The identification of Employer Enterprises Births (EEB)

According to the "new" suggested methodology, there are two conditions that allow to identify EEB in year xx:

(1) Real births in xx with employees>0
(2) Birth by growth i.e. active enterprises in xx with employees(xx)>0 existing as active units in the previous two years without employees

Let:

N_{xx}	active population (zero and >zero employees) in xx
$N(1)_{xx}$	active population (>zero employees) in xx
$N(0)_{xx-1}$	active population (zero employees) in xx-1
$N(1)_{xx-1}$	active population (>zero employees) in xx-1
$N(0)_{xx-2}$	active population (zero employees) in xx-2
$N(1)_{xx-2}$	active population (>zero employees) in xx-2

A merge by identification code of the three years populations determines the following pattern:

Table 1 - Pattern

N_{xx-2}	N_{xx-1}	N_{xx}	Result	Type
N(0)	**N(0)**	**N(1)**	**Yes**	Births by Growth
N(0)	N(1)	N(1)	No	-
N(0)	**missing**	**N(1)**	**Yes**	Births by Growth
N(1)	N(0)	N(1)	No	-
N(1)	N(1)	N(1)	No	-
N(1)	missing	N(1)	No	-
missing	**N(0)**	**N(1)**	**Yes**	Births by Growth
Missing	N(1)	N(1)	No	-
Missing	missing	N(1)	**Yes**	Real birth

Step 2c) In order to remove from Births by growth (2) some active units that grow because events of takeover, the following links for xx=2005 have been identified:

a) EEB_{2005} linked to $exits_{2004}$ that cease for events
b) EEB_{2005} linked to active $units_{2004\text{-}2005}$ that shrink for events
c) EEB_{2005} linked by continuity rules (RL) to $exits_{2004}$

However in the analysis of such events it is not clear whether the growing in terms of employees is due to the event or not. In fact what is feasible is to check just on the link, but to better investigate if that link causes a transfer of employment need different analysis. For that reason in someway maybe is better to not exclude such events. Data with and without events are then provided for further discussion and to better compare results among Countries.

The production of data

With regards to year 2005 - Nace C to K - in table 2 results from *standard* BD and *new* BD on employer enterprises are compared.

Table 2 – Results, absolute numbers

Nxx	Active population	3966758	BD standard
Rxx	Real births	308306	BD standard
Nxx in size class 1+	Active population in size class 1+	1209818	BD standard
N(1)	Active population with employees>0	1304591	BD New
(1)	Real births (xx) with employees>0	44442	BD New
(2)	Birth by growth including events	76943	BD New
(2)'	Birth by growth excluding events	74969	BD New
(1)+(2)	EEB including events	121385	EEB with events
(1)+(2)'	EEB excluding events	119411	EEB without events

Comments: Difference between N(1) and Nxx in size class 1+ is due to the different method of calculation of employment. In N(1) enterprises having on average in the year less then 0.50 employees are included, instead in the other population rounded figures are used (then 0.5 are excluded)

Table 3 – Births rates, year 2005 - comparison of methods

NACE	Birth rate Standard BD	Birth rate EEB*	Birth rate EEB**
C+E	5.3	4.6	4.5
D	4.9	5.8	5.7
F	10.1	12.5	12.4
G	6.3	9.0	8.8
H	7.4	10.8	10.5
I	7.3	9.7	9.6
J	8.7	8.3	8.2
K70.71.73	8.6	15.9	15.6
K72	10.0	8.0	7.8
K74	10.2	9.5	9.4
Total	7.8	9.3	9.2

* N1 obtained as N having average employees >0, events not excluded
** N1 obtained as N having average employees >0, events excluded

The identification of Employer Enterprises Deaths (EED)

According to the "new" suggested methodology, there are two conditions identifying EED in year xx:

(1) Real deaths (xx) with employees>0
(2) Deaths by decline i.e. active enterprises in xx with employees(xx)>0 existing as active in the following two years without employees

Step 2c: deaths by decline could be shrink because they split in another unit (new or existing) that takes part or the whole activity (measured in terms of employees) of the declining one. To identify those units a similar technique as births has been applied
Let:
N_{xx} active population (zero and >zero employees) in xx
$N(1)_{xx}$ active population (>zero employees) in xx

$N(0)_{xx+1}$ active population (zero employees) in xx+1
$N(1)_{xx-+1}$ active population (>zero employees) in xx+1
N_{xx+2} active population not available
$N'(1)_{xx+2}$ active units (>zero employees) in xx+2 as estimation from Social Security data
With regards to xx+2, in the current period of the year (October) it was possible to analyze the Social Security data referring to the number of employees in year xx+2=2006

Table 4 - Pattern

Nxx	Nxx+1	Nxx+2 Social security	Result	Type
N(1)	**N(0)**	**missing**	**yes**	Deaths by Decline
N(1)	N(0)	Y	no	
N(1)	**Missing**	**missing**	**yes**	Deaths by Decline
N(1)	N(1)	Y	no	
N(1)	N(1)	missing	No	
N(1)	**Missing**	**Missing**	**yes**	Real death

The production of data
With regards to year 2004 - Nace C to K - in table 5 results from *standard* BD and *new* BD are compared.
Table 5 – Results- absolute numbers

Nxx	Active population	3890937	BD standard
Dxx	Real deaths (xx)	282157	BD standard
Nxx in size class 1+	Active population in size class 1+	1170613	BD standard
N(1)	Active population with employees>0	1291735	New
(1)	Real deaths (xx) with employees>0	48237	New
(2)	Death by decline including events	65471	New
(2)'	Death by decline excluding events	63112	New
(1)+(2)	EED including events	113708	EED with events
(1)+(2)'	EED excluding events	111349	EED without events

Table 6 – Deaths rates, year 2004 - comparison of methods

NACE	Death rate Standard BD	Death rate EED*	Death rate EED**
C+E	4.1	4.5	4.2
D	6.2	7.1	7.0
F	8.1	11.4	11.3
G	7.4	8.5	8.3
H	6.2	9.3	8.9
I	7.6	9.1	9.0
J	9.1	6.9	6.7
K70.71.73	5.7	11.0	10.7
K72	8.6	7.5	7.3
K74	7.5	8.2	8.1
Total	7.3	8.8	8.6

* N1 obtained as N having average employees >0, events not excluded
** N1 obtained as N having average employees >0, events excluded

High Growth (HG) and gazelle

Identification of HG

Let Nxx be the active population in year xx, where xx=2005 and consider the three active population in xx-1 to xx-3

Step1: a merge by identification code produces AA=$N_{2002} \cap N_{2003} \cap N_{2004} \cap N_{2005}$

Step2: we exclude the real births in xx-3 i.e. Rb(2002) from AA, then Poten_population=AA-Rb(2002). Table 6 shows distribution of potential population for HG by size class at the beginning xx-3 (2002) and at the reference period xx (2005). The yellow area identify the potential population for HG having the threshold of 10 employees at the beginning period.

Table 6 – Potential population for HG

				Empl2005				
empl2002	**0**	**1**	**2-4**	**5-9**	**10-14**	**15-19**	**20+**	**Total**
0	1672471	111596	43514	6503	1479	526	851	1836940
1	70794	149104	68222	5597	789	204	224	294934
2-4	25676	44242	201802	43417	3432	689	633	319891
5-9	6621	3551	25942	88183	18540	2557	1448	146842
10-14	2311	716	2230	11952	28572	8252	2959	56992
15-19	986	239	552	1517	6002	9991	5903	25190
20+	1997	370	653	1076	1825	3896	48402	58219
Total	1780856	309818	342915	158245	60639	26115	60420	2739008

* Nace C to K in year=2002

Step3: If employees(2002) ≥ 10 then we obtain the population of reference for HG (140,410 for year 2005).

Step4: HG are obtained applying to the population of reference (step 3) the following rules to employees and to turnover:

HG_empl=yes if [empl(2005)/empl(2002)] $\geq$ **1.728**

HG_tur=yes if [turnover(2005)/turnover(2002)] $\geq$ **1.728**

Table 7 – HG enterprises, by type of growth

	hg_tur		
hg_empl	No	yes	total
no	0	8262	8262
yes	2749	3022	5771
Total	2749	11284	14033

Ratio= HG/pop_ref_hg= 14033**/**140410 *100= 10.0%

Step5: HG can grow because events of takeover. Looking for such events in the period 2002-2005, we have identified 979 (7%) of HG with events. By excluding such events we obtain the "pure" HG, HG*= 13054

Ratio*= HG*/pop_ref_hg= 13054**/**140410 *100= 9.3%

The difference between HG and HG* (having excluded events) is not so meaningful in terms of enterprises but to better understand effect on employment Table 8 shows the average of employment that refers to different HG subsets.

Table 8: average size of "pure" HG (coloured) and of HG with events

hgeve	hg_empl	hg_tur	_FREQ_	empl(2002) average	empl(2005) average
.	.	1	7961	48.167	55.343
.	1	.	2485	25.852	57.798
.	1	1	2608	38.276	98.819
1	.	1	301	253.801	305.887
1	1	.	264	84.322	198.101
1	1	1	414	110.972	454.391
Total			14033		

Gazelle

It is that subset of HG that are real births in xx-5 and xx-4. For xx=2005 gazelles are obtained by merging the real births 2000 and the real births 2001 with HG by identification code

Step1: potential gazelle are identified as the potential population of HG born in 2000 and 2001 (the subset of real births in 2000 and 2001 surviving in 2005, with employees(2002) ≥10).

Step2: Gazelle are then a subset of HG

Table 9 shows distribution of gazelle and potential gazelle by Nace*

NACE	poten_gazelle	%	gazelle	%	Gazelle Without events	%	Gazelle/ poten_gaz
C+E	54	0.74	17	1.01	17	1.05	31.5
D	2462	33.53	483	28.78	464	28.75	19.6
F	1390	18.93	348	20.74	343	21.25	25.0
G	888	12.09	173	10.31	163	10.1	19.5
H	485	6.62	70	4.17	68	4.21	14.4
I	694	9.45	214	12.75	207	12.83	30.8
J	64	0.87	19	1.13	18	1.12	29.7
K70.71.73	70	0.95	18	1.07	17	1.05	25.7
K72	258	3.51	53	3.16	47	2.91	20.5
K74	977	13.31	283	16.87	270	16.73	29.0
Total	7342	100.00	1678	100.00	1614	100	22.9

*Nace form C to K in year 2005