

# 14<sup>th</sup> Global Forum on Tourism Statistics

## 23-25 November 2014, Venice, Italy

### Session 1: Mobile phone data for tourism statistics

### When mobile network operators and statistical offices meet - integrating mobile positioning data into the production process of tourism statistics

Project team: Christophe Demunter, Gerdy Seynaeve,  
Youri Baeyens, Marc Debusschere, Freddy De Meersman, Pieter Dewitte,  
Patrick Lusyne, Fernando Reis, Hannes I. Reuter, Albrecht Wirthmann



# Outline of the presentation

- **Prehistory & project team**
- **Aims of the project**
- **Data sources**
- **Scope of the study**
- **Intermediate results**
- **Conclusions & outlook**

## Prehistory

- **Feasibility study on the use of mobile positioning data for tourism statistics (Eurostat) – 2012-2014**
- **Creation of Task Force on Big Data - 2013**



## Project team

- **Partnership between statistical offices and mobile network operator (MNO)**
  - ⇒ **Eurostat**
  - ⇒ **Proximus**
  - ⇒ **Statistics Belgium**

eurostat 

proximus

 **economie**  
FPS Economy, S.M.E.s, Self-employed and Energy

# Aims of the project

- **Explore partnerships and business models for cooperation between MNOs and NSIs**
  - ⇒ **Access**
  - ⇒ **Exploring, testing, pilots**
  - ⇒ **Regular data processing and production**
  - ⇒ **Continuity, viability of cooperation**
- **Cooperate on concrete, output-oriented projects**
  - ⇒ **Population statistics (present/resident population)**
  - ⇒ **Tourism statistics**
  - ⇒ **...**

# Data sources for the study

## ■ Mobile phone data

- ⇒ Data from one operator in Belgium: **PROXIMUS**
- ⇒ No longer based on Call Detail Records (CDR) but on **signaling data** → #observations x 10 [on the home network]
- ⇒ Better **temporal** (and geographical) **granularity**

## ■ Official tourism statistics

- ⇒ **Survey** based data on trips made by residents of Belgium
- ⇒ In line with the **Regulation concerning European statistics on tourism**
- ⇒ Quarterly interviews, annual **sample ± 10000 trips** (domestic + outbound trips with overnight stays)

# Known weaknesses

<b>Mobile phone data</b>	<b>Official tourism statistics</b>
<b>Selectivity bias</b>	<b>Non-response, non-contact,...</b>
<ul style="list-style-type: none"> <li>• <b>Extrapolation (inverse of market share)</b></li> </ul>	<b>Recall bias, memory effect</b>
<ul style="list-style-type: none"> <li>• <b>Socio-demographic composition of subscribers</b></li> </ul>	<b>Respondent burden</b>
<ul style="list-style-type: none"> <li>• <b>Intensity of use</b></li> </ul>	<b>Timeliness</b>
<b>Entirely algorithm based (choice of parameters?)</b>	<b>Entirely respondent based ('subjective opinion')</b>
<b>Continuity</b>	
<b>Privacy, public opinion</b>	

# Scope

- **Focus on outbound trips**

- ⇒ Mobile phone data: trips made **April – Sept 2016**
- ⇒ Official tourism statistics: trips made **April – Sept 2015**

- **Definition of an outbound trip**

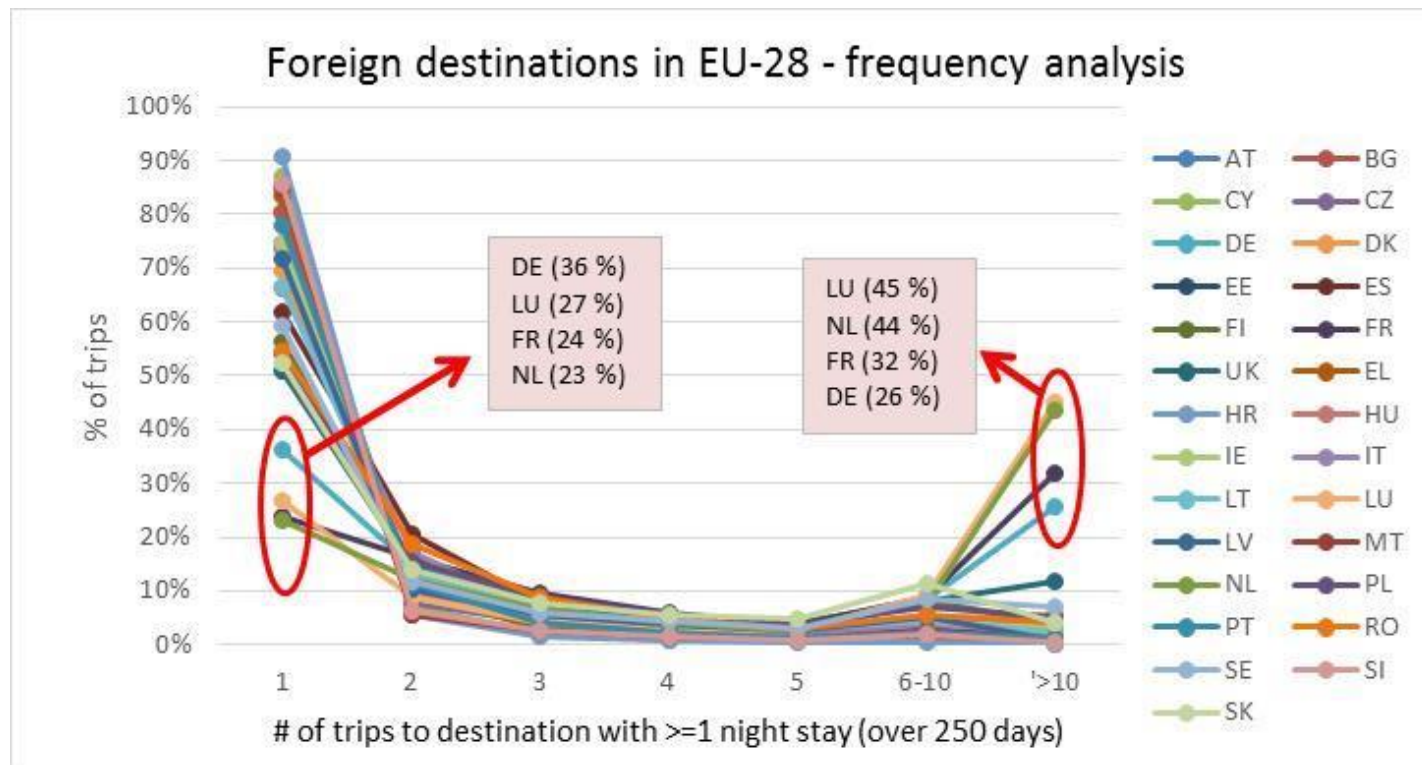
- ⇒ From leaving the home network to returning
- ⇒ Number of nights: number of hours divided by 24
- ⇒ Overnight stay: minimum 10 hours and return after 4am

- **Usual environment**

- ⇒ Duration (min. 10hrs + incl. 4am), border crossing (outbound)
- ⇒ Filtering of frequent trips to the same destination during a given reference period (250 days) → threshold = **5** (arbitrary)

# How often is a SIM seen in a given country?

***Distribution of SIMs, in terms of the number of times a SIM is observed during a 250 days period, by destination (MNO data)***



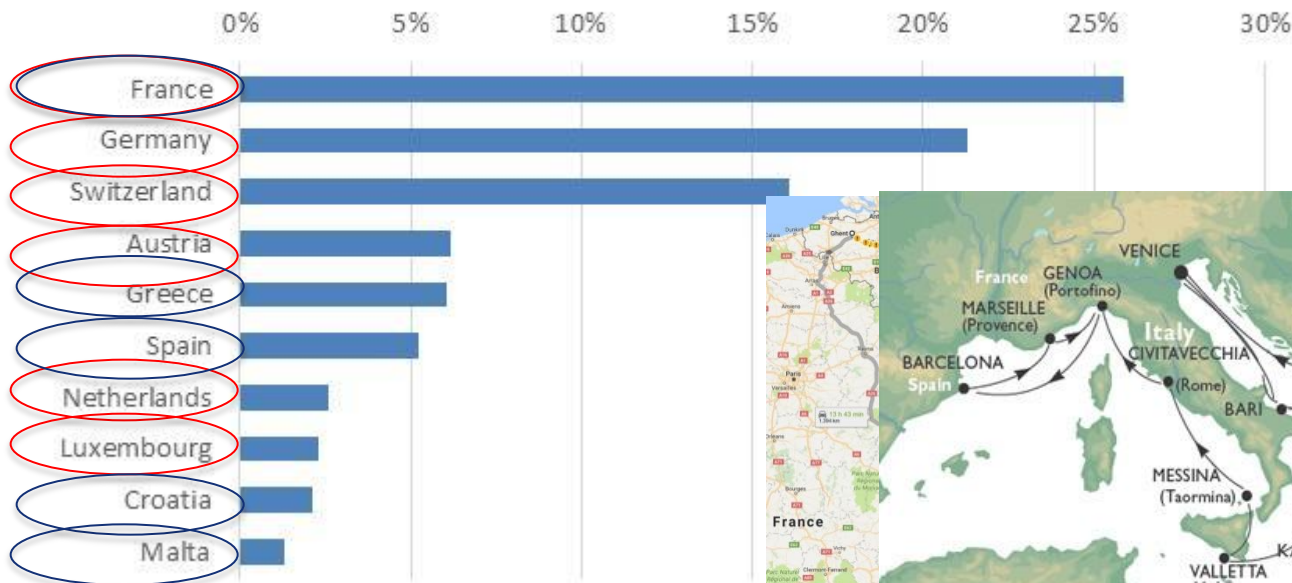


# Main and secondary destinations, transit

**Other countries visited during trips (of 7 nights) with Italy as main destination, top 10 (MNO data)**

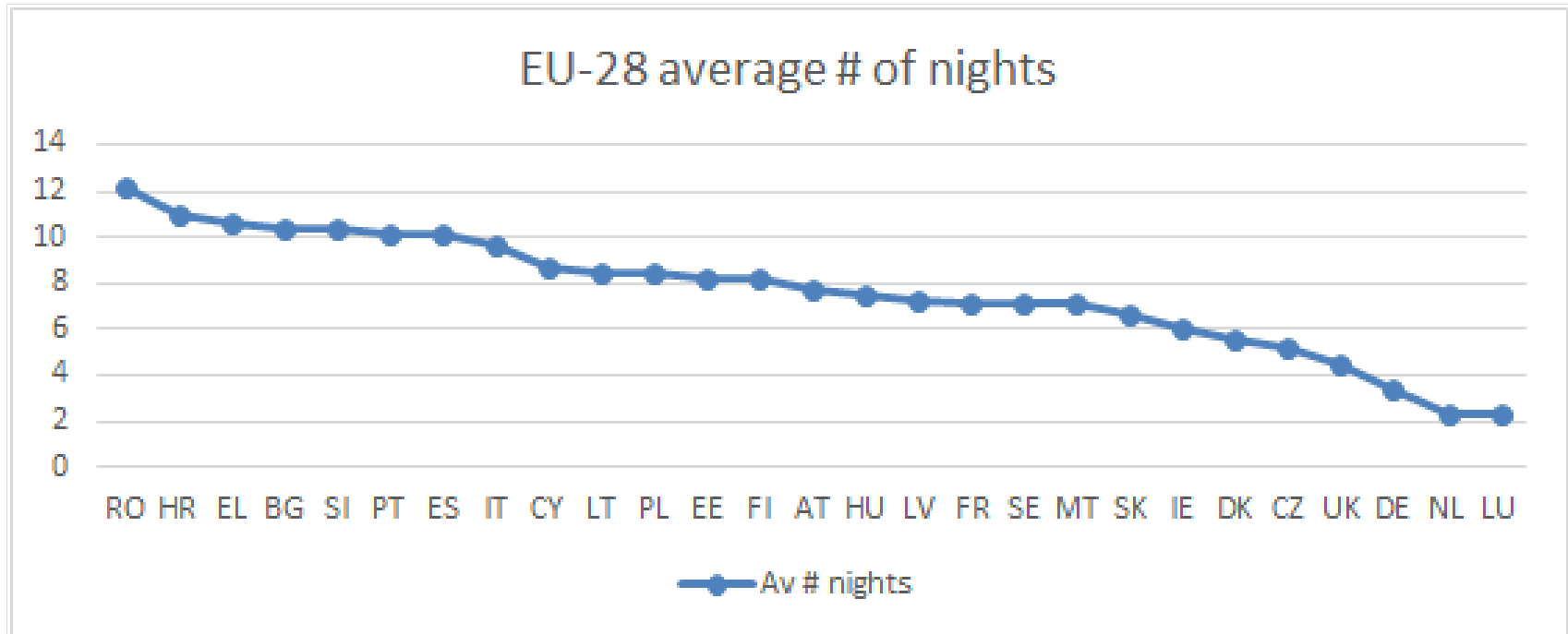
TOP 10 other countries visited during trips of 7 nights to Italy

Importance of country (staying time based)



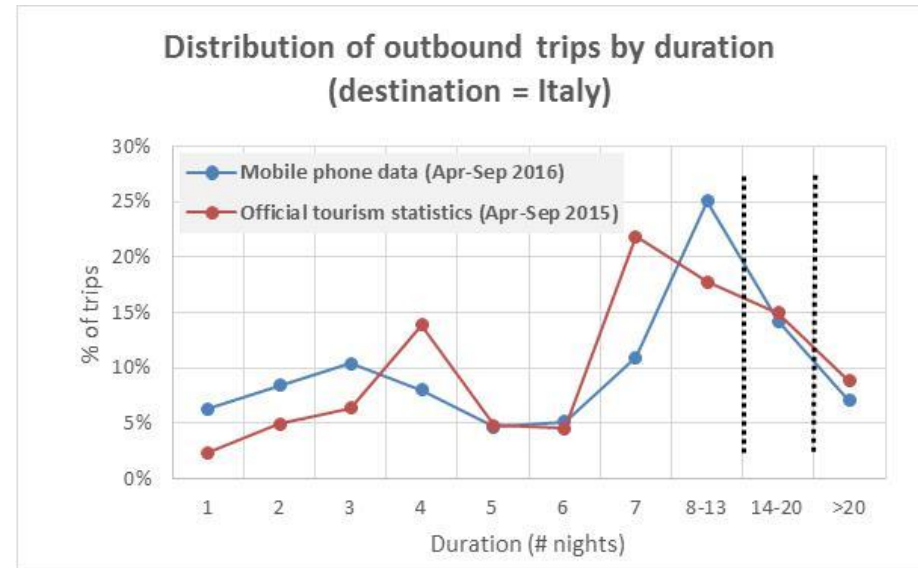
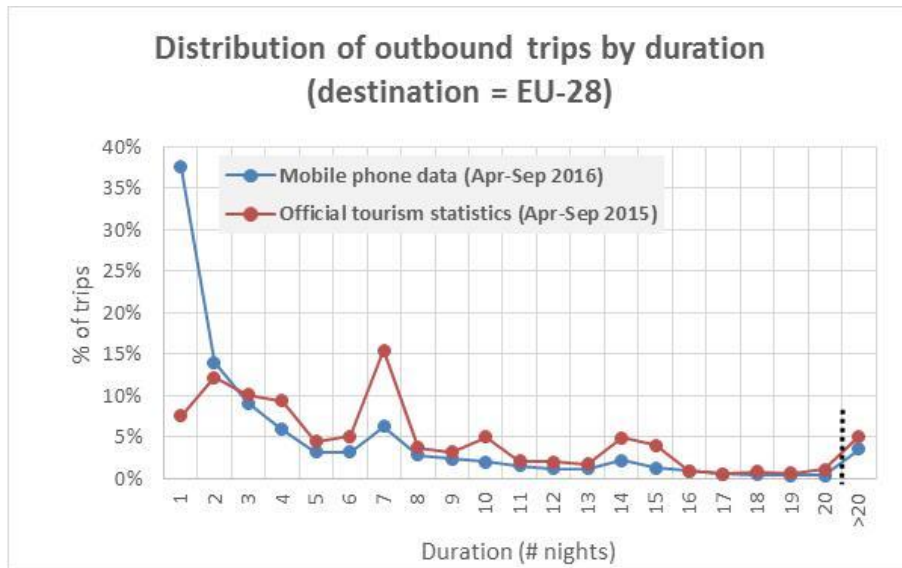
# Ranking of destinations

## *Ranking of EU-28 countries as destination for Belgian outbound trips (MNO data)*



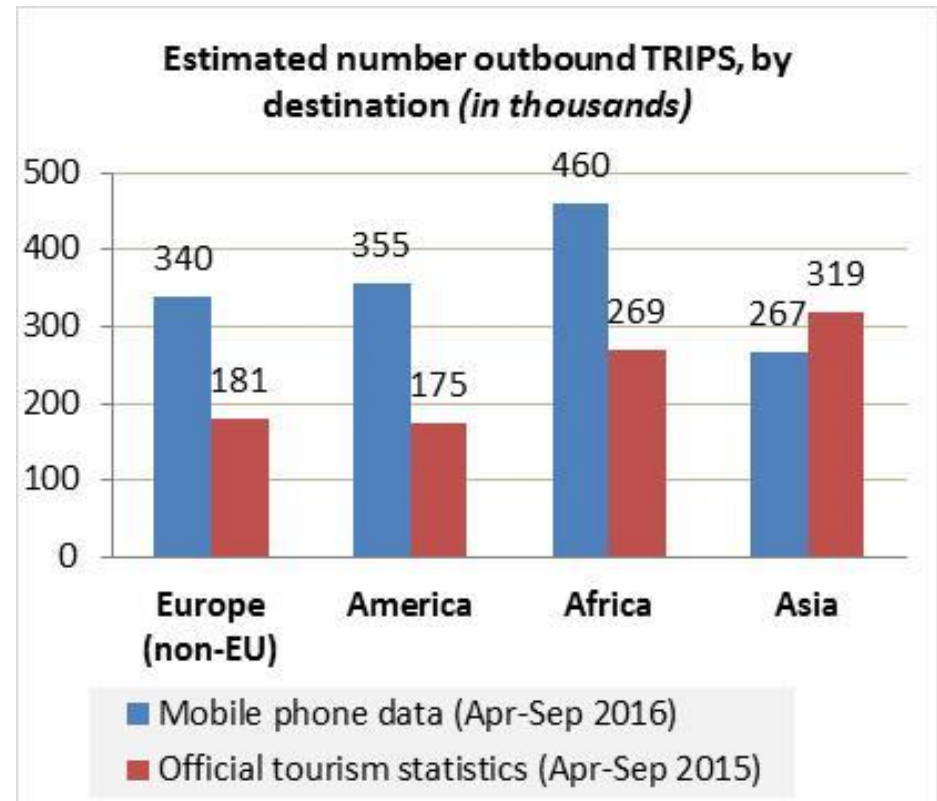
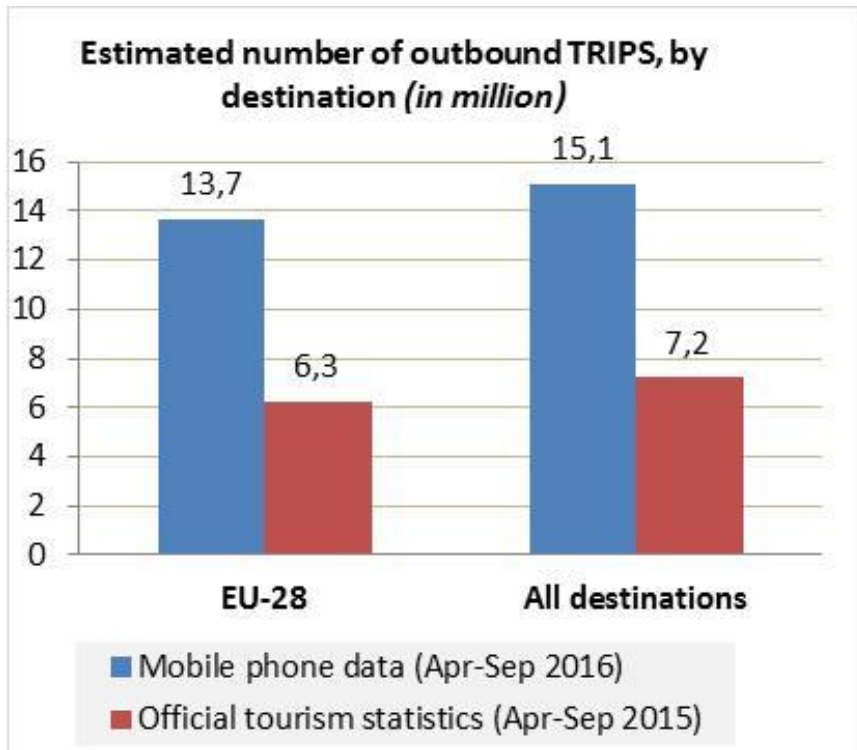
# Outbound trips by duration: comparison

***Comparison of the distribution of outbound trips to EU-28 and to Italy, by duration of the trips***



# Volume of trips and nights: comparison

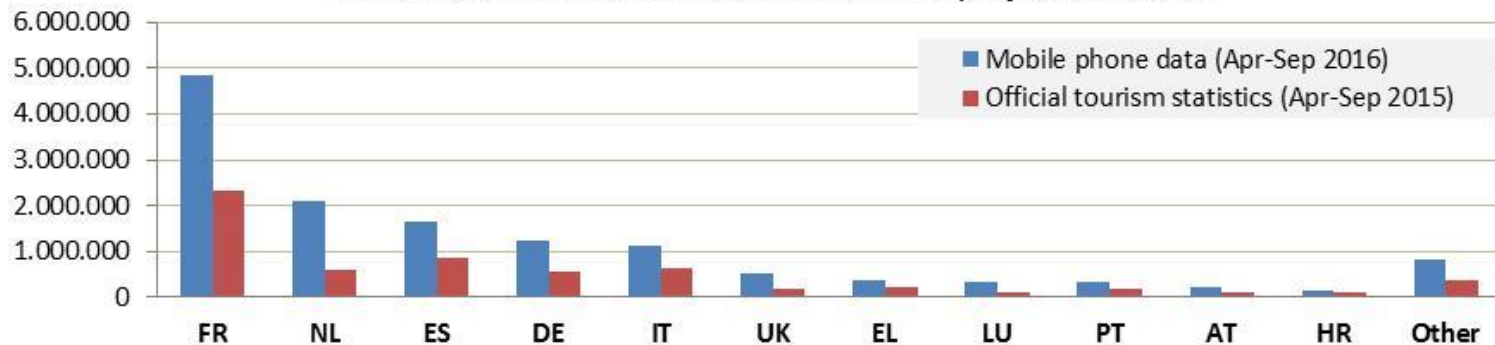
## Comparison of estimated number of outbound trips, by destination



# Volume of trips and nights: comparison (2)

***Comparison of estimated number of outbound trips, by destination: absolute values and ratio***

Estimated number of outbound TRIPS, by destination



Ratio estimated trips mobile phone data vs. official tourism statistics



# Volume of trips and nights: comparison (3)

## ■ Observations

- ⇒ **Big difference** between MNO and NSI estimates
- ⇒ **Systematic** nature

## ■ Understanding (and solving...) the deviations

- ⇒ Difference in **scope** (e.g. age limit) & reference year
- ⇒ **Selectivity bias** and impact on **extrapolations**
- ⇒ Intermediate results, **model & algorithm optimisation** ongoing (e.g. parameter setting for 'usual environment')
- ⇒ Dealing with **recall bias** in survey based data
- ⇒ Impact of **non-response** in surveys (structural bias?)

## ■ The project continues ...

# Conclusions

- **Positive & fruitful experience with the partnership**
  - ⇒ **Joining forces** (statisticians, data holders, data scientists)
  - ⇒ **Search for a win-win**
- **Promising results, but lots of homework**
  - ⇒ Mobile phone data clearly captures tourism concepts/definitions
  - ⇒ Currently: satisfactory for trends, not for estimating volumes
  - ⇒ How to make the series/sources converge to the unknown true values
  - ⇒ Extension to domestic tourism, to same-day visits
- **Further research to be encouraged** (other countries?)

# Outlook

## Mobile phone data for statistics: 3 stages



- **Test / Cross-validate**
  - Use mobile phone data to assess quality of current data
  - Interaction between mobile phone data and current data to develop new models or to fine-tune algorithms
- **Enrich / enhance**
  - Mobile phone data as auxiliary data, e.g. calibration of survey data
  - Mobile phone data as new data (previously not available variables or breakdowns)
- **Replace**
  - Partial (!) replacement of traditional surveys with mobile phone data





European  
Commission

**Thanks for  
your attention**

[christophe.demunter@ec.europa.eu](mailto:christophe.demunter@ec.europa.eu)

[gerdy.seynaeve@proximus.com](mailto:gerdy.seynaeve@proximus.com)