Trust in Official Statistics across Europe: Evidence from two waves of Eurobarometer using multilevel models

Serge Allegrezza
(STATEC)
serge.allegrezza@statec.etat.lu

Wolfgang Langer
(Martin Luther University of Halle-Wittenberg)
wolfgang.langer@soziologie.uni-halle.de

Majlinda Joxhe
(STATEC)
majlinda.joxhe@ext.statec.etat.lu

Paper prepared for the 37th IARIW General Conference

August 22-26, 2022

Opening Plenary Session: Trust, Well-being, and Productivity (live stream)

Time: Tuesday, August 23, 2022 [9:15-11:00 CEST]

Trust in Official Statistics across Europe: Evidence from two waves of Eurobarometer using multilevel models

Serge Allegrezza[1]

Wolfgang Langer[2]

Majlinda Joxhe[3]

15/07/2022

Abstract

This paper uses data from Eurobarometer for the years 2007 and 2015 to investigate the determinants of trust in official statistics across 28 countries and 2 regions in Europe. Our estimation follows a multilevel modelling approach, which allows us to distinguish *within*-country and *between*-country variation in individual trust in official statistics.

The econometric results show that trust in statistics is highly correlated with overall individual trust in national institutions. Within-country variation is mainly explained by individual-level statistical literacy and education, as well as occupational status. With respect to the variation between countries, we show that neither the level of GPD nor the index of inequality are important in explaining cross-country variation. Instead, EU membership history, i.e. the '***acquis communautaire***', is the main macro variable that explains the increase in trust in official statistics across Europe.

Keywords: *Trust in Official Statistics, Eurobarometer, Europe, Multilevel Models*

[1] STATEC, Director of the National Statistical Institute of Luxembourg, e-mail: serge.allegrezza@statec.etat.lu

[2] Martin-Luther-Universität Halle-Wittenberg, Institut für Soziologie, Associate Professor, e-mail: wolfgang.langer@soziologie.uni-halle.de

[3] STATEC, Researcher for the National Statistical Institute of Luxembourg, e-mail: majlinda.joxhe@ext.statec.etat.lu

1. Introduction

The literature has recognized three standard forms of trust: ***generalized, particularized and institutional***. Generalized trust usually refers to the type of social trust in which one trusts people they do not know. Particularized trust is trusting only the people one knows, such as family or friends. Institutionalized trust is a society's trust in government institutions. As described by Putnam (1995) when he writes *'features of social organization such as networks, norms, and social trust that facilitate coordination and cooperation for mutual benefit'*, social trust is the basis of social capital.

It is well documented that social trust influences political, social and economic outcomes, such as economic growth (see, among others, Bjornsok 2002). Arrow (1972) highlights that *'Virtually every commercial transaction has within itself an element of **trust**, certainly any transaction conducted over a period of time.'* Furthermore, social trust is also connected to the foundations of democracy and good governance; research has shown that in many countries, a high index of democracy correlates with greater social trust (Warren 2018). In this literature, in which the concept of trust is an important feature of society, certain other types of trust have been less investigated, however.

The notion of ***trust in statistics,*** a concept that connects individual trust to statistical outputs released by national and international institutions, is also an important aspect of social trust. Normally, the statistics produced by recognized institutes such as national statistics offices directly influence the perception of governance in a country. Charpin (2010) writes that *'trust in the institutions is actually grounded on the trust versus the experts that will analyse their quality based on statistical indicators of their governance'*. Moreover, these national experts are seen as a source of reliable, credible and well-timed statistical information and to be free of any political intervention.

Confidence in official statistics finds it foundations in the UN's Chart of Fundamental Rights, in which access to reliable statistical information is considered a human right. In this perspective, trust in statistics is positioned in an intermediate position between trust in institutions and trust in national experts. Given this framework, we claim that it is important to understand the mechanisms that influence individuals' trust in statistics.

In this study, we use data from the Eurobarometer surveys for most European countries to investigate the concept of trust in statistics. Eurobarometer is a series of multi-topic surveys carried out by the European Commission since 1970, covering attitudes towards European

integration, policies, institutions, social conditions, health, culture, the economy, citizenship, security, information technology, the environment and other topics.

We use two waves of the survey in which participants are asked to express their level of trust in official statistics (2007 and 2015). Using data from two different time periods and for all of Europe allows us to tackle both temporal variation and between-country variation. At first glance, the data show that the level of trust in statistics is very heterogeneous across Europe, with Nordic countries (such as Sweden or Finland) recording around 70% of the population expressing trust in official statistics. On the other hand, in countries like Spain, France and Italy, less than 50% of people report trusting official statistics. In addition, trust remains stable from 2007 to 2015[4] for some countries, such as Germany (42% in 2007 and 43% in 2015) but to have changed for others, such as the UK, where trust in official statistics increased from 33% in 2007 to 42% in 2015.

A few empirical studies have investigated trust in statistics and its determinants.[5] Usually, the consumers of statistical information have a variety of knowledge backgrounds as well as different expectations and financial constraints. These latter factors may determine different levels of trust in official statistics.When analysing trust in statistics, it is important to investigate the *'demand view'* to focus on the perspectives of the actual or potential users of statistical information. Access to data and information has changed dramatically with the increase in internet usage, which along with social media is considered the main channel of access to information nowadays, and to statistics in particular.

We investigate the determinants of trust in official statistics by employing an econometric modelling approach. Our estimation method follows a multilevel technique where we pool individual (level 1) and country (level 2) characteristics to explain variation in trust within and across countries. Our estimations show that at the country level, an individual's level of statistical literacy and education correlate significantly with trust in statistics. Individuals who trust their national government or parliament tend to also report a higher level of trust in official statistics.

With respect to variation between countries, our model shows that EU membership history is the main variable that explains cross-country variation. As a country becomes an EU member state, it is shown that there is an important impact on individuals' trust in statistics. We claim that national trust in European institutions (such as Eurostat, through European

[4] The economic crisis of 2008 might have had a big impact of the level of trust in statistics.
[5] See Alesina and Ferrara (2000) for a review of the determinants of trust in general.

statistical harmonization) is one of the main channels that might explain the significant effect of membership to the EU on trust in statistics.

In Sections 2 and 3, we outline the concept of trust and, in particular, the importance of trust in statistics. Sections 4 and 5 describe the data and the estimation methodology. In Section 6 we present our main results, and Section 7 offers some concluding remarks.

2. Institutional Trust and Trust in Statistics

Most of what we consider true or obvious is based on individual trust[6] through intuition. Intuition is a form of what David Kahneman calls fast, or ***System 1***,[7] thinking and we often base our decisions on what it tells us. *'We trust our intuitions even when they're wrong,'* Kahneman writes. However, we *can* trust our intuitions—as long as they are based on real knowledge. In addition, we develop knowledge through experiences that influence our level of trust. Thus, trust is a concept that involves different cognitive and non-cognitive aspects of human behaviour that are important to investigate.

The OECD (2011) formally defines trust as *'a person's belief that another person or institution will act consistently with their expectations of positive behaviour.'* In this perspective, trust is divided into two main forms, **interpersonal and institutional trust**. The concept of **institutional trust** encapsulates trust towards all types of institutions, including governmental and non-governmental institutions and institutions of justice or law. Institutional trust encompasses all types of public institutions, focusing mostly on specific institutions (i.e. the government, parliament, judicial system, police, civil service and national statistical offices). However, empirical analyses also suggest that a grouping can be made in terms trust in three types of institution: (i) political institutions, (ii) law institutions and (iii) non-governmental institutions (Schneider, 2017).

[6] The word 'trust' in Indo-European languages finds its origins in the root 'droust', which means 'solid' and 'lasting', whereas the synonym 'confidence' originates from the Latin word 'confidens', meaning a condition of 'faith' and 'hope' in others. In Greek epistemology, the word 'trust' is has its roots in the word 'pistis', which refers to the personification of 'good faith', whereas in Hebrew languages there are different words that can be translated as trust. For example, the word 'chasah'' means to 'lean on something or someone'. Instead, the word 'vera' in Slavic languages identifies 'faith' and 'believing' in something and is related to the word trust. In all of these languages, the word 'trust' has a very similar meaning in which relationships and faith between people is crucial. These old meanings of the word have also been well preserved to the present day, where trust (or confidence, psitis, chasah, etc.) is shown to be an important mechanism for social interactions and a key component in interpersonal relationships within families, between friends or within organizations and countries.

[7] Daniel Kahnemann (2013) *Thinking, Fast and Slow*: https://us.macmillan.com/books/9780374533557

Fukuyama (1995), one of the first authors to explore trust in institutions, considers that if a society has a limited radius of trust, referring to *'the circle of people among whom cooperation and mutual understanding exist,'* it is then classified as **'low trust'**. In this context, people end up trusting only those with similar characteristics, for instance, in terms of ethnicity. In contrast, in societies with a large radius of trust, or **'high trust**,' citizens develop more trust in the public sphere in social institutions, as they actively engage with other individuals.

Lipset and Schneider (1983) underscore that individual trust may function mostly as a predictor of **political trust**, in which case social capital becomes a source for institutional outcomes. An empirical study using data from the World Values Survey (WVS) conducted by Newton and Norris (2000) found a strong correlation between social (individual) and political trust. According to the authors, trust in governmental institutions is shaped through social relationships, and in this way, they also affect institutional performance.

In the political science literature one important theory focuses instead on the **institution-centred approach**, according to which political institutions shape social capital in the societies in which they operate. According to this theory, citizens are systematically affected by the behaviour of politicians: the existence of efficient and legal institutions makes a person less likely to believe that most other citizens engage in behaviour perceived as unfair. An empirical study by Kaase (1999) using data from the Eurobarometer and European/World Values Surveys for a set of European countries contests that interpersonal trust is due to the level of political trust, finding a positive but small statistical relationship between the two.

More generally, trust is embedded in a system of deep social preferences and beliefs, such as altruism, reciprocity and aversion to inequality, which partly shape attitudes towards institutions and society, including interpersonal trust and trust in institutions. Experts, including **official statisticians**, are the main protagonists in institutional trust in a society. Statistics has acted as a well-known channel of trust in modern societies, and the experts behind them—scientists, economists, mathematicians and statisticians—play a key role as their curators and defenders. As experts, they command trust and in turn provide numbers that others use to build their beliefs.

Official statistics may often appear highly abstracted from lived realities, but they are based on the legitimacy of statistical classifications and indicators such as GDP or employment rates traditionally used to represent demographic, social and economic changes. This legitimacy is required for individual citizens to trust institutions such as the government or parliament. The concept of trust in official statistics and the institutions that produce them involves an interaction between interpersonal and institutional trust, and showing confidence in official

data serves as a chain to connect the individual to the society in which he lives. Advanced and developing economies are nowadays very much concerned about trust in official statistics, as it may influence social and political actions.

3. The Importance of Trust in Official Statistics

Trust in official statistics is mostly understood in terms of confidence in the outputs of national statistical offices (NSOs), which operate as providers (suppliers) of statistical data.[8] More explicitly, NSOs are seen to be a source of reliable, credible and timely statistical communications that must be free of any political intervention. Statistical institutions, and in particularly NSOs and other international organisations, are essential as they provide important information to inform other decision-making organisations. More specifically, official statistics can affect large numbers of important decisions across governments and across countries. The main advantage of official statistics is that the principles of collecting and publishing data are consistent, being built on the *Fundamental Principles of Official Statistics* (UN FPOS),[9] and in this way, they can be seen to be trustworthy and reliable source of data and information. In Europe, Eurostat recommends national statistics offices to adhere to the *European Statistics Code of Practice*[10] in order to ensure consistent quality of statistical output among European countries. The code has the main objective of assuring independence, quality and impartiality in data dissemination among European countries.

Distrust in statistics is due to several reasons.[11] The first is certainly rooted in the awkward representation of society as a whole, structured around Quetelet's (1835) *'average man'* wiping out the uniqueness of individual characters, particular situations, contexts and biographies inherent in a single person. Second, statistics is a branch of mathematics and thinking statistically is difficult, as Nobel Prize winner Daniel Kahneman demonstrated through multiple experiments described in his famous book *Thinking, Fast and Slow*. Third, it is natural to recognize that trust in official statistics is linked to trust in science, and the attitudes of individuals towards science are potentially relevant for statistical analyses. There appears to be a general concern about the growing distrust in science, despite the increasing level of education throughout recent years.

[8] See the Paris 21 report 'Building Trust in Data' (2018)
[9] https://unstats.un.org/unsd/dnss/gp/FP-New-E.pdf
[10] https://ec.europa.eu/eurostat/web/quality/european-statistics-code-of-practice
[11] A recent report by France Stratégie (Agacinsky 2018), an official think tank, presents some recommendations discussed at length in reference to the French case.

A study conducted in the US by Drummond and Fischhoff (2017) analyses factors predicting belief in various topics such as climate change, human evolution, etc., and shows that the education level and knowledge of science of participants are weakly correlated. Nevertheless, respondents possessing higher levels of science literacy are more likely to agree with the scientific consensus. Moreover, the study finds that American public opinion is polarized along religious and political lines and that this polarization increases with the level of education. It is interesting to see quantitatively to what extent limitations in literacy and numeracy and the resistance to science are increasing[12,13] and how much this matters for trust in official numbers.[14]

The effort to increase trust in statistics has a long and painful history.[15] Many events have demonstrated the importance of individual trust in official statistics; for example, a survey in Japan found that almost 80% of people lost trust in the government's economic indicators when they reported false data about wages. In this framework, there is a double challenge for institutions disseminating statistical information: (i) to expand the integrity and relevance of official statistics but also (ii) to increase public confidence (trust) in official numbers. Given this two-sided challenge, in this paper we try to detect the determinants (socio-demographic individual characteristics or other macro variables like GPD or the inequality index of a country) that are relevant for shaping the level of trust in official statistics. We use both individual data collected by the Eurobarometer surveys and macro variables at the country level to detect the influence of both on trust. Our estimations bring some new results with respect to the literature on trust in statistics among European countries.

4. Data from Eurobarometer on Trust in Official Statistics

We use data from Eurobarometer (EB) Surveys **No. 67.2 and 83.3** conducted in 2007 and 2015.[16] The surveys are administered in all member states of the European Community using

[12] 'Statistics in a post-truth society. Determinants of confidence, independence and usage of official statistics'

[13] The Eurobarometer survey (2010) suggests that reluctance towards science is significant (in 2010, 38% reported thinking we depend too much on science and not enough on faith), but this proportion is decreasing.

[7] An OECD study (2016) on literacy and numeracy provides some perspective regarding information-processing skills and competences. The survey conducted from 2013 to 2015 in 33 countries showed that a significant proportion of adults have insufficient reading and numeracy skills (22.7%, on average). One in four adults have no or limited computer experience or lack confidence in their ability to use a computer. Literacy and numeracy skills peak around the age of about 25.

[15] In his book *Trust in numbers,* Theodor Porter (1996) makes it clear that official statistics cannot be properly understood if they are not examined through the lens of the history of science. As Porter shows, including in the engineering of official statistics, 'mechanical objectivity' is difficult to achieve fully because tacit knowledge, experience, wisdom, intuition, skills and craft play an important role in scientific activity.

[4] We use these two waves as the question regarding trust in statistics is included only in these two waves.

the same definition of target population,[17] with a standard sample size of 1,000 to 1,500 for large states and 500 respondents for small states. The national target population consists of national inhabitants aged at least 15 years of age. In order to create a representative sample, Eurobarometer uses a multistage random/probability sample to cover the regional distribution of the population. The interviews were conducted in all member states of the EU in the same period (EB67.2: April 10 to May 15, 2007; EB83.3: May 16 to 27, 2015).[18] In addition to the classical question related to trust in national and European institutions, Eurobarometer also introduced a question about individual trust in official statistics. The question captures a binary outcome, with respondents invited to *reply if they tend or do not tend to have trust in official statistics.* As initial statistics, we report the overall heterogeneity across Europe. In Figure 1 and Figure 2, we map the percentage of Europeans replying that *they tend to trust the official statistics* in 2007 and 2015, respectively.

As clearly evidenced, there is great heterogeneity among European countries, with Sweden, Finland and Luxembourg showing high percentages of the population (around 70%) declaring having trust in official statistics in the year 2007.[19] We also detect temporal variation when comparing the 2007 and 2015 surveys. Some countries show very high variability in terms of trust in official statistics across years, an example being Germany, where an average of around 42% of the population reported having trust in official statistics (for 2007 and 2015). In contrast, in the UK in 2007 only 33% of the population reported having trust in official statistics, but by 2015 this percentage had increased to 44%.

[17] Eurobarometer has the great advantage of using the same sampling frame and method, the same data collection method and nearly the same questionnaire. Data collection is conducted at the same time points for the two cross-sectionals and by the same public opinion research institute.

[18] The field work was conducted by TNS Opinion & Social, which is a consortium created between Taylor Nelson Sofres and EOS Gallup Europe.

[19] There is a 10% decrease for the year 2015.

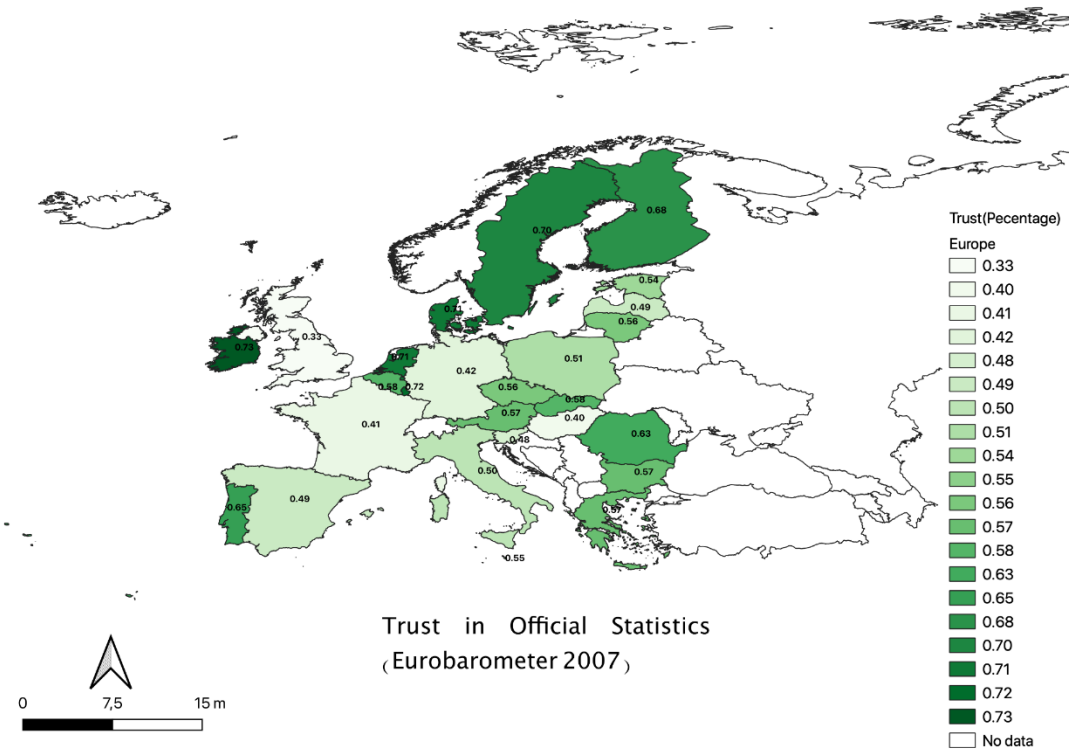

Figure 1: Europe: Trust in Official Statistics 2007

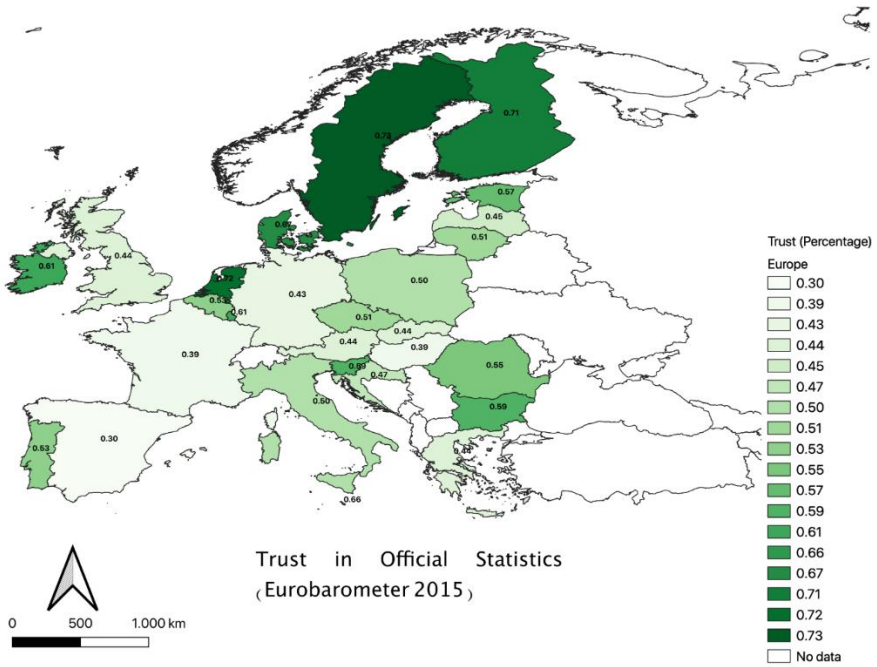

Figure 2: Europe: Trust in Official Statistics 2015

5. Theoretical Framework and Multilevel Modelling

We make use of the repeated survey using a binary multilevel logit model. For the within-country analysis, our model uses the exogenous variables to explain respondents' level of trust in official statistics. For the between-country part, the model uses exogenous variables measured at the national member-state level to predict the variation between time points and member states.

The multilevel analysis explicitly models the manner in which individuals are grouped within areas (such as regions, countries or schools) and has several advantages. Firstly, it allows analysts to obtain statistically efficient estimates of regression coefficients.[20] Secondly, by using the clustering information it provides correct standard errors, confidence intervals and significance tests, and these are generally more 'conservative' than the traditional ones obtained by ignoring the presence of clustering.[21] Thirdly, by allowing the use of covariates measured at any of the levels of a hierarchy, it enables the researcher to explore the extent to which differences in average results between groups (in our case, countries) can be explained by factors such as institutional and historical settings. For example, a study by Sitkin et al. (1998) focus their analysis explicitly on the micro-level and macro-level mechanisms that shape individual trust intentions.

For these reasons and given the nested structure of our data, we make use of a multilevel model (see Figure 3).

[20] On cluster-corrected robust standard errors, see Hox (2010: 260-263) and Raudenbush and Bryk (2002: 276-278).
[21] Just as Bennett's (2015) previously statistically significant results became non-significant upon reanalysis.

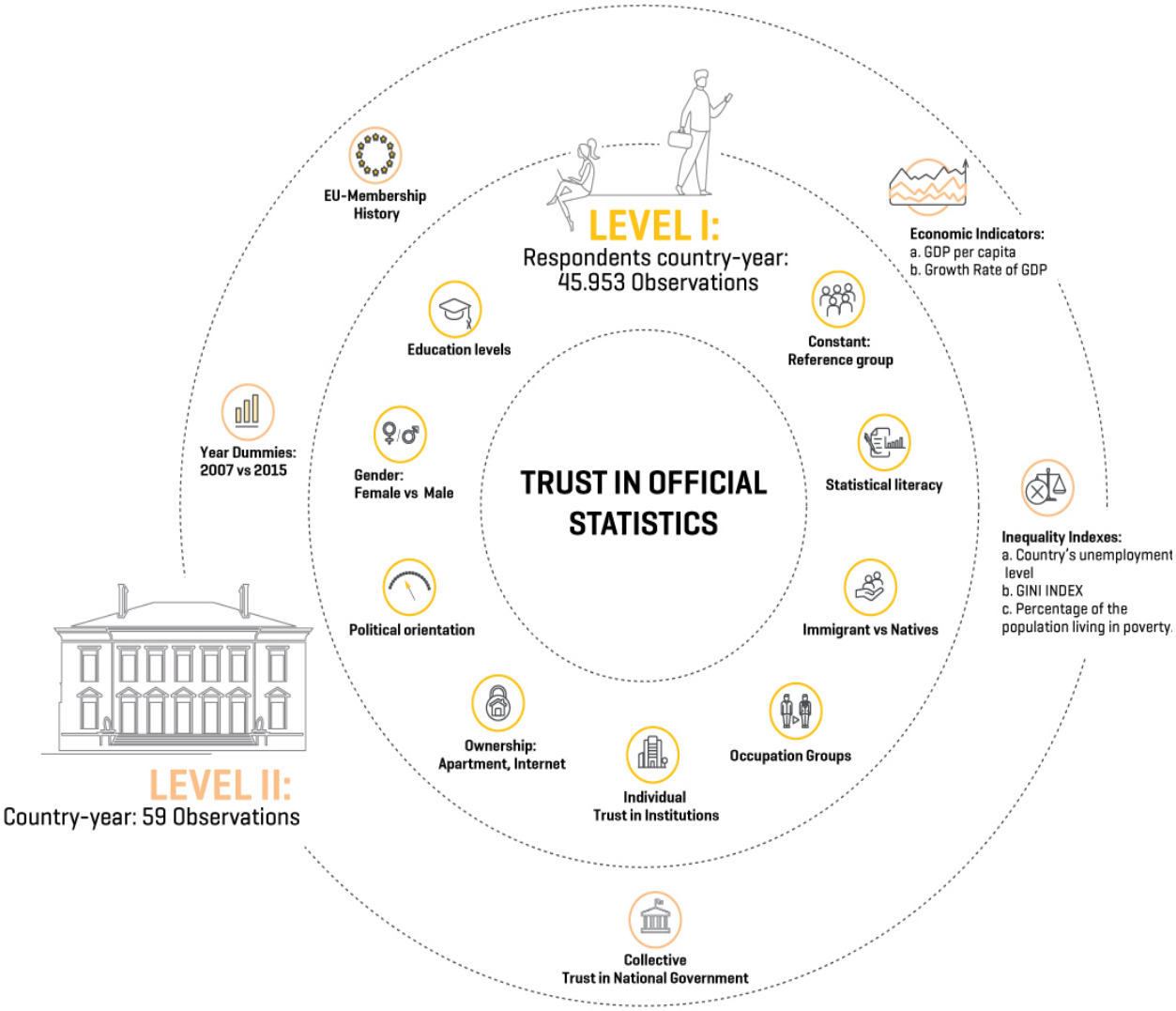

Figure 3: Multilevel Modelling to explain variation for Trust in Official Statistics in Europe

We estimate the following multilevel logit model:

$$Logit\ P(Y = 1) = \beta_{0j} + \beta_{ij}X_{ij} + \beta_j Z_j + u_{0j} + \epsilon_{ij}$$

where:

$\boldsymbol{Trust_{ij}} = dummy\ outcome\ of\ trust\ for\ individual\ i\ residing\ in\ country\ j,$
$0 = tends\ to\ not\ trust\ official\ statistics\ and\ 1 = tends\ to\ trust\ official\ statistics;$
$\boldsymbol{\beta_{0j}} = Expected\ logit\ of\ reference\ group\ at\ the\ grand\ mean\ of\ exogenous\ macro\ variables\ reference\ context\ for\ level\ 1\ and\ level\ 2;$
$\boldsymbol{\beta_{ij}} = logit\ slope\ of\ exogenous\ level\ 1\text{-}variables\ (X\_ij);$
$\boldsymbol{X_{ij}} = individual\ charateristics\ i\ for\ each\ country\ j\ (level\ 1);$

$\boldsymbol{\beta_j} = logit\ slope\ of\ exogenous\ level\ 2\ variables\ (Z);$
$\boldsymbol{Z_j} = exogenous\ level\ 2\ variables;$
$\boldsymbol{u_{0j}} = Country\text{-}specific\ random\ effect\ of\ the\ intercept;$
$\boldsymbol{\epsilon_{ij}} = Level\ (1 - Residual)\ error, assuming\ a\ mean\ of\ zero$
$and\ a\ variance\ of\ pi\text{^}2/3.$

This econometric model based on a multilevel approach allows us to distinguish how much of the individual characteristics contribute to explaining the variation of individual trust in official statistics within each country. A few studies analyse the factors that might influence individual trust in statistics, and here, we review some studies for trust in institutions in order to select covariates to use in our regressions. Our hypotheses are that trust is statistics is not influenced by gender or political orientation and that statistics are understood by everyone, independent of education level or geographical location, because statistics should ideally be a public good.

Studies at the individual level, such as Glaeser et al. (2000), use probit and OLS models show how trust in institutions varies (across the United States in their case). They show that trust 'is much lower for later cohorts' but 'much higher among richer and well-educated individuals'. They also find that males are more trusting than females. However, the study found higher levels of trust in institutions among members of 'more educated (or wealthier) religious denominations'. Another study by Guiso et al. (2003) used data from the World Value Survey referring to around 66 countries and established that some of the findings in the US study by Glaeser et al. (2000) could be generalized to this larger set of countries. By using a country fixed effects estimation, the study found that individuals' health, age, social status, income level and religious beliefs show positive and significant relationships with trust in institutions. Finally, the study by Wang and Gordon (2011) using a multilevel model with a global data set from the World Value Survey—a study with an approach very similar to ours—shows that at the individual level, age displays a U-shaped relationship with self-reported trust. Put differently, the stylized finding is that individuals gradually show less trust in the early and middle stages of their lives, with a deep decline until around age 30 followed by a gradual increase thereafter. People in their late 40s or older tend to trust others more. Male interviewees also show more overall trust in institutions than female interviewees. The study also found that religious denomination is crucial and explains much of the variation in trust. But for individual-level religious beliefs, the results indicate that individual beliefs in Muslim, Protestant or Catholic religious tenets is positively and significantly related with individual trust levels, although this is not the case for Orthodox denominations.

Taking into consideration all these studies, to explore individual variation we use exogenous socio-demographic variables such as age group, gender, education (age at which respondents left school) and dummies for occupation status (to capture social status, as in Glaeser et al. 2000). We also use dummies to distinguish whether a respondent is native or an immigrant, if he owns an apartment or if he has an internet connection. In addition, we include one's individual level of trust in the national government and their political orientation, which have both been shown to be very relevant for determining individual trust in official statistics. Finally, we construct an index of statistical literacy, with a variable that is generated as the sum of all correct answers to questions about economic indicators. The index ranges from 0-3, with 3 being the maximum number of correct answers that could be given by respondents. In order to separate false and 'don't know' answers, we use a dummy variable indicating at least one 'don't know' or refusal answer for the three questions. Tables 1 and 2 show average summary statistics for each individual characteristic for all European countries.

The tables reveal an important historical change in the trend regarding trust in institutions, including trust in official statistics. The decline in trust in public institutions experienced by several countries since the 2008 financial crisis has been a source of serious concern (Foster and Friden 2017). Indeed, trust in a broad range of public and private institutions fell the most in OECD countries hit hardest by the crisis (i.e. those that experienced the largest falls (or the smallest growth) in household income and earnings since 2005, as well as some of the largest increases in long-term unemployment; OECD 2017a). This decline in trust (which in some countries spanned several decades) has gone hand in hand with an increase in non-mainstream voting and populism in several countries (Inglehart and Norris 2016; Algan et al. 2017).

Table 1: Summary Statistics (baseline level 1), Eurobarometer Europe (2007)

	Mean	S.D.	Min.	p50	Max.	Count
Trust in Official Statistics (%)	**0.56**	0.50	0.00	1.00	1.00	22,317
Statistical literacy, sum of correct answers (%)	0.46	0.71	0.00	0.00	3.00	22,317
Statistical literacy, at least one non-response (%)	0.59	0.49	0.00	1.00	1.00	22,317
Left-right political orientation	3.54	1.52	1.00	3.00	6.00	22,317
Trust in national government (%)	0.47	0.50	0.00	0.00	1.00	22,317
Trust in national parliament (%)	0.49	0.50	0.00	0.00	1.00	22,317
Age groups (education)	2.27	0.89	1.00	2.00	5.00	22,317
Occupation scale	4.83	2.17	1.00	4.00	8.00	22,317
Male (%)	0.45	0.50	0.00	0.00	1.00	22,317
Age groups (6)	3.72	1.68	1.00	4.00	6.00	22,317
Immigrant vs native (%)	0.02	0.14	0.00	0.00	1.00	22,317
Type of community: village, small or large city	1.90	0.79	1.00	2.00	3.00	22,317
Ownership: apartment paid (%)	0.52	0.50	0.00	1.00	1.00	22,317
Ownership: apartment paying (%)	0.25	0.43	0.00	0.00	1.00	22,317
Ownership: internet connection (%)	0.49	0.50	0.00	0.00	1.00	22,317

Note: Average values for all countries in the panel

Table (2): Summary Statistics- (baseline level 1) Eurobarometer Europe (2015)

	Mean	S.D.	Min.	p50	Max.	Count
Trust in Official Statistics (%)	**0.51**	0.50	0.00	1.00	1.00	23,636
Statistical literacy, sum of correct answers (%)	0.35	0.55	0.00	0.00	3.00	23,636
Statistical literacy, at least one non-response (%)	0.37	0.48	0.00	0.00	1.00	23,636
Left-right political orientation (%)	3.45	1.50	1.00	3.00	6.00	23,636
Trust in national government (%)	0.36	0.48	0.00	0.00	1.00	23,636
Trust in national parliament (%)	0.35	0.48	0.00	0.00	1.00	23,636
Age groups (education)	2.35	0.84	1.00	2.00	5.00	23,636
Occupation scale	4.85	2.17	1.00	5.00	8.00	23,636
Male (%)	0.46	0.50	0.00	0.00	1.00	23,636
Age groups (6)	3.94	1.66	1.00	4.00	6.00	23,636
Immigrant vs native (%)	0.02	0.15	0.00	0.00	1.00	23,636
Type of community: village, small or large city	1.97	0.76	1.00	2.00	3.00	23,636
Ownership: apartment paid (%)	0.50	0.50	0.00	1.00	1.00	23,636
Ownership: apartment paying (%)	0.26	0.44	0.00	0.00	1.00	23,636
Ownership: internet connection (%)	0.74	0.44	0.00	1.00	1.00	23,636

Note: Average values for all countries in the panel

However, individual-level studies only measure individuals' socioeconomic characteristics and their impact on trust. To consider also national characteristics, previous empirical studies usually aggregated individual-level data to the national level. The study by Knack and Keefer (1997) aggregated World Value Survey data for 29 market economies and found that *'trust and civic norms are stronger in nations with higher and more equal incomes, and with better educated and ethnically homogeneous populations'*. In another paper, Zak and Knack (2001) used data for 41 countries to explore the relationship between trust and growth and found that *'institutions affect growth via their impact on trust'*. These authors' findings confirm that '*Trust is higher in more ethnically, socially, and economically homogenous societies, and where legal and social mechanisms for constraining opportunism are better developed, with high-trust societies exhibiting higher rates of investment and growth'*. In a country-level study, La Porta et al. (1997) conducted OLS estimations on World Value Survey data and reported validation of Putnam's (1993) and Fukuyama's theories. Their study stated that higher trust in institutions had a positive causal effect on the judicial system by lowering corruption and increasing bureaucratic quality. Tabellini (2010) also used European regions as the units of analysis and documented *'that both GDP per capita and growth are higher in those regions that exhibit higher levels of "good" cultural values like trust, beliefs in individual effort, generalized morality, and low obedience'*.

Following these studies, as level 2 macro variables we include GDP growth rate and GDP per capita, an index of inequality (GINI), the poverty rate and the unemployment rate for 2007 and 2015. We also include a macro variable that identifies collective trust in the national government and the parliament. Finally, in contrast to most of the previous studies, we also include time dummies referring to the entrance of countries into European Union.

To get an idea of how much trust in official statistics varies between member states and time points, Figure 4 shows the fraction of national respondents who tend to trust official national statistics.[22] The red horizontal line represents the grand mean of trust in official statistics. Independent of time and nation, 54% of respondents tend to trust official statistics.

[22] The data has been weighted by w1 the weight result from target to get representative results. Member states have been sorted according to their EU membership history, starting on the left side with the founding states from 1957 and through to the accession of Croatia in 2013.

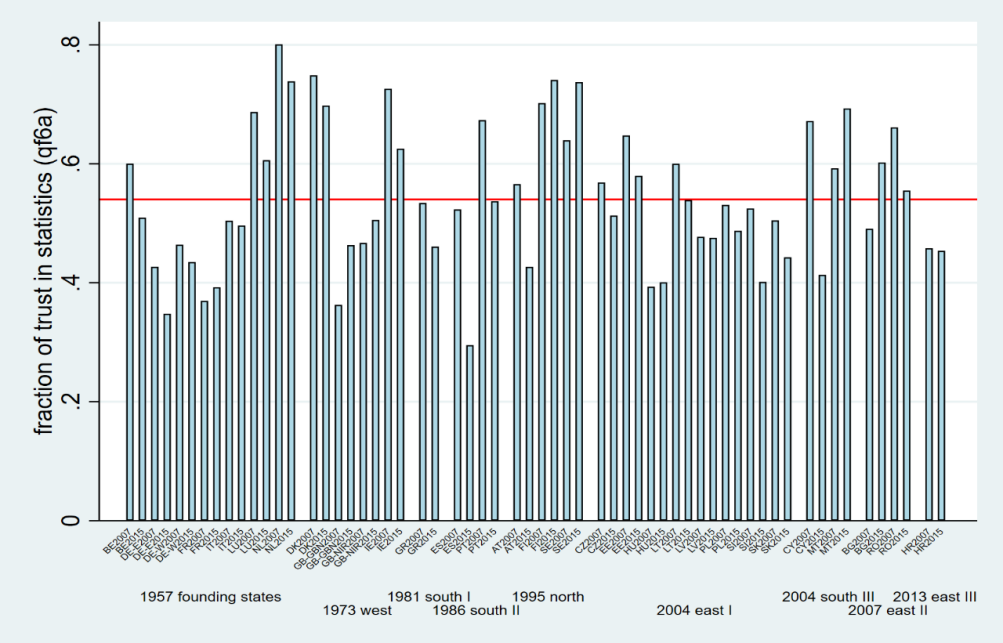

Figure 4: Fraction of population that trusts official statistics in EU member states and their membership history

6. Results of the Baseline Estimations

As a first step, we use the random-intercept-only logit model to assess the impact of the context of state membership and time points. The model estimates the intra-class-correlation (ICC), showing us in percentage terms how much the variation in trust in statistics can be maximally explained by the contextual units—the countries.

Table 3: RIOLM (random-intercept-only logit model) of trust in official statistics across Europe

	Trust in Official Statistics			
Constant	0.176**			
	(2.63)			
Country Var. (Constant)	0.255***			
	(6.04)			
N	45,953			
Level	**ICC**	Std. Err.	[95% Conf.	Interval]
Country	.0718314	.0110367	.0529832	.0966999

As shown in Table 3, the null model indicates that 7.2% of the variation in individual trust in statistics is due to the variation between countries. We see this as an important indicator indicating the need for the multilevel model.[23]

6.1 Baseline Modelling Including Level-1 Variables

Our first baseline model uses the exogenous level-1 variables measured at the individual level and explains the variation in trust in official statistics within countries and time points. We use a random-intercept logit model for our analysis. The reference group consists of 15-24 year-old native males who are manual workers without statistical literacy. They have a liberal political orientation, do not trust the national government or parliament and live in a small or mid-sized town. They do not own a house or apartment and have no internet connection. They left school at the age of 15.

To assess the effect sizes of the exogenous level-1 variables, we calculate their average marginal effects (AMEs)[24] as the deviation from the estimated probability of the reference group having trust in statistics. For the reference group, we calculate an average probability of 42.24% of trusting official statistics, which is significantly different from zero.

Respondents' trust in democratic institutions—the national government and parliament—is associated with a very high and significant increase in trust in official statistics. Both trust in government and trust in parliament lead to an increase of more than 16% in trust in official statistics.[25] In addition, the coefficient of statistical literacy indicates that being correctly informed about macro-economic indicators such as the GDP growth rate, inflation rate and unemployment rate corresponds with a significant increase in trust in official statistics (of around 11%). Political orientation is not significantly related to trust in statistics, whereas refusing to report one's personal political orientation is significantly negatively associated with trust in statistics.

Possessing a higher degree (tertiary education) is positively (but slightly) associated with trust in official statistics (3.9%), and having no full-time education is negatively associated with trust in official statistics (–7.1%). With respect to occupation status, compared

[23] Heck and Thomas (2009: 21) propose a 5% rule of thumb for the ICC to identify a substantial contextual effect and to justify the use of the multilevel model.

[24] We estimated our multilevel logit model and the average marginal effects of the exogenous variables using Stata 16. To plot the latter, we have used coefplot.ado, written by Ben Jann (2014).

[25] For a robustness check, given that trust in government and parliament can be highly correlated with trust in statistics, we have run regressions excluding both variables. Results remain the same and the magnitude is not significantly reduced. Results are available upon request.

to manual workers managers show higher trust in statistics (+2.7%), while being unemployed reduces trust in statistics by 6.3%.

With respect to demographic variables, we do not find significant gender differences. Age has a negative impact on trust in statistics. In comparison to the reference group (young individuals), trust in statistics decreases continually, up to –8.2 % for the 65 years and older group. Immigrants show higher trust in statistics compared to natives (+6.3%). Living in a rural area or a large town has no impact on trust in statistics. We use ownership of a house or apartment as a proxy indictor of socio-economic status. Owning or currently paying for a house or apartment significantly increases trust in statistics by 1.6%. We also use having an internet connection at home, as a proxy indicator of the willingness to inform oneself, and this leads to a significant increase in trust of 1.7%.

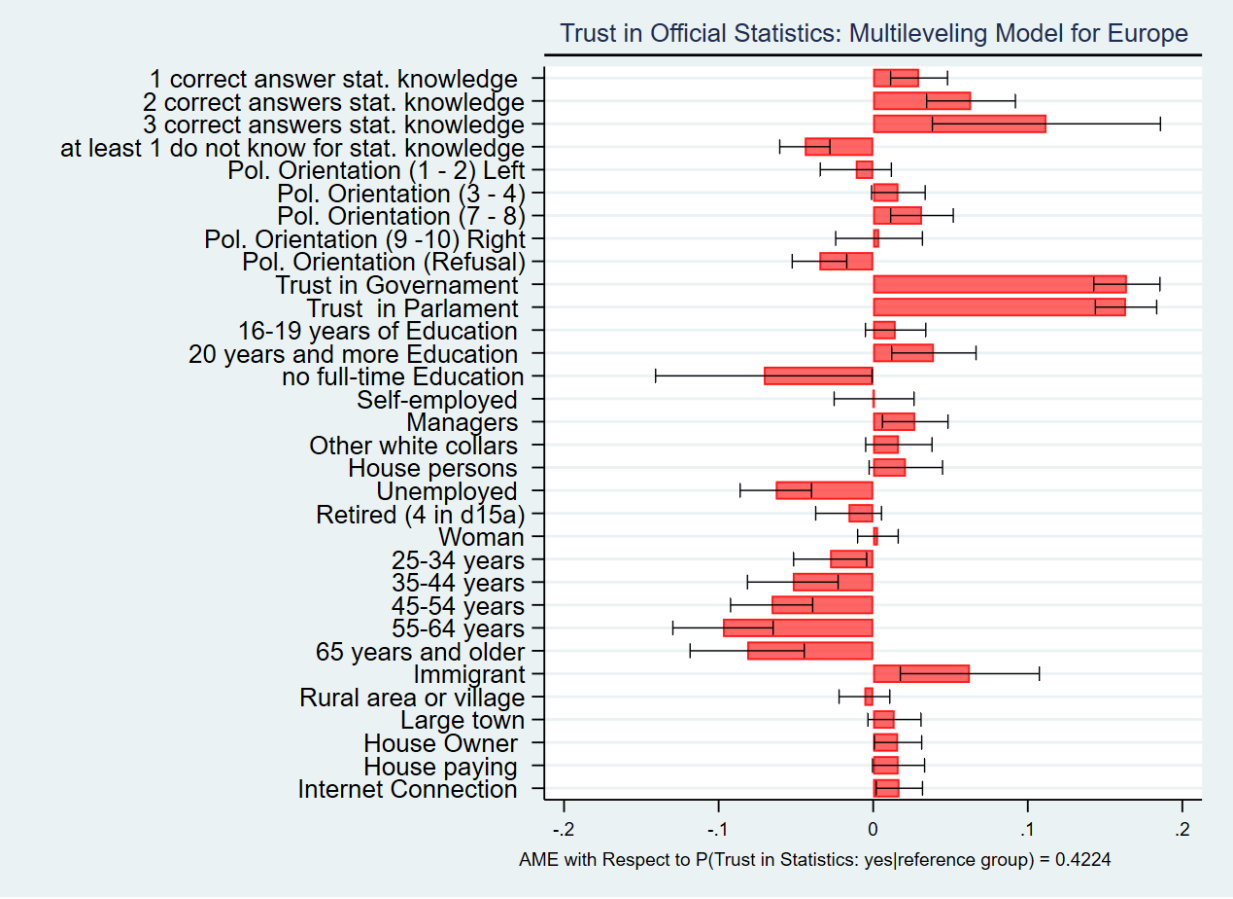

Figure 3: Trust in official statistics: multilevel model for Europe

***McKelvey and Zavoina pseudo-R^2 (fixed and random effects) = 0.1860**

***McFadden pseudo-R^2 (fixed and random effects) = 0.1109**

***LR-chi^2 test statistics (34) = 7030.73 Prob.> chi^2=0.0000**

6.2 Baseline Model Including Level-2 Variables

To explain the variation in trust in statistics between EU member states, we estimate an intercept-as-outcome logit model. In the *between* part, the exogenous level-2 variables explain the variation between states, while in the *within* part the exogenous level-1 variables explain the variation within EU member states. To explain the between variation, we use the collective trust in two national democratic institutions, government and parliament. For both variables, we find a very high correlation coefficient of +0.95 for level 2.

Therefore, we run a principle component analysis for the national means of both indicators and we merge their factor scores to the individual data set for use in further steps of our analysis.[26] To measure the economic wealth of member states, we use macro indicators such as the economic growth rate and the GDP per capita ratio in purchasing power standards (PPS) centred at the EU average, fixed at 100.[27] To measure social inequality between states, we use three indicators: the unemployment rate, the percentage of households living in poverty or marginalization and the Gini index of income distribution. At the state level, the Gini index correlates strongly with the unemployment and poverty rates; therefore, we exclude the Gini index from the set of exogenous level-2 variables. All metric indicators are z-standardized, and their marginal effects measure their magnitude by an increase of one standard unit. We also introduce the EU membership history of states on level 2, starting with the founding states in 1957 and ending with the accession of Croatia in 2013. Corresponding to the big enlargement in 2004, we differentiate between eastern and Mediterranean states. A dummy variable also measures whether change in time has taken place between the years 2007 and 2015.

To assess the fit of the multilevel logit model, we include the McKelvey and Zavoina and the McFadden pseudo-R^2 (see Langer 2020). The likelihood-ratio chi^2 test rejects the hypothesis that all fixed and random effects are zero in the population. The Bryk and Raudenbush level-2 PRE-R^2 shows us that our exogenous level-2 variables explain about 29.43% of the between-country variation in trust in official statistics.[28]

[26] The factor scores are z-standardized.
[27] We use the series prc_ppp_ind, nama_gdp_p, tec00114, ilc_di12, ilc_peps01-1 from the Eurostat database. https://ec.europa.eu/eurostat/web/main/data/database
[28] On problems regarding the assessment of the fitting estimation for multilevel models, please refer to Langer (2020) 'How to assess the fit logit models in STATA?' Swiss STTATA conference, University of Bern.

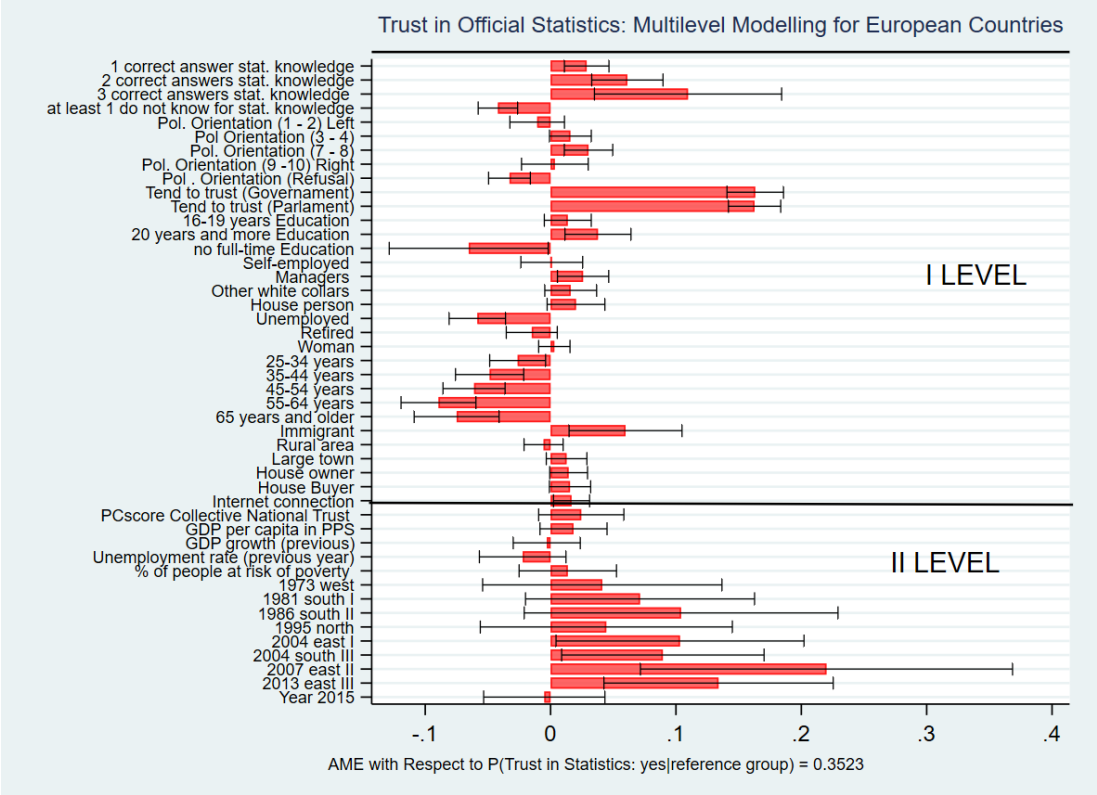

Figure 4: Trust in official statistics: multilevel model for European countries

***McKelvey and Zavoina pseudo-R^2 (fixed and random effects) = 0.1859**
*** McFadden pseudo-R^2 (fixed and random effects) = 0.1112**
**** LR-chi² test: H0: all fixed and random effects are zero in the population***
LR-chi² test statistic (48) = 7050.56 Prob. > chi² = 0.0000
****Bryk and Raudenbush level-2 PRE-R^2 = 0.2943***

We then look at the between-EU-member-state part of our multilevel logit model. Our exogenous level-2 variables should explain the variation in trust in official statistics between EU member states. The national wealth indicators show that neither the gross domestic product per capita nor the growth rate of GDP compared to the previous year have a statistically significant impact on trust in statistics. Even the two indicators of social inequality, the unemployment rate and the poverty rate, have no significant effect on trust in statistics. Collective trust in democratic institutions is associated with an additional increase in its individual counterpart (trust in statistics) of 2.4%, but this is not statistically significant. In contrast to these findings, EU membership history plays a strong role for the variation in trust in statistics between EU member states.

In comparison to the average founder state from 1957, trust in statistics increases significantly, by 10.3% on average, for the Eastern members that joined in 2004, by 9.0% for Malta and Cyprus, which joined in the same year. By 22.0% for Bulgaria and Romania, which became members in 2007, and by 13.4% for Croatia, which entered the European Union in

2013. Comparing the cross-sectional of the year 2015 with the national samples of 2007, we do not find a significant difference.

6.3 Interpreting the results

Our estimations reveal important policy implications for national and international statistical offices that are keen to understand the influence of individual trust on their governance and performance. Our analysis has uncovered that at the individual level, certain characteristics play a fundamental role in trust in official statistics. The level of education and statistical literacy (as measured by the sum of correct answers for simple statistical-knowledge questions) have a positive effect on trust in statistics. Occupation also plays an important role: the unemployed and home workers show a low level of trust in statistics. Gender and political orientation do not play a significant role, and these last results indicate that trust in official statistics is impartial to political views.

Our main results indicate that the European Union has played an important role in increasing trust in official statistics in member countries, a finding that, to the best of our knowledge, is new to this literature. The literature on trust in European institutions (e.g. Torcal and Christmann 2019; Muñoz et al. 2011) show that context matters when country-level characteristics are considered, including the level of corruption, welfare spending, and decision-making power in the EU. Member countries perceive the European Union to be a reliable and trustworthy institution, and this may have implications for new members. The enlargement of the EU is perceived as a process in which a higher level of governance may decrease corruption at the national level. Thus, the increase in national trust in the European Union and Eurostat can also explain the increase in individual trust in statistics through the story of Europe.

7. Conclusion

In this paper, we discuss a particular type of institutional trust, namely trust in official statistics, using two waves of the Eurobarometer survey (2007 and 2015). Summary data show that nearly 54% of Europeans tend to trust official statistics, even though there is great heterogeneity between European countries. We econometrically model individual trust in official statistics by the means of a multilevel model that considers the two-dimensional nature of our data (individual and country level). We find that with respect to level 1 (the individual), trust in

official statistics is mainly explained by education, statistical knowledge and occupation status. Variables such as gender and political orientation do not affect an individual's level of trust in statistics. With respect to variation between European countries, our analysis reveals an important result. We show that macro variables such as GDP growth, GDP per capita and indexes of inequality (unemployment or poverty) do not play a significant role in individual trust in official statistics. The only variable that is significant for trust in official statistics is the historical changes in European Union membership. More specifically, a country's accession to the EU always significantly increases trust in official statistics among its population. This result provides a significant contribution to the literature, but above all, it underscores the role of the EU (and Eurostat) in building credible statistics and thus influencing individuals' trust in national official statistics. Future research exploring, for example, trust in European statistics or Eurostat more specifically would be of great interest.

References:

Algan, Y, S. Guriev, E. Papaioannou, and E. Passari (2017), The European Trust Crisis and the Rise of Populism. *Brookings Papers on Economic Activity*, 309–382.

Arrow, K.J. (1972) Gifts and exchanges. **Philosophy & Public Affairs**, 343-362

Bjørnskov Ch. (2012) How Does Social Trust Affect Economic Growth? **Southern Economic Journal,** 78(4), 1346-1368

Budnik Ch. (2018) Trust, Reliance, and Democracy. **International Journal of Philosophical Studies,** 26(2), 221-239

Charpin J.-M. (2010) L'information statistique en perspective : six grands changements. **Revue d'économie financière,** 98/99, 15-26

Fukuyama, F. (1995) **Trust**. New York: Free Press

Foster C, Frieden J. Crisis of trust: Socio-economic determinants of Europeans' confidence in government. *European Union Politics.* 2017;18(4):511-535. doi:10.1177/1465116517723499

Glaeser, E.L., Laibson, D.I., Scheinkman, J.A. and Soutter, C.L. (2000) Measuring trust. **The Quarterly Journal of Economics,** 115(3), 811-846

Guiso, L., Sapienza, P. and Zingales, L. (2003) People's opium? Religion and economic attitudes. **Journal of Monetary Economics,** 50(1), 225-282

Hawley, K.J. (2017) Trustworthy groups and organisations. In P. Faulkner & T. Simpson (eds), **The Philosophy of Trust**. Oxford: Oxford University Press

Horsburgh, H.J.N. (1950) The Ethics of Trust. **The Philosophical Quarterly,** 10(41), 343-354

Gambetta, D. (1987) Were They Pushed or Did They Jump?: Individual Decision Mechanisms in Education **(Studies in Rationality and Social Change**). Cambridge: Cambridge University Press

Jann, B. (2014) Plotting regression coefficients and other estimates. **Stata Journal,** 14, 708-737

Jones, K. (1996) Trust as an affective attitude. **Ethics,** 107(1), 4-25

Knack, S. and Keefer, P. (1997) Does social capital have an economy payoff? A cross-country investigation. **Quarterly Journal of Economics,** 112(4), 1251-1288

Langer, W. (2020) How to assess the fit of Multilevel Logit Models with STATA? A project in progress. https://langer.soziologie.uni-halle.de/stata/index.html

La Porta, R., Lopez-de-Silanes, F., Shleifer, A. and Vishny, R. (1997) Trust in large organizations. **American Economic Review Papers and Proceedings,** 87(2), 333-338

Lovell D.W. (2001) Trust and the politics of postcommunism. **Communist and Post-Communist Studies**, 34(1), 27-38

Hox, J.J. (2010) **Multilevel Analysis. Techniques and Applications**. New York and Hove: Routledge

Heck, R.H. and Thomas, S.L. (2009) An Introduction to Multilevel Modeling Techniques. New York and Hove: Routledge

Inglehart, R., & Norris, P. (2017). Trump and the Populist Authoritarian Parties: The Silent Revolution in Reverse. ***Perspectives on Politics***, *15*(2), 443-454.

Kaufmann, D., Kraay, A. and Mastruzzi, M. (2010) The worldwide governance indicators: a summary of methodology, data and analytical issues. **World Bank Policy Research Working Paper,** 5430

Mark W. (2018) Trust and Democracy. In E.M. Uslaner (ed) **The Oxford Handbook of Social and Political Trust.** Oxford University Press

Miller, W.L., White, S. and Heywood, P. (1998) **Values and Political Change in Postcommunist Europe.** Basingstoke: Macmillan

Muñoz, J., Torcal, M. and Bonet, E. (2011) Institutional trust and multilevel government in the European Union: congruence or compensation? **European Union Politics,** 12(4), 551-574.

Oreskes N. (2019) **Why Trust Science?** Princeton University Press

Plasser, F. and Ulram, P. (1996) Measuring political culture in East Central Europe. Political trust and system support. In F. Plasser and A. Pribersky (eds), **Political Culture in East Central Europe** (pp. 3-33). Avebury Aldershot, Hants: Brookfield VT

Potter, N.N. (2002). **How Can I Be Trusted? A Virtue Theory of Trustworthiness**. Rowman & Littlefield Publishers

Putnam, R., Leonardi, R. and Nanetti, R.Y. (1993) **Making Democracy Work**. Princeton University Press: Princeton, NJ

Putnam, R.D., Helliwell, J. (1995) Economic Growth and Social Capital in Italy. **Eastern Economic Journal,** 21(3), 295-307

Sitkin, S.B. and Roth, N.L. (1993) Explaining the limited effectiveness of legalistic "Remedies" for trust/distrust. **Organization Science,** 4(3), 367-392

Tabellini, G. (2010) Culture and Institutions: Economic Development in the Regions of Europe. **Journal of the European Economic Association**, 8(4), 677-716

Torcal, M. and Christmann, P. (2019) Congruence, national context and trust in European institutions. **Journal of European Public Policy,** 26(12) 1779-1798

Ricciato, F., Wirthmann, A. and Hahn, M. (2020). Trusted Smart Statistics: How new data will change official statistics. **Data & Policy**, 2, E7. doi:10.1017/dap.2020.7

Raudenbush, S.W. and Bryk, A.S. (2002) **Hierarchical Linear Models. Application and Data Analysis Methods**. Thousand Oaks, CA: Sage

Warren, M.E. (2018) Trust and Democracy. In E.M Uslaner (ed) **The Oxford Handbook of Social and Political Trust.** Oxford University Press

Wang, L. and Gordon, P. (2011) Trust and institutions: A multilevel analysis. **The Journal of Socio-Economics**, 40, 583-593

Appendices:

Appendix A: Regression Model with Country Fixed Effects

While the Multilevel Model allows us to distinguish between-country variation and identify macro variables that affect this cross-variation, we also provide a robustness analysis using country fixed effects. We model the response in trust in official statistics with a logit framework in which we include all level-1 individual characteristics and 30 country and regional dummies. Figure 7 shows the average marginal effects for this regression. As can be clearly seen, the marginal effects are very similar to our baseline multilevel estimation in either sign or significance level. Thus, our identification in the baseline model is reasonable. We cannot identify macro variables at level 2 that explain cross-country variation in trust using country fixed effects. For this purpose, we have developed our multilevel model.

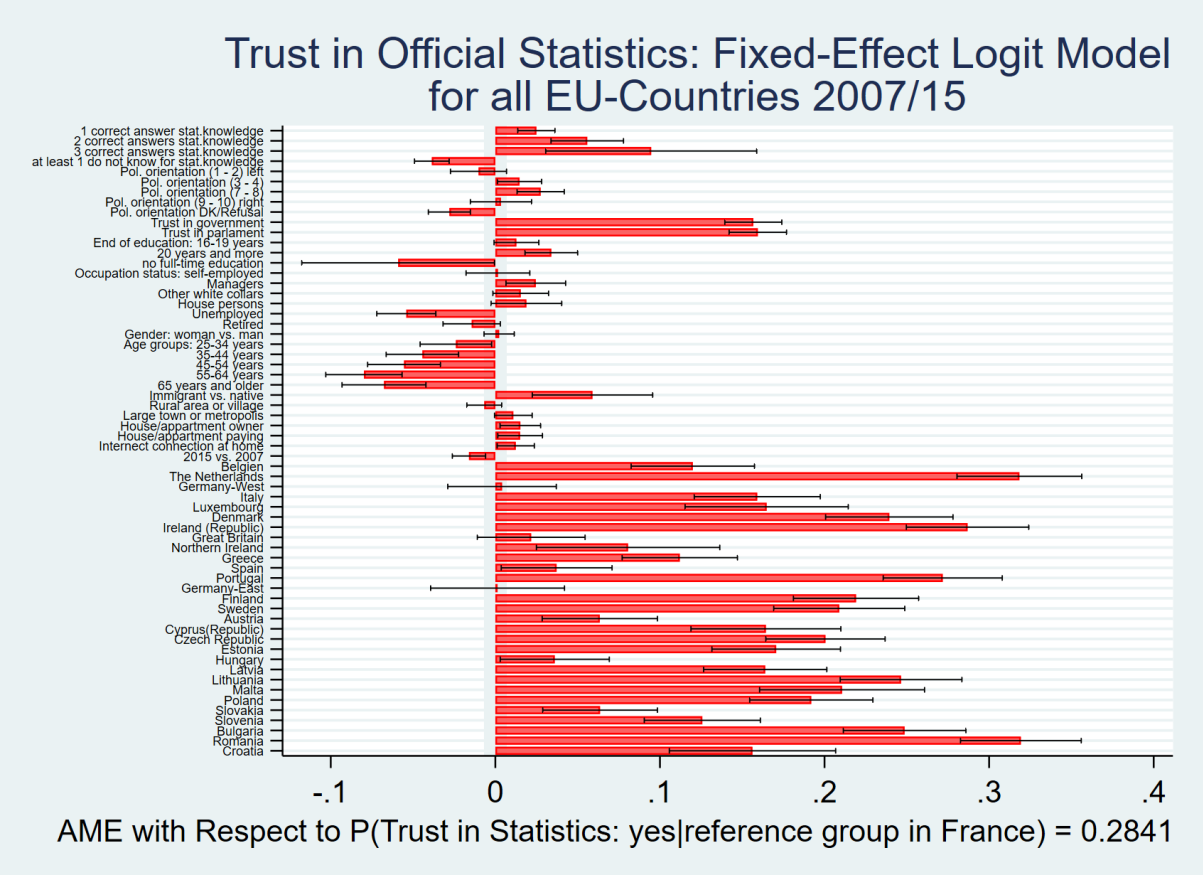

Figure 5: AMEs for the logit model with country dummies

McKelvey and Zavoina pseudo-R^2 = 0.1864

McFadden pseudo-R^2 = 0.1118

Wald-chi^2 (63) = 5114.30 Prob. > chi^2 = 0.000

N = 45,953

Appendix B: Separating trust in official statistics between Eastern and Western Europe

Historical trends in Europe have influenced the divergence of social trust between Eastern and Western Europe and, correspondingly, the trust of citizens in official statistics. Lowell (2001) emphasises how the notion of *'political culturale'* for post-communist countries is grounded on the concept of social capital and is quite different from the notion in Western countries. Political participation, multiparty systems and democracy are phenomena that became familiar in post-communist countries only at the beginning of the 1990s. Lowell (2001) further explains that *'Trust may be understood as a part of political culture'* and that in these countries, it was absent for half a century. Meanwhile, research regarding social trust in post-communism has found that *'these countries are characterised by low levels of trust in the new political institutions of democracy'* (Miller et al. 1998). From these studies, it can also be deducted that individual trust in the official statistics is quite different when comparing East and West, a topic that is investigated in this study.

In our last baseline regressions, in order to identify variation in trust in official statistics, we use as an econometric strategy the pooling of all European citizens in the Eurobarometer data. This gives us a general conclusion regarding the average individual variation within each country, although it may hide some issues related to the historical evolution of social trust in Europe. In particular, when considering social trust as part of social capital in most East European countries, it must be taken into account that this concept is new and very different from its conceptualization in Western societies. For nearly half a century, these countries experienced a political system based on complete central governance, where the concepts of democracy and social participation were absent. Social trust, and in particular trust in institutions, appeared as a new concept after the fall of the Iron Curtain, and the concept of trust in official statistics is even newer.

Research from political science has concluded that post-communist Eastern Europe *'is characterised by low levels of trust in the new political institutions of democracy'* (Miller et al. 1998). In contrast, trust in ordinary people *'was'* relatively high at 76% and uniform across the Eastern Countries of Europe (ECE). In many countries in the ECE, the churches and the army enjoy more trust than governments and parliaments (Plasser and Ulram 1996).

In this perspective, we need to account for this historical difference in social trust between Western and Eastern Countries.[29] We split our baseline level (1) regressions between Eastern and Western Europe and show the results in Figure 7 and Figure 8. With respect to our baseline-pooled

[29] **Eastern Europe**: Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Slovakia, Slovenia, Bulgaria, Romania, Croatia. **Western Europe**: Austria, Belgium, Germany (East and West), Luxembourg, France (reference category), the Netherlands, Spain, Portugal, Italy, UK, Denmark, Sweden, Finland, Greece, Malta, Cyprus.

regression, the estimations reveal an interesting difference. Political orientation plays a significant role in trust in statistics, and the marginal effects are quite different for Eastern and Western Europe. For Western countries, extreme political orientation (extreme left or extreme right) negatively and significantly influences trust in official statistics. In Eastern countries, centre orientation (neither left nor right) is positively and significantly correlated with trust in official statistics. All other variables follow the same trend for the two subsamples, although the significance and magnitudes differ.

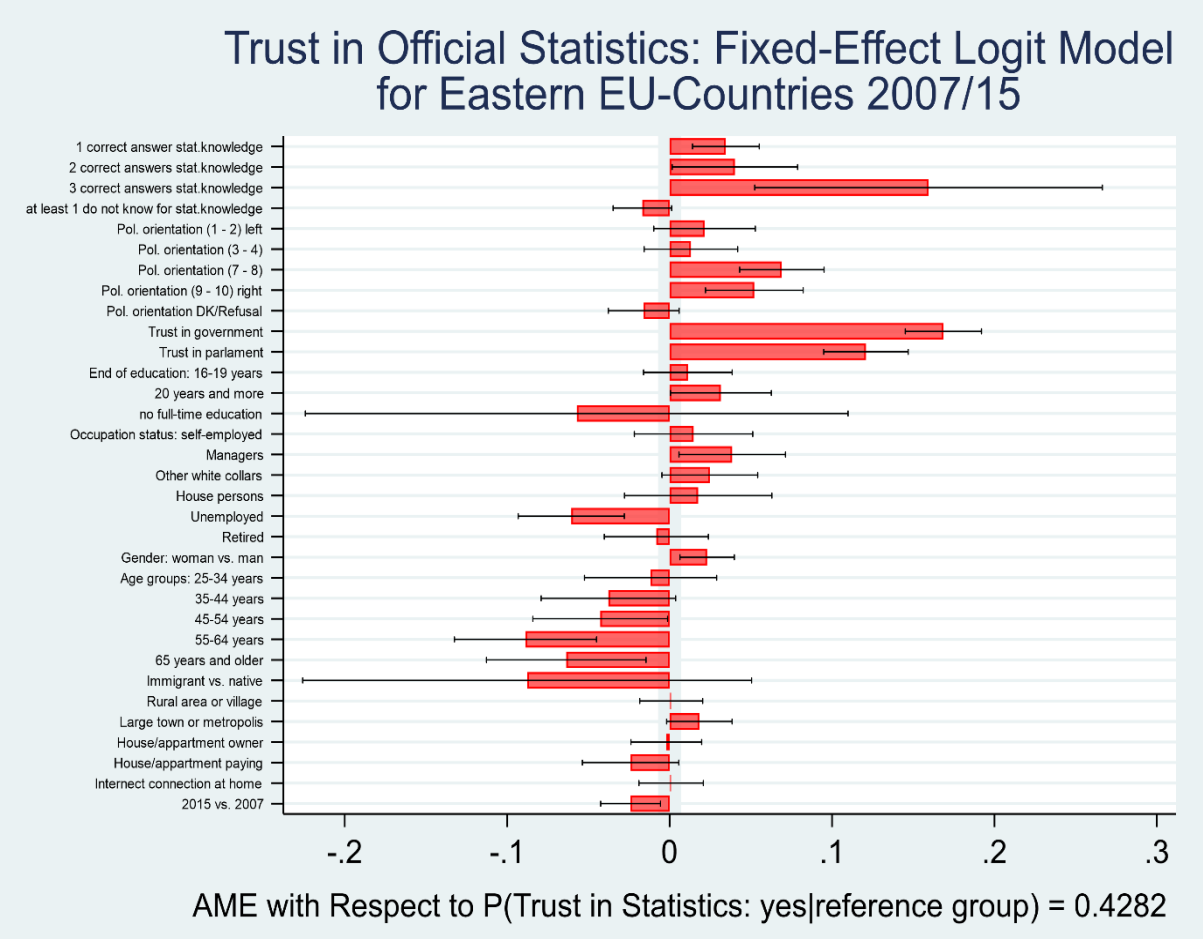

Figure 6: MLM of trust in official statistics for Eastern countries

***McKelvey and Zavoina pseudo-R^2 (fixed and random effects) = 0.0954**

***McFadden pseudo-R^2 (fixed and random effects) = 0.0563**

*** Wald-chi^2 test statistics (34) = 1049.29 Prob. > chi^2 = 0.000**

N = 17,910

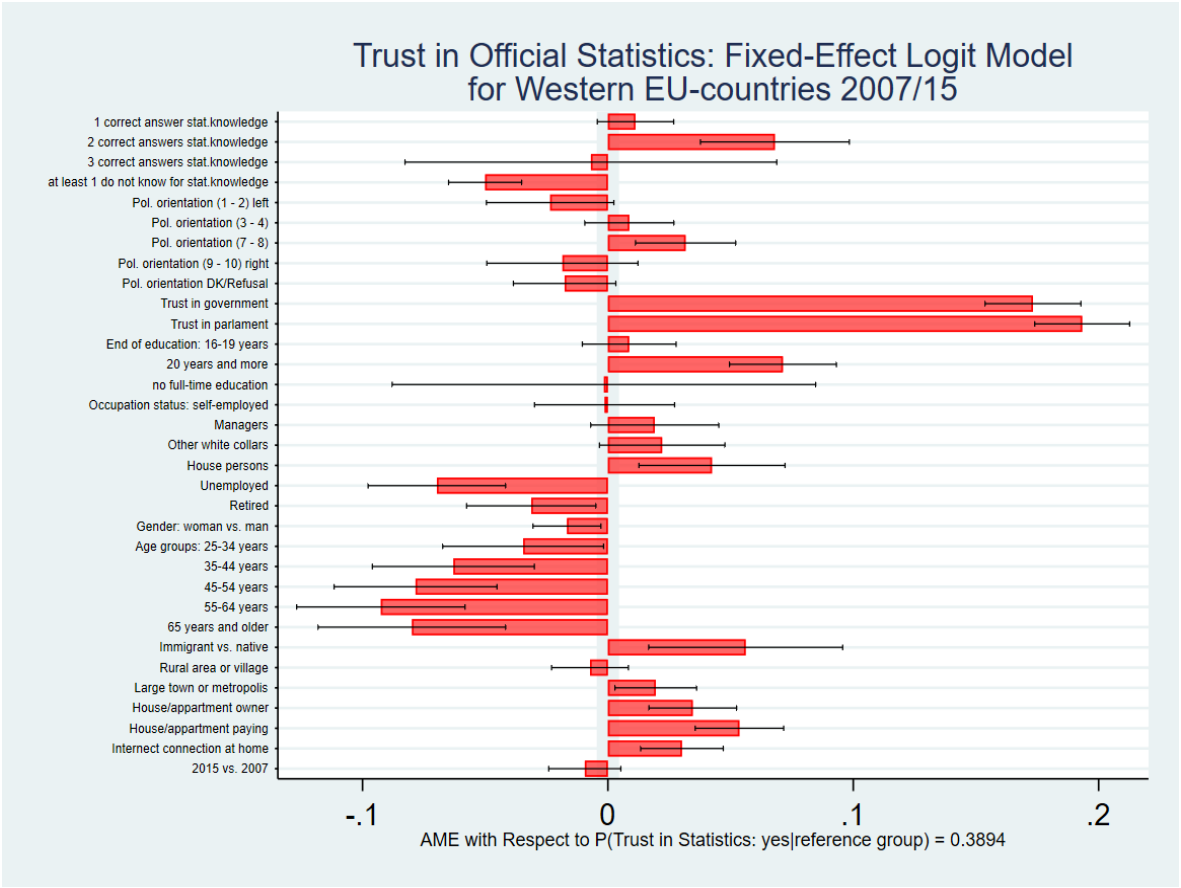

Figure 7: MLM for trust in statistics in Western countries

***McKelvey and Zavoina pseudo-R^2 (fixed and random effects) = 0.1787**

***McFadden pseudo-R^2 (fixed and random effects) = 0.1114**

*** Wald-chi^2 test statistics (34) = 3189.44 Prob. > chi^2 = 0.000**

N = 28,043