

# gesis

Leibniz Institute  
for the Social Sciences

5<sup>th</sup> European User Conference For EU- Microdata  
Mannheim, March 2-3, 2017 | Germany

## The Digital Divide in Europe in the 21<sup>st</sup> Century: a new methodological challenge for comparative social research

João Miguel Aguiar

PhD Candidate, University of Porto and UT Austin | Portugal Program

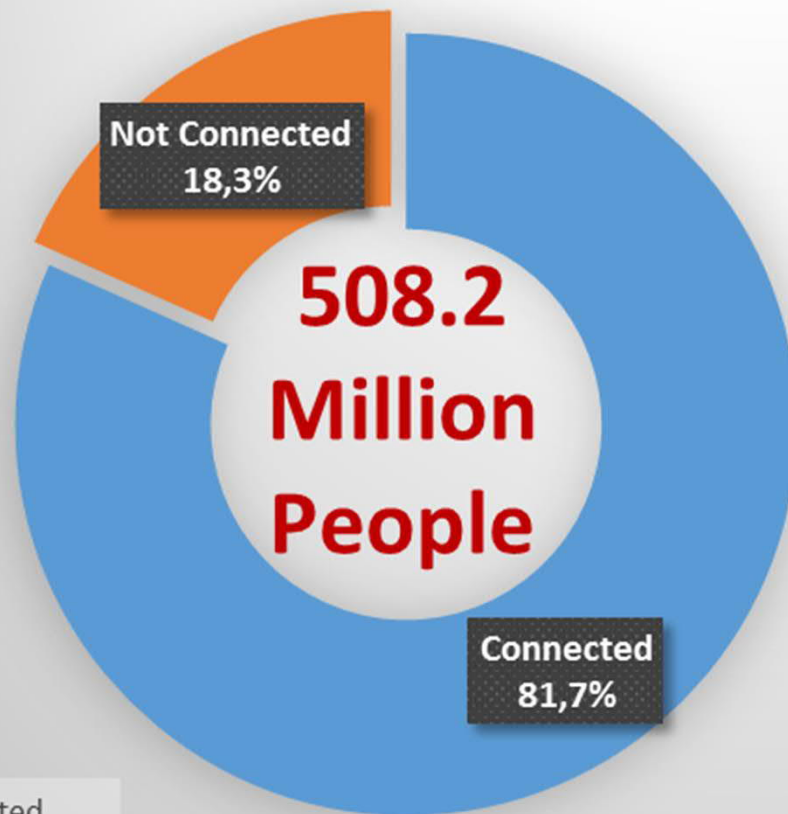
Pedro Campos

University of Porto | Statistics Portugal

José Azevedo

University of Porto

## EU INTERNET ACCESS (2016)



### Internet Penetration

. At least 9 out of every 10 individuals in DK, LU, NL, SE, FI, UK and DE used the internet.

. Less than two thirds of all individuals aged 16 to 74 used the internet in PT, EL, LT, BG and RO.

. 14% of the EU-28's population never used the internet.

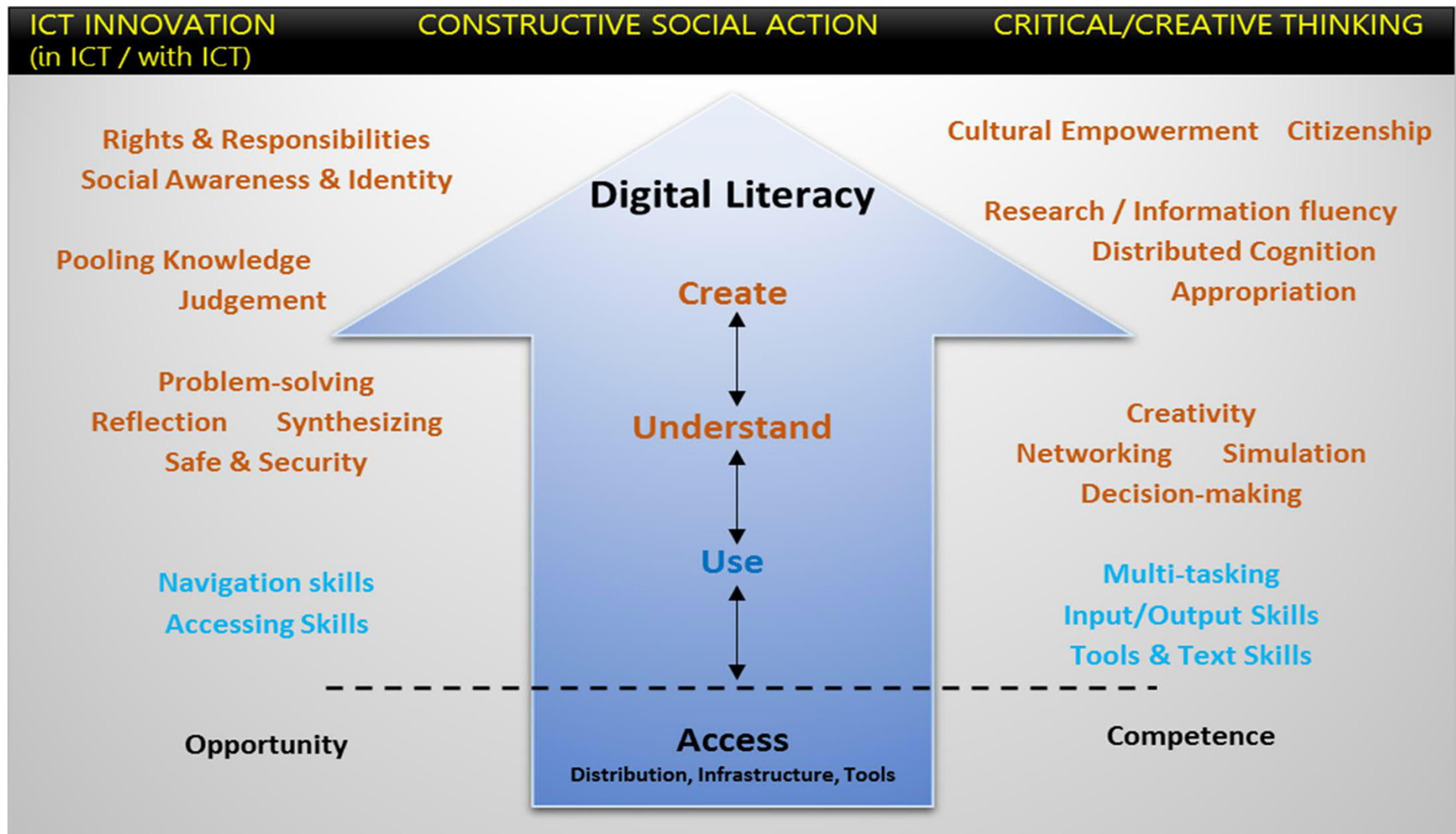
% of all individuals in the EU-28, aged between 16 and 74 years, used the internet (at least once within three months prior to the survey date).

Source: <http://ec.europa.eu/>

## Objectives of the study

---

- . Analyze the Digital Divide state in European Union and how it has evolved in the last few years (2008-2015)



This figure is based on models from the "Report of the Digital Britain Media Literacy Working Group" (March 2009), "DigEuLit - a European Framework for Digital Literacy" (2005), and Jenkins et al. (2006) "Confronting the Challenges of Participatory Culture: Media Education for the 21st Century".

## Research questions

---

**RQ1:** What is the current state of the Digital Divide among the European Union countries?

(Digital Divide Index - Chi-square Test - CATPCA)

**RQ2:** How did the Digital Divide evolved within the European Union in the last few years?

(Time-distance Methodology)

## Methodology / research framework

---

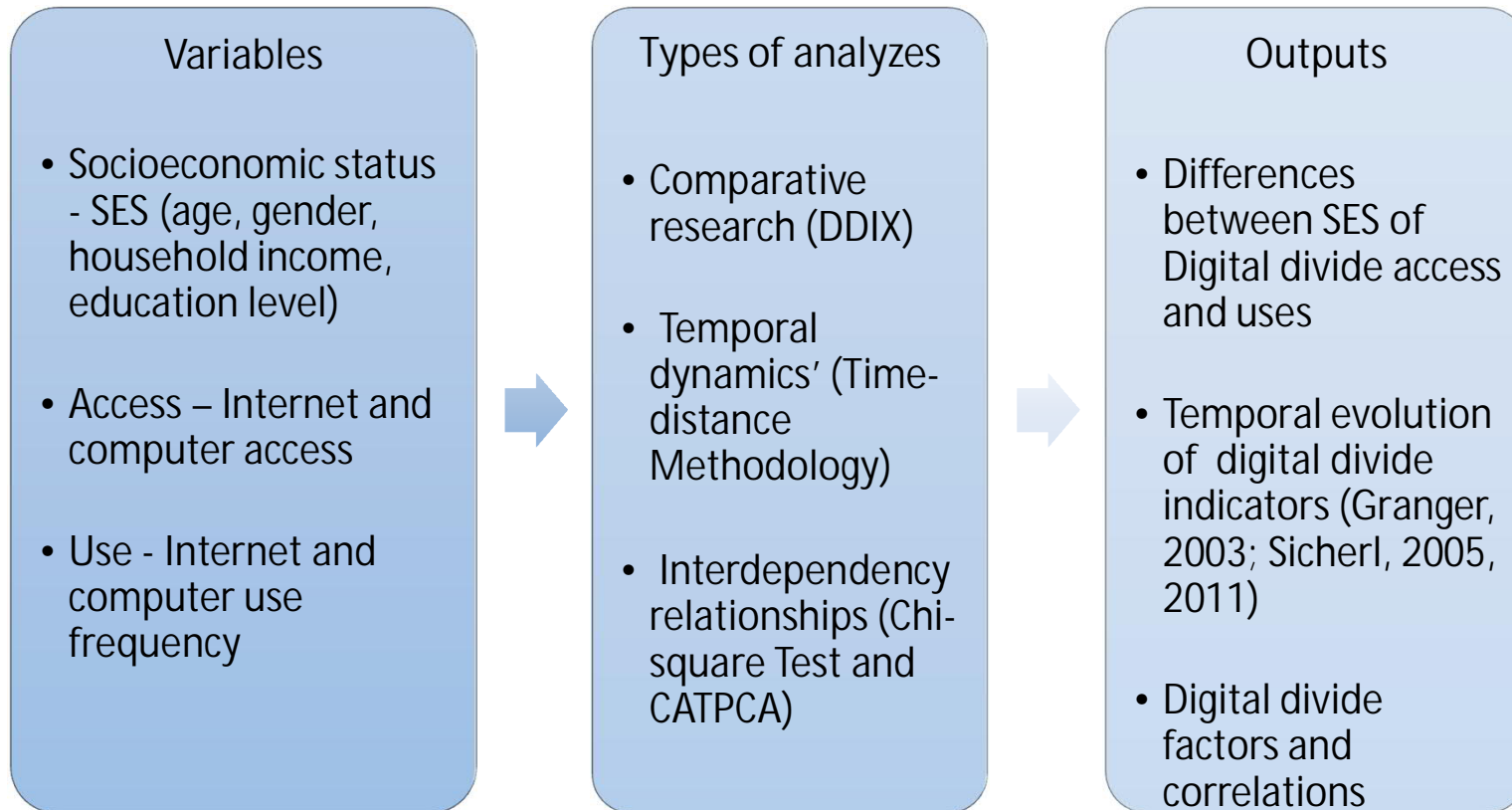


Fig. 2 – Digital Divide research framework: A multi-model approach (Centering on the 'user')

## Methodology / sample

---

Eurostat's Community survey on ICT usage in households and by individuals (2008-2015) micro-data.

The 2013 survey covered 148 290 households with at least one person aged 16-74, and 220 436 individuals aged 16-74 in the EU.

Source: [http://epp.eurostat.ec.europa.eu/statistics\\_explained/index.php/Information\\_society\\_statistics](http://epp.eurostat.ec.europa.eu/statistics_explained/index.php/Information_society_statistics)

# Methodology / sample

```
RGui (32-bit) - [C:\Users\Miguel Aguiar\Documents\DOCTORAMENTO\4º Ano 2º Semestre\DDIX Indicators Tables ALL\Codigo para p_value Europe.R]
File Edit Packages Windows Help
library(Rcmdr)

D<-matrix(nrow=5000, ncol=10)

setwd("C:/Dados")

países<-c("BE", "BG", "CY", "DK", "EE", "EL", "ES", "FI", "IS", "IT", "LU", "MT", "NL", "NO", "PT", "SE", "SI", "SK")
#.Test <- chisq.test(.Table, correct=FALSE)
l<-1

for (ano in 2008:2013) {

if (ano==2008) M<-read.csv("TOTAL EU_2008_DataForResearchers_CLEAN.csv", header=T, sep=";", dec=",")
if (ano==2009) M<-read.csv("TOTAL EU_2009_DataForResearchers_CLEAN.csv", header=T, sep=";", dec=",")
if (ano==2010) M<-read.csv("TOTAL EU_2010_DataForResearchers_CLEAN.csv", header=T, sep=";", dec=",")
if (ano==2011) M<-read.csv("TOTAL EU_2011_DataForResearchers_CLEAN.csv", header=T, sep=";", dec=",")
if (ano==2012) M<-read.csv("TOTAL EU_2012_DataForResearchers_CLEAN.csv", header=T, sep=";", dec=",")
if (ano==2013) M<-read.csv("TOTAL EU_2013_DataForResearchers_CLEAN.csv", header=T, sep=";", dec=",")

#ciclo países
  for (i in 1:18) {

#qualquer país
M_BE<-subset(M, M$i..COUNTRY==países[i])

T<-table(M_BESCOMP,M_BESSEX)
l<-1+l;D[1,1]<-países[i];D[1,2]<-ano; D[1,3]<-"V1"; D[1,4]<-"V2"; D[1,5]<-colPercents(table(M_BESCOMP,M_BESSEX)) [2,2]
D[1,6]<-totPercents(table(M_BESCOMP,M_BESSEX)) [2,3]
D[1,7]<-chisq.test(T, correct=FALSE)$p.value

T<-table(M_BESIACC,M_BESSEX)
l<-1+l;D[1,1]<-países[i];D[1,2]<-ano; D[1,3]<-"V1"; D[1,4]<-"V2"; D[1,5]<-colPercents(table(M_BESIACC,M_BESSEX)) [2,2]
D[1,6]<-totPercents(table(M_BESIACC,M_BESSEX)) [2,3]
D[1,7]<-chisq.test(T, correct=FALSE)$p.value

T<-table(M_BESCU,M_BESSEX)
l<-1+l;D[1,1]<-países[i];D[1,2]<-ano; D[1,3]<-"V1"; D[1,4]<-"V2"; D[1,5]<-colPercents(table(M_BESCU,M_BESSEX)) [1,2]
D[1,6]<-totPercents(table(M_BESCU,M_BESSEX)) [1,3]
D[1,7]<-chisq.test(T, correct=FALSE)$p.value

}
}

09:36
23/02/2017
```

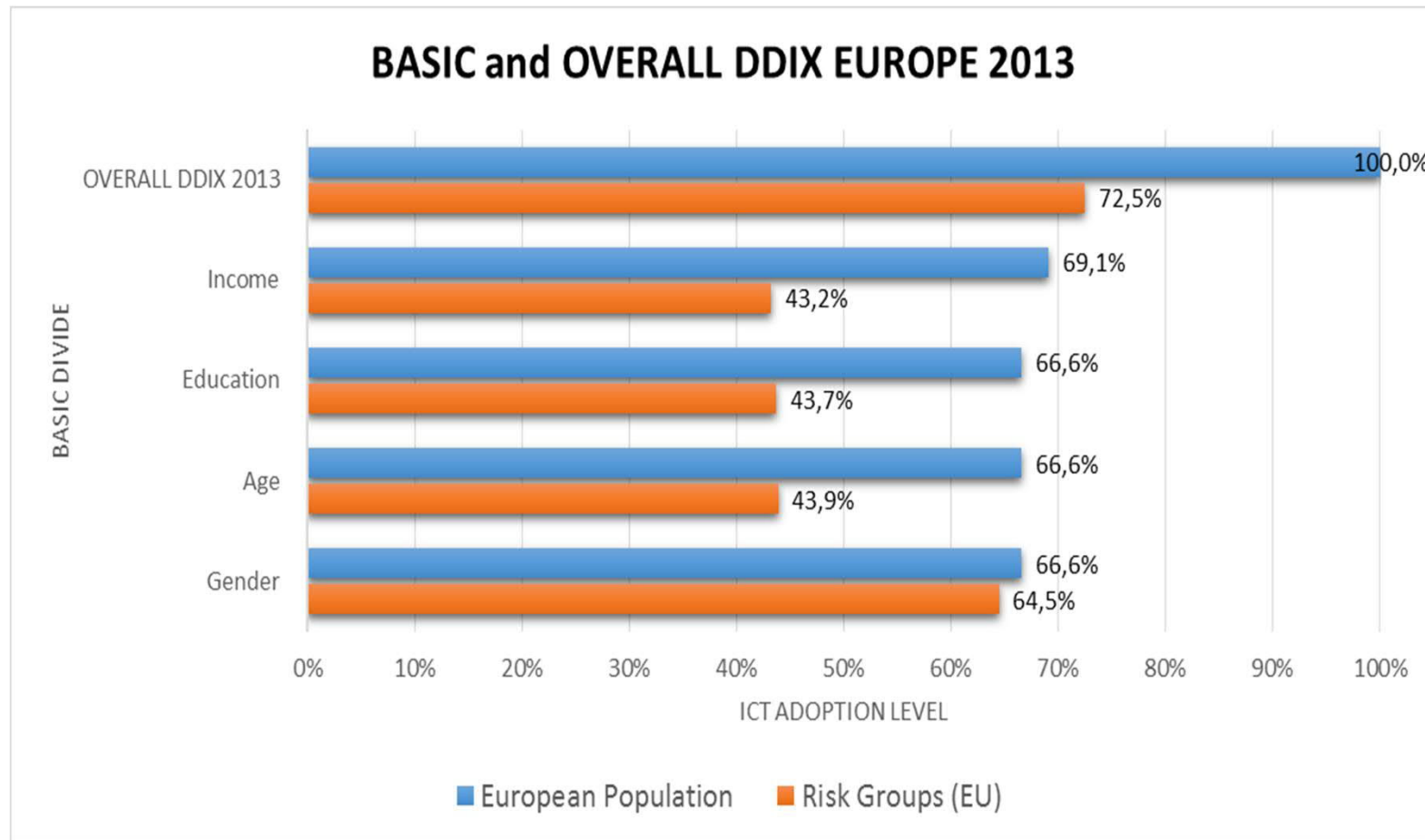


Table 3 - Chi-square Indendence Test - Risk Groups vs ICT Access Indicators (p-values)

<b>Access and ICT use indicators</b>	<b>2008</b>	<b>2009</b>	<b>2010</b>	<b>2011</b>	<b>2012</b>	<b>2013</b>
<b>Gender (female vs male)</b>						
A1: do you or anyone in the household have access to a computer at home?	0,0153 *	0,0000 **	0,0000 **	0,0371 *	0,0019 **	0,1597
A2: do you or anyone in the household have access to the internet at home? (by any device)	0,0028 **	0,0000 **	0,0000 **	0,2682	0,0724	0,5664
B1: when did you last use a computer (at home, at work any other place)? Within the last 3 months	0,0000 **	0,0000 **	0,0000 **	0,0000 **	0,0000 **	0,0000 **
C1: when did you last use the internet? Within the last 3 months	0,0000 **	0,0000 **	0,0000 **	0,0000 **	0,0000 **	0,0000 **
<b>Age (Individuals aged 54 or more vs others)</b>						
A1: do you or anyone in the household have access to a computer at home?	0,0000 **	0,0005 **	0,0000 **	0,0000 **	0,0000 **	0,0000 **
A2: do you or anyone in the household have access to the internet at home? (by any device)	0,0000 **	0,0001 **	0,0000 **	0,0000 **	0,0000 **	0,0000 **
B1: when did you last use a computer (at home, at work any other place)? Within the last 3 months	0,0000 **	0,0000 **	0,0000 **	0,0000 **	0,0000 **	0,0000 **
C1: when did you last use the internet? Within the last 3 months	0,0000 **	0,0000 **	0,0000 **	0,0000 **	0,0000 **	0,0000 **
<b>Education (Lower education level vs others)</b>						
A1: do you or anyone in the household have access to a computer at home?	0,0000 **	0,0000 **	0,0000 **	0,0000 **	0,0000 **	0,0000 **
A2: do you or anyone in the household have access to the internet at home? (by any device)	0,0000 **	0,0000 **	0,0000 **	0,0000 **	0,0000 **	0,0000 **
B1: when did you last use a computer (at home, at work any other place)? Within the last 3 months	0,0000 **	0,0000 **	0,0000 **	0,0000 **	0,0000 **	0,0000 **
C1: when did you last use the internet? Within the last 3 months	0,0000 **	0,0000 **	0,0000 **	0,0000 **	0,0000 **	0,0000 **
<b>Income (Lowest quartile vs others)</b>						
A1: do you or anyone in the household have access to a computer at home?	0,0000 **	0,0000 **	0,0000 **	0,0000 **	0,0000 **	0,0000 **
A2: do you or anyone in the household have access to the internet at home? (by any device)	0,0000 **	0,0000 **	0,0000 **	0,0000 **	0,0000 **	0,0000 **
B1: when did you last use a computer (at home, at work any other place)? Within the last 3 months	0,0000 **	0,0000 **	0,0000 **	0,0000 **	0,0000 **	0,0000 **
C1: when did you last use the internet? Within the last 3 months	0,0000 **	0,0000 **	0,0000 **	0,0000 **	0,0000 **	0,0000 **
*significant at the 5% level; **significant at the 1% level						

## Methodology / results

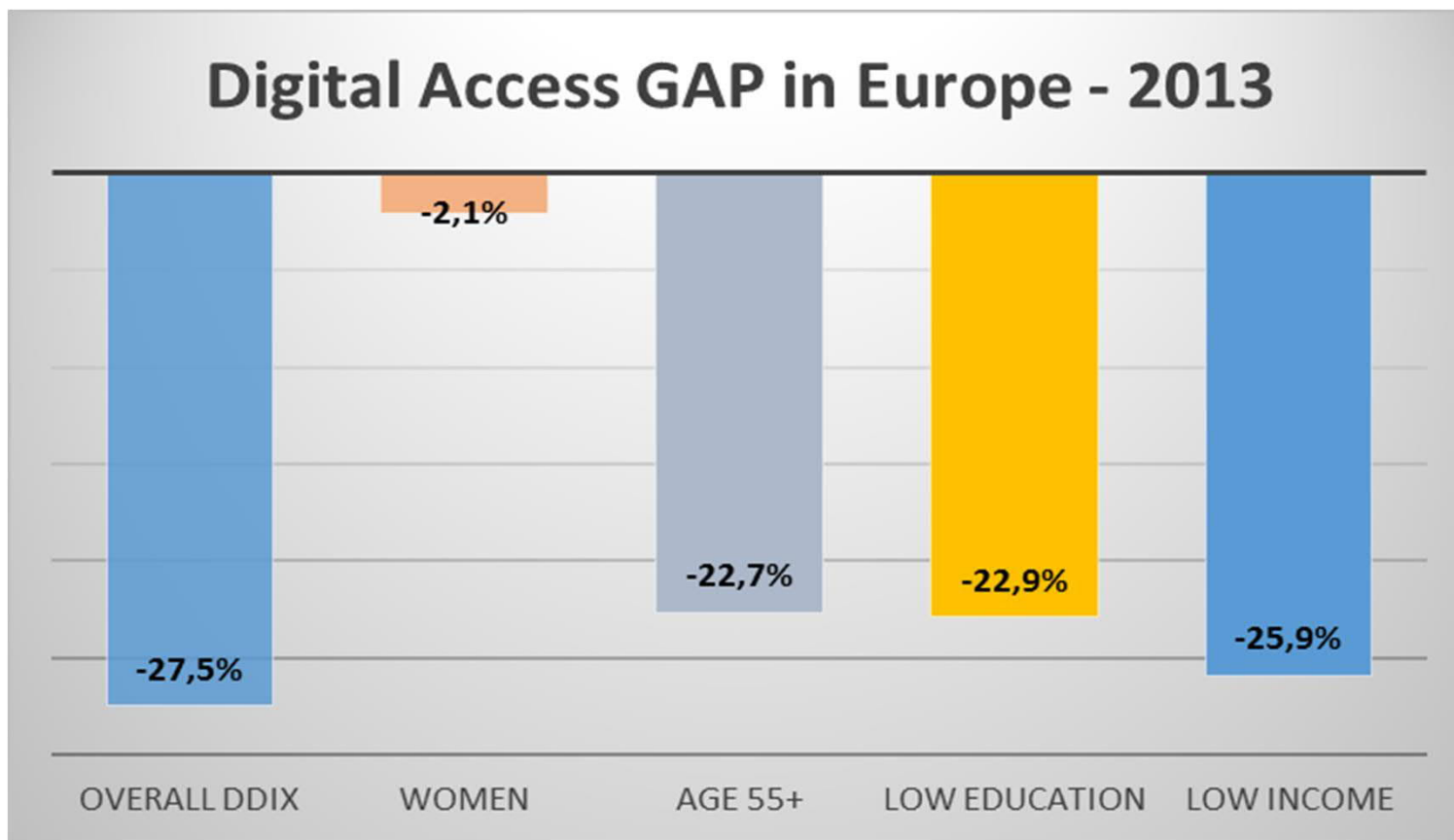
Comparative research: Digital Divide Index (DDIX)



## Methodology / results

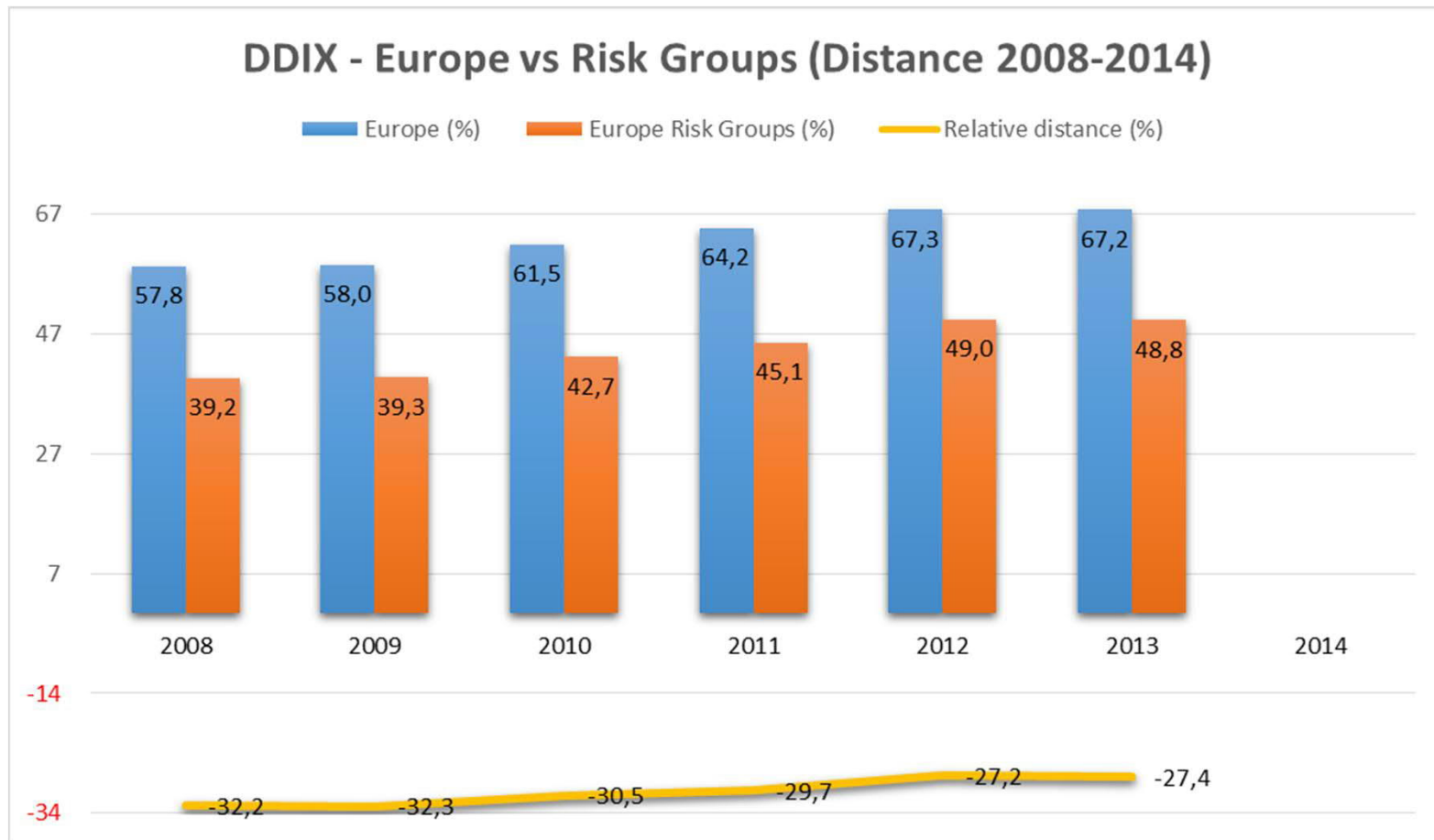
---

Comparative research: Digital Divide Index (DDIX)



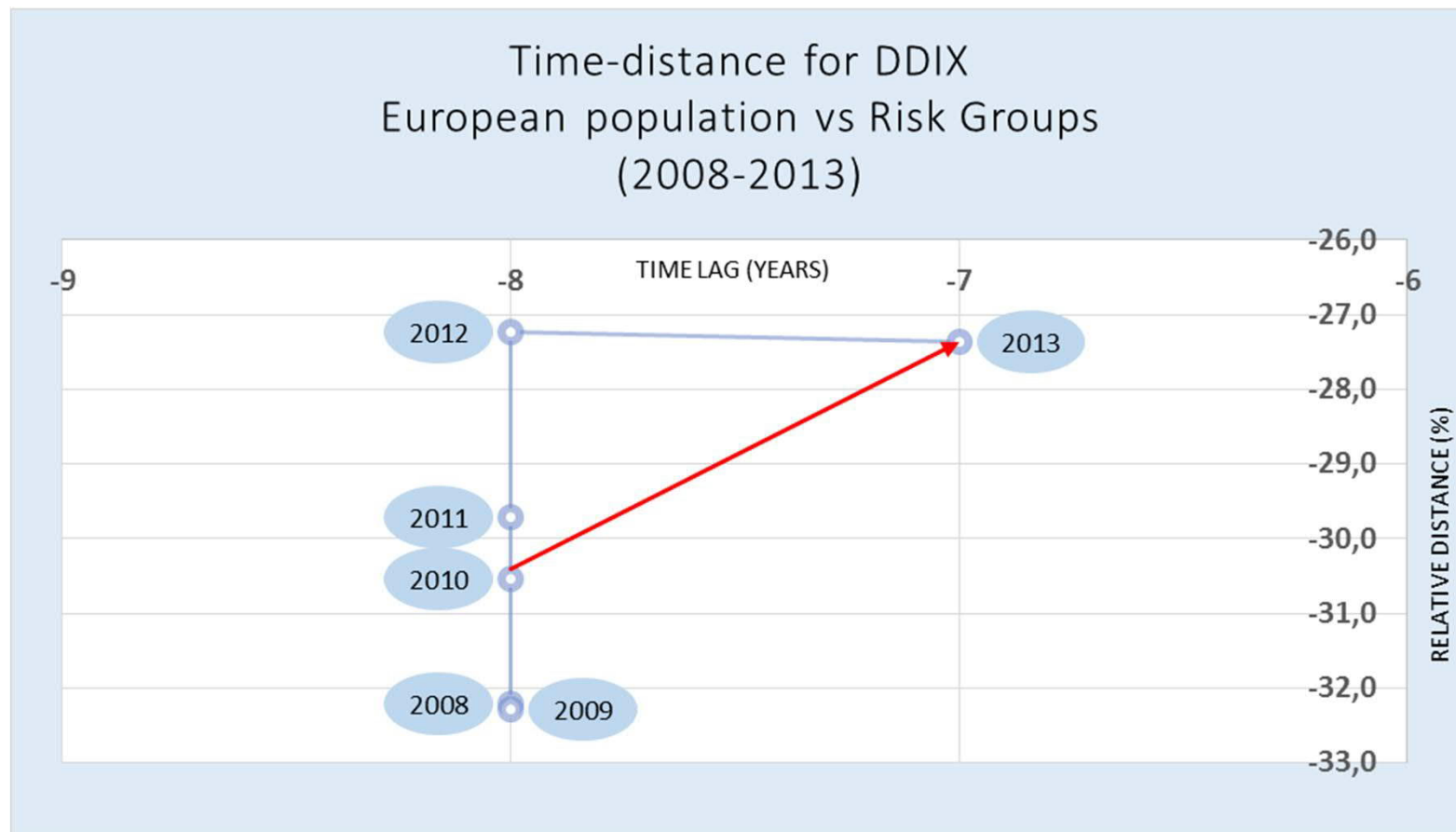
## Methodology / results

Comparative research: Digital Divide Index (DDIX)



## Methodology / results

Temporal dynamics': Time-distance methodology



## Methodology / CATPCA

---

### . CATPCA - Principal Components Analysis for Categorical Data

**Case Processing Summary**

Valid Active Cases	245867921,0
Active Cases with Missing Values	121982705,7
Supplementary Cases	0
Total	367850626,7
Cases Used in Analysis	367850626,7

Cases Weighted by INDIVIDUAL WEIGHT.

**Model Summary**

Dimension	Cronbach's Alpha	Variance Accounted For
		Total (Eigenvalue)
1	,859	4,020
2	,179	1,186
Total	,923 <sup>a</sup>	5,206

a. Total Cronbach's Alpha is based on the total Eigenvalue.

## Methodology / CATPCA

---

**Component Loadings**

	Dimension	
	1	2
HOUSEHOLD INCOME QUARTILE	-,325	,207
AGE CLASSES	,591	,528
SEX	,134	,591
EDUCATIONAL LEVEL	-,589	,343
EMPLOYMENT SITUATION	,646	,560
B1: WHEN DID YOU LAST USE A COMPUTER (AT HOME, AT WORK ANY OTHER PLACE)?	,954	-,162
C1: WHEN DID YOU LAST USE THE INTERNET?	,971	-,167
C2: ON AVERAGE HOW OFTEN DID YOU USE THE INTERNET IN THE LAST 3 MONTHS?	,965	-,173

Variable Principal Normalization.

## Methodology / CATPCA

---

**Component Loadings**

	Dimension	
	1	2
HOUSEHOLD INCOME QUARTILE	-,325	,207
AGE CLASSES	,591	,528
SEX	,134	,591
EDUCATIONAL LEVEL	-,589	,343
EMPLOYMENT SITUATION	,646	,560
B1: WHEN DID YOU LAST USE A COMPUTER (AT HOME, AT WORK ANY OTHER PLACE)?	,954	-,162
C1: WHEN DID YOU LAST USE THE INTERNET?	,971	-,167
C2: ON AVERAGE HOW OFTEN DID YOU USE THE INTERNET IN THE LAST 3 MONTHS?	,965	-,173

Variable Principal Normalization.



## Discussion | Contributions

---

1. The role of digital inclusion and digital literacy;
2. The conceptual and methodological improvement of several social sciences concepts to better understand a renewed public sphere;
3. Advanced methods that are crucial for digital divide measurements (and also for any type and level of divides).
4. Data limitations: no data available for Germany; many missing values for the income variable; delay in the delivery of data (2014-2015)

**João Miguel Aguiar**  
University of Porto  
UT Austin | Portugal Program in Digital Media  
(joaomiguelaguiar65@gmail.com)

**Pedro Campos**  
University of Porto and LIAAD – INESC TEC

**José Azevedo**  
University of Porto

**Thank You!**

**Obrigado!**

This research is supported by  
FCT – Fundação para a Ciência e Tecnologia and UT Austin | Portugal Program  
(SFRH / BD / 52327 / 2013)