Anders & Britt Wallgren
Statistics Sweden and Örebro University
ba.statistik@telia.com

Valmiera Workshop 27 August 2012

# Administrative Registers, Survey System Design and Quality Assessment

---

Survey Design aims at:

↓ Costs  Quality ↑

Dominating approach: One survey at a time

The production system (PS) at the National Statistical Institute (NSI)

| Traditional PS | Transition → | Register-based PS |
|---|---|---|
| ***Frames:*** | | ***Frames:*** |
| Maps, address lists | | Base registers |
| ***Data from:*** | | ***Data from:*** |
| Enumerators | → | Admin. registers |
| Interviewers | | The Register System |
| | | Mail questionnaires |
| | | Telephone interviews |

↓ Costs Quality **?**

## Six principles on how to use administrative registers for statistics

*Precondition 1:*
Unified systems of identity codes are used in all administrative systems. The same identity number should follow an object over its life-time.

*Precondition 2:*
A statistical office should have access to administrative registers kept by public authorities. This right should be supported by law as the protection of privacy.

*Methodology principle 1:*
These administrative registers should be transformed into statistical registers. Many sources should be used and compared during this transformation.

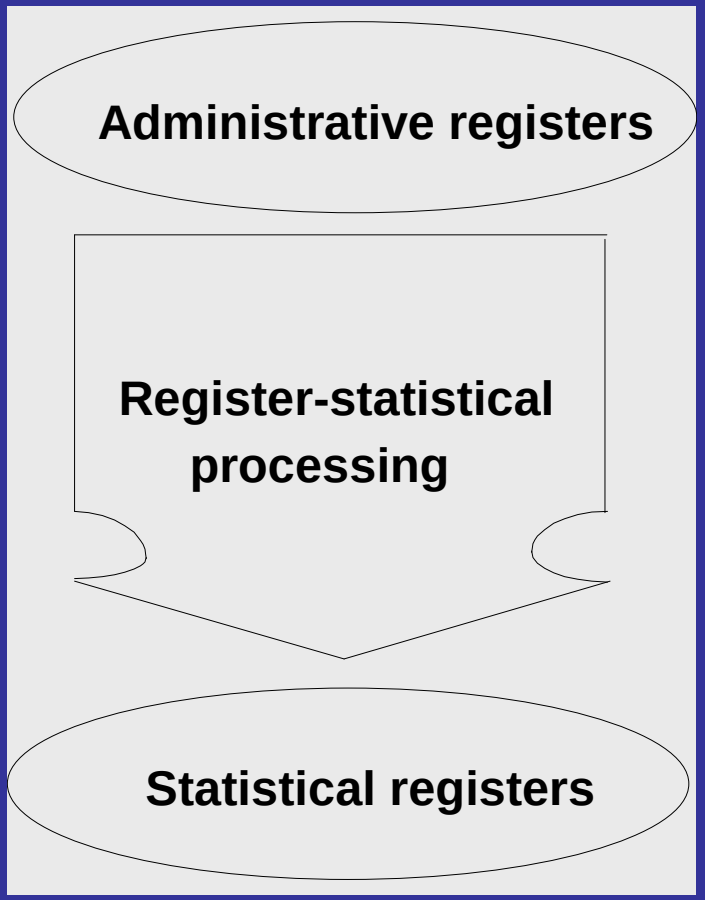
*Methodology principle 2:*
All statistical registers should be included in a coordinated register system. This system will ensure that all data can be integrated and used effectively.

*Methodology principle 3:*
Consistency regarding populations and variables are necessary for the coherence of estimates from different register surveys.

*Methodology principle 4:*
The register system should be used for quality assessment of statistical surveys based on microdata comparisons with other surveys in the production system.

*Methodology principle 1:*
These administrative registers should be transformed into statistical registers.
Many sources should be used and compared during this transformation.

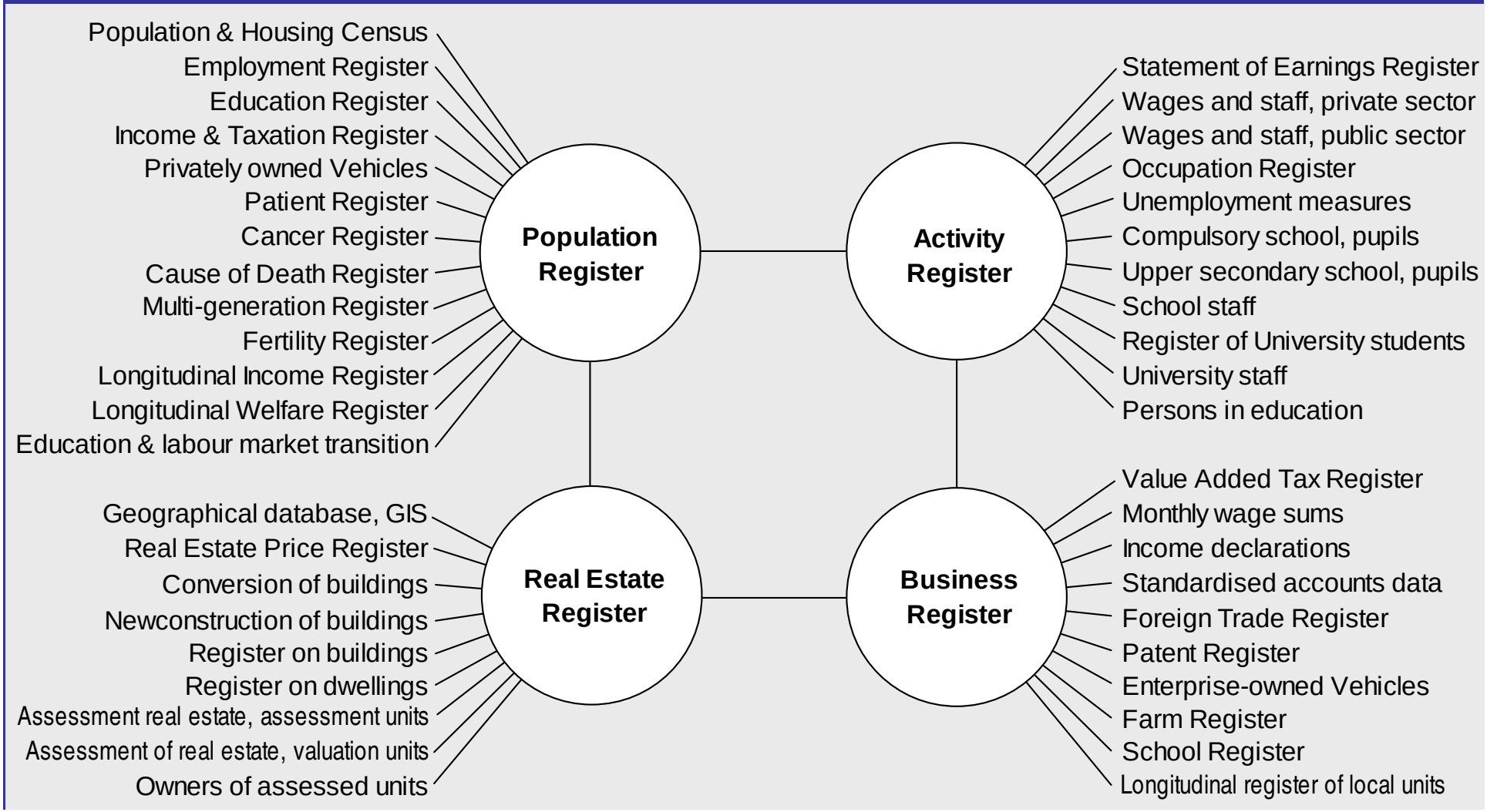
*Methodology principle 2:*
All statistical registers should be included in a coordinated register system.
This system will ensure that all data can be integrated and used effectively.



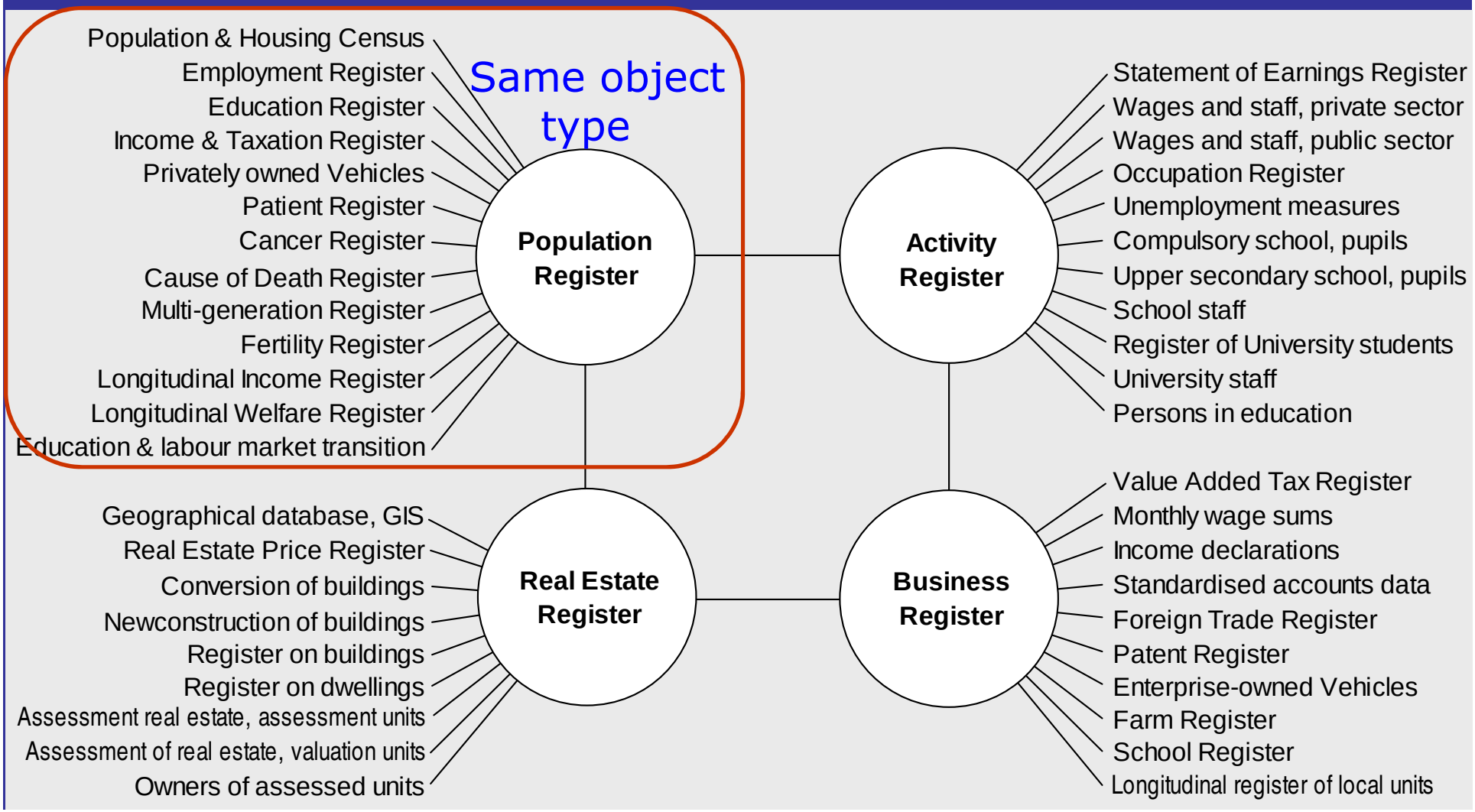

Administrative object set
Administrative units
Administrative variables

Statistical population
Statistical units
Statistical variables

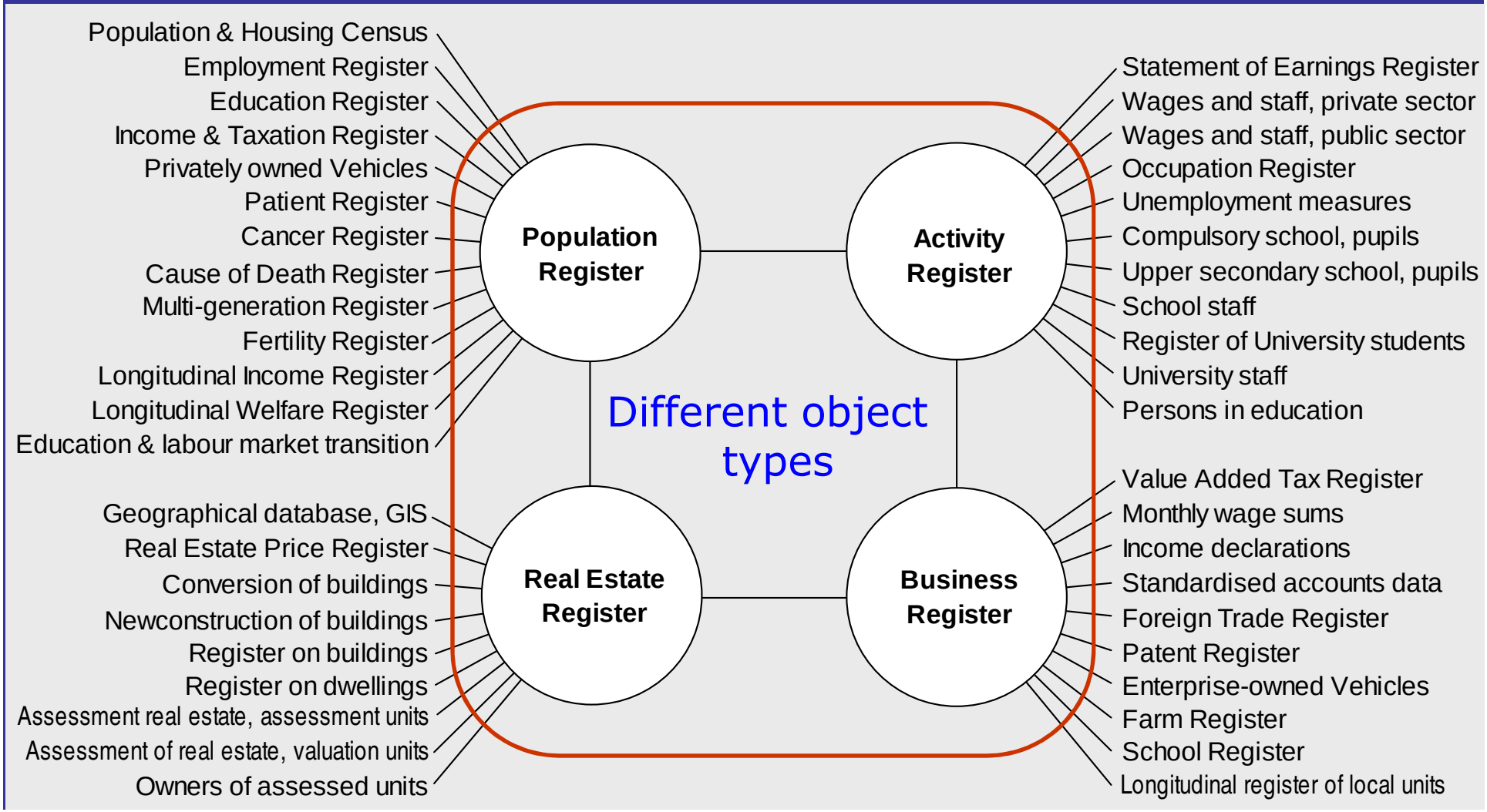
# The production systems in the Nordic countries are completely register-based. All sample surveys, censuses, register surveys are based on the Register System:



# The production systems in the Nordic countries are completely register-based. All sample surveys, censuses, register surveys are based on the Register System:

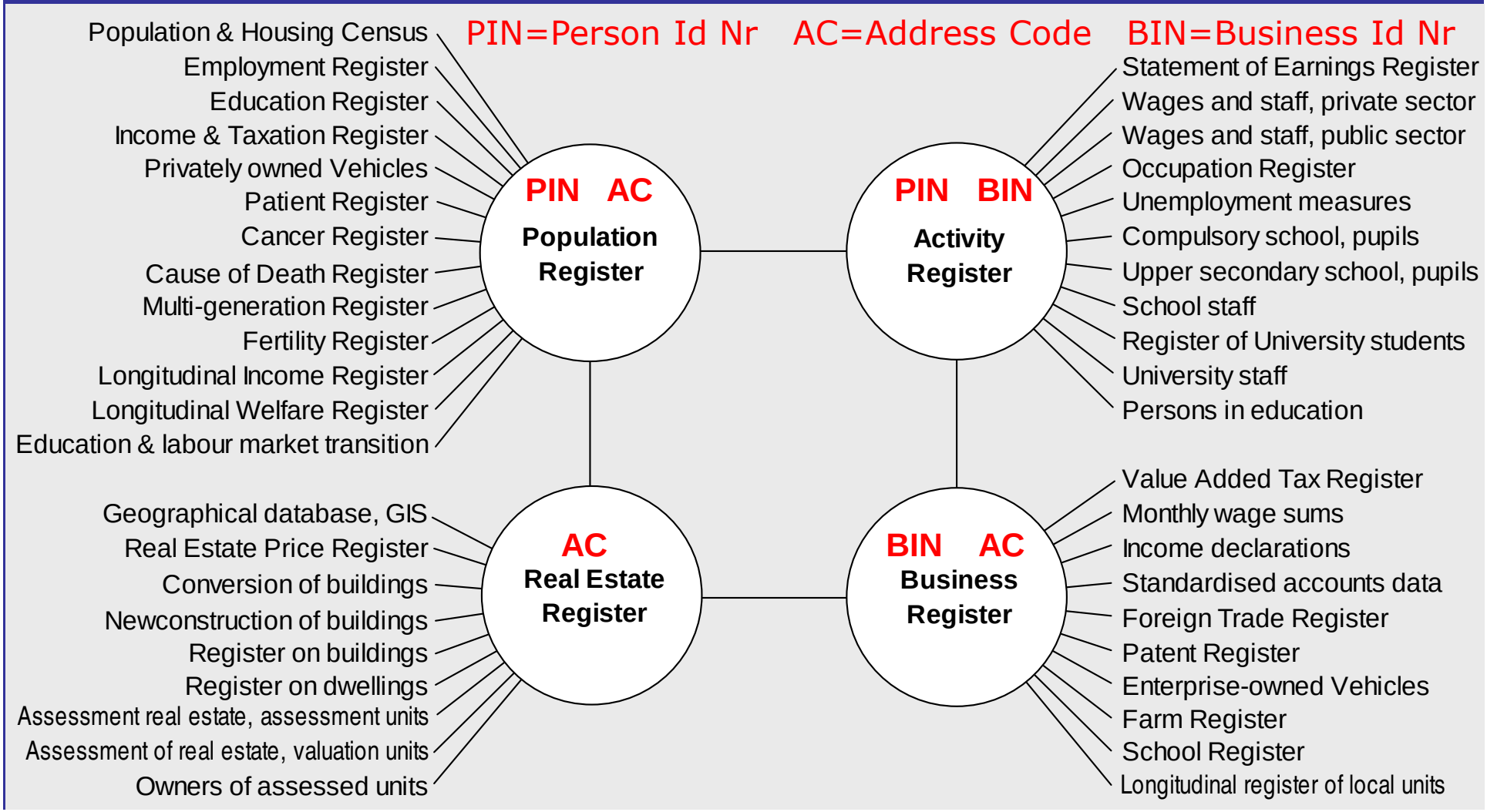


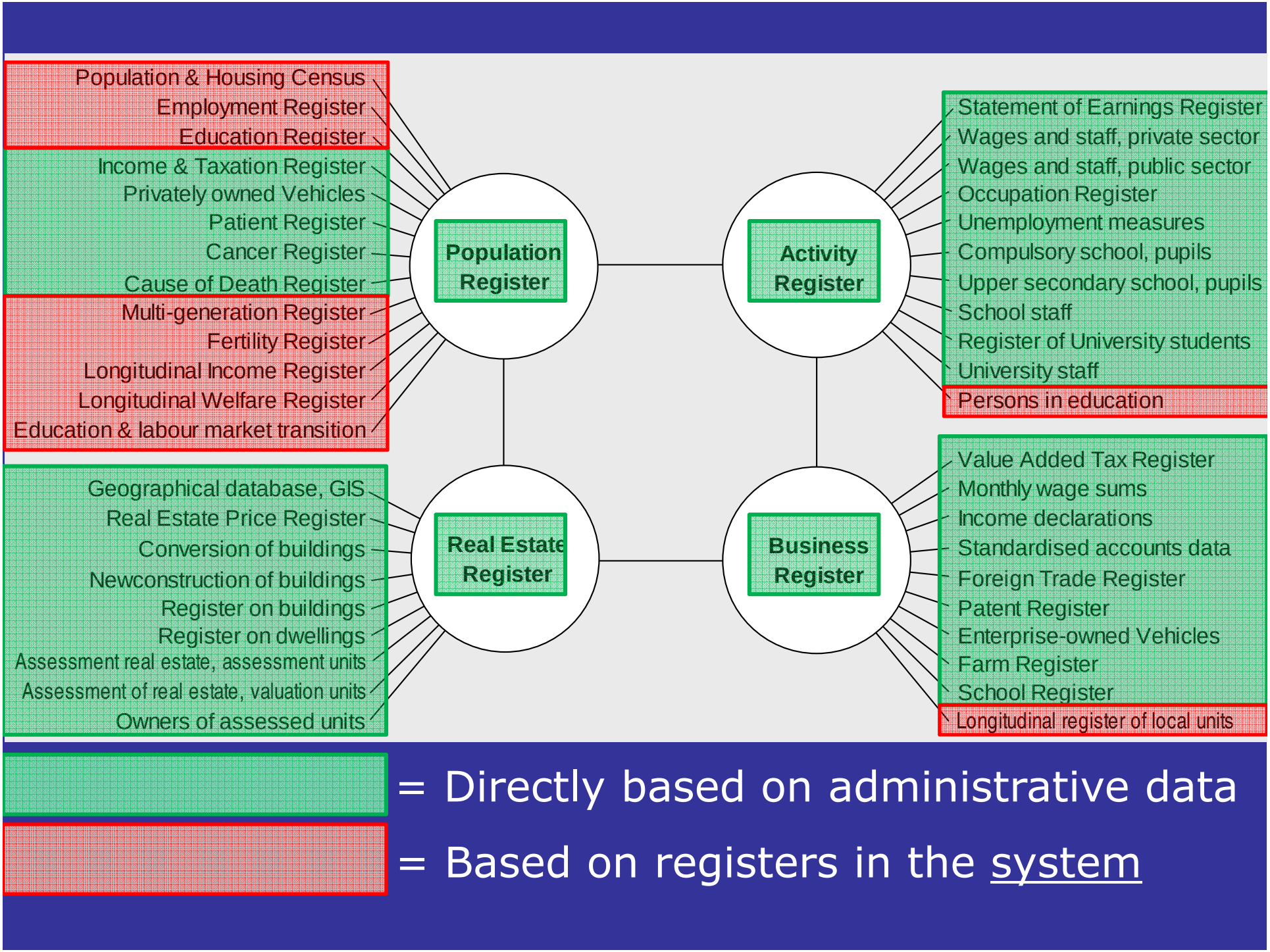
# The production systems in the Nordic countries are completely register-based. All sample surveys, censuses, register surveys are based on the Register System:



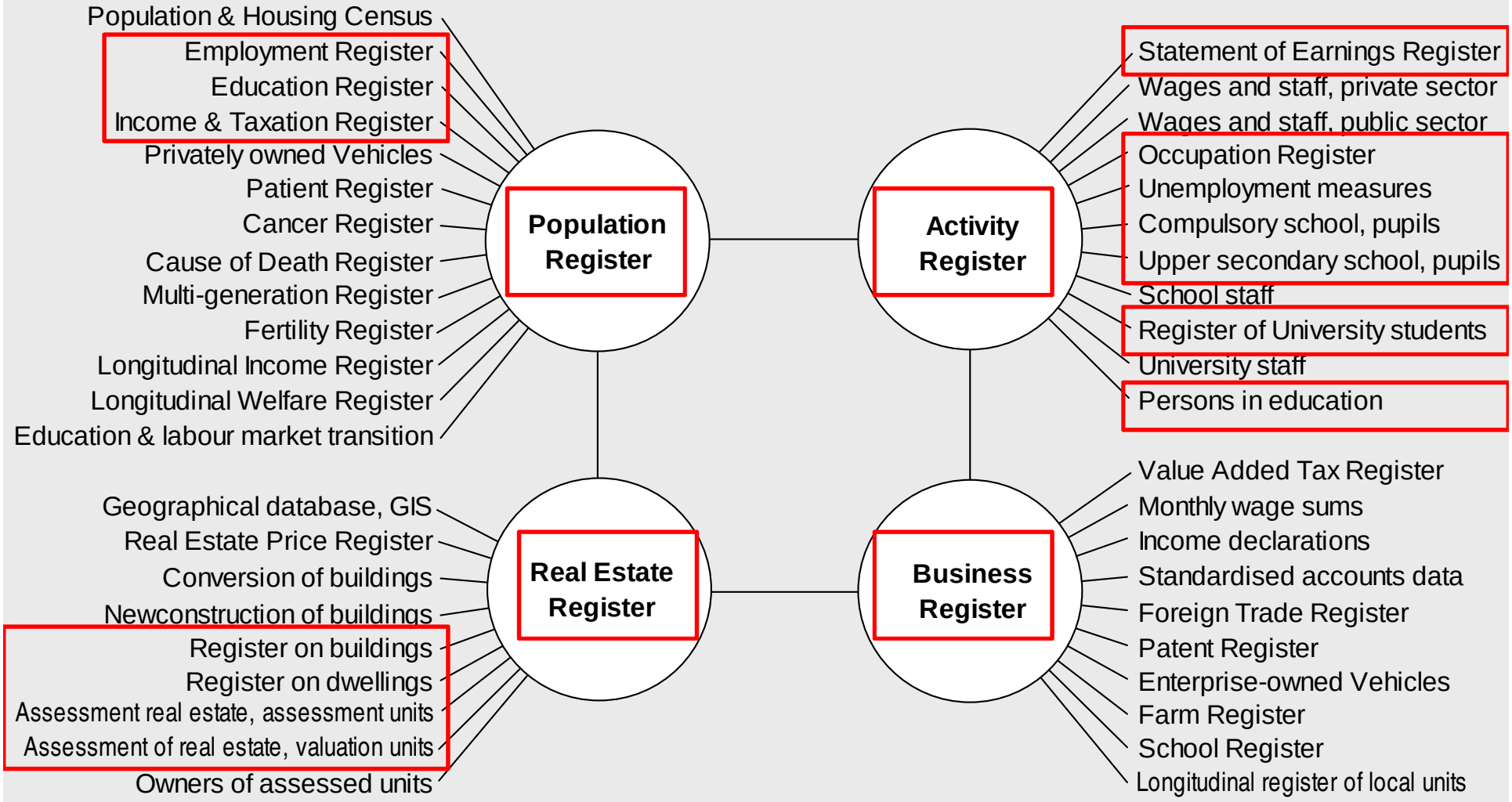
All registers can be linked with all other registers

All sample surveys can be linked with all registers




Population & Housing Census
Employment Register
Education Register
Income & Taxation Register
Privately owned Vehicles
Patient Register
Cancer Register
Cause of Death Register
Multi-generation Register
Fertility Register
Longitudinal Income Register
Longitudinal Welfare Register
Education & labour market transition
Population Register
Activity Register
Statement of Earnings Register
Wages and staff, private sector
Wages and staff, public sector
Occupation Register
Unemployment measures
Compulsory school, pupils
Upper secondary school, pupils
School staff
Register of University students
University staff
Persons in education
Geographical database, GIS
Real Estate Price Register
Conversion of buildings
Newconstruction of buildings
Register on buildings
Register on dwellings
Assessment real estate, assessment units
Assessment of real estate, valuation units
Owners of assessed units
Real Estate Register
Business Register
Value Added Tax Register
Monthly wage sums
Income declarations
Standardised accounts data
Foreign Trade Register
Patent Register
Enterprise-owned Vehicles
Farm Register
School Register
Longitudinal register of local units
= Directly based on administrative data
= Based on registers in the system

# A System of Statistical Registers – register by object type and subject field




## Register-based Census

*Methodology principle 3:*
Consistency regarding populations and variables are necessary for the coherence of estimates from different register surveys.

Coordinated versions of a number of registers are produced for the register-based census:

**Chart 2.13 Register-based statistics for one small municipality in Sweden 2003**

| **Population** | | **Employment** | | **Education** | | | | | **Income** Yearly earned, $ thousands | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Age | Number | Em-ployed | Not em-ployed | Com-pulsory | Upper secon-dary | Post-secon-dary | Post-gra-duate | Not known | 0 | 1–14 | 15–29 | 30–44 | 45+ |
| 0–15 | 1416 | - | - | - | - | - | - | - | - | - | - | - | - |
| 16–19 | 387 | 69 | 318 | 306 | 71 | 0 | 0 | 10 | 118 | 265 | 4 | 0 | 0 |
| 20–24 | 293 | 207 | 86 | 44 | 219 | 26 | 0 | 4 | 12 | 130 | 128 | 23 | 0 |
| 25–34 | 764 | 616 | 148 | 79 | 469 | 210 | 0 | 6 | 20 | 133 | 388 | 202 | 21 |
| 35–44 | 937 | 782 | 155 | 142 | 558 | 226 | 2 | 9 | 27 | 128 | 440 | 270 | 72 |
| 45–54 | 1002 | 847 | 155 | 259 | 510 | 225 | 4 | 4 | 14 | 90 | 501 | 318 | 79 |
| 55–64 | 1042 | 713 | 329 | 420 | 413 | 199 | 6 | 4 | 21 | 166 | 502 | 288 | 65 |
| 65+ | 1199 | 40 | 1159 | 333 | 168 | 78 | 3 | 617 | 3 | 552 | 535 | 90 | 19 |

**BUT:**

# Employees by economic activity November 2004, thousands

| | Business Register | | Employment | Labour Force Survey | |
|---|---|---|---|---|---|
| | Enterprises | Establishments | Register | | *Error margin* |
| **Economic activity** | ***(1)*** | ***(2)*** | ***(3)*** | ***(4)*** | ***(5)*** |
| Agriculture and forestry | 34 | 36 | 37 | 26 | *5* |
| Fishing | 1 | 1 | 0 | 0 | *1* |
| Mining and quarrying | 9 | 8 | 7 | 5 | *2* |
| Manufacturing | 679 | 629 | 710 | 635 | *23* |
| Electricity, gas and water | 21 | 22 | 28 | 29 | *5* |
| Construction | 197 | 209 | 215 | 199 | *14* |
| Wholesale and retail trade | 456 | 453 | 484 | 456 | *20* |
| Hotels and restaurants | 89 | 93 | 99 | 106 | *10* |
| Transport, communication | 240 | 242 | 243 | 236 | *15* |
| Financial intermediation | 83 | 77 | 85 | 78 | *9* |
| Real estate, business activities | 457 | 524 | 457 | 470 | *20* |
| Government | 139 | 215 | 239 | 230 | *15* |
| Education | 382 | 408 | 431 | 462 | *20* |
| Health and social work | 836 | 684 | 675 | 675 | *24* |
| Other service activities | 142 | 163 | 175 | 168 | *13* |
| Unknown activity | 0 | 0 | 38 | 4 | |
| **Total** | **3 763** | **3 763** | **3 924** | **3 778** | ***43*** |

# Employees by economic activity November 2004, thousands

| | Business Register | | Employment | Labour Force Survey | |
|---|---|---|---|---|---|
| | Enterprises | Establishments | Register | | *Error margin* |
| **Economic activity** | ***(1)*** | ***(2)*** | ***(3)*** | ***(4)*** | ***(5)*** |
| Agriculture and forestry | 34 | 36 | 37 | 26 | *5* |
| Fishing | 1 | 1 | 0 | 0 | *1* |
| Mining and quarrying | 9 | 8 | 7 | 5 | *2* |
| Manufacturing | 679 | 629 | 710 | 635 | *23* |
| Electricity, gas and water | 21 | 22 | 28 | 29 | *5* |
| Construction | 197 | 209 | 215 | 199 | *14* |
| Wholesale and retail trade | 456 | 453 | 484 | 456 | *20* |
| Hotels and restaurants | 89 | 93 | 99 | 106 | *10* |
| Transport, communication | 240 | 242 | 243 | 236 | *15* |
| Financial intermediation | 83 | 77 | 85 | 78 | *9* |
| Real estate, business activities | 457 | 524 | 457 | 470 | *20* |
| Government | 139 | 215 | 239 | 230 | *15* |
| Education | 382 | 408 | 431 | 462 | *20* |
| Health and social work | 836 | 684 | 675 | 675 | *24* |
| Other service activities | 142 | 163 | 175 | 168 | *13* |
| Unknown activity | 0 | 0 | 38 | 4 | |
| **Total** | **3 763** | **3 763** | **3 924** | **3 778** | ***43*** |

## Employees by economic activity November 2004, thousands

| | Business Register | | Employment | Labour Force Survey | |
|---|---|---|---|---|---|
| | Enterprises | Establishments | Register | | *Error margin* |
| **Economic acti** | ***(1)*** | ***(2)*** | ***(3)*** | ***(4)*** | ***(5)*** |
| Agriculture and | 34 | 36 | 37 | 26 | *5* |
| Fishing | 1 | 1 | 0 | 0 | *1* |
| Mining and qua | 9 | 8 | 7 | 5 | *2* |
| Manufacturing | 679 | 629 | 710 | 635 | *23* |
| Electricity, gas | 21 | 22 | 28 | 29 | *5* |
| Construction | 197 | 209 | 215 | 199 | *14* |
| Wholesale and | 456 | 453 | 484 | 456 | *20* |
| Hotels and rest | 89 | 93 | 99 | 106 | *10* |
| Transport, com | 240 | 242 | 243 | 236 | *15* |
| Financial interm | 83 | 77 | 85 | 78 | *9* |
| Real estate, bu | 457 | 524 | 457 | 470 | *20* |
| Government | 139 | 215 | 239 | 230 | *15* |
| Education | 382 | 408 | 431 | 462 | *20* |
| Health and soci | 836 | 684 | 675 | 675 | *24* |
| Other service a | 142 | 163 | 175 | 168 | *13* |
| Unknown activity | 0 | 0 | 38 | 4 | |
| **Total** | **3 763** | **3 763** | **3 924** | **3 778** | ***43*** |

1) 1th Industry Enterprises

2) 1th Industry Establish-ments

3) 1th Industry for the Establish-ment of the 1th Job of an employee

4) Same as (3)

One survey at a time thinking. No one has the time to compare. No one is responsible for the system.

## Employees by economic activity November 2004, thousands

| | Business Register | | Employment | Labour Force Survey | |
|---|---|---|---|---|---|
| | Enterprises | Establishments | Register | | *Error margin* |
| **Economic acti** | ***(1)*** | ***(2)*** | ***(3)*** | ***(4)*** | ***(5)*** |
| Agriculture and | 34 | 36 | 37 | 26 | *5* |
| Fishing | 1 | 1 | 0 | 0 | *1* |
| Mining and qua | 9 | 8 | 7 | 5 | *2* |
| Manufacturing | 679 | 629 | 710 | 635 | *23* |
| Electricity, gas | 21 | 22 | 28 | 29 | *5* |
| Construction | 197 | 209 | 215 | 199 | *14* |
| Wholesale and | 456 | 453 | 484 | 456 | *20* |
| Hotels and rest | 89 | 93 | 99 | 106 | *10* |
| Transport, com | 240 | 242 | 243 | 236 | *15* |
| Financial interm | 83 | 77 | 85 | 78 | *9* |
| Real estate, bu | 457 | 524 | 457 | 470 | *20* |
| Government | 139 | 215 | 239 | 230 | *15* |
| Education | 382 | 408 | 431 | 462 | *20* |
| Health and soci | 836 | 684 | 675 | 675 | *24* |
| Other service a | 142 | 163 | 175 | 168 | *13* |
| Unknown activity | 0 | 0 | 38 | 4 | |
| **Total** | **3 763** | **3 763** | **3 924** | **3 778** | ***43*** |

1) 1th Industry Enterprises

2) 1th Industry Establish-ments

3) 1th Industry for the Establish-ment of the 1th Job of an employee

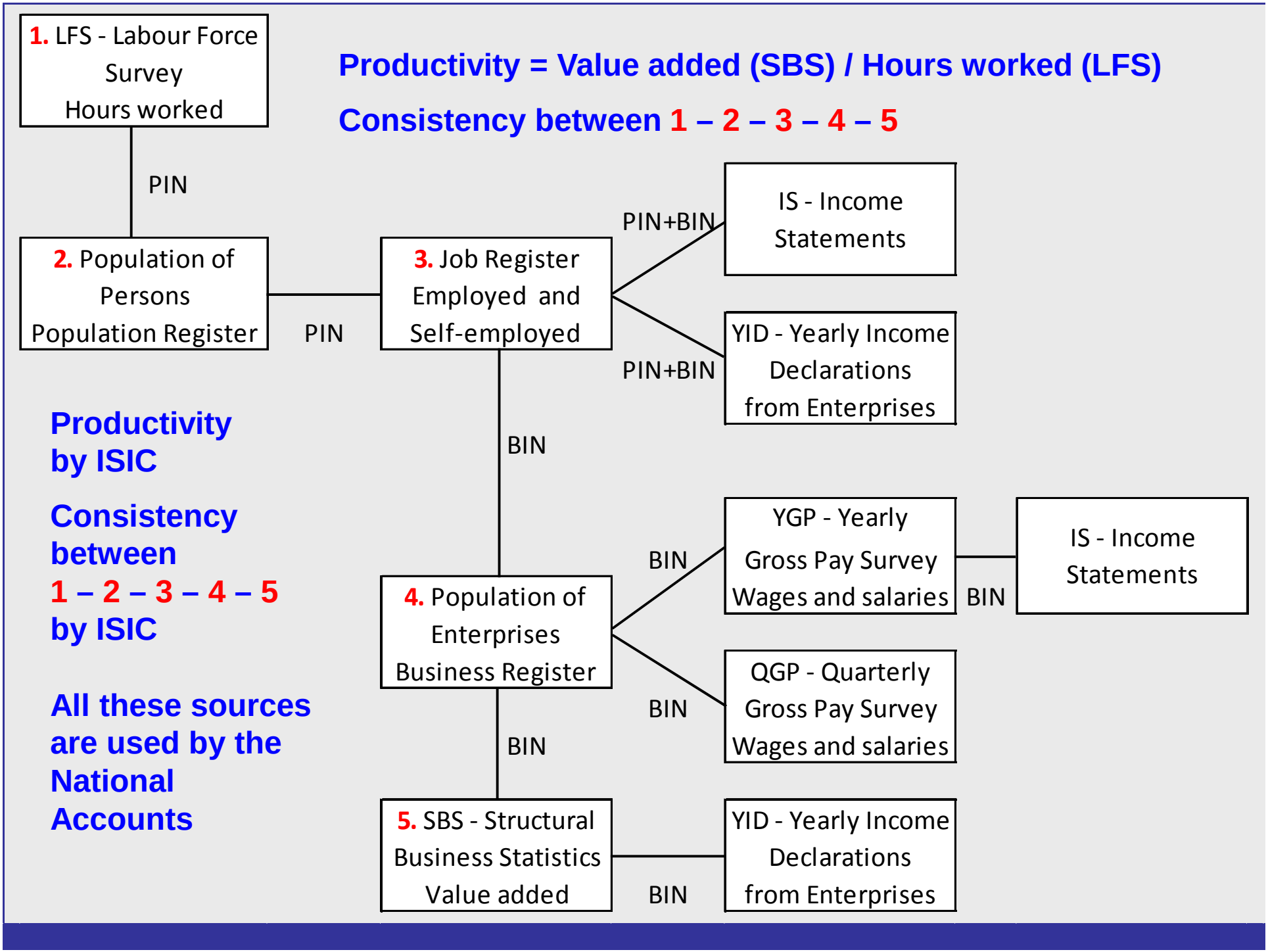
4) Same as (3)

The National Accounts receive macrodata that are inconsistent in this way. Economic activity!

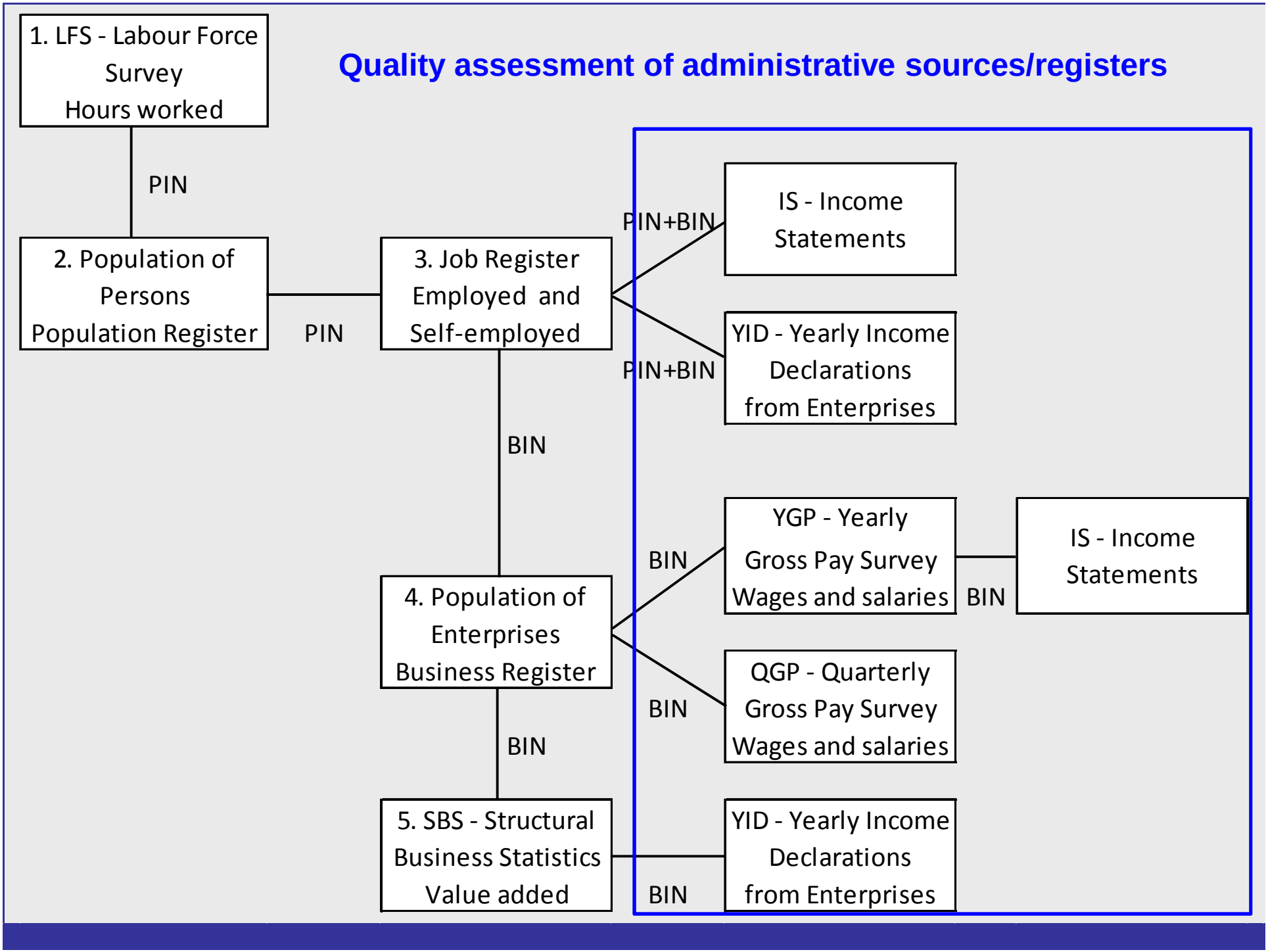
## Our present research project:

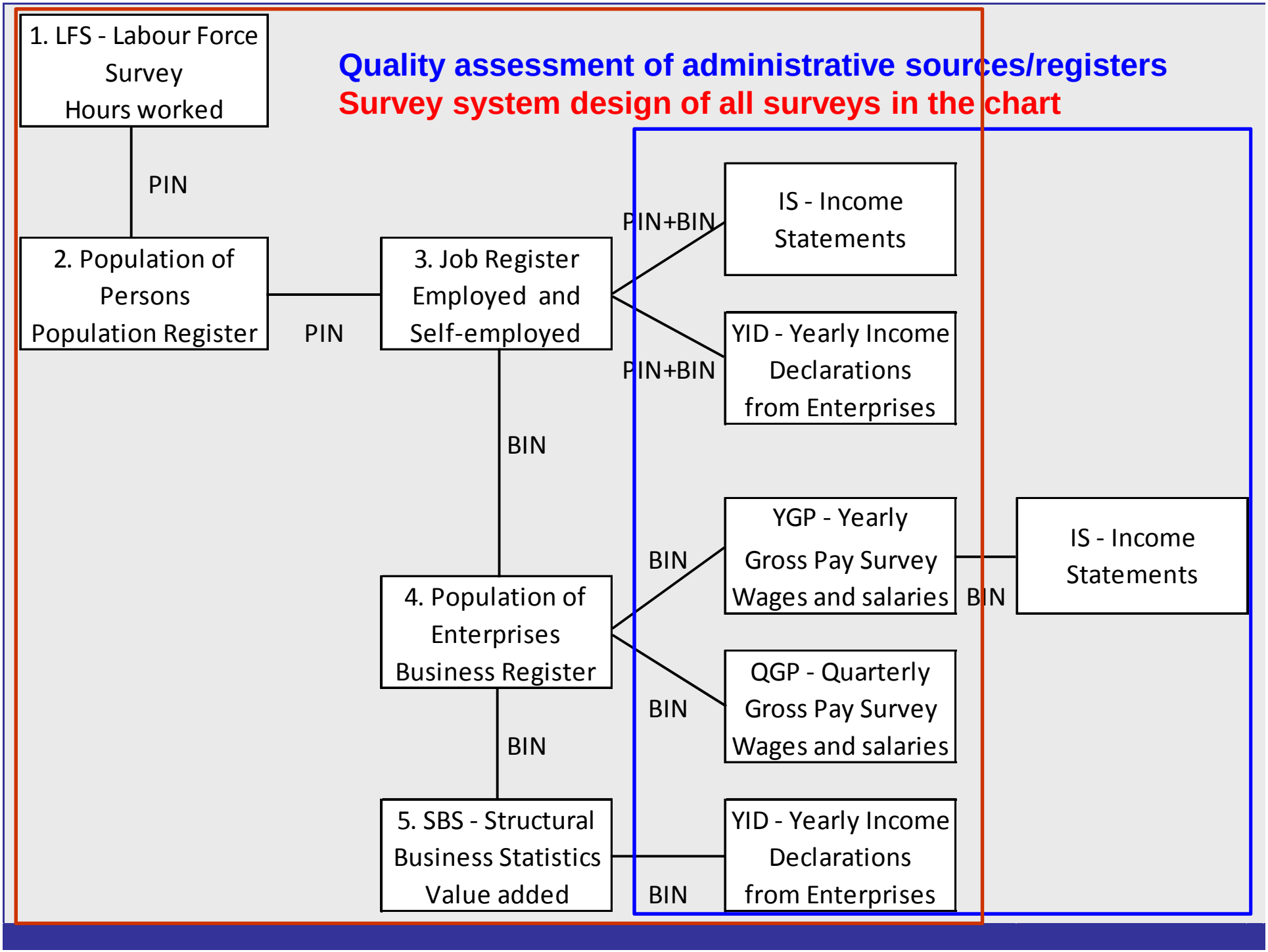
*Methodology principle 4:*
The register system should be used for quality assessment of statistical surveys based on microdata comparisons with other surveys in the production system.

- Productivity by institutional sector and economic activity
- Quality assessment of all administrative registers used
- Survey system design of all surveys in the system

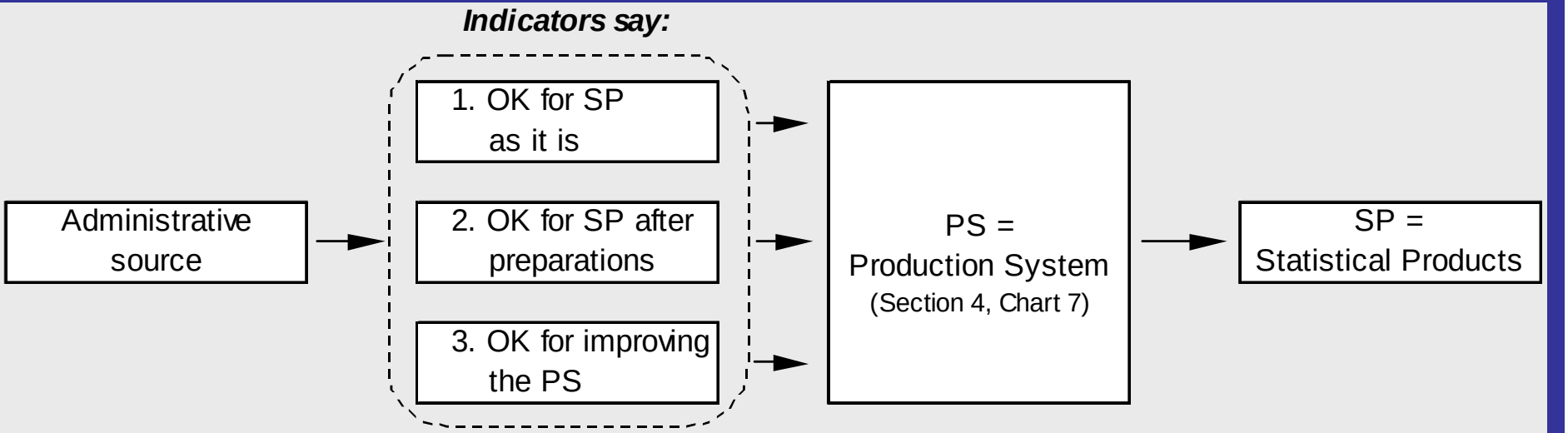

1. LFS - Labour Force Survey
Hours worked
PIN
2. Population of Persons
Population Register
PIN
3. Job Register
Employed and
Self-employed
PIN+BIN
IS - Income Statements
PIN+BIN
YID - Yearly Income Declarations from Enterprises
BIN
4. Population of Enterprises
Business Register
BIN
YGP - Yearly Gross Pay Survey
Wages and salaries
BIN
IS - Income Statements
BIN
QGP - Quarterly Gross Pay Survey
Wages and salaries
BIN
5. SBS - Structural Business Statistics
Value added
BIN
YID - Yearly Income Declarations from Enterprises
Productivity = Value added (SBS) / Hours worked (LFS)
Consistency between 1 – 2 – 3 – 4 – 5
Productivity by ISIC
Consistency between 1 – 2 – 3 – 4 – 5 by ISIC
All these sources are used by the National Accounts

# Quality assessment of administrative sources/registers

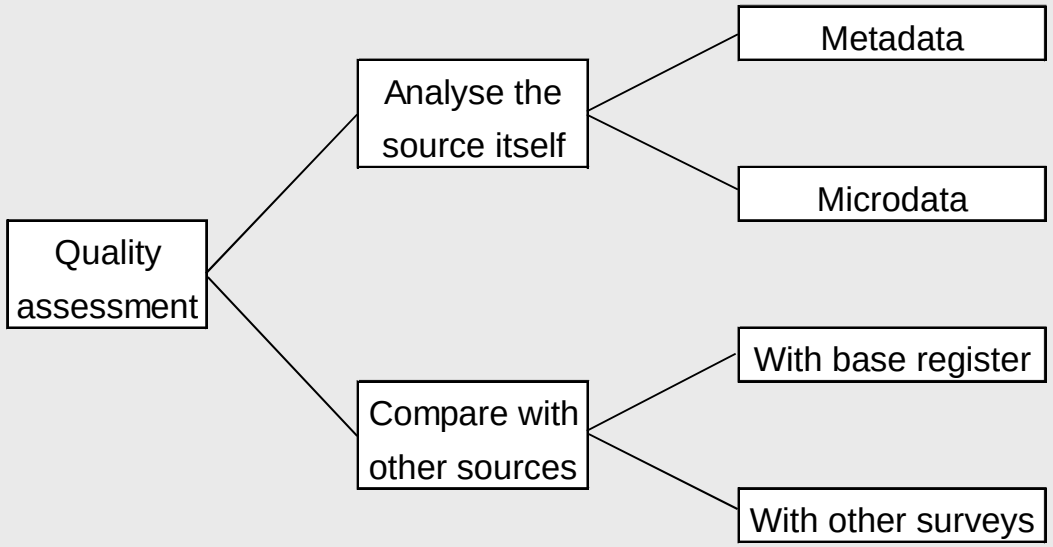



Quality assessment of administrative sources/registers
Survey system design of all surveys in the chart
1. LFS - Labour Force Survey Hours worked
PIN
2. Population of Persons Population Register
PIN
3. Job Register Employed and Self-employed
PIN+BIN
IS - Income Statements
PIN+BIN
YID - Yearly Income Declarations from Enterprises
BIN
4. Population of Enterprises Business Register
BIN
YGP - Yearly Gross Pay Survey Wages and salaries
BIN
IS - Income Statements
BIN
QGP - Quarterly Gross Pay Survey Wages and salaries
BIN
5. SBS - Structural Business Statistics Value added
BIN
YID - Yearly Income Declarations from Enterprises

## Chart 1. Three ways of using an administrative source in the Production System




## Chart 3. The four steps during the work with quality assessment




| Quality issue: | Group of indicators: |
|---|---|
| Quality of source? | A |
| Quality of source? | B |
| Quality of source? Quality of base register? Can the source improve the base register? | C |
| Quality of source? Quality of production system? Can the source be combined with other surveys? Can the source improve the production system? | D |

## 4.A Metadata – Information from the Administrative Authority

| Indicator | Quality factor | Description |
|---|---|---|
| A1 | Relevance of population | Definition of the administrative object set. Which administrative rules determine which objects are included? Is this set suitable as statistical population? |
| A2 | Relevance of units | Definition of the administrative units. Are these units suitable as statistical units? |
| A3 | Relevant keys | Are there primary keys and foreign keys in the source that are suitable for micro integration within the NSI? |
| A4 | Relevance of variables | Definitions of the administrative variables. Are these variables suitable as statistical variables? |
| A5 | Relevance of reference time | Are reference times suitable for statistical usage? What rules for accruing accounting data between months and years are used? |
| A6 | Study domains | Can the units be allocated between relevant study domains? Are there variables describing domains in the source or can the units be linked with domain variables in the Business Register? |
| A7 | Comprehen-siveness | Does the source contain a small/large part of an intended population? Does the source contain few/many statistically interesting variables? Can a small/large number of existing surveys benefit from the administrative source? |
| A8 | Updates, delivery time and punctuality | How often and at what time points is the administrative register updated? Time for delivery of the source from register holder to the NSI. Difference in time between delivery and agreed delivery time. |
| A9 | Comparability over time | Extent of changes in the content of the administrative register over time |

# Indicators of output and input data quality – relevance

## 4.B Analysis and Data Editing of the Source

| Indicator | Quality factor | Description |
|---|---|---|
| B1 | Primary key | Fraction of units with usable identities. The primary key should have correct format and reasonable values. |
| B2 | Foreign keys | Fraction of units with usable foreign keys. Foreign keys should have correct format and reasonable values. |
| B3 | Missing values | Fraction of missing values for the statistically interesting variables. |
| B4 | Wrong values | Fraction of wrong or unreasonable values for the statistically interesting variables. |
| B5 | Quality of preliminary data | Fraction of records corrected by the taxpayers. Estimates based on preliminary data are compared with estimates based on final data. |

# Indicators of output and input data quality – accuracy

## 4.C Integrate the Source with the Base Register

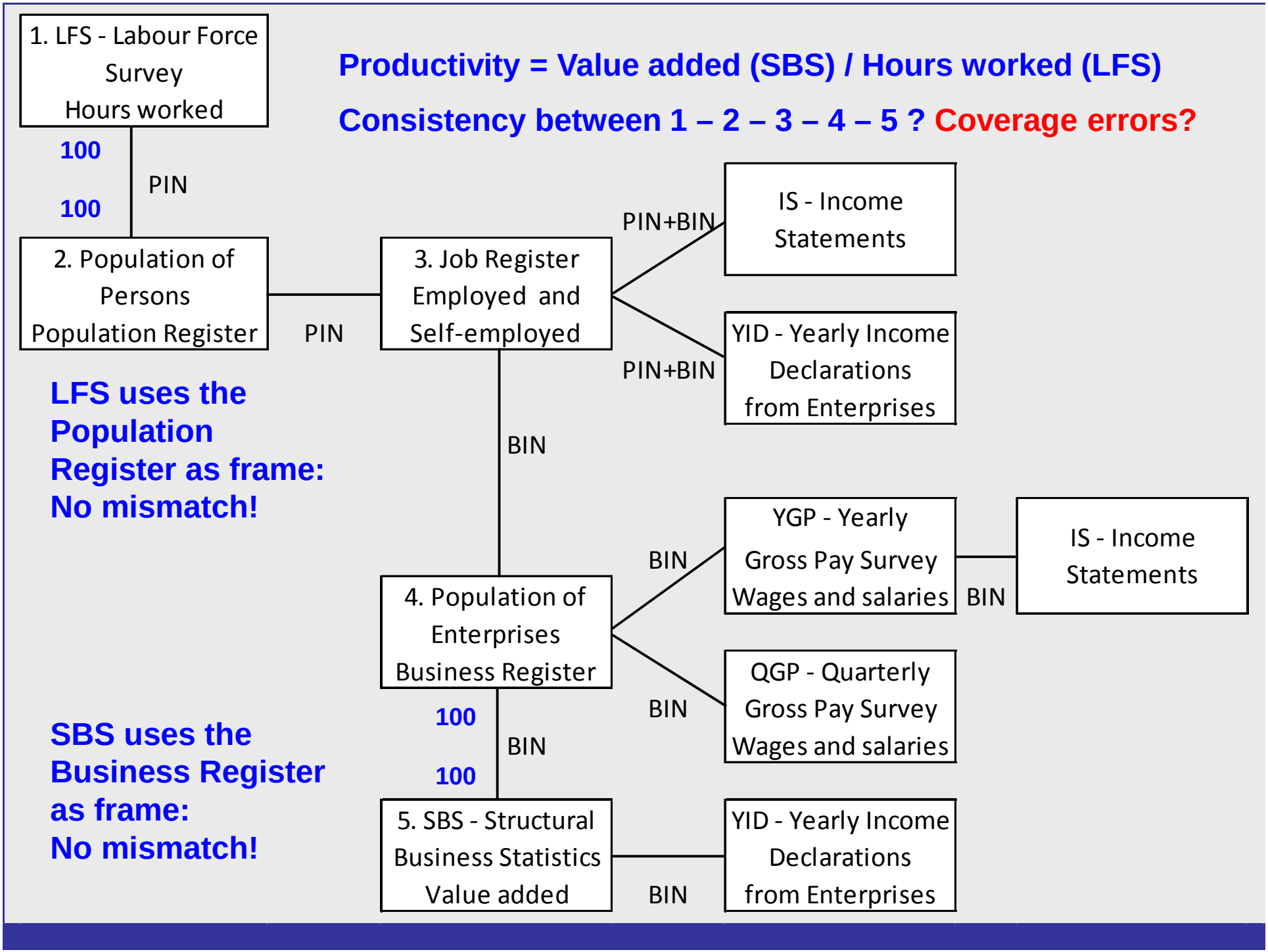
| Indicator | Quality factor | Description |
|---|---|---|
| C1 | Undercoverage in BR | Fraction of units: There are enterprises/units that have been active during the reference period but are missing in the BR or are coded as inactive in the BR. |
| C2 | Overcoverage in BR | Fraction of units: Enterprises/units are coded as active in the BR and belong to a category that is covered by the source, but they have no reported activity in the source. |
| C3 | Undercoverage in the source | Fraction of units: There are enterprises/units that have been active during the reference period according to the BR but are missing in the source. |
| C4 | Overcoverage in the source | Fraction of units: There are units in the source that belong to a category, or seem to belong to a category, that is not statistically relevant. |
| C5 | Can the source improve BR? | Here a more thorough analysis is required depending on the character of the source. The quality improvements should be measured. |

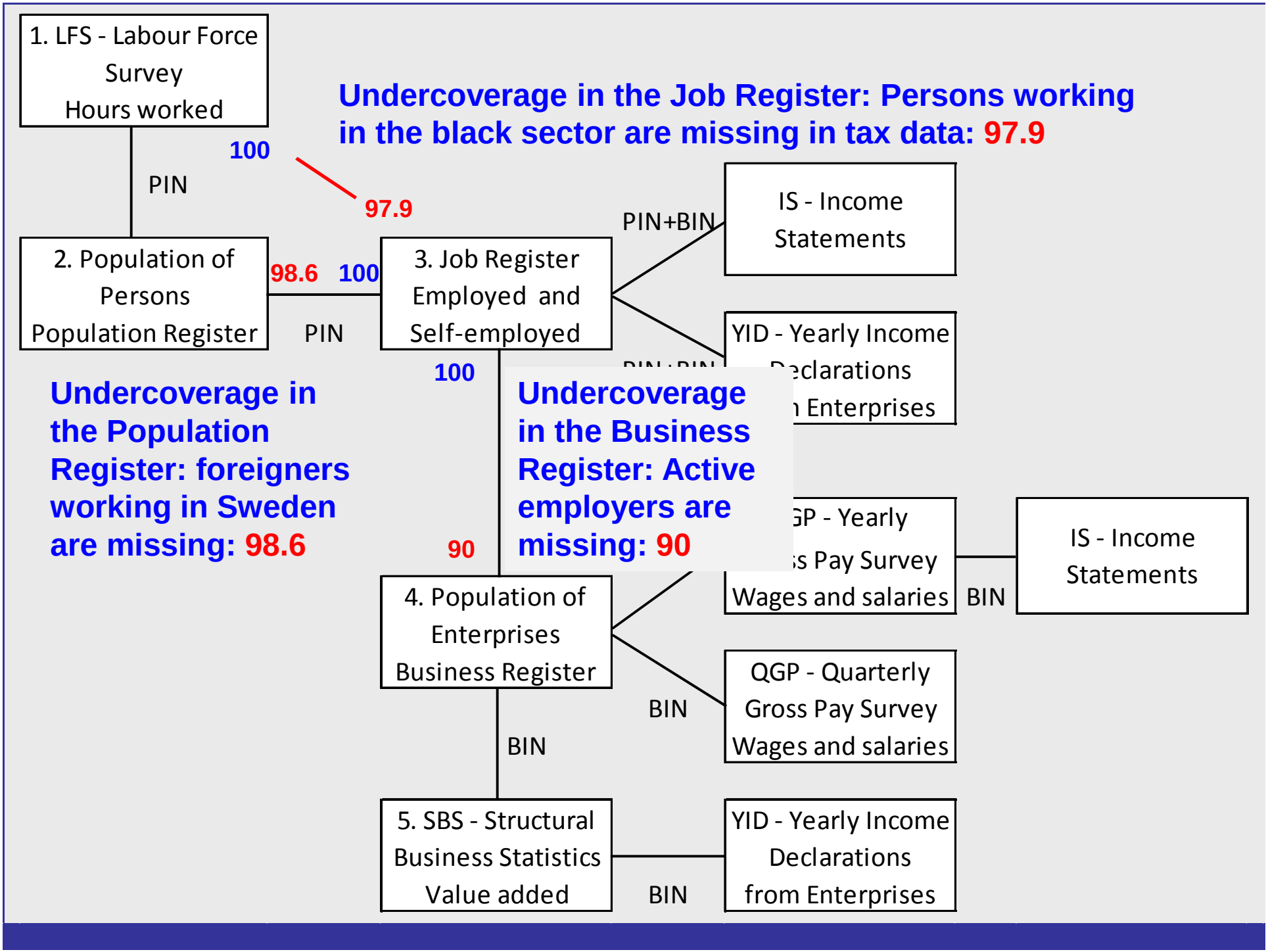
# Indicators of output and input data quality – accuracy

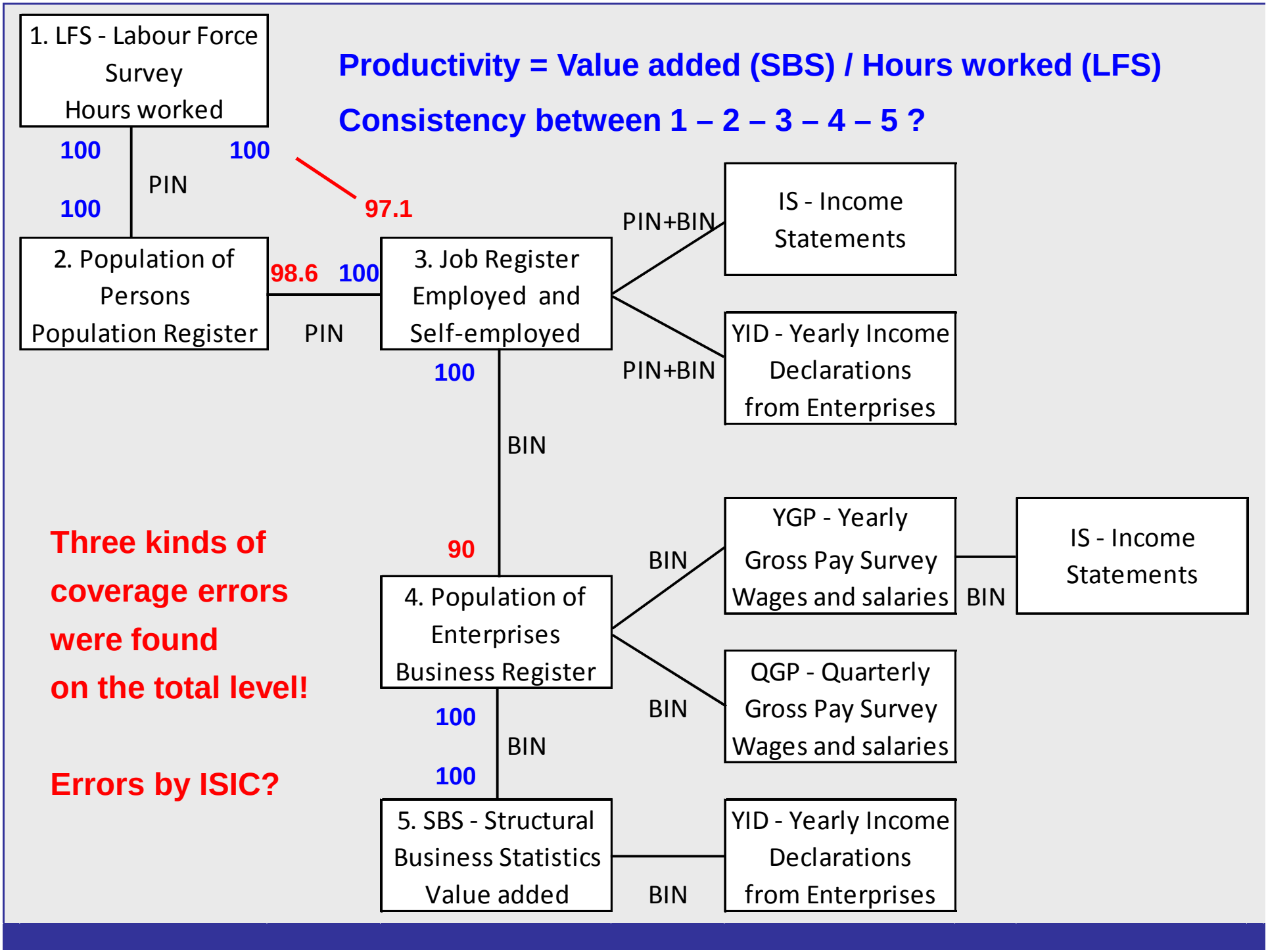
## 4.D Integrate the Source with Surveys with Similar Variables

| Indicator | Quality factor | Description | |
|---|---|---|---|
| D1 | Is the source good or bad? | a) Compare populations<br>b) Compare units<br>c) Compare variables | Measures production process quality |
| D2 | Is the production system good or bad? | a) Compare populations<br>b) Compare units<br>c) Compare variables | Measures production process quality |
| D3 | Can the source improve other surveys? | a) Will population be better?<br>b) Will units be better?<br>c) Will variables be better? | Measures production process quality |
| D4 | Can the source be combined with other sources? | a) Will population be better?<br>b) Will units be better?<br>c) Will variables be better? | Measures input data quality |

**Indicators on input data and production process quality**


1. LFS - Labour Force Survey
Hours worked
100
100
PIN
2. Population of Persons
Population Register
PIN
3. Job Register
Employed and
Self-employed
PIN+BIN
IS - Income Statements
PIN+BIN
YID - Yearly Income Declarations from Enterprises
BIN
4. Population of Enterprises
Business Register
BIN
YGP - Yearly Gross Pay Survey
Wages and salaries
BIN
IS - Income Statements
BIN
QGP - Quarterly Gross Pay Survey
Wages and salaries
100
100
BIN
5. SBS - Structural Business Statistics
Value added
BIN
YID - Yearly Income Declarations from Enterprises
Productivity = Value added (SBS) / Hours worked (LFS)
Consistency between 1 – 2 – 3 – 4 – 5 ? Coverage errors?
LFS uses the Population Register as frame: No mismatch!
SBS uses the Business Register as frame: No mismatch!


1. LFS - Labour Force Survey Hours worked
100
PIN
97.9
Undercoverage in the Job Register: Persons working in the black sector are missing in tax data: 97.9
2. Population of Persons Population Register
98.6
100
PIN
3. Job Register Employed and Self-employed
PIN+BIN
IS - Income Statements
YID - Yearly Income Declarations
Enterprises
100
Undercoverage in the Population Register: foreigners working in Sweden are missing: 98.6
Undercoverage in the Business Register: Active employers are missing: 90
90
4. Population of Enterprises Business Register
GP - Yearly
s Pay Survey
Wages and salaries
BIN
IS - Income Statements
BIN
QGP - Quarterly Gross Pay Survey Wages and salaries
BIN
5. SBS - Structural Business Statistics Value added
BIN
YID - Yearly Income Declarations from Enterprises


1. LFS - Labour Force Survey Hours worked
Productivity = Value added (SBS) / Hours worked (LFS)
Consistency between 1 – 2 – 3 – 4 – 5 ?
100
100
PIN
100
97.1
2. Population of Persons Population Register
98.6
100
PIN
3. Job Register Employed and Self-employed
PIN+BIN
IS - Income Statements
PIN+BIN
YID - Yearly Income Declarations from Enterprises
100
BIN
Three kinds of coverage errors were found on the total level!
90
4. Population of Enterprises Business Register
BIN
YGP - Yearly Gross Pay Survey Wages and salaries
BIN
IS - Income Statements
BIN
QGP - Quarterly Gross Pay Survey Wages and salaries
100
BIN
100
Errors by ISIC?
5. SBS - Structural Business Statistics Value added
BIN
YID - Yearly Income Declarations from Enterprises

## Example of integrated microdata from the LFS and the IS

| LFS PIN (1) | LFS Hours worked (2) | LFS Hours usually worked (3) | LFS Sector (4) | LFS ISIC (5) | LFS Weight (6) | IS PIN (7) | IS ISIC (8) | IS Sector (9) |
|---|---|---|---|---|---|---|---|---|
| PIN1 | 12 | 20 | 6 | 56100 | 32.2 | PIN1 | 56100 | 110 |
| PIN1 | 16 | 20 | 6 | 56100 | 28.8 | PIN1 | 56100 | 110 |
| PIN1 | 0 | 20 | 6 | 56100 | 27.9 | PIN1 | 56100 | 110 |
| PIN1 | 20 | 20 | 6 | 56100 | 33.1 | PIN1 | 56100 | 110 |
| PIN2 | 40 | 40 | 6 | 56100 | 32.4 | * | * | * |
| PIN2 | 40 | 40 | 6 | 56100 | 31.5 | * | * | * |
| PIN2 | 40 | 40 | 6 | 56100 | 33.2 | * | * | * |
| PIN3 | 40 | 40 | 1 | 01110 | 32.1 | PIN3 | 81300 | 320 |
| PIN4 | 10 | 10 | 6 | 01110 | 51.5 | PIN4 | 43320 | 611 |
| PIN5 | 45 | 40 | 6 | 01131 | 40.4 | PIN5 | 01500 | 611 |
| PIN6 | 30 | 30 | 6 | 01191 | 43.1 | PIN6 | * | * |
| PIN7 | 5 | 8 | 6 | 01191 | 45.7 | PIN7 | 01134 | 110 |
| PIN8 | 40 | 40 | 6 | 01199 | 48.1 | PIN8 | 01430 | 110 |
| PIN9 | 60 | 40 | 6 | 64190 | 47.1 | PIN9 | 55102 | 212 |
| PIN9 | 60 | 40 | 6 | 64190 | 44.7 | PIN9 | 55102 | 212 |

**Black work!**

## Example of integrated microdata from the LFS and the IS

| LFS PIN (1) | LFS Hours worked (2) | LFS Hours usually worked (3) | LFS Sector (4) | LFS ISIC (5) | LFS Weight (6) | IS PIN (7) | IS ISIC (8) | IS Sector (9) |
|---|---|---|---|---|---|---|---|---|
| PIN1 | 12 | 20 | 6 | 56100 | 32.2 | PIN1 | 56100 | 110 |
| PIN1 | 16 | 20 | 6 | 56100 | 28.8 | PIN1 | 56100 | 110 |
| PIN1 | 0 | 20 | 6 | 56100 | 27.9 | PIN1 | 56100 | 110 |
| PIN1 | 20 | 20 | 6 | 56100 | 33.1 | PIN1 | 56100 | 110 |
| PIN2 | 40 | 40 | 6 | 56100 | 32.4 | * | * | * |
| PIN2 | 40 | 40 | 6 | 56100 | 31.5 | * | * | * |
| PIN2 | 40 | 40 | 6 | 56100 | 33.2 | * | * | * |
| PIN3 | 40 | 40 | 1 | 01110 | 32.1 | PIN3 | 81300 | 320 |
| PIN4 | 10 | 10 | 6 | 01110 | 51.5 | PIN4 | 43320 | 611 |
| PIN5 | 45 | 40 | 6 | 01131 | 40.4 | PIN5 | 01500 | 611 |
| PIN6 | 30 | 30 | 6 | 01191 | 43.1 | PIN6 | * | * |
| PIN7 | 5 | 8 | 6 | 01191 | 45.7 | PIN7 | 01134 | 110 |
| PIN8 | 40 | 40 | 6 | 01199 | 48.1 | PIN8 | 01430 | 110 |
| PIN9 | 60 | 40 | 6 | 64190 | 47.1 | PIN9 | 55102 | 212 |
| PIN9 | 60 | 40 | 6 | 64190 | 44.7 | PIN9 | 55102 | 212 |

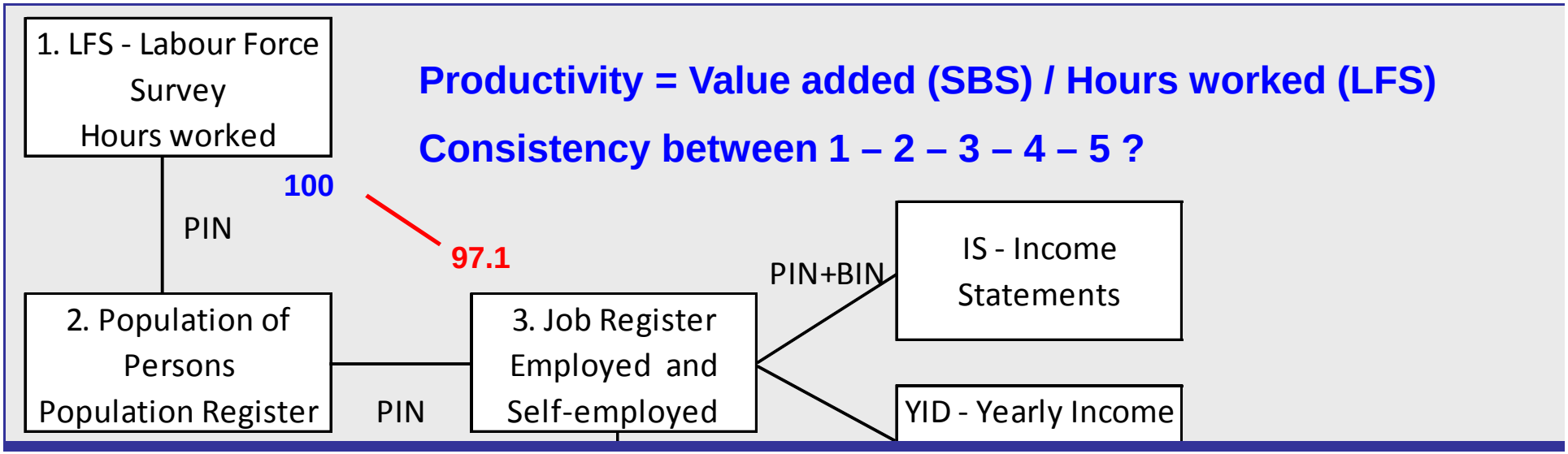
**Different ISIC!**

## Example of integrated microdata from the LFS and the IS

| LFS PIN (1) | LFS Hours worked (2) | LFS Hours usually worked (3) | LFS Sector (4) | LFS ISIC (5) | LFS Weight (6) | IS PIN (7) | IS ISIC (8) | IS Sector (9) |
|---|---|---|---|---|---|---|---|---|
| PIN1 | 12 | 20 | 6 | 56100 | 32.2 | PIN1 | 56100 | 110 |
| PIN1 | 16 | 20 | 6 | 56100 | 28.8 | PIN1 | 56100 | 110 |
| PIN1 | 0 | 20 | 6 | 56100 | 27.9 | PIN1 | 56100 | 110 |
| PIN1 | 20 | 20 | 6 | 56100 | 33.1 | PIN1 | 56100 | 110 |
| PIN2 | 40 | 40 | 6 | 56100 | 32.4 | * | * | * |
| PIN2 | 40 | 40 | 6 | 56100 | 31.5 | * | * | * |
| PIN2 | 40 | 40 | 6 | 56100 | 33.2 | * | * | * |
| PIN3 | 40 | 40 | 1 | 01110 | 32.1 | PIN3 | 81300 | 320 |
| PIN4 | 10 | 10 | 6 | 01110 | 51.5 | PIN4 | 43320 | 611 |
| PIN5 | 45 | 40 | 6 | 01131 | 40.4 | PIN5 | 01500 | 611 |
| PIN6 | 30 | 30 | 6 | 01191 | 43.1 | PIN6 | * | * |
| PIN7 | 5 | 8 | 6 | 01191 | 45.7 | PIN7 | 01134 | 110 |
| PIN8 | 40 | 40 | 6 | 01199 | 48.1 | PIN8 | 01430 | 110 |
| PIN9 | 60 | 40 | 6 | 64190 | 47.1 | PIN9 | 55102 | 212 |
| PIN9 | 60 | 40 | 6 | 64190 | 44.7 | PIN9 | 55102 | 212 |

**Different sector!**

**LFS not good for National Accounts!**




## Hours worked by employed in the LFS and the IS, millions per week 2009

| ISIC | All hours in LFS | Hours not in IS | Not in IS, % |
|---|---|---|---|
| Agriculture, forestry and fishing | 1.129 | 0.036 | 3.2 |
| Mining and quarrying | 0.275 | 0.025 | 9.3 |
| Construction | 7.447 | 0.196 | 2.6 |
| Wholesale and retail trade | 13.536 | 0.203 | 1.5 |
| Transport, storage | 6.469 | 0.174 | 2.7 |
| Hotels and restaurants | 3.070 | 0.143 | 4.6 |
| Administrative and support activities | 5.065 | 0.091 | 1.8 |
| Government | 7.784 | 0.048 | 0.6 |
| Arts, entertainment, recreation | 2.170 | 0.064 | 3.0 |
| Other service activities | 2.499 | 0.104 | 4.2 |
| ... | ... | ... | ... |
| All | 115.064 | 1.800 | 1.6 |

**Black work by ISIC!**
**Black hours in the LFS**
**But value added in SBS does not include black work**

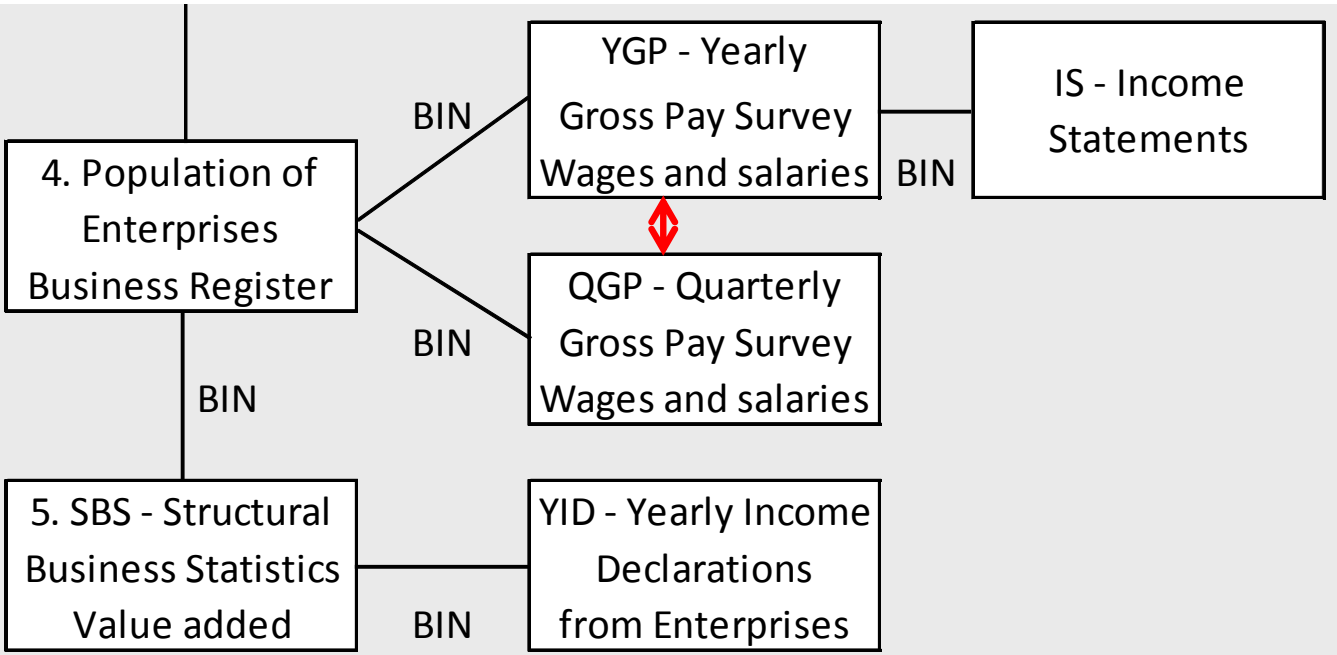
# ISIC classification differs between LFS – Job Register – Business Register – SBS

## Productivity by ISIC?

**Number of employed with one job by ISIC in the LFS, thousands**

| ISIC | ISIC | ISIC in Income Statements | Same code in LFS | Wrong code in LFS Persons | % |
|---|---|---|---|---|---|
| Manufacture of beverages | 11 | 3 687 | 2 146 | 1 541 | 41.8 |
| Manufacture of tobacco products | 12 | 496 | 381 | 115 | 23.2 |
| Pharmaceutical products | 21 | 12 227 | 5 728 | 6 499 | 53.2 |
| Computer, electronic, optical products | 26 | 32 366 | 17 426 | 14 940 | 46.2 |
| Electricity, gas, steam, air conditioning | 35 | 19 358 | 14 993 | 4 365 | 22.6 |
| Water supply | 36 | 2 740 | 2 161 | 579 | 21.1 |
| Sewerage | 37 | 1 950 | 1 461 | 489 | 25.1 |
| Wholesale trade | 46 | 142 865 | 127 928 | 14 937 | 10.5 |
| Retail trade | 47 | 175 486 | 162 891 | 12 595 | 7.2 |
| Business support activities | 82 | 32 015 | 16 766 | 15 249 | 47.6 |
| Public administration | 84 | 167 722 | 149 958 | 17 764 | 10.6 |
| Education | 85 | 310 805 | 286 170 | 24 635 | 7.9 |
| Residential care activities | 87 | 149 109 | 120 795 | 28 314 | 19.0 |
| Social work activities without accommodation | 88 | 106 343 | 67 330 | 39 013 | 36.7 |
| **...** | **...** | **...** | **...** | **...** | **...** |
| | **All:** | **2 868 809** | **2 530 335** | **338 474** | **11.8** |

## Compare Yearly Gross Pay and Quarterly Gross Pay



# Comparing gross pay in QGP and YGP, microdata

| BIN | NACE | Gross Pay, SEK millions | | |
|---|---|---|---|---|
| | | QGP | YGP | QGP-YGP |
| BIN1 | 65 | 5 956 | 265 | 5 692 |
| BIN2 | 65 | 1 455 | 310 | 1 145 |
| BIN3 | 65 | 817 | 1 | 816 |
| BIN4 | 65 | 328 | 8 | 320 |
| BIN5 | 41 | 259 | 663 | -404 |
| BIN6 | 43 | 115 | 0 | 115 |
| BIN7 | 43 | 112 | 0 | 112 |
| BIN8 | 42 | 175 | 0 | 175 |
| BIN9 | 29 | 25 | 110 | -85 |
| BIN10 | 25 | 84 | 0 | 84 |
| BIN11 | 47 | 681 | 731 | -50 |
| BIN12 | 46 | 50 | 0 | 50 |

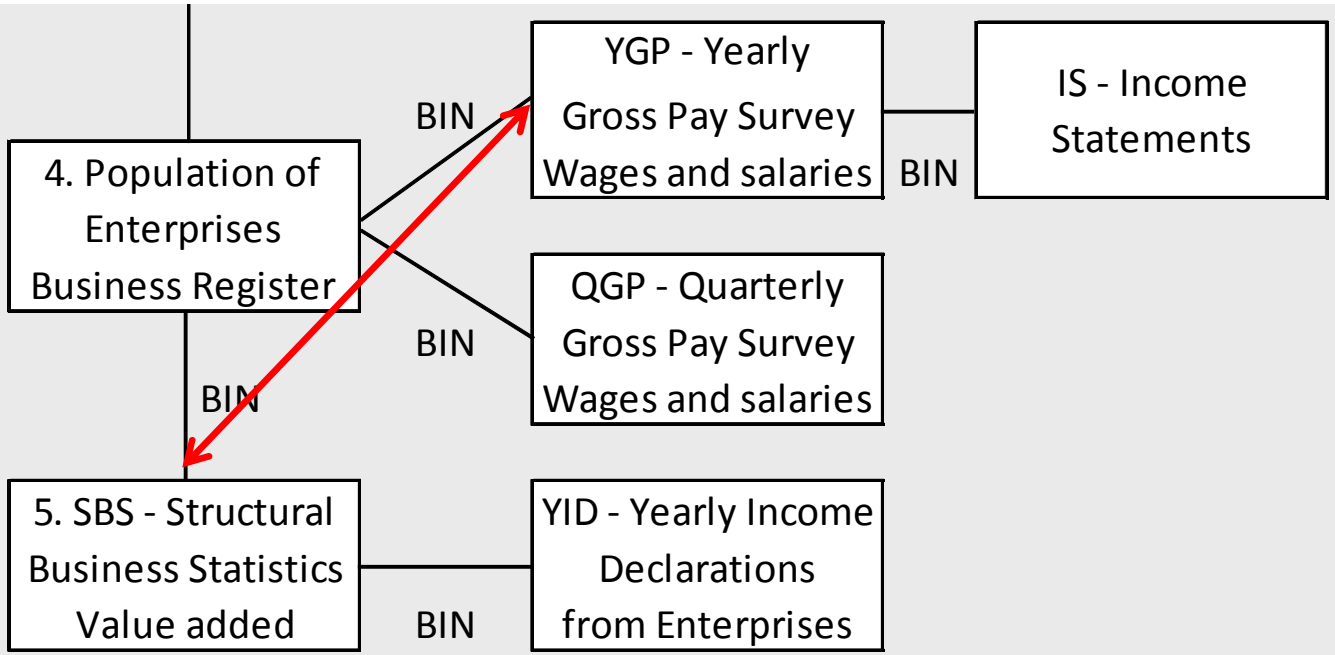
# Comparing gross pay in QGP and YGP, macrodata

| NACE | Gross Pay, SEK millions | | | Number of enterprises |
|---|---|---|---|---|
| | QGP | YGP | QGP - YGP | |
| 65 | 16 113 | 8 469 | 7 644 | 567 |
| 43 | 46 740 | 47 314 | -574 | 25 796 |
| 64 | 26 991 | 26 605 | 386 | 1 903 |
| 47 | 55 317 | 55 553 | -236 | 25 727 |
| 93 | 7 378 | 7 562 | -184 | 7 476 |
| 46 | 67 346 | 67 526 | -180 | 20 388 |
| 78 | 14 701 | 14 868 | -168 | 1 799 |
| 29 | 21 271 | 21 106 | 165 | 559 |
| ... | ... | ... | ... | ... |
| All | 1 246 593 | 1 241 138 | 5 454 | 307 230 |
| Adj.all | 1 238 949 | 1 241 138 | -2 189 | |

**Coverage errors in the SBS compared with YGP**

**Productivity by ISIC?**

**Next page!**



# Inconsistencies on the macro level between SBS and the Yearly Gross Pay Survey Gross Pay 2009 by industry in SEK Millions

| | SBS | | YGP | | Gross Pay |
|---|---|---|---|---|---|
| NACE | Gross Pay | Number of units | Gross Pay | Number of units | SBS/YGP |
| 70 | 14 731.0 | 26 520 | 22 845.8 | 14 384 | 0.64 |
| 08-09 | 823.3 | 429 | 942.0 | 338 | 0.87 |
| 50 | 3 799.0 | 653 | 4 249.1 | 414 | 0.89 |
| 52 | 7 032.9 | 2 303 | 7 553.1 | 1 780 | 0.93 |
| 68 | 14 788.5 | 18 971 | 15 832.7 | 24 500 | 0.93 |
| 94 | 1 714.5 | 1 067 | 1 832.3 | 798 | 0.94 |
| ... | ... | ... | ... | ... | ... |
| 49 | 25 593.7 | 18 490 | 24 249.3 | 14 994 | 1.06 |
| 31 | 4 409.6 | 1 427 | 4 169.5 | 925 | 1.06 |
| 80 | 3 400.0 | 608 | 3 204.6 | 434 | 1.06 |
| 07 | 160.8 | 16 | 151.0 | 13 | 1.06 |
| 23 | 3 538.6 | 1 108 | 3 286.8 | 632 | 1.08 |
| 78 | 8 880.5 | 2 380 | 8 157.4 | 1 810 | 1.09 |
| 32 | 3 106.5 | 2 273 | 2 827.7 | 1 194 | 1.10 |
| 29 | 5 328.9 | 706 | 4 835.0 | 545 | 1.10 |
| 27 | 4 031.6 | 695 | 3 638.0 | 554 | 1.11 |
| 20 | 3 941.7 | 516 | 3 556.2 | 402 | 1.11 |
| 17 | 3 367.2 | 324 | 3 030.8 | 264 | 1.11 |
| 36-37 | 692.1 | 186 | 618.4 | 149 | 1.12 |
| 61 | 1 702.3 | 467 | 1 520.7 | 309 | 1.12 |
| 72 | 4 922.0 | 1 857 | 4 351.5 | 1 151 | 1.13 |
| 33 | 4 939.6 | 3 112 | 4 315.3 | 1 978 | 1.14 |
| 11 | 360.9 | 85 | 310.7 | 62 | 1.16 |
| 03 | 178.1 | 754 | 144.7 | 244 | 1.23 |
| 35 | 5 322.1 | 822 | 3 965.8 | 582 | 1.34 |
| 12 | 186.3 | 11 | 82.4 | 8 | 2.26 |
| **All** | **570 888.0** | **486 741** | **567 001.1** | **286 003** | **1.01** |

**Conclusions:**

The systems approach (comparing many sources) is a very good way of finding errors

When you have found the errors, the work with correcting errors should follow

and that part is easier

Analyzing micro data and comparing different sources on the micro level is a difficult but interesting work – a new role for methodologists!