Istat
Istituto Nazionale di Statistica

Rilevazione statistica sull'innovazione nelle imprese CIS2008

2006-2008

Descrizione del file

Dicembre 2010

INDICE

1. Introduzione 3
2. Descrizione delle variabili 4
3. Metodologia statistica per la tutela della riservatezza.....6
4. Tutte le modalità della variabile NUTS sono state accorpate in una sola modalità; il dettaglio geografico dell'indirizzo dell'impresa sarà rilasciato soltanto a livello nazionale..... 7
4. Analisi del contenuto informativo.....9
5. Riferimenti bibliografici.....16
6. Contatti..... 16

1. Introduzione

I file di microdati sono collezioni campionarie di dati elementari relative ad alcune indagini svolte dall'Istat. I file di microdati per scopi di ricerca scientifica (MFR) possono essere rilasciati su richiesta motivata e previa autorizzazione del Presidente dell'Istituto, purché siano resi anonimi e privi di ogni riferimento che ne permetta il collegamento con singole persone fisiche e giuridiche. Per la rilevazione Statistica sull'innovazione nelle imprese (CIS4 e CIS3), è stata definita una metodologia statistica atta a garantire la tutela della riservatezza dei rispondenti. Due files MFR sono stati rilasciati sul sito Istat. La stessa metodologia è stata applicata per la rilevazione CIS2008.

Nella Sezione 2 sono fornite delle informazioni sintetiche sullo status di ciascuna variabile presente nel MFR. La metodologia per la tutela della riservatezza è brevemente descritta nella Sezione 3. Nella Sezione 4 alcuni effetti del metodo di riduzione del rischio di re-identificazione sono illustrati. Quest'ultime informazioni possono essere utilizzate per valutare l'impatto del metodo di riduzione del rischio.

In questo documento soltanto le modifiche dovute al metodo statistico di riduzione del rischio di violazione della riservatezza sono descritti.

Gli utenti del file di microdati per scopi di ricerca scientifica sono obbligati a rispettare le condizioni specificate nel modulo di richiesta.

2. Descrizione delle variabili

La protezione dei dati individuali è stata raggiunta modificando il contenuto informativo di alcune variabili. Questa sezione sintetizza i cambiamenti indotti dal metodo di protezione dei dati individuali.

Il MFR può essere fornito in formato fisso oppure in formato tab-delimited.
Il separatore decimale è ".".
Il carattere "." significa "missing oppure non applicabile".
Ciascun record del MFR contiene informazioni relative ad una singola impresa.
Il file MFR contiene 19904 records e 141 variabili.
La somma dei pesi (variabile *weightnr*) è pari a 208637.531.

Tabella 1. Lo status[1] di ciascuna variabile nel file MFR.

Variabile	Nome Variabile	SDC status	Breve descrizione
id	codice impresa Asia	**Modificata**	cancellata
NUTS	nuts livello 2 utilizzata per la costruzione dei domini di stima	**Modificata**	ricodificata a livello nazionale
nace	Attività economica (divisione NACErev.2)	**Modificata**	NACE.rev2 2-digit, NACE05-09 = NACE “B”; NACE11-12 = NACE “11-12”; NACE38-39 = NACE “38-39”
HO	sede dell'impresa capogruppo	**Modificata**	0: nessuna risposta; 1: Home country; 2: Europe + Paesi Candidati + Paesi EFTA; 3: il resto del mondo; NN: l'impresa non appartiene ad un gruppo; ZZ: l’impresa appartiene ad un gruppo ma il headoffice non è definito
rrdinx	spesa sostenuta per attività di r&s svolta all’interno dell’impresa nel 2008	**Modificata**	0.09% record sono stati perturbati
rrdexx	spesa sostenuta per l'acquisizione di servizi di r&s nel 208	**Modificata**	0.07% record sono stati perturbati
rmacx	spesa sostenuta per l'acquisizione di macchinari e attrezzature innovativi nel 2008	**Modificata**	0.08% record sono stati perturbati
roekx	spesa sostenuta per l'acquisizione di tecnologia non incorporata in beni	**Modificata**	0.06% record sono stati perturbati

[1] The status refers **only** to the application of the statistical disclosure control (SDC) methodology.

	strumentali nel 2008		
rprex	spesa sostenuta per il design, progettazione industriale e altre attività preliminari alla produzione e alla fornitura di servizi nel 2008	**Modificata**	0.07% record sono stati perturbati
rtrx	spesa sostenuta per le attività di formazione nel 2008	**Modificata**	0.09% record sono stati perturbati
rmarx	spesa sostenuta per il marketing di prodotti innovativi nel 2008	**Modificata**	0.07% record sono stati perturbati
rtot	spesa totale per l'innovazione sostenuta nel 2008	**Modificata**	0.11% record sono stati perturbati
turn06	fatturato 2006	**Modificata**	24.43% record sono stati perturbati
turn08	fatturato 2008	**Modificata**	24.30% record sono stati perturbati
EMP06	totale addetti 2006	**Modificata**	**"10-49":** 0-49 addetti; **"50-249":** 50-249 addetti; **"250+":** più di 250 addetti; **"50+":** soltanto per NACE 45, 65, 66, 68, 72, 77 e B, le categorie "50-249" e 250+" sono aggregate in una sola; **"ALL":** soltanto per NACE 53 tutte le categorie sono state aggregate.
EMP08	totale addetti 2008	**Modificata**	**"10-49":** 0-49 addetti; **"50-249":** 50-249 addetti; **"250+":** più di 250 addetti; **"50+":** soltanto per NACE 45, 65, 66, 68, 72, 77 e B, le categorie "50-249" e 250+" sono aggregate in una sola; **"ALL":** soltanto per NACE 53 tutte le categorie sono state aggregate.
STRB	domini di stratificazione	**Modificata**	cancellata
STRA	domini di stratificazione	**Modificata**	cancellata
weight	coefficiente di riporto all'universo senza correzione per non-rispondenti	**Modificata**	cancellata
weightnr	coefficiente di riporto all'universo utilizzato per la produzione delle stime (con correzione nrs)	Inalterata	

La sezione 3 fornisce più dettagli sul contenuto informativo delle variabili modificate.

3. Metodologia statistica per la tutela della riservatezza

3.1 Scenari di violazione della riservatezza

Le unità statistiche (le imprese) a rischio di violazione della riservatezza sono state identificate ipotizzando due scenari di violazione: uno definito dalle variabili strutturali ed il secondo definito dalle informazioni non-strutturali registrate nel file. I criteri utilizzati sono derivati da regole di frequenza, regole di dominanza e algoritmi di aggregazione definiti a partire dal concetto di densità.

3.2. Modifiche delle variabili

3.2.1. Soppressione delle variabili

Le seguenti variabili sono state cancellate dal file:

a. Nome dell'impresa (variabile *id*)
b. Informazioni sulla stratificazione (variabili *StRA*e *STRB*)
c. Coefficiente di riporto all'universo senza correzione per i non-rispondenti (variabile *weight*). Osservazione: coefficiente di riporto all'universo utilizzato per la produzione delle stime (con la correzione per i non-rispondenti) (variabile *weightnr*) non è stata cambiata.

Quando una variabile è stata cancellata, tutti i suoi valori sono stati sostituiti dal simbolo del valore mancante, cioè ".".

3.2.2. Ricodifica globale

La ricodifica globale è uno dei più semplici e noti metodi di protezione dei dati individuali. Si può applicare soltanto alle variabili categoriche. La ricodifica consiste nell'accorpamento di alcune modalità della variabile categorica. Ovviamente, le nuove modalità non hanno più il significato precedente.

Per il rilascio del MFR CIS2008 per scopi di ricerca scientifica, le seguenti variabili sono state modificate:

1. *L'attività economica* (*nace*) è stata ricodificata come segue:
   a. Il dettaglio gerarchico è stato mantenuto soltanto a livello di divisione (NACErev.2 2 digit);
   b. Le divisioni NACE 05, 06, 07, 08 e 09 sono state accorpate in una sola modalità denominata "NACE B";
   c. Le divisioni NACE 11 e 12 sono state accorpate in una sola modalità denominata "NACE 11-12";

d. Le divisioni NACE 38 e 39 sono state accorpate in una sola modalità denominata “NACE 38-39”.

2. *Numero di addetti* (*EMP08*) è stata ricodificata in una variabile categorica con tre modalità principali. Per il rilascio del MFR CIS2008, le seguenti modalità sono state ottenute:
    a. Piccole imprese: 10 – 49 addetti, modalità “10-49”;
    b. Medie imprese: 50 – 249 addetti, modalità “50-249”;
    c. Grandi imprese: più di 250 addetti, modalità “250+”;
    d. Soltanto per le divisioni NACE 45, 65, 66, 68, 72, 77 e B, la modalità “50-249” e la modalità “250+” sono state accorpate in una nuova modalità denominata “50+”;
    e. Soltanto per le divisioni NACE 53 tutte le modalità di *EMP08* sono state accorpate in una nuova modalità “ALL”.
3. *Numero di addetti all’inizio del periodo di riferimento* (*EMP06*) è stata ricodificata in una variabile categorica con tre modalità principali. Per il rilascio del MFR CIS2008, le seguenti modalità sono state ottenute:
    a. Piccole imprese: 10 – 49 addetti, modalità “10-49”;
    b. Medie imprese: 50 – 249 addetti, modalità “50-249”;
    c. Grandi imprese: più di 250 addetti, modalità “250+”;
    d. Soltanto per le divisioni NACE 45, 65, 66, 68, 72, 77 e B, la modalità “50-249” e la modalità “250+” sono state accorpate in una nuova modalità denominata “50+”;
    e. Soltanto per le divisioni NACE 53 tutte le modalità di *EMP08* sono state accorpate in una nuova modalità “ALL”.

4. Tutte le modalità della variabile NUTS sono state accorpate in una sola modalità; il dettaglio geografico dell’indirizzo dell’impresa sarà rilasciato soltanto a livello nazionale.

5. Le modalità della variabile *HO* sono state ricodificate ottenendo le seguenti modalità: 1 (Home), 2 (Europa + Paesi Candidati + Paesi EFTA), 3 (il resto del mondo), 0 (nessuna risposta), NN (l’impresa non appartiene a nessun gruppo di imprese).

3.2.3. Cifre significative per le variabili numeriche

Le variabili numeriche sono state arrotondate; la base di arrotondamento utilizzata è pari all'unità. Soltanto le variabili *TURNMAR, TURNIN* e *TURNUNG* sono rilasciate con due decimali. Soltanto i coefficienti di riporto all'universo (variabile *weightnr*) è rilasciata con tre decimali.

3.3. Protezione delle imprese

Un'ulteriore livello di protezione è stato raggiunto dalla perturbazione delle variabili numeriche considerate identificative nei due scenari di violazione della riservatezza. Soltanto per le variabili chiave o sensibili, i valori assunti sulle unità a rischio di re-identificazione sono stati modificati. Queste modifiche riguardano:

a) I valori della variabile *turn08* sulle unità a rischio di re-identificazione nelle scenario definito a partire dalle informazioni strutturali

- Per ciascuna combinazione di *nace* e *EMP08*, i valori della variabile *Turn08* sulle unità a rischio sono state modificate utilizzando l'imputazione dall'unità non a rischio più vicina. Per ciascuna combinazione di *Nace* e *EMP08*, sulle code della distribuzione di *turn08*, il metodo individual ranking è stato applicato. La perturbazione applicata alla variabile *turn06* è stata generalmente proporzionale alla perturbazione della variabile of *turn08*.

b) I valori delle variabili *rrdinx, rrdexx, rmacx, roekx, rprex, rtrx, rmarx, rtot* sulle unità a rischio di re-identificazione nello scenario definito a partire dalle informazioni non-strutturali

- Ad eccezione di un numero non significativo di valori, la perturbazione dei valori di *rtot* è stata proporzionale alla perturbazione di *turn08*. Lo stesso criterio di perturbazione è stato utilizzato anche per tutti i componenti di *rtot*.

4. Analisi del contenuto informativo

In questa sezione alcuni strumenti utili alla valutazione dell'impatto del metodo di perturbazione sono presentati.

Per il rilascio del file MFR CIS2008, tutti i totali pesati della variabile *turn08* sono stati mantenuti esattamente, per ciascuna combinazione di *nace* e *EMP08*.

I coefficienti di riporto all'universo (variabile *weightnr*) non sono stati modificati.

4.1. Numero di valori di turn08 modificati

Rispetto allo scenario di violazione della riservatezza definito a partire dalle informazioni strutturali, *turn08* è stata l'unica variabile modificata. La perturbazione è stata applicata soltanto alle unità a rischio di re-identificazione. I valori delle unità non a rischio di re-identificazione sono rilasciati non perturbati. Nella Tabella 2 le percentuali dei valori di *turn08* unità modificate sono presentate per ciascuna combinazione di *nace* and *EMP08*.

Tabella 2. Percentuale di unità modificate.

nace	*EMPp08*	Numero Osservazioni	Percentuale di unità modificate	*nace*	*EMP08*	Numero Osservazioni	Percentuale di unità modificate
10	250+	63	30.16	37	10-49	48	25.00
10	50-249	128	21.88	37	50-249	11	100.00
11-12	10-49	73	26.03	38-39	10-49	244	22.13
11-12	250+	11	100.00	38-39	250+	41	53.66
11-12	50-249	24	33.33	38-39	50-249	89	32.58
13	10-49	181	23.76	41	10-49	1546	30.47
13	250+	33	42.42	41	250+	21	33.33
13	50-249	91	23.08	41	50-249	160	28.75
14	10-49	282	10.28	42	10-49	242	29.75
14	250+	34	32.35	42	250+	22	40.91
14	50-249	81	29.63	42	50-249	83	33.73
15	10-49	137	25.55	43	10-49	2149	26.06
15	250+	10	100.00	43	250+	12	100.00
15	50-249	57	8.77	43	50-249	133	28.57
16	10-49	280	11.07	45	10-49	998	24.75
16	250+	9	100.00	45	50+	130	10.00
16	50-249	50	30.00	46	10-49	845	23.43
17	10-49	105	13.33	46	250+	89	20.22
17	250+	19	57.89	46	50-249	198	24.24
17	50-249	66	10.61	47	10-49	861	12.66
18	10-49	370	10.27	47	250+	152	21.05
18	250+	11	100.00	47	50-249	164	25.00
18	50-249	34	47.06	49	10-49	606	24.09
19	10-49	29	51.72	49	250+	93	39.78
19	250+	11	100.00	49	50-249	94	32.98
19	50-249	16	25.00	50	10-49	23	17.39

20	10-49	56	12.50	50	250+	22	27.27
20	250+	41	24.39	50	50-249	19	26.32
20	50-249	52	19.23	51	10-49	22	54.55
21	10-49	13	100.00	51	250+	6	100.00
21	250+	39	25.64	51	50-249	2	100.00
21	50-249	24	29.17	52	10-49	143	26.57
22	10-49	248	27.02	52	250+	86	10.47
22	250+	40	25.00	52	50-249	118	25.42
22	50-249	157	25.48	53	ALL	21	42.86
23	10-49	253	12.25	55	10-49	607	12.85
23	250+	51	23.53	55	250+	21	19.05
23	50-249	75	30.67	55	50-249	96	10.42
24	10-49	105	11.43	56	10-49	644	25.93
24	250+	42	16.67	56	250+	41	14.63
24	50-249	105	37.14	56	50-249	64	12.50
25	10-49	538	30.30	58	10-49	56	8.93
25	250+	51	45.10	58	250+	12	100.00
25	50-249	147	25.85	58	50-249	31	12.90
26	10-49	96	28.13	59	10-49	42	14.29
26	250+	28	42.86	59	250+	3	100.00
26	50-249	38	13.16	59	50-249	9	100.00
27	10-49	91	27.47	61	10-49	19	21.05
27	250+	46	10.87	61	250+	15	100.00
27	50-249	58	24.14	61	50-249	9	100.00
28	10-49	242	11.16	62	10-49	144	32.64
28	250+	121	21.49	62	250+	50	10.00
28	50-249	95	26.32	62	50-249	81	28.40
29	10-49	33	12.12	63	10-49	110	23.64
29	250+	59	30.51	63	250+	12	100.00
29	50-249	26	38.46	63	50-249	37	18.92
30	10-49	24	20.83	64	10-49	140	23.57
30	250+	24	12.50	64	250+	163	15.34
30	50-249	17	52.94	64	50-249	184	29.89
31	10-49	178	25.84	65	10-49	27	40.74
31	250+	19	31.58	65	50+	66	22.73
31	50-249	86	9.30	66	10-49	189	19.58
32	10-49	206	16.99	66	50+	34	11.76
32	250+	12	100.00	68	10-49	128	26.56
32	50-249	59	11.86	68	50+	24	58.33
33	10-49	351	23.36	71	10-49	186	10.22
33	250+	9	100.00	71	250+	15	100.00
33	50-249	47	14.89	71	50-249	54	12.96
35	10-49	106	42.45	72	10-49	45	37.78
35	250+	29	34.48	72	50+	20	55.00
35	50-249	49	36.73	77	10-49	72	11.11
36	10-49	25	36.00	77	50+	19	31.58
36	250+	16	25.00	B	10-49	172	31.98
36	50-249	39	20.51	B	50+	23	21.74

4.2. Confronto delle varianze

I rapporti fra varianze di *turn08* prima e dopo la protezione sono presentati

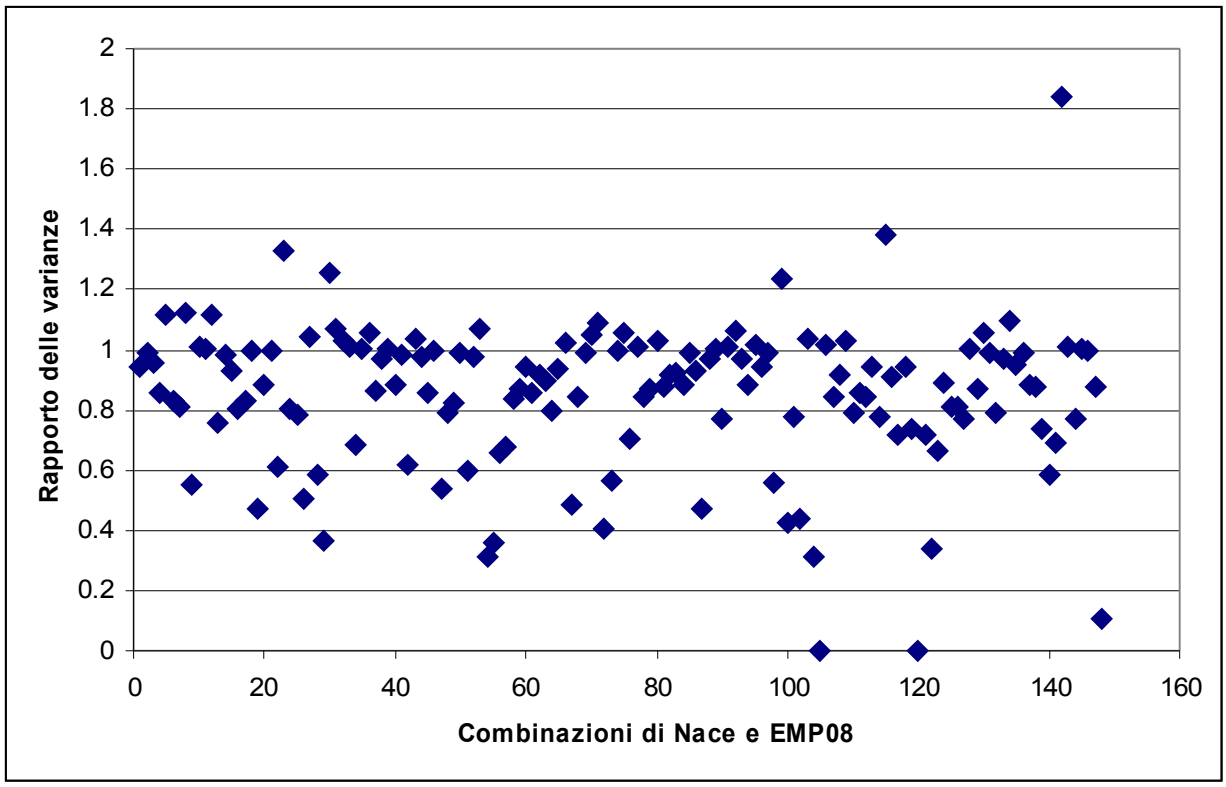
nella Tabella 3 e nella Figura 1.

Tabella 3. Rapporti delle varianze di *turn08*.

nace	*EMP08*	Rapporto delle varianze	*nace*	*EMP08*	Rapporto delle varianze
10	250+	0.94	37	10-49	1.06
10	50-249	0.99	37	50-249	0.70
11-12	10-49	0.96	38-39	10-49	1.01
11-12	250+	0.86	38-39	250+	0.84
11-12	50-249	1.12	38-39	50-249	0.87
13	10-49	0.83	41	10-49	1.03
13	250+	0.81	41	250+	0.88
13	50-249	1.12	41	50-249	0.92
14	10-49	0.55	42	10-49	0.92
14	250+	1.01	42	250+	0.88
14	50-249	1.01	42	50-249	0.99
15	10-49	1.12	43	10-49	0.93
15	250+	0.76	43	250+	0.47
15	50-249	0.98	43	50-249	0.97
16	10-49	0.93	45	10-49	1.00
16	250+	0.81	45	50+	0.77
16	50-249	0.83	46	10-49	1.01
17	10-49	1.00	46	250+	1.07
17	250+	0.47	46	50-249	0.97
17	50-249	0.88	47	10-49	0.88
18	10-49	1.00	47	250+	1.02
18	250+	0.61	47	50-249	0.94
18	50-249	1.33	49	10-49	0.99
19	10-49	0.80	49	250+	0.56
19	250+	0.78	49	50-249	1.23
19	50-249	0.50	50	10-49	0.43
20	10-49	1.04	50	250+	0.78
20	250+	0.58	50	50-249	0.44
20	50-249	0.37	51	10-49	1.03
21	10-49	1.26	51	250+	0.32
21	250+	1.07	51	50-249	0.00
21	50-249	1.03	52	10-49	1.02
22	10-49	1.01	52	250+	0.84
22	250+	0.68	52	50-249	0.92
22	50-249	1.00	53	ALL	1.03
23	10-49	1.06	55	10-49	0.79
23	250+	0.87	55	250+	0.86
23	50-249	0.97	55	50-249	0.84
24	10-49	1.01	56	10-49	0.95
24	250+	0.88	56	250+	0.78
24	50-249	0.98	56	50-249	1.38
25	10-49	0.62	58	10-49	0.91
25	250+	1.03	58	250+	0.71
25	50-249	0.97	58	50-249	0.94
26	10-49	0.85	59	10-49	0.73
26	250+	1.00	59	250+	0.00
26	50-249	0.54	59	50-249	0.72
27	10-49	0.79	61	10-49	0.34
27	250+	0.83	61	250+	0.67
27	50-249	0.99	61	50-249	0.89
28	10-49	0.60	62	10-49	0.81

28	250+	0.98	62	250+	0.81
28	50-249	1.07	62	50-249	0.77
29	10-49	0.31	63	10-49	1.00
29	250+	0.36	63	250+	0.87
29	50-249	0.66	63	50-249	1.05
30	10-49	0.68	64	10-49	0.99
30	250+	0.84	64	250+	0.79
30	50-249	0.87	64	50-249	0.97
31	10-49	0.94	65	10-49	1.10
31	250+	0.85	65	50+	0.95
31	50-249	0.92	66	10-49	0.99
32	10-49	0.90	66	50+	0.89
32	250+	0.80	68	10-49	0.88
32	50-249	0.94	68	50+	0.74
33	10-49	1.03	71	10-49	0.58
33	250+	0.48	71	250+	0.69
33	50-249	0.84	71	50-249	1.84
35	10-49	0.99	72	10-49	1.01
35	250+	1.05	72	50+	0.77
35	50-249	1.09	77	10-49	1.01
36	10-49	0.41	77	50+	1.00
36	250+	0.57	B	10-49	0.87
36	50-249	1.00	B	50+	0.10

Figura 1. Confronto delle varianze di *turn08* prima e dopo la protezione.



4.3. Confronto dei coefficienti di correlazione

Nella tabella 4 sono presentati alcuni confronti fra coefficienti di correlazione calcolati utilizzando valori originali e perturbati di *turn08* e *rtot*. Effetti simili sono stati osservati

utilizzando gli altri componenti di *rtot*. I missing sono dovuti al numero estremamente ridotto di valori non-missing.

Tabella 4. Coefficienti di correlazione.

nace	*EMP08*	cor(*turn08,turn08Protected*)	cor(*ttot,ttotProtected*)	cor(*turn08,ttot*)	cor(*turn08Protected,rtotProtected*)	*nace*	*EMP08*	cor(*turn08,turn08Protected*)	cor(*rtot,rtotProtected*)	cor(*turn08,rtot*)	cor(*turn08Protected,rtotProtected*)
10	10-49	0.97	1	0.19	0.16	36	50-249	1	1	0.14	0.16
10	250+	1	1	0.69	0.69	37	10-49	0.98	1	0.25	0.24
10	50-249	1	1	0.3	0.3	37	50-249	0.81	1	0.3	0
11-12	10-49	1	1	0.44	0.43	38-39	10-49	1	1	0.11	0.11
11-12	250+	0.93	1	0.7	0.59	38-39	250+	0.91	1	0.24	0.26
11-12	50-249	0.99	1	0.39	0.37	38-39	50-249	0.99	1	0.17	0.23
13	10-49	0.92	1	0.4	0.38	41	10-49	1	1	0.21	0.21
13	250+	0.94	1	0.69	0.65	41	250+	0.9	1	0.71	0.5
13	50-249	0.99	1	0.32	0.31	41	50-249	0.99	1	0.1	0.11
14	10-49	0.85	1	0.38	0.41	42	10-49	0.99	1	0.11	0.12
14	250+	1	1	0.72	0.71	42	250+	1	1	0.26	0.24
14	50-249	1	1	0.23	0.23	42	50-249	0.98	1	0.08	0.1
15	10-49	0.97	1	0.23	0.22	43	10-49	0.98	1	0.24	0.24
15	250+	0.87	1	0.64	0.56	43	250+	0.68	1	0.14	0.4
15	50-249	0.98	1	0.42	0.39	43	50-249	0.99	1	0.17	0.17
16	10-49	0.98	1	0.17	0.19	45	10-49	1	1	0.1	0.1
16	250+	0.85	1	0.67	0.34	45	50+	0.9	1	0.1	0.23
16	50-249	0.92	1	0.3	0.34	46	10-49	1	1	0.11	0.11
17	10-49	1	1	0.15	0.15	46	250+	1	1	0.21	0.21
17	250+	0.78	1	0.29	0.24	46	50-249	1	1	-0.03	-0.03
17	50-249	0.96	1	0.12	0.12	47	10-49	0.96	1	0.08	0.08
18	10-49	0.9	1	0.11	0.1	47	250+	0.99	1	0.31	0.3
18	250+	0.82	1	0.12	-0.01	47	50-249	0.97	1	-0.02	-0.02
18	50-249	0.95	1	0.18	0.11	49	10-49	1	1	0.14	0.14
19	10-49	0.93	1	0.08	0.1	49	250+	0.71	0.99	0.3	0.43
19	250+	0.89	1	0.69	0.35	49	50-249	0.99	1	0.38	0.39
19	50-249	0.71	1	-0.02	0.53	50	10-49	0.72	1	-0.09	0.07
20	10-49	0.95	1	0.27	0.29	50	250+	0.93	1	0.69	0.52
20	250+	0.86	0.95	0.79	0.83	50	50-249	0.81	1	-0.1	-0.12
20	50-249	0.69	1	0.63	0.29	51	10-49	0.96	1	0.24	0.13
21	10-49	0.8	1	0.13	0.1	51	250+	0.66	1	-0.16	0.45
21	250+	1	1	0.22	0.23	51	50-249	NA	NA	NA	NA
21	50-249	0.99	1	0.38	0.4	52	10-49	0.99	1	0.29	0.27
22	10-49	1	1	0.39	0.39	52	250+	0.91	1	0.38	0.38
22	250+	0.76	1	0.15	0.48	52	50-249	0.96	1	0.52	0.39
22	50-249	1	1	0.23	0.22	53	ALL	1	1	1	1
23	10-49	0.97	1	0.16	0.15	55	10-49	0.86	1	0.06	0.07
23	250+	0.96	1	0.45	0.48	55	250+	1	1	0.07	0.08
23	50-249	0.99	1	0.28	0.28	55	50-249	0.96	1	0.26	0.39

24	10-49	0.99	1	0.36	0.31	56	10-49	0.99	1	0.11	0.12
24	250+	0.91	0.97	0.8	0.66	56	250+	0.86	0.97	0.87	0.53
24	50-249	0.99	1	0.14	0.13	56	50-249	0.95	1	-0.01	-0.02
25	10-49	0.86	1	0.16	0.22	58	10-49	0.93	1	-0.07	-0.07
25	250+	1	1	0.46	0.44	58	250+	0.9	1	0.6	0.73
25	50-249	0.97	1	0.3	0.33	58	50-249	0.97	1	0.06	0.01
26	10-49	0.99	1	0.47	0.49	59	10-49	0.92	1	0	-0.03
26	250+	0.92	0.99	0.86	0.72	59	250+	NA	NA	NA	NA
26	50-249	0.88	1	0.82	0.69	59	50-249	0.67	1	-0.05	-0.1
27	10-49	0.93	1	0.41	0.42	61	10-49	0.86	1	0.03	0.36
27	250+	0.92	0.99	0.8	0.65	61	250+	0.81	0.99	0.92	0.58
27	50-249	0.99	1	0.54	0.5	61	50-249	0.84	1	0	-0.09
28	10-49	0.9	1	0.14	0.21	62	10-49	0.95	1	0.2	0.25
28	250+	0.99	1	0.87	0.86	62	250+	0.91	0.97	0.88	0.67
28	50-249	1	1	0.25	0.24	62	50-249	0.89	1	0.05	0.07
29	10-49	0.75	1	0.27	0.28	63	10-49	1	1	0.25	0.25
29	250+	0.59	0.95	0.16	0.44	63	250+	0.94	1	0.38	0.38
29	50-249	0.96	1	0.27	0.35	63	50-249	0.99	1	0.11	0.12
30	10-49	0.92	1	0.07	0.37	64	10-49	1	1	0.11	0.12
30	250+	0.88	0.99	0.75	0.62	64	250+	0.9	0.99	0.39	0.53
30	50-249	0.99	1	0.53	0.5	64	50-249	1	1	0.43	0.43
31	10-49	0.99	1	0.24	0.24	65	10-49	1	1	0.5	0.51
31	250+	0.98	1	0.36	0.45	65	50+	0.97	1	0.27	0.23
31	50-249	0.98	1	0.41	0.48	66	10-49	1	1	0.37	0.36
32	10-49	0.88	1	0.53	0.39	66	50+	0.93	0.99	0.72	0.6
32	250+	0.89	1	0	0.06	68	10-49	0.98	1	0.1	0.13
32	50-249	0.97	1	0.21	0.31	68	50+	0.89	1	0.19	0.23
33	10-49	1	1	0.11	0.11	71	10-49	0.94	1	0.1	0.14
33	250+	0.74	1	0.9	0.46	71	250+	0.83	0.97	-0.04	0.18
33	50-249	1	1	-0.05	-0.04	71	50-249	0.98	1	0.1	0.1
35	10-49	1	1	-0.02	-0.02	72	10-49	0.99	1	0.7	0.65
35	250+	0.99	1	0.25	0.26	72	50+	0.74	0.92	0.14	0.54
35	50-249	0.98	1	-0.02	-0.03	77	10-49	0.98	1	0.15	0.17
36	10-49	0.66	1	0.85	0.48	77	50+	1	1	0.06	0.06
36	250+	0.81	1	0.59	0.53	B	10-49	0.95	1	0.26	0.31
						B	50+	0.56	1	1	0.55

4.4. Variazioni rispetto all'inizio del periodo di riferimento

L'impatto del metodo di perturbazione sulla variabile *turn06* è stato valutato utilizzando i rapporti fra *turn06* e *turn08*. Nella tabella 5 sono presentati i confronti fra alcuni indicatori statistici calcolati sui rapporti *turn06/turn08* originali e sui rapporti *turn06/turn08* perturbati. Sono presentate soltanto le combinazioni di *nace* e *EMP08* che sono state modificate rispetto a questi criteri. Il rimanente 83% delle combinazioni è rimasto invariato rispetto alla distribuzione di *turn06/turn08*.

Tabella 5. Valutazione della perturbazione di *turn06/turn08* (valori arrotondati).

nace	EMP08	NbObs	MeanPerturbed/MeanOrig	VariancePerturbed/VarianceOrig	Corr(Perturbed,Orig)
11-12	250+	11	0.97	1.02	0.83
13	250+	33	0.98	1.43	0.89
15	250+	10	0.99	0.53	0.80
16	250+	9	1.00	0.71	0.79
18	250+	11	0.97	0.97	0.77
20	250+	41	1.00	1.00	1.00
22	250+	40	1.00	1.00	0.98
24	250+	42	0.99	1.00	1.00
26	250+	28	1.00	1.01	1.00
28	250+	121	1.00	1.04	0.99
31	250+	19	1.00	0.98	0.95
32	250+	12	1.00	0.96	0.96
33	250+	9	0.90	1.37	0.76
41	250+	21	1.02	1.81	0.88
43	250+	12	0.98	1.10	0.90
45	50+	130	1.00	1.01	1.00
50	250+	22	0.97	1.21	0.93
51	250+	6	0.96	0.92	0.82
53	ALL	21	1.37	19.07	0.17
55	250+	21	1.00	1.00	0.98
56	250+	41	0.99	1.03	1.00
63	250+	12	0.99	1.02	0.99
65	50+	66	1.00	1.08	0.98
66	50+	34	1.00	1.00	1.00
71	250+	15	0.97	1.16	0.86

4.5. I rapporti rtot/turn08 ed i altri componenti di rtot

Soltanto i rapporti *rtot/turn08*, *rrdinx/turn08*, *rrdexx/turn08*, *rmacx/turn08* sono stati presi in considerazione per questa valutazione dell'utilità.

Per ciascuna combinazione di *nace* e *EMP08*, i quantili dei rapporti *rtot/turn08*, *rrdinx/turn08*, *rrdexx/turn08*, *rmacx/turn08* sono stati calcolati sui dati originali e sui dati perturbati.

In seguito, per ciascuna combinazione di *nace* e *EMP08*, è stata calcolata la differenza fra i corrispondenti quantili originali e quantili perturbati.

La massima differenza fra quantili originali e perturbati è stata inferiore a 1.9, ma nella maggior parte dei casi, la differenza assoluta fra i quantili è stata inferiore a 0.5.

5. Riferimenti bibliografici

1. Ichim, D. (2007), "Microdata Anonymisation of the Community Innovation Survey Data: a Density Based Clustering Approach for Risk Assessment", *Documenti Istat*, 2, disponibile al sito www.istat.it.

2. Franconi, L. and Ichim, D. (2007), "Community Innovation Survey: comparable dissemination", Joint UNECE – Eurostat Worksession on Statistical Dislcosure Control, Manchester, Manchester, United Kingdom, 2007, disponibile al sito http://www.unece.org/stats/documents/2007.12.confidentiality.htm.

3. Ichim, D. (2009), "Disclosure Control for Business Microdata: a Density-Based Approach", *International Statistical Review*, (2009), 77, 2, 196–211.

4. Ichim, D. (2008), "Community Innovation Survey: a Flexible Approach to the Dissemination of Microdata Files for Research", European Conference on Quality in Official Statistics, disponibile al sito http://q2008.istat.it/.

5. ISTAT (2008), "Statistiche sull'innovazione delle imprese", Collana *Informazioni*, N. 1.

6. Contatti

Per eventuali informazioni sulla rilevazione CIS, rivolgersi a mastrost@istat.it.
Per eventuali informazioni sulla metodologia di tutela della riservatezza, rivolgersi a rilascio.microdati@istat.it.

Curatori dei capitoli

Il documento è stato redatto da Daniela Ichim.