

Quality controls and checks of data in the Finnish LFS production system

7th Workshop on LFS methodology
Madrid, 10-11 May 2012

Kalle Sinivuori / Statistics Finland

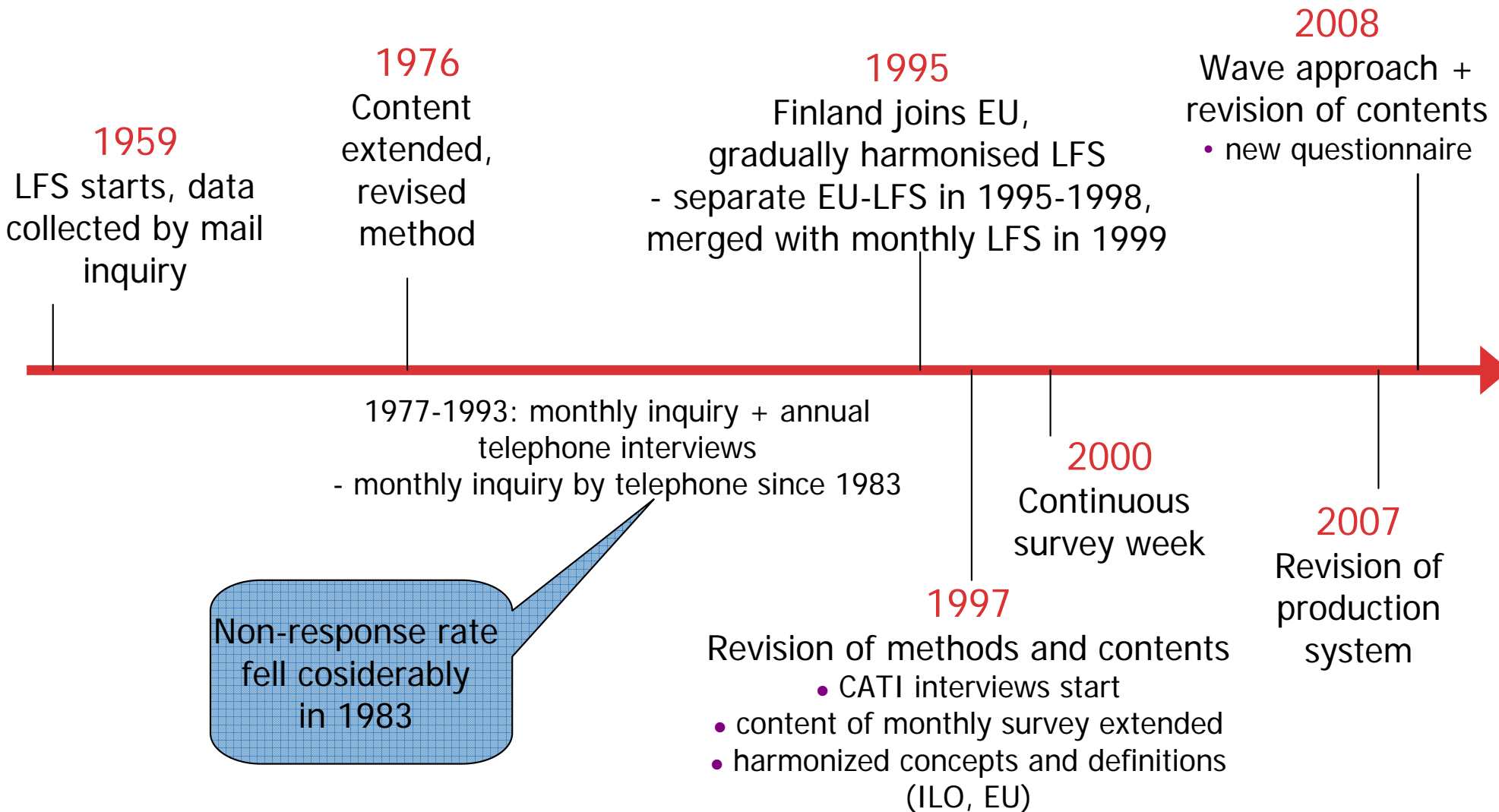
Content of presentation

- Background / Finnish LFS
- Production system from the user's perspective
- Check and controls in the LFS production process
- Evaluation

Finnish Labour Force Survey

- Started in 1959
- Based on individuals
- Data collection: CATI
- Sample size: 12 000 per month
- Publishing: monthly, quarterly, yearly.
- Major revisions:
 - 1995-97 EU-harmonisation; methods and contents
 - 2000 Continuous survey week
 - 2003-2007 Revision of production system
 - 2008 Wave approach + revision of contents

History of the LFS in Finland



'New' production system (from 2007)

- Moved from mainframe to open environment
- .NET application in the SQL database
- The aim for the new production system was that it would be reliable, transparent and managed by LFS experts.
- Easy to use application, with the following principle:

*Login => Selection of time period (year/month) =>
Selection of use case => Run! => Report on display =>
Acceptance/Rejection of report => Storing of data in
database => Next use case*

Muodosta korotuskertoimet
Vertaile korotuskertoimia

Työvoimatutkimus - Tulostusvaiheen sovellus

Sisältö

- [Pikaohje](#)
- [Tarkempia ohjeita](#)

Pikaohje

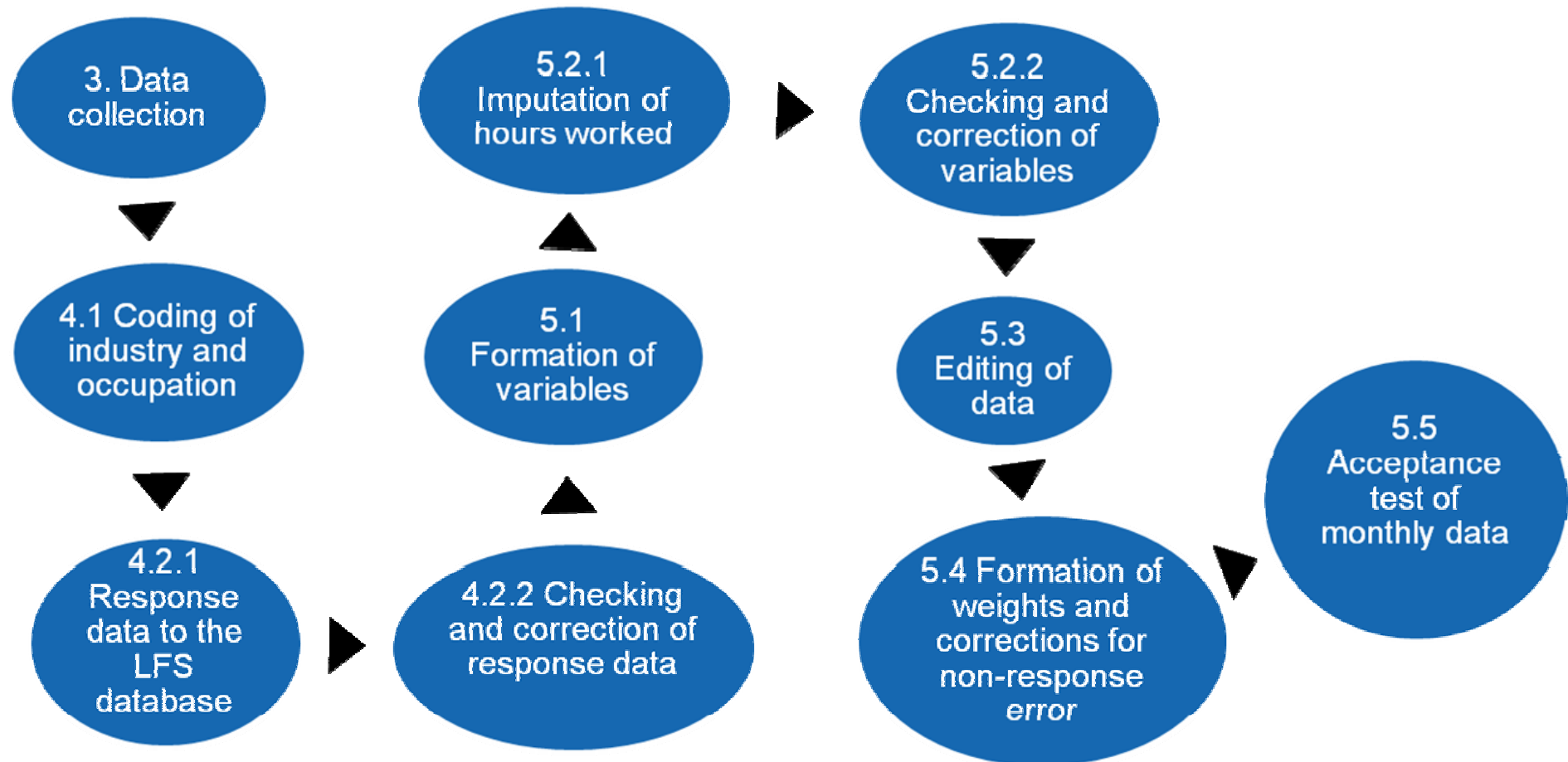
Käynnistyessään sovellus pyytää käyttäjää asettamaan tutkimusvuoden ja -kuukauden. Sen jälkeen ladataan tutkimus- ja rotaatioaikataulun sisältö muistiin näiden asetusten mukaisesti: mm. haetaan meneillään olevan jakson tehtävät ja niiden statustiedot kannasta.

Tietojensa perusteella ohjelma päättelee, mitä käyttäjä tässä vaiheessa saa/ei saa tehdä. Osa valikkojen toiminnoista saatetaan passivoida, eli voidaan käynnistää vain niitä toimintoja, joiden ajamiselle on edellytykset kunnossa.

Vastaustiedot pohjaksi tulostustiedoille: edellytykset eivät ole kunnossa.

Hyväksy Hylkää

Monthly production process



Processes which contain checkings in the Finnish LFS production system

■ Automatic checkings (+ corrections)

- Response data to the LFS database
- Imputation of hours worked

■ Manual checkings

- Coding of industry and occupation
- Checking and correction of response data
- Checking and correction of variables
- Acceptance test of monthly data

Automatic checkings and corrections/

4.2.1 Response data to the LFS database

- When data is moved from the interviewers' database to the LFS database, a set of automatic checks and corrections are made, such as:
 - Response data are formed or copied for disabled persons and conscripts
 - Education data are corrected for those aged 15 to 21 (education during the past four weeks)
 - If no responses to the first three questions, the respondent is moved to non-response

Automatic checkings and corrections/

5.2.1 Imputation of hours worked

- Unknown hours worked are imputed with the average data according to occupation and industry.
- Around 10 to 25 employees per month (less than 0.5% of all employees) => small effect on the total number of hours worked.
- User gets report on how many imputed values (plus same report from previous month as comparison) and on the basis of this information, he/she accepts or rejects this use case.

Manual checkings/

4.1 Coding of industry and occupation

- Is made with a separate application, the so-called coding application, that was taken into use a few years before the new LFS production system.
- The industry and occupation (+ socio-economic group and employer sector) are searched for all those interviewed for the first time and for those whose job has changed between the interview rounds
- Around 2,300 targets per month to be coded

Toimialan koodaus 2012/04

Työpaikan tiedot

nimi

Aakkosnimi Virallinen nimi Toiminimi

työpaikan numero

osoite

kunta

toimiala

työpaikan koko

TOL- koodi 2008 Oma Kyska

työnantajatyyppi Oma Kyska

Työntekijän tiedot

kohde ikä

sukupuoli Vuokratyöntekijä

tutkinto

työtehtävät

ammattiasema

ammatti

ammattikoodi vanha uusi

palkattua työvoimaa sose

maatalouden tuotantosuunta

Rajaukset

kunta

haastno

Perustiedot | Osoite- ja muut tiedot | Toimipaikat Rivejä haettu: 0

OYR-tunnus	Nimi	TOL08	TOL 08 nimike	TOL02	TOL 02 nimike	TmpLk	Kotiki	Kunnan nimi

Manual checkings/

4.2.2 Checking and correction of response data

- Value range checks and few logical checks
 - Dates (typing errors)
 - Relations.
For example: check and correction if workdays + sick days > 7
- Tool: Editor that brings all response data for the target to be checked on display => all corrections straight to the original data.
- Around 10-20 corrections per month.

Tietuenumero 4

Kentän nimi

Raportti **Korjaus**

kohdeno	04330062	at1		at2		at3		at4		at5	
ee1	3	ee10		ee11		ee12		ee13		ee14	
ee15		ee16		ee1b		ee2	2	ee21	2	ee3a	
ee3b	2	ee4		ee5		ee6		ee7		ee8	
ee901		ee902		ee903		ee904		ee905		ee906	
ee907		ee908		ee909		ee910		ee911		ee912	
ehd		ehdo		h1koodi		h1nimi		h2b		h3	2
h4		h5		h5jatk		h7		haaspvm	20120319	haastno	7751
haasvk	12	hkesto	5	huomio		ikanyt	55	ka1	2	ka10	
ka11		ka12		ka14t		ka15e	4	ka15t		ka2	2
ka3	2	ka4		ka5		ka6		ka7		ka8a	
ka8b		ka9		kl1	6	kl4		kr1a		kr1b	2
kr2		kr3		kr4		kr51		kr52		l2	
loptulo	01	m10	Rantamäki 3	m11koodi		m11nimi		m12	HOTELLI	m13	
m142		m14koodi	5246	m14nimi	TARJOILIJJA	m1	1	m15	1	m2	2009
m2b		m2c		m6	1	m7		m7p	2	m8	HOTELLI
m9koodi	049	m9nimi	ESPOO	p81		p82		perust		r12	2
r1ety	2	r1ty		r2ety		r2ty		s1		s10	
s2		s2p		s3		s4koodi		s4nimi		s5	
s6koodi		s6nimi		s7		s8		syykesk		t1	
t10		t11		t122		t12koodi		t12nimi		t12p	
t13		t2		t3u		t4		t5		t5p	
t6		t7koodi		t7nimi		t8		t9koodi		t9nimi	
ta1		ta10		ta11		ta12		ta13		ta14	
ta15		ta16		ta17		ta18		ta19		ta2	
ta20		ta21		ta22		ta23		ta24		ta3	
ta4		ta5		ta6		ta7a		ta7b		ta8	
ta9		tapa	2	te1		te2		te3		te401	
te402		te403		te404		te405		te406		te407	
te408		te409		te410		te411		te412		te5	
te6		te7		te8		te9		tep0		tep1	
tep2		tep3		tep4		tep5		tep6		tieant	1
tila		ts1		ts2		ts3		ts3y		ts4	
ts5		ts6		tutkviik	13	r12b					

Manual checkings/

5.2.2 Checking and correction of variables

- A similar correction process is made to the variables as to the response data.
- Logical relations between two variables are checked
 - for example, employer type with respect to occupational status: a self-employed person cannot have public sector as the employer type
- ✓ Two separate processes: Checking of national and checking of EU-variables
 - ✓ Risk of contradiction between national and EU-variables
 - ✓ No major changes made at this point

Manual checkings/

5.5. Acceptance test of monthly data

- Checking of publication tables
- The last test before acceptance
- If the figures seems to be in order
 - ⇒ accepting monthly data
 - ⇒ copying the data in the tabulation database.

Other controls (of quality) in the Finnish LFS production system

- Response data to the LFS Database
 - ⇒ Checking the amount of accepted answers and comparing it to the previous month.
 - ⇒ Preliminary distribution on employment/ unemployment and comparing it to the year before.

Response data taken to LFS database

Year 2012 month 3

Date	N
12.4.2012 11:19	2222
12.4.2012 11:24	2116
12.4.2012 15:28	2128
12.4.2012 15:32	2119
12.4.2012 15:37	2120

1:Employees	6370
2: Self-employed	878
3: Unpaid family workers	30
9: EOS	0
Total	7278
Missing	0

Employed	5286
Unemployed	453
Total	5739

1 Accepted answer	9036
2.Refusal	273
3.Sick/ unable to work/ answer	19
4: No-contact	1335
5. Language problems, etc	14
6. Died	9
7. Abroad	19
8: Other overlap	0
9.Unknown	0
TOTAL	10705

Response data taken to LFS database

Year 2012 month 2

Date	N
8.3.2012 14:52	2755
8.3.2012 14:58	2629
8.3.2012 15:34	2632
8.3.2012 15:47	2637

1:Employees	6446
2: Self-employed	920
3: Unpaid family workers	31
9: EOS	0
Total	7397
Missing	0

Employed	5473
Unemployed	405
Total	5878

1 Accepted answer	9130
2.Refusal	256
3.Sick/ unable to work/ answer	13
4: No-contact	1215
5. Language problems, etc.	20
6. Died	2
7. Abroad	17
8: Other overlap	0
9.Unknown	0
TOTAL	10653

Other controls (of quality) in the Finnish LFS production system

- Formation of variables
 - Frequencies on (ILO-) unemployment/ employment and main status.

- Editing of data
 - Increasing days worked and hours worked to the monthly level
 - Editing of preliminary population figures
 - Handling of the jobseeker register

Evaluation

- Improvements comparing to the old production system
 - All process stages go through an LFS-expert
 - Report storage. All the reports from monthly product process are easy to find in html-format.
 - Task management system. No chance to miss/ forget any stages of the process.

Evaluation...

- Still to improve
 - Dependency of IT-support during the production process
 - Timing of checks. All important checks should be done as early as possible (to the response data)
 - Manually corrected variables are not "flagged" => unnecessary double-entry bookkeeping (with excel)
 - Yearly maintenance
=> Under construction, during 2012-2013.

Thank You !!!

kalle.sinivuori@stat.fi