Valutazione della Qualità della Ricerca 2004-2010 (VQR 2004-2010)

Rapporto finale di area
Gruppo di Esperti della Valutazione dell'Area 13 (GEV13)

National Agency for the Evaluation of Universities and Research Institutes

anvur

Agenzia Nazionale di Valutazione del sistema Universitario e della Ricerca

Evaluation of Research Quality

vQr

Valutazione Qualità della Ricerca

1 INTRODUZIONE 2

1.1 LA COMPOSIZIONE DEL GEV 3
1.2 I TEMPI 4
1.3 DESCRIZIONEDELL'AREA 6

2 LA VALUTAZIONE DEI LAVORI 9

2.1 LA CLASSIFICAZIONE DELLE RIVISTE 9
2.2 LA VALUTAZIONE BIBLIOMETRICA 10
2.3 I REVISORI ESTERNI 12
2.4 LA VALUTAZIONE PEER 13
2.5 I LAVORI "NON VALUTABILI" 15
2.6 CRITICITÀ EMERSE E SUGGERIMENTI PER UN NUOVO CICLO DI VALUTAZIONE 16

3 I LAVORI VALUTATI 19

4 I RISULTATI DELLA VALUTAZIONE 20

4.1 SINTESI DEI RISULTATI DI AREA 20
4.2 LA VALUTAZIONE DELLE STRUTTURE 23
4.3 LA VALUTAZIONE DEI DIPARTIMENTI 27
4.4 I RISULTATI DEI SOGGETTI VALUTATI 29

5 ANALISI DEI RISULTATI 33

6 CONSIDERAZIONI FINALI 37

GLOSSARIO 38

ELENCO DELLE TABELLE 40

APPENDICE A: COMPARISON BETWEEN PEER REVIEW AND BIBLIOMETRIC EVALUATION 50

APPENDICE B: CRITERI PER LA VALUTAZIONE DEI PRODOTTI DI RICERCA 60

APPENDICE C: CLASSIFICATION, IMPUTATION OF MISSING VALUES AND RANKING OF JOURNALS 69

APPENDICE D: ANALISI DELLE CITAZIONI 104

APPENDICE E: LA VALUTAZIONE PEER REVIEW E I REVISORI ESTERNI 108

1 Introduzione

Dal punto di vista dei temi di ricerca, delle metodologie di studio e della collocazione editoriale dei lavori scientifici, l'Area 13 (Scienze Economiche e Statistiche) è intermedia tra le cosiddette "scienze dure" (le Aree 1-9) e le scienze umanistiche (le Aree 10-14). Queste caratteristiche dell'Area 13 si riflettono in una notevole eterogeneità delle sedi di pubblicazione, con una forte presenza sia di articoli su riviste internazionali censite dalle maggiori banche dati, sia di articoli su riviste non presenti in tali banche dati, di monografie e capitoli di libri, pubblicati in Italia o all'estero.

Il compito affidato al Gruppo di Esperti della Valutazione (GEV) è stato di valutare i lavori dei ricercatori dell'area utilizzando le griglie di classificazioni fissate dal bando ministeriale. Le griglie prevedono che ciascun lavoro inviato alla VQR sia collocato in una classe di merito (Eccellente, Buono, Accettabile, Limitato, Non valutabile, Caso di Plagio o Frode). Per individuare le classi di merito, il bando prevede che il GEV possa utilizzare sia l'analisi bibliometrica sia la peer review.

La riflessione preliminare del GEV è stata sui pregi e difetti che ciascuno dei due strumenti di valutazione presenta. In linea di principio, la peer review consente di valutare nel migliore dei modi il contenuto di un lavoro scientifico mediante il giudizio di uno o più esperti. D'altro canto, la dimensione stessa della VQR (circa 12.000 lavori inviati al GEV) ne riduce la praticabilità e forse l'affidabilità. Esiste cioè un chiaro conflitto tra qualità e quantità della peer review nella ricerca di revisori qualificati e nell'attenzione che ciascuno di essi può dedicare alla valutazione di un lavoro. Inoltre, la valutazione peer review non esclude del tutto i conflitti di interesse, né assicura una valutazione uniforme tra i diversi lavori, settori disciplinari, temi di ricerca, ecc. Infine, non è sempre facile ricondurre a una sintesi i giudizi difformi di due revisori, soprattutto quando i pareri sono molto discordi.

D'altro canto, la valutazione bibliometrica è possibile solo per articoli pubblicati su riviste, categoria nella quale ricade circa la metà dei lavori complessivamente conferiti al GEV. Ed anche la valutazione bibliometrica ha noti limiti, non solo per la scelta degli indicatori bibliometrici, ma anche per l'automatismo della valutazione, che assegna una fascia di merito

elevata a un lavoro per il solo fatto di essere pubblicato su una rivista molto citata, o per il solo fatto di avere ricevuto molte citazioni.

Come descritto di seguito, la soluzione adottata dal GEV è stata di valutare esclusivamente in peer review tutti i lavori scientifici non pubblicati su riviste (monografie, capitoli di libro, atti di congresso); invece, gli articoli su rivista sono stati valutati utilizzando entrambi gli strumenti di valutazione, senza alcuna gerarchia tra i due, e verificando in itinere la validità dei due strumenti attraverso un controllo reciproco di qualità. Il confronto statistico tra i due strumenti di valutazione (due revisori esterni, un gruppo di consenso interno, valutazione bibliometrica) rappresenta un contributo significativo che il GEV ha proposto nel corso del suo lavoro, la cui validità potrebbe avere in futuro anche altre applicazioni.

Vale la pena sottolineare che ogni valutazione ha margini di errore. La peer review, ad esempio, perché un giudizio potrebbe essere dato frettolosamente o senza adeguata competenza da parte del revisore; l'analisi bibliometrica perché le banche dati contengono inevitabilmente errori e perché le citazioni delle riviste e dei singoli articoli sono indicatori imperfetti della qualità di un lavoro scientifico.

Nel confermare la fiducia nei metodi adottati per gli scopi della VQR, cioè la valutazione di raggruppamenti consistenti di docenti che operano all'interno di strutture di ricerca con una dimensione minima significativa, il GEV ritiene quindi anche doveroso rappresentare gli oggettivi limiti di validità delle procedure adottate.[1]

1.1 La composizione del GEV

Il GEV ha valutato i lavori di tutti i 19 settori scientifico-disciplinari (SSD) dell'Area 13 (cfr. Tab. 1.1). Il GEV è composto di 36 membri (Tab. 1.2), organizzati in tre gruppi più ristretti (i sub-GEV, elencati in Tab. 1.3).

[1] Oltre alla relazione finale, il rapporto si compone di cinque appendici e un glossario. Le appendici A e C (sul confronto tra i criteri di valutazione e sulla classificazione delle riviste) potrebbero essere di interesse più generale e si è deciso di pubblicarle in lingua inglese. Nonostante sia stata posta la massima attenzione all'elaborazione dei dati, errori sono sempre possibili. Invitiamo pertanto a segnalarli per poterli correggere in una versione successiva.

Complessivamente 10 membri del GEV lavorano in università straniere (includendo tra queste l'Istituto Universitario Europeo); i rimanenti 26 provengono da 20 istituzioni o enti di ricerca diversi, tra cui 5 atenei del Mezzogiorno (Salento, Cagliari, Palermo, Catania, Napoli Federico II); 6 dell'Italia centrale (Pisa, Siena, S. Anna di Pisa, Perugia, Tor Vergata, Chieti-Pescara); 8 dell'Italia settentrionale (Torino, Milano Statale, Bocconi, Bolzano, Bologna, Venezia, Modena e Reggio Emilia, Padova), e un'istituzione di ricerca non universitaria (Banca d'Italia). Si tratta di un gruppo diversificato sul piano delle competenze e delle sedi di affiliazione, che ha affrontato la VQR con impegno, rigore, trasparenza e forte spirito di collaborazione. Per organizzare i propri lavori, il GEV ha creato gruppi di lavoro più ristretti (i sub-GEV) nelle aree economica, aziendale, e statistico-matematica. La numerosità dei componenti del sub-GEV riflette approssimativamente il numero di lavori scientifici attesi per ciascun gruppo di lavoro.

Il GEV ha tenuto complessivamente tre riunioni plenarie e un incontro tra coordinatori sub-GEV. In occasione del secondo e terzo incontro plenario si sono tenuti anche incontri separati tra sub-GEV. Molti altri incontri tra coordinatori e tra gruppi più ristretti sono avvenuti per via telematica. Tutte le decisioni sono state sottoposte all'approvazione dei membri GEV. I membri del GEV sono stati costantemente informati, con cadenza settimanale, degli sviluppi della valutazione e dei problemi incontrati e consultati per ogni decisione di rilievo. Tutte le decisioni più importanti sono state prese all'unanimità e in piena armonia all'interno del gruppo; le decisioni prese in via telematica sono state approvate nelle riunioni plenarie. Il Gruppo è stato coadiuvato da un'Assistente, la dott.ssa Carmela Anna Nappi, che ha lavorato con grande competenza, passione ed efficienza; a lei va il sentito ringraziamento di tutto il Gruppo.

1.2 I tempi

Il GEV ha lavorato complessivamente 18 mesi, da dicembre 2011 a maggio 2013. Dopo la riunione di insediamento del dicembre 2011, nei mesi di gennaio e febbraio 2012 è stato discusso e approvato il "Documento dei criteri per la valutazione dei prodotti di ricerca dell'Area 13", pubblicato sul sito ANVUR il 29 Febbraio 2012.

Tra marzo e aprile 2012 il GEV ha predisposto l'elenco delle riviste annunciato nel Documento dei Criteri e in aprile ha raccolto i dati bibliometrici su Google Scholar necessari per l'imputazione degli indicatori bibliometrici delle riviste non censite da WoS. Negli stessi mesi è stato predisposto un primo elenco di revisori esterni, consultando i membri del GEV e le società scientifiche. Il 30 aprile 2012 è stato pubblicato sul sito dell'ANVUR un primo elenco delle riviste e degli indicatori considerati dal GEV per l'analisi bibliometrica. L'elenco è rimasto a

disposizione della comunità scientifica (società scientifiche, responsabili di riviste, singoli studiosi) per commenti, revisioni e integrazioni fino alla fine di agosto 2012.

Tra maggio 2012 e agosto 2012 il GEV, in seguito ai molti commenti ricevuti, ha pubblicato sul sito ANVUR le correzioni adottate (errori materiali sul nome delle riviste e degli indicatori bibliometrici, classificazione delle riviste nelle aree di ricerca, codici ISSN, definizione di rivista italiana, ecc.) e inserito numerose altre riviste che rispettavano i requisiti di classificazione del GEV. Le correzioni sono state pubblicate sul sito ANVUR con i comunicati del 10 maggio 2012, 12 giugno 2012, 11 luglio 2012 e 4 settembre 2012. L'elenco definitivo delle riviste con i ranking delle quattro classi previste dal Bando VQR è stato pubblicato in data 4 settembre 2012. Da quella data, precedente all'inizio della valutazione peer review, non sono state apportate altre integrazioni o correzioni all'elenco.

Tra luglio e settembre 2012 i lavori da valutare in peer review sono stati assegnati a uno dei tre sub-GEV in base al SSD del soggetto proponente e ripartiti tra i membri del GEV in base all'argomento principale della ricerca. Come previsto dai criteri di valutazione del GEV, i lavori da valutare in peer review hanno avuto una prima valutazione interna al GEV da parte dei Gruppi di Consenso basata sui criteri previsti dal bando VQR per individuare i lavori non valutabili e quelli chiaramente limitati, cioè che si collocano nettamente sotto la mediana della scala di valore condivisa dalla comunità internazionale.

La valutazione peer review, dall'invio dei lavori ai revisori alla formulazione del giudizio di consenso da parte dei due membri GEV responsabili del lavoro, si è svolta da ottobre 2012 a marzo 2013. I tempi sono stati leggermente inferiori per l'area economica e statistica e superiori per l'area aziendale, a causa del maggior numero di lavori in peer review valutati dal sub-GEV dell'area aziendale.

A febbraio 2013 sono stati classificati in via preliminare i lavori "non valutabili". A marzo 2013 l'ANVUR, nello spirito di trasparenza e collaborazione con gli atenei e gli enti di ricerca, ha inviato a ciascuna struttura l'elenco dei lavori giudicati preliminarmente "non valutabili" insieme con la motivazione e la richiesta di trasmettere eventuali controdeduzioni. I dati delle valutazioni sono stati elaborati nei mesi di aprile e maggio 2013; è stato quindi redatto il Rapporto Finale di Area, discusso e approvato nella riunione plenaria del 13 maggio.

1.3 Descrizionedell'area

L'Area 13 (Scienze Economiche e Statistiche) è composta da circa 4.500 ricercatori che afferiscono complessivamente a 90 strutture, tra cui 86 università e 4 enti di ricerca o consorzi (Consiglio Nazionale delle Ricerche, Centro Euro-Mediterraneo per i Cambiamenti Climatici, Istituto Nazionale di Alta Matematica Francesco Severi e Fondazione Edmund Mach). L'area comprende quattro macro-settori: economia (circa il 30% dei ricercatori), economia aziendale (circa il 40%), storia economica (5%), statistica e metodi matematici per le decisioni (25%), distribuiti in 19 settori scientifico-disciplinari.

In termini di dimensioni, come evidenziato dalla Tab. 1.4, in 15 strutture la presenza dell'Area 13 è numericamente limitata (meno di 10 lavori attesi). Nelle rimanenti 75 strutture, 38 sono strutture medio-piccole (tra 11 e 150 lavori attesi) e 37 sono strutture grandi (oltre 150 lavori). Cinque atenei hanno oltre 400 lavori attesi: Torino, Cattolica di Milano, Bocconi, Roma La Sapienza e Bologna.

L'Area 13 è caratterizzata da una notevole eterogeneità e, allo stesso tempo, da una sovrapposizione di temi di ricerca tra alcuni SSD e da una presenza non trascurabile di lavori multidisciplinari. L'eterogeneità riguarda non solo i temi di ricerca, ma anche la tipologia e la collocazione dei lavori; ad esempio, studiosi di statistica e matematica per le decisioni pubblicano molto spesso i propri lavori su riviste internazionali in lingua inglese, mentre in altri settori, come la storia economica o alcuni settori di economia aziendale, è prevalente la monografia o la pubblicazione su volume.

In secondo luogo, in molti casi vi sono sovrapposizioni di interessi scientifici tra docenti che appartengono a settori disciplinari diversi. Ad esempio, molti ricercatori dei settori economici o della matematica per le decisioni hanno interessi di ricerca nel campo della finanza; molti economisti hanno interessi nel campo dell'analisi della struttura industriale e delle scelte delle imprese; spesso tra econometrici, statistici economici ed economisti applicati vi è convergenza di interessi di ricerca e sedi di pubblicazione.

Infine, molti lavori dell'Area 13 hanno una forte connotazione multidisciplinare e sono svolti in collaborazione non solo tra ricercatori che appartengono a diversi settori della stessa area, ma anche tra docenti di aree diverse. E' frequente, infatti, la collaborazione della statistica sperimentale con le scienze mediche e l'ingegneria, della statistica metodologica con le scienze matematiche, delle scienze merceologiche con le scienze chimiche e agrarie, dell'economia del lavoro con la sociologia, dell'economia pubblica con le scienze giuridiche, ecc.

Per quanto riguarda la tipologia delle pubblicazioni, è interessante notare che vi sono differenze non solo tra SSD, ma anche all'interno di ciascun SSD. In ciascuno dei SSD, ad esempio, sono presenti pubblicazioni su riviste, monografie, capitoli di libro e atti di congresso, ma le proporzioni cambiano significativamente a seconda dei settori.

Per quanto riguarda la disponibilità di banche dati bibliometriche, l'Area 13 si trova in una posizione intermedia tra le Aree 01-09 (per le quali sono disponibili banche dati degli articoli su rivista) e le discipline umanistiche, in cui mancano quasi del tutto banche dati. Per l'Area 13 le banche dati sono disponibili, ma incomplete, perché molte riviste (italiane e internazionali) non sono censite. Come descritto più avanti, preliminare alla vera e propria analisi bibliometrica dell'Area 13 è stata quindi la decisione di costruire una banca dati comprensiva della gran parte delle riviste scientifiche in cui pubblicano ricercatori dell'Area.

Le Tab. 1.5, 1.6. e 1.7 riportano la distribuzione dei lavori per tipologia di pubblicazione, SSD, sub-GEV e lingua di pubblicazione. Nel complesso dell'area, il 62% dei lavori è rappresentato da articoli su rivista, seguito da contributi in volume (20%) e monografie (13%). La percentuale di articoli su riviste raggiunge un minimo nel settore SECS-P/12 (27%) e un massimo nel settore SECS-S/02 (91%). Al contrario, la percentuale di monografie è massima nel settore SECS-P/07 e raggiunge valori inferiori al 5% in tutti i settori statistici e matematici (ad esclusione di SECS-S/04), in econometria, economia politica e scienze merceologiche.

Non per tutti i lavori la scheda riporta la lingua di pubblicazione. E' interessante notare che la lingua prevalente dei lavori inviati al GEV è l'inglese (57%), seguita dall'italiano (34%). Il numero di lavori pubblicati in altre lingue è trascurabile (1%); per l' 8% dei lavori manca l'informazione sulla lingua di pubblicazione.

All'interno dei SSD, la percentuale di lavori pubblicati in lingua inglese è superiore al 70% nei settori SECS-P/01, SECS-P/05, SECS-S/01, SECS-S/02, SECS-S/06; raggiunge valori inferiori al 40% nei settori SECS-P/07, SECS-P/09, SECS-P/11 e SECS-P/12. In generale, si nota una maggiore propensione a pubblicare articoli su rivista in lingua inglese nei settori economici e statistici, rispetto a quelli di economia aziendale e di storia economica. Tuttavia, l'internazionalizzazione dei settori aziendali e di storia economica non è affatto trascurabile: complessivamente quasi il 40% dei lavori presentati nell'area aziendale sono pubblicati in lingua straniera; tra gli articoli su rivista, i 2/3 sono pubblicati in lingua straniera.

Infine, è interessante osservare la distribuzione dei lavori per anno di pubblicazione e SSD (Tab. 1.8). Complessivamente, si nota un aumento degli articoli su rivista (dal 58% del 2004 al

63% del 2010) e dei capitoli di libri (dal 19 al 23%) e una riduzione piuttosto consistente della percentuale di monografie (dal 18 al 9%). Le dinamiche sono alquanto differenziate tra settori. In alcuni si nota un forte aumento del numero di lavori pubblicati su riviste. Ad esempio, nel SECS-P/01 la percentuale di articoli su rivista cresce dal 69% del 2004 al 80% del 2010; nel settore SECS-P/11, pur con oscillazioni, la percentuale cresce dal 40 al 52%. In altri settori si nota invece una sostanziale stabilità della percentuale di articoli su rivista, o anche una diminuzione. Allo stesso tempo, vi è una forte diminuzione della percentuale di monografie, particolarmente nelle scienze aziendali.

L'eterogeneità dell'area si riflette anche nell'organizzazione della ricerca all'interno dei singoli atenei. Complessivamente, dopo la riorganizzazione dipartimentale della Legge 240, ricercatori dell'area 13 sono presenti in 354 dipartimenti; 74 di questi dipartimenti hanno conferito tra 10 e 30 lavori, 56 tra 30 e 100 lavori e 51 oltre 100 lavori. In media, in ciascun ateneo i docenti dell'Area 13 sono quindi presenti in quattro dipartimenti distinti, anche se spesso si tratta di presenze di uno o due docenti. Prima della Legge 240 la frammentazione era maggiore. Docenti dell'area 13 erano presenti in 364 dipartimenti; tra questi, 55 avevano conferito tra 10 e 30 lavori, 103 tra 30 e 100 lavori e 33 più di 100.

La notevole eterogeneità della distribuzione dei docenti nei dipartimenti rende di particolare interesse anche i risultati raggiunti dai singoli dipartimenti, oltre che quelli degli atenei nel loro complesso, particolarmente per quanto riguarda gli atenei di maggiori dimensioni.

2 La valutazione dei lavori

Per la valutazione dei lavori di ricerca, il GEV ha adottato sia indicatori bibliometrici sia la valutazione peer review. La procedura ha seguito strettamente quella indicata nei criteri di valutazione pubblicati il 29 febbraio 2012 (cfr. Appendice B).

Tutti i lavori di ricerca non pubblicati su riviste (monografie, capitoli di libro, atti di congresso, ecc.) e tutti gli articoli pubblicati su riviste non classificate dal GEV sono stati valutati in peer review. Gli articoli pubblicati su riviste classificate dal GEV sono stati valutati in base agli indicatori bibliometrici scelti e al numero di citazioni ricevute nel 2004-2010 nella banca dati WoS.

Un campione casuale di articoli pubblicati su riviste classificate dal GEV è stato valutato anche con la peer review.

2.1 La classificazione delle riviste

Per la classificazione delle riviste, il GEV ha preliminarmente costruito una banca dati comprensiva della gran parte delle riviste di interesse per l'Area 13 e relativamente omogenea tra i diversi settori dell'Area. La procedura di classificazione delle riviste è descritta in dettaglio nell'Appendice C (Classification, Imputation of Missing Values and Ranking of Journals) e qui richiamata in sintesi.

L'elenco delle riviste considerate include le riviste ISI-WoS riconosciute come attinenti all'Area. Alle riviste selezionate sulla base delle Subject Category WoS, sono state aggiunte le riviste classificate a partire all'archivio U-Gov dell'Area 13 fornito dal Cineca, in cui almeno un ricercatore dell'Area 13 ha pubblicato almeno un articolo nel 2004-2010.

Le riviste sono state inizialmente classificate in 4 sotto-aree di ricerca: Economia Aziendale, Economia, Statistica e Storia Economica. Le riviste di carattere generale (come Nature e Science) sono state classificate a parte. Per tutte le riviste, è stato raccolto un indicatore bibliometrico su Google Scholar (h-index riferito al periodo 2004-2010).

L'elenco preliminare delle riviste è stato pubblicato il 30 aprile 2012 ed è rimasto a disposizione della comunità scientifica per controlli e integrazioni fino ad agosto 2012. I dati

mancanti relativi all'Impact Factor a 5 anni (IF5) e all'Article Influence Score (AIS) sono stati imputati sulla base di h-index, disponibile per tutte le riviste. Il GEV ha poi classificato le riviste per ciascuna sotto-area di classificazione in quattro classi di merito (E, B, A, L) sulla base dei percentili (20/20/10/50) della distribuzione di IF5 e AIS, utilizzando per il ranking finale delle riviste il valore massimo dei due ranking. Per quel che riguarda le riviste italiane, come previsto dai criteri, il ranking è stato integrato aggiungendo (in classe "Accettabile") un numero di riviste tale per cui almeno 20 di esse risultino almeno in classe "Accettabile".

Il 4 settembre 2012 è stato pubblicato sul sito ANVUR un documento che illustra la scelta del criterio di imputazione e degli indicatori bibliometrici, la lista riviste e il programma "Stata" utilizzato per l'imputazione dei dati mancanti.

2.2 La valutazione bibliometrica

Come previsto dai criteri di valutazione, la classe di merito finale di ciascun articolo dipende dalla classe di merito iniziale della rivista (calcolata sulla base di indicatori bibliometrici) e, solo per gli articoli con un "numero significativo di citazioni", dal numero di citazioni ricevute da ciascun articolo in WoS. La soglia che identifica gli articoli con un "numero significativo di citazioni" è stata determinata in 5 citazioni annue. Tali articoli sono stati classificati nella classe di merito successiva a quella iniziale. A causa della mancanza di dati affidabili per l'analisi delle citazioni dei singoli articoli, nessuna correzione è stata applicata per gli articoli pubblicati in riviste "non WoS". Nessuna "retrocessione" è stata applicata alla classe iniziale di merito di un lavoro di ricerca.

Come descritto in dettaglio in Appendice D (Analisi delle citazioni), vi è una forte correlazione tra ranking basato sugli indicatori bibliometrici delle riviste (IF5 e AIS), e numero di citazioni. Dei 614 articoli con almeno 5 citazioni complessive nel 2004-2010, 513 sono in classe A (84%), 77 in classe B, 17 in classe C e solo 7 in classe D. Dei 93 articoli con almeno 5 citazioni medie annue, 87 sono pubblicati in riviste già classificate in classe A sulla base degli indicatori bibliometrici delle riviste e 6 in classe B; in seguito all'analisi delle citazioni, questi 6 articoli sono stati quindi classificati in classe A. In pratica, quindi l'effetto di questa correzione sulla valutazione complessiva è trascurabile.

Per verificare la validità della valutazione bibliometrica è stato estratto un campione casuale di articoli pubblicati su riviste stratificato per sotto-aree; gli articoli estratti sono stati valutati con il metodo della peer review. Il confronto tra valutazione bibliometrica e peer review è descritto in

dettaglio nell'Appendice A di questo Rapporto e qui richiamato in sintesi, cfr. Appendice A (Comparison between peer review and bibliometric evaluation in Area 13).

Il campione casuale comprende il 10% degli articoli su riviste di Economia, Economia Aziendale e Statistica e il 25% degli articoli su riviste di Storia Economica, per un complesso di 590 osservazioni. Il campione è stato estratto a settembre 2012, prima dell'inizio della peer review.

Gli articoli selezionati sono stati valutati in peer review con le stesse modalità di tutti gli altri lavori in peer review. In particolare, sono stati collocati in una classe di merito da due revisori esterni al GEV scelti indipendentemente da due membri GEV. In caso di discordanza di valutazione dei due revisori, il Gruppo di Consenso ha provveduto alla valutazione di sintesi.

Il confronto tra la distribuzione bibliometrica nelle quattro classi di merito e le valutazioni peer review di sintesi mostra che la valutazione bibliometrica tende a classificare in classe Eccellente un maggior numero di articoli rispetto alla peer review. Viceversa, la peer review classifica con maggiore frequenza gli articoli in classe B rispetto all'analisi bibliometrica. La frequenza di attribuzione alle classi Accettabile e Limitato è simile con entrambi i metodi di valutazione. Complessivamente i due metodi di valutazione danno la stessa valutazione nel 53% dei casi e nel 89% dei casi differiscono al più di una classe.

Il confronto tra valutazione bibliometrica e peer review è stato condotto considerando due criteri: (1) il grado di concordanza tra valutazione bibliometrica e valutazione peer, misurato con l'indice kappa di Cohen; (2) la verifica di differenze sistematiche tra le medie campionarie della valutazione bibliometrica e peer review.

Per quanto riguarda il primo criterio, l'indice di concordanza kappa tra valutazione bibliometrica e valutazione peer review, calcolato con pesi lineari o con pesi non lineari dettati dalla VQR (1, 0.8, 0.5, 0) è pari a 0.54 per tutto il campione ed è statisticamente diverso da zero al livello del 1%. Per quanto riguarda il secondo criterio, la differenza tra le medie delle due valutazioni non è statisticamente diversa da zero.

Nel complesso del campione i risultati del confronto indicano dunque che i due metodi di valutazione forniscono una buona concordanza; non si evidenziano, dal punto di vista statistico, distorsioni significative. Non mancano peraltro differenze tra aree di ricerca, come evidenziato dal confronto tra i due metodi di valutazione (cfr. Appendice A). Il confronto nel complesso soddisfacente tra analisi bibliometrica e peer review ha confermato la fiducia nelle valutazioni

bibliometriche delle riviste e portato alla decisione di utilizzare la valutazione bibliometrica per tutti gli articoli pubblicati sulle riviste censite.

2.3 I revisori esterni

L'elenco dei revisori è stato costruito considerando standard di qualità scientifica, impatto sulla comunità scientifica internazionale, esperienza nella valutazione, competenza nelle rispettive aree di valutazione e copertura delle aree di ricerca dell'Area 13. Per limitare conflitti di interesse, il GEV ha cercato di coinvolgere, per quanto possibile, revisori di istituzioni straniere e di enti di ricerca non sottoposti alla valutazione.

Dal punto di vista operativo, il Presidente e i Coordinatori dei tre sub-GEV hanno inizialmente consultato la lista dei revisori di Area 13 del CINECA e l'elenco dei revisori che si sono autonomamente proposti attraverso il sito del CINECA; hanno inoltre chiesto ai membri GEV e alle società scientifiche di proporre, indipendentemente gli uni dagli altri, un numero minimo di nominativi (15-20) per le subject categories di competenza. Ogni nominativo proposto è stato accompagnato da informazioni sulla ricerca scientifica del potenziale revisore fornite sulla base di una scheda condivisa contenente le 3 migliori pubblicazioni, le subject categories di competenza e l'indice h delle citazioni.

Grande attenzione è stata posta all'anonimato dei revisori, sia nella fase di predisposizione dell'elenco dei revisori, sia nella fase operativa della valutazione. I singoli membri GEV hanno potuto consultare, per scegliere i revisori, solo l'elenco dei revisori relativi alle subject categories di propria competenza.

La lista dei revisori è stata integrata durante il processo di valutazione sulla base delle necessità emerse a valle della trasmissione dei lavori da parte delle strutture, quando sono emerse aree di competenza poco coperte dai revisori inizialmente contattati, la necessità di integrare il numero di revisori per subject categories in cui sono stati presentati molti lavori o sostituire revisori non più disponibili o che si sono resi disponibili a valutare un numero limitato di lavori. Le integrazioni dei revisori sono avvenute su proposta dei membri GEV e previa approvazione, caso per caso, da parte del Presidente e dei coordinatori sub-GEV, con la stessa procedura utilizzata per la composizione della lista di partenza.

Complessivamente, il GEV ha utilizzato 610 revisori esterni. Per ciascuno dei 3 sub-GEV quasi il 50% dei revisori è affiliato a un'istituzione straniera; circa ¼ dei revisori lavora nel Regno Unito o negli Stati Uniti e il restante quarto in altri paesi. Complessivamente l'80% dei

revisori è di genere maschile. La media complessiva dell'indice "h" dei revisori è 12, con valori leggermente più elevati per l'area statistica e leggermente inferiori per l'area aziendale.

In media, ciascun revisore ha valutato 17 lavori; la media è più elevata nell'area aziendale (23), che ha registrato anche il maggior numero di lavori in peer review. Il punteggio medio (normalizzato tra 0 e 1) attribuito dai revisori è 0,41, con valori molto simili per i tre sub-GEV (compresi tra 0,40 e 0,44). Anche la deviazione standard del punteggio attribuito dai revisori è simile nelle tre aree.

2.4 La valutazione peer

La valutazione peer è descritta in dettaglio nell'Appendice E di questo Rapporto e qui richiamata in sintesi (Appendice E: La valutazione peer review nell'Area 13).

Il GEV ha valutato in peer review complessivamente 6.046 lavori, che rappresentano il 54% dei lavori complessivamente valutati. Essi comprendono:

- tutti i lavori non pubblicati su riviste (monografie, capitoli di libro, atti di congresso, ecc.);
- tutti i lavori pubblicati su riviste non classificate dal GEV nell'elenco pubblicato sul sito ANVUR il 4 settembre 2012;
- un campione casuale stratificato per sotto-area di ricerca dei lavori pubblicati su riviste classificate dal GEV (590 articoli su rivista).

Ciascun lavoro da valutare in peer review è stato assegnato a uno dei tre sub-GEV in base al SSD e poi ripartito tra i membri del GEV in base all'argomento principale della ricerca. All'interno di ciascuno dei tre sub-GEV (Economia, Economia Aziendale, Statistica), sono state definite 33 aree di ricerca (subject categories) per facilitare la classificazione e la successiva assegnazione dei lavori ai membri GEV e ai revisori. La gestione della peer review è avvenuta in modo del tutto simile a quella di una rivista, con la differenza importante che ciascun lavoro è stato preso in carico da due membri GEV, e non da un solo editor come avviene normalmente per le riviste scientifiche.

I Coordinatori sub-GEV hanno assegnato ciascun lavoro a due membri dello stesso sub-GEV che non sono stati rivelati l'uno all'altro (ad eccezione di quanto notato più avanti per quanto riguarda i lavori dei membri del GEV). In altre parole, i due membri GEV responsabili del lavoro non conoscevano l'identità reciproca fino al termine della valutazione. Inoltre, ciascuna subject category è stata affidata ad almeno quattro diversi membri GEV (talvolta a più di quattro); anche

in questo caso l'identità dei membri GEV non era nota. L'assegnazione dei lavori ai membri GEV ha evitato tutti i conflitti di interesse con gli autori e la sede di ricerca degli autori. Sono stati inoltre esclusi casi di potenziali conflitti di interesse per progetti di ricerca comuni e affiliazioni precedenti dei membri GEV con la sede degli autori.

I lavori dei membri GEV e del Presidente GEV sono stati valutati seguendo le regole previste dalla VQR (Linee Generali per i Gruppi di Esperti della VQR, Sezione 5). In particolare, il Presidente ha assegnato a revisori esterni i lavori presentati dai componenti del GEV e il Coordinatore della VQR ha gestito i lavori del Presidente del GEV. Va evidenziato che questi sono stati gli unici casi di lavori assegnati a un solo membro GEV (il Presidente o il Coordinatore della VQR).

In media, ciascun membro GEV è stato responsabile di 338 lavori (si ricordi che ciascun lavoro in peer review è stato assegnato a due membri GEV). Il numero è maggiore nell'area aziendale, che ha registrato anche il numero maggiore di lavori valutati in peer review e minore nei settori economici e statistici in cui il numero di articoli pubblicati su riviste è maggiore. I membri GEV hanno effettuato una prima valutazione interna dei lavori per individuare i "lavori non valutabili" (lavori appartenenti a tipologie escluse dalla VQR, con allegati mancanti o non adeguati per la valutazione, pubblicati in anni precedenti o successivi al periodo di riferimento della VQR, lavori non pubblicati, lavori senza alcun contenuto di ricerca come libri di esercizi per il triennio, ecc.) e i lavori di qualità chiaramente limitata.

I lavori sono stati valutati internamente se entrambi i membri GEV – con scelta del tutto indipendente – avevano valutato i lavori come stabilito nei criteri di valutazione, e cioè "nettamente inferiore al 50% della scala di valore condivisa dalla comunità internazionale". Complessivamente, il GEV ha valutato al proprio interno con questa modalità (quindi in classe L) 653 lavori (il 10,4% dei lavori complessivamente valutati in peer review e il 5,2% dei lavori complessivi) per cui vi era consenso unanime.

In tutti gli altri casi i lavori sono stati inviati alla valutazione esterna. In particolare, i due membri GEV che avevano in carico ciascun lavoro hanno scelto ciascuno un revisore esterno in maniera indipendente (anche quando uno dei due membri aveva valutato preliminarmente il lavoro in classe Limitata). La scelta dei revisori è stata fatta evitando conflitti di interesse tra revisori e autori e tra revisori e sede di ricerca degli autori, anche in questo caso seguendo le regole della VQR. L'indipendenza dei revisori è stata garantita prestando attenzione alla sede di ricerca dei revisori, alle collaborazioni di ricerca e, ove possibile, alla nazionalità. Nei casi in cui

il revisore abbia evidenziato la presenza di conflitti di interessi non noti al GEV, la non competenza sullo specifico lavoro e la mancata disponibilità per motivi di tempo, il membro GEV che aveva proceduto all'assegnazione ha provveduto a modificarla e assegnare nuovamente il lavoro a un altro revisore.

I revisori hanno fornito una valutazione qualitativa dei lavori di ricerca utilizzando una scheda di valutazione predisposta dal GEV e costituita da tre domande tendenti a valutare le caratteristiche di rilevanza, originalità e internazionalizzazione previste dal bando e da un campo libero (la scheda è riportata in Appendice E). Il punteggio di sintesi è stato ottenuto sommando i punteggi delle tre risposte e attribuendo una classe di merito complessiva. Dato l'elevato numero di revisori stranieri, la scheda di valutazione è stata predisposta in inglese per tutti i lavori valutati.

Una volta ricevute le schede di valutazione dei valutatori, il GEV ha proceduto alle valutazioni di sintesi formando 6.047 Gruppi di Consenso, costituiti, per ciascun lavoro, dai due componenti del GEV che avevano assegnato il lavoro ai revisori esterni. I Gruppi di Consenso hanno definito la valutazione di sintesi all'unanimità oppure ricorrendo (in casi molto limitati) a un terzo revisore e sono stati integrati, solo in caso di conflitto di valutazioni tra i componenti del Gruppo di Consenso, dai Coordinatori sub-GEV o dal Presidente del GEV. Grande attenzione da parte dei Gruppi di Consenso e dei Coordinatori dei sub-GEV è stata posta all'esame di lavori con pareri discordi dei revisori, particolarmente nei casi di forte disaccordo tra revisori esterni. I casi di conflitto di valutazione di sintesi per cui è stato necessario integrare i Gruppi di Consenso sono stati complessivamente 34 su 6.046 lavori valutati. Il ricorso al terzo revisore esterno è stato limitato a soli 3 casi, grazie all'ampio consenso sulle valutazioni di sintesi nei Gruppi.

In alcune aree di ricerca (ad esempio, storia economica), si è verificato un grado di discordanza maggiore tra i due revisori esterni, relativamente ad altre aree, con giudizi di alcuni revisori particolarmente severi o generosi. In alcuni casi le differenze di opinione tra i due revisori riflette anche diverse visioni e approcci alla disciplina, oltre che giudizi di qualità differenti. In questi casi, i Gruppi di Consenso sono stati particolarmente utili ed hanno permesso di tenere conto, in misura diversa, dei pareri dei revisori esterni, per dare maggiore importanza al parere del revisore più competente nell'area di ricerca.

2.5 I lavori "non valutabili"

La definizione dei lavori "non valutabili" è stata particolarmente complessa ed ha richiesto molto tempo, per tre motivi. In primo luogo, il bando VQR prevede linee guida di carattere generale

circa la tipologia e l'ammissibilità dei lavori ed è stato interpretato in modo diverso dai singoli atenei e ricercatori. In secondo luogo, in alcuni casi (marginali) non è facile rendersi conto se la tipologia di pubblicazione sia una rivista, una monografia o altro, se un lavoro sia pubblicato o meno o quale sia l'anno di pubblicazione; la classificazione è dunque soggetta a margini di errore. In terzo luogo, non conoscendo ex-ante la tipologia di lavori presentati alla VQR, i Criteri stessi di valutazione del GEV non avevano delineato in maniera esauriente e sufficientemente chiara i casi in cui un lavoro sarebbe stato considerato "non valutabile". Questo potrebbe avere generato incertezze negli atenei a tra i ricercatori circa le tipologie "ammissibili".

I lavori individuati come "non valutabili" sono stati prima di tutto evidenziati da decisioni unanimi dei Gruppi di Consenso; successivamente sono stati esaminati dai Coordinatori sub-GEV e dal Presidente. Le motivazioni sono state inviate alle strutture per consentire le opportune verifiche e controdeduzioni. Dopo aver esaminato le controdeduzioni, il GEV ha considerato non valutabili 51 lavori.

Il GEV ha cercato di ridurre al minimo i casi di ambiguità; rientrano in questa tipologia, ad esempio, i lavori inviati più di una volta da una stessa struttura, quelli pubblicati in date diverse dal 2004-2010, ecc. Nel caso ambiguità fossero ancora presenti, sono state risolte nel senso di ammettere i lavori alla valutazione. Le principali motivazioni per l'esclusione sono: assenza di allegati e/o documenti adeguati per la valutazione, pubblicazione in anni precedenti o successivi al periodo di riferimento della VQR, mancanza di ISBN, duplicati di lavori valutati, tipologie non assimilabili a lavori di ricerca (ad esempio, libri di esercizi). Un elenco dettagliato è riportato in Appendice E.

2.6 Criticità emerse e suggerimenti per un nuovo ciclo di valutazione

Nella riunione plenaria del 13 maggio 2013, oltre ad aver analizzato i dati di sintesi della VQR, il GEV ha discusso le criticità emerse durante il processo valutazione, anche per fornire indicazioni per un futuro ciclo di valutazione. Di seguito si riportano alcuni dei temi più rilevanti emersi nella discussione.

Il GEV ha evidenziato che occorre chiarire meglio, sia in fase di bando sia in fase di elaborazione dei criteri di Area, le tipologie di lavori "non valutabili", ad esempio per quanto riguarda le pubblicazioni di dipartimento e i working papers (anche se dotati di ISSN o ISBN), gli abstract presentati ai congressi, i libri a esclusivo uso didattico, ed altre tipologie. Peraltro, andrebbero sottoposte a revisione le penalizzazioni dei lavori mancanti: con le regole attuali della VQR (centrate sulla presenza di ISSN e ISBN in riviste e monografie), è possibile che in un

futuro esercizio di valutazione si possa verificare una diminuzione dei lavori mancanti e, allo stesso tempo, una proliferazione di nuove riviste e pubblicazioni di bassa qualità.

E' opportuno riflettere sulle griglie previste dal bando VQR (20/20/10/50) che, dando un peso molto simile alle valutazioni "eccellenti" e "buone" (rispettivamente, 1 e 0,8), non consentono di valorizzare in modo adeguato i lavori migliori. Ad esempio, alcuni membri GEV considerano una griglia (10/20/20/50), con pesi maggiormente differenziati tra le prime due fasce, più adeguata ad individuare aree di eccellenza nella ricerca.

Andrà studiata la robustezza degli indicatori bibliometrici utilizzati dal GEV, in particolare per quanto riguarda la procedura di imputazione dei valori mancanti di IF5 e AIS, la definizione delle sotto-aree di ricerca e la classificazione delle riviste nelle sotto-aree, l'analisi delle citazioni ricevute dai singoli articoli. Alcuni membri GEV hanno espresso riserve sulla valutazione delle riviste italiane, ritenendo troppo generosa la promozione di alcune di queste riviste in fascia C (la procedura è descritta nell'Appendice C).

Tutti i membri GEV hanno sottolineato che in futuro sarà necessario migliorare la piattaforma del CINECA utilizzata per la gestione dei lavori, perché i tempi della valutazione e la disponibilità dei revisori esterni a collaborare con la VQR dipendono in modo cruciale dall'efficienza della piattaforma informatica. Pur apprezzando il lavoro dei tecnici del CINECA, non si può non rilevare che i tempi della valutazione, almeno per quanto riguarda l'Area 13, sono stati notevolmente influenzati dai ritardi del CINECA, in particolare per quanto riguarda la messa in opera della piattaforma per i Gruppi di Consenso e per la convalida finale delle valutazioni.

Anche se la distribuzione delle competenze all'interno del GEV riflette la composizione dei SSD all'interno dell'Area 13, la distribuzione del carico di lavoro tra i membri GEV dipende non solo dal numero di lavori attesi in ciascun SSD, ma anche dalla prevalenza o meno della valutazione bibliometrica nelle singole aree di ricerca. Il suggerimento è di poter integrare il GEV nel corso del processo di valutazione sulla base delle competenze che si rivelano necessarie per la gestione della valutazione. Andrebbe inoltre migliorata l'informazione a disposizione dei membri GEV sui revisori esterni, ad esempio rendendo disponibile, all'interno della piattaforma informatica, il curriculum e le migliori pubblicazioni dei revisori.

Alcuni membri GEV hanno sottolineato che sarebbe utile considerare per la valutazione anche altre informazioni sull'attività di ricerca dei singoli docenti (ad esempio, l'indice h). Molti

ricercatori infatti pubblicano ben più di tre lavori ogni sette anni, ma di ciò non si è tenuto conto nella VQR.

Infine, in futuro sarà necessario riflettere sulla regola del bando VQR, per cui un lavoro tra colleghi dello stesso dipartimento non può essere presentato due volte, mentre è ammesso alla valutazione un lavoro con un coautore di un'altra struttura. Ciò potrebbe indurre un incentivo a lavorare all'esterno dei dipartimenti, limitando le collaborazioni interne.

3 I lavori valutati

I lavori valutati di Area 13 sono stati complessivamente l1.941 rispetto ai 12.654 lavori attesi dall'Area. La percentuale di lavori valutati rispetto ai lavori attesi è quindi del 94% e quella di lavori mancanti del 6% (Tab. 3.1). La percentuale di lavori valutati è di oltre il 94% in tutti i SSD, ad eccezione di SECS-P/09 (89%) e SECS-S/06 (87%). Si noti che per 220 lavori inviati da personale di enti di ricerca o consorzi non è attributo il SSD del proponente; in questo caso la percentuale di lavori valutati raggiunge il valore minimo dell'Area (83%).

La Tab. 3.2 riporta la distribuzione del numero di lavori attesi dei 4.563 soggetti valutati. Secondo il bando VQR, soggetti valutati con doppia affiliazione (ateneo ed ente di ricerca) avrebbero dovuto inviare fino a 6 lavori. Dei 4.563 soggetti valutati, la gran parte (3.687) avrebbe dovuto inviare 3 lavori; complessivamente 811 soggetti valutati avrebbero dovuto inviare tra uno e due lavori. I soggetti valutati che afferiscono agli enti di ricerca o con doppia affiliazione sono complessivamente 65 (generalmente con più di 3 lavori da inviare), l'1,4% del personale di ricerca dell'Area.

La Tab. 3.3 riporta la distribuzione dei lavori valutati per GEV responsabile della valutazione. Dei 11.941 lavori valutati, 11.507 sono stati valutati dal GEV di Area 13 (96,3%), mentre 434 lavori sono stati valutati da altri GEV, in prevalenza dai GEV di Scienze matematiche (126 lavori), Scienze politiche e sociali (62 lavori), Scienze mediche (56 lavori) e Scienze storiche, filosofiche, pedagogiche e psicologiche (49 lavori).

La Tab. 3.4 indica che, complessivamente, il 53% dei lavori è stato valutato con peer review ed un'analoga percentuale con analisi bibliometrica. Il 6% dei lavori (corrispondente al campione casuale di articoli su rivista) è stato valutato sia con la peer review sia con la bibliometria. Nel caso di doppia valutazione, la valutazione finale del lavoro è stata quella bibliometrica.

Il dettaglio per settori della ripartizione tra valutazione peer e valutazione bibliometrica riflette il peso dei lavori pubblicati su rivista nei diversi settori, ed è di circa 2/3 nelle aree economiche e statistiche, e di circa 1/3 nelle discipline aziendali. Nel settore di Storia Economica solo l'11% dei 473 lavori è stato valutato con l'analisi bibliometrica. Ulteriori dettagli sulla valutazione peer review e le caratteristiche dei revisori esterni sono disponibili nell'Appendice E.

4 I risultati della valutazione

4.1 Sintesi dei risultati di area

Le tabelle che seguono presentano i risultati di sintesi della valutazione, con la distribuzione delle quattro classi finali di merito nel complesso dell'Area 13, per tipo di valutazione (peer o bibliometrica), per tipologia di lavoro (articolo su rivista, monografia, capitolo di libro, atto di congresso, ecc.), per ciascuno dei 3 sub-GEV e per ciascuno dei 19 SSD. La definizione di SSD e sub-GEV adottata nelle tabelle si riferisce al SSD del soggetto valutato, e non al SSD indicato per la valutazione dal soggetto valutato. Per questo motivo per i soggetti valutati degli enti di ricerca non è indicato il SSD e il sub-GEV; in ogni caso, i lavori sono stati valutati secondo l'indicazione di SSD fornita dal soggetto valutato e ripartiti all'interno del GEV secondo le medesime modalità di tutti gli altri lavori.

Sulla base del bando VQR, i singoli lavori ricevono un punteggio pari a 1, 0,8, 0,5 e 0 a seconda che il lavoro sia valutato Eccellente (20% superiore della scala di valore condivisa dalla comunità scientifica internazionale), Buono (60-80% della scala di valore), Accettabile (50-60% della scalda di valore) o Limitato (50% inferiore della scala di valore); ai lavori mancanti viene assegnato un punteggio pari a -0,5, ai lavori non valutabili è assegnato un punteggio -1, e in casi accertati di plagio o frode -2. I punteggi dei lavori di soggetti valutati che afferiscono a due strutture (ad esempio, al CNR e ad un ateneo) sono sommati per ciascun soggetto valutato e ripartiti in parti uguali tra le due strutture.

Indicando rispettivamente con niE, niB, niA, niL, niMIS, niNV e niPL il numero di lavori Eccellenti, Buoni, Accettabili, Limitati, Mancanti, Non Valutabili e Plagio della struttura i-esima, si ottiene il punteggio complessivo v_i della struttura *i* dell'Area 13 come:

$$v_i = n_{iE} + 0,8n_{iB} + 0,5n_{iA} - 0,5n_{iMIS} - n_{iNV} - 2n_{iPL}$$

Un primo indicatore di qualità della produzione scientifica è il punteggio medio della struttura *i*, $I_i = \frac{v_i}{n_i}$, dove $n_i = n_{iE} + n_{iB} + n_{iA} + n_{iL} + n_{iMIS} + n_{iNV} + n_{iPL}$. L'indice è pari a 1 nel caso in cui la struttura abbia presentato tutti i lavori attesi, e tutti abbiano ottenuto una valutazione eccellente.

Si noti che se una struttura presenta tutti i lavori attesi, e la distribuzione dei punteggi è uguale alla media della distribuzione internazionale, essa ottiene un punteggio medio pari a

$I = 0,2 + 0,8 \times 0,2 + 0,5 \times 0,1 = 0,41$.Ovviamente, un confronto dei risultati effettivi con questo valore di riferimento è valido solo se la valutazione peer e quella bibliometrica sono in grado di replicare esattamente, tramite giudizi soggettivi o analisi di citazioni, la "scala di valore condivisa" della ricerca internazionale. Il valore di riferimento va quindi trattato con molta cautela, ma è comunque un valore di confronto implicito nelle regole del bando VQR.

La Tab. 4.1 indica che il punteggio medio dei 12.654 lavori attesi dall'Area 13 è 0,32. Il punteggio medio è ottenuto da una percentuale di valutazioni "eccellenti" del 18%, di valutazioni "buone" del 12%, di valutazioni "accettabili" del 14% e di valutazioni "limitate" del 49%. Il punteggio riflette peraltro anche le penalizzazioni; queste rappresentano il 6% dei lavori attesi, e sono in larga misura da attribuire a lavori non conferiti e, in misura molto più limitata, ad un esiguo numero di lavori "non valutabili". Complessivamente il peso delle penalizzazioni riduce il punteggio medio di area del 13%: escludendo dal calcolo i lavori con penalizzazioni il punteggio medio sarebbe 0,37.

La Tab 4.1 registra differenze significative tra il punteggio dei 3 sub-GEV. In particolare, i settori economici e statistici registrano una valutazione superiore alla media (rispettivamente, 0,42 e 0,40), mentre l'area aziendale registra valori inferiori (0,18). La differenza emerge sia nel numero inferiore di valutazioni "eccellenti" del sub-GEV aziendale (8%), sia nel numero maggiore di lavori con giudizio "limitato" nella stessa area.

Il punteggio medio dei 265 lavori per cui non è attributo un particolare SSD (e quindi sub-GEV) è particolarmente basso (0,08). In questo gruppo di lavori, in cui pesa principalmente il CNR, il numero di giudizi "limitati è molto elevato (60%) e la percentuale di lavori con penalizzazioni è del 18%, un valore nettamente superiore alla media di Area.[2] Come si noterà più avanti, l'eterogeneità che si riscontra in Tab 4.1 è presente in tutte le tabelle che seguono, per singoli SSD, strutture valutate e dipartimenti.

La Tab. 4.2. evidenzia notevoli differenze anche tra SSD. Il punteggio medio oscilla infatti tra valori superiori a 0,5 per i settori SECS/P-05 (0,69) e SECS/S-02 (0,53) e valori inferiori a 0,15 per SECS-P/09 (0,11), SECS-P/11 (0,11) e SECS-P/07 (0,13). Si noti che forti differenze si

[2] Si noti che il fatto che non sia presente un SSD del soggetto valutato per gli enti di ricerca non significa che i lavori non siano stati valutati dai sub-GEV. Infatti la valutazione è avvenuta, salvo casi marginali descritti nell'Appendice E, sulla base del SSD indicato dal soggetto valutato per la valutazione.

verificano non solo tra SSD che appartengono a diversi sub-GEV, ma anche tra SSD che appartengono allo stesso sub-GEV; ad esempio, il SECS/P-05 ottiene una valutazione media elevata (0,69) mentre il SECS/S-03 e il SECS/S-05 registrano valutazioni medie (rispettivamente 0,29 e 0,27) inferiori sia a quella del sub-GEV di riferimento (0,40) sia a quella complessiva dell'Area 13 (0,32). E' da notare inoltre il punteggio largamente inferiore alla media degli enti di ricerca (ed in particolare del CNR), per i quali il SSD dei soggetti è indicato come "non disponibile".

La Tab. 4.3 riporta il punteggio medio per anno di pubblicazione del lavoro valutato e sub-GEV. Il confronto è interessante, perché la dinamica temporale potrebbe riflettere sia una diversa qualità dei lavori, sia una distorsione della peer review e dell'analisi bibliometrica rispetto alla data di pubblicazione dei lavori; ad esempio, i revisori potrebbero essere indotti a considerare meno innovativo un lavoro pubblicato nel 2004 piuttosto che nel 2010, oppure gli indicatori bibliometrici potrebbero penalizzare le riviste più recenti. In nessuno dei sub-GEV, tuttavia, si evidenziano dinamiche temporali significative, nel senso che il punteggio medio tende a oscillare intorno al punteggio medio del sub-GEV per ciascuno degli anni compresi tra il 2004 e il 2010. Il tema andrebbe certamente approfondito con maggiori strumenti di indagine; tuttavia è lecito affermare che l'evidenza non avvalora l'ipotesi che la valutazione dell'Area 13 abbia premiato o penalizzato i lavori sulla base della data di pubblicazione.

Le differenze dei risultati della valutazione bibliometrica e della peer review sono riportate nella Tab. 4.4 (si noti che la tabella esclude i lavori con penalizzazioni). I lavori valutati con l'analisi bibliometrica hanno ricevuto un punteggio medio pari a 0,57, circa tre volte superiore a quello dei lavori valutati in peer review (0,14). Inoltre, mentre nella peer review solo l'1% dei lavori ha ricevuto un giudizio "eccellente" e il 76% dei lavori ha ricevuto una valutazione "limitata", nella valutazione bibliometrica la percentuale di lavori "eccellenti" è del 35% e quella di lavori "limitati" è del 32%.

La Tab. 4.5 evidenzia differenze di valutazioni per tipologie di pubblicazione e indica che gli articoli su rivista hanno ricevuto mediamente valutazioni molto superiori (0,52) a quelle dei contributi in volume (0,14), delle monografie (0,12) e degli atti di congresso (0,08). Le Tab. 4.6 indica che differenze di valutazione tra tipologie di lavori sono presenti anche per i singoli SSD e conferma che in tutti i settori gli articoli su rivista ricevono una valutazione migliore delle altre tipologie.

Poiché il confronto tra analisi bibliometrica e peer review condotto sul medesimo insieme di lavori (il campione casuale degli articoli su rivista) non registra differenze sistematiche, se ne deduce che le differenze di punteggio tra tipologie di lavori sono verosimilmente da attribuire a differenze nella qualità dei lavori valutati, piuttosto che a differenze nelle modalità di valutazione. Tuttavia, sarebbe errato concludere che, ad esempio, gli articoli su rivista (valutati in larga misura con l'analisi bibliometrica) sono "di per sé" di qualità superiore a monografie, contributi in volume o atti di congresso (valutati esclusivamente in peer review).

La spiegazione più probabile delle differenze che emergono nelle Tab. 4.5 e 4.6 è che molti ricercatori scrivono sia monografie, sia capitoli di libri, sia articoli su rivista. Al momento di selezionare i propri lavori, i migliori ricercatori hanno scelto di inviare alla VQR prevalentemente articoli su riviste. Ciò riflette la tendenza prevalente nell'Area a pubblicare i propri migliori lavori su riviste internazionali, piuttosto che su libri, data la maggiore visibilità e diffusione delle riviste stesse. E' probabile anche che molti ricercatori dell'Area 13 assegnino alle monografie e ai libri una finalità diversa rispetto agli articoli su rivista, cioè di sintesi, documentazione e analisi di casi, piuttosto che di ricerca vera e propria (ciò non si applica ai lavori del SECS-P/12, dove vi è maggiore necessità di un libro per esporre i risultati di una vera ricerca storica). D'altronde, la qualità stessa dei lavori pubblicati su riviste si giova anche di uno sforzo maggiore di approfondimento e perfezionamento del lavoro dovuto al processo di peer review delle riviste, che spesso impone due o più cicli di revisione. Ne segue che per molti SSD le monografie (o altre tipologie di pubblicazione diverse da articoli su rivista) valutate dal GEV in peer review non sono rappresentative di tutte le monografie scritte dai soggetti valutati nel corso del 2004-2010.

4.2 La valutazione delle strutture

Per il ranking delle strutture si utilizza nel seguito un indicatore di qualità della ricerca che è indipendente dalla numerosità dei soggetti valutati della struttura. Si utilizza inoltre l'indicatore IRAS1 definito nel decreto ministeriale e nel Bando VQR; l'indicatore tiene conto sia della qualità della ricerca, sia della numerosità dei soggetti valutati della struttura. Gli indicatori si riferiscono a una struttura e un'area particolare e non sono quindi influenzati da eventuali difformità nelle valutazioni inter-area. Fatti salvi i 434 lavori di soggetti che afferiscono all'Area 13, ma che sono stati valutati dai GEV responsabili di altre aree (il 3,6% del totale), i ranking sono quindi ottenuti da una valutazione omogenea per tutta l'Area 13. Poiché gli stessi strumenti di valutazione sono stati utilizzati per ciascun sub-GEV, i risultati di tutte le strutture e di tutti i dipartimenti dell'area si possono ordinare secondo lo stesso metro di valutazione.

Indicando con $N = \sum_i n_i$ e $V = \sum_i v_i$, il rapporto $R_i = \frac{\frac{v_i}{n_i}}{\frac{V}{N}}$ indica la posizione della struttura rispetto alla media di area. Se R_i =1 la qualità della struttura coincide con la media di area; se R_i >1, la struttura ha una qualità superiore alla media di area; se invece R_i <1, la qualità è inferiore alla media.

L'indicatore $IRAS1_i = \frac{v_i}{V} = R_i \times \frac{n_i}{N}$ riflette sia la valutazione qualitativa, sia la dimensione della struttura (misurata dal numero di lavori attesi). L'indicatore può essere quindi utilizzato per una ripartizione delle risorse che pesa ciascun lavoro atteso per la sua qualità. Infatti, se in tutte le strutture i lavori ricevessero la stessa valutazione media (lo stesso R_i) l'indicatore IRAS1 rifletterebbe soltanto il numero relativo di lavori attesi, e quindi il peso numerico della struttura nell'area. Per costruzione, la somma di IRAS1 per tutte le strutture è uguale a 1, cioè $\sum_i IRAS1_i = 1$.

Nel seguito si presentano i risultati della valutazione per ciascuna struttura, la percentuale di lavori eccellenti, buoni, limitati e penalizzati, la numerosità relativa della struttura (n_i/N), il punteggio complessivo (v_i) e i due indici R_i e $IRAS1_i$.

Poiché la VQR non può diffondere i risultati delle valutazioni dei singoli lavori, le tabelle prendono in considerazioni solo strutture con almeno 10 lavori attesi (e quindi conferiti da almeno 3 soggetti valutati). Le tabelle evidenziano separatamene i risultati del CNR, l'unico ente di ricerca che in Area 13 supera la soglia dei 18 lavori attesi (corrispondente a tre soggetti valutati) e dell'unico consorzio presente in area 13 (CMCC di Venezia).

La Tab. 4.7a rappresenta il risultato più importante della valutazione VQR, con i ranking delle 75 strutture (di cui due enti di ricerca) ordinate secondo la dimensione (strutture medio-piccole, con 10-150 lavori attesi, e strutture grandi, con oltre 150 lavori attesi) e il punteggio medio ricevuto (l'indicatore I). Per motivi di affidabilità statistica degli indicatori, è opportuno considerare con cautela i valori degli indicatori di piccole strutture (orientativamente, con numero di lavori attesi inferiori a 30). Per costruzione, la somma di IRAS1 è pari a 1; in tabella tuttavia la somma è inferiore a 1 perché non sono riportati i valori delle strutture con meno di 10

lavori attesi. Il calcolo di, R, n/N ed IRAS1 è calcolato tra insiemi di strutture omogenee: Università, Enti di ricerca e Consorzi.

Tra le strutture grandi (oltre 150 lavori attesi) occupano le prime cinque posizioni le Università di Padova, Bocconi, Bologna, Siena e Roma Tor Vergata; tra le medio-piccole (10-150 lavori attesi, le prime cinque sono Lucca – IMT, S.Anna di Pisa, Milano Politecnico, Bolzano e Torino Politecnico. Indipendentemente dalla dimensione, emerge comunque una forte dispersione dell'indicatore I, che supera il valore di 0,5 in 9 strutture, di cui due di grandi dimensioni (Padova e Bocconi). All'estremo opposto, 12 strutture, registrano un punteggio inferiore a 0,2: tra queste, cinque grandi strutture (Genova, Messina, Bari, Salento e il CNR) e sette strutture piccole (Catanzaro, Tuscia, Reggio Calabria, Casamassima, Roma Mercatorum, Enna, Roma Marconi).

Una semplice analisi di correlazione non evidenzia un legame forte tra punteggio e dimensione della struttura; in altre parole, si registrano punteggi elevati sia per strutture di grandi dimensioni (come Padova e Bocconi) che per piccole strutture (come Pisa S. Anna), e viceversa. Anche se la dimensione della struttura non svolge, di per sé, un ruolo significativo nel punteggio medio, si evidenzia una relazione negativa tra dimensione delle strutture e percentuale di lavori con penalizzazioni; ciò è da interpretare nel senso che il mancato invio di lavori alla VQR pesa in modo leggermente maggiore nelle grandi strutture.

L'indicatore *R* riportato in tabella evidenzia in modo efficace i risultati delle strutture che ricevono valutazioni superiori alla media ($R>1$) o inferiori alla media ($R<1$). Infine, il confronto tra la numerosità relativa della struttura (n_i/N) e l'indicatore *IRAS1* evidenzia se una struttura "guadagna" o "perde" sulla base di una ipotetica graduatoria tiene conto della qualità della ricerca. Ad esempio, l'Università di Verona, con un punteggio medio pari a 0,34 e un valore di $R=1{,}01$, ha un risultato molto vicino alla media dell'area. Per questa ragione, il peso numerico, in termini di lavori attesi, è molto simile a quello corretto per la qualità (1,82% rispetto a 1,84%). Per altre strutture le differenze sono molto marcate. Per esempio l'Università di Padova ha un peso numerico di 2,84%, ma un indicatore *IRAS1* pari a 5,22%. Tra le strutture grandi, un caso opposto è quello dell'Università di Roma La Sapienza: il peso numerico è 5,25, che si riduce a 3,92 se si tiene conto della qualità della ricerca. Poiché la soglia dimensionale oltre la quale una struttura possa definirsi "media" o "grande" è in qualche misura arbitraria, la Tab. 4.7b riporta gli stessi dati della Tab. 4.7a senza distinguere per dimensione (i dati sono cioè ordinati solo secondo il voto medio). Nel formato excel la tabella consente anche di aggregare e ordinare i dati secondo soglie dimensionali diverse da quelle indicativamente prosposte.

La Tab. 4.8 riporta la stesse informazioni per i lavori dei soggetti valutati che afferiscono ai settori scientifici dei 3 sub-GEV; nel caso di valore mancante, si tratta dei soggetti valutati del CNR.[3] Per compattezza di presentazione, le graduatorie non sono ordinate per dimensione della struttura. Tuttavia, ciascuna tabella riporta il numero di lavori attesi di ciascuna struttura per ciascun sub-GEV; è quindi immediato ordinare i dati secondo che si tratti di strutture piccole, medie o grandi, scegliendo la soglia dimensionale più opportuna per il sub-GEV (o, di seguito, per il SSD).

Emergono anche in Tab. 4.8 forti eterogeneità nei ranking. La deviazione standard del punteggio oscilla tra 0,12 del sub-GEV delle discipline aziendali a 0,18 per il sub-GEV di statistica. Il punteggio ottenuto dalle migliori strutture (oltre il 90mo percentile della distribuzione) è almeno di 3 volte superiore a quello delle strutture che si collocano al di sotto del 10mo percentile della distribuzione. Per ciascun sub-GEV il punteggio ottenuto dalle prime strutture (oltre il 90mo percentile) è almeno del doppio superiore a quello delle strutture che hanno ricevuto le valutazioni meno positive (al di sotto del 10mo percentile).

Anche se vi è una notevole variabilità, i ranking tendono ad essere positivamente correlati tra i diversi sub-GEV. Ad esempio, tra le strutture presenti in tutti i ranking, Bocconi e Roma Luiss occupano uno dei primi 10 posti in graduatoria in ciascuno dei ranking; Padova occupa il primo e il terzo posto rispettivamente in statistica e nelle materie aziendali; molte università meridionali ottengono valutazioni inferiori alla media in ciascuna delle tre sotto-aree. In altri casi, tuttavia, la graduatoria per sub-GEV riportata in Tab. 4.8 evidenzia risultati diversi nelle singole sotto-aree.

La Tab. 4.9 considera ulteriori dettagli dei risultati della valutazione, riportando i punteggi ottenuti dai soggetti valutati di ciascun SSD. Si noti che molte delle strutture non sono presenti in tabella perché il numero di lavori attesi del SSD all'interno della struttura è inferiore a 10. Risultano quindi di non facile lettura i risultati dei SSD più piccoli, di cui sono riportati i risultati

[3] Le Tab. 4.8 e 4.9 riportano i dati sulle valutazioni per SSD e sub-GEV (gruppi di SSD): il numero di osservazioni relativo al CNR nelle Tab. 4.8 e 4.9 non coincide con quello della Tab. 4.7 perché i dati del CNR in tali tabelle sono disaggregati in base al SSD del soggetto valutato (e aggregazioni sub-GEV). Più in dettaglio, in Tab. 4.8 il CNR compare tre volte: per il sub-GEV di Economia (36 osservazioni), per il sub-GEV di Statistica (15 osservazioni) e nella riga n.a. (247 osservazioni). La riga n.a. contiene tutte le osservazioni prive del dato del SSD del soggetto valutato. Il totale delle osservazioni relative al CNR in tab 4.8 (298) è inferiore alle osservazioni del CNR in Tab. 4.7 (304) perchè sono state "oscurate" le righe con meno di 10 osservazioni. Lo stesso accade in Tab.4.9. Il confronto tra il punteggio in Tab. 4.7 e quello in Tab. 4.8 consente quindi di misurare approssimativamente il contributo alla qualità della ricerca del CNR da parte del personale con e senza il SSD del soggetto valutato.

solo per alcune strutture (ad esempio, dati relativi al settore SECS-P/05 sono riportati solo per le Università di Venezia, Bologna, Bocconi e Torino).

4.3 La valutazione dei dipartimenti

La VQR ha, tra i suoi compiti, anche di elaborare un ranking dei dipartimenti che possa essere utilizzato come strumento di informazione dagli organi decisionali delle strutture. I dipartimenti istituiti a valle della Legge 240 presentano tipologie diverse. Le più frequenti sono dipartimenti con docenti e ricercatori che appartengono tutti (o in gran parte) ad una stessa Area, ma sono assai diffusi anche dipartimenti più compositi, con docenti e ricercatori di due (o più) aree diverse. In entrambi i casi, occorre costruire indicatori di dipartimento a partire dalle valutazioni dei lavori associati ai soggetti valutati di ciascun dipartimento. Anche in questo caso, come per le strutture, è importante che il risultato finale non sia influenzato da difformità delle valutazione inter-area.

Di seguito si presentano gli indicatori per i dipartimenti calcolati con le medesime modalità delle strutture (l'indicatore IRD1 per i dipartimenti prende il posto di IRAS1). Si noti che gli indicatori si riferiscono esclusivamente ai ricercatori dell'Area 13. Pertanto, un dipartimento in cui tutti i membri appartengono all'Area 13 appare solo nel ranking dell'Area 13; invece, un dipartimento con soggetti valutati di aree diverse è presente nel ranking di due o più aree. Le informazioni sulle piante organiche dei dipartimenti sono state prese in considerazione se trasmesse al GEV entro il mese di maggio 2013.

In alcuni atenei un certo numero di lavori non è attribuito ad alcun dipartimento; questi lavori sono raggruppati nei codici "-999" (o "n.d."). I motivi per i valori mancanti sono vari: mancata trasmissione della pianta organica dei dipartimenti (pochissimi casi), docenti andati in pensione o trasferiti all'estero, che hanno quindi inviato lavori alla VQR ma non sono attualmente in servizio presso l'ateneo; riorganizzazione dell'ateneo in dipartimenti non ancora completata o non prevista.

Poiché la VQR non può diffondere i risultati delle singole valutazioni, anche in questo caso le tabelle prendono in considerazioni solo dipartimenti con almeno 10 lavori attesi (e quindi conferiti da almeno 3 soggetti valutati). Come per le strutture, le tabelle evidenziano anche gli istituti o dipartimenti degli enti di ricerca che superano la soglia dei 19 lavori attesi. Per motivi di affidabilità statistica, gli indicatori di dipartimenti con numero di lavori attesi compreso tra 10 e 30 vanno considerati con cautela.

Le Tab. 4.10a e 4.10b presentano i risultati della valutazione, per ciascuno dei 176 dipartimenti (di cui tre enti o consorzi) con almeno 10 lavori attesi.[4] Le informazioni contenute nelle due tabelle sono identiche, ma cambia l'ordine di presentazione dei dati. La Tab. 4.10a è ordinata secondo il punteggio medio ricevuto dai dipartimenti, mentre la Tab. 4.10b è ordinata secondo il nome dell'ateneo e, all'interno di questo, il punteggio medio dei dipartimenti. La variabilità degli indicatori di qualità è assai ampia, con valori superiori a 0,7 per 10 dipartimenti e inferiori a 0,1 per oltre 20 dipartimenti, tra cui alcuni di grande dimensione. In alcuni dipartimenti il punteggio medio è negativo, perché le valutazioni positive non compensano le penalizzazioni (-0,5 per ogni lavoro non inviato, -1 per ogni lavoro "non valutabile").

La Tab. 4.10b mette in evidenza che una forte variabilità si registra anche all'interno dei singoli atenei, inclusi quelli che registrano le migliori posizioni complessive. A titolo di esempio, vi sono 9 dipartimenti con almeno 10 lavori attesi nell'Università Roma La Sapienza. Tra questi, il Dipartimento di Scienze Statistiche riceve le migliori valutazioni (0,49), mentre altri quattro Dipartimenti dello stesso Ateneo hanno valutazione inferiore a 0,12. In generale, differenze significative di valutazione tra dipartimenti dello stesso ateneo sono la regola piuttosto che l'eccezione.

Le Tab. 4.11a (ordinata per punteggio) e 4.11b (ordinata per struttura e punteggio del dipartimento) riportano i risultati della valutazione raggruppando i lavori dei soggetti valutati per i dipartimenti in vigore prima della riforma introdotta nella Legge 240.[5] Le valutazioni non sono direttamente rilevanti per gli scopi istituzionali della VQR, ma sono di sicuro interesse per considerazioni interne a ciascun ateneo sull'efficacia della riorganizzazione dipartimentale successiva alla riforma.

Per completare il quadro, le due tabelle successive forniscono ulteriori dettagli sulla valutazione raggruppando i lavori di ciascun dipartimento attivato a valle della Legge 240 per sub-GEV (Tab. 4.12) e SSD (Tab. 4.13).[6] Con le dovute eccezioni, i dipartimenti con presenza più consistente di docenti nelle aree economiche e statistiche tendono a ricevere valutazioni

[4] Il numero di osservazioni relativo al Dipartimenti del CNR (292) in Tab. 4.10a, 4.10b, 4.11a e 4.11b non coincide con quello della Tab. 4.7 (304) perché alcuni soggetti valutati del CNR afferiscono ad altri dipartimenti. Inoltre, poiché il numero di lavori attesi di questi gruppi di ricerca è inferiore a 19, i dati non sono riportati in tabella.
[5] Per il,CNR si utilizza la stessa definizione di dipartimento che in Tab. 4.8.
[6] Per il CNR, cfr. nota 4.

migliori dei dipartimenti in cui è prevalente la componente aziendale, riproducendo quindi su scala locale le differenze generali emerse nel confronto tra SSD e sub-GEV.

4.4 I risultati dei soggetti valutati

La distribuzione del punteggio di una singola struttura presentata nelle sezioni 4.1 e 4.2 con la percentuale di lavori eccellenti, buoni, accettabili o limitati offre un quadro analitico che evidenzia il contributo di ciascun tipo di lavori al punteggio complessivo della struttura (o dipartimento). Tuttavia, la statistica si riferisce ai singoli lavori e non ai singoli soggetti valutati, quindi non permette di capire quanti ricercatori hanno effettivamente contribuito a scrivere lavori eccellenti, buoni, ecc.

Ad esempio, supponiamo che un dipartimento costituito da 30 ricercatori abbia presentato complessivamente 90 lavori (salvo eccezioni analizzate di seguito, la VQR richiede 3 lavori per ciascun soggetto valutato), e che 30 di questi siano stati valutati come "eccellenti". Una possibilità è che ciascuno dei 30 ricercatori abbia scritto un lavoro eccellente, e che quindi la qualità sia diffusa in modo omogeneo tra tutti i componenti del dipartimento; il caso opposto è quello in cui tutti e 30 i lavori valutati "eccellenti" siano stati presentati da 10 ricercatori. In altre parole, i dati sulla distribuzione delle valutazioni non permettono alcuna analisi sulla concentrazione (o diffusione) della qualità all'interno di ciascuna struttura.

Anche se i dati delle valutazioni individuali non possono essere resi noti, i dati possono essere aggregati per soggetto valutato, in modo da calcolare un punteggio medio per ciascuno di essi. Il soggetto valutato che ha inviato 3 lavori, ciascuno dei quali valutato come "eccellente", ha una valutazione media pari a 1; se due lavori sono "eccellenti" e uno "buono", il punteggio medio è pari a 0,93; se tutti i lavori hanno ricevuto una valutazione di "limitato" il punteggio medio è uguale a zero; se non ha inviato alcun lavoro, il punteggio medio è uguale a -0,5; e così via per tutte le combinazioni di valutazioni, incluso le penalizzazioni (la soglia inferiore è -1, corrispondente alla media di tre lavori "non valutabili"). E' importante sottolineare che nel caso di lavori con più di un autore, non è sempre facile distinguere il contributo di un singolo autore, particolarmente nel caso in cui gli autori siano molti. Nel seguito si suppone che ciascuno degli autori abbia contribuito nella stessa misura alla stesura del lavoro, attribuendo quindi la "qualità media dei lavori presentati" (in genere tre) alla "qualità del soggetto valutato". Nel caso dell'Area 13 molti lavori sono a firma singola o doppia, e lavori con oltre quattro autori rappresentano casi eccezionali; il problema è quindi meno rilevante che in aree in cui sono frequenti le collaborazioni di molte decine (a volte centinaia) di autori. In ogni caso, le statistiche vanno interpretate con cautela, e servono a evidenziare le tendenze che emergono tra gruppi

consistenti di ricercatori (strutture, dipartimenti, SSD, ecc.) e non per analisi o valutazioni di casi singoli.

La Tab. 4.14 riporta, per ciascuna struttura, la distribuzione dei punteggi medi ottenuti da ciascuno dei 4.587 soggetti valutati (vengono contati due volte i soggetti con doppia affiliazione rispetto alla tabella 3.2), raggruppati in 8 classi di punteggio.[7] L'ultima riga della tabella indica che il 7,3% dei soggetti ha ricevuto una valutazione negativa ed il 31% una valutazione pari a zero. Il 10% ha ricevuto una valutazione compresa tra 0 e 0,2, e percentuali analoghe per le successive quattro classi di punteggio (da 0,2 a 0,4; da 0,4 a 0,6; da 0,6 a 0,8; da 0,8 a 0,99). Infine, il 9,6% dei soggetti valutati (440 persone) ha ottenuto il punteggio massimo di 1. Tra questi, le ultime due colonne delle tabelle distinguono tra i 144 soggetti con un numero di lavori attesi inferiore a 3 (si tratta in massima parte di giovani ricercatori, assunti dopo il 2006 o dopo il 2008) e i 296 soggetti valutati con 3 valutazioni eccellenti (il 6,4% del totale).

In sintesi, la Tab. 4.14 evidenzia che quasi il 40% dei soggetti valutati dell'Area 13 non è attivo nella ricerca o riceve valutazioni uniformemente basse; un altro 30% dei soggetti valutati riceve valutazioni medie (comprese tra 0 e 0,6); un terzo ancora è in larga misura autore dei lavori classificati "buoni" o "eccellenti", e rappresenta quindi la parte più attiva e avanzata del corpo docente delle strutture prese in considerazione (tra questi, la percentuale dei migliori ricercatori è il 9,6% del totale).

L'analisi dei punteggi medi di ciascuna struttura evidenzia una forte concentrazione della qualità.[8] La percentuale di soggetti valutati con valutazioni negative raggiunge valori particolarmente elevati (superiori al 10%) a Roma La Sapienza, Teramo, Palermo, Trieste, L'Aquila, Bari, Salento, Roma Marconi, Roma LUISS, Roma UNISU, Milano Cattolica, Torino, Calabria, Messina, CNR e Roma UNISU (queste ultime due registrano i valori più elevati, con oltre il 20%). La percentuale di soggetti valutati con un punteggio medio pari a zero (tutti i lavori hanno ricevuto un giudizio "limitato") varia in maniera estremamente significativa, da valori prossimi allo zero a valori che si avvicinano, in alcuni casi, al 50%. L'analisi dei dati mette in

[7] Nella Tab. 4.14 il numero di soggetti valutati (4587) è leggermente diverso da quello della Tab. 3.2 (4.563). La differenza è dovuta al fatto che nella Tab. 3.2 i soggetti valutati con doppia affiliazione sono conteggiati un'unica volta (nei SSD) mentre quando si considera il numero di soggetti valutati per struttura (Tab. 4.14) i soggetti vengono contati due volte.

[8] Il punteggio medio dei soggetti valutati non è identico a quello delle strutture perché il punteggio della struttura è ottenuto come media dei punteggi di tutti i lavori della struttura, mentre il punteggio medio individuale è ottenuto come media, per ciascuna struttura, dei punteggi medi individuali (indipendentemente dal fatto che il numero di lavori attesi varia tra soggetti valutati).

evidenza che anche nelle strutture che ricevono punteggi medi molto elevati una quota significativa di ricercatori ha raggiunto risultati modesti nel campo della ricerca; ad esempio, nelle Università di Padova e Bocconi, che raggiungono complessivamente risultati molto buoni, tra il 15 e il 20% dei docenti non sono attivi nella ricerca o raggiungono risultati modesti.

Infine, la distribuzione dei migliori ricercatori (ultime due colonne della Tab. 4.14) evidenzia una notevole diffusione dei ricercatori "eccellenti". Questi sono presenti in 59 diverse strutture, ma solo nove grandi atenei (Padova, Bocconi, Bologna, Roma Tor Vergata, Venezia Cà Foscari, Milano Bicocca, Roma Tre, Torino, Roma La Sapienza) registrano la presenza di una massa critica di almeno 10 ricercatori eccellenti. La concentrazione della qualità tra soggetti valutati coesiste quindi con una frammentazione della presenza degli stessi ricercatori in molte università diverse.

I dati evidenziati nella Tab. 4.14 sono confermati, anche in forma più accentuata, nella Tab. 4.15, che riporta l'analisi dei punteggi individuali per i singoli dipartimenti. Si evidenzia ancor di più la frammentazione della distribuzione dei migliori ricercatori (presenti in 93 dipartimenti distinti), ma con una presenza significativa (oltre 10) in soli 8 dipartimenti (Dipartimenti di Scienze Statistiche e Scienze Economiche e Aziendali "Marco Fanno" di Padova, Dipartimenti di Economia, Management e Tecnologia, Analisi delle Politiche e Management Pubblico (PAM) della Bocconi, Dipartimento di Scienze Economiche di Bologna, Dipartimento di Scienze Economico-Sociali e Matematico-Statistiche di Torino e Dipartimento di Economia e Finanza di Roma Tor Vergata).

La Tab. 4.16 evidenzia un aspetto della valutazione che merita attenzione. Dato il peso dei ricercatori non attivi attribuito dal bando VQR, è interessante distinguere tra soggetti valutati che hanno inviato alla VQR tutti i lavori attesi (Attivi), meno del 50% dei lavori richiesti (Parzialmente Attivi), più del 50% ma meno di quelli richiesti (Parzialmente Inattivi), nessuno di quelli richiesti (Non Attivi). La Tab. 4.16 evidenzia come, in termini assoluti, le strutture che registrano il maggior numero di non attivi, parzialmente attivi o parzialmente inattivi sono le più grandi (oltre 15 a Milano Cattolica, Torino, Bologna, Milano Bocconi, Roma La Sapienza). Tuttavia, la statistica più interessante della tabella è riportata nell'ultima colonna, che indica la percentuale di ricercatori non attivi, parzialmente attivi o parzialmente inattivi. I valori più elevati (superiori al 15%) si registrano nel CNR, a L'Aquila, Roma Marconi, Trieste e Palermo.

Con le cautele ricordate all'inizio di questa sezione, le statistiche sui soggetti valutati possono essere messe in relazione con alcune informazioni disponibili sul ruolo accademico, il settore

scientifico disciplinare, l'età e l'anno di ingresso (o presa di servizio) nel sistema universitario o della ricerca. Le ultime tabelle del rapporto (4.17, 4.18 e 4.19) forniscono alcune indicazioni in merito.

La Tab. 4.17 riporta il punteggio medio e la distribuzione del punteggio distinto per ruolo universitario (ricercatore, professore associato, professore ordinario) e sub-GEV del soggetto valutato.[9] Nei sub-GEV di Economia e di Statistica il punteggio medio dei professori ordinari supera quello degli associati e dei ricercatoti. In particolare, nel sub-GEV di Economia il punteggio medio degli ordinari è 0,47, quello degli associati 0,43 e quello dei ricercatori 0,39; nel sub-GEV di Statistica i valori sono, rispettivamente, 0,44, 0,41 e 0,35. Nel sub-GEV di Economia Aziendale invece gli ordinari tendono ad avere punteggi inferiori sia agli associati sia ai ricercatori.

La Tab. 4.18 fornisce alcuni dati sulla relazione tra qualità della ricerca e anno di nascita del soggetto valutato, distinti nuovamente per sub-GEV (anche in questo caso si escludono gli enti di ricerca). Per le generazioni più anziane (nati prima del 1945) il punteggio medio è inferiore a 0,15 in ciascuno dei tre sub-GEV. Si registra un forte aumento della qualità della ricerca per le generazioni successive; per quanto riguarda il sub-GEV di Economia e di Statistica, l'aumento si verifica soprattutto a partire dalle generazioni di nati negli anni cinquanta; per il sub-GEV di Aziendale il miglioramento della qualità si manifesta circa dieci anni più tardi, a partire dai nati negli anni sessanta. Con l'eccezione del sub-GEV di statistica, i miglioramenti della qualità continuano anche per le generazioni più giovani (nati dopo il 1975), anche se con una dinamica più lenta che nei due decenni precedenti.

Un'altra informazione disponibile nei dati analizzati dal GEV è l'anno di ingresso (o presa di servizio) nel sistema universitario o della ricerca. Questo può essere misurato (con un certo grado di approssimazione perché non si tiene conto, ad esempio, dei periodi di congedo per maternità) dal numero di lavori attesi: tre per coloro che erano già in servizio prima del 2006, due per coloro che hanno preso servizio nel 2006 o nel 2007, uno solo per coloro che hanno preso servizio nel 2008 o nel 2009 (per gli enti di ricerca i valori sono 6, 4 e 2). Sulla base di questa informazione, la Tab. 4.19 riporta il punteggio medio ottenuto per anno di ingresso nelle

[9] Poiché la progressione di carriera negli enti di ricerca segue modalità diverse che negli atenei e poiché il numero di osservazioni in alcune celle è molto ridotto, la tabella esclude i soggetti valutati per cui non è disponibile il SSD (gli enti di ricerca).

strutture, separatamente per ciascun sub-GEV. La tabella evidenzia che, per ciascun sub-GEV, il punteggio dei ricercatori che hanno preso servizio dopo il 2006 è più elevato di coloro che erano già in servizio prima di quella data. Il dato può essere interpretato come un segnale che le assunzioni più recenti hanno riguardato personale di qualità media superiore a quella del personale in servizio, e quindi nel senso di un cambiamento strutturale dei risultati della ricerca. Tuttavia, così come la distribuzione per anno di nascita in Tab. 4.18, il dato potrebbe anche riflettere un effetto esclusivo dell'età (indipendentemente dall'anno di nascita, personale più giovane tende ad essere più attivo nella ricerca), effetto che nel corso del tempo sarà riassorbito.

Le dinamiche per anno di nascita e anno di presa di servizio possono dunque essere guardate con un certo ottimismo, anche se si devono attendere nuovi dati che, nel corso del tempo, confermino i risultati di coloro che hanno preso servizio negli anni più recenti. Il miglioramento rispetto a chi era in servizio prima del 2006 è evidente anche per gli ingressi più recenti, ovvero successivi al 2008, che però in certi casi arretrano rispetto agli ingressi del 2006-07, presumibilmente a causa di un limite al numero di lavori pubblicati nei primissimi anni della carriera.

5 Analisi dei risultati

Il GEV ha utilizzato le stesse modalità di valutazione per tutte le sotto-aree di ricerca. Il confronto statistico tra valutazione peer e valutazione bibliometrica rappresenta quindi un contributo metodologico del gruppo di lavoro dell'Area 13, e cioè uno strumento di controllo reciproco tra due modalità di valutazione che sono spesso viste come alternative. Anche se il mix tra analisi bibliometrica e valutazione peer è molto diverso tra settori e aree di ricerca, nel complesso dell'Area non si sono evidenziati scostamenti significativi tra le due modalità di valutazione per i lavori che è stato possibile sottoporre a entrambe le valutazioni (il campione casuale estratto degli articoli su rivista). Ne segue che nell'Area 13 le differenze registrate nei punteggi medi ottenuti da strutture, dipartimenti, SSD e singoli soggetti valutati riflettono in gran parte differenze nella qualità della ricerca, piuttosto che metodi di valutazione differenti.

L'analisi dei risultati evidenzia due elementi di riflessione. In primo luogo, si riscontra un notevole divario dell'Area nel confronto internazionale; in secondo luogo, una forte dispersione dei punteggi medi, e quindi della qualità della ricerca, tra strutture, dipartimenti, SSD e singoli soggetti valutati. Per quanto riguarda il confronto internazionale, si deve rilevare che il punteggio medio dell'Area è 0,32, a fronte di un valore di confronto internazionale (peraltro assolutamente

indicativo) implicito nel bando VQR di circa 0,4 (cfr. sezione 4.1). A questo riguardo, si deve anche ricordare che i lavori valutati dal GEV non sono un campione casuale dei lavori dei ricercatori del settore nel 2004-2007, ma un gruppo selezionato dei migliori lavori di ciascun soggetto valutato (generalmente 3).

La percentuale di lavori classificati come "eccellenti", sugli oltre 12.000 conferiti alla VQR nell'Area 13, è solo del 18%; la definizione stessa di "eccellente", secondo i criteri adottati nell'analisi bibliometrica, è alquanto generosa, comprendendo molte centinaia di riviste. All'estremo opposto, quasi metà dei lavori ha ricevuto una valutazione "limitata" (al di sotto cioè della media della distribuzione internazionale della ricerca, così come misurata dalla bibliometria e percepita dai revisori esterni); circa il 6% dei lavori attesi non è stato inviato alla VQR da parte di ricercatori non attivi o solo parzialmente attivi sul fronte della ricerca. Molto bassa è anche la percentuale di articoli su rivista con un numero di citazioni significativo: l'Appendice D indica che il 46% dei 3.475 articoli pubblicati su riviste censite da WoS e inviati al GEV non ha alcuna citazione nel 2004-2010; il 15% ha una sola citazione annua; il 22% ha un numero di citazioni annue compreso tra 2 e 4; solo 614 articoli (17,7% degli articoli pubblicati su riviste W0S) hanno un numero di citazioni annue superiore a quattro.

Il secondo dato che risulta evidente dall'analisi dei dati è la forte dispersione della qualità della ricerca tra strutture, dipartimenti, SSD e soggetti valutati. La distribuzione dei punteggi per struttura, ad esempio, evidenzia una differenza di 0,4-0,5 tra i punteggi più elevati e quelli più bassi. Differenze analoghe, e a volte più ampie, si registrano nelle differenze di punteggio tra dipartimenti o tra SSD. Vi sono notevoli differenze di qualità della ricerca anche all'interno degli atenei, con una forte variabilità di risultati anche all'interno di atenei che hanno complessivamente ricevuto valutazioni elevate. Ed infine, all'interno di ciascun dipartimento, i singoli ricercatori hanno contribuito in modo assai differenziato al punteggio medio, con valori di "eccellenza" anche nei dipartimenti che complessivamente hanno ricevuto valutazioni poco incoraggianti, e molti valori "limitati" (o casi di ricercatori non attivi) anche nei migliori dipartimenti.

Su quest'ultimo punto, l'aspetto da notare non è tanto il fatto che anche nei migliori dipartimenti vi siano ricercatori poco attivi sul fronte della ricerca. Le esigenze di insegnamento e organizzative richiedono docenti che possono voler dedicare la maggior parte del proprio tempo a queste attività. E' importante piuttosto domandarsi come motivare e valorizzare i ricercatori più attivi e capaci. Da questo punto di vista, la relazione ha messo in evidenza una forte dispersione dei migliori ricercatori (coloro che hanno ricevuto il punteggio massimo in

ciascuno dei lavori presentati), presenti in 93 dipartimenti diversi, e con una presenza significativa (oltre 10) in soli 8 dipartimenti. Nell'utilizzo dei risultati della valutazione, è importante quindi prestare attenzione a non scoraggiare i gruppi di ricercatori più attivi, presenti in modo diffuso nella gran parte degli atenei.

Dal punto di vista territoriale, pur con eccezioni di rilievo, particolarmente se l'analisi riguarda i dipartimenti e le differenze interne tra sub-GEV e SSD, il GEV osserva anche con preoccupazione il divario tra gli atenei del Nord e quelli del Mezzogiorno.

Il risultato dell'unico ente di ricerca con una presenza consistente di ricercatori (il CNR) è complessivamente deludente, sia nel punteggio dei lavori valutati, sia nella percentuale più elevata di lavori non inviati alla VQR (pesa in questo caso anche il numero più elevato di lavori richiesto); fanno eccezione, in senso positivo e con punteggi molto più elevati, i docenti universitari associati al CNR.

Mentre la dimensione non sembra spiegare, di per sé, la valutazione ottenuta dalle strutture, i ranking di alcune strutture potrebbero essere influenzati dalla diversa composizione del personale di ruolo (ordinari, associati, ricercatori) nei vari dipartimenti e atenei e, in alcune sedi, dalla concentrazione di personale anziano poco attivo. Anche l'età media del personale potrebbe essere uno dei fattori che spiega le differenze tra strutture e atenei, come suggerito dall'analisi dei punteggi medi per anno di nascita e anno di ingresso nelle strutture di ricerca. L'analisi dei punteggi per anno di nascita e di ingresso nelle strutture denota infatti un incoraggiante e progressivo miglioramento della qualità della ricerca tra le giovani generazioni. Si auspica che tale tendenza rappresenti un cambiamento strutturale e che in futuro sia possibile procedere ad un'analisi statistica approfondita per descrivere e interpretare meglio le tendenze in atto.

Per quanto riguarda le sotto-aree di ricerca, nel sub-GEV di Economia si notano forti eterogeneità tra SSD. In particolare, si notano i risultati dei settori SECS-P/02 (Politica Economica), SECS-P/03 (Scienza delle Finanze) e SECS-P/06 (Economia Applicata), settori in cui il risultato punteggio medio è inferiore rispetto al settore SECS-P/01 (Economia Politica), mentre risulta maggiore la percentuale di lavori conferiti, il che suggerisce un tasso di attività leggermente più elevato, ma di minore qualità. Per quanto riguarda i settori SECS-P/04 e SECS-P/12, si evidenzia la specificità della storia economica e del pensiero economico, che si discostano in parte dagli altri SSD dell'area economica per quanto riguarda le tradizioni e i metodi della ricerca, la tipologia e la collocazione editoriale delle pubblicazioni. In particolare, la percentuale di lavori in lingua inglese è limitata soprattutto per la storia economica, mentre in

ambedue i casi le valutazioni medie e la percentuale di articoli su rivista sono inferiori a quelle degli altri SSD del sub-GEV.

Anche all'interno del sub-GEV di Statistica e Matematica Applicata sono presenti forti eterogeneità sia tra SSD, sia all'interno di ciascun SSD. Si evidenziano, in particolare i buoni risultati dei settori di Statistica per la Ricerca Sperimentale e Tecnologica (SECS-S/02) e soprattutto di Econometria (SECS-P/05), e i punteggi inferiori alla media dei SSD compresi nel perimetro della statistica applicata, cioè Statistica Economica (SECS-S/03), Demografia (SECS-S/04) e Statistica Sociale (SECS-S/05). Il settore di Metodi Matematici dell'Economia e delle Scienze Attuariali e Finanziarie (SECS-S/06) raggiunge invece risultati complessivamente in linea con le medie di area; pesa negativamente in questo settore una percentuale relativamente elevata di ricercatori non attivi o parzialmente attivi.

Per quanto riguarda il sub-GEV di area aziendale, è molto evidente un risultato mediamente inferiore a quello generale di Area 13. Pur non essendo inatteso, il dato conferma che gli SSD dell'area aziendale non hanno raggiunto il livello di qualità e internazionalizzazione delle pubblicazioni dell'area economica e statistica; si evidenzia inoltre una quota molto elevata di lavori con valutazione "limitata". Il dato potrebbe riflettere il fatto che i ricercatori di area aziendale sono più orientati, per tradizione, a scrivere monografie in lingua italiana che, al contrario delle riviste internazionali, non sono classificate automaticamente nelle fasce di merito più elevate. Tuttavia, la spiegazione non è convincente. Infatti, tutti i libri e le monografie in lingua italiana sono stati valutati in peer review, ed in grandissima parte da revisori italiani, che li hanno classificati prevalentemente nella fascia di giudizio "limitato". Dunque, l'alta percentuale di lavori con valutazione "limitata" è il risultato di una valutazione della comunità scientifica italiana sulla comunità scientifica italiana, e non può essere ascritta alla tipologia di pubblicazione. Pur con differenze tra SSD, in particolare per le percentuali di lavori classificati "eccellenti" o "limitati", la VQR ha messo in evidenza che vi sono margini di miglioramento notevoli e diffusi in tutti i SSD del sub-GEV di Economia Aziendale.

Va infine sottolineato l'approccio multidisciplinare del SECS/P-13 (Scienze Merceologiche), peraltro comune ad altri SSD di area statistica, che trova conferma nelle diverse tipologie di lavori e quindi anche nelle modalità di valutazione. Insieme ad una percentuale abbastanza contenuta di valutazioni bibliometriche (per lavori di area chimico-merceologica e agrario-ambientale), si osserva anche una quota consistente di valutazioni in peer review. Ciò dipende anche dalla presenza relativamente limitata nel SSD di riviste di riferimento internazionali e

nazionali con indicatori bibliometrici significativi per tematiche multidisciplinari e spesso con competenze trasversali a più settori e aree di ricerca.

6 Considerazioni finali

Il dato più rilevante della valutazione in Area 13 è la forte dispersione della qualità della ricerca tra strutture, dipartimenti, SSD e soggetti valutati. Differenze ampie si registrano anche tra dipartimenti degli stessi atenei, particolarmente negli atenei di grande dimensione. Inoltre, all'interno stesso dei dipartimenti, i singoli ricercatori hanno contribuito in modo assai differenziato al punteggio medio, con valori di "eccellenza" anche in alcuni dipartimenti che complessivamente hanno ricevuto valutazioni poco incoraggianti, e molti valori "limitati" (o casi di ricercatori non attivi o solo parzialmente attivi) anche nei migliori dipartimenti.

L'analisi delle valutazioni per età e anno di ingresso nel mondo della ricerca denota un progressivo miglioramento della qualità della ricerca tra le giovani generazioni. Si auspica che questa tendenza rappresenti un cambiamento strutturale e che in futuro sia possibile proseguire l'analisi statistica per descrivere e interpretare le differenze registrate tra strutture, dipartimenti e settori disciplinari.

In tutto il corso della valutazione il GEV si è strettamente attenuto al mandato ricevuto. Nonostante i problemi emersi nel corso della valutazione ed evidenziati in questo rapporto finale, il GEV conferma la fiducia nei metodi adottati per il rilevamento, l'elaborazione e la valutazione della ricerca nell'Area 13 e confida che i risultati trovino applicazione nel processo di miglioramento della qualità dell'università e della ricerca.

Glossario

Article Influence Score (AIS): misura l'influenza degli articoli di una rivista nel corso dei primi cinque anni dopo la pubblicazione. È calcolato dividendo l'Eigenfactor Score per il numero di articoli della rivista, normalizzato come una frazione di tutti gli articoli in tutte le riviste. Il calcolo dell'Eigenfactor Score si basa sul numero di volte che gli articoli di una rivista pubblicati negli ultimi cinque anni sono stati citati, ma considera anche quali riviste hanno contribuito a queste citazioni nel senso che le riviste più citate influenzano l'indicatore più di riviste meno citate. Le citazioni di un articolo dalla rivista in cui è pubblicato vengono rimosse, in modo tale che l'Eigenfactor Score non sia influenzato dalle auto-citazioni di una rivista. AIS è fortemente correlato a IF5 perché si basa sul rapporto tra il numero di citazioni della rivista e il numero di articoli pubblicati negli ultimi cinque anni. Il valore medio di AIS è 1. Un punteggio superiore a 1 indica che la rivista ha un impatto superiore alla media. Un punteggio inferiore a 1 indica che la rivista ha un impatto inferiore alla media. I dati utilizzati dal GEV si riferiscono a dicembre 2010.

Gruppo di Consenso: gruppo di Esperti della Valutazione costituito da almeno due membri GEV con il compito di definire la valutazione di sintesi dei lavori di ricerca a loro assegnati ed individuare i lavori "non valutabili". I Gruppi di Consenso sono formati dai due membri GEV a cui è stato assegnato il lavoro dai Coordinatori dei sub-GEV. In caso di conflitto di valutazioni tra i componenti del Gruppo di Consenso, il Gruppo è integrato con il Coordinatore del sub-GEV o con il Presidente del GEV. I Gruppi di Consenso costituiti per verificare l'insieme dei "lavori non valutabili" sono stati sempre integrati dai Coordinatori sub-GEV o dal Presidente del GEV.

h-index: Una rivista ha un indice h se h degli N articoli pubblicati dalla rivista hanno almeno h citazioni ciascuno, e gli altri N-h articoli non hanno più di h citazioni ciascuno. L'indice h è stato calcolato in Google Scholar per il periodo 2004-2010 nel mese di aprile 2012, e controllato durante il mese di maggio 2012.

Impact Factor (IF): è una misura della frequenza con la quale un articolo di una rivista è stato citato in un particolare anno o periodo. L'Impact Factor annuale di una rivista è calcolato dividendo il numero di citazioni dell'anno corrente per il numero di articoli pubblicati nella rivista durante i due anni precedenti. I dati utilizzati dal GEV si riferiscono a dicembre 2010.

Impact Factor 5 anni (IF5): è il rapporto tra citazioni e articoli pubblicati in una rivista in un periodo di cinque anni. I dati utilizzati dal GEV si riferiscono a dicembre 2010.

Lavori attesi: numero di lavori di ricerca attesi calcolati sulla base della numerosità dei soggetti valutati accreditati dalle strutture tenendo conto delle esenzioni (ad esempio, per prese di servizio successive al 2006).

Lavori conferiti: numero di lavori sottoposti alla valutazione dalle strutture in base all' Area di appartenenza del soggetto valutato.

Sotto-aree di ricerca: Le riviste del GEV di Area 13 sono state divise sulla base dell'Area di ricerca prevalente e in quattro sotto-aree: Economia, Economia Aziendale, Storia Economica e Storia del Pensiero Economico, Statistica e Matematica per l'Economia e la Finanza.

Strutture: L'esercizio di Valutazione della Qualità della Ricerca 2004-2010, come previsto dal Bando del 7 novembre 2011, è rivolto alla valutazione dei risultati della ricerca scientifica delle seguenti strutture di ricerca:

a) Università statali;
b) Università non statali legalmente riconosciute;
c) Enti di ricerca pubblici vigilati dal MIUR;
d) Altri soggetti pubblici e privati che svolgono attività di ricerca, su esplicita richiesta e previa intesa che preveda la copertura delle spese relative.

SSD: Settore scientifico-disciplinare. I settori scientifico-disciplinari sono stabiliti dal Ministero dell'Istruzione, dell'Università e della Ricerca.

Sub-GEV: sottogruppo di esperti della valutazione con competenze disciplinari omogenee.

Subject category: categorie di classificazione dei lavori di ricerca sulla base dell'argomento principale di ricerca. Il GEV ha utilizzato le 33 subject categories definite in Appendice 5 per classificare i lavori da inviare in peer review e le competenze dei revisori esterni.

Elenco delle tabelle

Tabella 1.1: Settori scientifico-disciplinari (SSD) dell'Area 13 (Scienze Economiche e Statistiche).

Tabella 1.2: Composizione del Gruppo di Esperti della Valutazione Area 13.

Tabella 1.3: Organizzazione degli esperti in sub-GEV, corrispondenti SSD e distribuzione dei prodotti della ricerca. Per sub-GEV si intende un sottogruppo di SSD interni all'area con affinità disciplinari costituiti secondo i criteri del GEV.

Tabella 1.4: Dimensioni delle strutture (università, enti di di ricerca, consorzi) per numero di lavori presentati alla VQR nell'area 13.

Tabella 1.5: Distribuzione dei lavori di ricerca conferiti per tipologia e SSD. Per SSD si intende il settore scientifico disciplinare di afferenza del soggetto valutato. La tipologia "Altro" comprende tutte le tipologie di lavori non menzionate precedentemente. Tali tipologie sono state accorpate a causa del numero ridotto di lavori che includono.

Tabella 1.6: Distribuzione dei lavori di ricerca per lingua di pubblicazione e SSD. Per SSD si intende il settore scientifico disciplinare di afferenza del soggetto valutato. La categoria "Altra lingua" contiene le percentuali di lavori di ricerca pubblicati in lingue diverse da italiano e inglese. La categoria "Lingua non disponibile" contiene le percentuali dei lavori di ricerca per i quali la lingua di pubblicazione non è stata specificata.

Tabella 1.7: Distribuzione dei lavori conferiti per tipologia, lingua di pubblicazione e sub-GEV. Per sub-GEV si intende un sottogruppo di SSD interni all'area con affinità disciplinari costituiti secondo i criteri del GEV. La tipologia "Altro" comprende tutte le tipologie di lavori non menzionate precedentemente. Tali tipologie sono state accorpate a causa del numero ridotto di lavori che includono. La categoria "Altra lingua" contiene le percentuali di lavori di ricerca pubblicati in lingue diverse da italiano e inglese. La categoria "Lingua non disponibile" contiene le percentuali dei lavori di ricerca per i quali la lingua di pubblicazione non è stata specificata.

Tabella 1.8: Distribuzione dei lavori conferiti per anno di pubblicazione, tipologia e SSD. Per SSD si intende il settore scientifico disciplinare di afferenza del soggetto valutato. La tipologia "Altro" comprende tutte le tipologie di lavori non menzionate precedentemente. Tali tipologie sono state accorpate a causa del numero ridotto di lavori che includono.

Tabella 3.1: Numero di lavori conferiti e attesi, per SSD. Per SSD si intende il settore scientifico disciplinare di afferenza del soggetto valutato. Il numero di lavori attesi è stato calcolato sulla base del

SSD di afferenza dei soggetti valutati e del numero di lavori che da bando questi erano tenuti a inviare alla VQR. Il numero di lavori conferiti contiene il numero di lavori effettivamente sottomessi.

Tabella 3.2: Numero di soggetti valutati e relativo numero di lavori attesi. SV è l'acronimo di Soggetti Valutati. Il numero di lavori attesi è stato calcolato sulla base del SSD di afferenza dei soggetti valutati e del numero di lavori che da bando questi erano tenuti a inviare alla VQR.

Tabella 3.3: Numero e percentuale di lavori di ricerca conferiti da soggetti afferenti all'area, per area di valutazione. Per GEV si intende il GEV che ha valutato i lavori di ricerca.

Tabella 3.4: Percentuale di lavori di ricerca valutati attraverso la *peer review* , la bibliometria e con entrambi i metodi di valutazione. Per lavori valutati in bibliometria si intendono lavori che hanno una valutazione bibliometrica secondo i criteri del GEV valutante (anche non conclusiva, cioè IR). Per lavori con doppia valutazione si intendono i lavori che sono stati sottoposti ad entrambi i metodi di valutazione (*peer review* e bibliometria). Le percentuali della tabella non tengono conto dei lavori penalizzati. Per lavori penalizzati si intendono i lavori non valutabili e casi accertati di plagio o frode così come previsto dal bando VQR del 7 Novembre 2011, i prodotti mancanti (cioè attesi e non sottomessi), i prodotti identici presentati più volte dalla stessa struttura, i prodotti identici presentati più volte dallo stesso soggetto valutato per due strutture di tipologia differente (es. Università ed Ente di Ricerca).

Tabella 4.1: Punteggi ottenuti e distribuzione dei lavori nelle classi di merito (Eccellente -E-, Buono -B-, Accettabile –A-, Limitato -L-) nell'area e per sub-GEV. Per sub-GEV si intende un sottogruppo di SSD interni all'area con affinità disciplinari costituiti secondo i criteri del GEV. Per "somma punteggi (v)" si intende la valutazione complessiva del sub-GEV ottenuta sommando i punteggi dei lavori associati a soggetti valutati dell'area afferenti al sub-GEV. La categoria "lavori penalizzati" contiene i lavori ritenuti non valutabili e i casi accertati di plagio o frode così come previsto dal bando VQR del 7 novembre 2011, i lavori mancanti (cioè attesi e non sottomessi), i lavori identici presentati più volte dalla stessa struttura , i lavori identici presentati più volte dallo stesso soggetto valutato per due strutture di tipologia differente (es. Università ed Ente di Ricerca). Per # lavori attesi si intende il numero di lavori attesi calcolato sulla base del SSD di afferenza dei soggetti valutati e del numero di lavori che da bando questi erano tenuti a inviare alla VQR.

Tabella 4.2: Punteggi ottenuti e distribuzione dei lavori nelle classi di merito (Eccellente -E-, Buono -B-, Accettabile –A-, Limitato -L-) per SSD. Per "somma punteggi (v)" si intende la valutazione complessiva del SSD ottenuta sommando i punteggi dei lavori afferenti al SSD. La categoria "lavori penalizzati" contiene i lavori non valutabili e casi accertati di plagio o frode così come previsto dal bando VQR del 7 novembre 2011, i lavori mancanti (cioè attesi e non sottomessi), i lavori identici presentati più volte dalla stessa struttura, i lavori identici presentati più volte dallo stesso soggetto valutato per due strutture di tipologia differente (es. Università ed Ente di Ricerca). Per # lavori attesi si intende il numero di lavori

attesi calcolato sulla base del SSD di afferenza dei soggetti valutati e del numero di lavori che da bando questi erano tenuti a inviare alla VQR.

Tabella 4.3: Punteggi ottenuti e distribuzione dei lavori conferiti non penalizzati nelle classi di merito (Eccellente -E-, Buono -B-, Accettabile –A-, Limitato -L-) per sub-GEV e anno di pubblicazione. Per sub-GEV si intende un sottogruppo di SSD interni all'area con affinità disciplinari costituiti secondo i criteri del GEV. Per "somma punteggi (v)" si intende la valutazione complessiva del sub-GEV per l'anno considerato, ottenuta sommando i punteggi dei lavori afferenti al sub-GEV.

Tabella 4.4: Punteggi ottenuti e distribuzione dei lavori conferiti non penalizzati nelle classi di merito (Eccellente -E-, Buono -B-, Accettabile –A-, Limitato -L-)per modalità di valutazione (*peer review* o bibliometria). La tabella non considera i lavori penalizzati. Per "somma punteggi (v)" si intende la valutazione complessiva dei lavori valutati secondo la modalità indicata ottenuta sommando i punteggi dei lavori valutati.

Tabella 4.5: Punteggi ottenuti e distribuzione dei lavori conferiti non penalizzati nelle classi di merito (Eccellente -E-, Buono -B-, Accettabile –A-, Limitato -L-) per tipologia di pubblicazione. Per "somma punteggi (v)" si intende la valutazione complessiva dei lavori appartenenti alla tipologia indicata, ottenuta sommando i punteggi dei singoli lavori. La tipologia "Altro" comprende tutte le tipologie di lavori non menzionate precedentemente. Tali tipologie sono state accorpate a causa del numero ridotto di lavori che includono.

Tabella 4.6: Punteggi ottenuti e distribuzione dei lavori conferiti non penalizzati nelle classi di merito (Eccellente -E-, Buono -B-, Accettabile -A-, Limitato -L-) per tipologia di pubblicazione e SSD. Per "somma punteggi (v)" si intende la valutazione complessiva del SSD per l'anno considerato, ottenuta sommando i punteggi dei lavori afferenti al SSD. La tipologia "Altro" comprende tutte le tipologie di lavori non menzionate precedentemente. Tali tipologie sono state accorpate a causa del numero ridotto di lavori che includono. Alcune celle sono state omesse al fine di proteggere la riservatezza dei dati personali, per questomotivo i valori di sintesi riportati nell'ultima riga non corrispondono ai totali di colonna.

Tabella 4.7a. Graduatorie di strutture ordinate per voto medio (I=v/n) e distribuzione dei lavori nelle classi di merito (Eccellente -E-, Buono -B-, Accettabile -A-, Limitato -L-) per dimensione delle strutture (strutture medio-piccole -P-, con 10-150 lavori attesi, e strutture grandi -G- , con oltre 150 lavori attesi). Per "somma punteggi (v)" si intende la valutazione complessiva dei lavori appartenenti alla struttura ottenuta sommando i punteggi dei singoli lavori. Per # lavori attesi si intende il numero di lavori attesi calcolato sulla base del SSD di afferenza dei soggetti valutati e del numero di lavori che da bando questi erano tenuti a inviare alla VQR. La categoria "lavori penalizzati" contiene i lavori non valutabili e casi accertati di plagio o frode così come previsto dal bando VQR del 7 novembre 2011, i lavori mancanti (cioè attesi e non inviati), i lavori identici presentati più volte dalla stessa struttura e i lavori identici

presentati più volte dallo stesso soggetto valutato per due strutture di tipologia differente (es. Università ed Ente di Ricerca). R indica il voto medio normalizzato. (n/N) x 100 rappresenta la percentuale dei lavori attesi della struttura rispetto ai lavori attesi dell'area (N: numero lavori attesi dell'Area; n: numero lavori attesi della struttura). IRAS1 è il primo indicatore di qualità della ricerca definito nel Bando VQR come rapporto tra il punteggio complessivo raggiunto da una struttura in una data area e la valutazione complessiva dell'area stessa. Gli indicatori R ed IRAS1 riportati per gli Enti di Ricerca e Consorzi sono stati calcolati rispettivamente sull'insieme dei soli Enti di Ricerca o dei soli Consorzi. La tabella non include le strutture con meno di 10 lavori attesi.

Tabella 4.7b. Graduatorie di strutture ordinate per voto medio (I=v/n) e distribuzione dei lavori nelle classi di merito (Eccellente -E-, Buono -B-, Accettabile –A-, Limitato -L-). Per "somma punteggi (v)" si intende la valutazione complessiva dei lavori appartenenti alla struttura ottenuta sommando i punteggi dei singoli lavori. Per # lavori attesi si intende il numero di lavori attesi calcolato sulla base del SSD di afferenza dei soggetti valutati e del numero di lavori che da bando questi erano tenuti a inviare alla VQR. La categoria "lavori penalizzati" contiene i lavori non valutabili e casi accertati di plagio o frode così come previsto dal bando VQR del 7 novembre 2011, i lavori mancanti (cioè attesi e non inviati), i lavori identici presentati più volte dalla stessa struttura e i lavori identici presentati più volte dallo stesso soggetto valutato per due strutture di tipologia differente (es. Università ed Ente di Ricerca). R indica il voto medio normalizzato. (n/N) x 100 rappresenta la percentuale dei lavori attesi della struttura rispetto ai lavori attesi dell'area (N: numero lavori attesi dell'Area; n: numero lavori attesi della struttura). IRAS1 è il primo indicatore di qualità della ricerca definito nel Bando VQR come rapporto tra il punteggio complessivo raggiunto da una struttura in una data area e la valutazione complessiva dell'area stessa. Gli indicatori R ed IRAS1 riportati per gli Enti di Ricerca e Consorzi sono stati calcolati rispettivamente sull'insieme dei soli Enti di Ricerca o dei soli Consorzi. La tabella non include le strutture con meno di 10 lavori attesi.

Tabella 4.8. Graduatorie di strutture per sub-GEV ordinate per voto medio (I=v/n) e distribuzione dei lavori nelle classi di merito (Eccellente -E-, Buono -B-, Accettabile –A-, Limitato -L-). Per sub-GEV si intende un sottogruppo di SSD interni all'area con affinità disciplinari costituiti secondo i criteri del GEV. Per "somma punteggi (v)" si intende la valutazione complessiva dei lavori appartenenti alla struttura ottenuta sommando i punteggi dei singoli lavori. Per # lavori attesi si intende il numero di lavori attesi calcolato sulla base del SSD di afferenza dei soggetti valutati e del numero di lavori che da bando questi erano tenuti a inviare alla VQR. La categoria "lavori penalizzati" contiene i lavori non valutabili e casi accertati di plagio o frode così come previsto dal bando VQR del 7 novembre 2011, i lavori mancanti (cioè attesi e non sottomessi), i lavori identici presentati più volte dalla stessa struttura e i lavori identici presentati più volte dallo stesso soggetto valutato per due strutture di tipologia differente (es. Università ed Ente di Ricerca). Alcune celle sono state omesse al fine di proteggere la riservatezza dei dati personali, per questo motivo i valori di sintesi riportati nell'ultima riga non corrispondono ai totali di colonna.

Tabella 4.9: Graduatorie di strutture per SSD ordinate per voto medio (I=v/n) e distruzione dei lavori nelle classi di merito (Eccellente -E-, Buono -B-, Accettabile –A-, Limitato -L-). Per "somma punteggi (v)" si intende la valutazione complessiva dei lavori appartenenti alla struttura ottenuta sommando i punteggi dei singoli lavori. Per # lavori attesi si intende il numero di lavori attesi calcolato sulla base del SSD di afferenza dei soggetti valutati e del numero di lavori che da bando questi erano tenuti a inviare alla VQR. La categoria "lavori penalizzati" contiene i lavori non valutabili e casi accertati di plagio o frode così come previsto dal bando VQR del 7 novembre 2011, i lavori mancanti (cioè attesi e non sottomessi), i lavori identici presentati più volte dalla stessa struttura e i lavori identici presentati più volte dallo stesso soggetto valutato per due strutture di tipologia differente (es. Università ed Ente di Ricerca). Alcune celle sono state omesse al fine di proteggere la riservatezza dei dati personali, per questo motivo i valori di sintesi riportati nell'ultima riga non corrispondono ai totali di colonna.

Tabella 4.10a: Graduatorie dei dipartimenti post L. 240 ordinati per voto medio (I=v/n) e distribuzione dei lavori nelle classi di merito (Eccellente -E-, Buono -B-, Accettabile –A-, Limitato -L-). Per "somma punteggi (v)" si intende la valutazione complessiva dei lavori appartenenti alla struttura ottenuta sommando i punteggi dei singoli lavori. Per # lavori attesi si intende il numero di lavori attesi calcolato sulla base del SSD di afferenza dei soggetti valutati e del numero di lavori che da bando questi erano tenuti a inviare alla VQR. La categoria "lavori penalizzati" contiene i lavori non valutabili e casi accertati di plagio o frode così come previsto dal bando VQR del 7 novembre 2011, i lavori mancanti (cioè attesi e non sottomessi), i lavori identici presentati più volte dalla stessa struttura e i lavori identici presentati più volte dallo stesso soggetto valutato per due strutture di tipologia differente (es. Università ed Ente di Ricerca). R rappresenta il voto medio normalizzato. (n/N) x 100 rappresenta la percentuale dei lavori attesi dal dipartimento rispetto ai lavori attesi dell' area (N:numero lavori attesi dell'Area; n: numero lavori attesi del dipartimento). IRD1 è il primo indicatore di qualità della ricerca applicato ai dipartimenti definito nel Bando VQR come rapporto tra il punteggio complessivo raggiunto da un dipartimento in una data area e la valutazione complessiva dell'area stessa. Gli indicatori R ed IRAS1 riportati per le sottostrutture degli Enti di Ricerca sono stati calcolati rispettivamente sull'insieme dei soli Enti di Ricerca o dei soli Consorzi. La tabella non include i dipartimenti con meno di 10 lavori attesi.

Tabella 4.10b: Graduatorie dei dipartimenti post L.240 ordinati per voto medio (I=v/n) all'interno delle strutture (elencate in ordine alfabetico) e distribuzione dei lavori nelle classi di merito (Eccellente -E-, Buono -B-, Accettabile –A-, Limitato -L-). Per "somma punteggi (v)" si intende la valutazione complessiva dei lavori appartenenti alla struttura ottenuta sommando i punteggi dei singoli lavori. Per # lavori attesi si intende il numero di lavori attesi calcolato sulla base del SSD di afferenza dei soggetti valutati e del numero di lavori che da bando questi erano tenuti a sottomettere. La categoria "lavori penalizzati" contiene i lavori non valutabili e casi accertati di plagio o frode così come previsto dal bando VQR del 7 novembre 2011, i lavori mancanti (cioè attesi e non sottomessi), i lavori identici presentati più

volte dalla stessa struttura e i lavori identici presentati più volte dallo stesso soggetto valutato per due strutture di tipologia differente (es. Università ed Ente di Ricerca). R rappresenta il voto medio normalizzato. (n/N) x 100 rappresenta la percentuale dei lavori attesi dal dipartimento rispetto ai lavori attesi dell' area (N: numero lavori attesi dell'Area; n: numero lavori attesi del dipartimento). IRD1 è il primo indicatore di qualità della ricerca applicato ai dipartimenti definito nel Bando VQR come rapporto tra il punteggio complessivo raggiunto da un dipartimento in una data area e la valutazione complessiva dell'area stessa. Gli indicatori R ed IRAS1 riportati per le sottostrutture degli Enti di Ricerca sono stati calcolati rispettivamente sull'insieme dei soli Enti di Ricerca o dei soli Consorzi. La tabella non include i dipartimenti con meno di 10 lavori attesi.

Tabella 4.11a: Graduatorie dei dipartimenti ante L.240 ordinati per voto medio (I=v/n) e distribuzione dei lavori nelle classi di merito (Eccellente -E-, Buono -B-, Accettabile –A-, Limitato -L-). Per "somma punteggi (v)" si intende la valutazione complessiva dei lavori appartenenti alla struttura ottenuta sommando i punteggi dei singoli lavori. Per # lavori attesi si intende il numero di lavori attesi calcolato sulla base del SSD di afferenza dei soggetti valutati e del numero di lavori che da bando questi erano tenuti a inviare alla VQR. La categoria "lavori penalizzati" contiene i lavori non valutabili e casi accertati di plagio o frode così come previsto dal bando VQR del 7 Novembre 2011, i lavori mancanti (cioè attesi e non sottomessi), i lavori identici presentati più volte dalla stessa struttura e i lavori identici presentati più volte dallo stesso soggetto valutato per due strutture di tipologia differente (es. Università ed Ente di Ricerca). R rappresenta il voto medio normalizzato. (n/N) x 100 rappresenta la percentuale dei lavori attesi dal dipartimento rispetto ai lavori attesi dell' area (N: numero lavori attesi dell'Area; n: numero lavori attesi del dipartimento). IRD1 è il primo indicatore di qualità della ricerca applicato ai dipartimenti definito nel Bando VQR come rapporto tra il punteggio complessivo raggiunto da un dipartimento in una data area e la valutazione complessiva dell'area stessa. Gli indicatori R ed IRAS1 riportati per le sottostrutture degli Enti di Ricerca sono stati calcolati rispettivamente sull'insieme dei soli Enti di Ricerca o dei soli Consorzi. La tabella non include i dipartimenti con meno di 10 lavori attesi.

Tabella 4.11b: Graduatorie dei dipartimenti ante L.240 ordinati per voto medio (I=v/n) all'interno delle strutture (elencate in ordine alfabetico) e distribuzione dei lavori nelle classi di merito (Eccellente -E-, Buono -B-, Accettabile –A-, Limitato -L-). Per "somma punteggi (v)" si intende la valutazione complessiva dei lavori appartenenti alla struttura ottenuta sommando i punteggi dei singoli lavori. Per # lavori attesi si intende il numero di lavori attesi calcolato sulla base del SSD di afferenza dei soggetti valutati e del numero di lavori che da bando questi erano tenuti a sottomettere inviare alla VQR. La categoria "lavori penalizzati" contiene i lavori non valutabili e casi accertati di plagio o frode così come previsto dal bando VQR del 7 Novembre 2011, i lavori mancanti (cioè attesi e non sottomessi), i lavori identici presentati più volte dalla stessa struttura e i lavori identici presentati più volte dallo stesso soggetto valutato per due strutture di tipologia differente (es. Università ed Ente di Ricerca). R rappresenta il voto medio normalizzato secondo la formula (4). (n/N) x 100 rappresenta la percentuale dei lavori attesi dal dipartimento rispetto ai lavori attesi dell' area (N: numero lavori attesi dell'Area; n:

numero lavori attesi del dipartimento). IRD1 è il primo indicatore di qualità della ricerca applicato ai dipartimenti definito nel Bando VQR come rapporto tra il punteggio complessivo raggiunto da un dipartimento in una data area e la valutazione complessiva dell'area stessa. Gli indicatori R ed IRAS1 riportati per le sottostrutture degli Enti di Ricerca sono stati calcolati rispettivamente sull'insieme dei soli Enti di Ricerca o dei soli Consorzi. La tabella non include i dipartimenti con meno di 10 lavori attesi.

Tabella 4.12: Graduatorie dei dipartimenti post L.240 per sub-GEV ordinati per voto medio (I=v/n) e distribuzione dei lavori nelle classi di merito (Eccellente -E-, Buono -B-, Accettabile –A-, Limitato -L-). Per sub-GEV si intende un sottogruppo di SSD interni all'area con affinità disciplinari costituiti secondo i criteri del GEV. Per "somma punteggi (v)" si intende la valutazione complessiva dei lavori appartenenti al dipartimento ottenuta sommando i punteggi dei singoli lavori. Per # lavori attesi si intende il numero di lavori attesi calcolato sulla base del SSD di afferenza dei soggetti valutati e del numero di lavori che da bando questi erano tenuti a inviare alla VQR. La categoria "lavori penalizzati" contiene i lavori non valutabili e casi accertati di plagio o frode così come previsto dal bando VQR del 7 Novembre 2011, i lavori mancanti (cioè attesi e non sottomessi), i lavori identici presentati più volte dalla stessa struttura e i lavori identici presentati più volte dallo stesso soggetto valutato per due strutture di tipologia differente (es. Università ed Ente di Ricerca). Alcune celle sono state omesse al fine di proteggere la riservatezza dei dati personali, per questo motivo i valori di sintesi riportati nell'ultima riga non corrispondono ai totali di colonna.

Tabella 4.13: Graduatorie dei dipartimenti post L.240 per SSD ordinati per voto medio (I=v/n) e distribuzione dei lavori nelle classi di merito (Eccellente -E-, Buono -B-, Accettabile –A-, Limitato -L-). Per "somma punteggi (v)" si intende la valutazione complessiva dei lavori appartenenti al dipartimento ottenuta sommando i punteggi dei singoli lavori. Per # lavori attesi si intende il numero di lavori attesi calcolato sulla base del SSD di afferenza dei soggetti valutati e del numero di lavori che da bando questi erano tenuti a inviare alla VQR. La categoria "lavori penalizzati" contiene i lavori non valutabili e casi accertati di plagio o frode così come previsto dal bando VQR del 7 Novembre 2011, i lavori mancanti (cioè attesi e non sottomessi), i lavori identici presentati più volte dalla stessa struttura e i lavori identici presentati più volte dallo stesso soggetto valutato per due strutture di tipologia differente (es. Università ed Ente di Ricerca). Alcune celle sono state omesse al fine di proteggere la riservatezza dei dati personali, per questo motivo i valori di sintesi riportati nell'ultima riga non corrispondono ai totali di colonna.

Tabella 4.14: Graduatoria delle strutture per voto medio (VM) dei soggetti valutati (SV) e distribuzione in classi di voto medio. Per VM si intende il voto medio dei ricercatori della struttura calcolato come rapporto tra la somma del punteggio medio dei lavori presentati da ciascun ricercatore e il numero di soggetti valutati della struttura. Le 8 classi di voto medio contengono rispettivamente le percentuali dei soggetti valutati con VM<0, VM=0, 0<VM<=0.2, 0.2<VM<=0.4, 0.4<VM<=0.6, 0.6<VM<=0.8, 0.8<VM<=0.99, VM=1. Inoltre la tabella riporta il numero di soggetti valutati i cui lavori hanno tutti classe finale E (eccellente) assunti prima e dopo il 2004. Alcune celle sono state omesse al fine di

proteggere la riservatezza dei dati personali, per questo motivo i valori di sintesi riportati nell'ultima riga non corrispondono ai totali di colonna.

Tabella 4.15: Graduatoria dei dipartimenti ex L.240 per voto medio (VM) dei soggetti valutati (SV) e distribuzione in classi di voto medio. Per VM si intende il voto medio dei ricercatori del dipartimento calcolato come rapporto tra la somma del punteggio medio dei lavori presentati da ciascun ricercatore e il numero di soggetti valutati del dipartimento. Le 8 classi di voto medio contengono rispettivamente le percentuali dei soggetti valutati con VM<0, VM=0, 0<VM<=0.2, 0.2<VM<=0.4, 0.4<VM<=0.6, 0.6<VM<=0.8, 0.8<VM<=0.99, VM=1. Inoltre la tabella riporta il numero di soggetti valutati i cui lavori hanno tutti classe finale E (eccellente) assunti prima e dopo il 2006. Sulla base del Bando, e per i soli ricercatori delle università e ricercatori e tecnologi degli enti di ricerca, il numero di lavori attesi subiva delle riduzioni in funzione dell'anno della prese di servizio nei rispettivi ruoli. Per tutte le altre categorie, non erano previste riduzioni se non legate a determinate categorie di congedo. Nel costruire la tabella, si è identificato l'anno della presa di servizio con il numero di lavori attesi sulla base del Bando, senza tenere conte di eventuali congedi. Alcune celle sono state omesse al fine di proteggere la riservatezza dei dati personali, per questo motivo i valori di sintesi riportati nell'ultima riga non corrispondono ai totali di colonna.

Tabella 4.16: Numero soggetti valutati attivi, non attivi, parzialmente attivi e parzialmente inattivi per struttura. La colonna "# SV" indica il numero di soggetti valutati (SV) della struttura. Per soggetti valutati attivi (SV Attivi) si intendono i soggetti valutati che hanno conferito un numero di lavori uguale al numero di lavori attesi. Per soggetti valutati non attivi (SV NA) si intendono i soggetti valutati che non hanno presentato alcun lavoro. Per soggetti valutati parzialmente attivi (SV PA) si intendono i soggetti valutati che non hanno presentato tutti i lavori attesi ma un numero superiore alla metà. Per soggetti valutati parzialmente inattivi (SV PI) si intendono i soggetti valutati che hanno presentato al più la metà dei lavori attesi. La colonna "% (NA+PA+PI)" rappresenta la quota di soggetti valutati non pienamente attivi ottenuta considerando i non attivi (NA), i parzialmente attivi (PA) e i parzialmente inattivi (PI). La tabella non include i dipartimenti con meno di 4 soggetti valutati.

Tabella 4.17: Punteggio medio dei soggetti valutati (SV) e distribuzione del punteggio per ruolo universitario (ricercatore -RU, professore associato -PA, professore ordinario -PO) e sub-GEV del soggetto valutato. Per sub-GEV si intende un sottogruppo di SSD interni all'area con affinità disciplinari costituiti secondo i criteri del GEV. Per VM si intende il voto medio dei ricercatori del sub-GEV calcolato come rapporto tra la somma del punteggio medio dei lavori presentati da ciascun ricercatore e il numero di soggetti valutati del sub-GEV. Le 8 classi di voto medio contengono rispettivamente le percentuali dei soggetti valutati con VM<0, VM=0, 0<VM<=0.2, 0.2<VM<=0.4, 0.4<VM<=0.6, 0.6<VM<=0.8, 0.8<VM<=0.99, VM=1. Inoltre la tabella riporta il numero di soggetti valutati i cui lavori hanno tutti classe finale E (eccellente) assunti prima e dopo il 2006. Sulla base del Bando, e per i soli ricercatori delle università e ricercatori e tecnologi degli enti di ricerca, il numero di

lavori attesi subiva delle riduzioni in funzione dell'anno della prese di servizio nei rispettivi ruoli. Per tutte le altre categorie, non erano previste riduzioni se non legate a determinate categorie di congedo. Nel costruire la tabella, si è identificato l'anno della presa di servizio con il numero di lavori attesi sulla base del Bando, senza tenere conte di eventuali congedi.

Tabella 4.18: Punteggio medio dei soggetti valutati (SV) e distribuzione del punteggio per anno di nascita del soggetto valutato e sub-GEV. Per sub-GEV si intende un sottogruppo di SSD interni all'area con affinità disciplinari costituiti secondo i criteri del GEV. Per VM si intende il voto medio dei ricercatori del sub-GEV calcolato come rapporto tra la somma del punteggio medio dei lavori presentati da ciascun ricercatore e il numero di soggetti valutati del sub-GEV. Le 8 classi di voto medio contengono rispettivamente le percentuali dei soggetti valutati con VM<0, VM=0, 0<VM<=0.2, 0.2<VM<=0.4, 0.4<VM<=0.6, 0.6<VM<=0.8, 0.8<VM<=0.99, VM=1. Inoltre la tabella riporta il numero di soggetti valutati i cui lavori hanno tutti classe finale E (eccellente) assunti prima e dopo il 2006. Sulla base del Bando, e per i soli ricercatori delle università e ricercatori e tecnologi degli enti di ricerca, il numero di lavori attesi subiva delle riduzioni in funzione dell'anno della prese di servizio nei rispettivi ruoli. Per tutte le altre categorie, non erano previste riduzioni se non legate a determinate categorie di congedo. Nel costruire la tabella, si è identificato l'anno della presa di servizio con il numero di lavori attesi sulla base del Bando, senza tenere conte di eventuali congedi.

Tabella 4.19: Punteggio medio dei soggetti valutati per sub-GEV e anno di assunzione in ruolo. Sulla base del Bando, e per i soli ricercatori delle università e ricercatori e tecnologi degli enti di ricerca, il numero di lavori attesi subiva delle riduzioni in funzione dell'anno della prese di servizio nei rispettivi ruoli. Per tutte le altre categorie, non erano previste riduzioni se non legate a determinate categorie di congedo. Nel costruire la tabella, si è identificato l'anno della presa di servizio con il numero di lavori attesi sulla base del Bando, senza tenere conte di eventuali congedi. Per sub-GEV si intende un sottogruppo di SSD interni all'area con affinità disciplinari costituiti secondo i criteri del GEV. I periodi di presa servizio sono quelli riportati nella Tabella 3 del Bando VQR che stabilisce il numero di lavori attesi. Il punteggio medio è calcolato come rapporto del punteggio complessivo e il numero dei lavori attesi. Il numero di lavori attesi è stato calcolato sulla base del SSD di afferenza dei soggetti valutati e del numero di lavori che da bando questi erano tenuti a sottomettere.

Appendice A: Comparison between peer review and bibliometric evaluation

Appendice B: Criteri per la valutazione dei prodotti di ricerca

Appendice C: Classification, imputation of missing values and ranking of journals

Appendice D: Analisi delle citazioni

Appendice E: La valutazione peer review e i revisori esterni