FDZ

Research Data Centre (FDZ)
of the German Federal
Employment Agency (BA)
at the Institute for
Employment Research (IAB)

May 2011

Description of the error in the LIAB variable "Status of establishment number" (betr_st) in certain LIAB datasets

The variable "Status of establishment number" (betr_st) is a generated variable that classifies employment records as follows:

1	Establishment of IAB Establishment Panel, in this LIAB version on reference date (surveyed within the time period covered by the employment record)
2	Establishment of IAB Establishment Panel, in this LIAB version, not on reference date (not surveyed within the time period covered by the employment record)
3	Establishment of IAB Establishment Panel, not in this LIAB version
4	No establishment of IAB Establishment Panel

In certain LIAB versions up to 25 percent of the employment records were misclassified as 2 instead of 1 and as 1 instead of 3. The following table shows in which LIAB datasets the variable is incorrect and in which it is correct:

Update	QM1/QM2	LM1	LM2	LM3	LM4/PM1	LM5/MM1
2003		ok	ok			
2004	ok	ok	ok			
2005	ok	ok	ok	ok		
2006	ok	incorrect	incorrect	incorrect		
2007	ok	incorrect	incorrect	incorrect	(incorrect)	(incorrect)
9808	ok					

Versions in brackets are not documented yet, but were released for test users.

The next table lists the incorrect files and the filenames of the revised datasets:

Filenames of incorrect data	Filenames of revised data
LIAB_LM1_2006_pers.dta	LIAB_LM1_2006_v2_pers.dta
LIAB_LM2_2006_pers.dta	LIAB_LM2_2006_v2_pers.dta
LIAB_LM3_2006_pers.dta	LIAB_LM3_2006_v2_pers.dta
LIAB_LM1_2007_pers.dta	LIAB_LM1_2007_v2_pers.dta
LIAB_LM2_2007_pers.dta	LIAB_LM2_2007_v2_pers.dta
LIAB_LM3_2007_pers.dta	LIAB_LM3_2007_v2_pers.dta
LIAB_PM1_2007_pers.dta	LIAB_PM1_2007_v2_pers.dta
LIAB_MM1_2007_pers.dta	LIAB_MM1_2007_v2_pers.dta

Cross tabulations of the incorrect (rows) and the revised (columns) variables:

Frequencies of errors are printed in bold face.

LIAB LM1, 1993-2006, 30 April, 2008, data on individuals

```
          Status of |      Status of establishment number
establishment number | 1 EPiVoD  2 EPiVnotD   3 EPnotV     4 NoEP |     Total
---------------------+------------------------------------------------+----------
1 EPiVoD             | 3,206,492          0     73,565          0 | 3,280,057
2 EPiVnotD           | 6,009,978  8,060,971          0          0 |14,070,949
3 EPnotV             |         0          0  2,106,160          0 | 2,106,160
4 NoEP               |         0          0          0 13,100,436 |13,100,436
---------------------+------------------------------------------------+----------
               Total | 9,216,470  8,060,971  2,179,725 13,100,436 |32,557,602
```

LIAB LM2, 1993-2006, 30 April, 2008, data on individuals

```
          Status of |      Status of establishment number
establishment number | 1 EPiVoD  2 EPiVnotD   3 EPnotV     4 NoEP |     Total
---------------------+------------------------------------------------+----------
1 EPiVoD             | 9,675,475          0    242,164          0 | 9,917,639
2 EPiVnotD           | 6,401,614 12,890,162          0          0 |19,291,776
3 EPnotV             |         0          0  3,187,855          0 | 3,187,855
4 NoEP               |         0          0          0 22,435,976 |22,435,976
---------------------+------------------------------------------------+----------
               Total |16,077,089 12,890,162  3,430,019 22,435,976 |54,833,246
```

LIAB LM3, 1993-2006, 30 April, 2008, data on individuals

```
          Status of |      Status of establishment number
establishment number | 1 EPiVoD  2 EPiVnotD   3 EPnotV     4 NoEP |     Total
---------------------+------------------------------------------------+----------
1 EPiVoD             | 1,861,203          0     57,249          0 | 1,918,452
2 EPiVnotD           | 5,907,739  3,684,466          0          0 | 9,592,205
3 EPnotV             |         0          0  2,103,788          0 | 2,103,788
4 NoEP               |         0          0          0 11,371,921 |11,371,921
---------------------+------------------------------------------------+----------
               Total | 7,768,942  3,684,466  2,161,037 11,371,921 |24,986,366
```

LIAB LM1, 1993-2007, 19 August 2009, data on individuals

```
          Status of |      Status of establishment number
establishment number | 1 EPiVoD  2 EPiVnotD   3 EPnotV     4 NoEP |     Total
---------------------+------------------------------------------------+----------
1 EPiVoD             | 3,285,118          0     75,395          0 | 3,360,513
2 EPiVnotD           | 6,208,574  8,540,365          0          0 |14,748,939
3 EPnotV             |         0          0  2,304,861          0 | 2,304,861
4 NoEP               |         0          0          0 14,112,884 |14,112,884
---------------------+------------------------------------------------+----------
               Total | 9,493,692  8,540,365  2,380,256 14,112,884 |34,527,197
```

LIAB LM2, 1993-2007, 19 August 2009, data on individuals

```
         Status of |       Status of establishment number
establishment number | 1 EPiVoD 2 EPiVnotD  3 EPnotV   4 NoEP |    Total
---------------------+--------------------------------------------+----------
1 EPiVoD             | 9,942,929         0   245,387        0 |10,188,316
2 EPiVnotD           | 6,682,191 13,808,521        0        0 |20,490,712
3 EPnotV             |         0         0 3,569,527        0 | 3,569,527
4 NoEP               |         0         0        0 24,536,883 |24,536,883
---------------------+--------------------------------------------+----------
               Total |16,625,120 13,808,521 3,814,914 24,536,883 |58,785,438
```

LIAB LM3, 1993-2007, 19 August 2009, data on individuals

```
         Status of |       Status of establishment number
establishment number | 1 EPiVoD 2 EPiVnotD  3 EPnotV   4 NoEP |    Total
---------------------+--------------------------------------------+----------
1 EPiVoD             | 1,917,337         0    59,515        0 | 1,976,852
2 EPiVnotD           | 6,232,367 4,018,219         0        0 |10,250,586
3 EPnotV             |         0         0 2,400,005        0 | 2,400,005
4 NoEP               |         0         0        0 12,851,412 |12,851,412
---------------------+--------------------------------------------+----------
               Total | 8,149,704 4,018,219 2,459,520 12,851,412 |27,478,855
```

LIAB LM4/PM1, 1993-2007, 19 August 2009, data on individuals

```
         Status of |       Status of establishment number
establishment number | 1 EPiVoD 2 EPiVnotD  3 EPnotV   4 NoEP |    Total
---------------------+--------------------------------------------+----------
1 EPiVoD             | 4,650,093         0   200,177        0 | 4,850,270
2 EPiVnotD           | 3,520,509 5,380,422         0        0 | 8,900,931
3 EPnotV             |         0         0 2,214,719        0 | 2,214,719
4 NoEP               |         0         0        0 14,022,321 |14,022,321
---------------------+--------------------------------------------+----------
               Total | 8,170,602 5,380,422 2,414,896 14,022,321 |29,988,241
```

LIAB LM5/MM1, 1993-2007, 19 August 2009, data on individuals

```
         Status of |       Status of establishment number
establishment number | 1 EPiVoD 2 EPiVnotD  3 EPnotV   4 NoEP |    Total
---------------------+--------------------------------------------+----------
1 EPiVoD             | 7,612,891         0   203,174        0 | 7,816,065
2 EPiVnotD           | 6,142,193 30,278,730        0        0 |36,420,923
3 EPnotV             |         0         0   757,302        0 |   757,302
4 NoEP               |         0         0        0 25,831,434 |25,831,434
---------------------+--------------------------------------------+----------
               Total |13,755,084 30,278,730  960,476 25,831,434 |70,825,724
```

How did the error occur?

Two programming errors were made: Error 1 (value 2 instead of 1) is due to a merge without a key variable. Error 2 (value 1 instead of 3) reflects that all employment records that could be merged to a surveyed IAB Establishment Panel were classified as 1. Correct is only to classify records that merged to one of the firms selected for the specific LIAB version as 1, all other merges belong to status 3.

How to prevent such errors in the future?

We continuously try to improve the quality of our data and to reduce the risk of errors. One important step was taken in 2010 when we wrote down and committed ourselves to a list of measures to assure the quality of the data and documentation we produce. We believe that, by this means, the risk of data errors like the one described in this paper has been further reduced. An unavoidable minimum risk will remain.