

L'utilizzo dei microdati individuali di fonte Inps per la stima degli oneri sociali e del costo del lavoro dipendente

Elisa Grand e Roberto Quaranta

Sommario

Le informazioni statistiche sul costo del lavoro in Italia sono ricavabili dai dati pubblicati dall'Istat e da quelli elaborati da alcune associazioni imprenditoriali. Entrambe le fonti offrono un grado di informazione rilevante, ma limitato. I dati prodotti dall'Istat consistono in indici o indicatori molto sintetici, in alcuni casi disaggregabili per area, settore e qualifica, elaborati a partire dalle informazioni d'impresa; i secondi sono circoscritti ai settori e ai contratti nazionali rappresentati dall'associazione, nonché alle imprese che vi aderiscono. In questo lavoro vengono presentati i risultati di una sperimentazione che ha lo scopo di creare una banca dati sul costo del lavoro basata su microdati di natura amministrativa. Le variabili di interesse sono stimate attraverso l'integrazione delle informazioni sulle retribuzioni lorde individuali fornite dall'Inps e l'implementazione delle regole contributive specifiche per ogni settore di attività, dimensione d'impresa, area geografica, tipologia di contratto: si tratta di una complessa procedura che consente di ottenere micro-stime corrette del costo del lavoro. Questo rappresenta, secondo la nostra opinione, un significativo contributo all'ampliamento della ricchezza informativa del panorama statistico nazionale in tema di costo del lavoro.

Parole chiave: costo del lavoro, dati amministrativi, variabili derivate.

Abstract

Statistical information on labor costs in Italy comes from official data published by Istat and from data collected by some entrepreneurial associations. Both of these sources provide relevant but limited information. The Istat data consist of synthetic indexes, disaggregated by region, sector and skill level, originating from firm-level data; the second data source is restricted to a few industries and reflects national agreements between the association and the unions. In this paper we present the results of an experiment which aims at the creation of a database on labor costs primarily based on administrative micro-data. The relevant variables are obtained by integrating the information on individual gross wages provided by Inps and applying the contributive rules specific to each industry, firm size, geographical location, contract typology, a complex procedure that yields reliable micro-estimate of labor costs. This is, in our opinion, a significant enrichment over the existing statistical information on labor costs in the Italian economy.

Keywords: labour cost, administrative data, derived variables.

1. Introduzione

La predisposizione di dati statistici provenienti da fonti amministrative ha ampliato le possibilità di svolgere analisi empiriche sul mercato del lavoro come gli studi sulle dinamiche retributive, sulla distribuzione dei salari e sulla mobilità occupazionale. Allo stesso tempo la potenziale natura longitudinale delle informazioni amministrative ha offerto l'opportunità di analizzare e osservare in modo continuo questi fattori nell'evoluzione delle carriere dei lavoratori e delle imprese.

Se all'inizio il trattamento e la diffusione delle informazioni statistiche avveniva prevalentemente al di fuori dei contesti istituzionali - si veda ad esempio la creazione delle banche dati che utilizzano fonti di origine previdenziale come il *Work Histories Italian Panel* (WHIP), il *Veneto Worker Histories* (VWH), il panel Isfol-Inps e il panel sviluppato dalla fondazione Rodolfo Debenedetti, tutte basate sul *matching* di archivi gestionali dell'Inps relativi ad individui ed imprese - negli ultimi dieci anni anche istituzioni come l'Istat hanno avviato procedure per il trattamento e l'utilizzo di dati amministrativi allo scopo di produrre dati statistici, spesso ottenuti attraverso l'integrazione con dati provenienti da fonti d'indagine più tradizionali.

Le potenzialità offerte da queste banche dati risentono però della mancanza di altre informazioni importanti come la retribuzione netta e il costo del lavoro, assenti nelle fonti individuali da cui traggono origine. In particolare, gli unici dati disponibili per la determinazione del costo del lavoro dipendente, infatti, sono quelli estrapolabili dalle dichiarazioni contributive delle imprese trasmesse all'Inps (DM10): da queste è possibile risalire agli importi totali versati dall'impresa, escludendo tuttavia ogni possibilità di ricondurre i versamenti alle caratteristiche dei singoli occupati o dei rapporti di lavoro. Uno dei più recenti prodotti Istat, la rilevazione OROS (Occupazione, Retribuzioni e Oneri Sociali), è basata appunto sulle informazioni tratte dalle dichiarazioni contributive delle imprese, ma mette a disposizione soltanto numeri indice - dati per loro natura molto sintetici - sulle retribuzioni lorde, sugli oneri sociali e sul costo del lavoro disaggregabili solo per settore o per macro-settore.

La determinazione del costo del lavoro relativo al singolo occupato o allo specifico rapporto di lavoro risulta tuttavia possibile attraverso l'integrazione dei microdati individuali con le informazioni di impresa e l'implementazione delle norme che regolano la materia contributiva.

Questo lavoro presenta i risultati di una sperimentazione che ha lo scopo di valutare come la creazione di una banca dati sul costo del lavoro fondata principalmente su microdati individuali di natura amministrativa possa, da un lato, garantire una stima corretta della variabile di interesse e, dall'altro, determinare un notevole ampliamento della ricchezza informativa delle statistiche già disponibili. I risultati della sperimentazione mostrano, infatti, che le stime prodotte sono, a livello aggregato, coerenti con quelle provenienti dalle fonti statistiche ufficiali, risultando tuttavia disaggregabili secondo le caratteristiche dell'impresa, dell'individuo e del rapporto di lavoro.

Il lavoro è strutturato come segue: il primo paragrafo contiene una breve rassegna dei dati sul costo del lavoro attualmente disponibili in Italia; nel secondo e nel terzo paragrafo si descrivono i dati utilizzati dalla suddetta sperimentazione e si descrivono le modalità di imputazione delle voci che costituiscono il costo del lavoro dipendente; nel quarto si mostra la coerenza dei risultati prodotti rispetto alle fonti ufficiali al momento disponibili. Infine, nell'ultimo paragrafo si presentano, attraverso alcune statistiche, le potenzialità informative di una banca dati basata su microdati individuali e di impresa.

2. La disponibilità di dati sul costo del lavoro in Italia

In Italia le statistiche disponibili sul costo del lavoro prodotte e diffuse dall'Istat hanno caratteristiche diverse a seconda dei bisogni informativi da soddisfare e della fonte utilizzata - amministrativa e/o *survey* - ma risultano comunque strettamente legate alle caratteristiche delle informazioni provenienti dalle imprese.

L'indagine panel "Lavoro e retribuzioni nelle grandi imprese" (GI) raccoglie informazioni su tutte le imprese operanti nei settori privati non agricoli, a esclusione di quello relativo ai servizi sociali e personali - risultano dunque coinvolti i settori C-K della classificazione economica Ateco 2002 -, e aventi almeno 500 addetti nella media dell'anno base. Le statistiche da essa derivate sono gli indici del costo del lavoro e i valori medi del costo del lavoro mensile pro-capite riferiti all'anno base e disaggregabili per settore, sezioni e qualifica.

La struttura del costo del lavoro in Italia (LCS) è una rilevazione quadriennale che segue le disposizioni del Consiglio Europeo, Regolamento n. 530/1999; è la realizzazione italiana della *Labour Cost Survey* dell'Eurostat. Essa è condotta sulle imprese con almeno 10 addetti, attive nei settori dell'industria e dei servizi privati. Tra le informazioni rilevate compaiono, per tutte le figure professionali a esclusione degli apprendisti, quelle relative alla retribuzione lorda e al costo del lavoro dipendente, anche nella sua accezione più ampia¹, disaggregate per settore di attività, classe di addetti e area in base alla localizzazione dell'attività produttiva delle imprese.

Sempre per il settore privato, ma con l'esclusione del settore finanziario, è possibile ottenere informazioni sui costi del personale consultando i dati dei Conti economici delle imprese, disaggregati su base annua per regione e per settore, a loro volta ottenuti da due *survey* basate sulle imprese: la Rilevazione campionaria Istat sulle piccole e medie imprese industriali e dei servizi e la Rilevazione sul sistema dei conti delle imprese con almeno 100 addetti (SCI).

Una delle più recenti fonti di informazioni sul costo del lavoro dipendente è l'indagine OROS, che si rivolge a tutte le imprese private con dipendenti dei settori industriale e dei servizi (settori da C a K della classificazione Ateco 2002); la sua particolarità è quella di integrare dati amministrativi di fonte Inps con informazioni tratte dall'indagine mensile dell'Istat sul lavoro nelle grandi imprese (GI). OROS produce i numeri indice, trimestrali e stagionalizzati, sulle retribuzioni lorde, oneri sociali e costo del lavoro. La fonte primaria utilizzata è il modulo che le imprese compilano mensilmente per le denunce contributive dei dipendenti iscritti a libro paga (DM10).

Alcune informazioni sul costo del lavoro, sempre ottenute dall'integrazione di fonti di natura diversa, possono essere desunte dai dati della contabilità nazionale. Tali dati, prodotti in ottemperanza ai principi adottati dall'Unione Europea nel Sistema dei conti economici nazionali e regionali (Istat, 2004), nonostante il profondo rinnovamento delle metodologie di stima, rimangono comunque strumentali agli obiettivi propri della contabilità nazionale. I dati diffusi dalla contabilità nazionale riguardano il complesso dei redditi da lavoro dipendente, delle retribuzioni lorde erogate dalle imprese e degli oneri sociali a loro carico. Le stime prodotte riguardano sia la quota di lavoro dipendente regolare sia quella non regolare (economia sommersa): per quest'ultima il dato si riferisce alla sola retribuzione lorda, essendo gli oneri sociali omessi per definizione.

Infine, si annoverano nel panorama statistico nazionale le serie statistiche, dal considerevole potenziale informativo, messe a disposizione dalle associazioni imprenditoriali (si veda su tutte l'indagine di Federmeccanica sul costo del lavoro e le retribuzioni), la cui elaborazione non viene effettuata in modo sistematico da tutte le associazioni ed è circoscritta ai settori e ai contratti nazionali delle aziende afferenti alle singole associazioni.

3. Le potenzialità dei dati Inps per la stima del costo del lavoro dipendente

Nella definizione del costo del lavoro dipendente rientrano gli importi delle retribuzioni lorde degli oneri sociali obbligatori² e degli accantonamenti di fine rapporto (Tfr).

Gli archivi gestionali dell'Inps risultano idonei all'identificazione delle varie voci che costituiscono il costo del lavoro: i dati utili per la determinazione del costo del lavoro sono principalmente quelli tratti dai modelli DM10 (le denunce contributive che le imprese effettuano

¹ Nel costo del lavoro in senso ampio alla retribuzione lorda si sommano i contributi obbligatori, le spese della formazione ed i costi intermedi connessi all'attività produttiva. Rispetto alla definizione in senso stretto (retribuzione e contributi obbligatori) la differenza è nell'ordine di un punto percentuale (dati LCS, rilevazione anno 2000).

² Il complesso dei contributi versati agli enti di previdenza ed assistenza sociale (i contributi Inps e Inail).

mensilmente per i propri dipendenti) e quelli tratti dai modelli O1M³ (le denunce contributive annuali che il datore di lavoro compila per ogni lavoratore avuto alle dipendenze).

La prima fonte menzionata è la più completa: è questa la fonte da tempo sfruttata dall'Istat per la creazione della banca dati OROS. Nel modello DM10 sono rilevate tutte le informazioni di carattere amministrativo necessarie per la determinazione dei debiti e dei crediti contributivi dell'impresa nei confronti dell'Inps⁴. Gli importi relativi al Tfr e ai contributi dovuti all'Inail, a differenza di quelli relativi ai contributi da versare all'Inps, non sono rilevati nei DM10 e devono quindi essere stimati. In particolare in OROS il premio Inail è stimato da fonti esterne applicando delle aliquote medie per gruppo di attività economica pubblicate sul sito internet dell'istituto e utilizzando i dati desunti dall'Indagine Costo del lavoro e retribuzioni nette su base contrattuale, effettuata solo per l'anno 1995⁵. Un maggior livello di dettaglio nella disaggregazione delle statistiche, rispetto a quanto offrono l'archivio DM10 e i risultati attualmente pubblicati da OROS, potrebbe essere raggiunto integrando i dati provenienti dalle imprese, quelli che l'Istat utilizza attualmente per OROS, con quelli dei singoli lavoratori, producendo delle informazioni basate congiuntamente sulle caratteristiche dell'impresa, degli occupati e del rapporto di lavoro. L'utilizzo di microdati individuali consentirebbe anche di produrre indicatori statistici disaggregati territorialmente, superando un limite legato al solo utilizzo di fonti di impresa⁶: a causa di regole amministrative proprie dell'Inps, infatti, un'impresa plurilocalizzata può chiedere che le sia concesso l'accantonamento contributivo (ovvero la facoltà di versare i contributi presso una sola sede Inps), facendo sì che dai modelli DM10 dell'Inps risulti una distribuzione geografica dei dipendenti diversa da quella reale. Di questo problema non risente invece l'informazione derivante dalle denunce contributive individuali - né vi sono evidenze che confermino il contrario - che rilevano con uno specifico campo la zona in cui il rapporto di lavoro viene effettivamente svolto.

La seconda fonte utile per la determinazione del costo del lavoro è costituita, appunto, dai microdati individuali di cui sopra: essi sono tratti dalle denunce contributive annuali che il datore di lavoro compila per ogni lavoratore avuto alle dipendenze (O1M). Come i DM10, che registrano i dati mensili riferiti all'impresa, i modelli O1M registrano i dati annuali riferiti agli individui. Le due denunce riportano, per loro natura, informazioni corrispondenti, sono entrambe articolate in quadri ognuno caratterizzato da una specifica finalità espositiva, e sono relazionabili attraverso i codici identificativi di impresa. A differenza dei DM10 i moduli O1M non contengono le somme versate per il pagamento dei contributi pensionistici ed assistenziali, ma riportano quasi tutte le informazioni utili ai fini della loro determinazione⁷: i dati anagrafici del lavoratore (età, genere), le caratteristiche del rapporto di lavoro (la retribuzione lorda, l'elenco dei mesi, il numero di settimane e di giorni per cui il lavoratore ha ricevuto una retribuzione dal datore di lavoro, la provincia di lavoro, la qualifica, il tipo di contratto), le voci che indicano tipologie particolari di retribuzione o dati di particolari categorie di lavoratori (come ad esempio l'indennità sostitutiva del preavviso, eventuali versamenti a fondi complementari o a forme sostitutive dell'Inps, decontribuzioni previste per il lavoro straordinario o oneri dovuti al superamento del massimale contributivo).

Per poter stimare il costo del lavoro dipendente sui microdati individuali risulta quindi necessario integrarvi i dati di impresa (DM10) e implementare le regole contributive e quelle in materia di sgravi e fiscalizzazioni. L'informazione di impresa consente di determinare il profilo contributivo che l'Inps ha attribuito alle aziende (in particolare il codice statistico contributivo e i codici di autorizzazione⁸): da detto profilo dipendono la corretta applicazione delle aliquote per il calcolo degli oneri sociali obbligatori relativi ai rapporti di lavoro e la possibilità di identificare le imprese che hanno beneficiato di agevolazioni contributive (sgravi e fiscalizzazioni).

³ Negli anni le denunce contributive individuali sono state oggetto di modifiche, una delle quali ha riguardato il nome di riferimento del modello (O1M all'inizio, successivamente Modello CUD, poi 770, infine Emens), in questo lavoro è stato deciso di utilizzare il nome iniziale.

⁴ Per una dettagliata analisi delle informazioni dei campi del modulo DM10 si rimanda a Baldi et al., 2001.

⁵ Baldi et al., 2008.

⁶ Pacini, 2009.

⁷ Per maggiori dettagli sulle informazioni presenti in questo modulo si rimanda a Contini et al., 1994.

⁸ I codici di autorizzazione hanno la funzione di integrare il codice statistico contributivo (CSC) allo scopo di distinguere i datori di lavoro che presentano caratteristiche diverse da quelle tipiche del settore, della classe e della categoria d'appartenenza.

L'implementazione delle regole normative consente poi di tradurre, per ogni rapporto di lavoro afferente all'impresa, le informazioni sulle aliquote e sui regimi di agevolazione in veri e propri importi.

4. La sperimentazione: la stima del costo del lavoro dipendente sui microdati WHIP

La sperimentazione è stata condotta sugli archivi dei rapporti di lavoro dipendente (RL) della banca dati WHIP, prodotta dal Laboratorio R. Revelli⁹. WHIP è una banca dati longitudinale, basata su vari archivi gestionali dell'Inps dai quali è estratto un campione rappresentativo di individui: di questi sono osservati gli episodi lavorativi nel periodo compreso tra il 1985 e il 2004. La popolazione di riferimento è composta da tutti i lavoratori dipendenti occupati nelle imprese private nei settori di attività delle sezioni da C a K della classificazione Ateco 91 (dai settori dell'industria ai servizi orientati al mercato). Il rapporto teorico tra la dimensione campionaria e quella della popolazione di riferimento è di circa 1:90.

Gli archivi utilizzati sono strutturati per anno e contengono, per ogni individuo, un record per ciascun rapporto di lavoro dipendente attivo in qualunque momento dell'anno. Tra le informazioni riportate, quelle su cui l'imputazione del costo del lavoro si è fondata comprendono sia caratteristiche del lavoratore e del rapporto di lavoro, sia caratteristiche dell'impresa; le prime provengono dai modelli OIM, mentre le ultime sono il frutto di elaborazioni dei moduli DM10.

Gli RL non offrono tutto il potenziale informativo di cui si è trattato al paragrafo precedente: limitati sono, infatti, da una parte, i contenuti forniti dall'Inps, mentre dall'altra alcune scelte di costruzione di WHIP contribuiscono a ridurre ulteriormente l'informazione finale.

La lacuna informativa di cui al primo punto riguarda prevalentemente le informazioni di impresa: l'Inps non fornisce, infatti, il contenuto dei DM10, bensì i dati dell'Osservatorio delle imprese (Osservatorio) con codici che consentono l'identificazione longitudinale delle imprese¹⁰ e di conseguenza anche l'identificazione longitudinale dei rapporti di lavoro¹¹. Nell'Osservatorio non sono disponibili le informazioni relative a tutte le posizioni contributive dell'impresa¹², dalle quali, attraverso il CSC, fornito, e il codice di autorizzazione, non fornito, sarebbe possibile determinare in modo corretto profilo contributivo, crediti e debiti delle imprese nei confronti dell'Inps. Le informazioni disponibili per ogni impresa sono quelle relative alla sola posizione principale, scelta dall'Inps: questo comporta, come si vedrà nel prosieguo del lavoro, una certa distorsione, il cui grado non è possibile valutare, nelle stime che riguardano le imprese pluriposizionate¹³. Inoltre la fornitura di dati da parte dell'Inps è attualmente ferma al 2004 per quanto riguarda i dati individuali e al 2002 per i dati di impresa.

Relativamente al secondo punto, le scelte di costruzione di WHIP hanno ripercussioni sulla struttura informativa disponibile per quanto riguarda le imprese, i lavoratori dipendenti e i rapporti

⁹ Per ulteriori dettagli si rimanda al sito www.laboratoriorevelli.it/whip.

¹⁰ La procedura di identificazione longitudinale delle imprese riveste un ruolo di primaria importanza non solo nella produzione di dati di impresa coerenti, ma anche nella ricostruzione dei rapporti di lavoro dipendente. La necessità di mettere in atto una simile procedura è legata all'ampiezza della dimensione longitudinale di WHIP. WHIP è infatti costruito sulla base di forniture di un campione tratto dalle diverse, successive, edizioni dell'Osservatorio delle Imprese, che ha un orizzonte temporale di soli cinque anni. Tra le successive edizioni dell'Osservatorio i codici identificativi delle imprese possono cambiare, così come può cambiare la composizione stessa delle imprese per scoperti, fusioni e altre trasformazioni societarie: da qui la necessità di procedere all'identificazione longitudinale delle imprese all'interno dei vari rilasci dell'Osservatorio.

¹¹ Gli OIM forniti dall'Inps sono organizzati per dichiarazioni contributive prodotte a favore di un lavoratore. Non esiste però una relazione uno a uno tra gli episodi di contribuzione dichiarati e rapporti di lavoro instaurati dalla persona, così come sono carenti le informazioni sull'inizio e la fine del rapporto di lavoro, oggetto principale dell'identificazione longitudinale nella procedura di costruzione di WHIP (cfr. documentazione sul sito WHIP).

¹² Le informazioni di impresa elaborabili riguardano le caratteristiche in corrispondenza della posizione contributiva prevalente, scelta dall'Inps nella pubblicazione dell'Osservatorio: nel caso in cui una stessa impresa abbia più posizioni contributive per attività classificate in settori diversi questo può essere fonte di distorsione nell'applicazione delle aliquote contributive. La posizione assicurativa di un'impresa è determinata dal tipo di profilo contributivo assegnatole. Generalmente l'Inps assegna una sola posizione. In altri casi l'Inps assegna più posizioni contributive sia per esigenze dell'impresa sia per esigenze di norme o di legge. Per ulteriori dettagli si rimanda a Leombruni R. e Quaranta R., 2011.

¹³ Mentre nella realtà esse possono svolgere attività diverse, ognuna afferente a una diversa posizione previdenziale presso l'Inps, ognuna con un proprio settore di attività, un proprio CSC e un proprio inquadramento contributivo, la disponibilità di informazioni su di un'unica posizione obbliga a considerare quelle informazioni come rappresentative di tutta l'impresa, con una distorsione nelle stime tanto più rilevante quanto più le diverse attività svolte dalla stessa impresa si discostano in termini di inquadramento contributivo.

di lavoro. In particolare la costruzione di WHIP determina l'insorgenza di due criticità. La prima è collegata alla struttura degli RL: essendo questi archivi costruiti in modo da avere un solo record per ciascun rapporto attivo nell'anno, l'informazione relativa a un rapporto di lavoro che nel corso dell'anno modifica, senza interrompersi, una o più caratteristiche, come la qualifica (che può ad esempio cambiare da operaio a impiegato), o la durata dell'orario di lavoro (che può ad esempio passare da tempo parziale a tempo pieno), è quella relativa all'ultima modalità osservata in ordine cronologico; le informazioni riguardanti la retribuzione, il trattamento di fine rapporto, le giornate e le settimane lavorate vengono invece consolidate¹⁴. La seconda è relativa alla mancata elaborazione di alcune informazioni non strettamente legate all'identificazione dei rapporti di lavoro o alla produzione degli archivi RL: i dati non contengono, infatti, informazioni su particolari tipologie di versamenti o di dati utili per stimare correttamente alcune variabili di interesse (versamenti ad altri enti, quota sopra il massimale, quota di decontribuzione per il lavoro straordinario, integrazioni a carico del datore di lavoro, numero di settimane per le quali sono state erogate tali integrazioni).

Nonostante i limiti di cui sopra, l'utilizzo di WHIP offre l'indubbio vantaggio di poter disporre di una base dati statistica già sottoposta a procedure di trasformazione, normalizzazione e pulizia dei dati amministrativi Inps: le unità di analisi - come l'individuo, il rapporto di lavoro, l'impresa - e le informazioni ad esse collegate hanno subito un processo di identificazione e ricostruzione longitudinale che facilita l'applicazione delle procedure di imputazione degli oneri sociali, soprattutto in presenza di rapporti di lavoro beneficiari di sgravi e di fiscalizzazioni¹⁵.

4.1 L'imputazione del costo del lavoro dipendente

La stima è stata predisposta su tutti i rapporti di lavoro attivi nelle sezioni da C a K della classificazione Ateco 91, la cui qualifica registrata di apprendista, di operaio o di impiegato. La procedura di imputazione del costo del lavoro ai singoli episodi di lavoro è stata sviluppata per essere implementata sul periodo 1985 - 2008, ma per questioni diverse si è deciso di limitarla agli anni dal 1996 al 2002. Questa scelta non è casuale, ma guidata dai seguenti fattori:

- la difficoltà di reperire dati confrontabili per valutare le stime prodotte negli anni precedenti al 1996, mentre negli anni prescelti le stime sono comparabili con quelle ufficiali (OROS e LCS), e in alcuni casi la fonte utilizzata dalla sperimentazione è simile alla fonte ufficiale (OROS).
- la mancanza di osservazioni oltre il 2004;
- la mancanza per gli anni 2003 e 2004 di alcune informazioni basilari sulle imprese.

Come accennato sopra, il processo di stima mira a determinare l'ammontare degli oneri sociali obbligatori che, unito alla retribuzione lorda, già disponibile nei dati, costituisce il costo del lavoro.

La determinazione degli oneri sociali a carico delle imprese passa attraverso il calcolo di quelli teorici, dai quali si ottengono, sottraendo gli importi legati a eventuali agevolazioni contributive di vario genere, quelli effettivi.

La retribuzione lorda è la base da cui partire per determinare i singoli elementi che compongono il costo del lavoro: essa comprende, secondo la definizione, tutto quanto è stato erogato in forma di retribuzione, sia mensilmente, sia come mensilità aggiuntive, arretrati, premi e gratifiche, anche se concessi come una tantum. Inoltre la retribuzione da assumere come base per il calcolo dei contributi del dipendente non può essere inferiore ad un certo minimale giornaliero¹⁶: in caso contrario i contributi vengono calcolati sulla retribuzione minima; i lavoratori ai quali si applica il sistema contributivo sono invece soggetti a un limite di retribuzione oltre il quale i contributi pensionistici (IVS) non sono più dovuti, mentre i contributi minori (relativi a tutte le altre forme assicurative diverse da IVS) sono sempre dovuti sulla intera retribuzione imponibile¹⁷.

¹⁴ Leombruni et al., 2010.

¹⁵ In alcuni casi per imputare ai rapporti di lavoro gli importi riferiti a sgravi e fiscalizzazioni è necessario almeno, in mancanza dei codici di impresa specifici, essere a conoscenza di quando essi hanno avuto inizio (le informazioni di inizio e fine del rapporto di lavoro, ottenute attraverso il processo di stima elaborato nella fase di identificazione longitudinale dei rapporti di lavoro, sono presenti nei dati WHIP).

¹⁶ Il minimale giornaliero è stabilito dall'Inps all'inizio di ogni anno, ed è una percentuale dell'importo del trattamento minimo di pensione in vigore al primo gennaio (art. 7, Legge 11 novembre 1983, n. 638 e successive modificazioni).

¹⁷ Circ. Inps n. 177 del 7 settembre 1996.

La retribuzione contenuta negli RL rappresenta la retribuzione lorda, ovvero quella che per legge¹⁸ è da assoggettare a contribuzione¹⁹, dalle quale attraverso passi successivi è possibile determinare gli importi relativi al Tfr e agli oneri sociali, teorici ed effettivi.

Gli oneri sociali sono gli importi da versare all'Inps e all'Inail: essi sono calcolati applicando delle aliquote specifiche alla retribuzione lorda, computando al costo del lavoro la sola quota a carico del datore di lavoro, dal momento che quella a carico del lavoratore è già inclusa nella retribuzione lorda.

La stima del livello teorico degli oneri per ogni rapporto di lavoro è facilmente ottenibile; in WHIP sono già disponibili, infatti, le informazioni necessarie al calcolo: la retribuzione lorda annuale, il codice statistico contributivo (CSC), la quantità di lavoro prestata (settimane e giornate lavorate), la dimensione dell'impresa (quella media dell'anno), la qualifica del lavoratore e il tipo di contratto applicato (apprendistato, contratto di formazione lavoro, reinserimento...). Per giungere alla stima degli oneri sociali teorici è stato quindi sufficiente reperire e applicare le informazioni sulle aliquote contributive, previdenziali, assistenziali e assicurative.

Vediamo nel dettaglio le varie voci che riguardano gli oneri sociali e la loro strategia di applicazione. L'identificazione delle corrette aliquote per determinare gli importi da versare all'Inps avviene sulla base della natura dell'attività (CSC), della dimensione dell'azienda, della qualifica del lavoratore e del periodo considerato. Si richiama brevemente quanto introdotto al paragrafo precedente riguardo alla distorsione nella stima di oneri sociali e costo del lavoro per le imprese pluriposizionate: come accennato, i dati di impresa forniti dall'Inps per la costruzione di WHIP sono relativi ad una sola posizione contributiva, quella prevalente dell'azienda (la scelta della posizione da considerare prevalente è di competenza dell'Inps e avviene secondo meccanismi non noti). Tale scelta fa sì che un'impresa che svolge attività economiche diverse sia associata, nei dati WHIP, a un solo CSC: a tutti gli episodi di lavoro svolti presso quell'impresa, a parità di qualifica, sarà quindi imputata la stessa aliquota contributiva. Questo crea una distorsione, il cui livello e la cui direzione non siamo in grado di valutare, nella stima degli oneri sociali dovuti all'Inps, tanto maggiore quanto maggiore è la discrepanza tra le aliquote proprie dei diversi CSC associati alle varie posizioni contributive della stessa impresa. Si veda a titolo esemplificativo la Tavola 1 riportata sotto: si consideri un'impresa che svolge attività contraddistinte da due diversi CSC, entrambi del ramo industriale (1). Nel caso in cui l'attività prevalente, scelta dall'Inps, sia quella associata al CSC 1.01.01, mentre il codice proprio della seconda attività svolta dall'impresa sia il CSC 1.02.01, ai lavoratori che operano nella seconda attività verrà applicata un'aliquota contributiva Inps inferiore di quasi due punti.

La distorsione nella stima degli oneri sociali si rifletterà, anche se in misura minore, sulla stima del costo del lavoro. Più in generale, è corretto affermare che risulteranno affette da certa distorsione tutte le stime che si basano sull'individuazione delle imprese attraverso il CSC: anche l'attribuzione di sgravi e fiscalizzazioni, che nella maggior parte dei casi passa attraverso l'identificazione delle imprese beneficiarie per mezzo del CSC, non può dirsi certa per le imprese pluriposizionate.

Una ulteriore potenziale fonte di distorsione nella stima degli oneri sociali è costituita dalla non assoluta esaustività delle aliquote che sono state raccolte: pur individuando ben 350 profili contributivi²⁰ diversi e rilevanti (CSC o gruppi di CSC), come quelli mostrati nella tavola 1, le aliquote raccolte non coprono tutti i codici possibili. Nei casi in cui non è stato possibile reperire la corretta aliquota da applicare il CSC orfano è stato associato a un profilo contributivo adiacente nella classificazione, oppure è stata applicata un'aliquota media di settore²¹.

¹⁸ Articolo 12 della Legge n. 153/69, come sostituito dall'articolo 6 del D. Lgs. n. 314/1997. Sempre per legge, la retribuzione lorda assoggettata a contribuzione fa anche riferimento alla definizione di reddito da lavoro dipendente valida ai fini fiscali.

¹⁹ Da essa sono escluse tutte le integrazioni salariali erogate dall'Inps o da altri enti (CIG, indennità di malattia e maternità, assegni familiari, infortuni coperti dall'Inail), anche quando sono anticipate dal datore di lavoro.

²⁰ Dati da CSC, anno di validità, qualifica e dimensione d'impresa. Considerando l'intero periodo di copertura WHIP i profili contributivi individuati risulterebbero 900.

²¹ L'imputazione di un'aliquota media riguarda, in realtà, solo 5 CSC del ramo 7 (commercio, arti, professioni), che risultano associati a circa lo 0,7% dei rapporti di lavoro sui quali si applica il processo di stima. Agli episodi che si svolgono in imprese contrassegnate da uno di questi CSC è stata imputata l'aliquota media per qualifica del ramo 7. L'associazione di un CSC orfano a un profilo contributivo adiacente nella

Tavola 1 – Alcuni esempi di profili contributivi applicati nella sperimentazione su dati WHIP

Variabili di aggancio				Aliquote per l'imputazione degli oneri sociali			
Anno	CSC	Dimensione	Qualifica	Inps	FPLD	Altre Inps	SSN
1997	1.01.01	16-50	Operaio	33.66	22.89	10.78	9.60
1997	1.02.01	16-50	Operaio	35.47	22.89	12.58	9.60
1997	4.01.01	Tutte	Operaio	29.84	22.89	6.95	9.60
1997	7.04.01	<=50	Operaio	30.87	22.89	7.98	9.60
1997	7.07.01	Tutte	Operaio	30.87	22.89	7.98	9.60

Note: FPLD = contributo per il Fondo pensioni lavoratori dipendenti. Altre Inps = Contributo per l'assicurazione obbligatoria contro la disoccupazione involontaria (DS), contributo per il finanziamento degli assegni per l'assistenza agli orfani dei lavoratori italiani (ENAOI), contributo per la gestione case per lavoratori (GESCAL), contributo per il finanziamento dell'assegno per il nucleo familiare (CUAF), contributo assistenza malattia pensionati (Add. Pens.), contributo asili nido (Asili nido), contributo per l'indennità economica di malattia (Malattia), contributo per l'indennità economica di maternità (Maternità), contributo per l'assicurazione contro la tubercolosi (TBC), contributo per la cassa integrazione guadagni ordinaria (CIG), contributo per la cassa integrazione straordinaria (CIGS), contributo per il finanziamento dell'indennità di mobilità (Mobilità), contributo per il finanziamento del fondo di garanzia per il trattamento di fine rapporto (Tfr). SSN = Contributo per il Servizio Sanitario Nazionale.

Un'altra voce da imputare è il contributo al Servizio sanitario nazionale (SSN), la cui aliquota si applica sulla retribuzione lorda come per i versamenti dovuti all'Inps. Abolito alla fine del 1997 insieme ad altre imposte minori²², è sostituito dal 1° gennaio 1998 dall'Imposta Regionale sulle Attività Produttive (IRAP)²³.

Il premio Inail è stato determinato applicando delle aliquote medie effettive distinte per settore Ateco 91 alla retribuzione di ogni rapporto di lavoro. Le aliquote sono state calcolate, sui dati reperibili sul portale dell'istituto per i soli anni disponibili (dal 2002 al 2004), come rapporto tra il totale dei premi pagati dalle imprese e il monte retributivo dichiarato. L'attuale impossibilità di ottenere un dato più disaggregato²⁴, almeno per qualifica, e per un periodo più ampio, che permetta di tenere conto dell'andamento del premio, dovrebbe comportare distorsioni minime nelle stime aggregate, mentre è probabile che si riscontrino, per disaggregazioni approfondite, valori del premio Inail sottostimati per gli operai e sovrastimati per gli impiegati.

Per la stima degli oneri sociali è stata applicata solo la regola del minimale, mentre non è stato possibile applicare quella del massimale, così come altre regole (decontribuzione, lavoro straordinario, contributi figurativi) che non riguardano la totalità dei rapporti di lavoro, per mancanza delle variabili identificatrici dei casi specifici. Questa mancata identificazione dovrebbe comportare un effetto minimo (ad esempio, il massimale si applica solo ai lavoratori del regime contributivo, e in ogni caso nella sperimentazione non sono presenti i dirigenti) di sovrastima degli oneri sociali obbligatori. Il minimale è stato applicato confrontando la retribuzione giornaliera lorda ottenuta in WHIP con la retribuzione minima giornaliera stabilita per legge: le retribuzioni giornaliere inferiori al minimale sono quindi state sostituite, per il calcolo degli oneri sociali obbligatori, dalla retribuzione minima giornaliera. Da questa operazione sono stati esclusi i contratti di apprendistato e i lavoratori che nell'anno risultano fruitori di integrazioni salariali di maternità, malattia e CIG, poiché in questi casi la presenza, nella retribuzione annua, di integrazioni salariali, fa sì che la retribuzione media giornaliera risulti sistematicamente più bassa di quella effettivamente erogata dal datore di lavoro²⁵.

Stimate le diverse componenti che costituiscono gli oneri sociali teorici, si devono determinare le riduzioni contributive associate alle varie agevolazioni. Un attento vaglio delle regole legislative

classificazione ha invece riguardato tre casi, due dei quali ancora nel ramo 7 (in totale sono stati attribuiti CSC di un profilo adiacente allo 0,16% dei rapporti di lavoro coinvolti nella stima).

²² ILOR (Imposta locale sui redditi), ICIAP (Imposta comunale per l'esercizio di imprese, arti e professioni), imposta sul patrimonio netto delle imprese, tassa di concessione governativa sulla partita IVA, contributo per il servizio sanitario nazionale, contributi per l'assicurazione obbligatoria contro la tubercolosi (TBC), contributo per l'assistenza di malattia ai pensionati, tassa di concessione comunale e TOSAP (Tassa occupazione spazi ed aree pubbliche).

²³ D. Lgs. n. 446/1997.

²⁴ È possibile ottenere una disaggregazione maggiore per regione.

²⁵ La quota dei rapporti di lavoro coinvolti nel processo di stima con una retribuzione giornaliera inferiore al minimale si aggira, negli anni considerati, intorno al 2,5%. E' probabile che alcuni valori molto bassi di retribuzione, che vengono corretti con l'applicazione del minimale solo per il calcolo degli oneri sociali, siano dovuti a errori nei dati originari.

e delle norme esplicative contenute nelle circolari dell'Inps ha permesso di riconoscere e applicare particolari agevolazioni contributive. Quelle applicate ad alcuni tipi di contratto di lavoro, come i contratti di apprendistato, i contratti di formazione lavoro, di reinserimento, sono identificate da codici ai quali la normativa dell'Inps associa determinati importi di contribuzione o specifiche aliquote. Questa prima tipologia di agevolazione è, quando si sia in possesso delle informazioni che consentono di associare i codici al particolare trattamento contributivo previsto dall'Inps, di applicazione quasi automatica.

L'applicazione di sgravi²⁶ e fiscalizzazioni degli oneri sociali è invece più complessa; l'identificazione dei rapporti di lavoro ai quali applicare queste agevolazioni avviene sulla base di caratteristiche d'impresa (data di costituzione, settore di attività, inquadramento del datore di lavoro, dimensione di impresa mensile) unite a caratteristiche del rapporto di lavoro (area di lavoro, durata e collocazione temporale, età e genere del soggetto occupato, fruizione contemporanea di altri benefici). Tale identificazione risente comunque delle lacune informative di cui sopra: i benefici concessi in termini di riduzioni contributive potrebbero risultare in certa misura sovrastimati a causa della sovra-copertura della popolazione di imprese beneficiarie e in minima parte distorti per quanto riguarda le imprese pluripositonate, come spiegato in precedenza.

5. La validazione delle stime del costo del lavoro attraverso il confronto con le fonti ufficiali

La validazione delle stime prodotte passa necessariamente attraverso il confronto con quelle fornite da altre fonti. In particolare le fonti che appaiono più idonee per valutare la coerenza delle stime ottenute su WHIP sono l'indagine Labour Cost Survey (LCS) del 2000 e la rilevazione OROS.

Il confronto con i numeri indice prodotti da OROS consente di verificare in primo luogo se la procedura di stima messa in atto sui dati WHIP sia riuscita a catturare i fenomeni macroeconomici e le modifiche della normativa che influenzano la dinamica del costo del lavoro, e, in secondo luogo, se le stime prodotte diano luogo ad andamenti coerenti con quelli stimati da OROS. Il confronto con i dati tratti dalla LCS permette, invece, di verificare la coerenza dei livelli stimati del costo del lavoro in rapporto a quelli della retribuzione.

In entrambi i casi il confronto deve tenere conto, oltre che della presenza di differenze definitorie insite nella costruzione delle variabili, anche delle specifiche caratteristiche strutturali delle fonti²⁷, originali e trasformate.

Dal punto di vista delle caratteristiche strutturali i dati OROS si presentano molto simili a quelli WHIP, nonostante vi siano alcune differenze da menzionare.

La popolazione di riferimento è la stessa: le statistiche si riferiscono alle imprese private classificate nelle sezioni da C a K dell'Ateco 2002 e a tutti gli occupati²⁸ con l'esclusione dei dirigenti. Anche le definizioni delle variabili confrontate - retribuzione lorda, oneri sociali obbligatori e costo del lavoro - sono le stesse.

In particolare, la variabile retribuzione lorda ha in OROS e in WHIP gli stessi contenuti, anche se proviene da unità di rilevazione differenti (DM10 per OROS, O1M per WHIP). Il confronto su questa voce risulta di estrema importanza, in quanto per le sue caratteristiche è il benchmark perfetto per valutare la correttezza del confronto tra due banche dati di fonte simile che hanno però subito processi differenti²⁹ di trasformazione, pulizia e correzione delle variabili: è ragionevole

²⁶ I primi sgravi, influenti nel periodo considerato, risalgono alla metà degli anni Sessanta. Altri sono stati introdotti nei primi anni Ottanta e altri ancora negli anni Novanta. In alcuni casi si tratta di sgravi che rimandano o modificano regole introdotte nella metà degli anni Sessanta. In alcuni casi è prevista la possibilità di cumulo tra sgravi, così come tra sgravi e fiscalizzazioni.

²⁷ Ciascuna fonte statistica è costruita per soddisfare specifici bisogni informativi e le loro caratteristiche tecniche (concetti, definizioni, classificazioni, tecniche di rilevazione, forma delle variabili, ecc.) sono condizionate da tale particolare scopo.

²⁸ Sia OROS che WHIP rilevano i lavoratori interinali dalla parte della società fornitrice. Questi rientrano quindi nel settore K, servizi alle imprese, qualunque sia la loro occupazione.

²⁹ Per le procedure di controllo e correzione utilizzate in OROS si veda il lavoro di Baldi et al., 2008.

ritenere, infatti, che quanto più si discostano gli andamenti degli indici delle retribuzioni di WHIP e OROS, entrambi costruiti utilizzando la retribuzione e la quantità di lavoro, tanto più tenderanno a essere diversi i corrispondenti andamenti degli oneri sociali e del costo del lavoro, anche qualora il processo che li stima fosse corretto. Gli oneri sociali e il costo del lavoro sono determinati a partire dalle medesime definizioni, ma nel caso di WHIP le due variabili sono stimate sulla base di informazioni indirette, dal momento che la fonte utilizzata non contiene gli importi che ne consentano il calcolo.

Oltre agli aspetti sopra menzionati devono essere presi in considerazione altri due fattori:

- nelle due fonti le unità di riferimento sono diverse: nel caso di WHIP le stime sono prodotte per i singoli rapporti di lavoro, mentre in OROS l'unità di riferimento è l'impresa;
- in OROS le variabili di interesse sono trattate prima in modo distinto per posizione contributiva e sono successivamente aggregate per impresa, attribuendo alle imprese con più di una posizione assicurativa un'unica posizione principale. In WHIP gli oneri sociali, e di conseguenza il costo del lavoro, sono imputati sulla base delle caratteristiche d'impresa attribuite alla posizione prevalente, scelta dall'Inps e registrata nell'Osservatorio, oltre che sulla base delle caratteristiche del rapporto di lavoro. In entrambi i casi, dunque, le informazioni delle imprese sono quelle relative alla posizione prevalente, il cui processo di attribuzione avviene però in modo indipendente, con la conseguenza che differenti criteri di scelta della posizione prevalente possono portare a differenze nelle stime, a livello aggregato specie per settore, non dovute a errori.

Il confronto delle stime prodotte sui dati WHIP è stato effettuato utilizzando, di OROS, i soli numeri indice destagionalizzati e disaggregati per macro-settore, e non quelli trimestrali³⁰ forniti distintamente per i singoli settori. La destagionalizzazione operata dall'Istat genera però una perdita di dettaglio nelle stime prodotte; le procedure messe in atto, infatti, non sono in grado di depurare gli indici dagli effetti stagionali se non per aggregazioni di settori³¹. La scelta di questo tipo di dato per il confronto dipende dal fatto che i dati WHIP non seguono un andamento stagionale: il confronto è dunque avvenuto solo per macro-settore. Mentre gli indici OROS sono relativi agli andamenti della media per ULA³² della variabile considerata, quelli WHIP disegnano gli andamenti della media per giornata lavorata (giornata effettiva) della stessa variabile. La base di riferimento degli andamenti è l'anno 2000.

Per quanto riguarda le retribuzioni (tavola 2 - figura 1) dal confronto emerge che sul totale dei settori di attività considerati, da C a K, gli indici WHIP disegnano un andamento non molto discosto da quello indicato da OROS. Escludendo dal calcolo degli indici i valori estremi³³ della retribuzione giornaliera in WHIP le differenze diminuiscono, rimanendo attorno al punto percentuale, mentre si ampliano disaggregando per macro-settori, restando comunque sempre limitate. Si riscontra un picco anomalo nel 2002 nel settore dei servizi, dove peraltro gli scostamenti sono maggiori rispetto al settore dell'industria. Le differenze che emergono non sono legate al processo di imputazione degli oneri sociali e del costo del lavoro, ma sono riconducibili a tre generi di fattori: il confronto avviene tra valori ottenuti dalla popolazione delle imprese attraverso i DM10 e dall'indagine GI (OROS) e stime campionarie (WHIP); il processo di identificazione delle imprese e delle caratteristiche strutturali della classificazione delle attività economiche ad esse attribuite è differente; infine sono diverse anche le operazioni di trattamento e

³⁰ L'indagine OROS registra le voci secondo il criterio di cassa e non di competenza, ovvero attribuisce le erogazioni all'istante temporale in cui esse avvengono effettivamente, in cui si verificano circostanze di carattere episodico e accidentale o in cui si collocano particolari scadenze amministrative. Le stime prodotte in WHIP seguono l'andamento stagionale, in quanto risultano sempre aggregate su base annua.

³¹ Tuzi D., 2010.

³² Il numero di dipendenti desumibile dai modelli DM10 è pari al totale di quanti hanno percepito, nel mese di riferimento, compensi soggetti a retribuzione obbligatoria, cioè presentano almeno una giornata retribuita. Tale numero di occupati viene successivamente ridotto in unità di lavoro equivalenti a tempo pieno, attribuendo a ogni lavoratore part-time un peso pari al rapporto tra le ore che ha lavorato e le ore previste contrattualmente.

³³ Gli indici sono stati depurati dalla presenza di valori estremi, minori del 1° e maggiori del 99° percentile, della distribuzione della retribuzione giornaliera.

trasformazione messe in atto per la costruzione dei dati OROS e WHIP. Da questo primo confronto è possibile valutare quanto degli scostamenti che si rilevano negli indici delle due componenti stimate, oneri e costo, dipende dalle caratteristiche dei dati utilizzati.

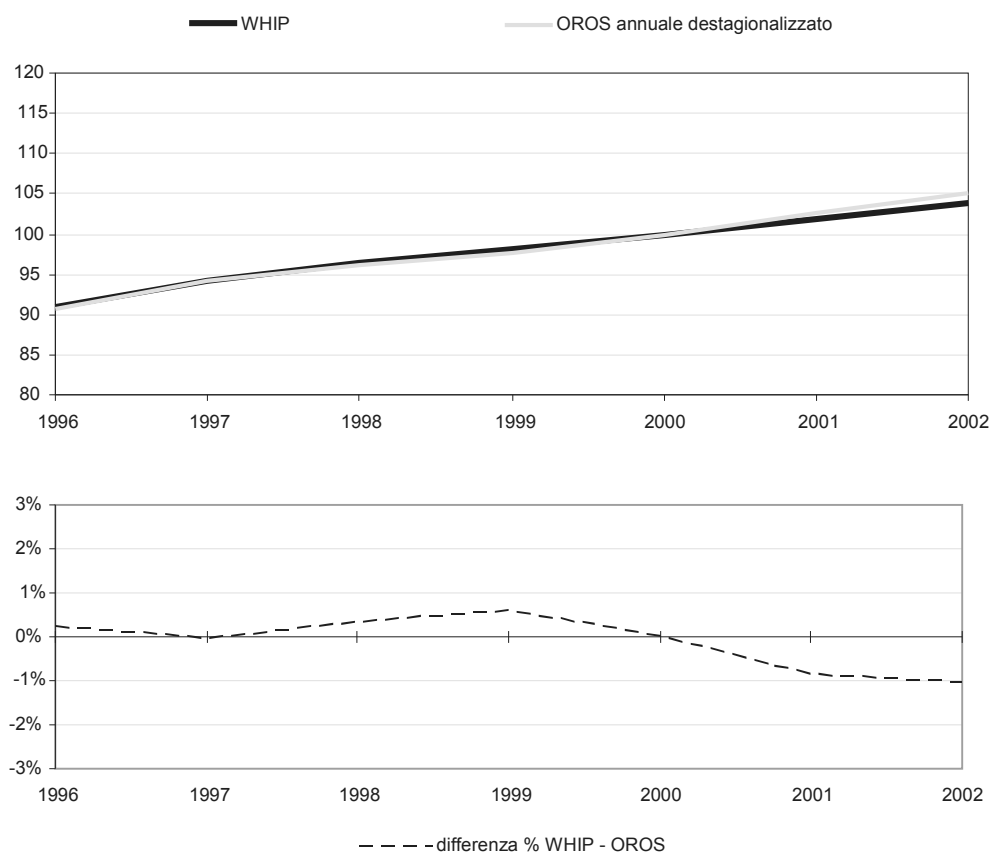
Tavola 2 – Confronto indici della retribuzione lorda

Settore	Anno	Indici OROS	Indici WHIP	Indici WHIP (a)
Industria (C-F)	1996	89.2	89.8	90.0
	1997	93.5	93.7	93.9
	1998	95.5	96.0	96.1
	1999	97.3	98.2	98.3
	2000	100.0	100.0	100.0
	2001	102.7	102.7	102.7
	2002	104.9	105.6	105.3
Servizi (G-K)	1996	93.0	92.2	92.4
	1997	95.7	100.1	95.0
	1998	97.6	99.7	97.3
	1999	98.5	98.4	98.3
	2000	100.0	100.0	100.0
	2001	102.5	100.4	100.3
	2002	104.8	102.3	101.9
Totale (C-K)	1996	90.6	90.6	90.8
	1997	94.2	96.1	94.2
	1998	96.2	97.4	96.5
	1999	97.7	98.2	98.3
	2000	100.0	100.0	100.0
	2001	102.7	101.8	101.8
	2002	105.1	104.3	103.9

Fonte: nostre elaborazioni su dati WHIP

(a) indici depurati dalla presenza di valori estremi, minori del 1° e maggiori del 99° percentile, della distribuzione della retribuzione giornaliera.

Figura 1 – Confronto indici della retribuzione lorda per il totale dei settori considerati (C-K)



Fonte: nostre elaborazioni su dati WHIP

Per quanto riguarda le stime degli oneri sociali e del costo del lavoro, si presentano due serie di indici WHIP: la prima è calcolata direttamente sui valori imputati (i_{whip}), la seconda ($i_{whip_corretti}$) moltiplicando la serie per il rapporto tra gli indici OROS e WHIP delle retribuzioni. Questa correzione consente di confrontare l'andamento degli oneri sociali medi e del costo del lavoro medio a partire dai dati WHIP con quello rilevato da OROS isolandolo dall'effetto distorsivo legato alle differenze, seppure ridotte, nell'andamento delle retribuzioni, che costituiscono la base per il calcolo degli oneri sociali e del costo del lavoro.

Nell'ambito del confronto sugli oneri sociali a livello complessivo, dove il problema dell'attribuzione del settore di attività economica alle imprese non è influente, gli scostamenti sono minimi, mentre si ampliano considerando distintamente i macrosettori dell'industria e dei servizi (tavola 3 - figura 2). Neutralizzare l'effetto delle retribuzioni permette di migliorare l'esito del confronto per gli anni a partire dal 2000, soprattutto per quanto riguarda il settore dei servizi. Per tutti i settori considerati tenere conto delle differenze negli andamenti delle retribuzioni aiuta a spiegare buona parte della differenza negli andamenti degli oneri sociali medi degli anni 2001 e 2002, mentre mette in luce la presenza di altri fattori il cui effetto risultava mitigato da quello delle retribuzioni (per esempio per gli anni 1998 e 1999).

Tavola 3 – Confronto indici degli oneri sociali

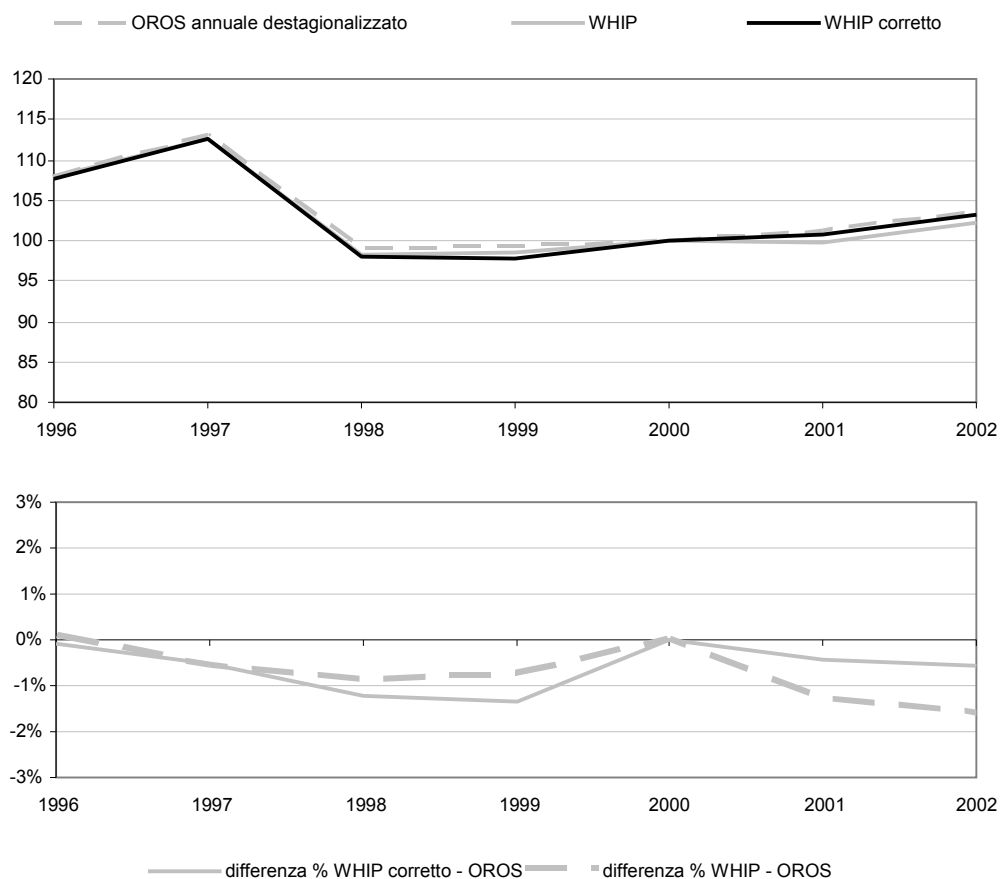
Settore	Anno	OROS	i_whip	i_whip_corretti	i_whip (a)	i_whip_corretti (a)
Industria (C-F)	1996	104.5	102.3	101.6	102.4	101.5
	1997	110.4	107.4	107.2	107.5	107.0
	1998	97.2	97.4	96.9	97.4	96.7
	1999	98.2	98.1	97.2	98.1	97.1
	2000	100.0	100.0	100.0	100.0	100.0
	2001	101.3	100.9	100.9	101.0	100.9
	2002	103.9	104.3	103.7	104.1	103.7
Servizi (G-K)	1996	112.2	116.8	117.7	117.0	117.8
	1997	116.6	127.8	122.2	120.6	121.5
	1998	101.3	102.4	100.3	99.6	99.8
	1999	100.5	98.9	99.0	98.8	99.1
	2000	100.0	100.0	100.0	100.0	100.0
	2001	101.2	98.3	100.3	98.2	100.3
	2002	104.1	99.9	102.4	99.4	102.3
Totale (C-K)	1996	107.9	107.8	107.8	108.0	107.8
	1997	113.2	115.2	113.0	112.5	112.6
	1998	99.1	99.4	98.1	98.2	97.9
	1999	99.2	98.4	97.8	98.4	97.8
	2000	100.0	100.0	100.0	100.0	100.0
	2001	101.1	99.8	100.7	99.8	100.7
	2002	103.8	102.5	103.2	102.1	103.2

Fonte: nostre elaborazioni su dati WHIP

(a) indici depurati dalla presenza di valori estremi, minori del 1° e maggiori del 99° percentile, della distribuzione della retribuzione giornaliera³⁴.

³⁴ I confronti su oneri sociali e costo del lavoro sono stati eseguiti sulla base dati depurata dai valori estremi della retribuzione giornaliera: la variabile retribuzione non subisce, nel processo di stima, alcuna manipolazione, rimanendo esattamente identica a quella fornita dall'Inps. In questo modo nessun filtro è applicato direttamente sugli importi degli oneri sociali e del costo del lavoro stimati in WHIP.

Figura 2 – Confronto indici degli oneri sociali per il totale dei settori considerati (C-K)



Fonte: nostre elaborazioni su dati WHIP

Le differenze che si osservano tra gli andamenti del costo del lavoro sono più limitate (tavola 4 - figura 3): gli scostamenti degli indici WHIP corretti non superano l'1.5% e, fatta eccezione per gli anni 1996 e 1997, assumono valori inferiori allo 0.5%. Tenere conto degli scostamenti di WHIP rispetto a OROS negli indici delle retribuzioni migliora i risultati del confronto sugli andamenti del costo del lavoro, specialmente per i servizi, fatta eccezione per gli anni 1996 e 1997.

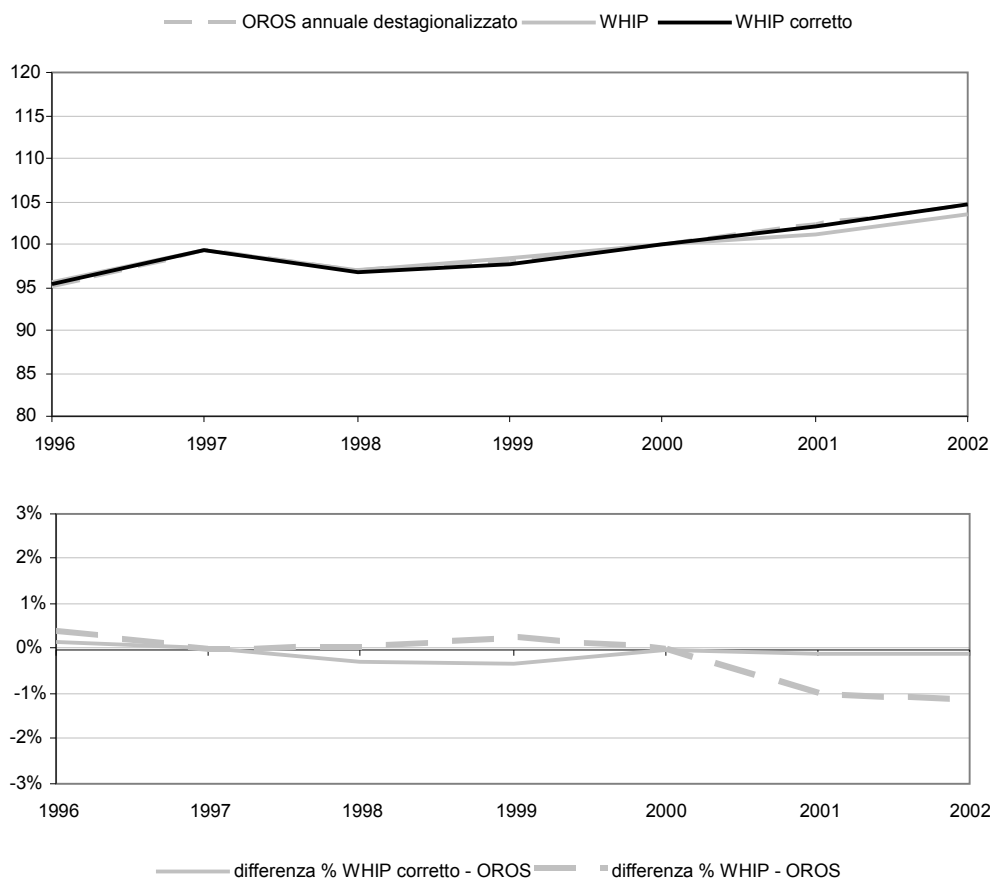
Tavola 4 – Confronto indici del costo del lavoro

Settore	Anno	OROS	i_whip	i_whip_corretti	i_whip (a)	i_whip_corretti (a)
Industria (C-F)	1996	93.5	93.3	92.7	93.5	92.6
	1997	98.2	97.5	97.3	97.7	97.3
	1998	95.9	96.4	95.9	96.5	95.8
	1999	97.5	98.2	97.3	98.3	97.3
	2000	100.0	100.0	100.0	100.0	100.0
	2001	102.3	102.2	102.2	102.2	102.2
	2002	104.6	105.3	104.6	105.0	104.6
Servizi (G-K)	1996	97.8	98.7	99.5	98.9	99.5
	1997	100.9	107.4	102.7	101.8	102.5
	1998	98.5	100.4	98.3	97.9	98.2
	1999	99.1	98.5	98.6	98.4	98.7
	2000	100.0	100.0	100.0	100.0	100.0
	2001	102.2	99.9	101.9	99.8	101.9
	2002	104.7	101.7	104.2	101.2	104.2
Totale (C-K)	1996	95.2	95.3	95.4	95.5	95.3
	1997	99.2	101.3	99.4	99.2	99.3
	1998	96.9	97.9	96.7	97.0	96.6
	1999	98.1	98.3	97.7	98.3	97.7
	2000	100.0	100.0	100.0	100.0	100.0
	2001	102.2	101.3	102.1	101.2	102.1
	2002	104.7	103.8	104.5	103.4	104.6

Fonte: nostre elaborazioni su dati WHIP

(a) indici depurati dalla presenza di valori estremi, minori del 1° e maggiori del 99° percentile, della distribuzione della retribuzione giornaliera.

Figura 3 – Confronto indici del costo del lavoro per il totale dei settori considerati (C-K)



Fonte: nostre elaborazioni su dati WHIP

L'indagine LCS si rivolge alle imprese con almeno dieci addetti³⁵ che operano nei settori di attività da C a K della classificazione Ateco 91. I dati, dichiarati dalle imprese, si riferiscono a tutte le categorie professionali (apprendisti, operai, impiegati e quadri, dirigenti). Le stime relative alle retribuzioni lorde, ai contributi obbligatori e al Tfr, prodotte in forma di medie annue, escludono gli apprendisti.

Gli elementi che costituiscono differenze strutturali tra i dati provenienti dalla LCS e quelli WHIP, alcuni dei quali già menzionati in precedenza, sono i seguenti:

- le unità di riferimento sulle quali vengono calcolati gli indicatori messi a confronto: nel caso di WHIP le stime sono prodotte sulla base delle informazioni dai rapporti di lavoro, nel caso di LCS le informazioni utilizzate sono quelle d'impresa;
- la classificazione per attività economica delle imprese: in WHIP ogni impresa è classificata nel settore proprio dell'attività associata alla posizione assicurativa prevalente, mentre in LCS l'impresa che esercita due o più attività afferenti a settori diversi sarà classificata nel settore dell'attività per la quale è maggiore "il valore aggiunto al costo dei fattori, oppure in termini di quota del fatturato o di numero medio annuo di addetti impegnati"³⁶.
- la popolazione di riferimento: tra le imprese rilevate LCS include anche le unità di natura pubblica operanti nelle sezioni da C a K, che invece sono escluse da WHIP.

³⁵ In particolare, la rilevazione è campionaria per le imprese fino a 250 addetti, totale per le imprese più grandi; essa è condotta attraverso l'invio di un questionario alle imprese selezionate.

³⁶ Questionario LCS 2000.

La scelta di rilevare solo le imprese con almeno 10 addetti, inoltre, può indurre un margine di errore nella selezione, in WHIP, dei rapporti di lavoro da includere per il confronto: non è qui disponibile, infatti, il dettaglio mensile della forza occupazionale nel 2000 ma solo il dato medio nell'anno. Allo stesso modo l'inclusione dei dirigenti nelle stime LCS rende più problematico il confronto con WHIP sui livelli di retribuzione e di costo del lavoro. Poiché la comparazione con altre fonti sarebbe stata comunque viziata da differenze ben più rilevanti (si veda ad esempio il dato della Contabilità Nazionale), il confronto delle stime tratte da WHIP con quelle fornite dall'indagine LCS è stato comunque effettuato, ma limitatamente al rapporto tra costo del lavoro e retribuzione lorda: questo permette di controllare la stabilità della relazione tra il costo del lavoro stimato e la retribuzione lorda, ossia il peso di una componente sull'altra, in modo disaggregato per sezione di attività economica.

Per simulare in WHIP l'estrazione di un campione coerente con quello utilizzato in LCS i rapporti di lavoro sono stati inclusi nella selezione se attivi nel 2000 in imprese con una dimensione media superiore ai 9 dipendenti, e con qualifica di operaio o di impiegato.

Dal confronto sul rapporto tra il costo del lavoro e la retribuzione lorda sul totale dei settori di attività considerati, da C a K, emerge una quasi generale corrispondenza tra le stime WHIP e quelle LCS (tavola 5). I valori WHIP risultano sovrastimare i corrispondenti valori LCS dello 0.10% non controllando per eventuali outlier³⁷ e dello 0.13% escludendo dal calcolo degli indici i valori estremi.

Nell'aggregato le discordanze possono essere ricondotte a due fattori che contribuiscono a elevare il rapporto costo - retribuzione in WHIP:

- l'esclusione dei dirigenti nel campione oggetto di stima tende a innalzare il rapporto costo - retribuzione in quanto, a parità di retribuzione lorda, per i dirigenti sono minori le aliquote di contribuzione Inps (almeno 1 punto percentuale in meno rispetto agli impiegati, 2 o 3 punti in meno rispetto agli operai), così come minori, ma con scostamenti più consistenti rispetto alle altre qualifiche, sono le aliquote Inail;
- il confluire, nella retribuzione lorda disponibile, di alcune voci, al momento non distinguibili, alle quali si dovrebbero applicare aliquote contributive inferiori a quelle che si applicano sulla retribuzione ordinaria (per esempio i compensi per lavoro straordinario): anche questo comporta un aumento del rapporto tra costo del lavoro e retribuzione.

Scendendo più nel dettaglio, è probabile che le differenze riconducibili alla classificazione delle imprese in settori di attività spieghino gli scostamenti maggiori. Infatti se per i macro-settori le differenze si mantengono ancora relativamente simili in valore assoluto e non superiori allo 0.4%, le differenze si ampliano quando si guarda alle singole sezioni.

Il rapporto costo - retribuzione ottenuto da WHIP per il settore dell'industria pare sovrastimare il suo corrispondente fornito da LCS dello 0.16%, mentre per i servizi l'indicatore WHIP sottostima quello LCS dello 0.34%. Nelle singole sezioni riconducibili all'industria l'indice WHIP sovrastima quello LCS del 2.71-2.72% per la sezione C, del 1.76-1.72% per la sezione E, mentre la sezione D rispecchia il comportamento dell'intero comparto industriale. Nel settore delle costruzioni (F) la sovrastima del rapporto costo - retribuzione delle stime su WHIP è del 2.45-2.39%. Nei servizi la direzione dello scostamento non è invece uniforme: accanto a valori negativi che si aggirano intorno allo 0.4%, se ne osservano altri più elevati nell'ordine dell'1.7% nel commercio e del 2.7% nel settore finanziario, mentre nei trasporti (in cui la presenza di imprese pubbliche rilevate da LCS potrebbe essere significativa) lo scostamento di WHIP da LCS è positivo e ammonta al 3.3%.

³⁷ Calcolati con la stessa modalità utilizzata per il confronto con OROS e sempre sulla base della retribuzione giornaliera.

Tavola 5 – Confronto indici del costo del lavoro

Settore	LCS	CL-WHIP	CL-WHIP (a)	WHIP-LCS %	WHIP(a)-LCS %
Industria s.s. (C-E)	1.40	1.41	1.41	0.16	0.15
Costruzioni (F)	1.41	1.44	1.44	2.45	2.39
Servizi (G-K)	1.38	1.37	1.37	-0.34	-0.32
Totale (C-K)	1.39	1.39	1.39	0.10	0.13
Sezioni					
C - Estrazione di minerali	1.40	1.44	1.44	2.71	2.72
D - Attività manifatturiere	1.41	1.41	1.41	0.12	0.12
E - Distr. energia, gas e acqua	1.37	1.40	1.40	1.76	1.72
F - Costruzioni	1.41	1.44	1.44	2.45	2.39
G - Commercio	1.39	1.37	1.37	-1.68	-1.70
H - Alberghi e ristoranti	1.38	1.37	1.37	-0.38	-0.37
I - Trasporti, magazzinaggio e comunicazioni	1.36	1.40	1.40	3.30	3.27
J - Intermediari finanziari e monetari	1.39	1.35	1.35	-2.66	-2.66
K - Att. Imm., informatica, servizi imprese	1.38	1.38	1.38	-0.28	-0.31

Fonte: nostre elaborazioni su dati WHIP

(a) indici depurati dalla presenza di valori estremi, minori del 1° e maggiori del 99° percentile, della distribuzione della retribuzione giornaliera.

6. Prime evidenze dai dati sul costo del lavoro

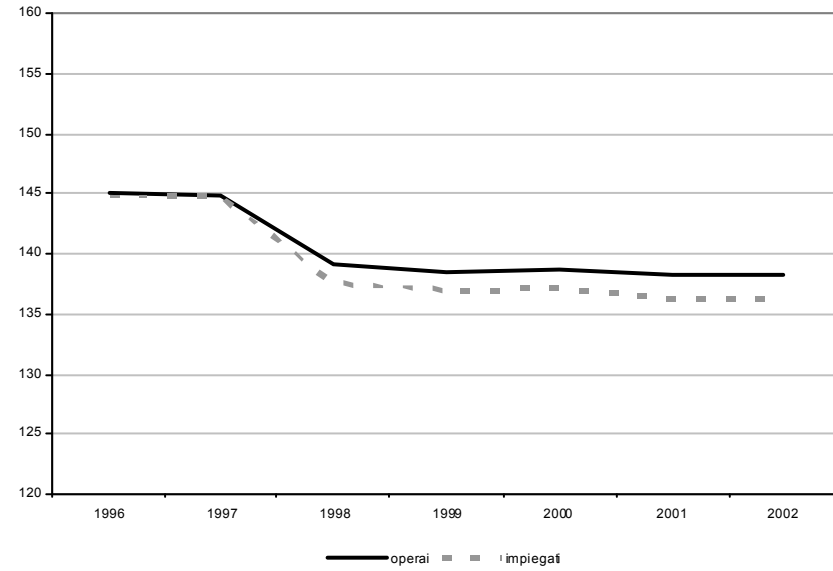
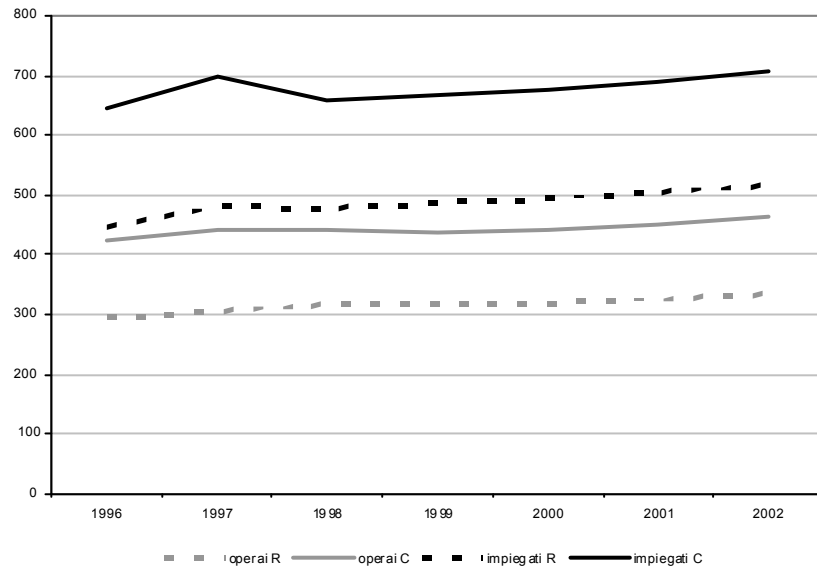
In questo paragrafo si presentano alcune statistiche tratte dalla banca dati sperimentale sul costo del lavoro dipendente. Come è stato spiegato in precedenza tale banca dati è stata costruita a partire dai microdati di fonte Inps contenuti nell'archivio WHIP, in particolare quelli relativi ai rapporti di lavoro dipendente. Le statistiche sono riferite alla retribuzione lorda e al costo del lavoro dipendente. Mentre i valori della retribuzione sono forniti direttamente dall'Inps, quelli del costo del lavoro sono il risultato del processo di stima i cui tratti salienti sono stati presentati nei paragrafi precedenti. Retribuzione lorda e costo del lavoro sono presentati in forma di importi medi settimanali. In tabella si riportano media e scarto interquartile della retribuzione lorda e del costo del lavoro relativi al 2000 per area, qualifica e classe di età: queste sono soltanto alcune delle dimensioni entro le quali il dato è disaggregabile. I grafici che seguono presentano alcuni andamenti della retribuzione e del costo del lavoro medi settimanali, o del rapporto tra costo e retribuzione, per area, comparto di attività economica, età, tipo di contratto e qualifica.

Tavola 6 – Valori settimanali in euro di retribuzione e costo del lavoro dipendente, anno 2000

Area di lavoro	Qualifica	Età	Retribuzione		Costo del lavoro	
			media	scarto interquartile	media	scarto interquartile
Nord	operaio	<25	284.3	81.8	391.4	132.0
Nord	operaio	25-40	326.3	106.7	458.8	157.6
Nord	operaio	>40	357.6	129.1	386.3	188.1
Nord	impiegato	<25	333.7	82.5	452.7	126.9
Nord	impiegato	25-40	465.9	211.3	644.1	297.2
Nord	impiegato	>40	629.9	321.9	873.3	446.0
Centro	operaio	<25	272.7	80.7	369.1	128.3
Centro	operaio	25-40	304.2	88.6	423.7	135.0
Centro	operaio	>40	339.4	119.8	478.9	178.1
Centro	impiegato	<25	318.3	81.1	424.1	130.2
Centro	impiegato	25-40	445.4	213.2	611.7	300.8
Centro	impiegato	>40	623.3	325.9	861.8	453.3
Sud	operaio	<25	257.1	81.3	329.1	128.8
Sud	operaio	25-40	286.7	80.3	385.3	143.9
Sud	operaio	>40	326.4	112.4	453.2	179.9
Sud	impiegato	<25	298.0	66.8	369.0	121.5
Sud	impiegato	25-40	395.0	171.3	522.7	260.0
Sud	impiegato	>40	589.7	309.4	809.4	423.5

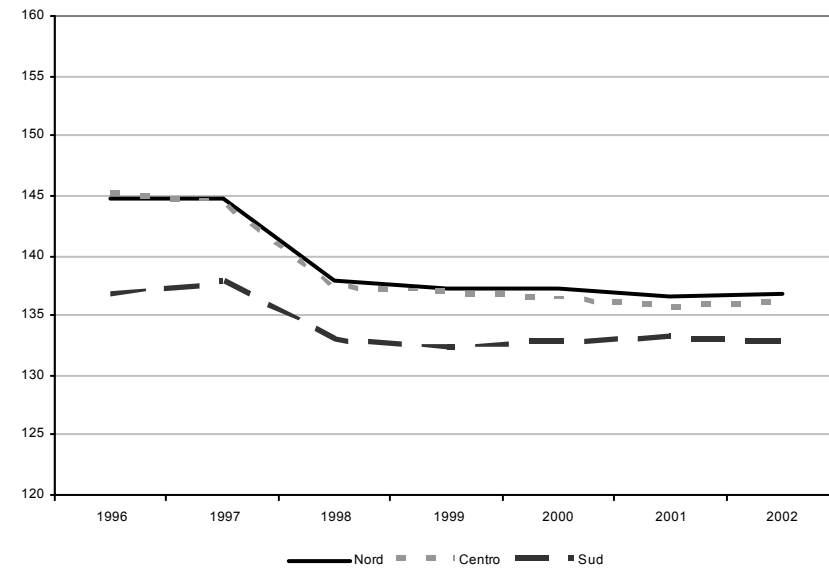
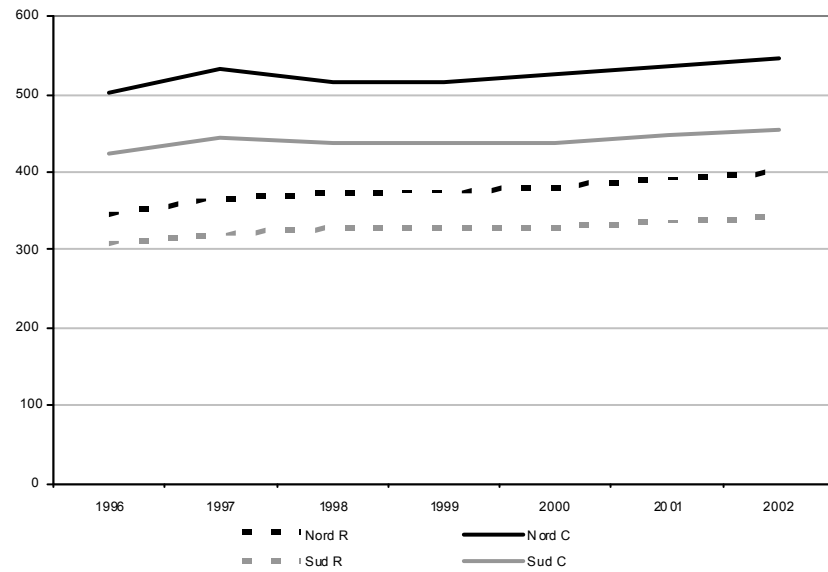
Fonte: nostre elaborazioni su dati WHIP

Figura 4 – Costo del lavoro (C) e retribuzione settimanali (R) per qualifica (sinistra); rapporto costo-retribuzione per qualifica (destra)



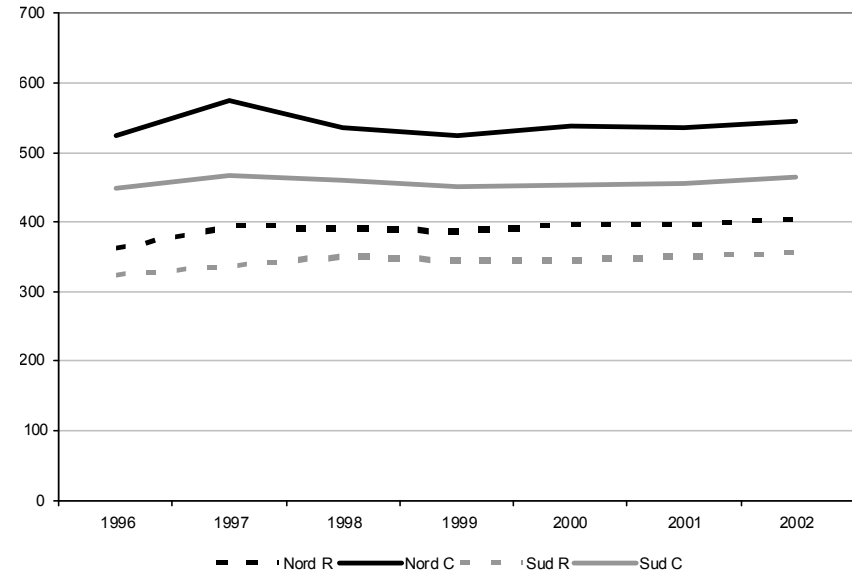
Fonte: nostre elaborazioni su dati WHIP

Figura 5 – Costo del lavoro (C) e retribuzione settimanali (R) per area di lavoro (sinistra); rapporto costo-retribuzione per area di lavoro (destra)



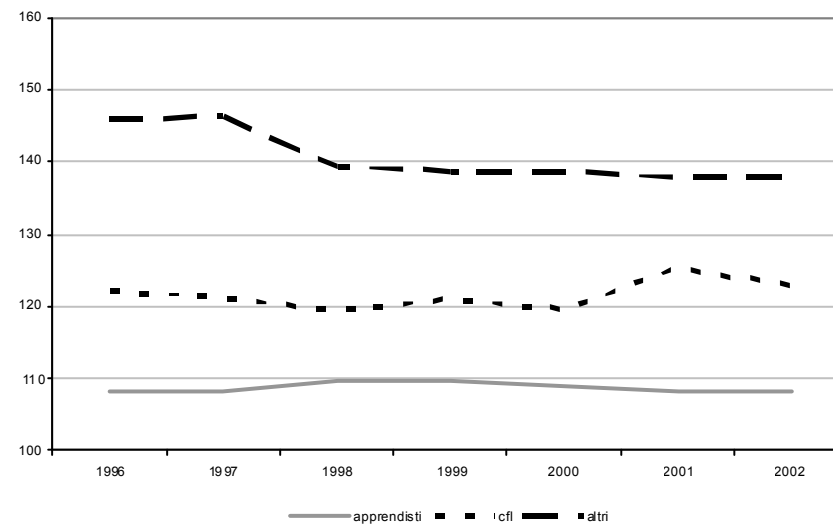
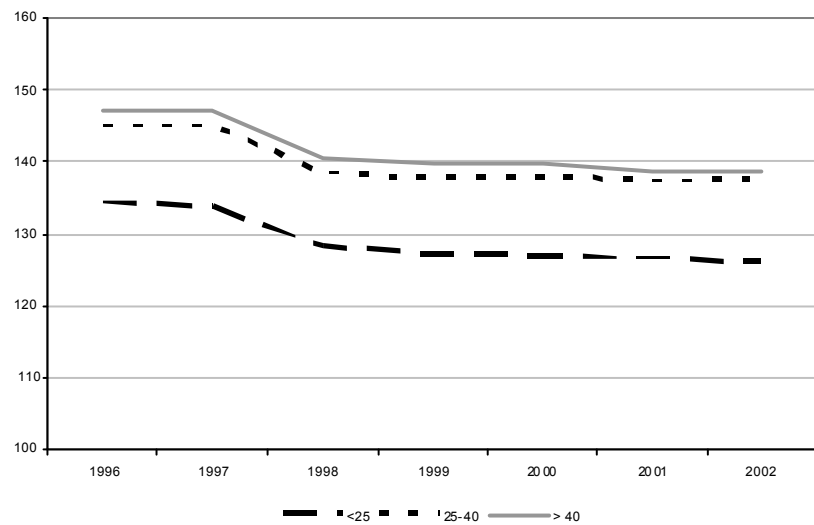
Fonte: nostre elaborazioni su dati WHIP

Figura 6 – Industria (sezioni C-F): costo del lavoro (C) e retribuzione settimanali (R) per area di lavoro (sinistra); Servizi (sezioni G-K): costo del lavoro e retribuzione settimanali per area di lavoro (destra)



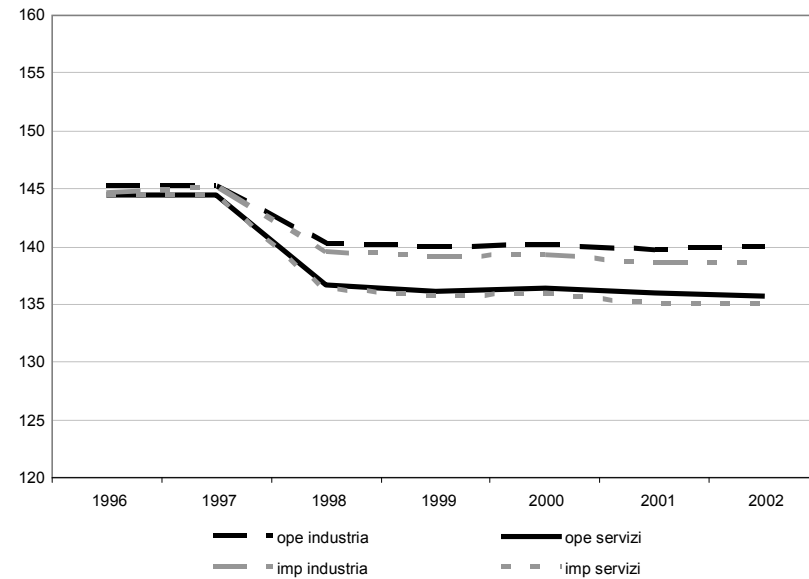
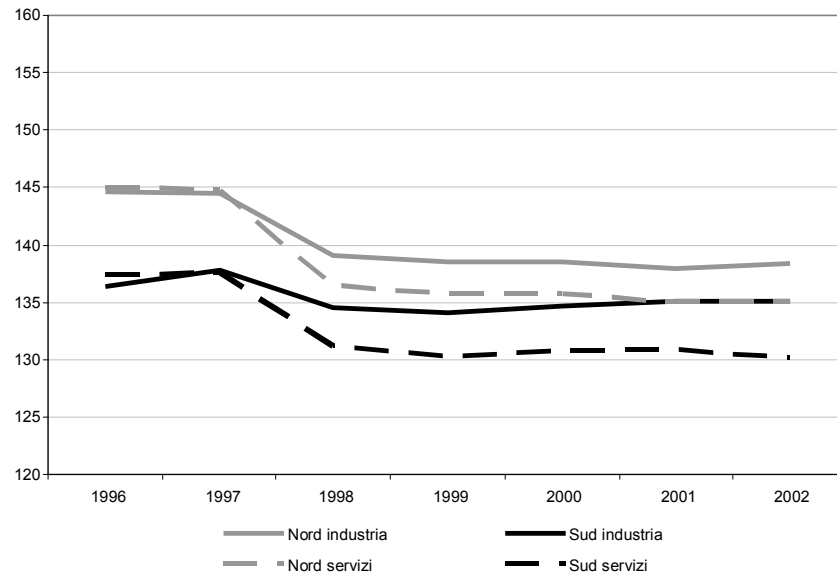
Fonte: nostre elaborazioni su dati WHIP

Figura 7 – Rapporto costo-retribuzione per età (sinistra); rapporto costo-retribuzione per tipo di contratto (destra)



Fonte: nostre elaborazioni su dati WHIP

Figura 8 – Rapporto costo-retribuzione per area e comparto (sinistra); rapporto costo-retribuzione per qualifica e comparto (destra)



Fonte: nostre elaborazioni su dati WHIP

Riferimenti bibliografici

- Baldi, C., E. Cimino, F. Rapiti, P. Minicucci, R. Succi e D. Tuzi. 2001. *L'utilizzo dei dati Inps per la stima trimestrale del numero dei dipendenti, le retribuzioni, il costo del lavoro e le ore lavorate*. Documenti Istat, 14. Roma.
- Baldi, C., F. Ceccato, E. Cimino, M.C. Congia, S. Pacini, F. Rapiti e D. Tuzi. 2008. *Il controllo e la correzione in una indagine congiunturale su dati amministrativi. Il caso della rilevazione Oros*. Contributi Istat, 13. Roma.
- Contini, B., L. Pacelli e R. Revelli. 1994. Struttura occupazionale e retribuzioni dai dati individuali degli archivi Inps. *CNEL – Norme e metodi sul mercato del lavoro*, 56, Vol. III. Roma.
- Forte, A. e B. Gallo. 2008. *Guida all'Inps*, IPSOA.
- Istat. 2004. Metodologie di stima degli aggregati di contabilità nazionale a prezzi correnti. Italia - Inventario SEC 95. *Metodi e Norme*, 21. Roma.
- Leombruni, R., R. Quaranta e C. Villosio. 2010. Note di pubblicazione WHIP 3.2. *Rapporti tecnici WHIP - Laboratorio R. Revelli*, 1: <http://www.laboratoriorevelli.it/index.php?t=2&p=3&s=6>.
- Leombruni, R. e R. Quaranta. 2011. La codifica di settore in WHIP. Problemi correnti e studio di un algoritmo di ricostruzione della codifica Ateco 2002. *Rapporti tecnici WHIP - Laboratorio R. Revelli*, 2: <http://www.laboratoriorevelli.it/index.php?t=2&p=3&s=6>.
- Pacini, S. 2009. *Indicatori territoriali sul mercato del lavoro utilizzando dati amministrativi: uno studio di fattibilità su retribuzioni lorde, oneri sociali e costo del lavoro*. Documenti Istat, 2. Roma.
- Tuzi, D. 2010. *L'individuazione di dati anomali nella Rilevazione Oros: un approccio in serie storiche basato su 'Tramo for Errors'*. Documenti Istat, 2. Roma.