

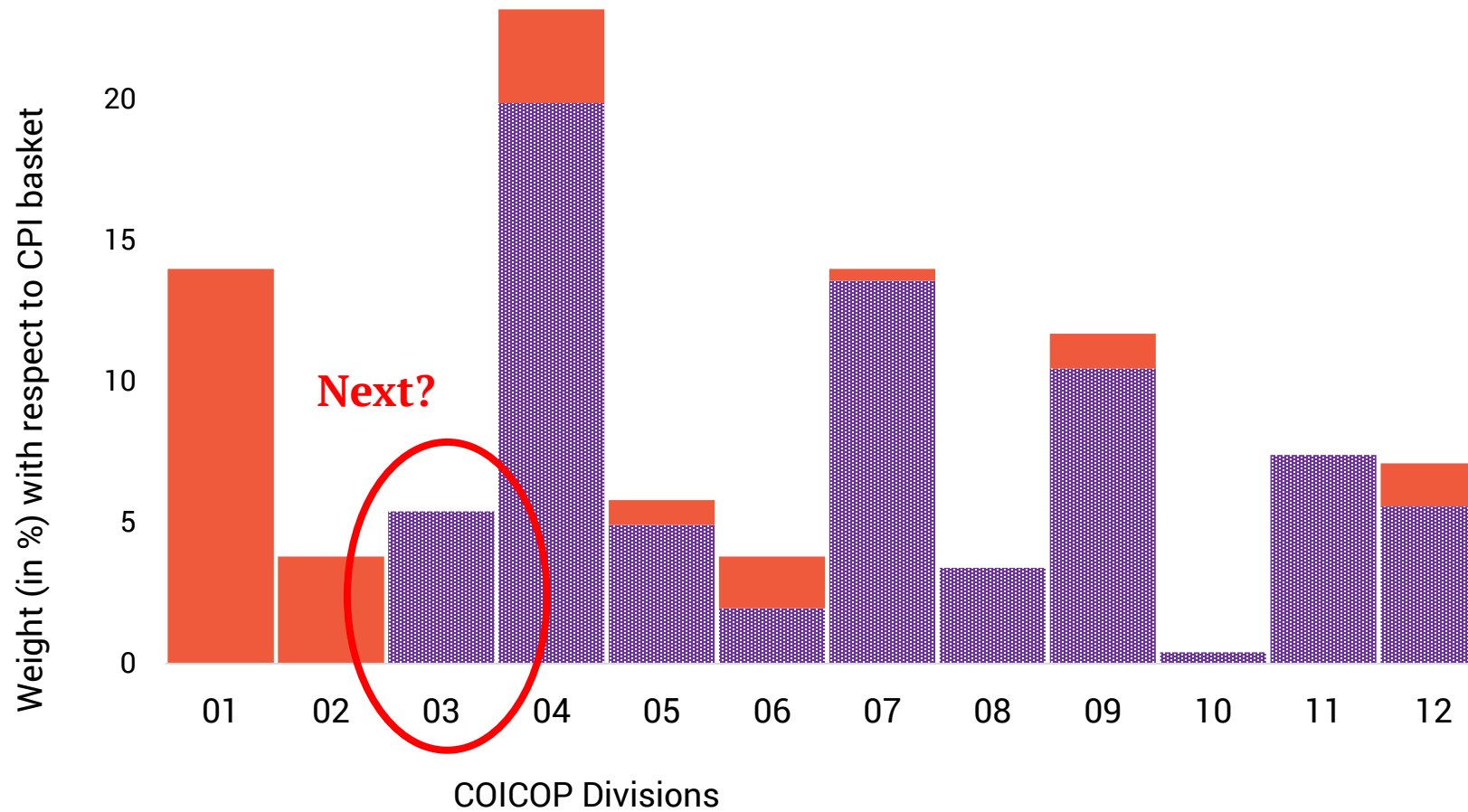
# Towards a Roadmap for Efficient Use of Electronic Transaction Data in the Swedish CPI

Olivia Ståhl, Peter Nilsson  
Price unit, Statistics Sweden  
Ottawa group meeting May 2019

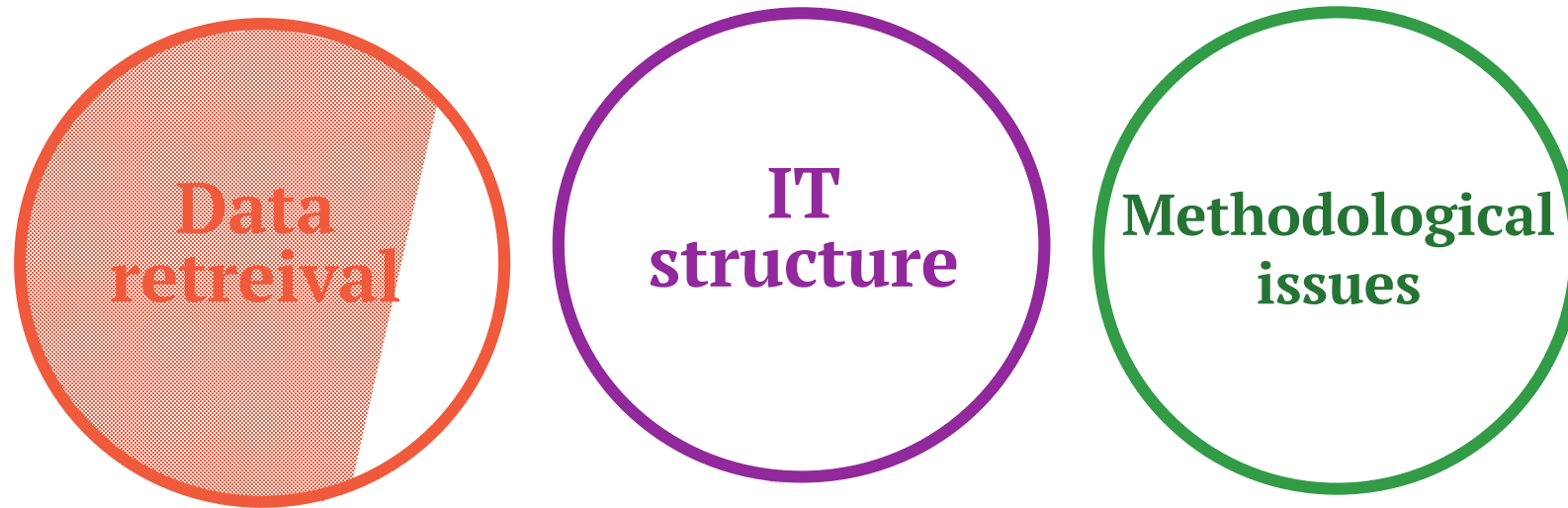




# Current situation



# Current situation



Challenge is to convince potential data providers that

- ... we are trustworthy organization

- ... it is in everyone's best interest to have a high quality CPI

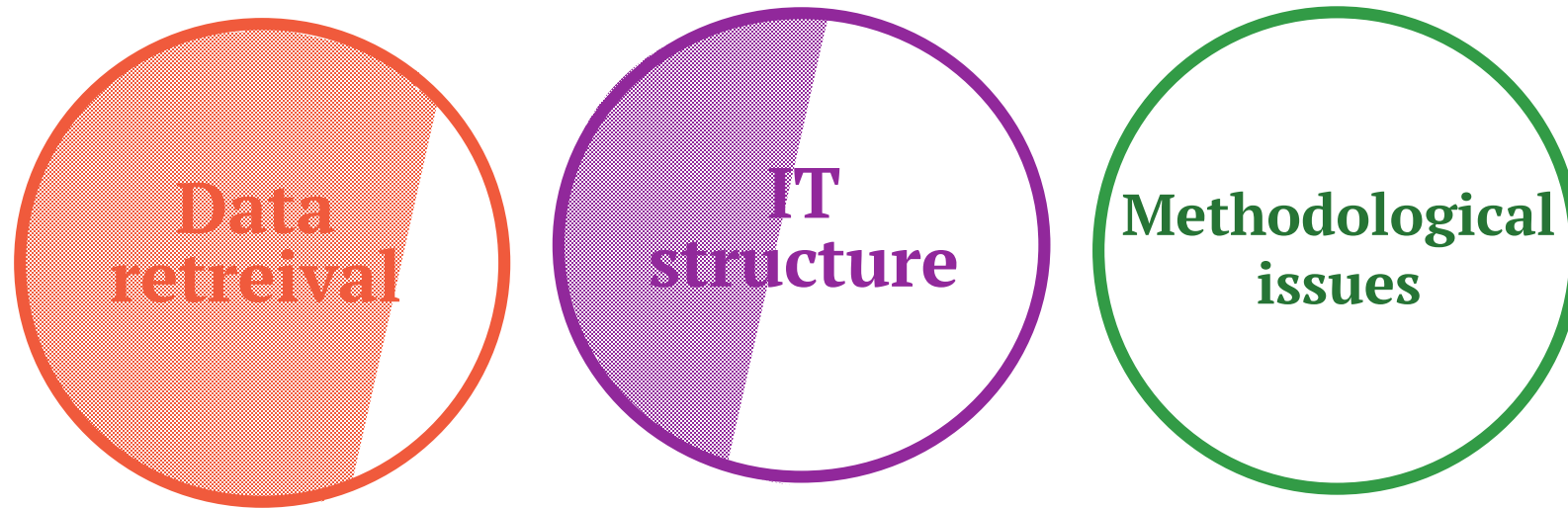


Quite successful so far in retrieving transaction data – But need standardized process

- ... future might hold many small data suppliers

- ... re-organization within Statistics Sweden

# Current situation



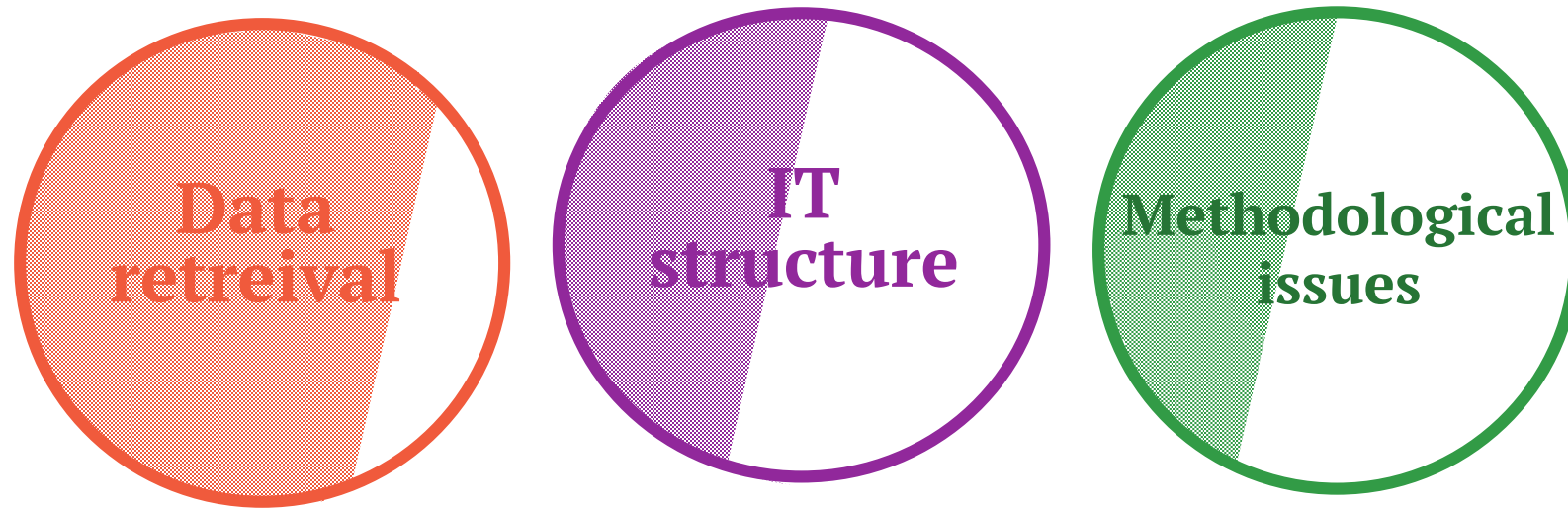
ETD sets new demands on IT systems...

... so far: incorporated ETD in existing structure

→ complex system

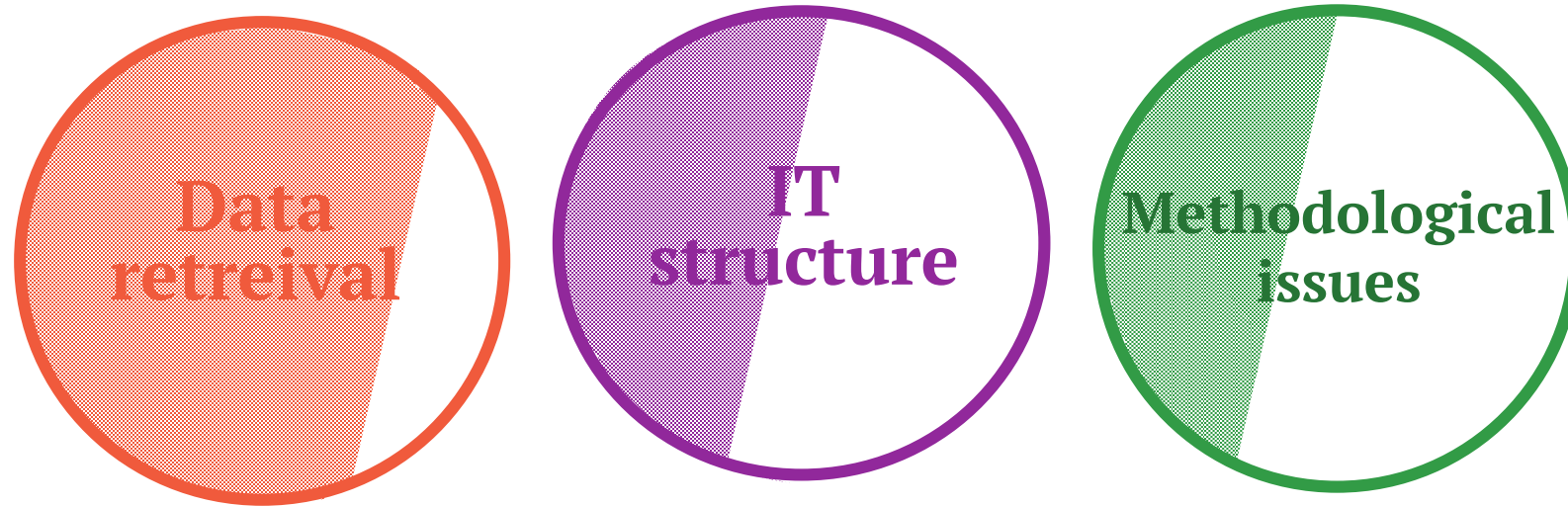
→ complicates methodological studies

# Current situation



- Somewhat diverging methodologies
- New methods (e.g. multilateral) have not been systematically evaluated
- Need to develop new "mindset" within the working group(s)

# Current situation



In summary:

Successful implementation of ETD in many product areas

... *but*

... lacking economies of scale

... not fully consistent

... perhaps not transparent to users

# Discussions prior to 2012

- Different ways of using scanner data
- Treatment of discounts
- Aggregation from weekly to monthly prices

# Discussions prior to 2012

## → Different ways of using scanner data

→ Treatment of discounts

→ Aggregation from weekly to monthly prices

- Use for quality control
- Use as auxiliary information in estimation
- Use to compute prices for a sample
- Use to compute prices and quantities for a sample
- Use to compute prices for full material
- Use to compute prices and quantities for full material



# Discussions prior to 2012

- Different ways of using scanner data
- **Treatment of discounts**
- Aggregation from weekly to monthly prices

One study showed: 49%  
of all prices included  
some form of discount

## HICP regulation:

" ... available to all potential  
consumers with no special  
conditions attached..."

# Discussions prior to 2012

- Different ways of using scanner data
- Treatment of discounts
- **Aggregation from weekly to monthly prices**



Geometric or  
arithmetic?



Weighted or  
unweighted?

# Discussions prior to 2012

- Different ways of using scanner data *Consensus (OK)*
- Treatment of discounts *Consensus (Eurostat reg. might change)*
- Aggregation from weekly to monthly prices *Consensus (Unit value)*

*Today (2019):*

- *What type of multilateral method?*
- *How to create homogeneous products?*

# Current situation

Rapid development these last 6-7 years

- with respect to the amount of ETD available to us...
- with respect to international methodological development...

Has led to somewhat diverging practices today





# Towards a practical "roadmap"

## Treatment not fully consistent between areas

- Sampling used for some product groups, not for others
- Manual replacements frequent in some product groups, not in others
- Editing processes differ (some "black boxes" do exist...)
- Homogeneous product definitions (unit value level)
- Different technical arrangements for data delivery
- Different structure for data storage
- Delivery agreements differ
- User communication practice

# Towards a practical "roadmap"

## Treatment not fully consistent between areas

- Sampling used for some product groups, not for others
- Manual replacements frequent in some product groups, not in others
- Editing processes differ (some "black boxes" do exist...)
- Homogeneous product definitions (unit value level)
- Different technical arrangements for data delivery
- Different structure for data storage
- Delivery agreements differ
- User communication practice

## Goal:

### More standardized treatment of Electronic Transaction Data

- ... in monthly production...
- ... when implementing new data...

# Towards a practical "roadmap"

TOPIC 1) Obtaining contact with potential data providers

TOPIC 2) Specifying requirements on experimental data sets

TOPIC 3) Systemizing continuous data delivery

TOPIC 4) Methodological issues

TOPIC 5) Implementation



# Topic 1

## Obtaining contact with potential data providers

**Until now: “Ad-hoc approach”.** Success formula *but* also person depending & has resulted in many different types of arrangements...

During 2017/2018, a draft **communication strategy** has been set up, in collaboration with Statistics Sweden’s communication division

How should we approach new potential data providers?



From 2019: Responsibility transferred to the division for data collection



# Topic 1

## Obtaining contact with potential

Until now: “Ad-hoc approach”. Success for  
resulted in many different types of arrangements

During 2017/2018, a draft communication  
collaboration with Statistics Sweden’s cooperation

How should we approach new potential distributors



From 2019: Responsibility transferred to Statistics Sweden

## Prisinsamling för Konsumentprisindex (KPI)

Hej XXXX,

Jag skriver till dig med anledning av att SCB arbetar med att utveckla nya sätt för att samla in priser till Konsumentprisindex (KPI). Syftet är att så tidseffektivt som möjligt för både er och oss kunna ta del av bästa tillgängliga prisdata från er. Idag samlar vi in priser från er på plats i era butiker.

KPI har en bred användning i samhället. KPI används som underlag för Riksbankens penningpolitiska beslut och vid beslut inom konjunktur- och stabiliseringspolitiken av Finansdepartementet. Det påverkar också ekonomiska förhållanden genom justering av priser i olika avtal och för att beräkna prisbasbeloppet.

Varje månad samlar SCB in miljontals prisnoteringar som används i beräkningen av KPI. Prisinsamlingen sker på olika sätt från olika företag, de senaste åren har det framförallt blivit vanligare att registerdata skickas automatiskt till SCB. ICA, Coop och Systembolaget har varit med och utvecklat den framgångsrika automatiska processen. Den traditionella prisinsamlingen i butiker utförs fortfarande till viss del men är en metod SCB på sikt kommer att upphöra med.

XXXX är en stor aktör inom XXXindustrin, och utgör en betydande del i KPI. Vi skulle därför vilja ha ett möte med er för att presentera vårt förslag till förbättrad prisinsamling och fördelarna med dem. Mötet kommer att vara ca 1 timme.

Jag ringer dig inom ett par dagar, för att hitta en lämplig tid för oss att träffas. Om vi bör kontakta någon annan inom er organisation, återkom gärna med namn och kontaktuppgifter till mig.

Har du några frågor är du välkommen att kontakta mig via e-post eller ring

## Meetings with enterprises

- Usually enough just to mention the local price collection of today...
- ... most companies reacts negative towards this fact and think its dated

## List of prioritized data deliverers

- Clothing (expensive)

# Prisinsamling för Konsumentprisindex (KPI)

Hej XXXX,

Jag skriver till dig med anledning av att SCB arbetar med att utveckla nya sätt för att samla in priser till Konsumentprisindex (KPI). Syftet är att så tidseffektivt som möjligt för både er och oss kunna ta del av bästa tillgängliga prisdata från er. Idag samlar vi in priser från er på plats i era butiker.

KPI har en bred användning i samhället. KPI används som underlag för Riksbankens penningpolitiska beslut och vid beslut inom konjunktur- och stabiliseringspolitiken av Finansdepartementet. Det påverkar också ekonomiska förhållanden genom justering av priser i olika avtal och för att beräkna prisbasbeloppet.

Varje månad samlar SCB in miljontals prisnoteringar som används i beräkningen av KPI. Prisinsamlingen sker på olika sätt från olika företag, de senaste åren har det framförallt blivit vanligare att registerdata skickas automatiskt till SCB. ICA, Coop och Systembolaget har varit med och utvecklat den framgångsrika automatiska processen. Den traditionella prisinsamlingen i butiker utförs fortfarande till viss del men är en metod SCB på sikt kommer att upphöra med.

XXXX är en stor aktör inom XXXindustrin, och utgör en betydande del i KPI. Vi skulle därför vilja ha ett möte med er för att presentera vårt förslag till förbättrad prisinsamling och fördelarna med dem. Mötet kommer att vara ca 1 timme.

Jag ringer dig inom ett par dagar, för att hitta en lämplig tid för oss att träffas. Om vi bör kontakta någon annan inom er organisation, återkom gärna med namn och kontaktuppgifter till mig.

Har du några frågor är du välkommen att kontakta mig via e-post eller ring

# Topic 2

## Specifying requirements on experimental data sets

**Aim: Checklist for type of information needed in experimental datasets**

Important for data collection unit staff

- Classification structures,
- Price-determining characteristics,
- Regular/Sales prices,
- Refunds (?)

# Topic 3

## Systemizing continuous data delivery

### **Flexibility**

- Minimize response burden
- Increase likelihood of obtaining high quality data

### **In-house systems must be able to convert and store incoming data and metadata in standardized way**

- Facilitate editing
- Facilitate ad-hoc analyses

Pre-study planned for end 2019

**All “delivery agreements” are being reviewed**





# Topic 4

## Methodological issues

### → Classification

Automatic methods will be tested (planned for end of 2019)

### → Stratification

Subjective – but quantitative measures e.g. the MARS score of help

### → Data cleaning

Filters, outlier detection methods

### → Editing

Macro focus? Probabilistic methods? Process indicators!

### → Index formula

Criteria for selecting a certain index number formula?

### → Sampling design

Variance Vs bias

# Topic 5

## Implementation

### **Goal: Increase total quality of CPI**

- *Relevance*
- *Accuracy*
- *Timeliness and punctuality*
- *Accessibility and clarity*
- *Comparability and coherence*

### **Draft “internal policy” to be discussed in CPI board (not only ETD)**

- Parallel series of at least one/two full years
- Estimated effects (based on historic data) in beginning of January

### **Leaving the fixed-basket approach for ETD?**

- What is really “transparent”..?





An aerial photograph of a dense forest with a curved path or road cutting through it. The trees are mostly green, with some yellowing, suggesting autumn. The path is a light brown color and curves from the left towards the right.

# Thank you

Olivia Ståhl, Peter Nilsson  
CPI unit, Statistics Sweden  
Ottawa group meeting May 2019

**SCB**