

# Data processing during and after the interview

Cor Kragt



Statistics Netherlands

## Structure

- Survey design Dutch LFS
- New process
- Data processing
  - during the interview
  - after the interview

## Survey Design Dutch LFS

- Panel: 5 waves
- 1 Year, every 3 months

	1999-2010	2010-2012	2012 onwards
Wave			
1	CAPI	CAPI & CATI	CAWI & CAPI & CATI
2	CATI	CATI	CATI
3	CATI	CATI	CATI
4	CATI	CATI	CATI
5	CATI	CATI	CATI

## New process

### Reasons

- Redesign social statistics
- Redundant software
- Dependent production lines
- New weighting model

## Data processing

During the interview

- Record per household

After the interview

- File per month all household responses

## Data processing during the interview

- Blaise questionnaire
- First wave:
  - empty records
  - all questions possible
- Consecutive waves:
  - data previous wave
  - only mutations

# Data processing during the interview

## Use of previous data

- Jobs
- Working hours
- Searching for work
- Occupation
- Economic activity
- Education

# Data processing during the interview

## Wave 1

### *UrDnst*

How many hours does your contract holds? 36

### *Periode*

Is this per week, month or year? per week



# Data processing during the interview

## Wave 2

### *T\_UrDnst*

Previous time you worked <36> hours <per week>.

Is this still the case?

Yes

**IF** WaveNr > 1 AND T\_UrDnst = Yes **THEN**

UrDnst := Prev\_UrDnst (**:=** 36)

Periode := Prev\_Periode (**:=** per week)

**ELSE**

UrDnst.ASK

Periode.ASK

**ENDIF**

## Data processing after the interview

- Automated
- Manipula (Blaise)
- Succession of steps
  - collecting the response records
  - conversion to a unique data model
  - derivation of variables
  - matching additional data (previous wave, registers, codes)
  - weighting
  - tabulation

S_PersID	H_M_V	Afl_Lft	W_Betwerk	W_UrDnst[1]	G_MaandGewicht	T_SOI_2006_HB	Afl_BB_Ned
3371053	Man	22	Ja	20			
3371054	Man	22	Ja	12			
3371055	Man	22	Nee				
3371057	Man	20	Nee				
3373070	Vrouw	45	Ja				
3373464	Vrouw	47	Nee				
3373465	Vrouw	14					
3373892	Man	54	Ja	28			
3374198	Vrouw	48	Ja	36			
3376012	Vrouw	32	Ja	16			

S_PersID	H_M_V	Afl_Lft	W_Betwerk	W_UrDnst[1]	G_MaandGewicht	T_SOI_2006_HB	Afl_BB_Ned
3371053	Man	22	Ja	20			
3371054	Man	22	Ja	12			
3371055	Man	22	Nee				
3371057	Man	20	Nee				
3373070	Vrouw	45	Ja				
3373464	Vrouw	47	Nee				
3373465	Vrouw	14					
3373892	Man	54	Ja	28			
3374198	Vrouw	48	Ja	36			
3376012	Vrouw	32	Ja	16			



S_PersID	H_M_V	Afl_Lft	W_Betwerk	W_UrDnst[1]	G_MaandGewicht	T_SOI_2006_HB	Afl_BB_Ned
3371053	Man	22	Ja	20	3862,015138	603311	werkzBB
3371054	Man	22	Ja	12	3862,015138	430128	werkzBB
3371055	Man	22	Nee		3862,015138	533311	niet_BB
3371057	Man	20	Nee		3862,015138	430120	werk1BB
3373070	Vrouw	45	Ja		4304,419166	604111	niet_BB
3373464	Vrouw	47	Nee		1313,494060	336413	niet_BB
3373465	Vrouw	14			1313,494060		nvt
3373892	Man	54	Ja	28	4270,986858	436322	werkzBB
3374198	Vrouw	48	Ja	36	5677,921708	413511	werkzBB
3376012	Vrouw	32	Ja	16	2636,033826	419114	werkzBB

RUNJC ""\SRSH1F\SAH2\Onderzoek\EBB\Productie\JC\PRODUCTIE\_EBB.jcp"

Job Parameters

Nummer Eerste job: 1

Nummer laatste job : 32

Start Publicatieperiode (EEJJMM): 201204

Eind Publicatieperiode (EEJJMM): 201204

Start Time: 17:29 uur, woensdag 18 apr 2012

Repeat Job: ☐ ma di wo do vr za zo

After Job: None

Interval: 24 uur

Until: 17:29 uur

Time Limit: 6 uur

Onexit Limit: 5 min.

Job Name: PRODUCTIE\_EBB

Job Group: I - default

Status Dir: \\SRSH1F\SAH2\Onderzoek\EBB\Productie\Monitor

Batch Queue: CBS

Config.File: \\SRSH1F\SAH2\Onderzoek\EBB\Productie\RCMD.DEF

Giga Bytes: 10

Output File: LIST

Parallel Jobs: No

History File:

E-Mail Msg: Errors

E-Mail User(s): SAHCKGT

Oper.System: WinNT

Terminal Msg: Always

Log Commands: ☐

Delete Status File: ☐

Eind Publicatieperiode (EEJJMM): { integer } waarde 201204

Start Job Annuleren Zoek Job... Bewerken... Help

RUNJC ""\\SRSH1F\\SAH2\\Onderzoek\\EBB\\Productie\\JC\\PRODUCTIE\_EBB.jcp"

Job Parameters

Start Time: 17:29 uur, woensdag 18 apr 2012 Repeat Job: ☐ ma di wo do vr za zo

After Job: None Interval: 24 uur Until: 17:29 uur

Time Limit: 6 uur Onexit Limit: 5 min. Job Name: PRODUCTIE\_EBB

Job Group: I - default Status Dir: \\SRSH1F\\SAH2\\Onderzoek\\EBB\\Productie\\Monitor

Batch Queue: CBS Config.File: \\SRSH1F\\SAH2\\Onderzoek\\EBB\\Productie\\RCMD.DEF

Giga Bytes: 10 Output File: LIST

Parallel Jobs: No History File:

E-Mail Msg: Errors E-Mail User(s): SAHCKGT

Oper.System: WinNT Terminal Msg: Always Log Commands: ☐ Delete Status File: ☐

Eind Publicatieperiode (EEJMM): { integer } waarde 201204

Start Job Annuleren Zoek Job... Bewerken... Help

RUNJC ""\\SRSH1F\\SAH2\\Onderzoek\\EBB\\Productie\\JC\\PRODUCTIE\_EBB.jcp"

Job Parameters

Start Time: 17:29 uur, woensdag 18 apr 2012 Repeat Job: ☐ ma di wo do vr za zo

After Job: None Interval: 24 uur Until: 17:29 uur

Time Limit: 6 uur Onexit Limit: 5 min. Job Name: PRODUCTIE\_EBB

Job Group: I - default Status Dir: \\SRSH1F\\SAH2\\Onderzoek\\EBB\\Productie\\Monitor

Batch Queue: CBS Config.File: \\SRSH1F\\SAH2\\Onderzoek\\EBB\\Productie\\RCMD.DEF

Giga Bytes: 10 Output File: LIST

Parallel Jobs: No History File:

Oper.System: WinNT E-Mail Msg: Errors E-Mail User(s): SAHCKGT

Terminal Msg: Always Log Commands: ☐ Delete Status File: ☐

Eind Publicatieperiode (EEJMM): { integer } waarde 201204

Start Job Annuleren Zoek Job... Bewerken... Help

## Data processing after the interview

Closer look to two steps

1. conversion to a unique data model
2. matching information from the previous wave



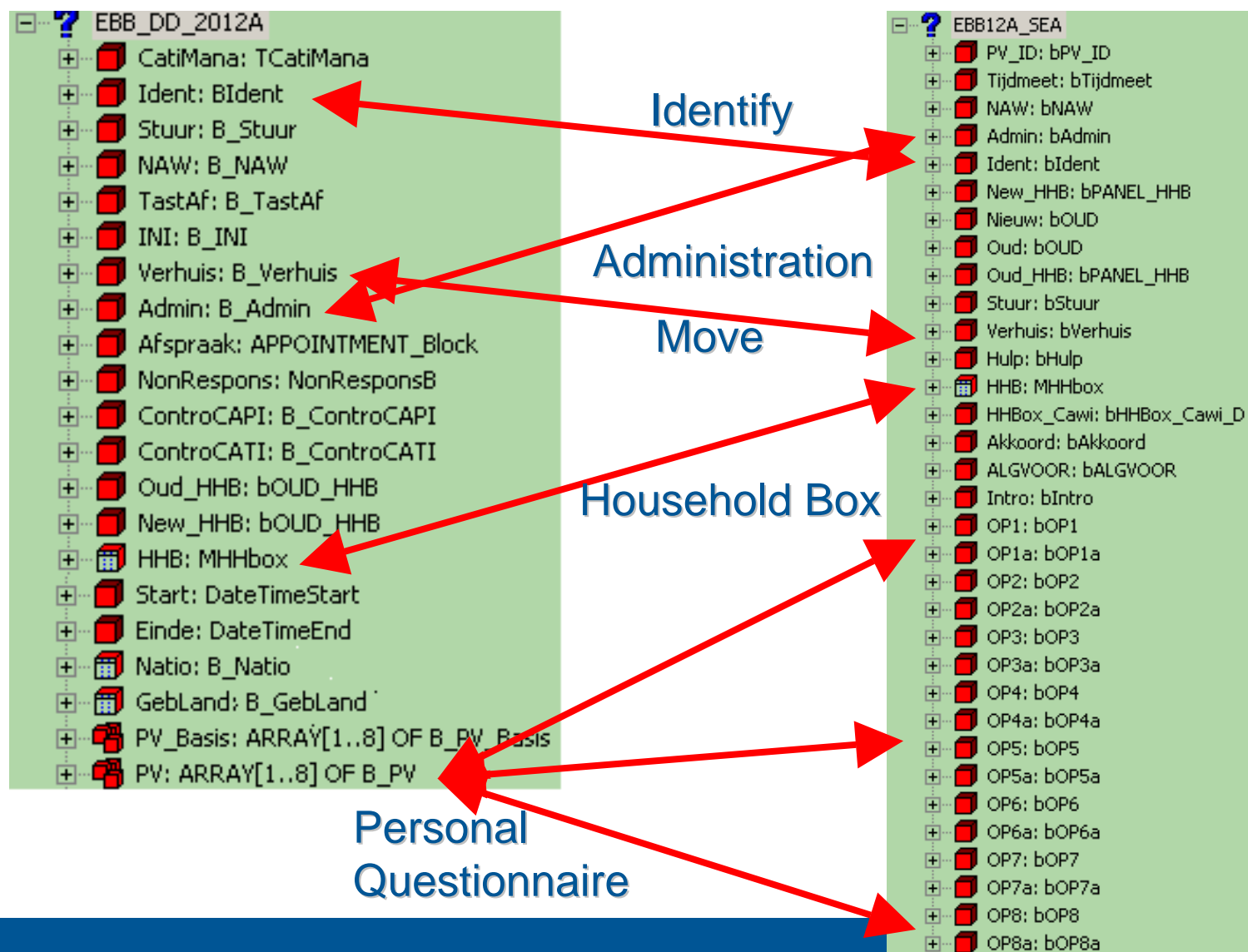
# 1. Conversion to a unique data model

## Pros

- Delete unnecessary variables
- Privacy
- Manipulation independent of mode

# Capi/Cati

# Cawi



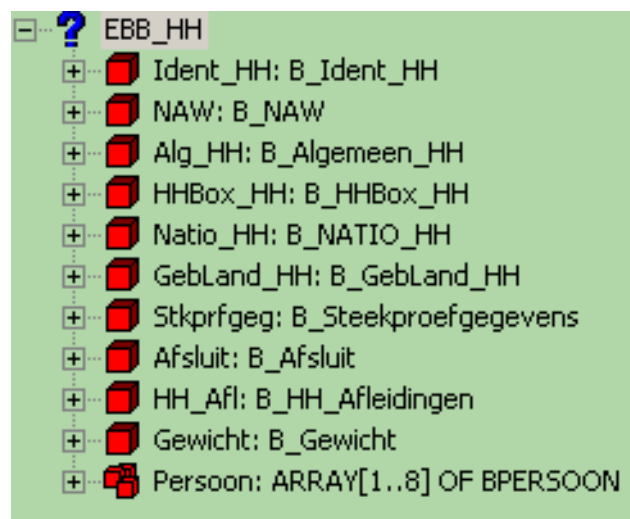
# Capi/Cati



# Cawi



Unique  
model



## 2. Matching information from the previous wave

- Retro
- Followed education
- Sample information
- Nationality
- Land of birth

## 2. Matching information from the previous wave

Match to file 3 months ago, but:

- Possibility to skip one wave
- Difference between statistical period and survey period

So:

- Match to file with records of last 8 months





## Additional slides



### 3. Matching codes for occupation, economic activity and education

- Coding is a separate process
- Available 3 weeks after survey period
- Unemployment figures
- First run without these codes
- Second run with codes

### 3. Matching codes for occupation, economic activity and education

- Only codes for new observed occupations, economic activities or educations are delivered
- Received codes => match these codes
- No code => same job, occupation or education?

### 3. Matching codes for occupation, economic activity and education

- ***T\_ZelfWrk***

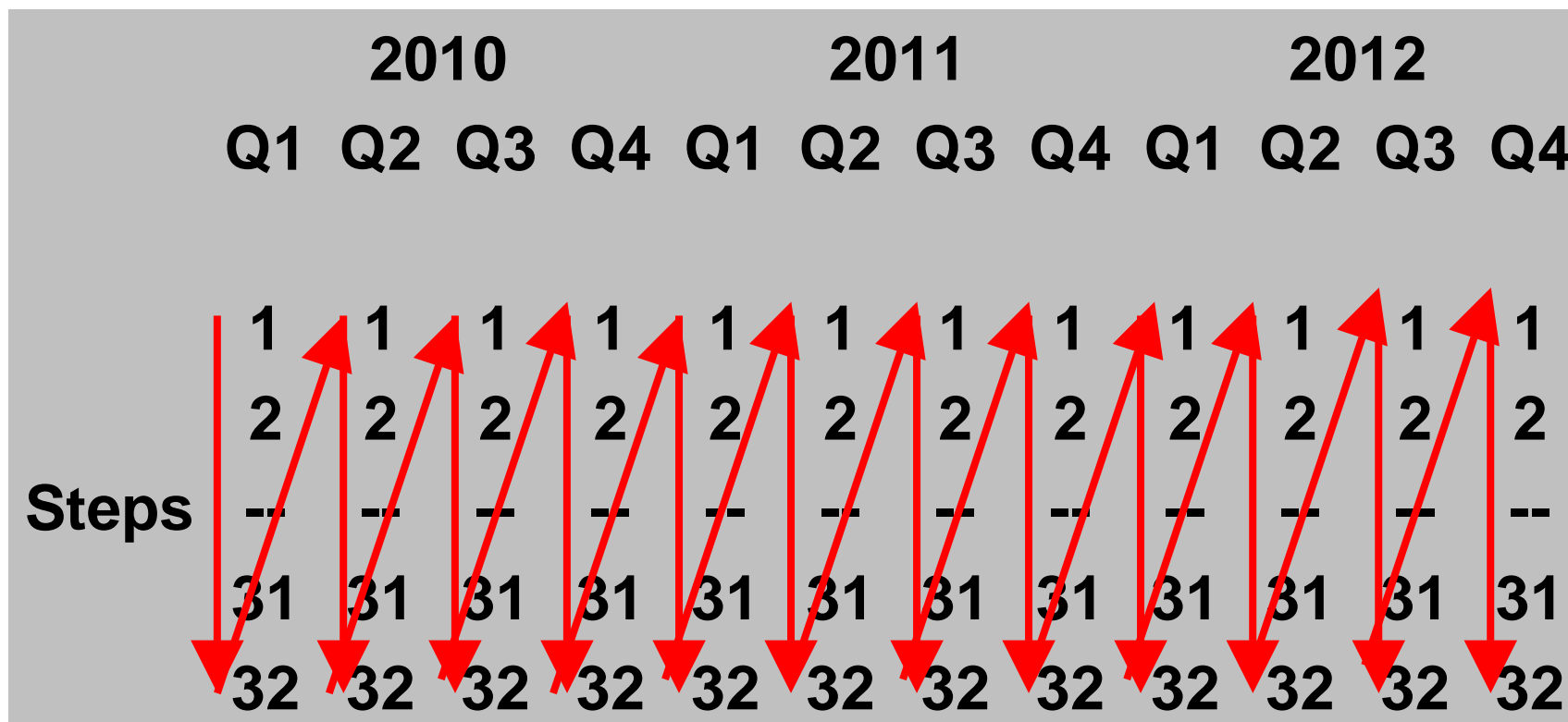
Previous time you worked <workname>. Is this still the case?

- ***T\_Beroep***

Do you have the same profession or function as the previous time?

- ***T\_Opleiding***

Previous time you were studying. Is this still the case?



# Survey Design Dutch LFS since 2010

Wave	Mode	2011					2012		
		Q1	Q2	Q3	Q4		Q2	Q3	Q4
1	CAPI/CATI	E1	F1	G1	H1		J1	K1	L1
2	CATI	D2	E2	F2	G2		I2	J2	K2
3	CATI	C3	D3	E3	F3		H3	I3	J3
4	CATI	B4	C4	D4	E4		G4	H4	I4
5	CATI	A5	B5	C5	D5		F5	G5	H5

## 2. Matching information from the previous wave

S_PersID	S_Peiling_nummer	CL_NatCode	CL_GebKode
3516284	2 0001	6030	

S_PersID	S_Peiling_nummer	CL_NatCode	CL_GebKode
3516284	2 0001	6030	
3516285	4 0001	6030	

S_PersID	S_Peiling_nummer	CL_NatCode	CL_GebKode
3516284	2 0001	6030	
3516285	4 0001	6030	

S_PersID	S_Peiling_nummer	CL_NatCode	CL_GebKode
3516284	2 0001	6030	
3516285	4 0001	6030	
3516286	4 0001	6030	
3516287	4 0001	6030	
3516288	4 0001	6030	
3516289	4 0001	6030	
3516290	4 0001	6030	
3516293	3 0339	6043	

Combine 8  
files to 1

S_PersID	S_Peiling_nummer	CL_NatCode	CL_GebKode
3516284	2 0001	6030	
3516285	4 0001	6030	
3516286	4 0001	6030	
3516287	4 0001	6030	
3516288	4 0001	6030	
3516289	4 0001	6030	
3516290	4 0001	6030	
3516293	3 0339	6043	