzione originaria del reddito su base locale. Da questa, si stimano le misure di sintesi per il reddito, necessarie per il calcolo dei parametri del modello: il reddito medio e quello

massimo locali che determinano il parametro $\bar{\gamma}_i$. Il reddito medio regionale (\bar{y}_i) si calcola come reddito pro capite “puro”³⁸.

Per assegnare un valore puntuale al reddito massimo locale - $y_{\max,i}$ - si è scelto di non selezionare l’individuo con il reddito più alto in assoluto nella regione, per due principali motivi. Tale valore, infatti, potrebbe essere stato rilevato o registrato non correttamente in sede di indagine³⁹. In secondo luogo, esso potrebbe portare ad avere stime distorte, data la sua possibile natura di valore *outlier*. In tal senso, si è preferito ricorrere ad una misura formalmente più corretta e robusta per indicare il reddito massimo, utilizzando i percentili della distribuzione del reddito: il 99° percentile per ogni distribuzione regionale⁴⁰. Questo ha permesso di avere misure meno sensibili per individuare il soggetto più ricco all’interno di ogni regione e, soprattutto, di avere valori più conformi al reddito medio regionale. A tale proposito, la *Tavola 1*, mostra che i valori di $y_{\max,i}$ più elevati si riscontrano in quelle regioni dove anche \bar{y}_i è più elevato.

Tavola 1: PARAMETRI REGIONALI

REGIONI	REDDITO MEDIO \bar{y}_i (in euro)	REDDITO MASSIMO $y_{\max,i}$ (in euro)	$\bar{\gamma}_i = \left(1 - \frac{\bar{y}_i}{y_{\max,i}}\right)$
Piemonte	13.600	57.810	0,7647
Lombardia	16.837	97.014	0,8264
Veneto	13.222	75.119	0,8240
Liguria	15.710	65.820	0,7613
Emilia Romagna	16.303	91.394	0,8216
Toscana	14.270	87.809	0,8375
Umbria	12.748	71.044	0,8206
Marche	12.624	51.996	0,7572
Lazio	12.743	66.473	0,8083
Abruzzo	9.710	50.773	0,8088
Molise	9.594	59.366	0,8384
Campania	7.385	41.695	0,8229
Puglia	9.702	49.468	0,8039
Basilicata	7.977	59.499	0,8659
Calabria	7.709	50.773	0,8482

Fonte: Elaborazioni proprie su microdati Banca d'Italia 2002

³⁸ Ossia come rapporto tra la somma pesata di tutti i redditi individuali lordi di una regione diviso per la popolazione campionaria.

³⁹ E' noto, infatti, il problema dell'*underreporting* dei redditi molto alti (o molto bassi) nella maggior parte delle *surveys* che raccolgono informazioni sui redditi individuali o familiari (Atkinson et al., 1995; Gottschalk e Smeeding, 1999).

⁴⁰ La scelta del 99° percentile è dovuta al fatto che tale misura è la più prossima al reddito massimo ed è anche quella che appiattisce meno la distribuzione del reddito, rispetto ad esempio all'utilizzo del 90° percentile. In aggiunta, tale valore fornisce anche una buona rappresentazione del reddito massimo in termini relativi, fornendo valori accettabili anche nell'ottica di confronto tra le regioni (ad esempio, l'individuo più ricco in assoluto risulta essere residente in Lombardia, mentre quello “meno ricco” in assoluto in Campania).

Dal calcolo dei $\bar{\gamma}_i$, emerge che essi sono piuttosto elevati e abbastanza prossimi all'estremo superiore. In aggiunta, il campo di variazione dei $\bar{\gamma}_i$ regionali è abbastanza ristretto: esso va da un minimo di 0,7572 (Marche) ad un massimo di 0,8659 (Basilicata).

Per la maggior parte delle regioni (Lombardia, Veneto, Emilia Romagna, Toscana, Umbria, Lazio, Abruzzo, Molise, Campania, Puglia) l'indice oscilla tra 0,80 e 0,83. Le "eccezioni" verso l'alto sono rappresentate da Calabria e Basilicata – valore più alto del campione - mentre quelle verso il basso riguardano Piemonte, Liguria e Marche.

In queste, il reddito massimo è meno distante dalla media rispetto a quanto accade per il resto del campione. Ciò implica che la distribuzione del reddito risulti meno sperequata. Ci si aspetta, dunque, che l'ammontare di spesa sanitaria da erogare in queste regioni sia inferiore rispetto a quanto accade nelle altre. Diversamente, nelle regioni in cui $\bar{\gamma}_i$ è particolarmente alto, si ha che la disuguaglianza tra i redditi è maggiore. Per cui, anche la politica sanitaria andrà nella stessa direzione: aliquote e spesa pro capite più elevate secondo la relazione (diretta) che lega g_i^* e t_i^* al parametro $\bar{\gamma}_i$.

5.2 Il parametro nazionale

Dai $\bar{\gamma}_i$ regionali, è possibile ottenere il parametro $\bar{\beta}$ secondo l'espressione (16), utilizzando i dati di *Tavola 1* (parametri regionali) e quelli di *Tavola A.1* (popolazione).

Il valore finale di $\bar{\beta}$ è pari a: 0,8137.

Confrontando questo dato con quelli di *Tavola 1*, si nota che solo poche regioni (Piemonte, Liguria, Marche, Lazio, Abruzzo e Puglia) hanno valori della disuguaglianza regionale al di sotto della media nazionale. Le restanti mostrano valori di $\bar{\gamma}_i$ superiori al valore di $\bar{\beta}$.

Infine, per completezza di analisi e per verificare la validità dell'approccio considerato, si è effettuato il calcolo del parametro $\bar{\beta}$ anche con il procedimento utilizzato per i vari $\bar{\gamma}_i$. In altri termini, dopo avere ricostruito la distribuzione del reddito per l'Italia a 15 regioni, si sono evidenziati il reddito medio (12.875 euro), il reddito massimo (tecnicamente, il 99° percentile: 67.977 euro) e si è calcolato il parametro $\bar{\beta}_{ITA}$ secondo la seguente espressione:

$$\bar{\beta}_{ITA} = \left(1 - \frac{\bar{y}_{ITA}}{y_{\max,ITA}} \right) \quad (16a)$$

Il risultato ottenuto per $\bar{\beta}_{ITA}$ è pari a 0,8106. Confrontando tale valore con quello precedente ($\bar{\beta}$), si nota che i due differiscono poco (a partire dalla terza cifra decimale). Si può affermare che la misurazione di tale indice con l'estensione del modello a 15 regioni è adeguata e non si discosta molto dal risultato ottenuto con il secondo metodo. Tuttavia, poiché questa scelta “forzerebbe” il modello, in quanto il parametro nazionale è ipotizzato essere una media dei parametri regionali, si è deciso di considerare il primo valore di $\bar{\beta}$ per le successive elaborazioni.

6. Calcolo delle soluzioni ottime del modello

Oltre ai valori che indicano i gradi di disuguaglianza ($\bar{\beta}$ e $\bar{\gamma}_i$), occorre assegnare un valore numerico anche agli altri parametri che definiscono l'aliquota e la spesa ottime.

6.1 Il parametro α^*

In particolare, la spesa sanitaria pro capite – centrale e locale – in entrambi gli scenari è funzione di un parametro α (uguale per entrambe), che nel modello rappresenta la preferenza per il reddito privato⁴¹. Purché compreso tra 0 e 1, il valore di α potrebbe essere scelto liberamente. Tuttavia, per definirlo si è scelto di utilizzare il meccanismo della calibrazione sui dati reali. In dettaglio, si è optato per quel valore di α che restituisce il valore effettivo della spesa sanitaria nazionale – a dire, di competenza del governo centrale - pro capite (g^*) secondo i dati reali (Istat). Tecnicamente, utilizzando il valore numerico precedentemente calcolato per $\bar{\beta}$, si tratta di invertire l'espressione g^* , assumendo che il valore teorico di tale spesa coincida proprio con il dato effettivo della stessa ($g^* = \hat{g}$):

$$\alpha^* = \frac{\bar{\beta}}{\hat{g}} \quad (17)$$

La decisione di calibrare α^* rispetto ai dati nazionali della spesa e non a quelli regionali è dovuta al fatto che nel modello si prevede un unico parametro, uguale su tutto il territorio, sganciato dalla dimensione regionale. Questo implica utilizzare il dato effettivo della spesa sanitaria pro capite nazionale - e non regionale - per assegnare un valore adeguato a tale parametro. In dettaglio, la spesa sanitaria nazionale pro capite effettiva (\hat{g}) è pari a (circa) 802 euro⁴². Questo valore

⁴¹ Si vedano a tale proposito le espressioni (3) e (7) per la *finanza autonoma*; (12) e (13) per la *finanza derivata*.

⁴² In realtà, il fabbisogno sanitario complessivo pro capite è pari a 1.376 euro. Tale valore si ottiene dal rapporto tra la somma della spesa per la funzione sanità nelle 15 regioni considerate e la somma della popolazione effettiva di tali regioni per l'anno 2002. Di questo ammontare, parte viene finanziato dal governo centrale (802 euro) e parte dalle

rappresenta l'equivalente monetario della spesa sanitaria (fornitura di ospedali, prestazioni sanitarie, personale medico ecc.), assunta in beni e servizi nel modello e di competenza del governo centrale. Dalla (17), si ottiene un valore di α^* pari a 0,0010.

6.2 I valori teorici: spesa pro capite e aliquota

A questo punto, per derivare le soluzioni teoriche della spesa pro capite regionale per la funzione sanità (g_i^*), è sufficiente sostituire il valore di α^* e il rispettivo valore di $\bar{\gamma}_i$ (Tavola 1) nell'espressione (3) o nella (13), che risultano uguali nei due scenari istituzionali.

La Tavola 2 mostra l'equivalente monetario della spesa sanitaria pro capite stanziata dalle RSO, a partire dai dati del campione e calcolata sulla base delle soluzioni del modello teorico.

Tavola 2: SPESA SANITARIA PRO CAPITE REGIONALE E NAZIONALE (in euro)
E ALIQUOTA REGIONALE E NAZIONALE (in percentuale)

REGIONI	$g_i^* = \frac{\bar{\gamma}_i}{\alpha^*}$ $g^* = \hat{g}$	$t_i^* = g_i^* \left(\frac{1}{y_i} \right)$ $t^* = g^* \left(\frac{1}{y} \right)$ $t_d^* = 2g^* \left(\frac{1}{y} \right)$
Piemonte	765	5,62
Lombardia	826	4,91
Veneto	824	6,23
Liguria	761	4,85
Emilia Romagna	822	5,04
Toscana	837	5,87
Umbria	821	6,44
Marche	757	6,00
Lazio	808	6,34
Abruzzo	809	8,33
Molise	838	8,74
Campania	823	11,14
Puglia	804	8,29
Basilicata	866	10,86
Calabria	848	11,00
Italia (governo centrale: g^*, t^*)	802	6,23
Italia (decisore centrale: t_d^*)		12,46

Fonte: Elaborazioni proprie su microdati Banca d'Italia 2002

Regioni (in media, 574 euro). La prima componente è stata stimata come differenza tra la spesa sanitaria totale delle 15 regioni e il gettito totale di quei tributi locali propri (Irap e Addizionale Irpef) utilizzati per coprire il fabbisogno sanitario regionale. In altri termini, la spesa sanitaria di "competenza" del governo centrale dovrebbe rappresentare quello che il governo centrale spende per la sanità nel complesso, utilizzando risorse proprie, raccolte su base nazionale. In questa categoria rientrano le compartecipazioni all'Iva e all'accisa sulla benzina. Tale suddivisione delle fonti di finanziamento sarà utile anche in seguito, per la determinazione delle aliquote effettive.

In particolare, le regioni in cui la quota di spesa risulta elevata rispetto al resto del campione sono quelle caratterizzate da redditi medi piuttosto bassi e da valori di $\bar{\gamma}_i$ piuttosto elevati (Basilicata e Calabria). Non sorprende osservare che laddove il parametro di disuguaglianza regionale è più alto, *ceteris paribus*, anche la quota di spesa sanitaria da erogare vada nella stessa direzione.

Dai valori teorici della spesa sanitaria pro capite regionale e nazionale (quest'ultima per ipotesi coincide con il dato effettivo) è possibile calcolare i valori teorici delle aliquote regionali, del governo centrale e del decisore, rispettivamente secondo le espressioni (4), (8) e (14).

Osservando la seconda colonna di *Tavola 2*, si nota che l'aliquota fissata dal decisore centrale – regime di *finanza derivata* - risulta essere il doppio di quella applicata dal governo centrale, quando opera contestualmente ai governi locali – *finanza autonoma*.

Dall'osservazione delle aliquote regionali, emerge che alcune regioni del centro (Abruzzo, Molise) e tutte quelle del sud (Campania, Puglia, Basilicata, Calabria) registrano i valori più elevati del campione. Al contempo, esse superano, in certi casi (Campania, Basilicata e Calabria) anche notevolmente il valore dell'aliquota stabilita dal governo centrale (t^*). In alcune di queste regioni, anche la spesa pro capite è relativamente elevata. Non sorprende, dunque, che la tassazione sia congrua all'ammontare di spesa erogata, dato il loro legame nel vincolo di bilancio.

Tuttavia, si verifica anche che tassazioni relativamente basse riescono a coprire il finanziamento di spese, in proporzione, più elevate. Questo è dovuto all'influenza del reddito pro capite regionale ovvero alle risorse disponibili che possiede la regione: laddove il reddito medio è più elevato, è sufficiente un prelievo fiscale più basso per garantire lo stesso ammontare di spesa. E' questo il caso della Lombardia e dell'Emilia Romagna che detengono i redditi medi più elevati del campione e al contempo le aliquote più basse.

7. Calcolo dei valori effettivi (le aliquote)

Come evidenziato in precedenza, il confronto più interessante è quello dal lato della tassazione. Dopo avere ricostruito i valori teorici delle aliquote a partire dai risultati del modello, si procede dunque al calcolo delle aliquote effettive. Per ottenerle, si considerano quei tributi, in particolare a livello regionale, volti principalmente al finanziamento della spesa sanitaria.

In dettaglio, il fabbisogno sanitario complessivo nelle RSO viene coperto principalmente con: l'addizionale regionale Irpef; la compartecipazione all'accisa sulla benzina; la compartecipazione al gettito Iva; il gettito derivante dall'Irap. Mentre Irap e addizionale regionale Irpef rappresentato

tributi propri locali, la compartecipazione all'Iva costituisce una risorsa proveniente dal prelievo nazionale sugli scambi e, come tale, di competenza del governo centrale⁴³.

Si assume che le aliquote effettive dei rispettivi livelli di governo siano legate alla competenza di spesa sanitaria che vanno a finanziare. Ciò presuppone che le aliquote effettive regionali si ottengano dal rapporto tra il gettito che esse producono e la base imponibile locale, dove il gettito regionale è dato dalla somma dei gettiti di Irap e addizionale regionale Irpef⁴⁴.

Formalmente, le aliquote effettive regionali (\hat{t}_i) sono date da:

$$\hat{t}_i = \frac{T_i}{Y_i} \quad (18)$$

con T_i che è uguale alla somma di Irap e addizionale regionale Irpef e Y_i che rappresenta la base imponibile locale (*Tavola A.2, colonna A*).

Diversamente, l'aliquota effettiva del governo centrale (\hat{t}) viene calcolata come rapporto tra il gettito derivante dalle compartecipazioni (Iva e accisa)⁴⁵ e la base imponibile nazionale (Y):

$$\hat{t} = \frac{\left(SST - \sum_{i=1}^{15} T_i \right)}{Y} \quad (19)$$

Tecnicamente, il gettito nazionale (al numeratore) è ottenuto come grandezza residuale, ovvero dalla differenza tra il fabbisogno sanitario complessivo delle RSO (SST = spesa sanitaria totale) e la parte di tale fabbisogno coperta dalle risorse regionali (somma di T_i per tutte le regioni). Tale valore fornisce la stima della spesa sanitaria di competenza del governo centrale che è circa uguale al totale delle compartecipazioni⁴⁶.

⁴³ Come descritto in precedenza, essa è stata introdotta in sostituzione dei trasferimenti erariali dallo Stato alle Regioni per garantire a queste ultime risorse per finanziare spese essenziali come la sanità.

⁴⁴ Poiché la voce delle compartecipazioni può essere assimilata ad un trasferimento di risorse dallo Stato alle Regioni o meglio ad un tributo "devoluto" - e non un vero tributo proprio - essa non viene inclusa tra i tributi che danno luogo al gettito regionale. In aggiunta, considerare le compartecipazioni come fonte di finanziamento della spesa sanitaria regionale avrebbe creato, in tale contesto, una sovrastima dell'aliquota effettiva locale e una duplicazione delle fonti di finanziamento della spesa locale.

⁴⁵ In dettaglio, il gettito "nazionale" viene stimato come differenza tra la spesa sanitaria complessiva e la somma dei gettiti regionali, destinati al finanziamento della quota di spesa di competenza delle regioni, come indicato dal numeratore dell'espressione (19). Detto altrimenti, la parte di fabbisogno sanitario, che non viene coperta dalle risorse regionali, viene finanziata dal governo centrale e dovrebbe rappresentare quanto lo Stato spende per la sanità nelle RSO. Il fatto che poi questo attribuisca - attraverso meccanismi di perequazione - tali risorse alle Regioni è solo un modo diverso di ripartire una spesa che, di fatto, è stata finanziata con un'imposta generale applicata su tutto il territorio.

⁴⁶ In realtà, esso è relativamente più grande della voce delle compartecipazioni all'Iva e all'accisa sulla benzina. Infatti, essendo stato calcolato come differenza tra la spesa sanitaria totale e quella di competenza delle Regioni, include anche altri introiti minori oltre alla voce principale delle compartecipazioni. I dati reali mostrano che il fabbisogno sanitario

La *Tavola 3* riporta le aliquote effettive regionali e nazionale.

Tavola 3: ALIQUOTE EFFETTIVE

(valori percentuali)

	Gettito centrale (stima) (miliardi di euro) $= SST - \sum_{i=1}^{15} T_i$	$\hat{t} = \frac{\left(SST - \sum_{i=1}^{15} T_i \right)}{Y}$
REGIONI	Gettito regionale (miliardi di euro) T_i	$\hat{t}_i = \left(\frac{T_i}{Y_i} \right)$
Piemonte	2,72	4,74
Lombardia	8,12	5,32
Veneto	3,35	5,57
Liguria	0,80	3,25
Emilia Romagna	3,06	4,68
Toscana	2,15	4,31
Umbria	0,37	3,51
Marche	0,81	4,34
Lazio	3,39	5,18
Abruzzo	0,48	3,93
Molise	0,04	1,41
Campania	1,42	3,36
Puglia	0,74	1,91
Basilicata	0,07	1,48
Calabria	0,17	1,10
Italia (governo centrale)	38,74	6,23

Fonte: Elaborazioni proprie su microdati Banca d'Italia 2002 e dati Ministero della Salute 2002

Nella prima colonna: i valori dei gettiti regionali e del gettito attribuibile al governo centrale calcolato come residuo; nella seconda: le aliquote effettive ottenute applicando le espressioni (18) e (19). Non sorprende osservare che le aliquote effettive più basse sono quelle delle regioni relativamente più povere (Basilicata e Calabria) e più piccole (Molise), che possono contare su un ammontare esiguo di tributi propri rispetto al resto del campione. Presumibilmente, queste saranno

complessivo (SST) per le RSO nel 2002 ammonta a circa 66,5 miliardi di euro. La somma complessiva dei gettiti regionali di Irap e addizionale Irpef è pari a 27,71 miliardi di euro. Il calcolo del residuo rappresenta il gettito nazionale o meglio la stima della spesa sanitaria di competenza del governo centrale, che è uguale a 38,747 miliardi di euro. Questo significa che circa il 40% del fabbisogno sanitario complessivo è coperto dalla fiscalità regionale, mentre il restante 60% dalla fiscalità nazionale attraverso le entrate relative alle imposte indirette, quali l'Iva, l'accisa e altri tributi minori.

anche le regioni che riceveranno una quota relativamente maggiore di compartecipazione all'Iva secondo i meccanismi del Fondo perequativo nazionale⁴⁷.

Ovviamente, anche il denominatore – rappresentato dal reddito disponibile lordo di Banca d'Italia – contribuisce ad alzare o abbassare l'aliquota effettiva, a parità del resto. Ad esempio, il fatto che la Lombardia abbia un'aliquota piuttosto simile a quella del Piemonte, benché il gettito regionale della prima sia circa 4 volte quello della seconda, è dovuto all'influenza del reddito lordo totale (3 volte quello del Piemonte) che spinge ad abbassare la tassazione in Lombardia quasi ai livelli del Piemonte, *ceteris paribus*.

Riguardo all'aliquota nazionale (6,23), si nota che essa è piuttosto elevata rispetto alle aliquote regionali. Questo è dovuto al fatto che la parte di spesa sanitaria stimata di competenza del governo centrale è notevolmente superiore a quella delle singole regioni.

In tal senso, anche il gettito complessivo della compartecipazione all'Iva e all'accisa è senza dubbio maggiore, essendo ottenuto da un prelievo su base nazionale.

A questo punto, è possibile confrontare la serie delle aliquote teoriche (*Tavola 2*) con quella delle aliquote effettive (*Tavola 3*) per il finanziamento della spesa sanitaria e valutare lo scostamento relativo tra le due misure. In particolare, l'idea è confrontare le aliquote non prese singolarmente, ma sulla base degli scenari proposti nel modello, cercando di capire quale sia quello più realistico per il caso italiano.

8. Il calcolo degli errori

In questo paragrafo vengono presentati i risultati conclusivi dell'analisi empirica. Si tratta di confrontare le misure di errore in seguito all'approssimazione dei valori effettivi delle aliquote con i valori teorici delle stesse rispettivamente nel caso di *finanza autonoma* e *derivata*. L'idea di fondo è che l'errore più piccolo tra i due dovrebbe indicare quale scenario sia più plausibile per l'Italia. Detto altrimenti, la migliore approssimazione evidenzierrebbe la direzione verso cui si muove il finanziamento della sanità nelle RSO.

Prima di procedere al calcolo degli errori, si è ritenuto opportuno testare la validità dell'ipotesi di scostamento tra la serie delle aliquote effettive e quella delle aliquote teoriche per entrambi gli scenari⁴⁸. In tal senso, si è effettuato preliminarmente il *test di Wilcoxon* che viene utilizzato per confrontare due serie di misure effettuate sugli stessi oggetti.

⁴⁷ Senza questo trasferimento, che si configura una risorsa determinante per sostenere la spesa sanitaria in tali regioni, il livello di tassazione raggiunge il valore minimo del campione.

⁴⁸ Se, infatti, le due tipologie di variabili risultassero coincidenti a priori, non si presenterebbe la necessità di valutare l'errore che si commette approssimando i valori teorici con quelli effettivi. Se, invece, si dovesse riscontrare una qualche differenza tra le due serie, avrebbe maggiormente senso analizzare lo scostamento relativo tra le due misure.

8.1 Analisi preliminare e test di Wilcoxon

Il *test di Wilcoxon* (Wilcoxon, 1945) è un test non parametrico, analogo al test *t-Student* (parametrico) per dati appaiati, che permette di verificare la significatività delle differenze tra le n coppie di due campioni⁴⁹. Tecnicamente, si effettua l'accostamento della misura teorica a quella effettiva per poi valutare la probabilità associata alla statistica test di accettare o rifiutare l'ipotesi nulla⁵⁰.

Il test di Wilcoxon è stato effettuato su due coppie di aliquote - teoriche ed effettive - diverse. Nel primo caso, si considerano le aliquote teoriche ottenute nell'ipotesi di *finanza autonoma* (Tavola 2, esclusa l'ultima riga) e le aliquote effettive (Tavola 3).

Nel secondo caso, si affiancano l'aliquota teorica nell'ipotesi di *finanza derivata* (Tavola 2, ultima riga) e le aliquote effettive (Tavola 3).

La Tavola 4 riporta il numero di osservazioni (**Obs**), il valore della statistica test (**z**) e la probabilità associata al valore di z (**Prob > |z|**).

Tavola 4: TEST DI WILCOXON

Test di Wilcoxon: Ipotesi nulla	Obs	z	Prob > z
$H_0: t_{teorica_finanza_autonoma} = t_{effettiva}$	16	-3,336	0,0008
$H_0: t_{teorica_finanza_derivata} = t_{effettiva}$	16	-3,516	0,0004

Fonte: Elaborazioni proprie su dati Banca d'Italia 2002

Dall'analisi di quest'ultima, si nota che i risultati di entrambi i test portano al rifiuto dell'ipotesi nulla: le differenze tra le due tipologie di aliquote sono sempre statisticamente significative⁵¹. In tal senso, è opportuno procedere al calcolo dell'errore che si commette approssimando i valori teorici con quelli effettivi nelle due ipotesi di finanziamento della sanità.

8.2 Risultati

La metodologia proposta prevede il calcolo delle seguenti espressioni di errore:

$$E_{FA} = (t^* - \hat{t})^2 + \sum_{i=1}^{15} \frac{N_i}{N} (t_i^* - \hat{t}_i)^2 \quad (20)$$

⁴⁹ Esso non richiede assunzioni circa la forma della distribuzione della misurazione e deve essere preferito allo *t-Student* test quando ci si trova in presenza di piccoli campioni (almeno per un numero di osservazioni inferiori a 50).

⁵⁰ Quest'ultima prevede che la mediana delle differenze tra coppie di dati sia pari a zero: $H_0: \delta = 0$. Accettare l'ipotesi nulla equivale a dire che le differenze tra tutte le coppie di valori x e y non sono statisticamente significative; rifiutare l'ipotesi nulla implica che le differenze siano statisticamente significative.

⁵¹ Il *p-value* associato al valore di z è nettamente inferiore alla soglia del 5%. Esso è addirittura inferiore al valore dell'1%.

$$E_{FD} = \sum_{i=1}^{15} \frac{N_i}{N} (t_d^* - \hat{t} - \hat{t}_i)^2 \quad (21)$$

Data l'ipotesi che in Italia è possibile individuare una tassazione effettiva regionale (\hat{t}_i) e una nazionale (\hat{t}) per la copertura del fabbisogno sanitario complessivo, la (20) misura la somma degli scarti al quadrato⁵² (considerando tutte le regioni con il rispettivo peso: N_i/N) tra i valori teorici (t^* e t_i^*) e quelli effettivi delle aliquote (\hat{t} e \hat{t}_i), nell'ipotesi di *finanza autonoma*.

La seconda espressione descrive l'errore che complessivamente si commette, ipotizzando che un unico decisore centrale fissi un'aliquota indifferenziata (t_d^*) su tutto il territorio per finanziare la spesa sanitaria a livello locale e centrale, rispetto al contesto reale in cui sono presenti entrambe le tipologie di prelievo (\hat{t} e \hat{t}_i).

L'errore più piccolo tra i due dovrebbe suggerire quale è la linea di interpretazione - fornita dal modello - più realistica per l'Italia riguardo al finanziamento della sanità.

Tecnicamente, nell'espressione (20) vengono inseriti i valori delle aliquote teoriche (Tavola 2) per le singole regioni (t_i^*) e per il governo centrale (t^*); per le aliquote effettive regionali (\hat{t}_i) e nazionale (\hat{t}), si considera la Tavola 3. La stessa metodologia viene applicata nel calcolo della (21), dove le aliquote effettive restano le stesse, mentre per il valore di t_d^* si ricorre alla Tavola 2. In entrambe le espressioni, si tiene conto del peso relativo di ogni regione, dato dal rapporto tra la dimensione della popolazione locale e quella complessiva (N_i/N), secondo i dati di Tavola A.1.

I risultati finali sono descritti nella Tavola 5.

Tavola 5: GLI ERRORI

FORMULA ERRORE	VALORE NUMERICO
$E_{FA} = (t^* - \hat{t})^2 + \sum_{i=1}^{15} \frac{N_i}{N} (t_i^* - \hat{t}_i)^2$	17,37
$E_{FD} = \sum_{i=1}^{15} \frac{N_i}{N} (t_d^* - \hat{t} - \hat{t}_i)^2$	5,70

Fonte: Elaborazioni proprie su dati Banca d'Italia 2002

⁵² Piuttosto che considerare lo scarto assoluto tra ogni coppia di valori, si preferisce prenderne il quadrato. Questo permette di avere una somma di soli numeri positivi, evitando che la presenza di valori negativi faccia ridurre la misura dell'errore nel caso che le aliquote effettive siano sempre superiori a quelle teoriche. In ogni caso, va detto che il risultato finale con il metodo degli scarti assoluti è del tutto analogo a quello ottenuto con il metodo dei quadrati, in genere preferito.

L'errore più piccolo è quello ottenuto sotto l'ipotesi di *finanza derivata* (E_{FD}): $5,70 < 17,37$. Questo risultato suggerisce che la situazione italiana (RSO) è più prossima ad uno schema in cui le decisioni per tutti i livelli di governo sono centralizzate, e si utilizza indirettamente il meccanismo del trasferimento di risorse alle unità sottostanti. In Italia, dunque, si è maggiormente vicini ad un meccanismo di *finanza derivata* – per il finanziamento della spesa sanitaria - benché la realtà istituzionale abbia recentemente fatto passi avanti verso un sistema di *finanza autonoma*, caratterizzato dalla crescente autonomia dei governi locali.

In generale, volendo dare una risposta ai quesiti in merito a quanto i risultati teorici si discostino dalla realtà e quale tra gli scenari del modello sia maggiormente rappresentativo per il contesto italiano, dall'analisi empirica si possono trarre due principali conclusioni.

Il *test di Wilcoxon* mostra che esiste una differenza statisticamente significativa tra i valori teorici e quelli effettivi. Questo risultato apre lo spazio ad un'ulteriore fase di indagine volta a capire quale, tra le situazioni proposte dal modello, sia la più “appropriata” per il caso italiano.

Di qui, la seconda conclusione. Dal calcolo degli errori, appare che la realtà italiana vada nella direzione in cui uno solo decide e lo fa per tutti i livelli, benché formalmente il quadro istituzionale si sia mosso verso il decentramento di alcune funzioni, in particolare quelle a maggior contenuto redistributivo (si pensi alla sanità qui trattata). Intuitivamente, i risultati ottenuti mostrano che in Italia il mancato “successo” di un sistema di *finanza autonoma* rispetto ad uno di *finanza derivata* è dovuto al fatto che la presenza degli standard sanitari nazionali può entrare in conflitto con i principi del federalismo, che prevedono il decentramento delle responsabilità per la definizione delle politiche sanitarie. In altri termini, in un contesto di governo decentrato, una volta ammessa l'esistenza di una forte propensione del centro a promuovere l'interesse nazionale – livelli minimi delle prestazioni sanitarie – è necessario trovare gli strumenti che consentano di raggiungere i risultati desiderati, evitando che la responsabilizzazione degli enti locali in sanità sia un pericolo per la garanzia degli standard nazionali. Se questi strumenti sono rappresentati, ad esempio, da quote di tributi erariali a favore delle Regioni (si pensi alla compartecipazione Iva) si verifica che la componente di finanziamento nazionale per la sanità continua ad essere forte, data appunto la natura di tributo “devoluto”, tipica della compartecipazione. Essa, infatti, è considerata dalle Regioni una risposta troppo debole alla loro richiesta di autonomia fiscale (Dirindin, 2001), dato che rappresenta un'imposta del governo centrale, rispetto alla quale i margini di effettivo intervento a livello locale sono praticamente nulli. In una certa misura, è normale dunque che continui a prevalere uno schema di finanziamento *derivato* piuttosto che *autonomo* per la funzione sanità in ragione dell'esistenza dei livelli minimi e dei relativi strumenti di prelievo utilizzati per garantirne la fornitura. Nonostante la volontà del legislatore di decentrare la responsabilità finanziaria della spesa sanitaria ad enti sub-centrali - in particolare, le Regioni - di fatto, si è mantenuto costante l'intervento e il controllo del

governo centrale, almeno per quanto riguarda i livelli di assistenza sanitaria fornita. Quello che in un certo senso è cambiato è la natura del finanziamento del SSN in quanto si è passati da trasferimenti specifici a compartecipazioni a tributi erariali senza vincoli di destinazione ma che, di fatto, servono a garantire che gli obiettivi del governo centrale vengano soddisfatti.

8.3 Analisi di robustezza

Data la diversità tra le misure per rappresentare il reddito disponibile lordo degli individui e la base imponibile (*Tavola A.2*), si è ritenuto opportuno procedere alla verifica di sensibilità dei risultati. Infatti, per le precedenti elaborazioni è stato utilizzato solo il dataset di Banca d'Italia⁵³, ma per rappresentare la base imponibile (Y_i e Y) è possibile utilizzare anche il dato Istat per il reddito disponibile lordo (*Tavola A.2, colonna B*), e quello relativo al PIL (stessa tavola, *colonna C*).

Dunque, le analisi di robustezza vengono condotte cambiando il dataset di riferimento, sul quale si calcolano di nuovo le variabili che determinano le aliquote teoriche ed effettive (reddito pro capite regionale e nazionale), fatta eccezione per i parametri α^* , $\bar{\beta}$ e $\bar{\gamma}_i$ che necessitano della distribuzione del reddito all'interno di ogni regione, ricostruibile solo attraverso i microdati di Banca d'Italia.

In Appendice, vengono replicati gli esercizi empirici per determinare tutte le aliquote, rispettivamente con il reddito disponibile lordo Istat e con il PIL (*Tavola A.3 e Tavola A.6*), per il calcolo del test di Wilcoxon (*Tavola A.4 e Tavola A.7*) e per la misura dell'errore (*Tavola A.5 e Tavola A.8*). Anche dopo queste verifiche, le conclusioni raggiunte in precedenza permangono.

Riguardo al test di Wilcoxon, in entrambi i casi si rifiuta l'ipotesi nulla con un livello di significatività inferiore al 5%; dal calcolo degli errori, emerge che la soluzione di *finanza derivata* produce nuovamente lo scarto minore rispetto al caso di *finanza autonoma*. In tal senso, anche con una diversa fonte di dati e, in particolare, con un diverso aggregato per rappresentare il reddito lordo, i risultati non subiscono variazioni. Si rafforza, dunque, la validità delle conclusioni raggiunte con il campione di Banca d'Italia.

9. Conclusioni

Con il modello teorico si analizza l'assegnazione della funzione sanità, ipotizzando due scenari istituzionali caratterizzati da una diversa autonomia di prelievo e di spesa da parte dei governi. La costruzione delle funzioni di benessere sociale prevede che in un caso governo centrale e governi locali decidano in modo autonomo e non coordinato la politica sanitaria da implementare (*finanza*

⁵³ Senza dubbio, era necessario partire dai microdati di Banca d'Italia per ricostruire la distribuzione del reddito in ogni Regione e calcolare i valori puntuali dei parametri $\bar{\beta}$ e $\bar{\gamma}_i$.

autonoma); nell'altro caso, si assume che un unico decisore centrale finanzi la spesa a tutti i livelli (*finanza derivata*). Partendo dalle soluzioni del modello, l'obiettivo è determinare se il finanziamento della spesa sanitaria in Italia (RSO) abbia caratteristiche istituzionali più simili al primo o al secondo scenario ipotizzato. In altri termini, si propone una verifica empirica dei risultati del modello, cercando una via di confronto con la situazione italiana.

Riguardo al quesito su quanto i risultati del modello teorico si discostino dalla realtà, il *test di Wilcoxon* suggerisce che, in generale, le variabili teoriche sono statisticamente diverse dalle variabili effettive.

In aggiunta, dal confronto tra i valori teorici e quelli effettivi è emerso che il contesto istituzionale italiano è più prossimo al modello di *finanza derivata*. Infatti, l'errore più piccolo (calcolato dal lato della tassazione) è quello in capo al decisore unico centrale. Dunque, l'evidenza empirica è più compatibile con uno schema derivato, dove le decisioni relative al finanziamento della sanità sono accentrate, "smentendo" potenzialmente il percorso di riforma istituzionale avviato in Italia nell'ultimo decennio e volto ad assegnare alle Regioni maggiori competenze e autonomia nella gestione di tale funzione. In tal senso, si potrebbe affermare che l'evidenza empirica si accosta maggiormente alle posizioni tradizionali sul federalismo fiscale (Musgrave, 1959; Oates, 1972) secondo cui non sia possibile né proficuo – per ragioni di efficienza e di equità - attuare il decentramento della funzione redistributiva, qui specificamente rappresentata dalla sanità. A tale proposito, sembra che non sia possibile conciliare l'interesse nazionale per il diritto ai livelli essenziali di assistenza sanitaria con il federalismo fiscale, specie se si "eliminano" i trasferimenti di tipo specifico o vincolato, indicati dalla teoria come lo strumento di finanziamento più adatto per perseguire obiettivi di interesse generale, quali appunto la definizione di priorità o di livelli minimi nella prestazione dei servizi da parte dei livelli inferiori di governo. In altri termini, fino a quando non si trova una soluzione adeguata per il finanziamento della spesa sanitaria e, più in generale, per il finanziamento delle spese legate alla fornitura dei livelli essenziali, è difficile pensare che un sistema di *finanza autonoma* possa emergere su uno di *finanza derivata*. In aggiunta, è altrettanto difficile pensare di poter coprire le esigenze derivanti da livelli di spesa garantiti costituzionalmente, negando la possibilità di far ricorso a trasferimenti specifici e introducendo, invece, la tipologia dei trasferimenti perequativi, che devono essere generali o senza vincolo di destinazione, quindi senza indicazioni circa la specificità della spesa da finanziare.

Appendice

Tavola A.1: POPOLAZIONE DOPO IL RIPROPORZIONAMENTO

REGIONI	NUMERO di INDIVIDUI (N_i)
Piemonte	4.222.300
Lombardia	9.071.100
Veneto	4.553.600
Liguria	1.571.100
Emilia Romagna	4.007.400
Toscana	3.506.700
Umbria	830.200
Marche	1.477.900
Lazio	5.131.400
Abruzzo	1.267.800
Molise	320.800
Campania	5.713.200
Puglia	4.021.700
Basilicata	597.100
Calabria	2.008.500
Italia (15 regioni)	48.300.800

Fonte: Elaborazioni proprie su microdati Banca d'Italia 2002 e Istat 2002

Tavola A.2: REDDITO DISPONIBILE LORDO CAMPIONARIO E REALE

(miliardi di euro)

REGIONI	(A) REDDITO DISPONIBILE LORDO TOTALE (Bdl)	(B) REDDITO DISPONIBILE LORDO TOTALE (Istat)	(C) PRODOTTO INTERNO LORDO (Istat)	(A) / (B) Valore percentuale (%)	(A) / (C) Valore percentuale (%)
Piemonte	57,42	74,93	105,21	77	55
Lombardia	152,7	172,64	270,65	88	56
Veneto	60,21	75,97	118,89	79	51
Liguria	24,68	27,67	36,05	89	68
Emilia R.	65,33	77,01	113,72	85	57
Toscana	50,04	60,06	87,29	83	57
Umbria	10,58	13,01	17,82	81	59
Marche	18,66	23,23	33,91	80	55
Lazio	65,39	87,82	137,18	74	48
Abruzzo	12,31	17,17	24,55	72	50
Molise	3,08	4,23	5,28	73	58
Campania	42,19	62,33	84,35	68	50
Puglia	39,02	45,28	60,17	86	65
Basilicata	4,76	6,83	9,39	70	51
Calabria	15,48	21,92	28,57	71	54
Media regionale	621,85	770,13	1.133,05	81	55

Fonte: Elaborazioni proprie su microdati Banca d'Italia 2002 e Istat 2002

Tavola A.3: ALIQUOTE TEORICHE ED EFFETTIVE (Reddito disponibile lordo, Istat)

(valori percentuali)

REGIONI	$t_i^* = g_i * \left(\frac{1}{y_i}\right)$ $t^* = g * \left(\frac{1}{y}\right)$ $t_d^* = 2g * \left(\frac{1}{y}\right)$	$\hat{t} = \frac{\left(SST - \sum_{i=1}^{15} T_i\right)}{Y}$ $\hat{t}_i = \left(\frac{T_i}{Y_i}\right)$
Piemonte	4,31	3,63
Lombardia	4,34	4,70
Veneto	4,94	4,41
Liguria	4,32	2,89
Emilia Romagna	4,28	3,97
Toscana	4,89	3,59
Umbria	5,24	2,85
Marche	4,82	3,49
Lazio	4,72	3,86
Abruzzo	5,97	2,82
Molise	6,36	1,02
Campania	7,54	2,27
Puglia	7,14	1,65
Basilicata	7,57	1,03
Calabria	7,77	0,78
Italia (governo centrale: t^*)	5,03	5,03
Italia (decisore centrale: t_d^*)	10,06	

Fonte: Elaborazioni proprie su dati Istat 2002

Tavola A.4: TEST DI WILCOXON (Reddito disponibile lordo, Istat)

Test di Wilcoxon: Ipotesi nulla	Obs	z	Prob > z
$H_0: t_{teorica_finanza_autonoma} = t_{effettiva}$	16	-3,336	0,0008
$H_0: t_{teorica_finanza_derivata} = t_{effettiva}$	16	-3,516	0,0004

Fonte: Elaborazioni proprie su dati Istat 2002

Tavola A.5: CALCOLO DEGLI ERRORI (Reddito disponibile lordo, Istat)

FORMULA ERRORE	VALORE NUMERICO
$E_{FA} = (t^* - \hat{t})^2 + \sum_{i=1}^{15} \frac{N_i}{N} (t_i^* - \hat{t}_i)^2$	9,33
$E_{FD} = \sum_{i=1}^{15} \frac{N_i}{N} (t_d^* - \hat{t} - \hat{t}_i)^2$	3,96

Fonte: Elaborazioni proprie su dati Istat 2002

Tavola A.6: ALIQUOTE TEORICHE ED EFFETTIVE (PIL, Istat)

(valori percentuali)

REGIONI	$t_i^* = g_i^* \left(\frac{1}{y_i} \right)$ $t^* = g^* \left(\frac{1}{y} \right)$ $t_d^* = 2g^* \left(\frac{1}{y} \right)$	$\hat{t} = \frac{\left(SST - \sum_{i=1}^{15} T_i \right)}{Y}$ $\hat{t}_i = \left(\frac{T_i}{Y_i} \right)$
Piemonte	3,07	2,59
Lombardia	2,77	3,00
Veneto	3,16	2,82
Liguria	3,32	2,22
Emilia Romagna	2,90	2,69
Toscana	3,36	2,47
Umbria	3,82	2,08
Marche	3,30	2,39
Lazio	3,02	2,47
Abruzzo	4,18	1,97
Molise	5,09	0,82
Campania	5,57	1,68
Puglia	5,37	1,24
Basilicata	5,50	0,75
Calabria	5,96	0,60
Italia (governo centrale: t^*)	3,42	3,42
Italia (decisore centrale: t_d^*)	6,84	

Fonte: Elaborazioni proprie su dati Istat 2002

Tavola A.7: TEST DI WILCOXON (PIL, Istat)

Test di Wilcoxon: Ipotesi nulla	Obs	z	Prob > z
$H_0: t_{teorica_finanza_autonoma} = t_{effettiva}$	16	-3,336	0,0008
$H_0: t_{teorica_finanza_derivata} = t_{effettiva}$	16	-3,517	0,0004

Fonte: Elaborazioni proprie su dati Istat 2002

Tavola A.8: CALCOLO DEGLI ERRORI (PIL, Istat)

FORMULA ERRORE	VALORE NUMERICO
$E_{FA} = (t^* - \hat{t})^2 + \sum_{i=1}^{15} \frac{N_i}{N} (t_i^* - \hat{t}_i)^2$	5,20
$E_{FD} = \sum_{i=1}^{15} \frac{N_i}{N} (t_d^* - \hat{t} - \hat{t}_i)^2$	1,71

Fonte: Elaborazioni proprie su dati Istat 2002

Riferimenti bibliografici

- Arzaghi, M., Henderson, J.V. (2005), "Why countries are fiscally decentralizing", *Journal of Public Economics*, vol. 89, pp. 1157-1189.
- Atkinson, A.B., Rainwater, L., Smeeding, T. (1995), *Income Distribution in OECD Countries: The Evidence from the Luxembourg Income Study (LIS)*, OECD, Parigi.
- Banca d'Italia (2004), "I bilanci delle famiglie italiane nell'anno 2002", in *Supplementi al Bollettino Statistico*, Anno XIV, n. 12, Marzo 2004.
- Banca d'Italia (2004), Archivio annuale, Indagine sul 2002, microdati.
- Barr, N. (2004), *The Economics of Welfare State*, 4^a edizione, Oxford University Press, Oxford.
- Brandolini, A., D'Alessio, G. (2001), "Household structure and income inequality", Working Paper n. 6, *Centre for Household, Income, Labour and Demographic Economics*.
- Brosio, G., Maggi, M., Piperno, S. (2003), *Governo e Finanza Locale*, Giappichelli, Torino.
- Buglione, E. (2005), "Un termometro del federalismo: decentramento e autonomia finanziaria negli anni 1996-2002", articolo presentato al convegno "Federalismo, equità e sviluppo", 6 Dicembre 2005, Roma.
- Cannari, L., D'Alessio, G. (2003), "La distribuzione del reddito e della ricchezza nelle regioni italiane", *Banca d'Italia*, Temi di Discussione, n. 482.
- Dirindin, N. (2001), "Federalismo fiscale e tutela della salute. Un percorso di responsabilizzazione delle regioni o il presupposto per cambiamenti strutturali?" in Dirindin et al., *Governare il federalismo. Le sfide per la sanità*, Il Pensiero Scientifico Editore, pp. 1-27.
- Faiella, I., Gambacorta, R. (2007), "The weighting process in the SHIW", *Banca d'Italia*, Temi di Discussione, n. 636.
- France, G. (2001), "Compatibilità fra il federalismo e gli standard sanitari nazionali: una sintesi delle esperienze estere e possibili lezioni per l'Italia", in Buglione et al., *Federalismo, regionalismo e standard sanitari nazionali. Quattro paesi, quattro approcci*, Giuffrè, Milano, pp. 169-197.
- Giarda, P. (2000), "Il federalismo fiscale in attuazione della legge n. 133/1999: aspetti tecnici, ragioni e problemi aperti", *Economia pubblica*, vol. 30, pp. 5-45.
- Goodspeed, T. J. (2000), "Tax Structure in a Federation", *Journal of Public Economics*, n. 75, pp. 493-506.
- Gottschalk, P., Smeeding, T. (1999), "Empirical Evidence on Income Inequality in Industrialized Countries", *Working Paper n. 154, Luxembourg Income Study*, pp. 1-76.
- Grazzini, L., Petretto, A. (2005), "Spesa pubblica per il welfare e crescita economica: una rassegna", *Working paper*, Dipartimento di Scienze economiche, Università di Firenze.
- ISTAT (2007), Conti economici regionali, Anni 2000 – 2005, Tavole.
- Liberati, P. (1999), *Il federalismo fiscale. Aspetti teorici e pratici*, Hoepli, Milano.
- Liberati, P. (2001), "Sanità e federalismo fiscale in Italia" in Buglione et al., *Federalismo, regionalismo e standard sanitari nazionali. Quattro paesi, quattro approcci*, Giuffrè, Milano, pp. 7-37.
- Marino, M. R., Rapallini, C. (2003), "La composizione familiare e l'imposta sul reddito delle persone fisiche: un'analisi degli effetti redistributivi e alcune considerazioni sul benessere sociale", *Banca d'Italia*, Temi di Discussione, n. 477.
- Messina, G. (2001), "Decentramento fiscale e perequazione regionale. Efficienza e redistribuzione nel nuovo sistema di finanziamento delle regioni a statuto ordinario", *Banca d'Italia*, Temi di Discussione, n. 416.
- Muraro, G. (2002), "Federalismo fiscale e sanità", *Working Paper SIEP*, Pavia.
- Musgrave, R.A. (1959), *The Theory of Public Finance*, McGraw-Hill, New York.
- Oates, W.E. (1972), *Fiscal Federalism*, Harcourt Brace Jovanovich, New York.

- Osberg, L., Schwabish, J., Smeeding, T. (2004), "Income Distribution and Public Social Expenditure: Theories, Effects and Evidence", in *Social Inequality*, a cura di Kathryn Neckerman, NY: Russell Sage Foundation.
- Pisauro, G. (2001), Intergovernmental Relations and Fiscal Discipline: Between Commons and Soft Budget Constraints, *IMF Working paper* n. 01/65.
- Pisauro, G. (2003), Fiscal decentralization and the budget process: a simple model of common pool and bailouts, *SIEP Working paper* n. 294, Pavia, Società Italiana di Economia Pubblica.
- Sacchi, A. (2008), "Il decentramento della funzione redistributiva: effetti sul benessere degli individui", di prossima pubblicazione in *Rivista di Diritto Finanziario e Scienza delle Finanze*
- Tresch, R.W. (2002), *Public Finance: A Normative Theory*, 2^a edizione, Elsevier Science, USA.
- Wilcoxon, F. (1945), "Individual Comparisons by Ranking Methods", *Biometrics*, vol. 1, pp. 80-83.