## Do You Want Cookies?

Trust dynamics and educational gaps in the datafied risk society

Angelica M. Maineri

**Open Press Tilburg University** 



#### © 🛈

This work is licensed under a Creative Commons Attribution 4.0 (CC BY 4.0) license, which means that you are free to copy and distribute this work in any medium or format, granted that you give appropriate credit to the author. For more information, see https://www.creativecommons.org/licenses/by/4.0/legalcode.

ISBN 978-94-036-7630-2 DOI 10.26116/ET28-VC69

Published by Open Press Tilburg University, Tilburg, the Netherlands https://www.openpresstiu.org

Typesetting & design by Thomas F. K. Jorna JOTE Publishers

© 2022 Angelica M. Maineri (https://orcid.org/0000-0002-6978-5278).

## Do you want cookies? Trust dynamics and educational gaps in the datafied risk society

Proefschrift ter verkrijging van de graad van doctor aan Tilburg University op gezag van de rector magnificus, prof. dr. W.B.H.J. van de Donk, in het openbaar te verdedigen ten overstaan van een door het college voor promoties aangewezen commissie in de Aula van de Universiteit op vrijdag 9 december 2022 om 10.00 uur door

> Angelica Maria Maineri geboren te Milaan, Italië

Promotor: prof. dr. P.H.J. Achterberg, Tilburg University

Copromotor: dr. A.R.C.M. Luijkx, Tilburg University & University of Trento

Promotiecommissie: prof. dr. G.A. Veltri, University of Trento prof. dr. Y.J. Park, Howard University prof. dr. M.E.H. van Reisen, Tilburg University prof. dr. W. de Koster, Erasmus University Rotterdam prof. dr. S.M.E. Wyatt, Maastricht University

#### Abstract

Ulrich Beck described the shift to the Risk Society, characterized by the uneven distribution of manufactured risks stemming from human activity. In the thesis I apply Beck's Risk Society perspective to the study of datafication, with a twofold aim: on the one hand, the Risk Society theory can explain the uneven acknowledgment of risks unfolding within the datafied society; on the other hand, the process of datafication constitutes an interesting case to test the empirical grounding of the reflexive modernization thesis, according to which, as modernization progresses, technological progress is increasingly questioned. Results of four empirical studies show how some elements of the reflexive modernization theory do not pass the empirical test. First, a country's level of digitalization does not deepen knowledge-based stratification mechanisms. Second, individuals' trust in data institutions does not drop when the pitfalls of datafication become visible, challenging the 'worked-and-won' dynamic of trust in the risk society. Nevertheless, I also show how the Risk Society perspective is beneficial to better understand some aspects of the datafication processes, as findings indicate the success of organized irresponsibility dynamics, as well as the important role of knowledge as a risk stratification mechanism at the individual level.

### Contents

1

2

Abstract	IX
List of Figures	XV
List of Tables	XIX
Introduction	I
Datafication and Risks	Ι
The Risk Society	5
The contribution of the thesis:	
understanding the datafied risk society	7
In Zuck We Trust?	23
Introduction	24
Theoretical background	27
Data and methods	32
Results	40
Conclusion and Discussion	50

3	Public acceptance of a COVID-19 Health Pass	55
	Introduction	56
	Accepting the COVID-19 Health Pass	58
	The present study	62
	Data and Methods	67
	Results	72
	Conclusion and Discussion	79
4	The Closing Educational Gap in E-privacy Management in E	uropean
	Perspective	85
	Introduction	86
	The educational gap in e-privacy management	89
	Data and methods	92
	Results	IOI
	Summary of findings and Discussion	106
5	Switch on the Big Brother!	115
	Introduction	116
	Explaining acceptance of surveillance	118
	Data and Method	125
	Results	132
	Discussion	136
6	Conclusions	145
	The Datafied Risk Society	145

	Limitations and future work Concluding remarks	154 156
А	Appendix to In Zuck We Trust?	159
В	Appendix to Public acceptance of a COVID-19 Health Pass	167
	Test of proportional odds	168
	Question wording	169
	Additional results	174
С	Appendix to The Closing Educational Gap in E-privacy Mana European Perspective	gement in 179
	Non-Internet users	179
	Alternative operationalization of Reflexive mindset	181
D	Appendix to Switch on the Big Brother!	185
	Bibliography	214
	Summary	215
	Acknowledgements	219

xiii

## List of Figures

1.1	Share of internet users by country between 2015 and 2017. Own elaboration of data from Ourworldindata.org.	16
2.1	Timeline of the study (source of Google search volumes: Google Trends)	35
2.2	Correlation plot among institutional trust items	37
2.3	Predicted probabilities of change in trust in social media and 95% confidence intervals by institutional trust (estimated from the model in Table 2.4).	48
3.1	Predicted Probabilities (with 95% Confidence Intervals) of Will- ingness to use COVID-19 HP by institutional trust (PPO model), N = 1,464	73
3.2	Predicted Probabilities of Willingness to use HP with 95% Confi- dence intervals by experimental conditions, estimated via ordinal logistic model (PPO model), N = 1,454	76
3.3	Predicted Probabilities of Willingness to use Health Pass and 95% Confidence intervals by experimental conditions by the purpose of use of the HP, estimated via ordinal logistic model (PPO model), $N = 1,454$ . Full models available in Appendix B $\ldots$ .	77

3.4	Predicted Probabilities of Willingness to use Health Pass and 95% Confidence intervals by experimental conditions by levels of trust in government, estimated via ordinal logistic model (PPO model), N = 1454. Only extreme categories are displayed for parsimony, full results can be found in Appendix B	78
3.5	Predicted Probabilities of Willingness to use Health Pass and 95% Confidence intervals by experimental conditions by levels of trust in science, estimated via ordinal logistic model (PPO model), N = 1454. Only extreme categories are displayed for parsimony, full results can be found in Appendix B	80
4.1	Percentage of mentions for each e-privacy management action.	94
4.2	Mean of years spent in education by country (error bars represent 95% confidence intervals)	96
4.3	Distribution of Digital Economy and Society Index (DESI) 2017 across EU countries.	100
4.4	Average number of e-privacy management activities by country. The dotted line represents the grand-mean across countries.	102
4.5	Predicted values of e-privacy management and 95% confidence intervals by 10 <sup>th</sup> , 50 <sup>th</sup> and 90 <sup>th</sup> percentile of years spent in full- time education and age categories (controlling for all variables included in M5)	105
4.6	Predicted values of e-privacy management and 95% confidence intervals by 10 <sup>th</sup> , 50 <sup>th</sup> and 90 <sup>th</sup> percentile of years spent in full-time	
	education and DESI	108
5.1	Conceptual model	123
5.2	Distribution of the dependent variables (N = 48,047). Source: EVS (2020), own calculations.	127

5.3	(a) Distribution of the change in KOF Globalization Index between 2007 and 2017. Source: KOF Globalization Index (1970-2020), own elaboration (b) Distribution of the NRI across countries. Source: World	
	Economic Forum (2016)	130
5.4	Average acceptance of surveillance by country with 95% Confi-	
	dence intervals. Source: EVS (2020). Own calculations	133
5.5	Predicted Acceptance of Surveillance by education and change in KOF Globalization Index (based on M6 in Table 5.2). Source:	
	EVS (2020). Own calculations.	139
5.6	Predicted Acceptance of Surveillance by education and NRI	
	(based on M7 in Table 5.2). Source: EVS (2020). Own calculations.	. 140
B.1	Frequency distribution of willingness to use the COVID-19 HP,	
	N = 1,454	167
B.2	Example of vignette	171
B.3	Correlation matrix of variables included in the regression models.	174
C.1	Proportion of non–internet users by country. The dotted line represents the grand mean across countries	180
C.2	Predicted values of non-internet use and 95% confidence inter-	
	education and DESI.	181
D.1	Distribution of the ES-ISCED educational attainment by countries (N = $48,047$ ) Source: EVS (2020), own calculations	186
D.2	Average score on item measuring the preference for a strong leader across countries with 95% confidence intervals. Source: EVS	
	(2020), own calculations	187
D.3	Average score of institutional knowledge across countries with 95% confidence intervals. Source: EVS (2020), own calculations.	190

xvii

D.4	Predicted Acceptance of Surveillance by education (based on M1	
	in Table 5.1). Source: EVS (2020). Own calculations	191
D.5	Fixed and random slopes (estimated from M5, Table 5.2). Dashed	
	lines represent the slopes of education for the pooled sample	
	whereas the solid lines represent the slopes for each country.	
	Source: EVS (2020). Own calculations.	192

## List of Tables

2.1	Factor loadings and uniqueness after principal components factor analysis (N = 1,504)	42
2.2	Ordinal logistic regression of trust in social media on relevant covariates	45
2.3	Cross Tabulation of Trust in Social Media in the pre-test and in the post-test.	47
2.4	Multinomial logistic regression of change in trust in social media (Reference = No Change)	49
3.1	Summary of experimental conditions and hypotheses	70
3.2	Descriptive statistics	72
3.3	PPO models of Willingness to use the COVID-19 HP by individ- ual characteristics and experimental conditions	74
4.1	Factor loadings and uniqueness of principal components factor analysis (N = $21,450$ ).	95
4.2	Descriptive statistics of individual characteristics (N = 21,177)	98
4.3	Multilevel linear regression analyses of e-privacy management on individual characteristics (N=21,177)	104
4.4	Multilevel linear regression analyses of e-privacy management on individual and country characteristics (N=21,177)	107

xix

5.1	Multivariate multilevel linear regression of AoS on individual characteristics (N=48,047 in 33 countries).	134
5.2	Multivariate multilevel linear regression of AoS on individual and	
	country characteristics (N=48,047 in 33 countries)	137
A.1	Descriptives of all trust-items	160
A.2	Descriptives of the continuous variables	161
A.3	Descriptives of the categorical variables	162
A.4	Ordinal logistic regression of trust in social media on relevant covariates, including trust in government instead of institutional	
	trust	163
A.5	Multinomial logistic regression of change in trust in social media (Reference = No Change), using trust in government instead of	
	institutional trust.	165
B.1	Testing the proportional odds assumption	168
B.2	Question wording and answer categories per variable	172
B.3	Partial proportional odds models of Willingness to use the COVID-19 HP by individual characteristics and experimental	
	conditions, with interaction terms.	176
C.1	Multilevel linear probability model of non-internet use on indi-	
	vidual characteristics	183
C.2	Multilevel linear regression analyses of e-privacy management on	
	individual characteristics	184
D.1	Percentage of missing values (Don't know + I prefer not to answer)	
	per variable	188
D.2	Descriptive statistics of individual-level variables	188
D.3	Factor loadings after promax rotation (N = 54,689)	189

To G., 'cause this gift will last forever

Ron Swanson: "APRIL! Listen I was trying to buy this hand crafted mahogany wood model of a B-25 Mitchell Panchito aircraft [...] I went to this website and this ad popped up, 'Hey Ron Swanson, check out this great offer.'" April Ludgate: "What's your question?" RS: "My question is, what the hell?" AL: "Like how do they know who you are?" RS: "Yeah," AL: "OK, um, there are these things called cookies, where like if you go to a site and buy something it will remember you and create ads for other stuff you might wanna buy." RS: "So it learns information about me? Seems like an invasion of privacy." AL: "Dude, if you think that's bad, go to Google Earth and type in your address." (Ron takes his computer to the dumpster)

Ron Swanson on Parks & Recreation, season 4 episode 9 "The Trial of Leslie Knope", 2011 - watch the scene at https://edu.nl/3ycjt.

# Introduction

#### I • DATAFICATION AND RISKS

The expansion of ICTs has tremendously enhanced life opportunities: an enormous amount of information is available to citizens, governments, and firms on all corners of the world, which use it in combination with powerful digital tools to make processes more efficient, and – in many cases - safer. In recent years, the advent of Big Data has nurtured the process of datafication of society, or 'the process by which subjects, objects, and practices are transformed into digital data' (Southerton, 2020, p. 1), which now involves virtually every aspect of social life. Think of the quantification of friendships enabled by Facebook via likes and comments, but also –- in the framework of the COVID-19 pandemic – the quantification of exposure risk via the collection of information on social encounters via digital contact tracing.

In current understandings, datafication consist of two elements (Southerton, 2020): the translation of social/human life into machine-readable data, and 'the generation of different kinds of value from data' (Mejias & Couldry, 2019, p. 3). The first element, i.e. rendering social life into quantified bits, has to do with a general tendency to quantification common to all modern societies (cf. Mennicken & Espeland, 2019). Nowadays, however, it is boosted by inexpensive data storage solutions, powerful machine learning tools, and algorithms, that enable to process and analyze large volumes of data at relatively little cost and effort. In this sense, datafication has led to the quantification of