Register-based statistics production

Administrative data used for statistical purposes

Bo Sundgren

2010

Part 2: Extra material

Tables with means and ratios for different domains of study

Chart 7.5 Table forming the basis for several different tables

Sex	Level	$\sum w_i$	$\sum w_i \cdot \text{salary}_i$	$\sum w_i \cdot \text{extent}_i$	$\sum w_i \cdot \text{salaryfull}_i$	(6)/(3)	(4)/(5)	(5)/(3)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Wom.	1	4677	40635041	3061	61701517	13 191	13 274	0.65
Wom.	2	33084	360320838	25085	474211547	14 334	14 364	0.76
Wom.	3	7762	106558725	6416	129063365	16 627	16 607	0.83
Wom.	4	10763	175129111	9313	202168595	18 784	18 805	0.87
Wom.	5	817	17415674	771	18333690	22 431	22 583	0.94
Men	1	869	10905658	797	11795151	13 579	13 690	0.92
Men	2	5758	73156416	4898	85083524	14 777	14 936	0.85
Men	3	2717	42580548	2532	45632570	16 796	16 815	0.93
Men	4	5930	125232247	5430	136222043	22 971	23 063	0.92
Men	5	675	18050648	655	18594809	27 531	27 564	0.97

Chart 7.6

Mean salary by sex and level of competences

Level	Women	Men
1	13 274	13 690
2	14 364	14 936
3	16 607	16 815
4	18 805	23 063
5	22 583	27 564

The table above is formed using column (8) in the table in Chart 7.5

Chart 7.7 Calculation of standardised mean salaries

Age	ISCO	Average salary, women	Average salary, men	Number of women	Number of men	Total number	Standard weighting	Women:	Men:
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8) = (7) / 5 688	(3) · (8)	(4) · (8)
17–24	2330	13 660	14 100	276	75	351	0.0062	84.41	87.13
...	...	...	...	...	...	...	...	...	...
60–64	7130	13 826	13 900	10	63	74	0.0013	18.09	18.19
Total		**15 680**	**18 860**	4 523	1 165	5 688	1.0000	**16 256**	**16 505**

Calibration of weights in register-based surveys: compensating for missing values

Chart 7.10 Register on persons from two small regions

(1) PIN	(2) Sex	(3) District	(4) Employed	(5) Industry	(6) Education	(7) d_i	x_{1i} Sex=F	x_{2i} Sex=M	x_{3i} District=1	x_{4i} Employed=1	w_i
1	F	1	0	null	Low	1	1	0	1	1	0.98276
2	M	1	1	A	Low	1	0	1	1	1	1.15517
3	F	1	1	A	Low	1	1	0	1	1	1.13793
4	M	1	1	A	Medium	1	0	1	1	1	1.15517
5	F	1	1	A	Medium	1	1	0	1	1	1.13793
6	M	1	1	***Missing***	Low	0	0	1	1	1	0.00000
7	F	1	1	D	Medium	1	1	0	1	1	1.13793
8	M	1	1	D	High	1	0	1	1	1	1.15517
9	F	1	1	D	Medium	1	1	0	1	1	1.13793
10	M	1	0	null	Medium	1	0	1	1	1	1.00000
11	F	2	0	null	Low	1	1	0	0	0	1.00000
12	M	2	1	D	Low	1	0	1	0	0	1.17241
13	F	2	1	D	Low	1	1	0	0	0	1.15517
14	M	2	1	D	Medium	1	0	1	0	0	1.17241
15	F	2	1	D	***Missing***	0	1	0	0	0	0.00000
16	M	2	1	A	Low	1	0	1	0	0	1.17241
17	F	2	1	A	Medium	1	1	0	0	0	1.15517
18	F	2	1	A	Medium	1	1	0	0	0	1.15517
19	M	2	0	null	Medium	1	0	1	0	0	1.01724

Two ways of handling missing values

Chart 7.11 Persons by Education and Industry

	Industry A Number of persons	Industry D Number of persons	Industry A Per cent	Industry D Per cent
High education	0	1	0.0%	16.7%
Medium education	4	3	57.1%	50.0%
Low education	3	2	42.9%	33.3%
All	7	6	100.0%	100.0%

- Based on the shaded columns in 7.10, giving weight 0 to persons with missing values

Chart 7.12 Persons by Education and Industry, adjusted for missing values

	Industry A, weighted number of persons	Industry D, weighted number of persons	Industry A Per cent	Industry D Per cent
High education	0.0	1.2	0.0%	16.7%
Medium education	4.6	3.4	57.0%	49.7%
Low education	3.5	2.3	43.0%	33.6%
All	8.1	6.9	100.0%	100.0%

- Based on adjusted weights after calibration

Estimation methods to correct for overcoverage: the problem

Overcoverage in the Population Register

The first sign that there is overcoverage in Statistics Sweden's Population Register came from demographic studies on mortality. Among a few categories of foreign-born persons, mortality was strangely low. Furthermore, it was found that the share of families, with no information on disposable income, was high among certain categories of immigrants.

Overcoverage in the Swedish Population Register has been estimated by Greijer (1995, 1996, 1997a, 1997b), who analysed nonresponse in the Labour Force Surveys and in a census on foreign-born persons based on a postal questionnaire. Using this information, it was possible to estimate overcoverage among different categories of foreign-born persons.

Data in statistical registers can also be used to give indications of overcoverage. A foreign-born person without income in any register can have left Sweden without reporting this to the tax authorities.

Overcoverage can cause serious errors in register-based statistics. For instance, the average income for those born in different countries can be misleading. For persons born in certain countries, the underestimation can be around 20%.

Estimation methods to correct for overcoverage: strategy

How should we control overcoverage and improve quality? The strategy for correcting errors caused by overcoverage can include the following:

1. By being watchful when carrying out macro editing, it is possible to find unreasonable estimates in the register-based statistics. The question should be asked, whether overcoverage could be the cause of these extreme estimates.
2. If overcoverage is suspected, available sample surveys and other sources can be used to help estimate this overcoverage.
3. Overcoverage can be estimated for different categories in the register once enough information on the extent and character of the overcoverage has been collected.
4. The weights can then be adjusted to correct for the estimated overcoverage. Before adjustment, all weights are equal to 1; after adjustment, the weights for the different categories for which there is overcoverage will be less than 1. Use *calibration methods* (Section 7.5) to adjust the weights when overcoverage is described by many variables.
5. The adjusted weights are stored in the base register (in this case the Population Register).
6. All other statistical products using the base register will then use the weights. In this way, all the statistics produced will be consistently corrected for the estimated effect of overcoverage.

Estimation methods to correct for overcoverage: example of solution

Chart 8.17 Estimated overcoverage for different categories

		Number of persons before correction (1)	Estimated overcoverage (2)	Number of persons after correction for overcoverage (3)
Country of birth	Europe	*584*	6.7%	**545**
	Not Europe	*416*	14.7%	**355**
	Total	1000	10.0%	900
Years in Sweden	Few	*819*	7.2%	**760**
	Many	181	22.7%	140
	Total	1000	10.0%	900
Income	Low	101	40.6%	60
	High	*899*	6.6%	**840**
	Total	1000	10.0%	900

Chart 8.18 Register with calibrated weights

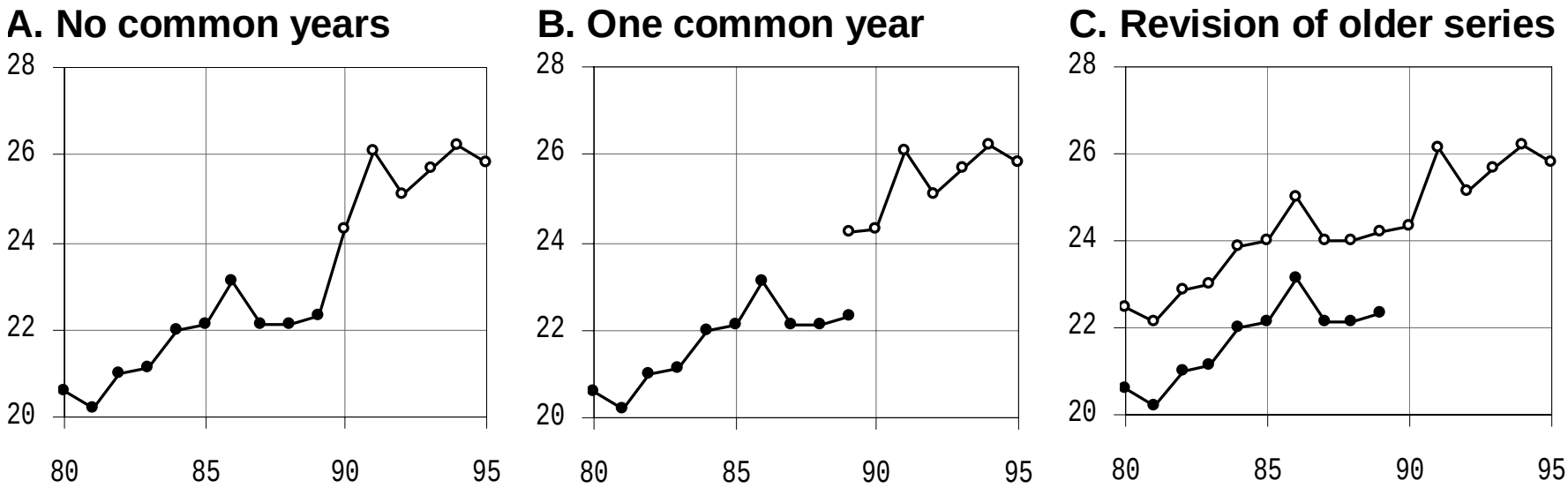
PIN	Country of birth	Years in Sweden	Income	Weights d_k	x_{1i} Country Europe	x_{2i} Country Not Eur	x_{3i} Few years	x_{4i} High income	Adjusted weight, w_i
1	Europe	Few	High	1	1	0	1	1	0.992
2	Not Eur	Few	High	1	0	1	1	1	0.916
3	Europe	Few	Low	1	1	0	1	0	0.657
4	Europe	Few	High	1	1	0	1	1	0.992
5	Not Eur	Many	High	1	0	1	0	1	0.770
6	Not Eur	Few	Low	1	0	1	1	0	0.581
...	...	...	...	...	...	...	...	...	...
1000	Not Eur	Few	High	1	0	1	1	1	0.916
Total				**1 000**	**584**	**416**	**819**	**899**	**900**

Condition

t_x	$\sum_k d_k x_{ki}$
545	*584*
355	*416*
760	*819*
840	*899*

Methods to correct for level shifts in time series

Chart 8.19 Three ways of reporting level shifts in time series



Methods to correct for level shifts in time series

- Correcting for level shifts in time series is called linking time series. We differentiate between linking at macro level and linking at micro level.
- Linking at macro level has the disadvantage that it can be difficult to ensure consistency among several linked series, and there may be many of them.
- Linking at micro level does not have this disadvantage, since all possible time series are linked at the same time, and all these series will be consistent. Linking at micro level is carried out by calculating revised weights in data matrices for the years or periods that have been revised. The old time series values are given by the original weights, while the revised weights can be used to calculate the linked series.

Estimation with combination objects: calendar year registers

art 9.2 Calendar year register for the population of persons during 2005

rson	Address	Municipality	From date yyyymmdd	To date yyyymmdd	Weight = Time at the address, years
IN1	Address 1	1	20050101	20050517	136/365 = 0.37
IN1	Address 2	2	20050518	20051231	229/365 = 0.63
IN2	Address 3	1	20050101	20051231	365/365 = 1.00
IN3	Address 4	2	20050101	20050314	73/365 = 0.20
IN3	Address 5	1	20050315	20050925	194/365 = 0.53
IN3	Address 6	2	20050926	20051231	365/365 = 0.27
IN4	Address 7	2	20050101	20050605	156/365 = 0.43
IN4	Address 8	1	20050606	20051231	209/365 = 0.57

art 9.3 Average population 2005

nicipality	Absolute frequency	Relative frequency
1	0.37 + 1.00 + 0.53 + 0.57 = 2.47	62%
2	0.63 + 0.20 + 0.27 + 0.43 = 1.53	38%
Total	4.00	100%

Estimation with combination objects: occupation (1)

Chart 9.7 Job Register with occupational data

Job id	Person	Legal unit	Occupation	ISCO	Extent
J1	PIN1	LeU1	Statistician	2211	100
J2	PIN1	LeU2	Farmer	6111	15
J3	PIN1	LeU3	Politician	1110	10
J4	PIN2	LeU4	Hospital orderly	5132	30
J5	PIN2	LeU5	Cleaner	9122	20
J6	PIN3	LeU6	Shop assistant	5221	10
J7	PIN4	LeU6	Shop assistant	5221	50
J8	PIN5	LeU6	Shop assistant	5221	20
J9	PIN6	LeU6	Shop assistant	5221	100
Total					

Chart 9.8 Traditional register on persons with occupational information

Person	Legal unit	Principal occupation	ISCO	Extent	Weight alt 1
PIN1	LeU1	Statistician	2211	100	1
PIN2	LeU4	Hospital orderly	5132	30	1
PIN3	LeU6	Shop assistant	5221	10	1
PIN4	LeU6	Shop assistant	5221	50	1
PIN5	LeU6	Shop assistant	5221	20	1
PIN6	LeU6	Shop assistant	5221	100	1
Total					**6**

Chart 9.9 Employed persons by occupation, traditional alternative 1

Main occupation	ISCO	Number	Per cent
Statistician	2211	1	16.7
Hospital orderly	5132	1	16.7
Shop assistant	5221	4	66.7
Total		**6**	**100.0**

Estimation with combination objects: occupation (2)

Chart 9.10 Register on persons with occupational data

Combination object	Person	Occupation	Extent	Weight alternative 1	Weight alternative 2
1	PIN1	Statistician	100	1	0.80
2	PIN1	Farmer	15	0	0.12
3	PIN1	Politician	10	0	0.08
4	PIN2	Hospital orderly	30	1	0.60
5	PIN2	Cleaner	20	0	0.40
6	PIN3	Shop assistant	10	1	1.00
7	PIN4	Shop assistant	50	1	1.00
8	PIN5	Shop assistant	20	1	1.00
9	PIN6	Shop assistant	100	1	1.00
Total				6	6.00

hart 9.11 Employed persons by occupation according to two alternatives

		Alternative 1		Alternative 2	
Occupation	ISCO	Nr.	Per cent	Nr.	Per cent
Politician	1110	0.00	0.0	0.08	1.3
Statistician	2211	1.00	16.7	0.80	13.3
Hospital orderly	5132	1.00	16.7	0.60	10.0
Shop assistant	5221	4.00	66.7	4.00	66.7
Farmer	6111	0.00	0.0	0.12	2.0
Cleaner	9122	0.00	0.0	0.40	6.7
Total		6.00	100.0	6.00	100.0

Estimation with combination objects: occupation (3)

Chart 9.12 Register on persons with occupational data

Person	Occupation	Weight alt 1	Weight alt 2	Weight alt 3
PIN1	Statistician	1	0.80	1.00
PIN1	Farmer	0	0.12	0.15
PIN1	Politician	0	0.08	0.10
PIN2	Hospital orderly	1	0.60	0.30
PIN2	Cleaner	0	0.40	0.20
PIN3	Shop assistant	1	1.00	0.10
PIN4	Shop assistant	1	1.00	0.50
PIN5	Shop assistant	1	1.00	0.20
PIN6	Shop assistant	1	1.00	1.00
Total		**6**	**6.00**	**3.55**

Chart 9.13 Persons and full-time employed by occupation, three alternatives

		Alternative 1		Alternative 2		Alternative 3	
Occupation	ISCO	Nr.	Per cent	Nr.	Per cent	Nr.	Per cent
Politician	1110	0.00	0.0	0.08	1.3	0.10	2.8
Statistician	2211	1.00	16.7	0.80	13.3	1.00	28.2
Hospital orderly	5132	1.00	16.7	0.60	10.0	0.30	8.5
Shop assistant	5221	4.00	66.7	4.00	66.7	1.80	50.7
Farmer	6111	0.00	0.0	0.12	2.0	0.15	4.2
Cleaner	9122	0.00	0.0	0.40	6.7	0.20	5.6
Total		**6.00**	**100.0**	**6.00**	**100.0**	**3.55**	**100.0**

Estimation with combination objects: industrial classification (1)

Chart 9.14a Business Register year 1: Data matrix for local units

Local unit	Industry 1	%	Industry 2	%	Industry 3	%	Nr. of employees
LU1	DJ	100					218
LU2	DH	51	DJ	49			293
LU3	DJ	40	DH	30	DK	30	156

Chart 9.14b Business Register year 2: Data matrix for local units

Local unit	Industry 1	%	Industry 2	%	Industry 3	%	Nr. of employees
LU1	DJ	100					221
LU2	DJ	52	DH	48			314
LU3	DJ	36	DH	34	DK	30	143

Chart 9.14c Number of employees by industry, traditional estimates

Industry	Year 1	Year 2
DH	293	0
DJ	374	678
DK	0	0
Total	**667**	**678**

The number of employees is sorted by principal industry, which is the most common way of presenting time series based on industrial classification from the Business Register.

This leads to abrupt changes in the series here.

Estimation with combination objects: industrial classification (2)

Chart 9.15 Data matrix with combination objects: Local unit · Industry

Year 1				
Local unit	Industry	Weight, w_i	Nr. empl, yi	$w_i\, y_i$
LU1	DJ	1.00	218	218
LU2	DH	0.51	293	149.43
LU2	DJ	0.49	293	143.57
LU3	DJ	0.40	156	62.4
LU3	DH	0.30	156	46.8
LU3	DK	0.30	156	46.8
Total		**3.00**		**667**

Year 2				
Local unit	Industry	Weight, w_i	Nr. empl, yi	$w_i\, y_i$
LU1	DJ	1.00	221	221
LU2	DH	0.48	314	150.72
LU2	DJ	0.52	314	163.28
LU3	DJ	0.36	143	51.48
LU3	DH	0.34	143	48.62
LU3	DK	0.30	143	42.9
Total		**3.00**		**678**

Chart 9.16 Number of employees by industry, estimated with combination objects

Industry	Year 1	Year 2
DH	196.23	199.34
DJ	423.97	435.76
DK	46.80	42.90
Total	**667.00**	**678.00**

The time series in Chart 9.16 have been calculated with the weights w_i.

The series here have a higher quality than those in Chart 9.14c, with relevant changes and no level shifts.

Estimation with combination objects: transformation of weights

Chart 9.17 Transformation of weights

Register of local units			Aggregated data		Register of local units		
Year 1		Weights	Models for different industries		Transformed weights based on model		
Local unit	Industry	based on turnover	Employees / Turnover SEK m		adapted for estimation of number of employees		
LU3	DJ	0.4	DJ	0.5	LU3	DJ	$\frac{0.4 \cdot 0.5}{(0.4 \cdot 0.5 + 0.3 \cdot 0.6 + 0.3 \cdot 0.7)} = \mathbf{0.34}$
LU3	DH	0.3	DH	0.6	LU3	DH	$\frac{0.3 \cdot 0.6}{(0.4 \cdot 0.5 + 0.3 \cdot 0.6 + 0.3 \cdot 0.7)} = \mathbf{0.30}$
LU3	DK	0.3	DK	0.7	LU3	DK	$\frac{0.3 \cdot 0.7}{(0.4 \cdot 0.5 + 0.3 \cdot 0.6 + 0.3 \cdot 0.7)} = \mathbf{0.36}$

Estimation with combination objects: importing many multi-valued variables: traditional methodology

Chart 9.18a Population Register

Person	Sex	Age
PIN10	F	32

Chart 9.18b Education Register

Person	Educ 1	Points 1	Educ 2	Points 2
PIN10	Ed1	180	Ed2	120

PIN10 has two degrees on the same level in different fields; Educ 2 is the most recent.

Chart 9.18c Activity Register, with extent of job in November

Person	Local unit	Extent
PIN10	LU11	80%
PIN10	LU12	20%

Traditionally, only the local unit of the principal activity is used.

Chart 9.18d Occupation Register

Person	Local unit	Occup.
PIN10	LU11	Oc1
PIN10	LU12	Oc2

Traditionally, only the occupation of the principal activity is used.

Chart 9.18e Business Register

Local unit	Industry	Weight 1	Industry	Weight 2
LU11	DH	70%	DJ	30%
LU12	DK	100%		

Chart 9.18f Activity Register, industry and occupation are imported

Person	Local unit	Industry	Occup.
PIN10	LU11	DH	Oc1
PIN10	LU12	DK	Oc2

Traditionally, only the local unit and occupation of the principal activity are used.

Chart 9.18g Employment Register, data for person PIN10

Person	Sex	Age	Education	Occupation	Local unit	Industry	EmpNov
PIN10	F	32	Ed2	Oc1	LU11	DH	Yes

Estimation with combination objects: importing many multi-valued variables: all information used

Chart 9.19a Population Register

Person	Sex	Age
PIN10	F	32

Chart 9.19b Education Register

Person	Education	w_{Edu}
PIN10	Ed1	0.6
PIN10	Ed2	0.4

Weights for education are created using the length of the educational programme expressed as education 'points'.

Chart 9.19c Activity Register, with extent of job in November

Person	Local unit	w_{LU}
PIN10	LU11	0.8
PIN10	LU12	0.2

For the object *Person*, weights for the multi-valued variable *Local unit* are created using the variable *Extent*

Chart 9.19d Occupation Register

Person	Local unit	Occup.
PIN10	LU11	Oc1
PIN10	LU12	Oc2

Occupation is linked to the relation between *Person* and *Local unit*, the weight for *Occupation* is the same as that for *Local unit*.

Chart 9.19e Business Register

Local unit	Industry	w_{Ind}
LU11	DH	0.7
LU11	DJ	0.3
LU12	DK	1.0

Using information in the Business Register, a register is created with the combination object *Local unit* · *Industry* and the weights for different industries.

Chart 9.19f Combination objects: Person · Education · Local unit · Industry

Person	Sex	Age	Educ.	Local unit	Occup	Industry	EmpNov	w_{Edu}	w_{LU}	w_{Ind}	$w_{CombObj}$
PIN10	F	32	Ed1	LU11	Oc1	DH	Yes	0.6	0.8	0.7	0.336
PIN10	F	32	Ed1	LU11	Oc1	DJ	Yes	0.6	0.8	0.3	0.144
PIN10	F	32	Ed1	LU12	Oc2	DK	Yes	0.6	0.2	1.0	0.120
PIN10	F	32	Ed2	LU11	Oc1	DH	Yes	0.4	0.8	0.7	0.224
PIN10	F	32	Ed2	LU11	Oc1	DJ	Yes	0.4	0.8	0.3	0.096
PIN10	F	32	Ed2	LU12	Oc2	DK	Yes	0.4	0.2	1.0	0.080
Total											1.000

Estimation with combination objects: importing many multi-valued variables: comparison of methods (1)

Chart 9.20a Number of gainfully employed persons in November, by age and sex

Traditional estimation

Age	F	M	Total
20–49	1	0	1
50–64	0	0	0
65–	0	0	0
Total	1	0	

The estimates for single-valued variables such as sex and age are not affected by the weights that are formed for the multi-valued variables.

Estimation with weights

Age	F	M	Total
20–49	1	0	1
50–64	0	0	0
65–	0	0	0
Total	1	0	1

Chart 9.20b Number of gainfully employed in November by occupation

Traditional estimation

Occup.	Number
Oc1	1
Oc2	0
Total	1

Estimation with weights:
0.8 =
0.336 + 0.144 + 0.224 +0.096

Estimation with weights

Occup.	Number
Oc1	0.8
Oc2	0.2
Total	1

Chart 9.20c Number of gainfully employed in November by education

Traditional estimation

Educ.	Number
Ed1	0
Ed2	1
Total	1

Estimation with weights:
0.6 = 0.336 + 0.144 +0.120

Estimation with weights

Educ.	Number
Ed1	0.6
Ed2	0.4
Total	1

Chart 9.20d Gainfully employed in November by industrial classification

Traditional estimation

Industry	Number
DH	1
DJ	0
DK	0
Total	1

Estimation with weights:
0.56 = 0.336 + 0.224

Estimation with weights

Industry	Number
DH	0.56
DJ	0.24
DK	0.20
Total	1

Estimation with combination objects: importing many multi-valued variables: comparison of methods (2)

Chart 9.20e Gainfully employed in November by occupation and education

Traditional estimation

Educ.	Oc1	Oc2	Total
Ed1	0	0	0
Ed2	1	0	1
Total	1	0	1

Estimation with weights:
0.48 = 0.336 + 0.144

Estimation with weights

Educ.	Oc1	Oc2	Total
Ed1	0.48	0.12	0.60
Ed2	0.32	0.08	0.40
Total	0.80	0.20	1

Chart 9.20f Gainfully employed in November by education and industry

Traditional estimation

Industry	Ed1	Ed2	Total
DH	0	1	1
DJ	0	0	0
DK	0	0	0
Total	0	1	1

Estimation with weights:
0.336 is taken directly from the data matrix in Chart 9.19f

Estimation with weights

Industry	Ed1	Ed2	Total
DH	0.336	0.224	0.560
DJ	0.144	0.096	0.240
DK	0.120	0.080	0.200
Total	0.600	0.400	1

Chart 9.20g Gainfully employed in November by occupation and industry

Traditional estimation

Industry	Oc1	Oc2	Total
DH	1	0	1
DJ	0	0	0
DK	0	0	0
Total	1	0	1

Estimation with weights
0.56 = 0.336 + 0.224

Estimation with weights

Industry	Oc1	Oc2	Total
DH	0.56	0.00	0.56
DJ	0.24	0.00	0.24
DK	0.00	0.20	0.20
Total	0.80	0.20	1

Estimation with combination objects: consistency between estimates from different registers (1)

Chart 9.21 Available information in four registers before integration

1. Register on persons

Person
PIN1
PIN2
PIN3
PIN4
PIN5

2. Job Register

Job	Enterprise	Local unit	Person	Ext_J
J1	EU1	LU11	PIN1	0.3
J2	EU1	LU11	PIN2	1.0
J3	EU2	LU21	PIN3	1.0
J4	EU2	LU21	PIN4	1.0
J5	EU2	LU22	PIN5	1.0
J6	EU2	LU22	PIN1	0.2

Person PIN1 has two jobs, J1 and J6, PIN1 works 50% of a full-time employed position.

The information of the proportions of Industries within each Local Unit will be imported into registers 1 and 2 when combination objects are created and the aggregated weight will be called w_{Ind} in Chart 9.23 below.

3. Local unit Register

Local unit	Enterprise	Industry 1	%	Industry2	%
LU11	EU1	A	60	B	40
LU21	EU2	C	100		
LU22	EU2	D	100		

4. Enterprise Register

Enterprise	Local unit 1	Local unit 2	Industry1	%	Industry2	%
EU1	LU11		A	60	B	40
EU1	LU21	LU22	C	62.5	D	37.5

Estimation with combination objects: consistency between estimates from different registers (2)

Chart 9.22 Traditional estimation in a register system after integration

1. Employment Register – persons

Person	Industry	Ext_P	weight
PIN1	A	0.5	1
PIN2	A	1.0	1
PIN3	C	1.0	1
PIN4	C	1.0	1
PIN5	D	1.0	1
Total		**4.5**	**5**

2. Job Register

Job	Enterprise	Local unit	Person	Ext_J	Industry	weight
J1	EU1	LU11	PIN1	0.3	A	1
J2	EU1	LU11	PIN2	1.0	A	1
J3	EU2	LU21	PIN3	1.0	C	1
J4	EU2	LU21	PIN4	1.0	C	1
J5	EU2	LU22	PIN5	1.0	D	1
J6	EU2	LU22	PIN1	0.2	D	1
Total				**4.5**		**6**

In all the registers in Chart 9.22 the variable Industry refers to the *principal* Industry. Every register also has only one row per object (person, job, local unit or enterprise unit).

The variable Ext_J, *Extent for Job*, in the Job Register is imported into all other registers. Derived variables Ext_P, *Extent for Person*, Ext_{LU}, *Extent for Local Unit and* Ext_{EU}, *Extent for Enterprise Unit*, are created by aggregation.

With these four registers, volume of work by Industry, can be estimated by summing up the variables:

Ext_P , Ext_J , Ext_{LU} *and* Ext_{EU}

Due to different aggregation errors, the estimated tables with volume of work by Industry will be different if different registers are used.

These different tables are compared in Chart 9.24 below.

3. Local unit Register

Local unit	Industry	Ext_{LU}	weight
LU11	A	1.3	1
LU21	C	2.0	1
LU22	D	1.2	1
Total		**4.5**	**3**

4. Enterprise Register

Enterprise	Local unit	Industry	Ext_{EU}	weight
EU1	LU11	A	1.3	1
EU2	LU21	C	3.2	1
Total			**4.5**	**2**

Estimation with combination objects: consistency between estimates from different registers (3)

Chart 9.23 Consistent estimates with weights in a register system

1. Employment Register – persons

Person	Industry	Ext_p	w_{ind}	w_{Job}	$w_p = w_{ind} \cdot w_{Job}$	$Ext_p \cdot w_p$
PIN1	A	0.50	0.6	0.6	0.36	0.18
PIN1	B	0.50	0.4	0.6	0.24	0.12
PIN1	D	0.50	1	0.4	0.4	0.20
PIN2	A	1.00	0.6	1	0.6	0.60
PIN2	B	1.00	0.4	1	0.4	0.40
PIN3	C	1.00	1	1	1	1.00
PIN4	C	1.00	1	1	1	1.00
PIN5	D	1.00	1	1	1	1.00
Total					5	4.5

2. Job Register

Job	Enterprise	Local unit	Person	Ext_J	Industry	w_{ind}	$Ext_J \cdot w_{ind}$
J1	EU1	LU11	**PIN1**	0.30	A	0.6	0.18
J1	EU1	LU11	**PIN1**	0.30	B	0.4	0.12
J2	EU1	LU11	PIN2	1.00	A	0.6	0.60
J2	EU1	LU11	PIN2	1.00	B	0.4	0.40
J3	EU2	LU21	PIN3	1.00	C	1	1.00
J4	EU2	LU21	PIN4	1.00	C	1	1.00
J5	EU2	LU22	PIN5	1.00	D	1	1.00
J6	EU2	LU22	**PIN1**	0.20	D	1	0.20
Total						6	4.5

3. Local unit Register

Local unit	Industry	Ext_{LU}	w_{ind}	$Ext_{LU} \cdot w_{ind}$
LU11	A	1.3	0.6	0.78
LU11	B	1.3	0.4	0.52
LU21	C	2.0	1	2.00
LU22	D	1.2	1	1.20
Total			3	4.5

4. Enterprise Register

Enterprise	Local unit	Industry	Ext_{EU}	w_{ind}	$Ext_{EU} \cdot w_{ind}$
EU1	LU11	A	1.3	0.6	0.78
EU1	LU11	B	1.3	0.4	0.52
EU2	LU21	C	3.2	0.625	2.00
EU2	LU22	D	3.2	0.375	1.20
Total				2	4.5

Comments on the Job Register:

The register contains data on six jobs corresponding to 4.5 full-time jobs.

Jobs J1 and J2 are divided into two combination objects each, as LU11 is active in both Industry A and B. The weights 0.6 and 0.4 for these two industries are taken from the *Local unit Register*. Ext_J refers to the extent of the work for each job.

Comments on the Employment Register:

If $Ext_J \cdot w_{ind}$ is summed up for person PIN1 in the *Job Register*, the result obtained is 0.18 + 0.12 + 0.20 = 0.50. This value becomes Ext_p for PIN1 in the *Employment Register*.

Three combination objects for three industries are formed for PIN1. Both industry and job/local unit are multi-valued variables for persons.

The weights for combination objects are formed by multiplying w_{ind} with w_{Job}, where w_{ind} is taken from the *Job Register* and w_{Job} is calculated as every job's share of all jobs that the person has. For PIN1, job J1 has the weight (.18+.12)/(.18+.12+.20) = 0.6

Estimation with combination objects: consistency between estimates from different registers (4)

Chart 9.24 Comparison of different types of estimation

Table 1
Number of full-time employees by Industry

Industry	Traditional estimation, Chart 9.22				Estimation with combination objects
	Employment Register	Job Register	Local unit Register	Enterprise Register	Register 1–4 in Chart 9.23
A	1.5	1.3	1.3	1.3	0.78
B	0.0	0.0	0.0	0.0	0.52
C	2.0	2.0	2.0	3.2	2.00
D	1.0	1.2	1.2	0.0	1.20
Tot	**4.5**	**4.5**	**4.5**	**4.5**	**4.5**

Table 2
Number of persons by Industry

Traditional, Chart 9.22	With combination objects
Employment Register	Employment Register in Chart 9.23
2	0.96
0	0.64
2	2.00
1	1.40
5	**5.00**

The traditional estimation method thus results in inconsistencies in the register-based statistics produced. Furthermore, differences in population and variable definitions should lead to further inconsistencies in real registers.

There is a fourth reason for why register-based statistics from different products can be inconsistent. The table above shows the effects of content-related differences. Table 1 describes *full-time employees*, while Table 2 describes *persons*. With statistics on persons and labour market statistics, it is common to describe persons but, with economic statistics, it is more common to measure volumes and full-time employees.

Estimation with combination objects: linking of time series

Chart 9.31 Translation of old codes to new ones

Relationship between old and new codes	Old code	Code key	New code	Comments
One-to-one	1		A	No problems, the old code 1 is recoded to the new code A
Many-to-one	2 3		B	No problems, the old codes 2 and 3 are combined to the new code B
One-to-many	4		C D	Causes problems, how should the old code 4 be divided into the new codes C and D?
Many-to-many	5 6		E F	Same problem as mentioned above, how should: – old code 5 be divided up into E and F? – old code 6 be divided up into E and F?

One condition for the method we describe in this section is that there exists a register for a particular point in time or period in which every object is classified according to both the new and the old classifications.

Chart 9.14a Business Register year 1: Data matrix for local units

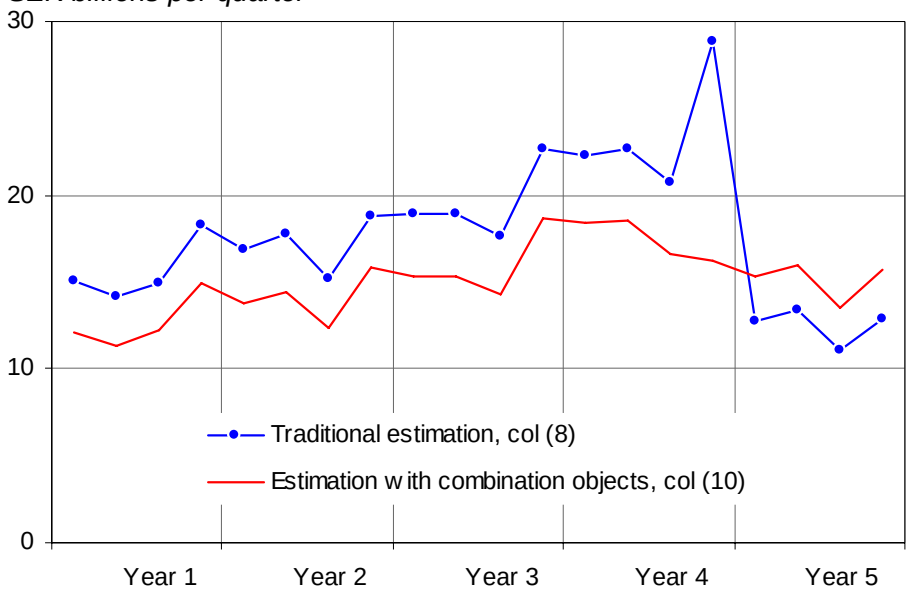
Local unit	Industry 1	%	Industry 2	%	Industry 3	%	Nr. of employees
LU1	DJ	100					218
LU2	DH	51	DJ	49			293
LU3	DJ	40	DH	30	DK	30	156

Chart 9.14b Business Register year 2: Data matrix for local units

Local unit	Industry 1	%	Industry 2	%	Industry 3	%	Nr. of employees
LU1	DJ	100					221
LU2	DJ	52	DH	48			314
LU3	DJ	36	DH	34	DK	30	143

Chart 9.26 Turnover in an industry, two estimates

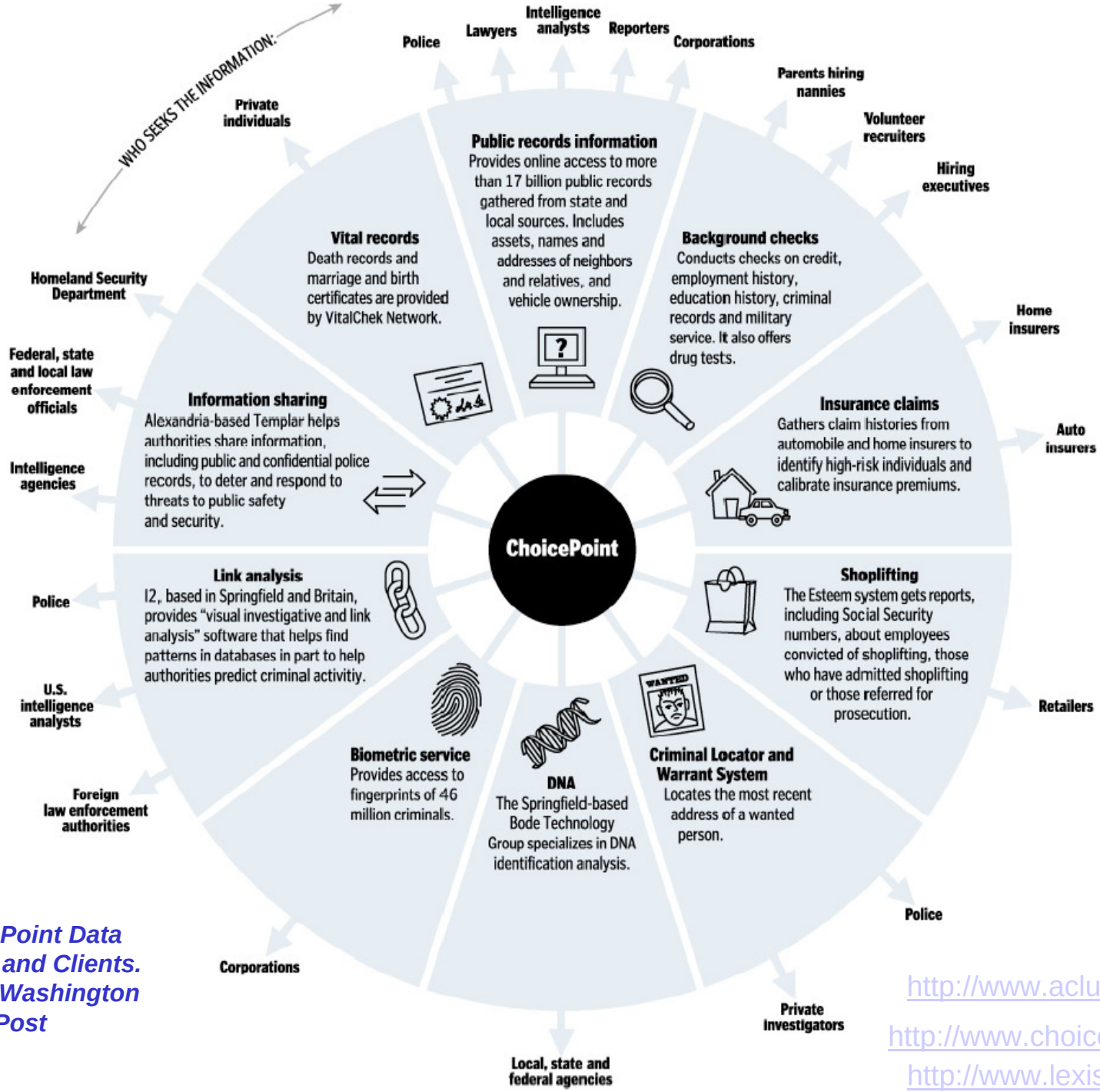
SEK billions per quarter



Registers...

- *… US version: ”war against terrorism”*

http://www.aclu.org/pizza/
http://www.choicepoint.com/
http://www.lexisnexis.com/



ChoicePoint Data Sources and Clients. Source: Washington Post

http://www.aclu.org/pizza/

http://www.choicepoint.com/

http://www.lexisnexis.com/