

# Paradata: Het gebruik van audit trails

Mattijn Morren  
Ger Snijkers

Centraal Bureau voor de Statistiek, Heerlen, Nederland  
Universiteit Utrecht



## Audit trails

### 1. Wat zijn audit trails?

- Een log van hoe respondenten / interviewers stap-voor-stap een elektronische vragenlijst invullen
- Paradata op micro niveau

### 2. Wat kun je ermee?

- Geeft inzicht in het invulgedrag/invulproces  
De respons is het resultaat van dit proces
- Verbeteren van de vragenlijst
- Verbeteren van de benaderingsstrategie

### 3. Een aantal analyses met audit trails van de elektronische vragenlijst voor de Productiestatistiek



11 maart 2010

NPSO, Brussel, België

**CBS Enquête Productiestatistiek Algemene burgerlijke en utiliteitsbouw 2005**

- ✓ Inleiding
- ✓ Invulinstructie
- ✓ Contactpersoon
- ✓ Boekjaar
- Werkzame personen
  - Werkzame personen
  - Specificatie werkgelegenheid
  - Bedrijfsopbrengsten
  - Bedrijfslasten
  - Resultaten
  - Afsluiten en Verzenden

**Medewerkende eigenaren en gezinsleden** X  
 Betreft niet in loondienst zijnde medewerkende eigenaren en/of medewerkende familieleden, firmanten en vennoten.

**Versturen voor:**  
11 mei 2006

**Correspondentienummer:**  
4508 0001

**Werkzame personen**

**Werknemers op de eigen loonlijst per 30 september**

Totaal aantal

Omgerekend in voltijdequivalenten (VTE, op één decimaal nauwkeurig)

**Overige werkzame personen per 30 september**

Medewerkende eigenaren en gezinsleden

Uitzendkrachten / gedetacheerd personeel (van uitzendbedrijven)

Overig ingeleend personeel (van andere bedrijven)

**Subtotaal aantal werknemers en overige werkzame personen**

Aantal van het Totaal aantal opgegeven personeel dat was uitgeleend (aan andere bedrijven)

**Totaal aantal personen werkzaam in uw bedrijf**

**Totaal aantal personen werkzaam in uw bedrijf omgerekend in voltijdequivalenten (VTE, op één decimaal nauwkeurig)**

Print

Akkoord +

Centraal Bureau voor de Statistiek

Verzenden Afdrukken Opslaan Stoppen Calculator Rekenveld Informatie

## Audit trail fragment

	id	date	time	action	field	value
1	12345	22-MAY-2007	10:43:17	0	0	
2	12345	22-MAY-2007	10:43:58	4	40	
3	12345				40	01-04-2005
4	12345				40	
5	12345				27	
6	12345				21	
7	12345	22-MAY-2007	10:44:51	1	48	
8	12345	22-MAY-2007	10:45:15	3	48	
9	12345	22-MAY-2007	10:50:20	8	48	
10	12345	22-MAY-2007	10:51:05	10	48	
11	12345	22-MAY-2007	10:51:05	99	0	
12	12345	20-JUN-2007	8:52:28	0	0	
13	12345	20-JUN-2007	8:53:09	8	15	
14	12345	20-JUN-2007	8:54:50	13	15	
15	12345	20-JUN-2007	8:54:53	11	15	
16	12345	20-JUN-2007	8:55:02	9	15	
17	12345	20-JUN-2007	9:35:58	99	0	
18	12345	29-JUN-2007	15:10:38	0	0	

- 11,5 miljoen records
- 41.274 respondenten

**Actiecodes:**

- 0 = Start programma
- 1 = In veld gaan
- 2 = Veld verlaten
- 3 = Klik op 'help'
- 7 = Verzenden
- 8 = Afdrukken
- 9 = Opslaan
- 10 = Stoppen
- 11 = Calculator
- 13 = Informatie
- 14 = Klik index niveau 1
- 99 = Afsluiten progr.

## Audit trails

- A. Wat doen respondenten stap-voor-stap?
- B. Wanneer werken respondenten aan de vragenlijst?



11 maart 2010

NPSO, Brussel, België

## Audit trails – invulgedrag

- A. Wat doen respondenten stap-voor-stap:
  - In één keer invullen of in meerdere sessies?
  - Hoe lang is men ermee bezig?
  - Bekijkt men eerst de vragenlijst of begint men meteen met invullen?
  - Gebruikt men de print-functie? Hoe?

### Invulgedrag:

- Invulstatistieken: invulmaten, navigatiematen
- Invulgedrag naar achtergrondkenmerken
- Invulprofielen



11 maart 2010

NPSO, Brussel, België

## Invulmaten

	Gem	SD	Min	Max
<b>Sessies (n = 67.479)</b>				
Aantal acties	165	156	2	2209
Sessieduur	0:41:20	0:59:27	0:00:01	10:50:15
<b>Respondent (n = 41.274)</b>				
Aantal acties	270	210	42	2209
Aantal sessies	1,6	1,4	1	35
Invultijd	1:07:34	1:58:11	0:00:20	59:05:57



11 maart 2010

NPSO, Brussel, België

## Navigatiematen

Navigatie	% v. R gebruikt	Gem. Keer gebruikt	SD	Min-max
Start programma	100	1,6	1,4	1 - 32
Help	45	5,3	6,2	1 - 82
Verzenden	100	1,3	0,7	1 - 13
Printen	43	1,7	1,3	1 - 27
Opslaan	29	2,6	4,2	1 - 104
Rekenmachine	5	1,9	2,4	1 - 38
Informatie	13	1,4	1,5	1 - 97
Volgende	46	13,5	20,5	1 - 346
Vorige	65	14,8	19,8	1 - 325
Bevestigen	100	33,1	18,6	12 - 316

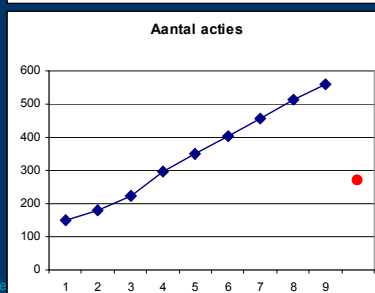
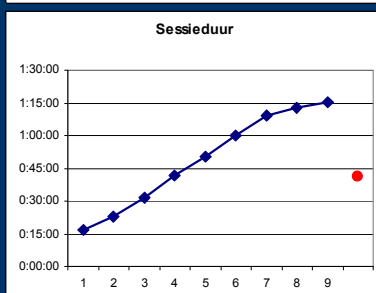
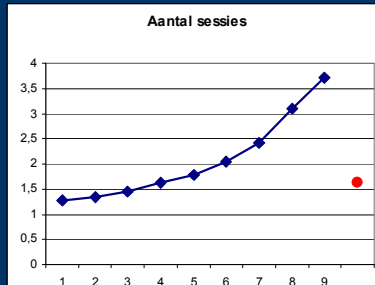
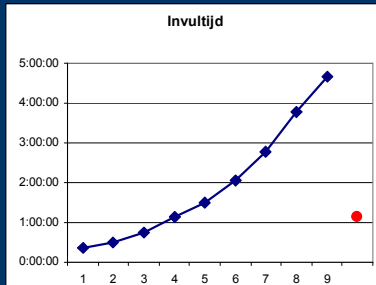
N = 41.274 respondenten



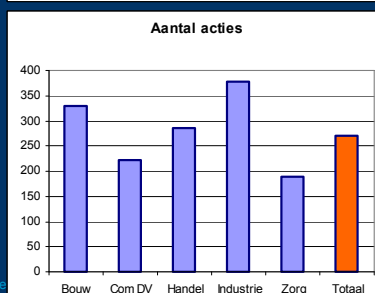
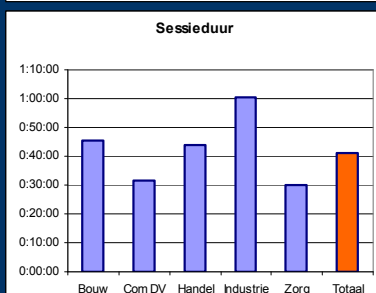
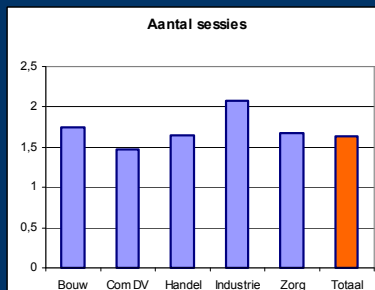
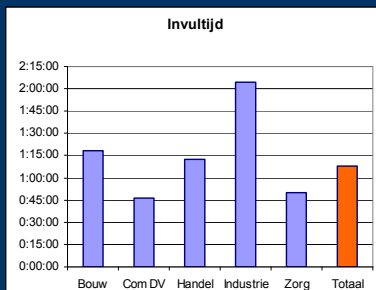
11 maart 2010

NPSO, Brussel, België

## Invulmaten naar grootteklasse



## Invulmaten naar branche



## Audit trails – invulgedrag

### A. Wat doen respondenten stap-voor-stap:

- In één keer invullen of in meerdere sessies?
- Hoe lang is men ermee bezig?
- Bekijkt men eerst de vragenlijst of begint men meteen met invullen?
- Gebruikt men de print-functie? Hoe?

### Invulgedrag:

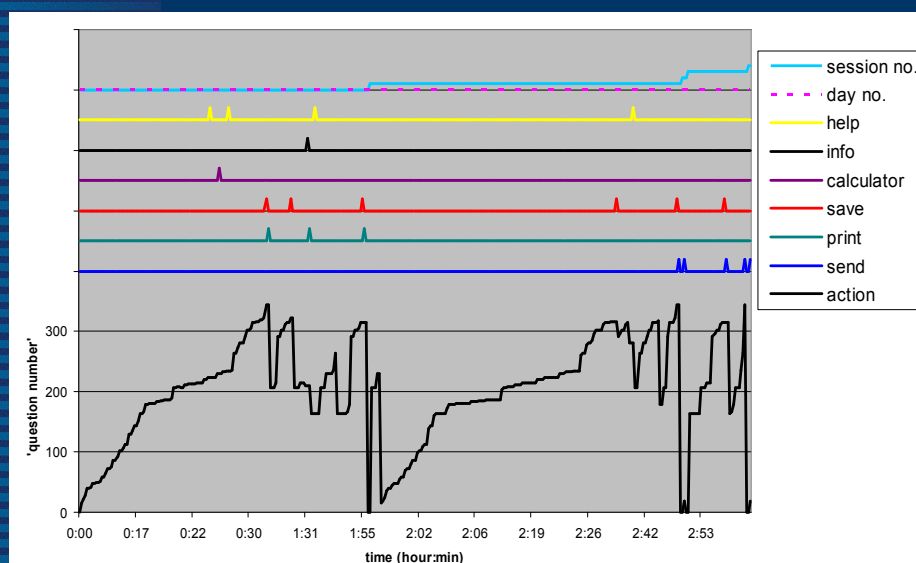
- Invulstatistieken: invulmaten, navigatiematen
- Invulgedrag naar achtergrondkenmerken
- **Invulprofielen**



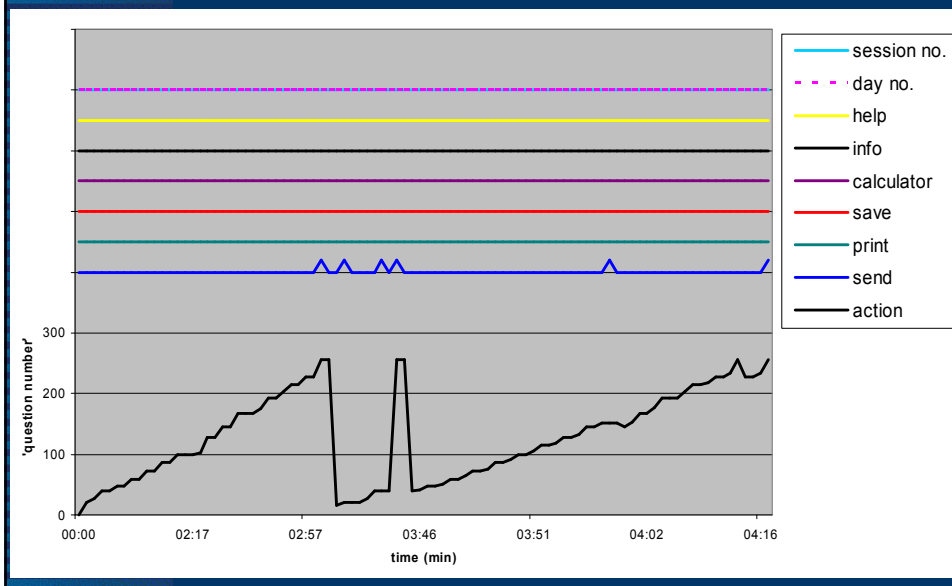
11 maart 2010

NPSO, Brussel, België

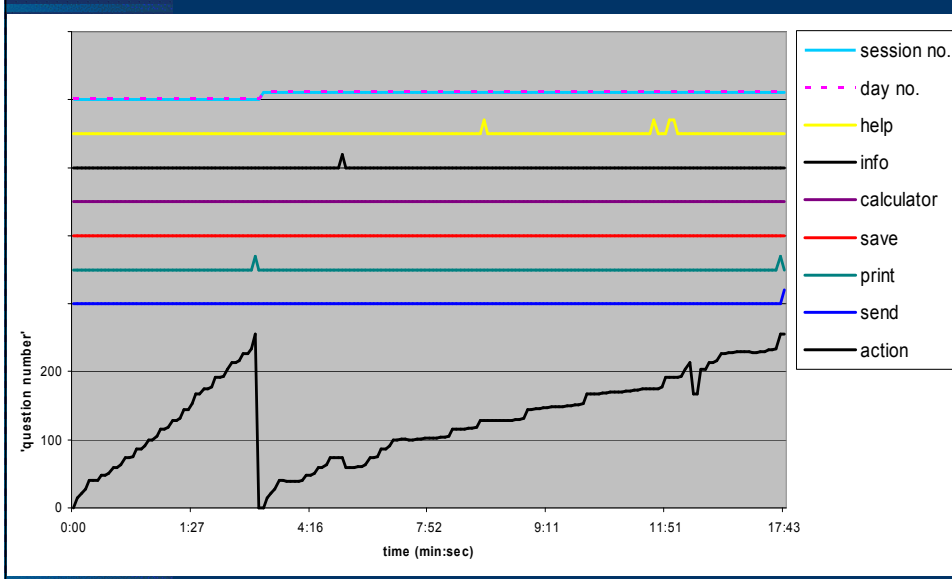
## Voorbeeld profiel: Zorgvuldig



## Voorbeeld profiel: "Quick 'n' dirty"



## Voorbeeld profiel: printen



## Audit trails – invulverloop

A. Wat doen respondenten stap-voor-stap

B. Wanneer werken respondenten aan de vragenlijst:

- Als ze vragenlijst ontvangen, of als ze de gegevens hebben?
- Werken rappelbrieven?

Invulverloop:

1. Invulincidentie: wanneer beginnen Rs aan de V?
2. Invulprevalentie: wanneer werken ze er aan?

- Aanleiding: responsverloop



11 maart 2010

NPSO, Brussel, België

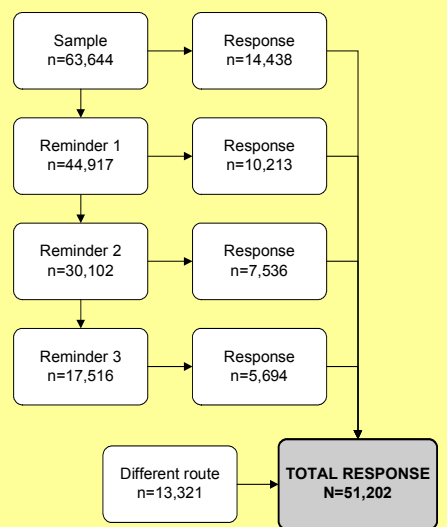
## PS-2006: Response analyse

• 100%

• >2/3

• <1/2

• >1/4



• Verzending in 7 clusters (april/Mei)

• 9-13 weken later (juni/aug)

• 2-6 weken later (juni/aug)

• 3-13 weeks later (sept/okt)

• 80% respons  
 - 80% e-Q (42.902)  
 - Gem. 3,5 mnd responstijd

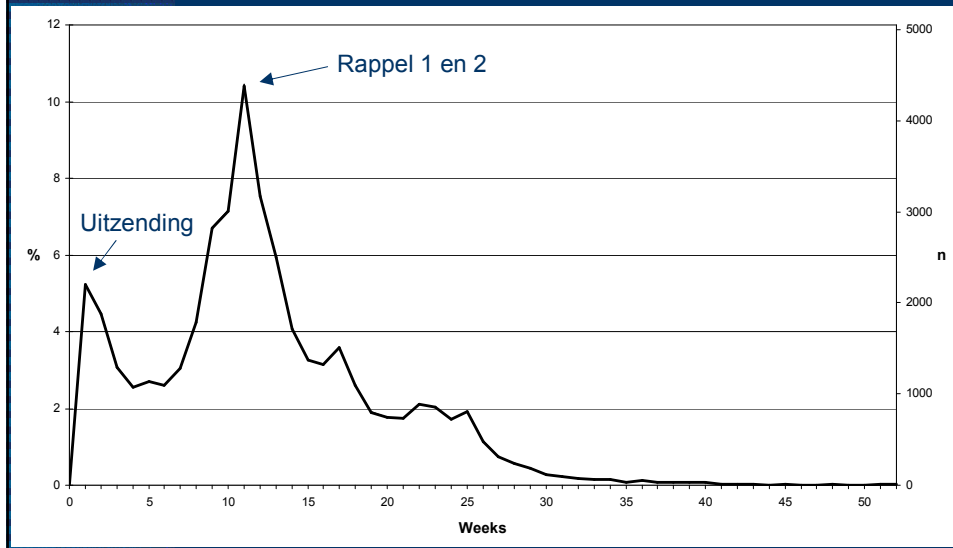


11 maart 2010

NPSO, Brussel, België



## Responsverloop na uitzending



## Onderzoeksvragen

1. Wanneer vullen bedrijven de vragenlijst in?
  - als ze de gegevens en de middelen hebben
  - als het CBS met hen in contact treedt (rappel)
  - Responsanalyse: wanneer gegevens ingestuurd
  - Audit trails: wanneer Rs de vragenlijst invullen
2. Invloed van rappels op respons- en invulverloop
3. Rappels als deadline of motivatie om de PS in te vullen en te responderen?



## Respons- en invulverloop

### Responsverloop:

De respons verkregen in achtereenvolgende tijdperiodes na het uitzenden van de vragenlijst.

### Invulverloop:

De tijd en inspanning besteedt aan het invullen van de vragenlijst in deze tijdperiodes.



11 maart 2010

NPSO, Brussel, België

## Maten van invulverloop

### 1. Invulincidentie

Het percentage respondenten dat in elke tijdperiode (van bijv. dagen of weken) *voor het eerst* aan het invullen van de vragenlijst heeft gewerkt

### 2. Invulprevalence

Het percentage respondenten dat in elke achtereenvolgende tijdperiode aan het invullen van de vragenlijst heeft gewerkt

### 3. Invultijd

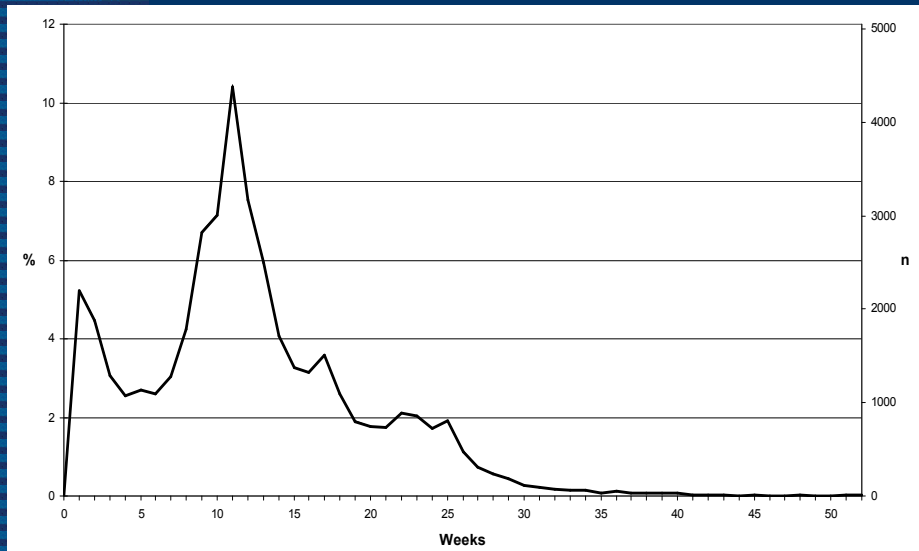
De tijd die respondenten in elke achtereenvolgende periode besteden aan het invullen van de vragenlijst



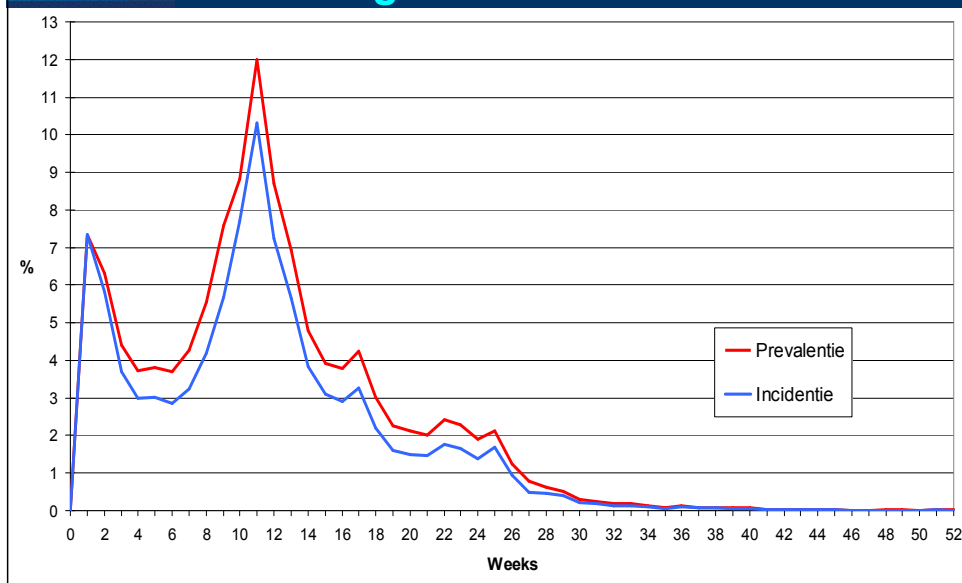
11 maart 2010

NPSO, Brussel, België

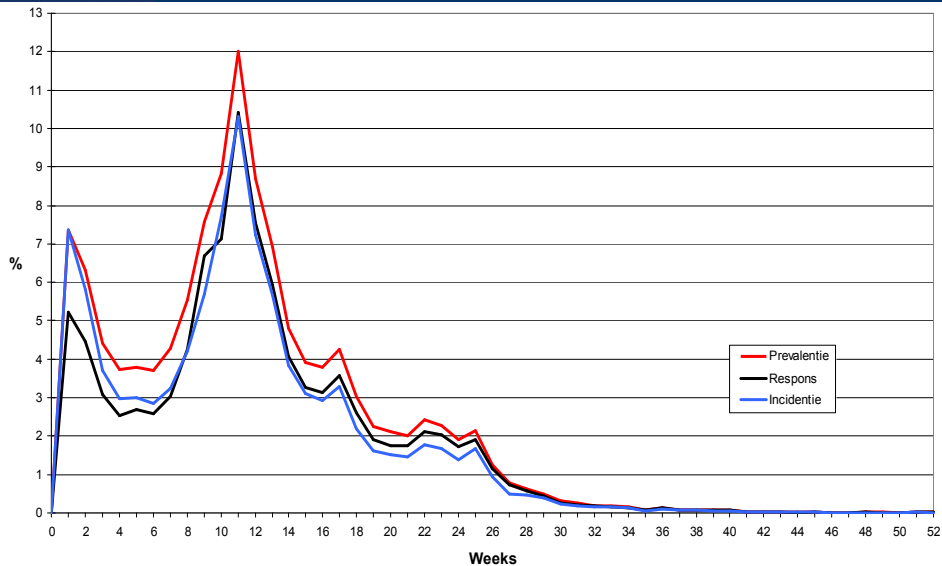
## Respons na uitzending



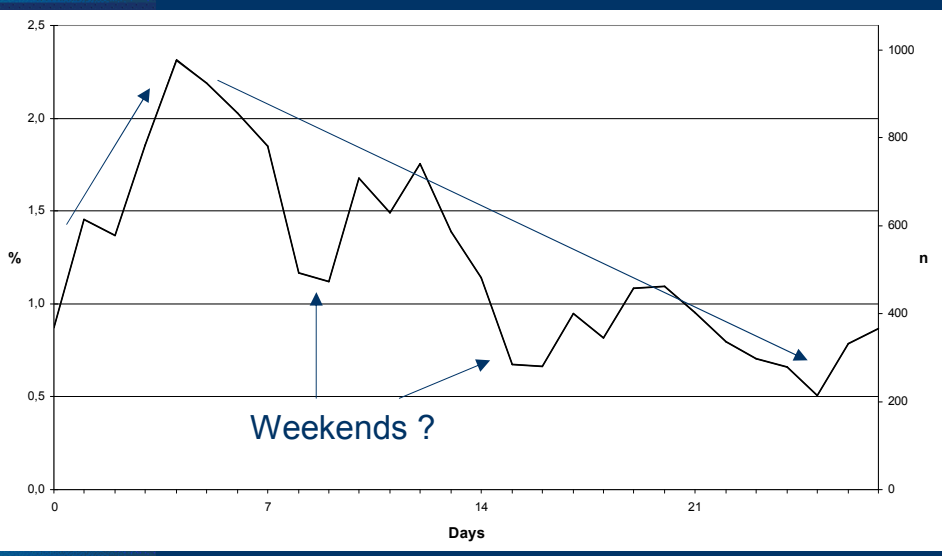
## Incidentie en prevalentie na uitzending



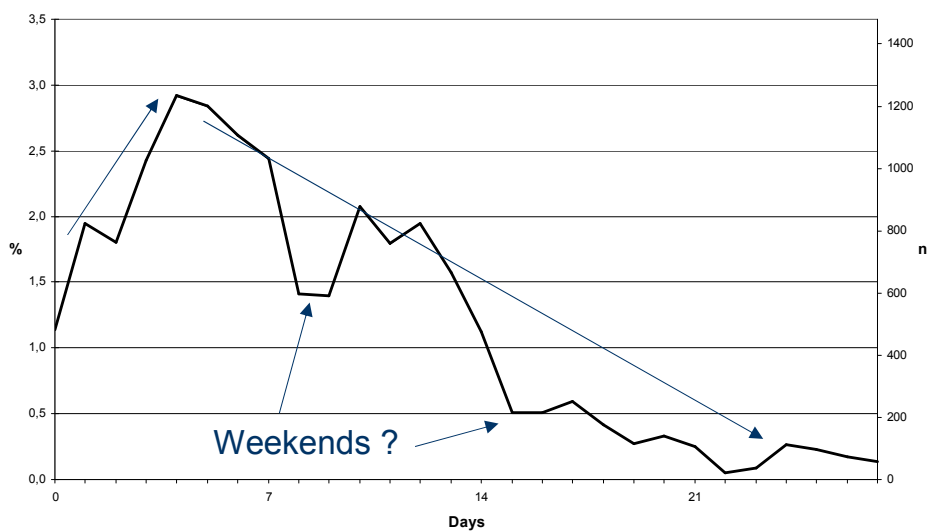
## Respons, invulincidentie en invulprevalentie na uitzending



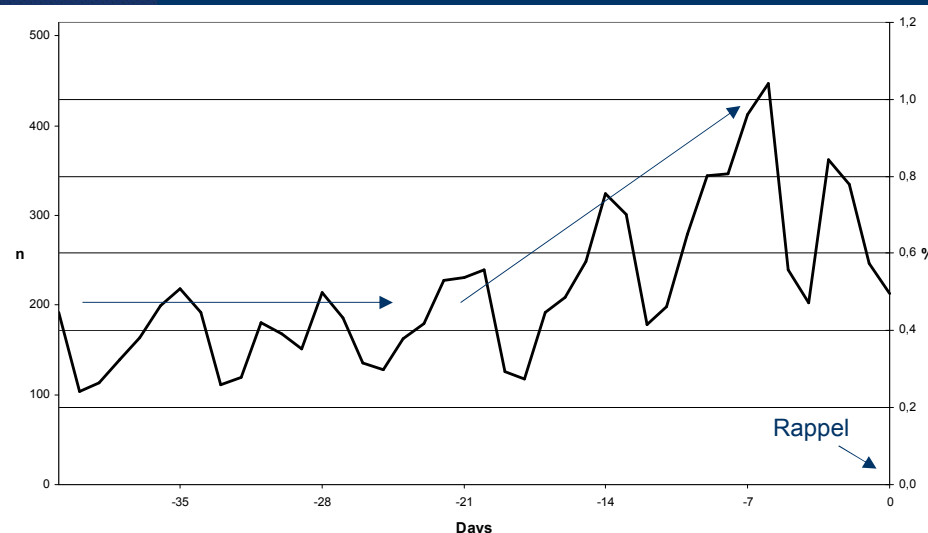
## Responsverloop na rappel 1



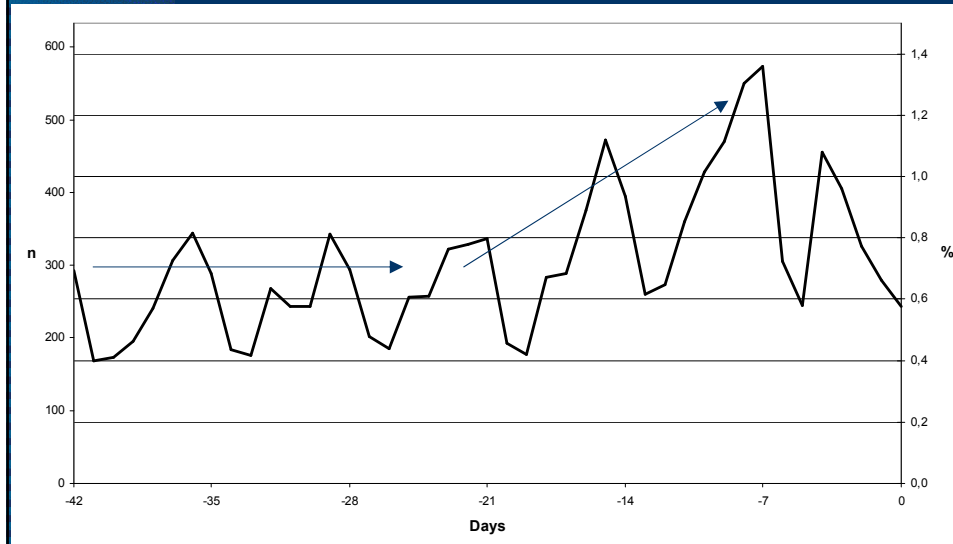
## Invulprevalentie na rappel 1



## Responsverloop voor rappel 1



## Invulprevalentie voor rappel 1



## Tweede rappel

- Vergelijkbaar met 1e rappel
- Lagere, smallere pieken met steilere hellingen
- Er waren minder bedrijven die het vragenlijstprogramma voor de 1e keer openden
- De totale tijd besteedt aan het invullen van de vragenlijst was minder bij rappel 2 dan bij rappel 1



11 maart 2010

NPSO, Brussel, België

## Resultaten samengevat

- Een meerderheid van de bedrijven vult de PS in en respondeert snel nadat ze de vragenlijst voor het eerst openen
- De invulactiviteit en respons van bedrijven neemt toe zowel in de weken voor als na de rappels:
  - Respons
  - Invulincidentie (beginnen met invullen)
  - Invulprevalentie
  - Invultijd



11 maart 2010

NPSO, Brussel, België

## Resultaten

### Wanneer vullen bedrijven de vragenlijst in?

- Als het hen uitkomt!  
Als ze de gegevens en de middelen hebben
  - Als het CBS met hen in contact treedt (rappel)
- Beide zijn van toepassing:
- Ze beginnen niet voordat ze de data en de middelen hebben
    - Hierop hebben we geen invloed!
  - Ze beginnen als ze gerappelleerd worden
    - Kan worden beïnvloed!



11 maart 2010

NPSO, Brussel, België

## Gevolgen voor contactstrategie

- **Rappels zijn een effectieve manier om respons te verkrijgen. Ze functioneren als:**
  - Motivatie tot invullen en responderen
  - Deadline
- **Uitzenden vragenlijsten liever later dan vroeger:**
  - Bedrijven beginnen niet met invullen omdat vroeg wordt uitgezonden
- **Intervallen tussen rappels van >4 weken lijken het meest effectief**



11 maart 2010

NPSO, Brussel, België

## Andere mogelijkheden met audit trails

- **Tellen van invulprofielen:**
  - Welke profielen komen vaak voor?
  - In relatie tot achtergrondkenmerken en invulverloop?
  - Verschillen tussen nieuwe en ervaren respondenten?
- **Aanpassen van antwoorden:**
  - Bij welke vragen: Wat zijn moeilijke vragen?
  - Hoe vaak gebeurt dit?



11 maart 2010

NPSO, Brussel, België



## Wat leert ons dit?

- Hoe kunnen we de vragenlijst aanpassen aan het invulgedrag
- Wat is een optimale benaderingsstrategie



11 maart 2010

NPSO, Brussel, België