Università degli Studi di Modena e Reggio Emilia

DOTTORATO DI RICERCA IN

Lavoro, Sviluppo e Innovazione

CICLO DEL CORSO DI DOTTORATO

XXXII ciclo

Changing societies in the post-crash era: essays in political economy

Michele Cantarella

Supervisor: Prof.ssa Tindara Addabbo

Co-supervisor: Prof.ssa Chiara Strozzi

Coordinator: Prof.ssa Tindara Addabbo

Declaration

I certify that the thesis I have presented for examination for the PhD degree in Labour, Development and Innovation of the University of Modena and Reggio Emilia is solely my own work other than where I have clearly indicated that it is the work of others. Where specified, the extent of any work carried out jointly by me and any other person is clearly identified in it. The copyright of this thesis rests with the author. Quotation from it is permitted, provided that full acknowledgement is made. This thesis may not be reproduced without the prior written consent of the author. I warrant that this authorisation does not, to the best of my belief, infringe the rights of any third party.

I confirm that chapter 1 was jointly co-authored with Chiara Strozzi, that chapter 2 was jointly co-authored Nicolò Fraccaroli and Roberto Volpe, that I am the sole author of chapter 3, and that chapter 4 was jointly co-authored with Andrea Neri and Maria Giovanna Ranalli.

I certify that Chiara Strozzi and I jointly wrote chapter 1. I attest that I collected, merged and harmonised the survey data, that the empirical model was jointly designed by Chiara and myself and that the results were also jointly analysed.

I certify that Nicolò, Roberto and I jointly wrote chapter 2. I certify that I designed the econometric model, while Nicolò developed the text analysis techniques, which have been jointly integrated into the final identification strategy. Roberto provided exploratory work in the search for an observational study, while also informing the research by detailing the political and demographic features of the population used in the study. The data was jointly collected by all authors. I attest that I provided most codes used in the production of the final data-set — harmonising text analysis codes from Nicolò — and in the estimation of the model. The results were jointly analysed by Nicolò and me, while I developed the final theoretical model.

I certify that Andrea Neri, Maria Giovanna Ranalli and I jointly wrote chapter 4. I attest that Andrea and Maria Giovanna and I jointly developed the proposed allocation method. with Andrea and Maria Giovanna providing input on survey calibration methodologies, while I developed Pareto tail estimation methodologies and integrated these methods into the final calibration model.

Abstract

Over the last decade, the financial crisis, accompanied by ongoing processes of globalization and digitalisation, have accelerated major societal disruptions and transformations, shaking the foundations upon which modern democracies have been built upon. Unchecked wealth accumulation has led to unprecedented inequalities, and it is unclear whether these have been recently amplified by quantitative easing monetary policies pursued by central banks. New atypical forms of work have led to the deterioration of labour market laws and offered the potential for massive disruptions on the supply side. Social media filter bubbles have challenged our understanding of how individuals inform their voting decisions, and many have questioned which role online social networks have played in the rise of populism. Intoxicated by their own success, populist policy-makers themselves have openly violated those human rights at the basis of democratic societies, as exemplified by the recent migration policies implemented in certain EU Mediterranean states.

Over the course of my PhD studies, I have pursued diverse research paths across these different, yet interlaced, topics in political economy, focusing on new labour markets, post-truth politics, forced migration and wealth inequality, and using the tool-set from applied econometrics to provide concise answers to far-reaching questions.

The present work collects four empirical papers in the field of political economy. The first chapter focuses on online micro-task labour markets. In the paper, joint with Chiara Strozzi, "Workers in the crowd, the labour market impact of the online platform economy", I focus on online micro-task service outsourcing (or, "Crowdwork") and its impact on earnings and working conditions. This research aims at estimating (i) the impact of these work arrangements in terms of hourly earnings and working conditions, and (ii) how much these differences in job quality are caused, on the supply side, by self-selection into crowdwork, or how much they rather relate to other characteristics of these markets. We found crowdworkers to earn around 70% less than traditional workers of similar ability, and their human capital to be highly under-utilised. We conclude that earnings in online micro-task markets are

indifferent to both observed and unobserved ability and that the remaining differences can be attributed to (i) increased competition from workers, on the supply side, (ii) low demand for these kinds of tasks and/or (iii) the monopsony power held by these online platforms.

The second chapter focuses on misinformation on social media and its effects on electoral behavior. Co-authored with Nicolò Fraccaroli and Roberto Volpe, the paper "Does fake news affects voting behaviour" has focused on the estimation of a causal connection between the spread of online hoaxes and the electoral success of populist parties, controlling for self-selection into "misinformation bubbles". We do so by exploiting a natural experiment typical of the Trento and Bolzano/Bozen autonomous provinces in Italy, where the presence of different linguistic groups shields the German-speaking minority of Bolzano from exposition to most sources of online misinformation. Combining municipality-level data from various sources, and after introducing a novel index based on text mining techniques to measure populism, our results eventually indicate that misinformation had a negligible effect on populist vote in Trentino and South Tyrol during the Italian 2018 general elections, meaning that exposition to misinformation in the region is mostly explained by an already present, yet latent, demand for populism.

The third chapter, "#Portichiusi: the human costs of migrant deterrence in the Mediterranean", focuses on the migration policies pursed in Italy between June 2018 and July 2019, and how these policies have affected mortality rates in migration. Using daily data on forced migration from the IOM, I compared trends in flows and mortality across three major migration routes in the Mediterranean, analysing the effects of the introduction of these policies. Controlling for exogenous shocks which affect push and pull factors in mobility, along with sea state conditions and route-day fixed effects, I found that the reduction in refugee migration flows in the Central Mediterranean has been modest, at best. At the same time, these policies have generated a permanent increase in daily mortality rates in the Central Mediterranean, having grown by more than 4 deaths per day. Finally, I investigated whether variations in mortality are sufficient to offset migration flows. Increases in mortality rates, however, are only accompanied by a short-term negative displacement effect, as migration attempts are delayed by increases in absolute mortality, rather than being prevented.

Over the course of my research activities at the European Central Bank, as an active member of the Linking Micro-Macro Household Statistics expert group, I have also undertaken major research in the development a novel allocation method combining distributional information from a Pareto power law model with re-weighting and imputation procedures, in order to fill the gap between survey data from the Eurosystem Household Finance and Consumption Survey and macro statistics from the National Accounts. Indeed, surveys typically suffer from a number of biases, frequently caused by differential non-response across the wealthiest households and under-reporting behaviour, leading to severe mismatches with macroeconomic aggregates. The fourth chapter of this thesis proposes a novel allocation methodology for survey data, allowing for significant improvements in the measurement of wealth inequality, but also providing the significant advantage of retaining the micro-data structure of the data-set, enabling further analysis on the micro level.

The four chapters are at different stages of development. An earlier version of chapter 1 has been published as an IZA working paper (IZA DP No. 12327), and is currently under review for publication. An early version of chapter 2 has been published in the DEMB working paper series (DEMB WP No. 146), and an updated version is close to being submitted for publication at international journals. Chapter 3 has been published as a Household in Conflict Network working paper (HiCN WP No. 317). Chapter 4 is in its final draft status, and will soon be submitted to the ECB Working Paper series.

Italian translation

Nell'ultimo decennio, la crisi finanziaria, accompagnata da processi di globalizzazione e digitalizzazione, ha portato all'accelerazione di processi trasformativi in ambito socio-economico, scuotendo le fondamenta su cui le democrazie moderne sono costruite. L'accumulazione incontrollata di ricchezza ha portato a disparità senza precedenti, e non è chiaro se queste disuguaglianze siano state amplificate dalle recenti politiche monetarie di quantitative easing perseguite dalle banche centrali nel mondo. Con l'emergere di nuove forme atipiche di lavoro, si assiste al deterioramento dei diritti dei lavoratori e a trasformazioni inedite nel mercato del lavoro. Le "filter bubbles" nei social media hanno messo in dubbio la nostra comprensione del mod in cui gli individui sviluppano le proprie preferenze politiche, e molti si chiedono quale ruolo i social network abbiano avuto nel contribuire alla crescita del populismo. Intossicati dal loro stesso successo, gli stessi politici populisti stessi hanno apertamente violato diritti umani alla base di società democratiche, come esemplificato dalle recenti politiche migratorie attuate in alcuni stati mediterranei dell'UE.

Nel corso dei miei studi di dottorato, ho perseguito vari percorsi di ricerca su questi diversi, ma intrecciati, temi di economia politica, concentrandomi sui nuovi mercati del lavoro, sulla politica della "post-verità", sulla migrazione forzata e sulla disuguaglianza di ricchezza, usando il set di strumenti dall'econometria applicata per fornire risposte concise a domande di ampia portata.

Questo lavoro raccoglie quattro saggi empirici nel campo dell'economia politica. Il primo capitolo si concentra sui mercati del lavoro online con micro-task. Nel saggio, in collaborazione con Chiara Strozzi, "Workers in the crowd, the labour market impact of the online platform economy", mi concentro sui servizi di micro-task outsourcing (o "Crowdwork") e sul loro impatto sui guadagni e sulle condizioni di lavoro. Questa ricerca mira a stimare (i) l'impatto di questi accordi di lavoro in termini di retribuzione oraria e condizioni di lavoro, (ii) quanto queste differenze nella qualità del lavoro siano causate, dal lato dell'offerta, dall'auto-selezione nel lavoro su piattaforma, e quanto invece siano dovute a caratteristiche proprie di questi mercati. Abbiamo rilevato che i crowdworker guadagnano circa il 70 per cento in meno rispetto a lavoratori "tradizionali" dotati di abilità simili, e che il loro capitale umano è in gran parte sotto-utilizzato. Concludiamo che i guadagni nei mercati del lavoro online sono indifferenti sia a capacità osservate che inosservate e che le rimanenti differenze di salario possano essere attribuite a (i) l'aumento della concorrenza tra lavoratori, dal lato dell'offerta, (ii) dalla scarsa domanda per questo tipo di servizi e / o (iii) il potere monopsonistico detenuto da queste piattaforme online.

Il secondo capitolo si concentra sulla disinformazione sui social media e sui suoi effetti sul comportamento elettorale. Scritto in collaborazione con Nicolò Fraccaroli e Roberto Volpe, l'articolo "Does fake news affects voting behaviour" si concentra sulla stima di un rapporto causale tra la diffusione di bufale online e il successo elettorale dei partiti populisti, controllando per auto-selezione all'interno di "misinformation bubbles". Per farlo, sfruttiamo le caratteristiche delle province autonome di Trento e Bolzano all'interno di un esperimento naturale dove, grazie alla presenza di gruppi linguistici diversi nello stesso territorio, la minoranza di lingua tedesca è esposta alle fake news in misura minore rispetto alla popolazione di lingua italiana. Combinando dati a livello comunale provenienti da varie fonti, e dopo aver introdotto un nuovo indice di text analysis per poter quantificare il populismo, i nostri risultati indicano che la disinformazione ha avuto un effetto trascurabile sul comportamento di voto – in termini di preferenze verso partiti populisti – in Trentino e Alto Adige nel corso delle elezioni generali italiane del 2018, a significare che l'esposizione alla disinformazione nella regione è per lo più spiegata da una domanda latente di populismo.

Il terzo capitolo, "#Portichiusi: the human costs of migrant deterrence in the Mediterranean", si concentra sulle politiche migratorie perseguite in Italia tra giugno

2018 e luglio 2019, e come queste politiche abbiano influenzato i tassi di mortalità nei tentativi di migrazione via mare. Utilizzando dati giornalieri sui flussi migratori dall'IOM, ho confrontato i cambiamenti nei tentativi di migrazione e nei tassi di mortalità attraverso tre principali rotte migratorie nel Mediterraneo, analizzando gli effetti dell'introduzione di queste politiche. Controllando per shock esogeni che influenzano i "pull" e "push factors" nella mobilità migratoria, insieme alle condizioni dello stato del mare e ad effetti stagionali e di rotta, i miei risultati indicano che la riduzione dei flussi migratori di rifugiati nel Mediterraneo centrale è stata, nella migliore delle ipotesi, modesta. Allo stesso tempo, queste politiche hanno generato un aumento permanente dei tassi di mortalità giornaliera nel Mediterraneo centrale, essendo cresciuto di oltre 4 morti al giorno. Infine, ho studiato se le variazioni nella mortalità sono sufficienti per compensare i flussi migratori. Gli aumenti dei tassi di mortalità, tuttavia, sono soltanto accompagnati da una riduzione dei flussi nel breve termine che viene riassorbita in pochi giorni, in quanto i tentativi di migrazione vengono solo ritardati da aumenti nella mortalità, piuttosto che essere prevenuti.

Nel corso delle mie attività di ricerca presso la Banca Centrale Europea, in qualità di membro attivo dell'EG-LMM (Expert Group on Linking Micro-Macro Household Statistics), ho anche intrapreso importanti ricerche nello sviluppo di un nuovo metodo di allocazione, combinando informazioni distributive da un modello Pareto con procedure di calibrazione, al fine di colmare il divario tra i dati dell'indagine HFCS (Eurosystem Household Finance and Consumption Survey) con statistiche macroeconomiche dai conti finanziari nazionali. Infatti, queste indagini in genere soffrono di una serie di distorsioni, spesso causate da difficoltà nell'intervistare le famiglie più ricche e ad errori di misurazione, che portano a gravi disallineamenti con le statistiche macroeconomiche. Il quarto capitolo di questa tesi propone quindi una nuova metodologia di allocazione per i dati dell'indagine, offrendo miglioramenti significativi in termini di misurazione della disuguaglianza della ricchezza, e fornendo anche il vantaggio significativo di conservare la struttura individuale del data-set, consentendo ulteriori analisi a livello micro.

I quattro capitoli sono in diverse fasi di sviluppo. Una versione precedente del capitolo 1 è stata pubblicata come working paper IZA (DP IZA n. 12327) ed è attualmente in fase di revisione. Una prima versione del capitolo 2 è stata pubblicata nella DEMB working paper series (DEMB WP n. 146) e una versione aggiornata sarà presto proposta per la pubblicazione su riviste internazionali. Il capitolo 3 è stato pubblicato come working paper dell'Household in Conflict Network (HiCN WP No. 317). Il capitolo 4 è in uno stadio finale e sarà presto sottoposto alla serie di working paper della Banca Centrale Europea.

Ad Eleonora,

Contents

1 Workers in the crowd: the labour market impact of the online platform economy **13**

1.1 Introduction . . . 14

1.2 The online micro-task labour market . . . 16

1.3 Literature review . . . 17

1.4 Data . . . 19

1.4.1 Crowdworkers and traditional workers . . . 20

1.4.2 Selected labour market indicators and controls . . . 23

1.5 Model specification . . . 25

1.5.1 Instrumental variable identification . . . 26

1.5.2 Gender bias in caregiving . . . 27

1.5.3 Potential earnings and violations of the exclusion restriction . 29

1.6 Discussion of results . . . 31

1.7 Robustness checks . . . 44

1.7.1 Instrumental variable specification . . . 44

1.7.2 Model specification . . . 48

1.8 Conclusions . . . 50

Appendices **57**

1.A Summary statistics . . . 57

1.B Returns to observable skills and first stage IV regressions . . . 60

2 Does Fake News Affect Voting Behaviour? **62**

2.1 Introduction . . . 63

2.2 Literature review . . . 65

2.3 Background . . . 68

2.3.1 Trentino and South Tyrol: political and sociolinguistic background . . . 68

2.3.2 Fake news in the 2018 Italian general elections . . . 73

2.4 Data . . . 77
2.4.1 Electoral data . . . 77
2.4.2 Socio-demographic and internet connectivity data . . . 78
2.4.3 Social media data . . . 79
2.5 A text mining approach to measuring the populist content of parties . 80
2.6 Econometric model . . . 84
2.7 Results . . . 87
2.8 Discussion . . . 93
2.8.1 A simple model for misinformation and policy preferences . . . 93
2.9 Conclusions . . . 98

Appendices **105**
2.A Measuring exposure to fake news . . . 105
2.B Electoral gains for specific parties . . . 110
2.C correlation between text-based populist scores and CHES data . . . 112

3 #Portichiusi: the human costs of migrant deterrence in the Mediterranean **114**
3.1 Introduction . . . 115
3.2 Overview of the literature . . . 118
3.3 Data . . . 120
3.4 Empirical Model . . . 122
3.5 Deterrence and migration flows . . . 128
3.6 Deterrence and human costs . . . 133
3.7 Deterrence of mortality . . . 139
3.8 Conclusions . . . 143

Appendices **148**
3.A Reduced form regressions . . . 148

4 Mind the wealth gap: a new allocation method to match micro and macro statistics for household wealth **149**
4.1 Introduction . . . 150
4.2 Data . . . 152
4.3 The missing wealth . . . 154
4.3.1 Pareto tail estimation . . . 154
4.3.2 Methods for adjusting survey data . . . 159
4.3.3 Tail households estimation . . . 162
4.4 Allocating the remaining wealth . . . 168

4.5 Results 172
4.6 Conclusions 176

Chapter 1

Workers in the crowd: the labour market impact of the online platform economy

Abstract

In this paper, we compare wages and labour market conditions between individuals engaged in online platform work and in traditional occupations by exploiting individual-level survey data on crowdworkers belonging to the largest micro-task marketplaces, focusing on evidence from the United States and Europe. To match similar individuals, survey responses of crowdworkers from the US and EU have been harmonised with the American Working Conditions Survey (AWCS) and the European Working Conditions Survey (EWCS). Our findings indicate that traditional workers retain a significant premium in their earnings with respect to online platform workers, and that those differences are not explained by the observed and unobserved ability of individuals. This holds true also taking into account similar levels of routine intensity and abstractness in their jobs, as well as the time spent working. Moreover, labour force in crowdworking arrangements appears to suffer from high levels of underutilisation, with crowdworkers being more likely to be found wanting for more work than comparable individuals.

Keywords: *crowdwork, online platform economy, micro-tasks, routine intensity, labour market conditions*

JEL codes: J31, J42, F66

1.1 Introduction

Among the "mega-trends" which characterise the future of work, the growth of the online platform economy has been steady and fast in the recent years and has been contributing to the changing nature of work (OECD, 2016, Harris and Krueger, 2015).[1] Technological progress and digitalisation are at the basis of its current development. Due to the overall exponential growth of internet facilities, indeed, recent years have shown an increasing number of workers participating in online micro-task labour markets, within what is described as the gig, on-demand, or platform-based economy (Degryse, 2016, Prassl and Risak, 2015).

These workers are usually called crowdworkers, where crowdwork is defined as an "employment form that uses an online platform to enable organisations or individuals to access an indefinite and unknown group of other organisations or individuals to solve specific problems or to provide specific services or products in exchange for payment" (Eurofound, 2015).

The economic conditions of crowdworkers have been analysed in a number of recent descriptive studies (e.g. Berg et al., 2018, Berg, 2015, Difallah et al., 2018, Hara et al., 2018, Pesole et al., 2018) showing how these workers suffer from the erosion of fundamental labour rights, the loss of social protections and difficulties in exercising collective action. However, it would be a mistake to assume, solely based on the evidence from these descriptive studies, that this deterioration of working condition is to be fully attributed to a platform effect, as it could be argued that the characteristics of crowdworkers are intrinsically different from the characteristics of workers in traditional professions. More definitive answers are needed, especially in light of the 2030 Agenda for Sustainable Development and the goals of the United Nations and the European Parliament in terms of decent work and social rights.[2]

Given the possibility that the online platform economy will further expand in the coming years, it is crucial for governments and social partners to take an active role in designing labour market institutions (e.g. minimum wages, employment protection, health and safety regulations) that can ensure labour and social rights

[1] According to the OECD (2016), the online platform economy is the economic activity which enables transactions - partly or fully online - of goods, services and information.

[2] During the UN General Assembly in September 2015, the four pillars of the Decent Work Agenda – employment creation, social protection, rights at work, and social dialogue – became part of the new UN 2030 Agenda for Sustainable Development (United Nations, 2015, Transforming our world: the 2030 Agenda for Sustainable Development). At the same time, the European Parliament resolution of 19 January 2017 recognised the need to set a European Pillar of Social Rights also for 'atypical or non-standard forms of employment, such as temporary work, involuntary part-time work, casual work, seasonal work, on-demand work, dependent self-employment or work intermediated by digital platforms' (European Parliament, 2017, European Parliament resolution of 19 January 2017 on a European Pillar of Social Rights).

for this type of workers. This is especially urgent for platform workers involved in the so-called micro-tasks (a series of small tasks which together comprise a large unified project and can be performed independently over the Internet in a short period of time), which are more exposed to risks concerning low pay, precariousness and poor working conditions.[3]

In light of these critical issues, in this paper we analyse a large fraction of the available evidence on earning and working conditions of micro-tasks crowdworkers. We focus on the evidence from the United States and Europe, aiming to answer to the following questions: Are individuals involved in online micro-task service outsourcing intrinsically different from traditional salaried workers involved in comparable occupations, and are there differences between micro-task crowdworkers from the US and from Europe? Is it possible to estimate the real impact of micro-task crowdwork on wages and working conditions of platform workers? We focus on the supply side of these labour markets and intend to measure how much individual characteristics influences those differences in outcomes.

Our contribution is based on an empirical analysis of cross-sectional data collected from three different surveys and harmonised in order to obtain the greatest degree of comparability. As our aim is to provide an unbiased comparison of earnings and working conditions of platform workers and traditional workers, we supplement data from general working conditions surveys with responses from specific surveys on online workers, creating two groups of 'traditional' and 'crowd' workers, and analysing variations in outcomes conditionally on participation into crowdwork markets.

For both the US and Europe, the crowdwork group includes information on workers from different online platforms – namely, Amazon Mechanical Turk (AMT), Crowdflower, Clickworker, Microworkers and Prolific Academic – coming from two dedicated surveys distributed by the International Labour Organization, while the control groups include information from available general surveys on working conditions in the US and the EU (namely, the American Working Conditions Survey, and the European Working Conditions Survey).

Our findings indicate that earnings in crowdwork are mostly indifferent to skills, and that crowdworkers earn about 70% less than traditional workers with comparable ability, while working only a few hours less per week. Also, platform workers appear to be uninterested in looking for other forms of occupation, while still expressing the desire to work more than what they currently do. These results suggest

[3]On the contrary, individuals participating in online freelancing marketplaces (such as UpWork) are involved in job projects which are usually larger in scope and can enjoy more favourable conditions.

that most crowdworkers are similar to a form of idle workforce, which is excluded from traditional employment and is still under-utilised.

To the best of our knowledge, this is one of the first attempts to provide an unbiased comparison of platform and traditional workers in terms of earnings and working conditions by matching different surveys. Moreover, in contrast with most studies on the online platform economy, which aim their attention at specific regional settings, we focus on both the United States and Europe simultaneously.

The rest of the paper is organised as follows. Section 1.2 outlines the online micro-task labour market, Section 1.3 is dedicated to a review of the literature, Section 1.4 describes the data used for our empirical analysis, Section 1.5 outlines our empirical specification and Sections 1.6 and 1.7 show our results and robustness checks. Finally, in Section 1.8 we discuss our conclusions. The Appendix is dedicated to additional descriptive statistics and regressions.

1.2 The online micro-task labour market

Phenomena such as crowdwork do not exist in a vacuum, but are fostered and facilitated by wider socio-economic trends, and the development of "virtual work" can surely be identified as one of these. The term virtual work has been used by many authors to describe all of the various forms of work characterised by the execution of work through the Internet, computers, or other IT-based tools. However, not all digital jobs are necessarily a novelty *per se*, and not all new jobs are digital. While new forms of employment have surfaced, pre-existing ones have acquired a new role and relevance, thanks to the influence of new technologies.[4]

Crowd employment is one of these new forms of work and transcends traditional arrangements by de facto requiring a tripartite relationship in which an intermediary agent - the platform - manages workers - or, rather, service providers - not only by matching them with clients but also controlling pay levels, providing ratings and generally exercising many other functions that affect workers directly. Within the platform, through an open call, client companies can offer online tasks, which are performed by contractors in exchange for remuneration (see, e.g., Eurofound, 2015). Because the majority of online platforms explicitly deny the existence of any employment relationship between the parties, individuals in crowdwork are generally characterised as independent contractors, performing their work in a discontinuous or intermittent basis.

[4]Eurofound (2015) has identified nine distinct new forms of employment: employee sharing, job sharing, interim management, casual work, ICT-based mobile work, voucher-based work, portfolio work, crowd employment and collaborative employment.

Crowd employment can then be identified as a phenomenon that essentially entails a new, and substantially cheaper, way of outsourcing tasks to a large pool of workers through IT-based platforms (Prassl and Risak, 2015) and, because of this, it has also been defined as "crowdsourcing".[5] By requiring platforms as intermediate actors, crowdwork manages to reduce most transaction costs, thus allowing for a flexible and potentially global workforce to enter the labour market and maximise the use of under-utilised assets such as human capital.[6]

Crowdwork arrangements may vary greatly: skill requirements for outsourced jobs may range from high to low and, while tasks with high routine intensity and low abstract content are prevalent – as, for example, most tasks in Amazon Mechanical Turk (AMT), Clickworker and Figure-Eight – complex and even creative activities are also present. Amazon Mechanical Turk easily stands as a prime example of a crowdwork platform, being widely recognised as one of the most popular ones (see Harris and Krueger, 2015). The short and repetitive tasks offered in AMT, as in the many other platforms, often include: image/video processing, translation, data verification, information gathering and processing, audio and visual editing, amongst many others.[7]

1.3 Literature review

Tackling the issues related to micro-task crowdsourcing has proven to be a multi-faceted effort which, so far, has seen the intervention of different disciplines such as law, information technology and economics. Until recently, the body of research on the economics of crowdsourcing has been, so far, remarkably thin, compared to other areas of study: a glaring lacuna, considering the growing size of the platform economy.

As suggested by Hara et al. (2018), this scarcity of literature is mostly attributable to the absence of publicly available data on crowdwork platforms and their workers. Nonetheless, as discussed by Horton et al. (2011), Paolacci et al. (2010) and Berinsky et al. (2012), crowdwork platforms potentially present themselves as an ideal environment for empirical studies, in particular those based on experimental research. In this regard, Horton and Chilton (2010) offer one the first

[5] This term which was first used by Jeff Howe in the article "The Rise of Crowdsourcing", Wired Magazine, 14, 2006.

[6] The ability to provide services online significantly enlarges the scope of crowdwork markets, thus enabling services to be provided globally, as opposed to the local focus of the services offered by work-on-demand platforms (such as Uber, Foodora, or Taskrabbit), which are characterised by the physical and tangible nature of the tasks being offered.

[7] As described in AMT website: https://www.mturk.com/ (last accessed: 19th September 2018).

attempts to obtain empirical evidence on reservation wages in crowd employment from an experimental framework.

Several additional descriptive studies have been provided. Harris and Krueger (2015) document the development of the platform economy and call for the recognition of an independent worker status, while other studies, receiving support from international institutions such as ILO (Berg, 2015 and Berg et al., 2018) and FEPS (Huws et al., 2017), have contributed to the literature with a thorough overview of the demographics of crowdsourcing. Hara et al. (2018) document wage and working time amongst AMT crowdworkers, discussing the necessity of including the time spent searching for tasks in working time indicators, while a recent paper from Difallah et al. (2018) summarises the main take-aways from a longitudinal survey on AMT workers.[8]

Another important contribution on the analysis of the platform economy in US comes from Katz and Krueger (2018), where the two economists, in the context of studying the evolution of all alternative work arrangements from 2005 to 2015, estimate that, out of all occupations, 0.5% involve the direct selling of activities and services mediated by an online intermediary – a figure that can proxy the size of the so called gig-economy (see Harris and Krueger, 2015).

Crowdwork can be considered as another form of service outsourcing. Some – such as Degryse (2016) – suggest that crowd employment could be equated to a form of digital migration and, in this regard, Ottaviano et al. (2013) offer a valuable study of the labour market effects of migration and task offshoring. Proxying substitutability through routine intensity of tasks – a concept originally introduced by Autor and Dorn (2013) which spurred a novel body of literature focusing on the task-based approach to labour markets – Ottaviano et al. (2013) find that service outsourcing, while having no effect on employment, has changed the task composition of native workers.

A few recent works, however, have focused on a number of supply and demand factors which contribute to the deterioration of earnings in online labour markets. Dube et al. (2018) address monopsony in online labour markets, finding that their peculiar structure allows platforms to impose a considerable markup on workers' productivity, leading up to a 20% contraption in their earnings. Looking at the supply of online workforce, the relationship between unemployment and micro-task labour markets was further explored in Borchert et al. (2018), where labour demand shocks have been found to affect temporary participation in online labour markets.

[8]The survey contains data on country, gender, age, income from AMT, time spent on AMT, marital status, household income and household size of Mechanical Turk workers, and can be accessed at the address: http://demographics.mturk-tracker.com/

Negative spill-over effects from crowdwork markets may be less obvious, but cannot be excluded. Focusing on on-demand labour platforms, Berger et al. (2018) explore the effect of introduction of Uber across taxi drivers, finding a negative association with their hourly earnings. Finally, the effects of digital labour markets on high skilled service flows are investigated in Horton et al. (2017), who focus on the UpWork freelancing platform.

While these studies all improve our understanding of important factors contributing to wage deterioration of online platform workers, none of these contributions focuses on the issue of self-selection into crowdsourcing, leaving the effect of individual ability unmeasured. We believe that a complete picture on working conditions in online crowdsourcing can only be achieved by comparison with other forms of work, and measuring how much individual characteristics of online platform workers contribute to these conditions, a task which we intend to pursue with this paper.

1.4 Data

The identification of crowdworkers in existing general working conditions surveys is not trivial. The European Working Conditions Survey (EWCS) (European Foundation For The Improvement Of Living And Working Conditions, 2017) and the American Working Conditions Survey (AWCS) (Maestas et al., 2017) both contain comparable information on wages, job quality and skills but, in both instances, it is often not possible to disentangle platform workers from any freelancer working from home. As micro-task crowdsourcers tend to perform specific, routine intensive activities, we expect that equating them to any freelancer working from home will likely pose as a serious source of bias. Also, due to the current size of the platform economy, platform workers, even if correctly identified, will naturally be under-represented in general surveys.

Dedicated surveys on crowdworkers can assist with bridging this gap. However, while there is currently plenty of information on work on digital platforms – acquired either through online questionnaires (e.g. Berg, 2015, Berg et al., 2018, Huws et al., 2017, Difallah et al., 2018) or web plug-ins (e.g. Hara et al., 2018) – the methodologies behind the collection of this data often differ significantly, with the resulting surveys varying not only in their sample sizes but also in terms of item comparability.

With the aim to provide a reliable empirical analysis of the effects of crowdwork on labour market conditions in United States and in Europe, only data sources which maximised comparability, while retaining a satisfactory pool of observations and key variables, were selected.

1.4.1 Crowdworkers and traditional workers

Our crowdwork sample uses information on European and US crowdworkers from the two rounds of the ILO Survey on Crowdworkers (Berg, 2015 and Berg et al., 2018). Thanks to the similarities in terms of the relevant variables of analysis, a group of 'traditional' workers was constructed using data from the American Working Conditions Survey and from the European Working Conditions Survey. We harmonise the ILO Survey on Crowdworkers with these general working conditions surveys in our attempt to put these new forms of work into a comparative and global perspective.

The dataset from Berg (2015) and Berg et al. (2018) consists of two consecutive surveys conducted on major online micro-task platforms[9] in 2015 and 2017 and covers crowdworkers from both the United States and Europe, along with other countries. The 2015 round of the survey provides cross-sectional data on earnings, demographics and working quality indicators for 1,167 crowdworkers from all over the world. The 2017 round similarly provides this information for a much larger number of workers ($n = 2350$), while also supplying a number of crucial variables that can be used to reconstruct the task composition of online platform work.

Using information from both rounds of the survey, we extracted a group of 1,393 US crowdworkers and 1,000 European[10] crowdworkers, where dimensions such as earnings, working hours, work quality and proxies for labour utilisation were all recorded along with demographical characteristics including gender, age, education, health condition, marital status and household size. The survey also includes items which allowed us to identify whether crowdwork constituted the respondent's main source of income.[11] Thanks to the design of the ILO survey, its contents have been easily harmonised with data from the 2015 rounds of the European Working Conditions Survey (EWCS) and the American Working Conditions Survey (AWCS) in a single cross-section.

[9]In detail: Amazon Mechanical Turk (US, EU), Crowdflower (EU), Clickworker (EU), Microworkers (EU) and Prolific Academic (EU).

[10]The European data include 852 observations from the European Member States, and 148 observations from EWCS guest countries (Norway, Switzerland, Albania, the former Yugoslav Republic of Macedonia, Montenegro, Serbia and Turkey).

[11]Further details on the sampling methodology followed in the ILO surveys are available in Berg et al. (2018).

Table 1.1: Differences across crowdworkers and traditional workers in the US and EU

	US			EU		
	Traditional	Crowdwork	diff.	Traditional	Crowdwork	diff.
Hourly nominal earnings (USD)	30,774	7,208	-23.566***	17,058	6,585	-10.473***
	(207,851)	(7,483)		(91,886)	(28,970)	
Hourly nominal earnings (USD)†	30,774	5,433	-25.341***	17,058	3,901	-13.157***
	(207,851)	(5,079)		(91,886)	(18,574)	
Weekly working hours	39,056	21,180	-17.876***	37,176	14,697	-22.479***
	(11,655)	(20,511)		(11,901)	(24,137)	
Weekly working hours†	39,056	28,266	-10.789***	37,176	19,903	-17.273***
	(11,655)	(26,422)		(11,901)	(32,601)	
Age	41,024	35,027	-5.997***	42,207	35,543	-6.663***
	(12,615)	(10,934)		(11,390)	(11,137)	
Female	0,463	0,476	0.013	0,478	0,426	-0.051***
	(0,499)	(0,500)		(0,500)	(0,495)	
Married or living with a partner	0,516	0,434	-0.082***	0,697	0,493	-0.204***
	(0,500)	(0,496)		(0,459)	(0,500)	
No. of people in household	3,063	2,665	-0.398***	2,882	2,819	-0.063
	(1,672)	(1,429)		(1,268)	(1,260)	
Main earner in household	0,603	0,789	0.186***	0,595	0,815	0.220***
	(0,489)	(0,408)		(0,491)	(0,389)	
Educ.: no high school diploma	0,064	0,009	-0.055***	0,161	0,052	-0.109***
	(0,244)	(0,092)		(0,367)	(0,222)	
Educ.: high school diploma	0,502	0,374	-0.128***	0,448	0,309	-0.139***
	(0,500)	(0,484)		(0,497)	(0,462)	
Educ.: technical/associate	0,097	0,157	0.061***	0,147	0,102	-0.045***
	(0,296)	(0,364)		(0,354)	(0,303)	
Educ.: bachelor's degree	0,208	0,348	0.141***	0,127	0,322	0.195***
	(0,406)	(0,477)		(0,333)	(0,468)	
Educ.: master's degree	0,094	0,097	0.003	0,108	0,165	0.056***
	(0,292)	(0,296)		(0,311)	(0,371)	
Educ.: higher	0,036	0,015	-0.021***	0,009	0,051	0.042***
	(0,185)	(0,122)		(0,092)	(0,219)	
Health: Very Good	0,132	0,244	0.112***	0,261	0,257	-0.003
	(0,338)	(0,429)		(0,439)	(0,437)	
Health: Good	0,407	0,534	0.128***	0,532	0,523	-0.008
	(0,491)	(0,499)		(0,499)	(0,500)	
Health: Fair	0,345	0,180	-0.165***	0,185	0,178	-0.007
	(0,475)	(0,384)		(0,389)	(0,383)	
Health: Poor	0,099	0,037	-0.062***	0,020	0,033	0.013**
	(0,299)	(0,190)		(0,140)	(0,178)	
Health: Very Poor	0,018	0,005	-0.013	0,002	0,008	0.006*
	(0,132)	(0,071)		(0,048)	(0,090)	

*N*otes: Mean-comparison t-tests across crowdworkers (ILO data) and traditional workers (AWCS and EWCS data) from the US and EU. Standard errors in parentheses. Summary statistics and t-test are calculated from weighted US and EU reference samples. The sample is restricted to employed and self-employed individuals in working age. †: adjusted for time spent in unpaid activities.

*p<.05; **p<.01; ***p<.001

We used information from the EWCS and AWCS to construct a baseline group for traditional workers. The AWCS surveys a sample of 3,109 individuals from the US, sharing several dimensions in common with the ILO data. Raked post-stratification weights conforming to the Current Population Survey (CPS) target population are already provided with the survey, and we restricted our sample to employed working

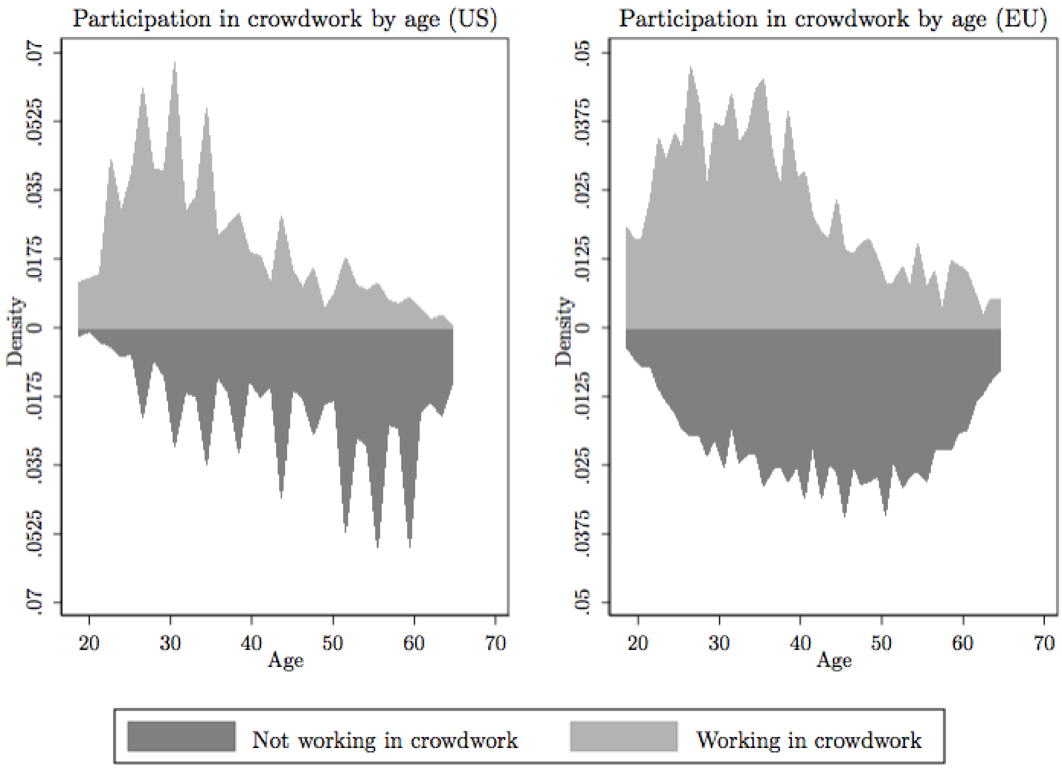


Figure 1.1: Participation in crowdsourcing versus traditional occupations by age

Notes: The figure shows the probability density functions of age by type of work across the US and European samples. Control sample is restricted to employed and self-employed individuals in working age, excluding freelancers working from home.

age population (n = 1,946).[12] Similarly, a sample of 32,429 employed working-age individuals from the EU28 area was extracted from the EWCS, weighted, and paired as a group of traditional workers to the data on European crowdworkers. All data was finally aggregated into a single dataset, providing a shared set of common variables and adjusting earnings for inflation and purchasing power parity.

Summary statistics for all relevant variables are reported in Tables A.1, A.2

[12]For most our estimates, we decided not to narrow our sample of traditional workers based on the respondent's profession. While an analysis of earnings and outcomes across comparable tasks (for example, in terms of routine intensity, as suggested in Autor and Dorn, 2013 and Ottaviano et al., 2013) will not be disregarded, our estimates focus on comparing workers while controlling for their ability, disregarding any bias-inducing factor – such in the case of occupations – that could affect our estimates. For similar reasons, a small number of individuals, which have been reporting to do freelancing work from home as their main occupation, has been omitted from the estimations. This being considered, we restrict our group of traditional workers to individuals in occupations with comparable routine and abstract task-intensity in Table 1.3, so to provide a more complete picture of the crowdworking phenomenon: the results included in said table, for all the aforementioned reasons, are included for descriptive purposes and should be intended void of any causal interpretation.

and A.3 in the Appendix. Weighted mean comparison t-tests for a number of key dimensions across the crowdwork and traditional work groups are shown in Table 1.1 (United States: n = 3,339 and Europe: n = 33,281). Mean comparison t-tests between the two groups, restricted to the employed working age population, reveal differences in earnings, age, education and marital status across forms of work. While earnings, as expected, appear to be lower for online platform workers, their demographic composition also shows significant differences with traditional workers from both the US and the EU, with the typical crowdworker being more likely to be younger, single and more educated overall. These differences are likely explained by the younger relative age of platform workers, being years of schooling and marital status obviously correlated with age. Notably, Figure 1.1 pictures participation in crowdsourcing conditional on age for both forms of work, showing how platform workers tend to occupy those younger age cohorts where individuals are more likely to be excluded from traditional forms of employment.

This age differential affects the likeliness of not being married or having children, explaining the higher propensity of being the main earner in the household and the smaller household size amongst crowdworkers. The condition of caring for children or disabled relatives, as will be discussed later, also appears more common to platform workers.

Looking at each region, differences in earnings also appear to be much more pronounced in the United States than in Europe, where the differential with traditional occupations increases from 10.47 USD in EU to 23.56 USD in US (US: -76.58%, EU: -61.39%). Also, while hourly earnings between crowdworkers in the two regions average at similar levels, the hourly rate of pay among platform workers in the EU is subject to much higher variability, presumably because of differences in the platforms. Similarly, European crowdworkers, on average, appear to work fairly less than their US counterparts. Other disparities emerge in terms of gender (where a male majority is statistically significant in EU), health status and education.

1.4.2 Selected labour market indicators and controls

In order to compare crowdworkers and traditional salaried workers, we selected a number of key labour market indicators. With our data being extracted from different sources, a number of variables have been subjected to re-coding, for the sake of harmonisation. Keeping the changes at a minimal level, the final coding sometimes differs across the US and European samples. In many cases, the changes have been negligible, but will nonetheless be reported when needed.[13]

Hourly nominal earnings have been selected as our key variable of interest. Given that crowdworkers do not have fixed working hours, and considering how some individuals already in employment may work on online labour platforms in their free time to gain access to an auxiliary source of income, hourly rate of pay will not suffer from distortions caused by the number of hours worked per week, allowing for a comparison between platform and traditional workers. This does not imply that hourly earnings are indifferent to crowdwork being the main – or only – source of income for the respondent, a variable for which we will also control.

Another crucial dimension of interest is weekly working hours. Thanks to the ILO survey, we were able to estimate how much time crowdworkers spend on the platform between paid and unpaid tasks. This allowed us to investigate the differential in our earnings estimates between crowdworkers and traditional workers when accounting for unpaid working hours. In all instances, availability of weekly working hours proved essential for computing hourly earnings, as all surveys do not report the hourly rate of pay, but rather weekly, monthly or yearly absolute earnings.[14]

Additionally, along with indicators of skill use and job satisfaction, the EWCS, AWCS and ILO surveys contain items for identifying if the surveyed individuals would like to work more than what they currently do or whether they are currently looking for another occupation,[15] serving as proxies for labour use in the platform. This enabled us to identify involuntary crowdwork as a dimension that goes beyond standard employment statistics.

In our analysis we consider a number of controls. We first control for age, gender and education and, from there, we add other predictors. In the literature, returns to education on earnings have been widely documented,[16] while gender pay gaps have also been studied thoroughly.[17] We can also expect marital status and the number of people living in the household to affect earnings and working conditions in general. Finally, we control for state specific effects and for whether the respondent is the main earner of his household. Another fundamental variable in our analysis is caregiving, indicating whether the respondent has been involved in full-time caring for children or disabled/elderly relatives. The implications of this variable for our

[13] This is the case for education, where achievements were grouped to the closest common title, while other similar adjustments were made to marital status.

[14] While the ILO survey reports weekly earnings, AWCS reports yearly earnings, and EWCS lets the respondent to choose the measure he/she is most comfortable with. Hourly rate was then computed by dividing weekly nominal earnings by weekly working hours.

[15] This last item was however only recorded in the AWCS and ILO.

[16] See, e.g., Angrist and Krueger (1991) and Card and Krueger (1992).

[17] E.g., Blau and Kahn (2003) and, Altonji and Blank (1999).

2SLS model will be discussed later.

1.5 Model specification

We estimate the effect of working in online platforms on labour market outcomes comparing earnings and working conditions between platform and 'traditional' workers, which we treat as two distinct groups. In our case, the first group is composed by crowdworkers interviewed in the ILO survey, while the second group includes workers from the AWCS and EWCS surveys. From this point of view, our approach has drawn inspiration from LaLonde (1986), while our identification strategy is not dissimilar from previous studies on part-time employment which instrumented hours of work through household size and fertility (such as Ermisch and Wright, 1993, and Hotchkiss, 1991).

As platform workers are usually paid by task, and not by hour, hourly earnings are determined first by the demand for those specific skills and characteristics over which clients can discriminate upon (factors which we can mostly control for with our set of observable covariates) and, on the supply side, by the ability of each individual worker to complete these tasks efficiently (which is mostly unobserved).

Simple descriptive analyses may then produce biased results, potentially overestimating the effect of the platform economy on wages and working conditions. Indeed, it could be argued that individuals in crowdsourcing arrangements possess characteristics which make them qualitatively different from traditional workers, thus leading to a problem of self-selection into online labour markets. To account for this potential selection bias and offer a more appropriate comparison between the different outcomes, we initially compare outcomes across types of workers, controlling for observable characteristics with an OLS model, and later offer further controls for unobservable skills by adopting an instrumental variable approach. For our instrumental variable model, we choose the following specification:

(1) $Y_i = \alpha_2 + \hat{T}_i\lambda + X_i'\gamma_2 + F_i\varphi_2 + e_{2i}$

(2) $T_i = \alpha_1 + Z_i\phi + X_i'\gamma_1 + F_i\varphi_1 + e_{1i}$

where i refers to each individual, Y is the set of our outcome variables (natural logarithm of hourly earnings and of hourly earnings adjusted for unpaid activities), while X is a vector of k–2 controls, and F is a dummy which indicates whether the respondent is female.[18] The full set of controls in the X vector are age, age

[18]The need for this specification, with the gender dummy appearing outside the X vector, will be explained in subsection 5.1, as the coefficient φ_2 will be used to adjust split-sample estimates to the whole population.

squared, number of people in household, main earner (i.e. if an individual is the main earner in the household), main source of income (i.e. if the reported earnings refers to the individual's main source of income), education level, marital status and state/country of residence.

In the first stage regression (2), working in crowdwork T (a dummy which equals 1 when crowdwork is the individual's main paid activity) is regressed on our chosen instrument Z plus the same controls we use in the second stage regression (1). Using the predicted value of T (the estimated linear probability of working in the platform) in (1), we obtain the impact of crowdwork on our desired outcome through the coefficient λ. In case crowdwork T is really assigned exogenously conditionally on Z, then the coefficient on λ will not suffer from selection bias.

1.5.1 Instrumental variable identification

A number of exogenous variables, such as age or health condition, are significantly correlated with crowdwork (age: -0.1643***; poor health: 0.0193***).[19] Their adoption as instrumental variables, however, would potentially lead to a violation of the exclusion restriction, biasing our estimates downwardly: younger workers typically earn less than older individuals, while workers in poor health may take longer times to complete their work activities, leading to a reduction in hourly earnings.

We then considered a third instrument: time spent in caregiving at home. This variable is potentially highly correlated with crowdwork. The underlying reasoning is that people may be more involved in crowdwork if they are compelled to stay at home to look after children or elderly relatives: this type of work, indeed, can be a reasonable source of extra income to these individuals, given their circumstances.

Both the ILO and the AWCS-EWCS datasets capture time spent in caregiving at home, although in different ways. While caregiving appears as a dummy in the ILO dataset (where the respondent is asked whether this activity constituted a full-time commitment before entering crowdwork), it is treated as a continuous variable in the AWCS and EWCS (where the respondent is asked how many hours per week/per day has been engaged in these activities). We harmonised the two variables by identifying both a 40 and 15 hours-per-week effort as a full-time caring activity, following the findings from the Gallup-Healthways Well-Being Survey. Indeed, according to this study, caregivers working at least 15 hours per week have declared that this activity significantly affected their work life.[20]

[19] Sidak-adjusted pairwise correlations. Survey question: : "Do you have any illness or health problem which has lasted, or is expected to last, for 6/12 months or more?".

[20] For details about the Gallup-Healthways Well-Being Index, see https://www.gallup.com/175196/gallup-healthways-index-methodology.aspx. For

Caregiving appears to be highly correlated with crowdwork in our US sample (estimated correlations: caregiving 15h = 0.0521***; caregiving 40h = 0.1698***). This relationship is similar in Europe where caregiving also reveals itself as a significant predictor of platform work, but only at higher thresholds (caregiving 40h=0.0933***). These differences hint at the possibility of welfare-biased differential effects of caregiving, as caregivers may have access to more labour law safeguards in Europe than in US, reducing the need for auxiliary earnings from crowdwork. Evidence from Germany (Bick, 2016), indicates that a large fraction of working mothers in part-time would work full-time if they had greater access to subsidised child care. It is then not unreasonable to expect labour market policies to similarly influence participation in crowdwork.

While the connections between crowdsourcing and caregiving are theoretically plausible and empirically proven, the choice of this instrument, however, can raise concerns with regards to its endogeneity and to the risk of violation of the exclusion restriction. These concerns, however, can be overcome, as discussed below.

1.5.2 Gender bias in caregiving

Caregiving appears to be consistently correlated with the gender of the respondent: females are over-represented among crowdworkers who are caregivers, with the correlation between being in caregiving (40h) and crowdwork raising from a full sample (US+EU) correlation coefficient of 0.1920*** to 0.2502*** for the female population. This differential may support prior evidence on men's caregiving being a complex phenomenon influenced by endogenous socio-economic determinants,[21], uncovering a potential obstacle in our identification strategy. Indeed, while a number of studies finds caregiving to be exogenous to the female population (see, as discussed later Ciani, 2012, and Schmitz and Westphal, 2017), the effect on the male population is less unambiguous.

Nonetheless, we trust that these complications can be overcome by assuming that platform work has no intrinsic effect on gender-dependant outcomes, arguing that, after controlling for individual's characteristics and ability, crowdwork arrangements do not tend to reinforce discrimination based on the sex of the worker, due to the relative anonymity that service providers enjoy on the platform:[22] clients are, indeed, usually unable to ascertain the gender of online service providers. Should

the Gallup evidence about the relevant threshold levels for caregiving, see https://news.gallup.com/poll/148640/one-six-american-workers-act-as-caregivers.aspx.

[21]See, for example, Gerstel and Gallagher (2001).

[22]As found in Adams, Abi and Berg, Janine, (2017) "When Home Affects Pay: An Analysis of the Gender Pay Gap Among Crowdworkers"

this assumption hold, all differences between genders will then be linked to common structural trends across traditional and platform forms of work which can be identified linearly, and the interaction between gender and the selected instrument can be added to the instrument pool in the first stage of the estimation process. In other words, if the interaction term between gender and crowdwork yields a zero effect on earnings, said interaction can be added to the instrument pool without expecting violations of the exclusion restriction.

Additionally, the 2SLS estimates that can be drawn from the pool of female workers can be also said to hold for the rest of the sample. Given that earnings are estimated by a log wage equation, the non significance of interaction effect (which we will denote as ζ, omitting, from now on, the second-stage index from equation 2) allows the non-interacted gender effect to be fully absorbed by the constant term in the split sample estimate, leading, after controlling for all observables, to:

(3) $\lambda \approx \lambda_f + \zeta$

meaning that λ_f, the effect of platform work on the female population as predicted by our model will approximate the full sample coefficient λ minus the interaction term ζ. If this interaction term is not statistically different from zero, λ_f will also closely approximate the baseline effect of platform work on the selected dependent variable. As our 2SLS estimation will be based on the full US-EU sample,[23] region-specific differential gender effects can also be isolated by the coefficient of the interaction between gender and the regional dummy, and then applied to the final estimates using a similar procedure, if needed. In first part of our analysis, we will show that the coefficient of the interaction term between crowdwork and gender is not statistically different from zero when controlling for other observables, allowing us to generalise the common structural term predicted with φ.

Split sample instrumental variable models – or TS2SLS – have already been explored in the past by Angrist and Krueger (1995) and Inoue and Solon (2010), who address those events when the instrument and the outcome are not measured in the same sample. In our case, however, the two subsamples – male and female – are not homogeneous. It is vital, then, to assume the differences between the two subsamples to be linear and, most importantly, to assume the structural relations within them to remain the same.

[23] Since our chosen instrument affects participation in crowdwork but is not intended to randomise regional assignment, differential effects across countries become a second-order priority. Hence controls for specific regional differences are sufficient for the estimation of these effects, with the 2SLS estimation benefiting from the increase in sample size for all groups.

1.5.3 Potential earnings and violations of the exclusion restriction

The chosen instrumental variable – caregiving – could also pose as a threat to our identification strategy in terms of violations of the exclusion restriction. Indeed, it is reasonable to believe that the amount of time a worker spends in caregiving is endogenous to the wage he/she could earn in the market. While this condition should clearly affect working hours and – by extension – total earnings (as documented in Wakabayashi and Donato, 2005, and Earle and Heymann, 2012), the effects on hourly earnings are less obvious. If the wage is high enough, individuals could, in fact, purchase care for either a child or a relative and, in such a case, transition to caregiving will be biased towards lower salaries.

To properly account for this endogenous variation, we would need to have access to a measure of unobserved potential earnings, i.e., the hourly salary a worker would earn before being engaged in caregiving. In case potential earnings are available, it holds that:

(4) $E[Y_{1,0}|Y_{0,0}, T] - E[Y_{1,c}|Y_{0,c}, T] = \xi.$

This means that the expected value of earnings of workers in caregiving $E[Y_{1,c}]$ would equal the expected value of earnings of individuals not in caregiving $E[Y_{1,0}]$, minus the caregiving bias ξ, once we control for occupation T (traditional or crowdwork) and unobserved potential earnings Y_0. This, of course, implies that $Y_{0,0} = Y_{0,c}$, meaning that, all else being equal, potential earnings are indifferent to caregiving.

However, we are only able to observe effective earnings. While we cannot observe potential earnings, we can nonetheless proxy for them through our set of controls X, which we believe can correctly predict them. If this is the case, Y_0 can then be replaced by our set of controls X. This will also allow us to filter out other endogenous aspects of caregiving linked, for example, with household size and marital status.

As it will be discussed later, our findings – see Table 1.6 below – suggest that hourly earnings are unaffected by caregiving after controlling for other observables. This implies that not only this bias is absent, but also, as long as there is no reason to suspect ξ to be positive, that the chosen set of controls correctly proxies for potential earnings as in equation 4.

In other words, while access to caregiving is most probably related to the inability to purchase formal care, we believe our set of control to correctly predict potential earnings and appease concerns concerning the violation of the exclusion restriction, given that the effect of caregiving on the hourly rate of pay appears null after controlling for these these variables.

Evidence from the literature provides support to these findings, as Ciani (2012) shows that caregiving, while affecting labour market participation, can be, in most cases, assumed as exogenous. Similarly, Leigh (2010) finds a direct effect on participation, while observing no impact on hourly wages, while Schmitz and Westphal (2017) also find female caregiving to be related to a decrease in probability in working full time, with no short-term effect on hourly wages.

Another potential concern comes from the fact that caregiving could influence skills, and thus the returns to them. However, when potential earnings are properly controlled for, caregiving is likely unable to influence ability. Caregiving could affect the opportunity to work more, not the relative skills of an individual – or how much the labour market rewards these skills.

The test presented in Table 1.6 is comparable to the zero-first stage test presented in Bound and Jaeger (2000), Altonji et al. (2005), and Angrist et al. (2010). Essentially, as we also present results from the effect of caregiving splitting the sample conditionally on the work type (crowdwork or traditional work), we are testing the effect of the instruments on samples where there is no first stage, as assignment into online working arrangements is undefined. Accordingly, individuals in caregiving who retain their status in traditional occupations should earn a lower hourly salary than comparable workers if transition to caregiving is only linked to the ability to purchase formal care, and cannot be properly controlled for by the other observables. This, however, does not happen, and suggests that the choice of caregiving as an instrument does not violate the exclusion restriction. As caregiving influences participation in crowdwork, crowdworkers in caregiving can, however, earn more than crowdworkers not in caregiving: in this case, the zero coefficient identified by those equations will then suggest, anticipating our final interpretation, that earnings in crowdwork are indifferent to individual ability.

While these are convincing arguments with regards the validity of caregiving as an instrument for hourly earnings, similar reasonings, however, prevent the use of the same instrument for the estimation of the effects of crowdwork on other outcomes. Indeed, as the crowdwork 'complier' group[24] will include individuals spending a significant amount of time in caregiving, we can expect the 2SLS estimates of working hours and weekly earnings to suffer from a downward bias, as will be discussed in the next section.

[24] We here define as 'compliers' all individuals in caregiving who participate in crowdwork arrangements and all individuals not in caregiving who stay in traditional forms of work.

1.6 Discussion of results

Columns (1) to (3) from Table 1.2 present initial OLS results, using a sample of US crowdworkers from the ILO survey and regular workers from the AWCS. The dependent variable is hourly earnings and additional controls are added with each specification, with an initial sample including a total of 3,128 workers.[25] The coefficient for the dummy variable for working in crowdwork will denote the earnings differential between occupations.

Additional key controls are: gender, age (and its squared term), number of people in the household, marital status (whether the respondent is married or lives with a partner), and two dummies indicating whether the respondent is the main contributor to the household's income and whether crowdwork is his/her main source of income. We also take into account a set of control dummies for the different US and EU28 states of residence (with a total of 79 states) and for the level of education (distinguishing among six different education levels). Standard errors are also robust to clustering on the level of the federal or member state.

As shown in the table, the effect of crowdwork on earnings is always negative and significant. The effect of the female dummy also remains negative and significant, confirming the presence of a gender pay gap in all labour markets. In the third column we present our full specification: all the relevant regressors, controls and interactions are included. The regression shows that crowdwork has a negative and significant effect (indicating a 63.6% reduction in earnings),[26] while both dummies for being the main earner in the family and for the surveyed occupation being the respondent's main job are positive and significant.

Controlling for all other observables, the interaction term between gender and crowdwork is not statistically different from zero, while, most notably, the coefficient on gender alone retains its magnitude and significance, showing a negative linear effect on earnings (-16.5%) and no notable variation between specification (2) and (3), where the interaction is introduced. This finding provides support to our hypothesis that crowdwork platforms do not generate any intrinsic gender discriminatory effect other than reaffirming common structural gaps and, as discussed earlier, provide support to our identification strategy.

Columns (4) to (6) present the estimates for the effect of crowdwork on hourly earnings on the European sample. Here the initial number of complete observations

[25] Observations with missing values are excluded from the estimation.

[26] Given the magnitude of the effect of crowdwork on earnings, it should be noted that log normal interpretations might be incorrect since the parameters are far above the 0.1 threshold and must then be exponentiated.

Table 1.2: OLS estimates of the effect of online platform work on earnings in the US and EU

VARIABLES	(1) US OLS	(2) OLS	(3) OLS	(4) EU OLS	(5) OLS	(6) OLS	(7) US+EU OLS
Working in crowdwork	-1.032***	-1.010***	-1.012***	-1.198***	-1.116***	-1.067***	-1.007***
	(0.036)	(0.043)	(0.055)	(0.072)	(0.043)	(0.049)	(0.043)
Female	-0.245***	-0.179***	-0.181***	-0.127***	-0.074***	-0.071***	-0.195***
	(0.042)	(0.040)	(0.061)	(0.010)	(0.009)	(0.010)	(0.061)
$Crowdwork \times Female$			0.004			-0.103*	-0.043
			(0.069)			(0.051)	(0.054)
$EU \times Female$							0.129**
							(0.059)
Age	0.052***	0.030**	0.030**	0.023***	0.011***	0.011***	0.014***
	(0.010)	(0.014)	(0.014)	(0.004)	(0.004)	(0.004)	(0.003)
Age squared	-0.001***	-0.000*	-0.000*	-0.000***	-0.000	-0.000	-0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
No. of people in household		-0.032**	-0.032**		0.002	0.002	-0.004
		(0.013)	(0.013)		(0.007)	(0.007)	(0.006)
Married or living with a partner		0.252***	0.252***		0.099***	0.100***	0.115***
		(0.036)	(0.038)		(0.010)	(0.010)	(0.011)
Main earner in household		0.348***	0.348***		0.137***	0.137***	0.155***
		(0.050)	(0.050)		(0.013)	(0.013)	(0.014)
Main source of income		0.147***	0.146***		0.127*	0.133*	0.156***
		(0.042)	(0.042)		(0.066)	(0.068)	(0.043)
Observations	3,218	3,217	3,217	27,758	27,676	27,676	30,893
Adjusted R-squared	0.361	0.389	0.389	0.367	0.377	0.377	0.378
State controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Education controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

*N*otes: State clustered standard errors in parentheses. Control sample restricted to employed and self-employed individuals in working age, excluding freelancers working from home.

*p<.05; **p<.01; ***p<.001

is 27,578, referring to the total number of EU28 workers included in the ILO and EWCS sample. The sign and magnitude of the crowdwork coefficient is always negative and significant and, after controlling for all covariates in column (6), the effect is now much closer to our estimate for the US sample, equalling to a 65.5% reduction in hourly earnings. The effect of the gender dummy is also negative and significant, this time indicating a smaller reduction in earnings (-6.8%). A negative gender effect can also be found across European crowdworkers, albeit with a 5% statistical significance.

A significative improvement in our estimates is offered in column (7), where a full sample (US+EU) specification is presented. The difference in general region-specific gender effects is isolated by the coefficient of the $EU \times Female$ interaction term, whose positive effect counteracts the negative sign of the $Female$ term, now referring to the baseline US sample.[27] Most importantly, the $Crowdwork \times Female$ interaction turns again not significant, as its effect seems to be recaptured by the regional gender effects, confirming that crowdwork platforms do not generate any intrinsic gender discrimination on earnings. Finally, the effect of crowdwork on PPP-adjusted net hourly earnings is estimated up to a 63.5% reduction. Also, in all instances, the negative effect of working in digital labour market is slightly reduced when crowdwork is the main source of income.

We also test for the presence of differential returns to observable skills in crowdwork by interacting crowdwork with education and maintaining the same specification from Table 1.2. The results suggest that disparities between traditional and platform workers persist and increase with the level of education. Most importantly, crowdwork arrangements appear to offer almost no return to observable skills, as the negative interaction coefficients mostly cancel out the returns to education in traditional occupations (Table B.1, Appendix).

Table 1.3 presents the results of our OLS regressions where we take into account the degree of routine intensity and abstractness of the tasks performed, with reference to both the traditional work and crowdwork samples. It is worth pointing out that, while occupation could be considered a poor choice for a control and, by inducing bias in the estimates, certainly cannot be used in the 2SLS estimation stage unless a different instrument is chosen, it is however true that an analysis which focuses only on the individuals who perform similar occupations can enhance our ability to explore the actual wage premium of traditional workers with respect to platform workers.

[27]The regional dummy for EU (not significant, as its effects are fully captured by the state controls) is omitted from the table.

Table 1.3: OLS estimates of the effect of online platform work on earnings in the US and EU

VARIABLES	(1) US OLS routine tasks	(2) OLS abstract tasks	(3) OLS a+r tasks	(4) EU OLS routine tasks	(5) OLS abstract tasks	(6) OLS a+r tasks	(7) US+EU OLS routine tasks	(8) OLS abstract tasks	(9) OLS a+r tasks
Working in crowdwork	-1.377***	-1.276***	-1.183***	-1.093***	-1.025***	-1.045***	-1.117***	-1.032***	-1.043***
	(0.105)	(0.163)	(0.150)	(0.050)	(0.052)	(0.048)	(0.046)	(0.054)	(0.057)
Female	-0.305**	-0.637**	-0.331	-0.074***	-0.082***	-0.087***	-0.157**	-0.197***	-0.135**
	(0.115)	(0.269)	(0.345)	(0.015)	(0.016)	(0.014)	(0.062)	(0.073)	(0.056)
$Crowdwork \times Female$	0.124	0.454	0.145	-0.098*	-0.088	-0.082*	-0.068	-0.036	-0.076*
	(0.117)	(0.279)	(0.350)	(0.049)	(0.054)	(0.046)	(0.048)	(0.065)	(0.045)
$EU \times Female$							0.082	0.112*	0.047
							(0.060)	(0.064)	(0.053)
Age	0.015	0.016	0.011	0.011**	0.009**	0.012***	0.011**	0.010***	0.012***
	(0.011)	(0.011)	(0.011)	(0.005)	(0.003)	(0.004)	(0.004)	(0.003)	(0.004)
Age squared	-0.000*	-0.000*	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000*	-0.000*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
No. of people in household	-0.063***	-0.057***	-0.049***	0.007	-0.001	0.003	-0.001	-0.005	-0.004
	(0.012)	(0.012)	(0.009)	(0.010)	(0.007)	(0.009)	(0.009)	(0.007)	(0.009)
Married or living with a partner	0.093**	0.092	0.070	0.080***	0.107***	0.103***	0.086***	0.109***	0.108***
	(0.044)	(0.061)	(0.051)	(0.017)	(0.015)	(0.017)	(0.016)	(0.014)	(0.016)
Main earner in household	0.036	0.009	-0.028	0.127***	0.113***	0.117***	0.125***	0.110***	0.111***
	(0.061)	(0.079)	(0.076)	(0.019)	(0.019)	(0.020)	(0.017)	(0.018)	(0.019)
Main source of income	0.065	0.053	0.041	0.120*	0.117*	0.102	0.120***	0.113***	0.097**
	(0.046)	(0.044)	(0.044)	(0.069)	(0.063)	(0.067)	(0.039)	(0.037)	(0.038)
Observations	1,658	1,484	1,415	15,006	20,341	11,107	16,664	21,825	12,522
Adjusted R-squared	0.377	0.178	0.132	0.422	0.373	0.426	0.434	0.376	0.427
State controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Education controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

*N*otes: State clustered standard errors in parentheses. Control sample restricted to occupations whose routine and abstract task content is comparable to the 5th and 95th percentile of crowdwork occupations by their routine and abstract task content.

*p<.05; **p<.01; ***p<.001

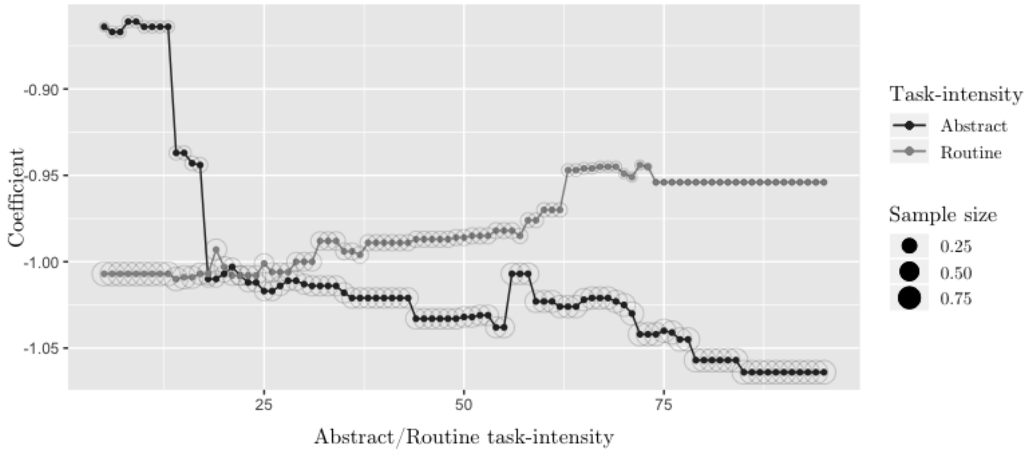


Figure 1.2: Estimated OLS coefficients from varying task-intensity splits (US+EU)

Notes: OLS coefficients for the 'Working in crowdwork' dummy after restricting the control sample (US+EU) by increasing routine task-intensity and decreasing abstract task-intensity. Sample sizes from each estimation are reported as a percentage of the full control sample. Control sample restricted to employed and self-employed individuals in working age, excluding freelancers working from home.

To this aim, we break down the two groups of workers finding regions of common support based on the degree of routine task intensity, abstractness, and a combination of the two indicators. We assign routine and abstract task intensity scores to individuals in the traditional occupations using the indicators from Autor and Dorn (2013), where each occupation is given a score based on O*NET task measures. We then compute, using a similar methodology, the same scores from the ILO sample, disaggregating each observation into the five most common tasks, and assigning each task a score based on the routine and non-routine cognitive O*NET measures, as reported in Acemoglu and Autor (2011), and then averaging the scores after re-weighting each task by its relative frequency. Finally, we restrict the group of traditional workers to those observations whose routine and non-routine task intensity falls within the range of scores obtained in the crowdwork sample.

Our results show that the coefficients do not diverge excessively from our initial results, displaying a negative – and slightly stronger – effect on earnings for platform workers, in all the regressions considered (US, EU, US+EU), indicating that the routine and abstract content of micro-task jobs might not capture the reduction in earnings from traditional professions in any way.

As we cannot ascertain the full comparability of the routine and abstract task-intensity scores between crowdworkers and traditional workers, we provide a further robustness check in Figure 1.2, where we restrict the control sample by decreasing abstract and increasing routine task-intensity scores, and estimate the 'Working in crowdwork' coefficient (y-axis) using the same least squares specifications from Table 1.3 (columns 7 and 8). The x-axis indicates the minimum abstract task-intensity and the maximum routine task-intensity score used for the sample split.

The figure suggests that, the more the maximum abstract intensity of traditional occupations is lowered, the more the effect of crowdwork on earnings is reduced. A similar decrease is found when we raise the minimum routine content for regular occupations. Nevertheless, our previous interpretation is not invalidated: these contractions in the effect of crowdwork on earnings remain minimal, as we consider that the coefficient fully maintains its sign and significancy, and that the estimated effect ranges from 57.8 to 65.5% only when performing splits on abstract intensity, and from 63.5 to 61.5% when increasing the minimum routine content. The great majority of the earnings differential between platform and traditional work remains then unexplained by the abstract and routine task-intensity of crowdsourcing.

OLS estimates for working hours indicators are shown in Table 1.4. When investigating time spent on the platform, the estimates appear particularly sensitive to the way working hours are computed. In particular, in columns (1), (4) and (7) we find that, on average, when only paid activities are considered, working in crowdwork reduces the number of weekly working hours by 16 hours, also indicating a 7 hours differential between US and the EU platform workers. When crowdwork is also the main source of income, these figures are further reduced, and all crowdworkers appear to be working circa 7 hours less than traditional workers, all else being equal.

If, however, the indicator is adjusted for the time spent in unpaid tasks – as in columns (2), (5) and (8), Table 1.4 – the magnitude of the coefficient changes again, showing a 9 hours increase in working hours across the US and the EU. For individuals whose main occupation is crowdwork, the differential with the control is reduced even more, to the point that, on average, US crowdworkers appear to be working even more than comparable workers. Significant disparities with the European sample remain, indicating that, for EU workers, there is no discernible difference in working hours between platform and traditional workers when crowdwork consists in the main source of income of an individual.

Moving to factor utilisation, we are presented with some intriguing figures. In (3), (6) and (9), Table 1.4, our OLS model suggest that most platform workers would like to work more than they currently do in either crowdwork or in other forms of employment, suggesting a degree of factor under-utilisation. While not shown in the table, we also found out that these figures are halved when respondents are

Table 1.4: OLS estimates of the effect of online platform work on working hours in the US and EU

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	US			EU			US+EU		
	Work Hours	Work Hours†	More work	Work Hours	Work Hours†	More work	Work Hours	Work Hours†	More work
VARIABLES	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS
Working in crowdwork	-13.792***	-3.009*	0.537***	-21.180***	-15.064***	0.299***	-16.268***	-7.208***	0.452***
	(1.207)	(1.742)	(0.037)	(1.992)	(2.815)	(0.051)	(1.546)	(2.308)	(0.042)
Female	-3.985***	-3.842***	-0.020	-5.516***	-5.521***	-0.001	-4.661***	-4.757***	0.011
	(0.719)	(0.747)	(0.039)	(0.564)	(0.566)	(0.014)	(0.806)	(0.893)	(0.045)
$Crowdwork \times Female$	3.340**	3.170**	0.093**	5.989***	7.578***	0.013	4.686***	5.384***	0.050
	(1.330)	(1.557)	(0.042)	(2.067)	(2.307)	(0.040)	(1.361)	(1.562)	(0.040)
$EU \times Female$							-0.771	-0.662	-0.015
							(0.885)	(0.978)	(0.042)
Age	0.934***	1.150***	-0.011**	0.614***	0.632***	0.000	0.664***	0.714***	-0.001
	(0.267)	(0.332)	(0.005)	(0.111)	(0.111)	(0.003)	(0.107)	(0.113)	(0.003)
Age squared	-0.011***	-0.013***	0.000*	-0.007***	-0.008***	0.000	-0.008***	-0.008***	0.000
	(0.003)	(0.004)	(0.000)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)
No. of people in household	-0.345**	-0.408*	0.024**	-0.267*	-0.267*	0.011**	-0.301**	-0.327**	0.013**
	(0.164)	(0.207)	(0.009)	(0.146)	(0.146)	(0.005)	(0.125)	(0.127)	(0.005)
Married or living with a partner	0.775	-0.115	-0.137***	1.565***	1.558***	-0.017	1.487***	1.391***	-0.026
	(0.881)	(1.012)	(0.025)	(0.226)	(0.225)	(0.016)	(0.235)	(0.264)	(0.016)
Main earner in household	5.000***	4.985***	-0.126***	3.701***	3.694***	0.028**	3.828***	3.815***	0.017
	(0.950)	(1.038)	(0.023)	(0.682)	(0.665)	(0.011)	(0.621)	(0.612)	(0.012)
Main source of income	10.423***	15.656***	0.025	5.376***	7.252***	0.048*	9.234***	13.581***	0.101***
	(1.051)	(1.654)	(0.020)	(1.872)	(2.104)	(0.026)	(0.996)	(1.491)	(0.022)
Observations	3,217	3,197	3,216	27,676	27,649	27,129	30,893	30,846	30,345
Adjusted R-squared	0.303	0.168	0.371	0.218	0.175	0.075	0.242	0.173	0.100
State controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Education controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: State clustered standard errors in parentheses. Control sample restricted to employed and self-employed individuals in working age, excluding freelancers working from home. †: adjusted for time spent in unpaid activities

*p<.05; **p<.01; ***p<.001

asked whether they would prefer to work in non-crowdwork occupations (even when crowdwork is the main source of income). These findings partially confute the perception of platform work as a temporary form of occupation for the underemployed, configuring it as a rather stable condition with unremarkable mobility towards other forms of employment – for many, at least.

However, even if not actively looking for a job, this status presents some uncanny similarities to the ones of involuntary part-timers or inactive persons with labour force attachment, where individuals would like to work more but are unable or too discouraged to look for other forms of employment, and, for that, crowdwork could be found to be related to slack in the labour market, possibly linked to a scarcity in demand.

These results are consistent with the interpretation of Katz and Krueger (2018), who find slack in online platform work to be mostly involuntary and linked to economic reasons. Also, the idiosyncratic relationship between working nearly as many hours as traditional workers while still desiring to work more, alongside with the largely low earnings, may corroborate the findings from Horton and Chilton (2010), if we inductively assume that platform workers are usually unable to meet their earnings targets. It should be noted, however, that while these remarks could reflect the condition of many online workers, crowdwork could still represent a convenient source of auxiliary income for many others.

We now turn turn to our IV estimates for the effect of platform work on hourly earnings, which are displayed in Table 1.5, together with OLS estimates for both the full sample and a female-only sample.[28] In the 2SLS regressions the estimates for the full sample and the female sample show both weak predictive power when instrumenting caregiving with a 15 hours weekly threshold (columns 3 and 4): while the first stage displays a high R-squared, the crowdwork coefficient is never statistically different from zero and the instrument always fails to pass the F score test for excluded instruments.

The 40 hours threshold generates instead much more reasonable coefficients for working in crowdwork (columns 5 and 6), predicting a general and statistically significant reduction (-63.46%; coeff.: -1.007) in hourly earnings. While very close to our OLS estimates, it could be argued that these estimates still suffer from bias due to endogenous caregiving in the male sub-sample. Restricting our study to the female population, working on crowdwork platforms reduces earnings by 60.07% (column 6, coeff.: -0.918) over working age women, all else being equal. This is

[28] In the former, caregiving and the its interaction with gender is instrumented; in the latter, only caregiving is.

Table 1.5: 2SLS estimates of the effect of online platform work on earnings in the US and EU

	(1)	(2)	(3)	(4)	(5)	(6)
	US+EU	US+EU	US+EU	US+EU	US+EU	US+EU
			Caregiving (15h)		Caregiving (40h)	
	OLS	OLS	2SLS	2SLS	2SLS	2SLS
VARIABLES	full sample	female only	full sample	female only	full sample	female only
Working in crowdwork	-1.028***	-1.055***	0.518	0.902	-1.007***	-0.918***
	(0.041)	(0.056)	(1.060)	(1.158)	(0.247)	(0.236)
Female	-0.212***		-0.284***		-0.213***	
	(0.045)		(0.059)		(0.049)	
EU × Female	0.145***	1.348***	0.207***		0.146***	
	(0.046)	(0.089)	(0.053)		(0.048)	
Age	0.014***	0.015***	0.020***	0.018***	0.014***	0.016***
	(0.003)	(0.006)	(0.003)	(0.006)	(0.004)	(0.005)
Age squared	-0.000**	-0.000*	-0.000***	-0.000**	-0.000**	-0.000*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
No. of people in household	-0.004	-0.009	-0.006	-0.010	-0.004	-0.009
	(0.006)	(0.008)	(0.005)	(0.008)	(0.006)	(0.008)
Married or living with a partner	0.115***	0.104***	0.121***	0.081***	0.115***	0.102***
	(0.011)	(0.018)	(0.011)	(0.020)	(0.011)	(0.018)
Main earner in household	0.155***	0.124***	0.122***	0.075**	0.154***	0.120***
	(0.014)	(0.017)	(0.024)	(0.031)	(0.016)	(0.019)
Main source of income	0.153***	0.156***	1.503*	1.818*	0.171	0.271
	(0.042)	(0.057)	(0.894)	(0.975)	(0.227)	(0.211)
Observations	30,893	15,921	30,893	15,921	30,893	15,921
Adjusted R-squared	0.378	0.366	0.151	0.051	0.255	0.231
State controls	Yes	Yes	Yes	Yes	Yes	Yes
Education controls	Yes	Yes	Yes	Yes	Yes	Yes
F-Test			3.968	4.657	12.40	23.25
First Stage R^2			0.738	0.712	0.742	0.722

*N*otes: State clustered standard errors in parentheses. Control sample restricted to employed and self-employed individuals in working age, excluding freelancers working from home.

*p<.05; **p<.01; ***p<.001

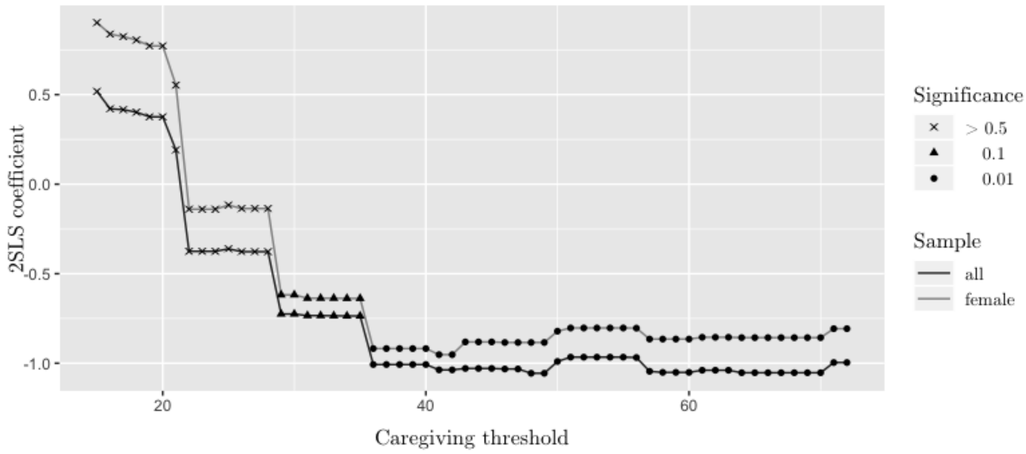


Figure 1.3: Estimated 2SLS coefficients from varying full-time caregiving thresholds (US+EU)

*N*otes: Second-stage coefficients for the "Working in crowdwork" dummy instrumented through a caregiving instrument with increasing weekly hours threshold. Control sample restricted to employed and self-employed individuals in working age, excluding freelancers working from home.

Table 1.6: Effect of caregiving on hourly earnings (US+EU)

	(1)	(2)	(3)	(4)	(5)	(6)
	Full sample			Female only		
	C+T	T	C	C+T	T	C
VARIABLES	OLS	OLS	OLS	OLS	OLS	OLS
Caregiving (15h)	0.008	0.005	0.032	0.033*	0.026	0.116
	(0.012)	(0.012)	(0.060)	(0.019)	(0.020)	(0.076)
Caregiving (40h)	-0.015	-0.009	0.032	0.017	-0.011	0.116
	(0.030)	(0.029)	(0.060)	(0.032)	(0.030)	(0.076)
Observations	30,893	28,699	2,194	15,921	14,921	1,000
Control covariates	Yes	Yes	Yes	Yes	Yes	Yes

*N*otes: "C+T"(crowdwork and traditional work samples), "C" (crowdwork sample), "T" (traditional work sample). Notes: State clustered standard errors in parentheses. Dependent variable: natural logarithm of hourly PPP adjusted nominal earnings (US dollars). The dummy caregiving is first set at the 15h and then at 40h threshold, and the sample is reduced to the traditional work (AWCS+EWCS) groups in (2) and to the crowdwork (ILO) group in (3). Covariate list: age, age squared, number of people in household, main earner, main source of income, education, marital status, health status and state controls.

*p<.05; **p<.01; ***p<.001

well below the -1.05 (-65.18%) log points that the least squares model would predict over the female sample (column 2). In both cases, anyway, all instruments pass the F score tests for excluded instruments, with the first-stage partial R^2 also yielding remarkable results (see Bound et al., 1995). Complete first stage regressions are shown in Appendix B.

As discussed earlier, while the exogeneity of the instrument on the male population can be disputed, the literature points at caregiving being exogenous to the female population, implying that, if randomisation is achieved though this channel, the -0.918 coefficient could be considered close to an unbiased parameter of the effect of crowdwork on the earnings of the whole population, given that these online labour platforms do not seem to generate further gender gaps in earnings. After generalising the split-sample estimates as in equation (3), we obtain a baseline reduction in earnings of 60.07%, raising our confidence in the results from the previous full sample specification. This interpretation holds even if we assume presence of gender based self-selection into the crowdworker population: should this hypothesis be true, then only full sample estimates would be biased. Since, however, we are now interested in the effect of earnings, irrespective of gender, this estimate could be considered appropriate for both men and women if the sample conforms to the target population.

In order to achieve a better understanding of the variability of the 2SLS estimates as the instrument changes its threshold, and to reduce the conceptual differences between the definitions of full time caregiving between the two groups of platform and traditional workers, Figure 1.3 plots the selected threshold against the estimated effect of working in crowdwork, together with their significance level. It is evident from the figure that, with caregiving becoming a significant predictor of crowdwork at its 36 hours per week threshold, the estimated coefficients also follow a more reliable pattern with little variation in their sign and statistical significance. Most importantly, full and split sample estimates conform to very similar trends, providing evidence that our instrument choice adequately controls for gendered bias in caregiving.

We do not report 2SLS estimates for working hours. The reason is that the condition of caregiving may prevent crowdworkers from working more or from pursuing other sources of income, whereas the desire to work more may be biased by the complications associated with the transition to caregiving. In this case, our interpretations from Table 1.4 should then be understood as not robust to unobserved heterogeneity, and alternative instruments should be considered for this specific analysis.

Table 1.7: 2SLS estimates of the effect of online platform work on net hourly earnings, adjusted for unpaid tasks, in the US and EU

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	US		EU		US+EU		US+EU	
							Caregiving (40h)	
	OLS	OLS	OLS	OLS	OLS	OLS	2SLS	2SLS
VARIABLES	full sample	female only	full sample	female only	full sample	female only	full sample	female only
Working in crowdwork	-1.271***	-1.271***	-1.455***	-1.514***	-1.323***	-1.359***	-1.224***	-1.144***
	(0.051)	(0.060)	(0.069)	(0.043)	(0.046)	(0.053)	(0.255)	(0.250)
Female	-0.182***		-0.071***		-0.205***		-0.224***	
	(0.061)		(0.010)		(0.062)		(0.048)	
Crowdwork × Female	-0.002		-0.084		-0.032			
	(0.068)		(0.052)		(0.059)			
EU × Female					0.140**	-0.248***	0.157***	
					(0.059)	(0.038)	(0.047)	
Age	0.029**	0.057***	0.011***	0.009*	0.014***	0.015***	0.014***	0.016***
	(0.014)	(0.015)	(0.004)	(0.005)	(0.003)	(0.006)	(0.003)	(0.005)
Age squared	-0.000*	-0.001***	-0.000	-0.000	-0.000**	-0.000*	-0.000**	-0.000*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
No. of people in household	-0.028**	-0.048***	0.002	0.001	-0.004	-0.008	-0.004	-0.008
	(0.013)	(0.017)	(0.007)	(0.008)	(0.006)	(0.008)	(0.006)	(0.008)
Married or living with a partner	0.258***	0.270***	0.100***	0.083***	0.117***	0.105***	0.118***	0.102***
	(0.039)	(0.068)	(0.010)	(0.016)	(0.011)	(0.018)	(0.011)	(0.018)
Main earner in household	0.359***	0.272***	0.137***	0.112***	0.155***	0.123***	0.153***	0.118***
	(0.049)	(0.077)	(0.013)	(0.016)	(0.014)	(0.017)	(0.016)	(0.019)
Main source of income	0.084*	0.070	0.095	0.127*	0.119***	0.119*	0.217	0.302
	(0.046)	(0.074)	(0.074)	(0.064)	(0.044)	(0.061)	(0.229)	(0.220)
Observations	3,200	1,696	27,653	14,206	30,853	15,902	30,853	15,902
Adjusted R-squared	0.465	0.476	0.420	0.399	0.428	0.414	0.315	0.287
State controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Education controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F-Test							12.29	22.99
First Stage R^2							0.742	0.722

Notes: State clustered standard errors in parentheses. Control sample restricted to employed and self-employed individuals in working age, excluding freelancers working from home. †: adjusted for time spent in unpaid activities

*p<.05; **p<.01; ***p<.001

Caregiving certainly influences weekly earnings through two distinct channels: first, as more time is allocated to caregiving, the total number of maximum weekly working hours is reduced; secondly, this activity may also generate costs for the caregiver which influence how much he or she will necessitate to earn each week. As argued earlier, our focus on hourly earnings allows us to filter most of these issues out under the assumption that, in a static setting such as in our cross-section, platform workers are unable to individually influence their hourly salary, which is only determined by how efficiently they work. However, the less obvious implication stemming from this reasoning, as also discussed in Section 1.5, is that these caregiving costs may lead to a lowering of the reservation wage, which in turn could also affect participation in online labour markets and raise concerns with regards to violations of the exclusion restriction.

While this mechanism is expected and motivates our identification strategy by providing a theoretical justification for transition into crowdwork for individuals in caregiving, whether ability is linked to the level of prior and posterior reservation wages is, instead, a source of concern. In other terms, if individuals previously outside of the workforce are entering the labour force because of caregiving and are only able to join crowdworking arrangements because of their ability, then self-selection into online labour markets cannot be excluded, and estimates will suffer from bias. At the same time, caregiving may not affect participation in crowdwork for individuals who already have access to other forms of income or who can purchase formal care.

Following from this reasoning, a final test for our instrument is provided in Table 1.6, where hourly earnings are regressed over the instrument and the full set of control covariates across partitions of our sample.[29] The analysis is performed for different specifications including, first, the 15 hours and, then, the 40 hours caregiving threshold, effectively showing reduced form estimates for the instrumental variable model. If heterogenous ability factors which we cannot already control for affect reservation wages and, in turn, participation in crowdwork, then we should see differential effects of caregiving in our reduced form estimates across the two groups of workers: non-caregivers in traditional occupations should earn more, on average, than their counterparts in caregiving, and caregivers who crowdwork should similarly earn less than other crowdworkers not in caregiving.

Our results, however, tell us a different story, indicating that our covariate selection already controls for these effects relatively well. While caregiving, under the 15

[29] This analysis can be seen as an extension of the zero first-stage test for the validity of the exclusion restriction presented in Bound and Jaeger (2000), Altonji et al. (2005), and Angrist et al. (2010).

hours threshold, appears to have a negative and slightly significant effect on earnings in our full sample of female workers, these effects are rendered insignificant when performing the same regressions over the crowdwork and traditional work groups, indicating that the negative sign of that initial coefficient is entirely linked to the first-stage relationship between caregiving and crowdwork. Most importantly, when caregiving is set at its 40 hours, no significant effect on hourly earnings is found in any of the specifications presented in the Table.

Notably, in no case the caregiving coefficient reaches any level of statistical significance once modelling the same regressions on the full sample (men and women). As the income bias described in section 5.2 is accounted for, the exogenous variation left by the caregiving instrument will yield the income-indifferent individual propensity to assist a relative needing for care.

Last but not least, we model hourly earnings again while accounting for time spent in unpaid activities in Table 1.7. As a consequence, hourly earnings – columns (1), (3) and (5) – fall well below our previous estimates, displaying a coefficient of -1.323 (-73.3%), with the prediction moving to -70.6% when instrumenting participation in crowdwork in column (7). Comparable results also apply to the female population (columns 2, 4, 6, and 8), where IV estimates point at a 68% reduction in the hourly rate of pay.

1.7 Robustness checks

1.7.1 Instrumental variable specification

In this section we perform robustness checks for our 2SLS model. The choice of caregiving in the female population as an instrument for participation in crowdwork calls indeed for a number of robustness checks, as it could be argued that the effect of caregiving on participation in crowdwork may change with time, or that caregiving affects the participation in crowdwork but not the duration of crowdwork arrangements. Differences in survey items may then cause issues with identification of caregivers when these individuals have been working on the platform for a long time.

While the EWCS and AWCS surveys inquire how much time does the respondent currently spent in caregiving, the ILO survey records whether the respondent was engaged in full-time caregiving right before starting to work on the platform. The design of the ILO survey then allows us to maintain the causal channel between caregiving and platform work (back when they started working online), while the controls enable us to identify whether comparable individuals in the complier group are still employed in traditional forms of work. This approach, however, imposes that, if caregiving is an exogenous determinant of crowdworking, we should reasonably assume that crowdworkers who entered this form of employment due to caregiving are still engaged in this activity.

To account for these issues, we control in Table 1.8 for time spent in the current occupation, a control that was previously excluded from the final model due to its – obvious – correlation with participation in crowdwork.

In the final models from Tables 1.5 and 1.7, we made the assumption that most crowdworkers have not been engaged in this form of employment for a long time and the ones acting as caregivers when starting platform work are still engaged as such, based on the finding that 75.51% of crowdworkers have not been engaged in this form of employment for more than two years. We now relax this assumption in Table 1.8, where we run the same final IV specification from Table 1.5, adding dummies for years spent in current occupation along with all prior covariates in columns (1) and (5).[30] In the subsequent specifications – columns (2) to (4) and (5) to (8) – we perform a similar analysis by restricting the sample to people who have been working for less than 4 years, 2 years and finally 1 year. By comparing workers that have been working in their current occupation for similar time, the more we reduce the years they have been spending in their current occupation, the more our assumption that these workers are still in caregiving is made reasonable: in this way, we believe to be able to filter out the effects of time spent in a given occupation through the first stage of the 2SLS model. The trade-off is that, the more we reduce our sample size, the more our estimates lose in precision. Nevertheless, the interpretation of our results stays relatively unchanged, with the coefficients retaining their signs and significance. The magnitude of our coefficient for platform work, however, seems somewhat sensible to the sample reduction: in any case, it never overestimates the coefficient of the OLS model, while remaining relatively stable after individuals with more than 5 years of employment have been accounted for. After generalising for split-sample trends, as in equation (3), we can reasonably argue that working in crowdwork generates a negative effect on earnings ranging between 67.2 and 55.02% less than for comparable workers after controlling for time spent in current occupation.

Finally, as mentioned earlier, it could be argued that the inability to distinguish between different forms of caregiving may pose as a source of bias. Indeed, differences between surveys have led to the inability to disentangle caring for children from caring for elderly or disabled relatives. Evidence from studies such as Kremer

[30]The results are reported for both the 15h and 40h caregiving thresholds.

Table 1.8: 2SLS estimates of the effect of online platform work on hourly earnings in the US and EU

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	US+EU				US+EU			
		Caregiving (40h)				Caregiving (40h)		
	2SLS				2SLS			
VARIABLES	full sample	<= 4	<= 2	<= 1	female only	<= 4	<= 2	<= 1
Working in crowdwork	-1.101***	-1.036***	-1.116***	-1.005***	-0.854***	-0.799***	-0.865***	-0.826***
	(0.230)	(0.227)	(0.269)	(0.227)	(0.236)	(0.229)	(0.267)	(0.236)
Female	-0.082***	-0.077***	-0.087***	-0.096***				
	(0.011)	(0.017)	(0.023)	(0.022)				
Age	0.007*	0.015***	0.012**	0.019**	0.007	0.017*	0.013	0.031***
	(0.004)	(0.006)	(0.006)	(0.009)	(0.006)	(0.010)	(0.013)	(0.012)
Age squared	-0.000	-0.000**	-0.000	-0.000**	-0.000	-0.000	-0.000	-0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
No. of people in household	-0.004	-0.022***	-0.025***	-0.018	-0.007	-0.032***	-0.043***	-0.025**
	(0.006)	(0.007)	(0.008)	(0.013)	(0.008)	(0.008)	(0.008)	(0.012)
Main earner in household	0.143***	0.127***	0.110***	0.102***	0.113***	0.084***	0.050	0.040
	(0.015)	(0.016)	(0.019)	(0.031)	(0.018)	(0.026)	(0.033)	(0.039)
Main source of income	0.028	0.059	-0.013	0.066	0.266	0.267	0.180	0.217
	(0.208)	(0.183)	(0.211)	(0.160)	(0.208)	(0.182)	(0.199)	(0.173)
Married or living with a partner	0.100***	0.087***	0.088***	0.097***	0.089***	0.075**	0.075**	0.065**
	(0.010)	(0.014)	(0.017)	(0.025)	(0.018)	(0.032)	(0.036)	(0.032)
Observations	30,673	12,763	8,848	4,104	15,805	6,589	4,570	2,110
Adjusted R-squared	0.265				0.243			
State controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Education controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Years in occupation controls	Yes	No	No	No	Yes	No	No	No
F-Test	13.58	20.72	20.93	42.21	24.39	36.14	37.74	61.96
First Stage R^2	0.742	0.747	0.752	0.759	0.722	0.735	0.744	0.768

Notes: State clustered standard errors in parentheses. Dependent variable: natural logarithm of hourly nominal earnings (US dollars). Control sample restricted to employedand self-employed American individuals in working age excluding freelancers working from home. From columns (2) to (4) and (7) to (9), the sample is restricted to individuals who have been working in their current occupation for less than 4 years, 2 years and finally 1 year.

*p<.05; **p<.01; ***p<.001

Table 1.9: 2SLS estimates of the effect of online platform work on hourly earnings in the US and EU

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	US+EU				US+EU			
	Caregiving (Elderly only, 40h)				Caregiving (Elderly only, 40h)			
	2SLS				2SLS			
VARIABLES	full sample	<= 4	<= 2	<= 1	female only	<= 4	<= 2	<= 1
Working in crowdwork	-1.772***	-1.299**	-1.393*	-1.192**	-1.354***	-0.994**	-1.178	-1.114
	(0.567)	(0.506)	(0.844)	(0.496)	(0.483)	(0.453)	(0.793)	(0.939)
Female	-0.075***	-0.073***	-0.081***	-0.093***				
	(0.012)	(0.018)	(0.026)	(0.021)				
Age	0.006	0.014**	0.012**	0.019*	0.007	0.017*	0.013	0.031***
	(0.004)	(0.006)	(0.006)	(0.010)	(0.006)	(0.010)	(0.014)	(0.012)
Age squared	-0.000	-0.000**	-0.000	-0.000*	-0.000	-0.000	-0.000	-0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
No. of people in household	-0.003	-0.021***	-0.024***	-0.017	-0.007	-0.031***	-0.041***	-0.023*
	(0.006)	(0.007)	(0.009)	(0.013)	(0.008)	(0.008)	(0.010)	(0.013)
Main earner in household	0.160***	0.140***	0.126***	0.115***	0.126***	0.094***	0.071	0.060
	(0.020)	(0.027)	(0.046)	(0.044)	(0.021)	(0.033)	(0.060)	(0.072)
Main source of income	-0.547	-0.157	-0.241	-0.080	-0.151	0.112	-0.068	0.002
	(0.503)	(0.438)	(0.701)	(0.392)	(0.431)	(0.390)	(0.634)	(0.709)
Married or living with a partner	0.100***	0.087***	0.090***	0.099***	0.096***	0.080**	0.087*	0.078
	(0.011)	(0.014)	(0.019)	(0.026)	(0.019)	(0.035)	(0.050)	(0.056)
Observations	30,673	12,763	8,848	4,104	15,805	6,589	4,570	2,110
Adjusted R-squared	0.238				0.238			
State controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Education controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Years in occupation controls	Yes	No	No	No	Yes	No	No	No
F-Test	12.68	10.62	6.485	14.79	24.68	22.05	13.35	14.34
First Stage R^2	0.740	0.742	0.745	0.752	0.716	0.725	0.729	0.751

Notes: State clustered standard errors in parentheses. Dependent variable: natural logarithm of hourly nominal earnings (US dollars). Control sample restricted to employed and self-employed American individuals in working age excluding freelancers working from home. From columns (2) to (4) and (7) to (9), the sample is restricted to individuals who have been working in their current occupation for less than 4 years, 2 years and finally 1 year.

*p<.05; **p<.01; ***p<.001

and Chen (2002) suggests that fertility may be influenced by a number of social drivers. While we believe that our controls are able to filter these influences out,[31] we here intend to relax this assumption and treat fertility as endogenous. Even if, as discussed, conflicting survey designs prevent us from fully separating individuals caring for children from the ones caring for disabled or elderly relatives, we can nonetheless identify individuals in caregiving who, at the same time, do not have kids – and, therefore, are most surely not caring for children. We then switch our instrument with the new one ("Caring for elderly or disabled relatives only" and present our results in Table 1.9, adopting the same approach used for the robustness checks in Table 1.8. The reductions in the "complier" group for crowdworkers leave to an increase in the variability of our estimates which appear particularly sensible to the reduction in sample size. Since this time we are only able to compare individuals with no children, some kind of bias can still be expected: in fact, while our estimates maintain their sign and do not diverge too much from our results in Table 1.8, they surely suffer from some level of overestimation. In any case, these results do not contradict our previous findings.

1.7.2 Model specification

All robustness checks we previously presented rely on the correct specification of the IV estimator. In this section, instead, we address the concerns related to this approach by relying on an alternative specification for the estimation of the effects of crowdwork on earnings.

An interesting result from our IV estimates is that first-stage regressions produce comfortably high R-squared statistics, meaning that our set of observables adequately predicts assignment into platform work. If we have a correct specification for the probability to work in online labour platforms, then a binomial model can be used to compute propensity scores, which can be used to re-weight observations across the two groups of workers. Re-weighting can be achieved through inverse probability weighting (first proposed by Rosenbaum, 1987; see Austin, 2011, for a methodological review of uses of propensity scores in quasi-experimental settings), where new weights are produced by assigning each observation the inverse of the conditional probability of its treatment status.

This means that, in our case, individuals in crowdwork will receive a weight equal to $1/p_i$, while traditional workers will be weighted $1/(1-p_i)$, where p_i indicates the propensity score; in other words, it indicates each individual probability $P(T=1|X)$

[31] In particular, we believe that controls for education, marital status and household size can adequately capture these endogenous variations.

Table 1.10: Effect of online platform work on earnings in the US and EU

	(1) Coeff.	(2) Std.Err.	(3) z-score	(4) n
	Earnings (natural log)			
ATE				
Working in crowdwork	-0.877	0.294	-2.985	43,643
ATT				
Working in crowdwork	-0.928	0.316	-2.936	2,380
	Earnings†(natural log)			
ATE				
Working in crowdwork	-1.191	0.306	-3.886	43,643
ATT				
Working in crowdwork	-1.192	0.328	-3.629	2,380

Notes: IPWRA estimator of the average treatment effect (ATE) and average treatment effect on the treated (ATT) of online platform work on earnings. †: adjusted for time spent in unpaid activities. Control sample restricted to employed and self-employed individuals in working age, excluding freelancers working from home.

to be working in crowdsourcing, given a set of covariates X. The differences of inverse propensity scores weighted averages will yield the effect of platform work, under the caveat that the underlying propensity score model is correct.

To overcome this issue, an inverse-probability-weighted regression adjustment (IPWRA), first covered by Robins et al. (1994) and further developed in Wooldridge (2007) is proposed, where inverse probability weighting is combined with regression adjustments in order to produce a doubly robust estimator. In IPWRAs, regression models are fit on inverse probability weighted observations according to their treatment status (meaning that the model is fit on two separate treatment and control samples), and the parameters from these models are used to predict counter-factual outcomes on an individual level, for all observations. The difference in means between treatment and control predicted outcomes will then yield the ATE.

The IPWRA estimator ensures that, as long as one of the two models, one for predicting assignment, and the other one for modelling outcome, is correct, then the results will not suffer from bias. We then use a binomial logistic model to calculate propensity scores,[32] and then assign the new weights to each observation so that a log-normal model for earnings can be fit across the two groups of workers, using the same covariate specification from Table 1.2, column 7 (omitting, for obvious reasons,

[32] Covariates list: Female, Female*EU, Age, Age squared, No. of people in household, Main earner in household, Married or living with a partner, Health condition and Education controls.

the “Working in crowdwork” dummy). The results from our test can be found in Table 1.10.

Our estimates show that, after controlling for these different models, working in crowdwork still produces a statistically significant -69.62% reduction on earnings (adjusted for unpaid tasks). Extending our double robust approach to the estimation of the ATT, we find the effect on the treated to be close to -70% as well (with comparable statistical significance). These results are remarkably similar to the ones obtained by our previous instrumental variable approach (and OLS, by extension), and reinforce our finding that working conditions in crowdwork are generally unaffected by the characteristics of individuals working in these arrangements.

1.8 Conclusions

In this paper we have provided an empirical analysis of the effect of crowdwork on working conditions in both the United States and Europe. We assemble data from different sources, harmonising responses from an online survey on crowdworkers with observations from two general surveys on workers' conditions in the US and EU, and then comparing outcomes across forms of work. To the best of our knowledge, this is one of the first attempts to provide an unbiased comparison of platform and traditional workers in terms of earnings and working conditions.

In our contribution, we focused on the effects of individual ability on earnings in the platform economy, finding that most of the differences between platform workers and traditional workers are unexplained by individual characteristics. As we show that the effect of crowdsourcing on earnings is even larger as it could be expected from simple differences in means, our estimates cast a dark shadow over platform work: crowdsourcers earn 70.6% to 68.1% less than comparable workers in terms of ability, while spending nearly as much time working in the platform as their counterparts do in traditional occupations. Most importantly, labour force in crowdworking arrangements appears to be highly under-utilised, with all crowdworkers being more likely to be left wanting for more work than comparable individuals. All these findings, along with the fact that these individuals do not appear to be looking for other jobs more than traditional workers, suggest crowdworkers to belong to a new category of idle workers whose human capital is not being fully utilised nor adequately compensated.

It should be noted that while these results hold for US and EU platform workers, the external validity of our estimates is threatened by the nature of crowdwork platforms themselves and, while our conclusions may be extended to routine-task intensive platforms such as Crowdflower or Clickworker, our analysis may not hold in other contexts where more diversified tasks, requiring specific skills and creative input from service providers, are offered, such as in the case of 'macro-task' freelance marketplaces like UpWork.

The observed disparities should then be attributed to factors other than individual ability. We were able to rule out the possibility that most of these differences are caused by the routine and abstract content of online platform jobs, as workers with comparable routine and abstract tasks still retain most of their salary premium, indicating that the relative simplicity and repetitiveness of these tasks does not necessarily lead to a sizeable decrease in earnings. This leads us to believe that this effect could be better explained by the following factors:

1. competition from equally skilled but cheaper labour from other countries within the same platform;
2. scarcity and heterogeneity in demand for these kind of activities;
3. lack of labour rights and minimum standards stemming from the status of independent contractors.

In the first case, the earnings effect of platform work can be attributed to excess supply: indeed, the influx of "digital immigrants" may lead to an increase in labour supply and infra-task competition, lowering remunerations due to the low complementarity of these workers. Indeed, Borchert et al. (2018) have found that unemployment shocks, leading to increased participation in online markets, can have a positive effect on wage elasticities in crowdwork.

In the second case, it could be argued either that firms and clients are mostly uninformed about the possibility of outsourcing through online platforms, or that the sample of clients which employs online labour is intrinsically different in its nature from other firms employing traditional workers, generating scarcity in demand. While panel data are necessary to study these effects, the lack of particular differences between crowdworkers in 2015 and 2017 – in the ILO quasi-panel – indicates that, so far, the demand for these services has seen little growth. Also, while Katz and Krueger (2018) estimate a general rise in participation in the platform economy between 2005 and 2015 (from 10.7 to 15.8% in the US), evidence from Farrell and Greig (2017) could support the claim that these markets have, overall, reached their peak in 2016. Still, persistence of slack and factor under-utilisation in these markets is indicative of the presence of a mismatch between supply and demand which, if not found to change over the next years, could be described as a structural condition of crowdsourcing as a consequence of the nature of its clients.

In the third and final case, the monopsonistic nature of platforms, linked with the general lack of labour standards, enables the imposition of a heavy markup over online workers, allowing clients to operate at prices well below the market's marginal costs. These considerations are consistent with the results of Dube et al. (2018). As our results refer to year 2015, the influence of these factors could change in the future, in parallel with the evolution of the platform economy. In any case, we believe that the poor working conditions crowdsourcers have to live with are the result of an interplay between these elements, and it is up to future research to test each of these hypotheses individually, disentangling the effect of each of these factors from the others.

Acknowledgements

The author would like to thank Nicolò Fraccaroli, Anzelika Zaiceva, Matteo Villa and Roberto Volpe for their invaluable advice and support. The author also wishes to thank the participants of the BISA Postgraduate Research Workshop on International Migration Politics, and Teodora Tsankova in particular, for their useful comments and suggestions. All remaining errors are mine.

Bibliography

Acemoglu, D. and Autor, D. (2011). Skills, tasks and technologies: Implications for employment and earnings. In *Handbook of Labor Economics*, pages 1043–1171. Elsevier.

Altonji, J. G. and Blank, R. M. (1999). Race and gender in the labor market. In *Handbook of Labor Economics*, pages 3143–3259. Elsevier.

Altonji, J. G., Elder, T. E., and Taber, C. R. (2005). An evaluation of instrumental variable strategies for estimating the effects of catholic schooling. *The Journal of Human Resources*, 40(4):791–821.

Angrist, J., Lavy, V., and Schlosser, A. (2010). Multiple experiments for the causal link between the quantity and quality of children. *Journal of Labor Economics*, 28(4):773–824.

Angrist, J. D. and Krueger, A. B. (1991). Does compulsory school attendance affect schooling and earnings? *The Quarterly Journal of Economics*, 106(4):979–1014.

Angrist, J. D. and Krueger, A. B. (1995). Split-sample instrumental variables estimates of the return to schooling. *Journal of Business & Economic Statistics*, 13(2):225.

Austin, P. C. (2011). An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research*, 46(3):399–424.

Autor, D. H. and Dorn, D. (2013). The growth of low-skill service jobs and the polarization of the US labor market. *American Economic Review*, 103(5):1553–1597.

Berg, J. (2015). Income security in the on-demand economy: Findings and policy lessons from a survey of crowdworkers. *Comparative Labor Law and Policy Journal*, 37:543.

Berg, J., Furrer, M., Harmon, E., Rani, U., and Silberman, M. S. (2018). Digital labour platforms and the future of work: Towards decent work in the online world. International Labour Office. ISBN: 978-92-2-031025-0.

Berger, T., Chen, C., and Frey, C. B. (2018). Drivers of disruption? Estimating the Uber effect. *European Economic Review*, 110:197–210.

Berinsky, A. J., Huber, G. A., and Lenz, G. S. (2012). Evaluating online labor markets for experimental research: Amazon.com's Mechanical Turk. *Political Analysis*, 20(03):351–368.

Bick, A. (2016). The quantitative role of child care for female labor force participation and fertility. *Journal of the European Economic Association*, 14(3):639–668.

Blau, F. D. and Kahn, L. M. (2003). Understanding international differences in the gender pay gap. *Journal of Labor Economics*, 21(1):106–144.

Borchert, K., Hirth, M., Kummer, M., Laitenberger, U., Slivko, O., and Viete, S. (2018). Unemployment and online labor. *ZEW Discussion Paper No. 18-023.*

Bound, J. and Jaeger, D. A. (2000). *Do compulsory school attendance laws alone explain the association between quarter of birth and earnings?*, volume 19, pages 83–108. Emerald Group Publishing Limited.

Bound, J., Jaeger, D. A., and Baker, R. M. (1995). Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. *Journal of the American Statistical Association*, 90(430):443–450.

Card, D. and Krueger, A. B. (1992). Does school quality matter? Returns to education and the characteristics of public schools in the united states. *Journal of Political Economy*, 100(1):1–40.

Ciani, E. (2012). Informal adult care and caregivers' employment in europe. *Labour Economics*, 19(2):155–164.

Degryse, C. (2016). Digitalisation of the economy and its impact on labour markets. *ETUI Research Paper - Working Paper 2016.02.*

Difallah, D., Filatova, E., and Ipeirotis, P. (2018). Demographics and dynamics of Mechanical Turk workers. In *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining*. ACM Press.

Dube, A., Jacobs, J., Naidu, S., and Suri, S. (2018). Monopsony in online labor markets. NBER Working Paper Series 24416, National Bureau of Economic Research.

Earle, A. and Heymann, J. (2012). The cost of caregiving: Wage loss among caregivers of elderly and disabled adults and children with special needs. *Community, Work & Family*, 15(3):357–375.

Ermisch, J. F. and Wright, R. E. (1993). Wage offers and full-time and part-time employment by British women. *The Journal of Human Resources*, 28(1):111–133.

Eurofound (2015). New forms of employment. *Publication Office of the European Union, Luxembourg.*

European Foundation For The Improvement Of Living And Working Conditions (2017). European working conditions survey, 2015. UK Data Service. doi: 10.5255/UKDA-SN-8098-4.

Farrell, D. and Greig, F. E. (2017). The online platform economy: Has growth peaked? *Jp Morgan Chase & Co. Institute.*

Gerstel, N. and Gallagher, S. K. (2001). Men's caregiving. *Gender & Society*, 15(2):197–217.

Hara, K., Adams, A., Milland, K., Savage, S., Callison-Burch, C., and Bigham, J. P. (2018). A data-driven analysis of workers' earnings on Amazon Mechanical Turk.

In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, pages 449:1–449:14, New York, NY, USA. ACM.

Harris, S. D. and Krueger, A. B. (2015). A proposal for modernizing labor laws for twenty-first-century work: The "independent worker". The Hamilton Project Discussion Papers 2015 2.

Horton, J., Kerr, W. R., and Stanton, C. (2017). Digital labor markets and global talent flows. NBER Working Paper Series 23398, National Bureau of Economic Research.

Horton, J. J. and Chilton, L. B. (2010). The labor economics of paid crowdsourcing. In *Proceedings of the 11th ACM conference on Electronic Commerce*. ACM Press.

Horton, J. J., Rand, D. G., and Zeckhauser, R. J. (2011). The online laboratory: conducting experiments in a real labor market. *Experimental Economics*, 14(3):399–425.

Hotchkiss, J. L. (1991). The definition of part-time employment: A switching regression model with unknown sample selection. *International Economic Review*, 32(4):899–917.

Huws, U., Spencer, N. H., Syrdal, D. S., and Holts, K. (2017). *Work in the Gig Economy - Employment in the Era of Online Platforms*. FEPS Studies, Foundation for European Progressive Studies, Brussels.

Inoue, A. and Solon, G. (2010). Two-sample instrumental variables estimators. *Review of Economics and Statistics*, 92(3):557–561.

Katz, L. F. and Krueger, A. B. (2018). The rise and nature of alternative work arrangements in the United States, 1995–2015. *ILR Review*, 72(2):382–416.

Kremer, M. and Chen, D. L. (2002). Income distribution dynamics with endogenous fertility. *Journal of Economic Growth*, 7(3):227–258.

LaLonde, R. J. (1986). Evaluating the econometric evaluations of training programs with experimental data. *American Economic Review*, 76(4):604–620.

Leigh, A. (2010). Informal care and labor market participation. *Labour Economics*, 17(1):140–149.

Maestas, N., J. Mullen, K., Powell, D., von Wachter, T., and B. Wenger, J. (2017). The American Working Conditions Survey data: Codebook and data description. Santa Monica, CA: RAND Corporation. doi: 10.7249/TL269.

OECD (2016). New forms of work in the digital economy. *OECD Digital Economy Papers*, (260).

Ottaviano, G. I., Peri, G., and Wright, G. (2013). Immigration, offshoring and american jobs. *American Economic Review*, 103(5):1925–59.

Paolacci, G., Chandler, J., and Ipeirotis, P. G. (2010). Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making*, 5:411–419.

Pesole, A., Urzí Brancati, M., Fernández-Macías, E., Biagi, F., and González Vázquez, I. (2018). Platform workers in Europe. Evidence from the COLLEEM survey. Technical report, Joint Research Centre, European Commission.

Prassl, J. and Risak, M. (2015). Uber, Taskrabbit, and co.: Platforms as employers-rethinking the legal analysis of crowdwork. *Comparative Labor Law and Policy Journal*, 37:619.

Robins, J. M., Rotnitzky, A., and Zhao, L. P. (1994). Estimation of regression coefficients when some regressors are not always observed. *Journal of the American Statistical Association*, 89(427):846–866.

Rosenbaum, P. R. (1987). Model-based direct adjustment. *Journal of the American Statistical Association*, 82(398):387–394.

Schmitz, H. and Westphal, M. (2017). Informal care and long-term labor market outcomes. *Journal of Health Economics*, 56:1–18.

Wakabayashi, C. and Donato, K. M. (2005). The consequences of caregiving: Effects on women's employment and earnings. *Population Research and Policy Review*, 24(5):467–488.

Wooldridge, J. (2007). Inverse probability weighted estimation for general missing data problems. journal of econometrics, 141, 1281-1301. *Journal of Econometrics*, 141:1281–1301.

Appendix

1.A Summary statistics

Table A.1: Descriptive statistics on US workers employed in traditional occupations, AWCS 2015

	count	mean	sd	min	p5	p50	p95	max
Hourly nominal earnings (USD)	1847	30.77	207.9	0	2.301	17.58	58.81	10547.9
Weekly working hours	1910	39.06	11.65	0	20	40	60	112
Age	1941	41.02	12.61	18	21	41	61	64
Female	1941	0.463	0.499	0	0	0	1	1
Married or living with a partner	1941	0.516	0.500	0	0	1	1	1
No. of people in household	1941	3.063	1.672	1	1	3	6	12
Main earner in household	1891	0.603	0.489	0	0	1	1	1
Educ.: no high school diploma	1941	0.0638	0.244	0	0	0	1	1
Educ.: high school diploma	1941	0.502	0.500	0	0	1	1	1
Educ.: technical/associate	1941	0.0966	0.296	0	0	0	1	1
Educ.: bachelor's degree	1941	0.208	0.406	0	0	0	1	1
Educ.: master's degree	1941	0.0944	0.292	0	0	0	1	1
Educ.: higher	1941	0.0356	0.185	0	0	0	0	1
Health: Very Good	1891	0.132	0.338	0	0	0	1	1
Health: Good	1891	0.407	0.491	0	0	0	1	1
Health: Fair	1891	0.345	0.475	0	0	0	1	1
Health: Poor	1891	0.0991	0.299	0	0	0	1	1
Health: Very Poor	1891	0.0176	0.132	0	0	0	0	1
Caregiving (15h/week)	1941	0.149	0.356	0	0	0	1	1
Caregiving (40h/week)	1941	0.0824	0.275	0	0	0	1	1

Notes: Weighted summary statistics for workers in traditional occupations from the US (AWCS 2015). Sample restricted to employed and self-employed individuals in working age, excluding freelancers working from home.

*p<.05; **p<.01; ***p<.001

Table A.2: Descriptive statistics on EU workers employed in traditional occupations, EWCS 2015

	count	mean	sd	min	p5	p50	p95	max
Hourly nominal earnings (USD)	26991	17.06	91.89	0.00319	3.935	11.83	29.77	5687.8
Weekly working hours	31650	37.18	11.90	1	15	40	55	126
Age	32429	42.21	11.39	15	23	43	60	64
Female	32429	0.478	0.500	0	0	0	1	1
Married or living with a partner	32429	0.697	0.459	0	0	1	1	1
No. of people in household	32312	2.882	1.268	1	1	3	5	10
Main earner in household	32429	0.595	0.491	0	0	1	1	1
Educ.: no high school diploma	32316	0.161	0.367	0	0	0	1	1
Educ.: high school diploma	32316	0.448	0.497	0	0	0	1	1
Educ.: technical/associate	32316	0.147	0.354	0	0	0	1	1
Educ.: bachelor's degree	32316	0.127	0.333	0	0	0	1	1
Educ.: master's degree	32316	0.108	0.311	0	0	0	1	1
Educ.: higher	32316	0.00856	0.0921	0	0	0	0	1
Health: Very Good	32400	0.261	0.439	0	0	0	1	1
Health: Good	32400	0.532	0.499	0	0	1	1	1
Health: Fair	32400	0.185	0.389	0	0	0	1	1
Health: Poor	32400	0.0201	0.140	0	0	0	0	1
Health: Very Poor	32400	0.00228	0.0477	0	0	0	0	1
Caregiving (15h/week)	32429	0.170	0.375	0	0	0	1	1
Caregiving (40h/week)	32429	0.0197	0.139	0	0	0	0	1

*N*otes: Weighted summary statistics for workers in traditional occupations from the EU (EWCS 2015), EU member states only. Control sample restricted to employed and self-employed individuals in working age, excluding freelancers working from home. Earnings are adjusted for purchasing power parity.

*p<.05; **p<.01; ***p<.001

Table A.3: Descriptive statistics, US and EU crowdworkers, ILO (2015, 2017)

	count	mean	sd	min	p5	p50	p95	max
Hourly nominal earnings (USD)	2341	7.166	18.72	0.0489	0.568	4.888	17.39	568.4
Hourly nominal earnings (USD)†	2302	4.697	11.72	0	0.300	3.125	12	357.1
Weekly working hours	2369	19.36	23.69	0	2	13	50	168
Weekly working hours†	2320	26.03	30.56	0	2	18	70	336
Age	2393	35.03	10.93	18	21	33	57	83
Female	2393	0.448	0.497	0	0	0	1	1
Married or living with a partner	2393	0.455	0.498	0	0	0	1	1
No. of people in household	2393	2.768	1.377	1	1	3	5	10
Main earner in household	2393	0.806	0.396	0	0	1	1	1
Educ.: no high school diploma	2391	0.0247	0.155	0	0	0	0	1
Educ.: high school diploma	2391	0.356	0.479	0	0	0	1	1
Educ.: technical/associate	2391	0.132	0.339	0	0	0	1	1
Educ.: bachelor's degree	2391	0.334	0.472	0	0	0	1	1
Educ.: master's degree	2391	0.125	0.330	0	0	0	1	1
Educ.: higher	2391	0.0284	0.166	0	0	0	0	1
Health: Very Good	2392	0.258	0.437	0	0	0	1	1
Health: Good	2392	0.528	0.499	0	0	1	1	1
Health: Fair	2392	0.174	0.379	0	0	0	1	1
Health: Poor	2392	0.0347	0.183	0	0	0	0	1
Health: Very Poor	2392	0.00585	0.0763	0	0	0	0	1
Caregiving (15h/week)	2393	0.166	0.372	0	0	0	1	1
Caregiving (40h/week)	2393	0.166	0.372	0	0	0	1	1

Notes: Summary statistics for crowdworkers from the US and EU (ILO), pooled 2015 and 2017 survey waves. Earnings are deflated to the 2015 reference period (local currency) and then adjusted for purchasing power parity. †: adjusted for time spent in unpaid activities.

*p<.05; **p<.01; ***p<.001

1.B Returns to observable skills and first stage IV regressions

Table B.1: Returns to education in crowdwork in US and EU

VARIABLES	(1) US OLS	(2) EU OLS	(3) US+EU OLS
Working in crowdwork	-0.558	-0.837***	-0.806***
	(0.486)	(0.135)	(0.135)
Crowdwork × High school diploma	-0.260	-0.062	-0.056
	(0.472)	(0.113)	(0.115)
Crowdwork × Technical/associate degree	-0.370	-0.455***	-0.217*
	(0.474)	(0.107)	(0.117)
Crowdwork × Bachelor's degree	-0.674	-0.425***	-0.355***
	(0.473)	(0.139)	(0.120)
Crowdwork × Master's degree	-1.102**	-0.402***	-0.510***
	(0.490)	(0.116)	(0.133)
Crowdwork × Higher	-1.202**	-0.852***	-0.859***
	(0.534)	(0.165)	(0.162)
High school diploma	0.313	0.111***	0.109***
	(0.226)	(0.025)	(0.028)
Technical/associate degree	0.493**	0.264***	0.267***
	(0.207)	(0.034)	(0.035)
Bachelor's degree	0.758***	0.399***	0.413***
	(0.221)	(0.027)	(0.031)
Master's degree	0.998***	0.505***	0.520***
	(0.225)	(0.023)	(0.026)
Higher	1.111***	0.731***	0.754***
	(0.229)	(0.067)	(0.063)
Observations	3,217	27,676	30,893
Adjusted R-squared	0.408	0.380	0.382
Control covariates	Yes	Yes	Yes

*N*otes: State clustered standard errors in parentheses. Control sample restricted to employed and self-employed individuals in working age, excluding freelancers working from home.

*p<.05; **p<.01; ***p<.001

Table B.2: 2SLS First and Second Stage coefficients of the effect of online platform work on earnings in the US and EU (Table 1.5)

	(1) US+EU	(2) US+EU	(3) US+EU	(4) US+EU	(5) US+EU	(6) US+EU	(7) US+EU	(8) US+EU
	Caregiving (15h)				Caregiving (40h)			
VARIABLES	First Stage full sample	Second Stage full sample	First Stage female only	Second Stage female only	First Stage full sample	Second Stage full sample	First Stage female only	Second Stage female only
Female	0.042***	-0.284***			0.027**	-0.213***		
	(0.012)	(0.059)			(0.012)	(0.049)		
EU × Female	-0.040***	0.207***			-0.024*	0.146***		
	(0.012)	(0.053)			(0.013)	(0.048)		
Age	-0.004***	0.020***	-0.002*	0.018***	-0.004***	0.014***	-0.003**	0.016***
	(0.001)	(0.003)	(0.001)	(0.006)	(0.001)	(0.004)	(0.001)	(0.005)
Age squared	0.000***	-0.000***	0.000	-0.000**	0.000***	-0.000**	0.000	-0.000*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
No. of people in household	0.001	-0.006	-0.001	-0.010	0.000	-0.004	-0.002	-0.009
	(0.001)	(0.005)	(0.002)	(0.008)	(0.001)	(0.006)	(0.002)	(0.008)
Main earner in household	0.021***	0.122***	0.024***	0.075**	0.022***	0.154***	0.026***	0.120***
	(0.006)	(0.024)	(0.006)	(0.031)	(0.006)	(0.016)	(0.006)	(0.019)
Main source of income	-0.873***	1.503*	-0.849***	1.818*	-0.866***	0.171	-0.835***	0.271
	(0.029)	(0.894)	(0.038)	(0.975)	(0.030)	(0.227)	(0.039)	(0.211)
Married or living with a partner	-0.004*	0.121***	0.011*	0.081***	-0.005**	0.115***	0.011*	0.102***
	(0.002)	(0.011)	(0.006)	(0.020)	(0.002)	(0.011)	(0.006)	(0.018)
Caregiving	-0.006		0.017**		0.044*		0.123***	
	(0.004)		(0.008)		(0.023)		(0.026)	
Caregiving × Female	0.024***				0.077**			
	(0.008)				0.077**			
Working in crowdwork		0.518		0.902		-1.007***		-0.918***
		(1.060)		(1.158)		(0.247)		(0.236)
Observations	30,893	30,893	15,921	15,921	30,893	30,893	15,921	15,921
Adjusted R-squared	0.737	0.151	0.710	0.051	0.742	0.255	0.720	0.231
State controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Education controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F-Test		18.33		23.35		47.20		90.36
First Stage R^2		0.738		0.712		0.742		0.722

*N*otes: State clustered standard errors in parentheses. Control sample restricted to employed and self-employed individuals in working age, excluding freelancers working from home.

*p<.05; **p<.01; ***p<.001

Chapter 2

Does Fake News Affect Voting Behaviour?

Abstract

Over the last decade, the erosion of trust in public institutions and traditional media sources have been proceeding in parallel. Recent developments in media consumption have led to a proliferation of politically charged online misinformation. In this paper we investigate whether the spread of fake news has affected the results of recent elections, contributing to the growth of populist platforms. We aim to quantify the causal effect of the spread of misinformation over electoral outcomes in the 2018 Italian general elections. The presence of Italian and German linguistic groups in the Trento and Bolzano/Bozen autonomous provinces offers a unique source of exogenous variation, as it assigns individuals into distinct filter bubbles each differently exposed to misinformation. We introduce a novel index based on text mining techniques to measure populism. We construct a novel database with social media content of each party and their leaders over the course of the electoral campaign for the 2013 and 2018 elections. Our results indicate that misinformation had a negligible effect on populist vote in Trentino and South Tyrol during the Italian 2018 general elections.

Keywords: *Fake News, Political Economy, Electoral Outcomes, Populism*
JEL codes: C26, D72, P16.

2.1 Introduction

Over the last years, fake news has proliferated online, in parallel with the expansion of online communities. As recent elections across the globe have been characterised by a swing towards populism – from the 2016 US presidential elections to the 2018 Italian parliamentary elections – it is no surprise that the spread of fake news has been held responsible for affecting electoral results.[1] The question then arises as to whether exposure to fake news actually affects voting behaviour or if it only reinforces predetermined political beliefs. In this paper, we attempt to provide an answer to this question by collecting data on voting preferences and fake news exposure from the Italian elections of 2013 and 2018 and analysing it in a quasi-experimental setting.

We intend to follow the path laid out by Allcott and Gentzkow (2017), and study fake news as intentionally fabricated information characterised by its politically charged content. Fake news in Italy has enjoyed quite a large exposure in anticipation of the 2018 elections. A recent report from the newspaper Il Sole 24 Ore estimated that Five Star Movement (M5S) and Lega Nord (LN or Lega) voters believe (52% and 49%, respectively) and share (22% and 11%) fake news in much higher proportions than individuals who did not vote for a 'populist' party.[2] As empirical evidence points at an increasing internet and social network usage in Italy growing across all ages and educational cohorts, with middle aged and less educated individuals experiencing the greatest increase (Istat-Fub, 2018), it is fair to argue that many more Italians had access to false information in 2018 than in 2013, when the previous general elections took place.

Whether the diffusion of fake news has affected electoral outcomes is then a legitimate question that deserves to be answered. In this regard, Italy – and the Trentino-Alto Adige/Südtirol region, specifically – presents the ideal research setting for the study of the phenomenon. First of all, the multiparty nature of the Italian political system allows us to analyse electoral outcomes with much more granularity than in countries with majoritarian systems. Moreover, the last two elections in 2013 and 2018 have been characterised by strong electoral performance for political parties whose platform has been often described as 'populist': the M5S, which topped the

[1] See Parkinson, H. J. (2016); Click and elect: how fake news helped Donald Trump win a real election; The Guardian; available at: https://www.theguardian.com/commentisfree/2016/nov/14/fake-news-donald-trump-election-alt-right-social-media-tech-companies; last accessed: 15 February 2019

[2] Il Sole 24 Ore (2018); Fake news: quando le bugie hanno le gambe lunghe; il-sole24ore.com; Available at: http://www.infodata.ilsole24ore.com/2018/05/04/fake-news-le-bugie-le-gambe-lunghe/; last accessed: 15 February 2019

poll in each of the two general elections it contested, and the more well-established Lega, whose gradual transformation from a Northern autonomist party to a country-wide 'national-populist' outfit was met by unprecedented success in 2018. Also, as discussed in Campante et al. (2017), part of the growth of these forces can be attributed to a surge in participation of previously excluded voters, fostered in turn by increased access to broadband internet connections.

Several areas of Italy are home to significant linguistic minorities. In Trentino-Alto Adige/ Südtirol, in particular, both German and Italian-speaking communities are represented in sizeable numbers. Remarkably, studies show that only a small portion of the local population is functionally bilingual (Abel et al., 2012; Ebner, 2016), which suggests a certain degree of separation between linguistic communities. We intend to exploit the language differences across the two *province autonome* that make up the region. Alto Adige/Südtirol, in English 'South Tyrol', located on the border with Austria, has a majority of German-speaking population, and overwhelmingly so in rural areas. Conversely, Trentino is dominated by the Italian-speaking population. We will hence exploit this difference as an exogenous source of variation in exposure to fake news.

Controlling for the electoral trends between the Italian and German-speaking population, and considering how fake news has been known to spread through channels filtered by an individual's 'echo chambers' (Allcott and Gentzkow, 2017, Boutyline and Willer, 2016), we then advance the hypothesis that German-speaking citizens have been exposed to fake news concerning the Italian elections in a lower magnitude when compared to their Italian-speaking counterparts.

We assume – and this assumption will be tested empirically – that the German-speaking population in Trentino-Alto Adige/Südtirol, while comparable to its Italian-speaking counterpart in terms of economic and demographic conditions, is exposed to a peculiar filter bubble where exposure to fake news concerning Italian politics is limited. Indeed, in line with the approach of Allcott and Gentzkow (2017), fake news sources may be assumed to hold economic or agenda-driven incentives to spread false information. We believe that these incentives are not met in this case: from such a small population as the German speaking population in South Tyrol, website accesses may not generate enough advertising revenue, and the impact on national elections may be considered negligible as well. In this way, after controlling for electoral trends specific to each language group, we believe we are able to assess the impact of fabricated news over electoral outcomes. This is, to the best of our knowledge, the first attempt to study this phenomenon using a quasi-experimental methodology.

The remainder of the paper is organised as follows. Section 2 provides an overview of the literature on misinformation, filter bubbles and electoral outcomes. Section 3 covers background information on the demographic features and political traditions of the Trentino-Alto Adige/Südtirol region, along with contextualising the spread of fake news within the 2018 Italian general elections. Section 4 describes our data sources, and Section 5 develops a text mining methodology for measuring populism. Our econometric model and results are contained in Sections 6 and 7, respectively. Section 8 concludes and proposes a simple theoretical model assisting us in understanding and discussing our results.

2.2 Literature review

Due to the nature of the phenomenon and the growing interest from both academics and policy-makers, the field of empirical research on online misinformation has witnessed a considerable growth. Yet a number of challenges linked with developing a successful research design leave many aspects of this phenomenon unexplored.

Fake news and the ability to correctly recognise its mendacity in correlation with prior political beliefs have been studied in a seminal paper by Allcott and Gentzkow (2017). The authors investigate the effects of exposure to fake news – assessing recall rates through the use of placebo headlines – and the nature of the sources of false information, developing a database of fake news and a post-election survey in the process. Most importantly, the authors identify a correlation between exposure to misinformation and decision to support Donald Trump in the 2016 US presidential elections.

This connection was corroborated by further evidence from Bovet and Makse (2019). They found that 25% of news-related tweets spread misinformation and noted that the activity of Trump supporters on social media influenced and preceded the activity of the top 100 misinformation disseminators. Other empirical works include Guess et al. (2018), who uncover evidence of heterogeneous effects conditional on partisan beliefs.[3] Employing a survey, they estimate that one out of four Americans had been visiting fake news websites in the weeks preceding their interview, and find that fact-checking websites had a limited reach on fake news consumers.

Indeed, long before the rise of the 'fake news epidemics', Nyhan and Reifler (2010) provided evidence that corrections are rarely successful in rectifying misconceptions

[3] As most of the traffic on fake news websites appears to originate from individuals with extreme conservative views.

among targeted groups. Similarly, Barrera Rodriguez et al. (2017) find that, while misleading statements from political figures are able to influence perceptions, once individuals update their preference for policy or candidates, those preferences are unaffected by fact checking.

Overall, these studies reveal heterogeneous patterns in access to quality information. Kennedy and Prat (2019) find a significant connection between socioeconomic inequality and information inequality: notably, poorly educated and low-income individuals have access to fewer media sources, and might then be vulnerable targets to misinformation. In this context, it is certainly worth mentioning the work of Roozenbeek and van der Linden (2018), which tests in a controlled experimental setting educational instruments that would allow the public to discern between true and false information.

Other studies, such as Törnberg (2018) and Azzimonti and Fernandes (2018), also attempt to model the spread of fake news and its effects on political polarization, providing a much needed theoretical anchoring to the study of misinformation. Finally, a number of studies has also focused on tracking the patterns of diffusion of misinformation (such as Shin et al., 2018, and Allcott et al., 2019), and on developing algorithms for detecting fake news (Shu et al., 2017).

While most of the aforementioned research points to an evident connection between fake news and populism, the assessment of a causal relationship between misinformation and voting behaviour has proven rather difficult. Due to a number of endogeneity issues, exacerbated by the difficulties in finding a proper quasi-experimental setting, it is still unclear how much online misinformation contributed to the rise of populism, and how much populist sentiments are linked to the inability to recognise fabricated information.

A first step in this direction has been made in Gunther et al. (2019), who find that, among Obama voters in the 2012 US presidential elections, those who believed in fake news were more likely to vote for Trump in 2016. These results are far from definitive: reverse causality may have led to an overestimation of these figures. A number of confounding factors correlated, for example, with the ability to recognise false information, could have led these people to update their political preference. The necessity to control for these unobservable trends and characteristics, along with the bias produced by focusing only on the fraction of the population who voted for Obama, leave the question open for further research.

Any study on the effects of online misinformation cannot disregard the work of Sunstein (2002, 2018) and Pariser (2011) on online echo chambers and filter bubbles, which ultimately highlights how the spread of misinformation is facilitated by online interactions taking place in extremely personalised social media environments. Past experiments on non-online interactions already showed how peer-effects in homogeneous groups can affect political beliefs and the perception of reality (see Schkade et al., 2007). More recently, Boutyline and Willer (2016) find that people sharing conservative or extreme political views tend to seek reaffirmation in their views associating themselves with similar individuals in their online interactions, while Del Vicario et al. (2016) show how false information is propagated through homogeneous online clusters – echo chambers, essentially – each characterised by different cascading dynamics.

There is evidence that social media echo chambers have already been exploited in a political context. Liberini et al. (2018) found that targeted political advertisements on Facebook effectively convinced specific groups of electors to vote for Donald Trump in 2016. The permeability of the social network bubbles between language groups living in the same area has not yet been investigated, and we intend to further explore this topic in the present article.

A number of works rely on quasi-experimental approaches to study political processes and voting behaviour empirically. Our research design draws inspiration from this literature and in particular from the works of Madestam et al. (2013), and Martin and Yurukoglu (2017). The first studied the effect of political protests on electoral outcomes in the US, taking advantage of the random variation of rainfall as an instrumental predictor for participation in Tea Party rallies. The second estimated the effect of exposure to biased news on television over voting behaviour by instrumenting channel positions as a predictor for viewership. In a similar fashion, Durante et al. (2019) studied the effect of entertainment television on electoral outcomes in Italy, exploiting the staggered introduction of Berlusconi's commercial TV network in the country since the 1980s. They found that, while municipalities exposed to entertainment TV displayed higher support for Berlusconi's party, in 2013 such support shifted to the populist Five Star Movement (M5s), suggesting that exposure to entertainment TV made voters generally more supportive of populist parties.

2.3 Background

2.3.1 Trentino and South Tyrol: political and sociolinguistic background

In the context of the 2018 Italian general elections, the *provincia autonoma* of Bolzano/Bozen ('South Tyrol' in the following[4]) serves as a natural experiment. Previously part of the Austrian Empire, it became part of the Italian unitary state in 1919. This Alpine province, home to little more than 500,000 inhabitants (527,750, Astat, 2017), is the only part of Italy where a sizeable majority of the population is not Italian-speaking.

In 2011, 69.7% of South Tyrol's population declared German as their first language (Astat, 2011). Italian speakers represent slightly less than one quarter of all inhabitants (118,000 people). They are concentrated in the largest municipalities – including the provincial capital, Bolzano/Bozen, where they make up almost three quarters of the population – and in the southernmost part of the province. Conversely, in smaller municipalities, a virtual totality of the population (usually around 95-97%) are German speakers. A third minority language, Ladin, also has legal status in the province: it is widely spoken as a first language in a number of small municipalities in the east.

Like South Tyrol, Trentino was integrated into the Italian unitary state only after the First World War: before that, it constituted the southern third of the Austrian county of Tirol. Other than shared history, the provinces feature striking similarities. Trentino, with 540,000 inhabitants (Istat, 2018) is only slightly more populated than South Tyrol; both provinces are highly rural, with a large share of the population living outside of the few mid-sized urban centres,[5] scattered across hundreds of very small municipalities. Most crucially, both enjoy a large degree

[4]Due to its complex history and strong nationalistic currents on either side of the language barrier, place name choices have long been a matter of contention in this territory. Owing to common usage in English-language sources, in this study we will adopt the name South Tyrol and the demonym South Tyrolean to refer to the territory and population of the officially trilingual *provincia autonoma* di Bolzano – Alto Adige (IT)/Autonome Provinz Bozen – Südtirol (DE)/Provinzia autonoma de Bulsan – Südtirol (Ladin). In Italian-language sources, 'Alto Adige', a name with a strong historical association with Italian nationalism, and the demonym 'altoatesini' are relatively more common. Together with the provincia autonoma of Trento (commonly referred to as Trentino), South Tyrol is part of the Trentino–Alto Adige/Südtirol region. In the following, we will adopt the official, bilingual denomination in full when referring to the region as a whole. For municipalities and other place names in South Tyrol we will generally use the official denomination (bilingual, Italian name first, separated by a slash, e.g. in Bolzano/Bozen). For the sake of simplicity, in our primary data municipalities are listed under their Italian-language name.

[5]Istat (2015); Principali dimensioni geostatistiche e grado di urbanizzazione del Paese; istat.it; Last accessed: 25/05/2019; Available at: https://www.istat.it/it/archivio/137001

of self-government compared to other Italian local authorities, with the *provincia autonoma* having significant legislative, fiscal and budgetary autonomy.[6]

The main divergence between the two provinces is therefore language: with limited exceptions, Trentino is essentially monolingually Italian. This is, to some extent, an intended effect of the post-Second World War settlement that confirmed the sovereignty of Italy over the area, fending off separatist tendencies among the German-speaking population. The De Gasperi-Gruber agreement between Italy and Austria (1946) enshrined a set of safeguards for the German community (such as native language education, public sector employment quotas, reversal of assimilationist practices) and is still the basis of the current institutional setup. It soon led to the transfer of a few German-majority municipalities from the province of Trento to that of Bolzano/Bozen. Indeed, subsequent pieces of legislation increased the degree of autonomy of South Tyrol not just from Rome, but from Trento too: tellingly and uniquely in Italy, legislative functions are given to the provincial authority (hence, *provincia autonoma*) and not to the Trentino–Alto Adige/Südtirol region, which is currently little more than a coordinating body.

All of this means that the border between the two provinces is akin to a linguistic frontier, with no meaningful German community to the south of it. Even though there are recognised and protected minority languages in Trentino – the aforementioned Ladin, plus two archaic Germanic dialects, Cimbro and Mocheno – their level of diffusion and their legal status is not comparable to that enjoyed by German (and Ladin) in South Tyrol. Nonetheless, census data shows that a few municipalities in the north-east of the province have a non-Italian (generally Ladin) majority (Servizio Statistica Provincia Trento, 2012).

South Tyrol acts as an unusual filter bubble in the Italian context, clearly showing evidence of linguistic segregation. *Dolomiten*, the main German-language newspaper, had in 2016 a circulation of 42,103 copies:[7] this is more than four times higher than that of its main, local Italian-language counterpart, *Alto Adige*, and over ten times higher than that registered by the most common mainstream newspapers of

[6]For an overview of the competencies of the *province autonome*: Trentino: Consiglio della provincia autonoma di Trento; "Le competenze legislative secondo lo statuto"; consiglio.provincia.tn.it; Last accessed: 25/05/2019; Available at: https://www.consiglio.provincia.tn.it/istituzione/l-autonomia/il-regime-delle-competenze-legislative-e-amministrative/Pages/il-quadro-delle-competenze-legislative-secondo-lo-statuto.aspx. South Tyrol: Amministrazione Provincia Bolzano; "Competenze e finanziamento dell'autonomia"; provincia.bz.it; Last accessed: 25/05/2019; Available at: http://www.provincia.bz.it/politica-diritto-relazioni-estere/autonomia/competenze-finanziamento-autonomia.asp

[7]Accertamenti Diffusione Stampa (ADS) (2016), *Dati medi annuali territoriali per testata: diffusione cartacea Italia*, available at: http://www.adsnotizie.it/_grafici.asp? (Last accessed August 2018)

the country ('Corriere della Sera' and 'La Repubblica'). Alongside Rai, the national broadcasting operator – which also maintains German-speaking radio and TV programming – a Province-funded broadcaster relays transmissions from surrounding German areas. Although both Italian and German are compulsory subjects for members of both language groups from age six onwards, and public sector jobs require high proficiency in either language, effective bilingualism is not widespread. Past studies, cited in Ebner (2016), point to a relatively low second language proficiency (L2) of the South Tyrolean population. In particular, less than 10% of 17-18 year old high school pupils of either language group are proficient in the other language (above C1 level in the Common European Framework of Reference for Languages).

The extent of the language divide is emphasised in the political dynamics of the two provinces. While devolution certainly plays a part in Trentino's politics, its voting trends, especially in national elections, do not dramatically diverge from those of neighbouring areas. Like in much of the Italian north-east, Trentino's dominant force used to be Democrazia Cristiana (DC), which scored over 50% of the votes in every general election until the 1980s. Since then, the province has experienced an increase in competition between the centre-left and the centre-right blocs. In fact, leaving 2018 aside – a year that has seen a decisive win for the centre-right parties, Lega in particular – the centre-left bloc fared relatively well in Trentino compared to the rest of Triveneto. This may be attributed to an increased prominence of localist Catholic-centrist parties, a force to be reckoned with in provincial elections since the 1990s, which have generally been aligned with centre-left parties at the national level (either running in coalition or backing nationwide lists). Indeed, no localist pro-autonomy party has ever reached double digits in general elections in Trentino.

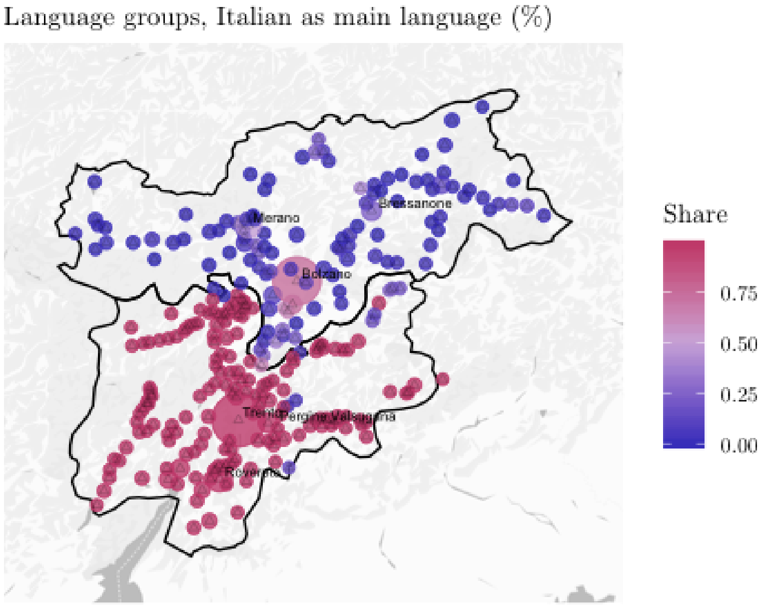
In South Tyrol, conversely, the situation is that of a parallel party system segmented across linguistic divides, to an extent that is not experienced anywhere else in Italy. Since the birth of the Italian Republic (1946), South Tyrolean politics has been marked by the dominance of the ethnic German South Tyrolean People's Party (Südtiroler Volkspartei, 'SVP' in the following). A typical catch-all party, Catholic-based, featuring both conservative and social-democratic wings, SVP obtained a majority of the votes cast in the province in every election between 1948 and 2008, scoring landslides (often more than 90% of the vote) in several of the purely German-speaking municipalities. A remarkable exception is the area of Bolzano/Bozen, where the SVP has rarely exceeded 25% of the votes. Nevertheless, the SVP can hardly be considered a populist platform and, in its history, has always linked itself with traditional political forces such as the DC before 1992 and the Partito Democratico (PD) and its predecessors afterwards.

In general elections, SVP has mostly faced provincial-wide competition from Italian parties only. In the last decades, however, several regionalist and separatist parties, which have long been a relevant force in local and provincial elections, have also ran. In 1996 – an election in which, due to a quirk of the electoral system then in force, the SVP logo did not appear on the ballot on its own right – a conservative splinter got almost 20% of the votes.

In 2008 and 2013 – elections that were fought with a proportional-based system, *Die Freiheitlichen* (DF), a radical-right separatist party akin to Austria's Freedom Party, received 9.5% and 15% of the vote, respectively. Consequently, in parallel with the electoral decline of the centre-left in Trentino, the provincial vote of the SVP has been progressively eroded, now stopping well short of 50% overall. Under the electoral law used in those years, a regionalist party with concentrated support could, at least theoretically, aspire to obtain a seat in the lower house of the Italian Parliament (Camera dei Deputati). Moreover, as the main feature of this system was the award of a "majority bonus" to the coalition of parties obtaining the largest share of the vote nationwide, South Tyrolean votes had also a national significance. However, also due to a high effective threshold, no German-speaking party other than the SVP won seats to the Italian Parliament. An alternative strategy, running candidates in the list of an Italian-based party, proved indeed successful for another local party, the Greens of South Tyrol, which won a single seat in 2013 in the Sinistra Ecologia Libertà list.

A change in the law before the 2018 elections slanted electoral competition even further towards SVP. Under the new system ("Rosatellum"), Trentino-Alto Adige/Südtirol elects six MPs to the lower house by a first-past-the-post system; the system used for the Senate has an even more dominant FPTP component. As a consequence, in 2018, DF did not field candidates in the national election, openly justifying the decision as an effect of the new electoral system.[8]

[8] Alto Adige (2017), *I Freiheitlichen*, altoadige.it, available at: http://www.altoadige.it/cronaca/bolzano/i-freiheitlichen-1.1478219



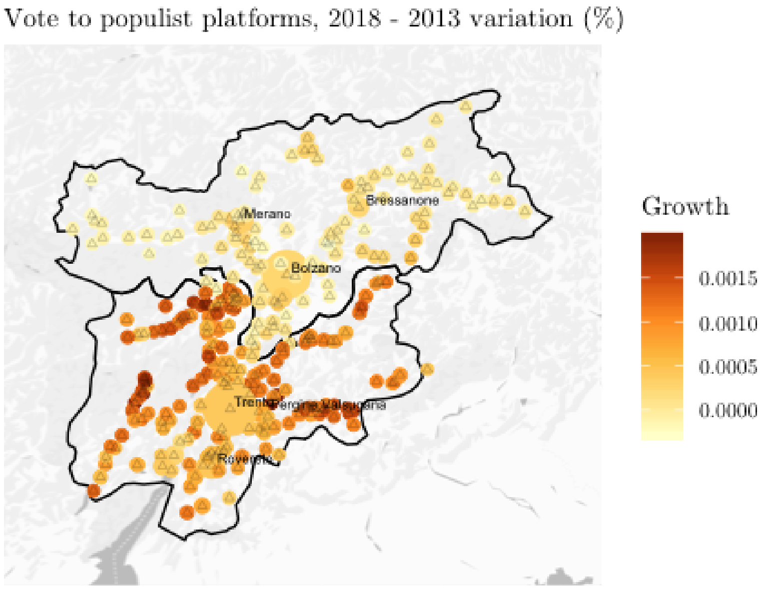


Figure 2.3.1: Language groups shares and growth of populist parties (2013-2018) in Trentino-Alto Adige/Südtirol. The size of the circle around each municipality indicates average electorate size between the two periods. Language groups shares are computed for the period in-between the two elections, estimated through a non-parametric linear interpolation of Aitchison's (1986) log-ratios for the 2001 and 2011 census data shares. Populist scores are computed by multiplying votes for each party for the relative score obtained through the 'assertive' dictionary method, summing all resulting scores by municipality, and then dividing for the number of voters.

Left as the only German party, the SVP scored more than 60% of the valid votes in majority-German constituencies. The vote share of Italian-speaking parties in the area - in particular, M5s and Lega - modestly increased, even though themselves were disadvantaged by the strongly majoritarian election law. The most significant effect was however registered on turnout: in the areas with the highest representation of German-speaking voters the drop in participation was very remarkable, from over 80% to slightly more than 50%. An unusually large number of spoilt and blank ballots was also recorded[9] It can thus easily be assumed that a rather large share of German-speaking voters opted to file their "protest vote" in this way.

Figure 2.3.1 displays the share of Italian speaking voters for each municipality in the Trentino-Alto Adige/Südtirol region,[10] along with the variation in populist vote between the two elections.[11]

While populist vote has been increasing across both linguistic groups, it is evident that populist platforms have had a greater appeal towards Italian-speaking voters. This trend seems to persist even in the areas where the two groups are more evenly mixed, and the increase appears particularly marked in Italian-speaking rural municipalities.

Of course, there are confounding factors to control for, and without information on exposure to misinformation the connection between fake news and populism growth cannot yet be quantified. Still, these figures prove that linguistic filter bubbles have had differential effects with regard to the rise of populist platforms.

2.3.2 Fake news in the 2018 Italian general elections

According to several journalistic[12] and institutional[13] sources, the campaign period leading to the 2018 Italian general election saw a remarkable spread of 'fake news', i.e. entirely baseless news stories published by non-institutional outlets, on social media.

These claims have been easily demonstrated by a recent investigation performed

[9] Italian Ministry of Interior, Archivio storico delle elezioni, Dipartimento per gli Affari Interni e Territoriali. Available at: https://elezionistorico.interno.gov.it/

[10] Source: authors' calculations from 2001 and 2011 census data; available at https://astat.provincia.bz.it/downloads/mit38_2012.pdf; and http://www.statistica.provincia.tn.it/statistiche/ societa/popolazione/; last accessed: March 9, 2019

[11] Source: Italian Ministry of Internal Affairs, Historical Archive of Italian Elections; available at: https://elezionistorico.interno.gov.it/index.php?tpel=C; last accessed: March 9, 2019

[12] See Buzzfeed News (2017); One Of The Biggest Alternative Media Networks In Italy Is Spreading Anti-Immigrant News And Misinformation On Facebook; available at: https://www.buzzfeed.com/albertonardelli/one-of-the-biggest-alternative-media-networks-in-italy-is; last accessed: 25 May 2019

[13] See Autorità per le Garanzie nelle Comunicazioni (2018)

by an Italian news channel[14]. Tracking a limited set of politically-charged keywords via a content analysis tool (see footnote 12), it has been found that among the top 100 articles in Italian for social media engagement, 5 were hoaxes, while another 10 were judged ‘highly problematic’ by the authors. While the ‘purely false’ news has a predominantly ‘anti-establishment’ character (and, by extension, anti-Democratic Party, the incumbent party), many of the ‘problematic’ (i.e. misleading) news in the sample focuses on immigration. According to an extensive report (Giglietto et al., 2018), immigration and security were dominant themes on social media during the run-up to the 2018 Italian general election.[15]

While few in number, the exposure reached by these hoaxes was highly significant. The second-most shared news in the SkyTg24 database, published on the day before the election and consisting of an entirely unsubstantiated report of voter fraud in Sicily to the advantage of PD, received more than 140,000 interactions, mostly on Facebook. As pointed out by some observers, the news seemed to target a public which was friendly to the 5 Star Movement, and was widely shared among M5S supporters (including an MP).

In addition, the name of the originating website, the now offline ‘ilfatto.it’, is a potentially misleading reference to *Il Fatto Quotidiano*, a daily newspaper popular among M5S voters. This ‘spoofing’ tactic of mirroring more reputable news sources was exploited by a number of partisan outlets that occasionally managed to create an engagement on social media comparable to their ‘genuine’ peers. Unsurprisingly, their URL also appeared in ‘blacklists’ compiled by the most popular fact-checking websites in the country (Butac.it, Bufale.net).[16]

Giglietto et al. (2018) provides a detailed classification of news sources based on partisanship of their news content, evidencing that a vast majority of ‘non-institutional’ websites feature some form of bias towards of the largest Italian populist parties, Lega and M5S. Crucially, comparable biased sources supporting pro-establishment (and smaller anti-establishment) parties captured much less social media attention than the pro-M5S and Lega networks.

The study however stops short of establishing a link between the spread of false

[14]Bruno, Nicola (2018); Satira e fake news: gli articoli più condivisi delle Elezioni 2018; Available at: https://tg24.sky.it/politica/2018/03/08/fake-news-elezioni-2018.html

[15]This was also influenced by a real, contentious news story involving the alleged murder of an Italian teenager in Macerata, Central Italy, which was followed by a racially motivated attack against a group of Africans in the town. See also: The New York Times (2018); This Italian Town Once Welcomed Migrants. Now, It’s a Symbol for Right-Wing Politics; available at: https://www.nytimes.com/2018/07/07/world/europe/italy-macerata-migrants.html; last accessed: 25 May 2019

[16]These black lists are available at the following links: https://www.butac.it/the-black-list/ and https://www.bufale.net/the-black-list-la-lista-nera-del-web/. Last accessed: 29 May 2019.

information in the electorate and the support for Italian populist parties and their policy stances. Nonetheless, recent research by Avaaz[17] provides evidence in support of this link, at least as far as Facebook is concerned. The report uncovered an extensive network of Facebook pages and fake accounts (with a reach of millions of interactions) that, in blatant violation of the platform's terms of use, pushed misinformation in support of Lega, M5S and other anti-establishment and fringe causes – including antisemitism and racism.[18]

As a complement to these findings, we provide further evidence for this anti-establishment bias by analysing the topics occurring in a sample of fake news propagated between March 2016 and March 2018, the month of the elections. We do so by analysing entry metadata scraped from an independent Italian debunking website (Butac.it) and retrieving from it the text of all fake news headlines it reported.[19]

Using a simple text mining technique on fake news headlines,[20] we search for recurring mentions of parties and leaders of incumbent/establishment and challenging/anti-establishment platforms, along with terms denoting topics which have been predominant during the 2018 electoral campaign: namely, immigration, foreign policy and the European Union, and vaccinations.[21]

Figure 2.3.2 plots the relative frequency of these topics over a two-month period, showing that the challenging platform received much less attention than any other topic. Moreover, fake news' mentions to anti-establishment parties and leaders were overall less frequent than mentions to incumbent ones, providing additional descriptive evidence in support of fake news' anti-establishment bias.

[17]La Repubblica (2019); Facebook chiude 23 pagine italiane con 2.4 milioni di follower: diffondevano fake news e parole d'odio; available at: https://www.repubblica.it/tecnologia/social-network/2019/05/12/news/facebook_chiude_23_pagine_italiane_con_2_4_milioni_di_follwers_diffondevano_fake_news_e_parole_d_odio-226098817/; last accessed: 13 May 2019

[18]As a result, in May 2019 Facebook took down 23 of the pages reported by Avaaz. While some of these pages explicitly portrayed themselves as unofficial supporters of the two parties, others used a more subtle approach. The most popular page in this subset, 'I Valori Della Vita' (1.5m followers), ostensibly a lifestyle page, was actually part of a bigger network, sharing in a coordinated manner content from a right-wing, pro-Lega news site.

[19]The website was scraped in January 2019.

[20]We use a dictionary technique which computes a score for each topic based on the occurrence of key terms related to each topic in the textual database. For this purpose, dictionary techniques require the creation of list of key words that relate to a specific topic and that the algorithm will search in the text subject to the analysis, that is the fake news headlines in our case.

[21]The text bags used to construct these indicators are based on the available fake news headlines. The codes and data used in Figure 2.3.2 (including all text bags) are available in the online data archive.

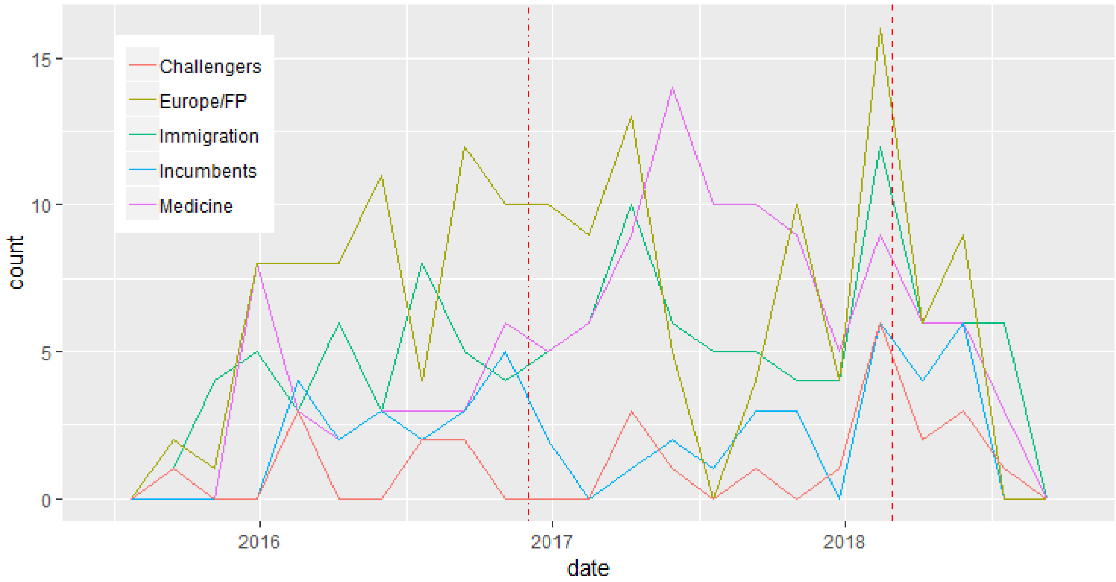


Figure 2.3.2: bimonthly frequency of pieces of fake news by topic. The dot-dashed and dashed red lines indicate the date of the 2016 Constitutional Referendum and the 2018 General Elections respectively.

In addition, it should be noted that, when a false fact pertains to the challenger, the topic does not necessarily indicate the bias of the piece of misinformation: while most headlines pertaining to the incumbents, the EU, immigration and vaccinations show an obvious anti-establishment bias, the same cannot be said for the challenging platform. On the contrary, most of the few pieces of misinformation concerning the Five Star Movement and the Lega take a favourable and supportive stance with regard to these parties, meaning that the actually damaging headlines are even sparser than what the figure would suggest. Indeed, within the time window we analysed, only 6 among all headlines reported by Butac.it could be said to have an anti-M5S or anti-Lega slant, while 2 of them were directed against the neo-fascist parties Casapound and Forza Nuova.[22] It is also interesting to note that each topic experienced differential increments in their frequency in proximity to the 2016 Constitutional Referendum and the 2018 General Elections, suggesting that the propagation of fake news may be driven by electoral purposes.

While it is not our intention to assess whether the spread of 'fake news' was an integral part of Italian populist parties' strategy and campaign, there is enough

[22]The 4 headlines with an anti-M5S stance are available at these links: https://www.butac.it/di-maio-emiliano-e-matera/; https://www.butac.it/sindaco-grillino-attacca-zanardi/; https://www.butac.it/sahaja-yoga-e-la-candidata-del-m5s/; https://www.butac.it/piccole-perle-di-facebbok-i-soldi-del-m5s/. The other 2 headlines concerning the Lega are here available: https://www.butac.it/salvini-i-giovani-padani-e-quelli-del-mezzogiorno/; https://www.butac.it/democratica-salvini-calabresi/. Last accessed: 29 May 2019

empirical evidence to assume that misinformation and outright hoaxes have been exploited by partisans of Lega and M5S on social media in the run-up to the 2018 general election, and that this effort was much more successful than that of their pro-establishment counterparts *in terms of social media exposure*. The goal of our study is to assess whether this social media success causally translated to shifts in voting behaviour.

2.4 Data

Due to the lack of individual-level micro-data on fake news exposure and voting preferences, we used municipalities, where much more information is available, as the principal unit of analysis of this study. We collected information on voting preferences, exposure to fake news, and important socio-economic features such as linguistic group shares, income levels and internet coverage from a number of different sources. Moreover, we included information on populist stances for each party running for election.

Between the two elections some municipalities were suppressed and new ones were formed from their merger; accordingly a few adjustments had to be made. For all suppressed municipalities from 2013, values for the corresponding 2018 entities were imputed either by summing the totals or, in the case of per capita figures, by averaging the population-weighted sum of the suppressed municipalities.

The present section provides an overview on how these sources were gathered and harmonised. Summary statistics on the core variables used in our analysis are presented in Table 2.4.1. Further information on how misinformation data was collected and processed is available in Appendix A. All data sources and codes used to construct our final dataset and perform our analyses are available in the online data archive.

2.4.1 Electoral data

Official municipality-level data on general election results from the Trento and Bolzano/Bozen provinces are obtained from the Italian Ministry of the Interior,[23] where the complete election history of the Italian Republic is available. We extracted information on electoral results from the 2013 and 2018 general elections.

Other than indicating how many votes each party received in every election, the data set also includes a number of auxiliary variables disaggregated by municipality.

[23] Italian Ministry of the Interior, Historical Archive of Italian Elections; available at: https://elezionistorico.interno.gov.it/index.php?tpel=C; last accessed: March 9, 2019

Among these, electorate size and total number of abstentions or invalid votes were of particular interest for our research. While voting outcomes have been later used to construct our dependent variable, these other variables have found important applications in our work. Electorate size, other than being used for weighting other variables, has proven an important predictor of populist preferences. The same can be said for non-voting behaviour or abstention, which has also been used to control for potential voters of the Die Freiheitlichen party across the German-speaking population.

2.4.2 Socio-demographic and internet connectivity data

The proportion of language groups by municipality is a key variable in our identification strategy. Language group shares for the region can be obtained through census data.[24] However, as census data for Italy is only released once every 10 years, the latest figures available date back to 2011. To compensate for differing trends in population growth across language group which may have affected our analysis, we perform a small and simple adjustment by interpolating the figures from 2001 and 2011 to predict the group-specific shares for the years 2013 and 2018, using Aitchison's (1986) log-ratio transformation to preserve the compositional form of the data. While the great majority of locations in Trentino are overwhelmingly Italian-speaking, language groups shares can vary significantly within South Tyrol: consequently, this indicator allows us to address the correlations between language groups and exposure to misinformation with much more precision than a binary discriminant between the two provinces would allow.

To reconstruct per capita income by municipality, we used tax data from the Italian Ministry of Economy and Finances for the years 2012 and 2017.[25] Those calculations are based on self-declared taxable income from the *Imposta sul reddito delle persone fisiche* (IRPEF), and per capita figures are already available by municipality. The connection between income and voting preference is not unambiguous, but has been discussed extensively in Galbraith and Hale (2008), Lewis-Beck and Nadeau (2011), and Hersh and Nall (2015), among others. Income differences can also arise across language groups, motivating our choice to control for them in our econometric specification.

[24]Province-specific census data on language shares is available at: https://astat.provincia.bz.it/downloads /mit38_2012.pdf; and http://www.statistica.provincia.tn.it/statistiche/societa/popolazione/; last accessed: March 9, 2019

[25]Available at: https://www1.finanze.gov.it/finanze3/analisi_stat/index.php?tree=2013 and https://www1.finanze.gov.it/finanze3/analisi_stat/index.php?tree=2018; last accessed: March 31, 2019

Two different sources were used to construct statistics for broadband internet connectivity in 2013 and 2018. As mentioned previously, Campante et al. (2017) found that increases in internet connectivity have had an effect on voting preferences in Italy, suggesting that a measure for connection should be used as a control variable in our econometric model or even as an alternative to our language group instrument. Also, as Italian regions – including the two *province autonome* of Trentino and Alto Adige/Südtirol – possess considerable autonomy in the implementation of broadband infrastructures, this indicator then plays a vital role in controlling for differences in connectivity arising from staggered local legislation.

The 'digital divide' in Italy, defined as the share of households not covered by broadband connection, was originally covered by a set of municipality-specific indicators released by the 'Agenzia per la Coesione Territoriale'.[26] This indicator, which defined any landline connection whose speed exceeded 2Mbps as broadband, was then subtracted from unity to compute the share of low latency connections by municipality. As these publications were discontinued, the collection duty was then moved to the 'Autorità per le Garanzie nelle Comunicazioni' (AgCom), which has released internet connectivity indicators for 2018.[27] The new variables released by AgCom were slightly different from the previous ones, but we managed to reconstruct a fully comparable digital divide indicator using the available information on the number of households with a landline connection faster than 2Mbps.

2.4.3 Social media data

Information on fake news exposure and on the social media campaigns of Italian parties was all scraped from Facebook. Our fake news exposure indicator shows, for each municipality and for each year of election, the estimated number of Facebook likes being held by all Facebook fan pages that are known to spread politically-charged misinformation.[28] Due to the unavailability of granular data on the spread of each piece of false information, we decided to focus our attention on their 'disseminators', measuring – in a given municipality – the effect of each unitary increase in their social media following on aggregate electoral preferences. Our estimates then include both the 'intensive' and 'extensive' margin of fake news exposure, under the implicit assumption that individuals who 'liked' these pages also play an active part

[26] Available at: http://old2018.agenziacoesione.gov.it/it/arint/OpenAreeInterne/index.html; last accessed: March 31, 2019

[27] Available at: https://maps.agcom.it/#; last accessed: March 31, 2019

[28] Likes from all pages which appeared in black lists compiled by debunking websites Butac.it and Bufale.net (as discussed in section 3.2) were estimated. See Appendix A for the full list of pages used in the process.

Table 2.4.1: Summary statistics by province and year

	(1) Bolzano/Bozen			(2) Trento		
	2013	2018	Average	2013	2018	Average
Populist score (total)	5.972	6.526	6.249	4.123	5.899	5.011
	(13.376)	(15.823)	(14.621)	(12.074)	(15.465)	(13.883)
Exposed to fake news	148.424	376.876	262.650	199.858	521.874	360.866
	(869.421)	(2366.432)	(1782.496)	(1206.976)	(3459.335)	(2592.059)
Broadband connections	2469.032	3007.955	2738.494	1987.041	2217.626	2102.334
	(7542.113)	(7661.528)	(7590.388)	(6763.404)	(7030.532)	(6889.395)
Italian speaking voters	879.493	907.630	893.562	2173.312	2226.701	2200.007
	(5321.100)	(5435.457)	(5366.947)	(6837.192)	(6963.666)	(6890.933)
Income per capita	18721.313	20692.652	19706.982	17193.012	18434.036	17813.524
	(2534.504)	(2728.883)	(2807.309)	(2188.215)	(1891.333)	(2134.696)
Electorate size	3245.284	3332.853	3289.069	2276.602	2330.716	2303.659
	(7548.803)	(7683.594)	(7600.120)	(6986.644)	(7112.244)	(7039.726)
Abstentions and invalid votes	677.569	1345.802	1011.685	524.455	649.989	587.222
	(1668.889)	(2502.587)	(2148.627)	(1499.638)	(1966.976)	(1747.624)
Observations	232			352		

*N*otes: Mean coefficients, standard errors in parentheses.

in spreading misinformation. The lengthy data collection and estimation process behind the construction of this variable is reported in Appendix A.

Our research question also required us to transpose categorical voting decisions into a continuous scale measuring affinity to populist discourse by municipality. The construction of such an indicator required access to propaganda content used over the course of the two electoral campaigns preceding the 2013 and 2018 elections. We turned, again, to Facebook, from which we scraped all posts from running parties and their leaders, covering all three months preceding the elections in 2013 and 2018, and coinciding with the beginning and the end of each electoral campaign. A more detailed description of the data and calculations leading to the production of our final indicators is provided in the next section.

2.5 A text mining approach to measuring the populist content of parties

As described in Section 3, descriptive evidence suggested that fake news benefits populist parties. In order to investigate the electoral impact of fake news, the dependent variable of our empirical model will be represented by the electoral performance of these platforms. Nevertheless, defining the degree of populism of each party in an objective manner is not a trivial exercise.

To overcome the issue of objectivity, we apply text analysis to the Facebook posts of political parties and leaders that ran in the 2013 and 2018 Italian elections. In doing so, we define populists as those parties that (1) share an anti-elite rhetoric (Albertazzi and McDonnell, 2008; Pauwels, 2011; Rooduijn and Pauwels, 2011; Kaltwasser et al., 2017) and that (2) tend to adopt a particularly emotional language in their campaign (Taggart, 2000; Rooduijn, 2014; Caiani and Graziani, 2016; Bischof and Senninger, 2018). These two elements are generally complementary and do not exclude one another. As found by Guiso et al. (2017), in fact, the supply of populist parties tends to be higher where disappointment with traditional parties is greater. This suggests a correlation between the rise of populist parties and an emotional narrative against incumbent politicians perceived as part of the establishment.

Following these considerations, we introduce a new methodology to measure the degree of populist rhetoric in each party based on text analysis. In particular, we create two dictionaries that capture (1) an anti-establishment rhetoric, and (2) a more emotional or assertive tone, and match them with the text produced on Facebook by all the Italian parties that ran in the 2013 and 2018 elections and their leaders.[29] The first dictionary contains 23 words which capture a populist tone in the Italian political language; for example, the words 'establishment' or 'caste'. The second dictionary simply counts the number of exclamation marks as a proxy of an emotional tone.[30] While previous works applied dictionary techniques to measure populism on party manifestos (Rooduijn and Pauwels, 2011; Pauwels, 2011), this paper is the first to our knowledge that use them to measure the degree of populism in political communication on social media.

We then web-scrape the text of all the Facebook posts of Italian parties (12.159 posts) and their political leaders (8.164 posts) that posted three months before the Italian elections in 2013 and 2018. After pre-processing the text,[31] we first match the words of our two dictionaries against all Facebook posts collected. Formally, we

[29]We decided to focus on the rhetorical aspect of populism rather than on party-specific policy stances. This choice was led by the fact that, while on the one hand defining populism as linked to a specific policy stance is controversial, as it has often been applied to very different contexts (Caiani and Graziani, 2016), on the other hand, populism can be more easily identified by its forms of communication rather than by specific ideological stances (de Vreese et al., 2018).

[30]Previous works in computer science, such as Kumar and Sebastian (2012), found exclamation marks to be a good proxy for the prediction of emotional statements on social media.

[31]In order to allow for matching, we tokenise the text, remove Italian stopwords, punctuation (excluding exclamation marks in the second dictionary), numbers and white spaces and transform all terms in lower case. We also record the number of social media interactions for each post, a duplicate each text-bag in German so that scores can be computed for the SVP and DF parties. With the exception of the text bag capturing assertive tones, which relies on the computation of exclamation marks, we also remove punctuation from the posts.

compute the following indicator for each text bag:

$$Popscore_{tp} = \sum_{d=1}^{D} (Freq_{tpd}/T_{tpd}) * Engagement_{tpd} * 10 \tag{2.1}$$

where $Freq$ is the number of words matched for each text bag in each Facebook post d published by each party/leader $p = \{1, ..., P\}$ three months ahead of each election $t \in [2013, 2018]$. T is the total number of words contained in each Facebook post d, and $Engagement$ is the sum of all Facebook shares and likes a given post has received.

In other words, we sum the number of matches ($Freq$) and weight them by the total number of words[32] (T) and by the number of interactions ($Engagement$) for each party. Values of $Popscore$ for each party and elections are displayed in Fig. 2.5.1 for both the dictionary (panel on the left) and the exclamation marks (right). To obtain an indicator of populism at the municipal level, we multiply the populist score of party p in election t ($Popscore_{tp}$) by the votes gained by the same party p in municipality i during election t. Formally, we aggregate the scores as follows:

$Y_{ti} = \sum_{p=1}^{P} Popscore_{tp} * Votes\_received_{tpi}$

where $Votes\_received$ is the number of votes received by each party p in each election t in each municipality i. Y_{ti} will be our indicator of populism at municipal level and our main dependent variable. Given these transformations, our scores should not be assessed by looking at their absolute values, which are necessarily low, but rather by focussing on the relative distance in the scores between parties.

The two indexes display similar trends, further confirming the belief that anti-establishment discourses and emotional tones are generally correlated in political communication. In particular, the Lega clearly features as the most populist party in the election of 2018 in both cases, as a result of a steep increase in populist language from 2013. On the other hand, the 5 Star Movement (M5s) shows a lower degree of populist language in 2018 than in 2013, for both cases. In line with our expectations, the incumbent Democratic Party (PD) displays lower levels of populism. The ability of our methodologies to proxy for a populist language is further reinforced by the proximity of fringe parties, which are generally anti-establishment, to the Lega and the Five Star Movement, such as the far-left Potere al Popolo (PaP) and the neofascist Casapound (CP).

[32] This practice aims to avoid that long posts inflate the populist score. Intuitively, posts containing larger number of words are more likely to contain a higher number of words that are matched.

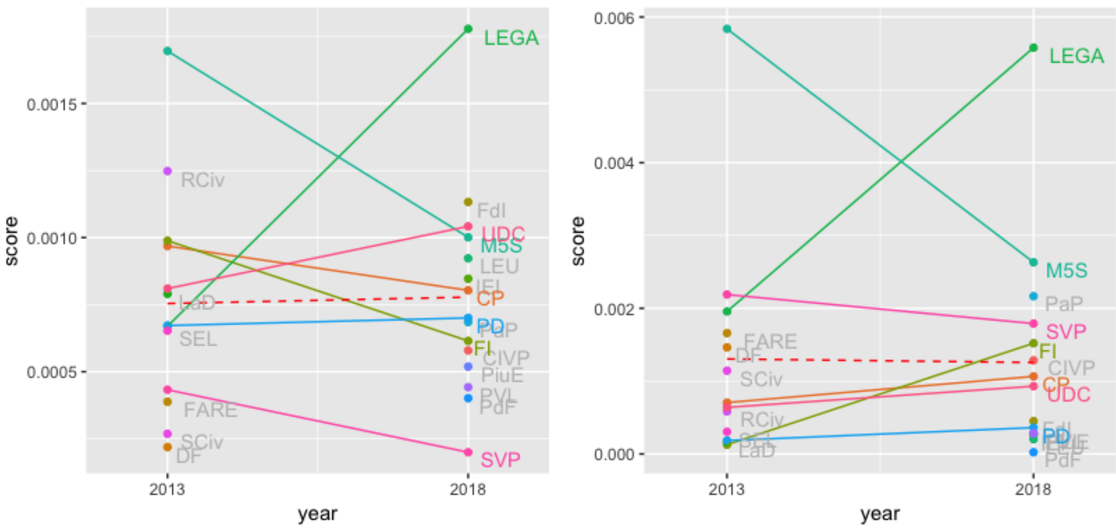


Figure 2.5.1: Text analysis scores of social media posts from parties and their leaders during the 2013 and 2018 elections campaigns. The left figure refers to the scores obtained with the populist text-bag, while the right figure computes the same score as the frequency of exclamation marks in the text. Parties in grey have only took part in one of the two elections. The red dashed line refers to the election specific average.

However, such proximity is not constant as, for example, in the first indicator PaP has a similar score to the PD, suggesting that the potential of the indicators might be weaker for minor formations. On the other hand, some of these counter-intuitive results might find an explanation through a deeper study of political dynamics. The clearest example is UDC, a centrist moderate party, which presented particularly high scores in 2018. This shift might be explained by a transformation in the communication style of UDC following the change of coalition in the 2018 election. While in 2013 it ran in coalition with Scelta Civica, a party founded by the former technocratic prime minister Mario Monti, which in fact scores relatively low in both indicators, in 2018 UDC joined a centre-right wing coalition formed by the Lega, Forza Italia (FI) and Brothers of Italy (FdI), which all report high scores in 2018. By inspecting the words matched under the first methodology, terms like ‘scandal’ or ‘shame’ appear much more frequently in 2018, reflecting this change in communication style during the two elections.

Despite their relative limitations, these results confirm the widespread categorisation of parties such as the Lega and the M5S as populist. As a further test, we compared our results to the values assigned by political experts in the Chapel Hill Expert Survey (Polk et al., 2017)[33] to a number of parties for the variable ‘People vs the Elites’, which measures the ‘salience of anti-establishment and anti-elite rhetoric’.

[33] Data are available at the following link: https://www.chesdata.eu/.

Text analysis scores display a positive correlation with the anti-establishment indicator based on experts' perceptions, in particular displaying the Lega and the Five Star Movement among the most populist parties in Italy (see Figures 2.C.1 and 2.C.2 in Appendix C).

2.6 Econometric model

In an ideal experimental setting, voters would be randomly assigned into two groups: in the 'treatment' group, individuals would be exposed to fake news, while voters in the 'control' group would have access to reputable sources of information only. In this case, differences in voting behaviour would be explained by the assignment to the experimental unit and the average treatment effect would be obtained by the simple difference in means across groups. Such an experiment, however, would be difficult to replicate, as we believe that prior exposure to politically charged misinformation can have lasting effects on voting behaviour that could affect the outcome of the experiment and bias estimates downward. Indeed, random assignment cannot control whether individuals in the control group have already been exposed to fake news in the past. Preventing access to the control group based on exposure to fake news would also invalidate the research, as such exposure could be linked to factors that also influence voting preferences. For this experiment to be unadulterated by environmental factors influencing our object of study, every element of it – treatment and outcome included – has to be detached from their connections with real-world politics, something very difficult to achieve.

This is then one of the few instances when a quasi-experimental setting would be preferable to an experimental one. The South Tyrol province is an ideal setting for a case study, as individuals are randomly assigned at birth in one of two different linguistic groups. It is not unreasonable to argue that each of these groups is exposed to misinformation concerning Italian national politics in a different way: as mentioned earlier, and following Allcott and Gentzkow (2017), fake news disseminators may have little to no incentive to produce misinformation across the German-speaking population, whose spread would also be isolated by the linguistic echo chamber each voter belongs to.

There are, however, some complications that need to be addressed. The first is the different electoral patterns across the two linguistic groups. As the SVP has consistently proven to be a popular voting choice across the German-speaking population, our estimates would suffer from extreme upward bias when not controlling for previous elections. To account for this issue, we would first employ a difference in differences framework, taking advantage of the different filter bubbles generated across the two linguistic groups in a mixed language area, constructing the treatment group from the Italian-speaking individuals in the sample, and the control group from their German-speaking counterparts. We would then investigate differences in voting outcomes, exploiting the changes between the 2013 and 2018 elections. In this way, diverging electoral trends can be accounted for. Our diff-in-diff specification would take the following form:

(6.1) $Y_{ti} = \alpha + (Z_{ti} \times P_t)\lambda + Z_{ti}\delta + P_t\zeta + X'_{ti}\gamma + e_{ti}$

where i represents each municipality and $t \in [2013; 2018]$ stands for each election ranked by their date. The variable Z proxies for the exposure to fake news for voters in municipality i. In this specification Z is the share of Italian-speaking population in one municipality, i.e. the share of population potentially exposed to fake news due to their linguistic characteristics. P is a dummy that captures time fixed effects for the two elections,[34] whereas X is a vector of covariates controlling for demographic

[34]The inclusion of time fixed effects is fundamental to control for unobservable differences between one election and the other. The time dummy will hence include the potential effect of the change in the electoral law in 2017 (informally known as 'Rosatellum'). Under the older electoral law (so called 'Legge Calderoli'), in force between 2006 and 2013, the mechanisms used for the lower house ('Camera dei Deputati') and the upper house ('Senato della Repubblica') significantly diverged. The system chosen for Camera was a form of proportional representation with a majoritarian correction: the party or coalition of parties obtaining the largest share of the vote nationwide were assigned automatically 55 percent of the seats; the remainder was distributed among lists either crossing 4 percent at the national level (with exceptions for lists running in coalitions) or reaching 20 percent in a single electoral district.

The latter threshold implies that a regionalist party with concentrated support could have at least a slim chance to win seats; however, in the South Tyrolean case, the nature of electoral competition (i.e. the dominance of the SVP) made this event unlikely, and arguably a sub-optimal strategy compared to e.g. running candidates within the lists of Italian parties with a national voter base. Indeed, the South Tyrolean Greens, which ran on a joint list with Sinistra Ecologia Libertà, managed to obtain one seat through this strategy, while Die Freiheitlichen, which generally outperforms the Greens in provincial elections, won none via their stand-alone list. It is also noteworthy that Trentino-Alto Adige/Südtirol were included in the same district, further penalising the parties of the German-speaking community not running in coalition - as their vote share was diluted by competition with Italian-speaking parties in Trentino.

As mandated by the Italian Constitution, the Senato cannot be elected via constituencies that comprise more than one regional authority (i.e., only regional and sub-regional constituencies are allowed), which prevented the introduction of a 'nationalised' election system as described above: that means that 'region-wide', and not nationwide majority bonuses were present. However, Trentino-Alto Adige/Südtirol deviated even further: six out of seven seats assigned to the region were elected in single-member constituencies under a first-past-the-post (FPTP) system. This system favours the SVP even more, which did not even need to run in a formal coalition to win the two rural seats in South Tyrol.

The 'Rosatellum' scraps the majority bonus, introducing a mixed formula which is approximately 2/3 proportional and for 1/3 based on FPTP constituencies. In this sense, only minor differences exist between the system for the lower and the upper house (except, in practice, constituency size and magnitude, as the Senato has half the seats of the Camera).

However, for Trentino-Alto Adige/Südtirol only, the balance between proportional and FPTP seats is slanted towards the latter: in the Camera, six members are elected in single member and economic characteristics. As long as assignment to the Italian linguistic group – identified by Z – correctly proxies for exposure to misinformation, and that exposure to fake news took place only prior to the latest elections, then the coefficient λ of the interaction term $Z_i \times P_t$ indicates the effect of fake news on populist preferences Y in each municipality.

However, as there are reasons to believe that the two language groups are not entirely isolated from each other, we reckon that a simple diff-in-diff specification may not suffice: random assignment to the language group may rather affect the intensity of exposure to fake news. To account for 'permeability' of linguistic filter bubbles, we then turn to a two stage least squares specification where we make use of language group as an instrumental variable, while controlling for both time variant and invariant determinants. We then propose an alternative specification, for $t \in [2013; 2018]$:

(6.2.1) $F_{ti} = \alpha_1 + (Z_{ti} \times P_t)\lambda_1 + Z_{ti}\delta_1 + P_t\zeta_1 + X'_{ti}\gamma_1 + e_{1ti}$

(6.2.2) $(F_{ti} \times P_t) = \alpha_2 + (Z_{ti} \times P_t)\lambda_2 + Z_{ti}\delta_2 + P_t\zeta_2 + X'_{ti}\gamma_2 + e_{2ti}$

(6.2.3) $Y_{ti} = \alpha_3 + (\widehat{F_{ti} \times P_t})\lambda_3 + \hat{F}_{ti}\delta_3 + P_t\zeta_3 + X'_{ti}\gamma_3 + e_{3ti}$

In the present specification, we apply an alternative definition of exposure to fake news, based on the number of likes to Facebook pages that diffuse fake news in municipality i at time t. This exposure to fake news is represented by F in regressions (6.2.1) and (6.2.2), where its value – and the value of its interaction with the year of election in (6.2.2) – is predicted by the instrument Z (indicating, again, assignment to the Italian linguistic group) and the other covariates. The fitted values $\hat{F}$ and $F_i \hat{\times} P_t$ from first stage regressions (6.2.1) and (6.2.2) are then plugged into equation (6.2.3) to predict populist preferences Y. These preferences are computed by multiplying the number of votes each party received in each municipality and for each election by the party and election-specific populist scores we obtained earlier.

What we develop is essentially a diff-in-diff model where treatment – exposure to fake news – is predicted by the assignment to the language group. In this way, the relationship between linguistic groups and exposure to fake news can be tested in the first stages of our model (6.2.1) and (6.2.2), rather than naively assuming that the

constituencies and five in the proportional remainder - in which only lists crossing a 20 percent threshold at the regional level (or three percent nationwide) could participate. In the Senato, the system is identical to that used in the previous law: six seats out of seven are FPTP.

While significant, the changes described above only marginally altered the nature of electoral competition in Trentino-Alto Adige/Südtirol, and among German-speaking parties in particular. The arise of a competitor to the SVP for German- and Ladin-speaking voters was already strongly constrained in 2013 and rendered almost impossible in 2018. This led, as pointed out in the background section, the strongest separatist party, Die Freiheitlichen, to sit out the latter election entirely, something that is reflected by the decrease in turnout - and accounted for in our methodology.

German-speaking population is completely unexposed to fake news as in our initial diff-in-diff framework, as we believe that membership to a linguistic community influences the intensity of exposure rather than fully determining it. If randomisation is achieved through assignment in a linguistic community and if we are able to control for individual qualities and common trends across the two groups, then the coefficient λ_3 of the instrumented interaction term between predicted exposure and year of election will capture the causal effect of fake news exposure on electoral behaviour. An interesting implication of this new setting is that equation (6.1) will stand as the reduced form of this model, explaining the direct effect between the instrument and our outcome of interest.

2.7 Results

Table 2.7.1: OLS estimates of the effect of misinformation on populist vote

VARIABLES	(1) OLS	(2) OLS	(3) OLS	(4) OLS
Exposed to fake news	-0.002	-0.003***	-0.003***	-0.003***
	(0.002)	(0.001)	(0.000)	(0.000)
Exposed to fake news × Year of election	0.002*	0.003***	0.003***	0.003***
	(0.001)	(0.000)	(0.000)	(0.000)
Year of election	0.778***	0.211**	0.114	0.728***
	(0.073)	(0.100)	(0.107)	(0.077)
Italian speaking voters	0.002***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Broadband connections		0.002***	0.002***	-0.000
		(0.000)	(0.000)	(0.000)
Income per capita (natural log)			1.251***	0.167
			(0.352)	(0.161)
Electorate size				0.002***
				(0.000)
Abstentions and invalid votes				-0.000
				(0.000)
Observations	584	584	584	584
Adjusted R-squared	0.934	0.992	0.992	0.997

*N*otes: OLS estimates for the effect of misinformation on populist vote. Populist scores computed using the 'Assertive' text bag. Standard errors robust to clustering by municipality in parentheses.

*p<.05; **p<.01; ***p<.001

Table 2.7.1 shows the results from an initial linear model, using, as for all subsequent specifications, the natural logarithm of total populist scores by municipality as the outcome variable. This has been computed by aggregating text analysis scores obtained through the 'Assertive' dictionary and weighting them by the size of the electorate. Estimates from columns (1) to (4) indicate that the positive correlation between exposure to fake news and populist vote is robust to the addition of various controls.

Table 2.7.2: 2SLS estimates of the effect of misinformation on populist vote

	(1)	(2)	(3)	(4)
	DiD	First Stages		2SLS
VARIABLES		Exposition	Interaction	
Italian speaking voters	-0.0001	0.131**	-0.009	
	(0.0001)	(0.054)	(0.023)	
Italian speaking voters × Year of election	0.0005***	0.297***	0.455***	
	(0.0000)	(0.039)	(0.057)	
Year of election	0.4336***	-207.135***	-332.732***	0.778***
	(0.0689)	(33.100)	(46.819)	(0.098)
Broadband connections	-0.0000	0.021	0.013	-0.000
	(0.0001)	(0.079)	(0.043)	(0.000)
Electorate size	0.0019***	0.022	-0.019	0.002***
	(0.0001)	(0.053)	(0.051)	(0.000)
Income per capita (natural log)	0.4947***	-143.422	-37.960	0.491
	(0.1908)	(171.636)	(124.660)	(0.299)
Abstentions and invalid votes	-0.0003	-0.047	0.063	-0.000
	(0.0003)	(0.143)	(0.158)	(0.000)
Exposed to fake news				-0.000
				(0.001)
Exposed to fake news × Year of election				0.001***
				(0.000)
Observations	584	584	584	584
Adjusted R-squared	0.9972	0.931	0.930	0.994
Partial R-squared		0.342	0.397	
F-Test		45.03	59.19	

*N*otes: IV estimates (including reduced form - DiD - and first stages) for the effect of misinformation on populist vote. Populist scores computed using the 'Assertive' text bag. F-tests for excluded instruments for the individual instrument (voters in the Italian-speaking language group) and its interaction with year of election are reported as F-Test (exposition) and F-Test (interaction), respectively. Standard errors robust to clustering by municipality in parentheses.
*p<.05; **p<.01; ***p<.001

Column (1) presents our baseline model, including exposure to fake news, year of election (a dummy that equals 1 for 2018 and 0 for 2013), their interaction, and the number of Italian-speaking voters by municipality. As expected, the increase in populist vote is mostly explained by the year dummy, as 2018 coincided with the electoral success of populist platforms. Nonetheless, the interaction between year and exposure, indicates a positive and significant effect of misinformation on voting in 2018.

Column (2) adds the number of landline low-latency connections to the covariates. The introduction of the variable in the model increases the model precision from a 93.4% to a 99.2% R-squared statistic. The effect of exposure to fake news, given again by the interaction term, remains significant, and surpasses the coefficients for language group and internet connectivity in magnitude. The interaction coefficient is again unaffected from the introduction of controls for income per capita in Column (3).

In Column (4) we add electorate size and abstention behaviour as further controls, increasing the adjusted R-squared to 99.7%. Although it is not significant, the negative link between abstention and populist vote is not unexpected and indicates that populist platforms may have attracted a number of individuals disenchanted by traditional politics. Most importantly, the log of the electoral size reveal itself as one of the most significant predictors in the model, indicating that the presence of a non-linear relationship between population density and populist vote. The effect of exposure on voting remains statistically significant, indicating that, for each additional like to a misinformation disseminator in a given municipality, populist scores increase by around 0.003%. While the magnitude of this effect is small, it should be noted that the interaction between exposure to fake news and the 2018 elections is still a significant predictor of populist vote. Moreover, the order of magnitude of its coefficient is still comparable to the ones of other covariates.

Our OLS results show a correlation between populist preference and exposure to misinformation. However, these estimates may still suffer from bias, as they give no information on the direction of causation. Indeed, we do not know whether the significance of the coefficient is actually showing the impact of fake news on voting or, on the contrary, that access to misinformation bubbles is linked to individual characteristics that may already determine a populist preference, either through self-selection or online recruitment. These endogeneity issues cannot be addressed by simple correlations, and motivate our instrumental variable design.

Results from our two-stage-least-squares model are shown in Table 2.7.2 Column 4, including reduced form (Column 1) and first stage (Columns 2 and 3) estimates. Here, the number of Italian-speaking voters by municipality and its interaction with year of election are instrumented to predict the exposure to fake news and its interaction with year. With the exception of the instrument and the endogenous exposure variable, the specifications in Table 2.7.2 retain the same control group used in our

Table 2.7.3: OLS estimates of the effect of misinformation on populist vote

VARIABLES	(1) OLS	(2) OLS	(3) OLS	(4) OLS
Exposed to fake news	-0.000	-0.001***	-0.001***	-0.001***
	(0.001)	(0.000)	(0.000)	(0.000)
Exposed to fake news × Year of election	0.001	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
Year of election	0.176***	0.007	0.005	0.187***
	(0.024)	(0.030)	(0.032)	(0.032)
Italian speaking voters	0.001***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Broadband connections		0.001***	0.001***	0.000***
		(0.000)	(0.000)	(0.000)
Income per capita (natural log)			0.029	-0.214
			(0.141)	(0.145)
Electorate size				0.000***
				(0.000)
Abstentions and invalid votes				-0.000**
				(0.000)
Observations	584	584	584	584
Adjusted R-squared	0.972	0.997	0.997	0.998

Notes: OLS estimates for the effect of misinformation on populist vote. Populist scores computed using the 'Anti- establishment/aggressive' text bag. Standard errors robust to clustering by municipality in parentheses.

*p<.05; **p<.01; ***p<.001

final OLS model, as in column (4) of Table 2.7.1.

Reduced form estimates are displayed in column (1) of Table 2.7.2. As discussed earlier, this specification corresponds to the difference in differences model discussed in equation (1), where the effect of the exogenous variable on the outcome is given by the interaction term between language group shares and year of election.

Diff-in-diff estimates indicate a positive and statistically significant effect of the Italian language group on voting in the 2018 election. Accordingly, after controlling for all observables, the populist content of the vote in each municipality increased by 0.00001% for each additional Italian-speaking voter, during the 2018 general elections. These figures indicate a much smaller coefficient compared to the one of fake news exposure in the final OLS specification (Table 2.7.1, column 4). Reduced form figures are proportional to the causal effect of interest, therefore the positive sign might still indicate the presence of an effect. Nonetheless, unless we assume that exposure to misinformation is completely proxied by language groups, the coefficient

Table 2.7.4: 2SLS estimates of the effect of misinformation on populist vote

	(1) DiD	(2) First Stages	(3)	(4) 2SLS
VARIABLES		Exposition	Interaction	
Italian speaking voters	0.000***	0.131**	-0.009	
	(0.000)	(0.054)	(0.023)	
Italian speaking voters × Year of election	0.000***	0.297***	0.455***	
	(0.000)	(0.039)	(0.057)	
Year of election	0.090***	-207.135***	-332.732***	0.174***
	(0.033)	(33.100)	(46.819)	(0.058)
Broadband connections	0.000***	0.021	0.013	0.000
	(0.000)	(0.079)	(0.043)	(0.000)
Electorate size	0.000***	0.022	-0.019	0.000***
	(0.000)	(0.053)	(0.051)	(0.000)
Income per capita (natural log)	-0.114	-143.422	-37.960	0.173
	(0.152)	(171.636)	(124.660)	(0.432)
Abstentions and invalid votes	-0.000**	-0.047	0.063	-0.000
	(0.000)	(0.143)	(0.158)	(0.000)
Exposed to fake news				0.002**
				(0.001)
Exposed to fake news × Year of election				-0.001*
				(0.001)
Observations	584	584	584	584
Adjusted R-squared	0.998	0.931	0.930	0.978
Partial R-squared		0.342	0.397	
F-Test		0.714	0.830	

*N*otes: IV estimates (including reduced form - DiD - and first stages) for the effect of misinformation on populist vote. Populist scores computed using the 'Anti- establishment/aggressive' text bag. F-tests for excluded instruments for the individual instrument (voters in the Italian-speaking language group) and its interaction with year of election are reported as F-Test (exposition) and F-Test (interaction), respectively. Standard errors robust to clustering by municipality in parentheses.

*p<.05; **p<.01; ***p<.001

of the interaction term is not a substitute for the effect of exposure to fake news on voting.

Turning at our instrumental variable estimates, first stage regressions predicting exposure to fake news and its interaction with year are shown in columns (2) and (3). The positive effect of the interaction between language group and year of election confirms our assumptions on differential exposure to fake news based on linguistic grouping. Results from column (3) indicate that, for each additional Italian-speaking voter, the number of likes to fake news disseminators increased by 0.45. Both models predict exposure with relative precision, boasting, overall, a

94.1% and a 93.7% adjusted R-squared. Partial R-squared statistics also indicate that, in both cases, most of the variation unexplained by the control covariates is captured by the instrument. Most importantly, both instruments pass the F-test for excluded instruments (as suggested in Bound et al., 1995), increasing our confidence in our estimates.

Our final estimates are presented in column (4), showing that fake news had no statistically significant effect on contributing to the rise of populist platforms. Indeed, the interaction coefficient between fake news exposure and year is not only lower than the OLS model in Table 2.7.1, but also not statistically different from zero. These results take any credit for the success of populist parties in 2018 away from the spread of fake news, and support the hypothesis that voters self-select into misinformation bubbles and consume fake news because of their prior preference for populist platforms, and not the other way around.

The year of election dummy is good enough for our language group instrument not to affect our estimates. We therefore conclude that fake news has had a negligible effect on voting outcomes in Trentino-Alto Adige/Südtirol over the course of the 2018 elections.

These results may seem counterintuitive, given that linguistic filter bubbles clearly had an effect in influencing voting. To understand these figures, we have to look at the local average treatment effect (LATE) of exposure to misinformation. In an IV setting, we are effectively comparing Italian-speaking voters, who were exposed to misinformation, against German-speaking voters, who were unexposed. This means, as hinted in Figure 2.3.1, that rural Italian-speaking areas where populist preferences grew the most were also the ones which were relatively less exposed to misinformation. The larger the exposure relative to the Italian linguistic group, the more the real effect of misinformation is revealed.

Moreover, we present alternative estimates using the alternative populist score based on the 'Anti-establishment/aggressive' text bag (as exposed in Section 2.5). The two tables replicate the ones presented in this section, with Table 2.7.3 showing OLS estimates and Table 2.7.4 displaying IV estimates using language group as instrument. Our results and interpretations remain mostly unchanged. Table 2.7.3 shows that a positive correlation between populist preference and exposure is still present. Again, these results are overturned with our IV estimates in Tables 2.7.4, showing that exposure had a negative and significant, but yet negligible, effect on voting when treatment is assigned through random variations in language groups.[35]

[35]First stages are reproduced for clarity purposes but, for obvious reasons, they are identical to the ones shown in Table 2.7.2.

Finally, table 2.B.1 in the Appendix shows estimates under different outcome variables, investigating whether specific parties benefited from the spread of misinformation. In this figure, we leave populist scores aside and estimate the direct effect of misinformation on the electoral outcomes of the two major anti-establishment party platforms in the region. This model retains the same specifications from Tables 2.7.2 and 2.7.4, but focuses on the empirical connections – discussed in section 3.2 – between misinformation and the platforms of LEGA and M5S, testing whether these parties have benefited from any electoral gain by the introduction of fake news. This analysis also investigates whether misinformation has lead to electoral gains for the LEGA party against the M5S, and also discusses whether the PD party has lost any votes because of fake news.

In all cases, the OLS results are overturned and, in the first column, the language group instrument even shows a -0.56 reduction for each additional like to a disseminator. Once we control for the votes that the SVP has received in both elections, we obtain null results for the effect of misinformation across all instrumental variable specifications, indicating that the spread of hoaxes has not benefited either the LEGA or the M5s. Also, these results suggest that misinformation has not affected voting preferences between the two platforms, and that, in turn, 'establishment' party platforms such as the PD have not suffered electoral losses due to misinformation.

2.8 Discussion

2.8.1 A simple model for misinformation and policy preferences

In this section we propose a simple model for misinformation and voting that could assist us in contextualising our results.

Our model draws on the median voter theorem, an approach not dissimilar to the one followed by Madestam et al. (2013). While simple, it can provide some useful conceptualisations which could help better understand the relationship between misinformation and voting preferences.

Suppose that two party platforms are contending for election, and that the position of these platforms can be drawn along a continuous policy axis representing support for the current establishment. The 'incumbent' platform sits on the left side of this axis, holding a pro-establishment stance, while the 'challenger' platforms sits on the opposite direction with an anti-establishment/populist stance.

Suppose each voter i has a prior preference g_i mirroring the same pro/anti establishment continuum characterising the two platforms. We assume these preferences are normally distributed with mean $\overline{g}$ and standard deviation σ.

Suppose also that each voter possesses an intrinsic ability a_i to ascertain the veracity of a given piece of information. This ability can follow any symmetric distribution as long as it is orthogonal to g_i. For clarity purposes, we will assume a_i follows a beta distribution with equal shape parameters, so that $A \sim Beta(\alpha, \beta)$ will be approximately normal while still being supported on the interval $[0, 1]$. a_i will then denote the percentage of fabricated facts which voter i will misreport as truthful, with $a_i = 0$ designating perfect ability to recognise a fake fact and $a_i = 1$ indicating that the voter will believe any piece of information he or she is presented with.

Vicinity determines voting preference. Given the two party platforms setting, there will be a level of support for policy g_{pop}, after which a voter will prefer to vote for the anti-establishment platform. Since g_i is normally distributed, the median voter 'bliss' preference will correspond to the mean of the population preference, meaning that, if the mean $\overline{g}$ equals g_{pop}, then the challenger will win the elections.

Now assume each voter is exposed to the same information stream, and receives a number of F facts. Each fact $j : [1, 2, ..., F]$ can either offer support for the anti-establishment movement (and be coded as $f_j = 1$) or show evidence against it (and be coded as $f_j = 0$). Information affects final support for the policy conditional on prior beliefs. We model final preference p_{i0} as a function of prior beliefs and information held by each individual $f(g_i, \sum_{j=1}^{F} f_{ji})$. We make the simple assumption that p_{i0} is directly proportional to g_i, given the share of facts in support of the anti-establishment, meaning that, if the information stream is the same for everyone, $p_{i0} = g_i(\sum_{j=1}^{F} f_j)/F$. Since a_i is bounded to the $[0, 1]$ interval, p_{i0} cannot be bigger than g_i, meaning that post-information distribution of policy preferences will have smaller mean and variance of the distribution of g_i. In any case, the normal distributional form of final preferences is unaffected, and the mean estimator will again indicate which platform the median elector will vote for. No further influx of misinformation will achieve support to the challenger platform if g_{pop} is already larger than the median value for g_i.

Now, suppose that a 'disseminator' is able to spread misinformation in support of the anti-establishment platform. We already discussed the empirical links between populism and misinformation, and we believe these factual connections can be justified by a number of factors. Indeed, it is possible for the pro-establishment platform to suffer from severe penalties from disseminating blatantly false facts in its favour: due to its position as an institutional force, trust in the institutions might stand as a distinctive signifier of the political identity and communication of the platform. As a consequence, said platform might be particularly sensitive to erosion of public trust in traditional media and institutions, rendering a disinformation strategy potentially counterproductive. The pro-establishment platform, due to its incumbent position, may also have better access to traditional media, making the use of new communication technologies the only sensible choice for the challenger. In this case, the use of 'alternative facts' as a source of political legitimacy may devalue the role of traditional media sources, which instead stand as a vital asset for a pro-establishment platform.

In any case, we assume platforms are faced with a different set of incentives that make the dissemination of misinformation a less than optimal strategy for the incumbent platform, meaning that all misinformation will be in support for the challenger. We leave any further consideration on the nature of these asymmetries to future research, resting, for our intents and purposes, on the factual evidence supporting the empirical link between anti-establishment movements and fake news, as discussed in Section 3.2.

The disseminator is presented with a simple problem, to shift the median voter to g_{pop} by introducing misinformation. The disseminator has no control over who is going to be exposed to its fake facts, but it can affect the number of anti-establishment facts in the system. The insertion of these facts will increase the total stock of facts F by K, and will affect the individual share of perceived supported facts conditional on a_i. The formula yielding $p_{i,1}$, the misinformation-adjusted final preference, is given by equation 8.1.

$$p_{i,1} = \sum_{i=1}^{N} g_i \frac{\sum_{j=1}^{F} f_j + a_i K}{N(F+K)} \tag{2.2}$$

Gullible voters – for a_i approaching 1 – will take a share a_i of the total false facts K as true, and add it to their stock of perceived anti-establishment facts. Note that the introduction of fake news might generate a penalty: as the total stock of facts in the denominator is still increased by K, less gullible voters see their final misinformation-adjusted preference $p_{i,1}$ reduced when compared to $p_{i,0}$.

For large quantities of voters, the disseminator can replace a_i in equation 8.1 with its expectation $E[a_i]$. An unbiased estimator for $E[a_i]$ would then be given by the mean of a_i, as in equation 8.2. Given that, by design, $g_i \perp a_i$ and the distribution of a_i is symmetric, $p_{i,1}$ will still be symmetrically distributed, and its mean $\overline{p}_1$ will still reveal median voter preference $\tilde{p}_1$. An implication of this model is that $p_{i,1}$

and a_i are not orthogonal, so that, this time, support for the anti-establishment will positively correlate with ability to recognise truthful facts.

$$\tilde{p}_1 \hat{=} \overline{g}\frac{\sum_{j=1}^{F} f_j + \overline{a}K}{F + K} \tag{2.3}$$

The misinformation-adjusted distribution of p_{i1} will then have mean as in the right side of equation 8.2 and standard deviation in $\sigma(\sum_{j=1}^{F} f_j + \overline{a}K)/(F + K)$. As $\overline{g}(\sum_{j=1}^{F} f_j + \overline{a})/(F + 1)$ will denote the marginal effect of a single piece of misinformation being introduced, the number of fake facts to be introduced in order to shift the median voter to the anti-establishment platform will be given, after equating the right side of equation 8.2 to g_{pop} and rearranging, by:

$$K \hat{=} \frac{\overline{g}\sum_{j=1}^{F} f_j - g_{pop}F}{g_{pop} - \overline{ga}} \tag{2.4}$$

The disseminator will then only need to supply K fake news into the system for the anti-establishment platform to win the elections. In the context of linguistic groups affecting exposure, the subscript $l : [0, L]$ could be added, with K_1 set to zero[36] for voters belonging to linguistic minorities. In that case, when only $p_{i,1}$ is observed for at least two time periods, differences in the causal effect of exposure will be given by the difference between the change in final preference before $(t - 1)$ and after (t) the introduction of misinformation across the two linguistic groups, as discussed in section (6): $\lambda = (\overline{p}_{1,t,0} - \overline{p}_{1,t-1,0}) - (\overline{p}_{1,t,L} - \overline{p}_{1,t-1,L})$.

While factual evidence still suggests that fake news have been overwhelmingly characterised by anti-establishment bias, our results, however, suggest that the electoral behaviour envisaged under this mechanism has not taken place during the Italian general elections in Trentino-Alto Adige/Südtirol.

There are at least three possible explanations for this outcome. First of all, mean ability might be too low. Should the distribution of A be skewed towards zero, or bounded by a support $[0, < 1]$, then lower values for $\overline{a}$ might lead to $\overline{p}_{i,1} < \overline{p}_{i,0}$, indicating that increases in support for the anti-establishment platform will be obtained for negative values of K. While factual evidence from the Italian elections suggests that $K > 0$, it could be argued that the anti-establishment platform possessed incomplete information on $\overline{a}$, overestimating the parameter. In these cases, a negative – or even null – effect of misinformation on final preferences is certainly possible. The preference-indifferent value for $\overline{a}^*$ can be obtained by solving for the first order condition of the partial derivative of $\overline{p}_1$ with respect to K in equation

[36] For simplicity, exposure is here set to zero for the least prevalent language group. Of course, degrees of exposure above zero are possible and have been accounted for in our regression model.

(8.2).[37] For values of mean ability lower than $\overline{a}^*$, any injection of misinformation will have detrimental effects on final preferences.

A second explanation relies on a different interpretation on the way facts are perceived. Our initial assumption on how facts are accounted when determining final political preferences might not correspond to reality: maybe facts are not accounted at all, or maybe voters only select facts which already confirm their prior beliefs. Either way, if individuals are completely indifferent to evidence opposing their prior preferences (therefore, $p_{i0} = g_i$), then misinformation is also ineffective. The idea that confirmation bias already affects the information individuals assimilate is not new to the literature, and experimental evidence for biased assimilation traces back to Lord et al. (1979).

Finally, our third explanation rests on a reasoning which transcends the model developed so far. Drawing on the echo chamber and filter bubble theories from Sunstein (2018) and Pariser (2011), it could be argued that a different mechanism should be introduced in order to explain the increased support to populist/anti-establishment platforms.

Our assumptions on information exposure may, indeed, be too unrealistic. There are no reasons to believe that every individual is exposed to the same stock of true facts as the next one. It could be argued that each voter is exposed only to a fraction of information $N < F$, and that the share of true facts in support of a given platform is functional to the set of prior preferences of the voter. If $(\sum_{j=1}^{N} f_{ji})/N = f(g_i)$, then each voter is effectively exposed to the information he or she wants to believe, and the two platforms will play an entirely different game.

While modelling such a mechanism goes well beyond the scope of our study, the implications of this reasoning are almost obvious. In our model, already, as long as $\overline{a}$ is smaller than 1, then the marginal effect of an additional true fact in support of the anti-establishment will always generate greater electoral gains than an additional fabricated fact. Platforms may then have no incentive to disseminate misinformation if the stream of true information can already be tailored to the elector. In this context, fake news would arise because of a demand for facts supporting partisan views of the world, as an entirely natural process arising from the increased fragmentation and segregation of political opinion caused by the personalisation of social media filter bubbles. In this case, the relationship between populist platform and the disseminator needs not to be symbiotic, and could as well be parasitic.

[37] Meaning that solving $(K + F)^2 = -\overline{g}(-\overline{a}F + \sum_{j=1}^{F} f_j)$ for $\overline{a}$ will yield the average ability level needed for fake news to have no effect. Preference-indifferent values for the share of true facts supporting the anti-establishment, holding ability and the other determinants of preference as fixed, can be similarly computed by solving this equation for $F_P = \sum_{j=1}^{F} f_j$.

The disseminator, while still creating false facts in support of the anti-establishment platform, may be driven entirely by economic motives – such as obtaining advertising revenue or increasing the value of his/her domain or fan page – with little to no voting externalities arising from his/her activities.

2.9 Conclusions

The measurement of the influence of fake news on electoral behaviour has, so far, escaped empirical assessment. With this study, we set to fill this void, identifying both the presence of this influence and, in that case, its magnitude.

In order to account for the inevitable reverse causality issues between voting preferences and exposure to misinformation, we proceed with a quasi-experimental approach. Gathering municipality-level data on electoral outcomes, demographics and social media usage, we exploited the autonomous provinces of Trentino and South Tyrol in the 2018 Italian elections to randomise exposure to fake news.

In our contribution to the literature, we believe that our work sheds more light on the relationship between the spread of fake news and populist echo chambers. Mining the text from social media posts of parties and their leaders during their electoral campaigns, we produced an indicator for populist content, allowing us to study populism as a phenomenon that eschews the political dimensions of left and right.

We showed that misinformation had a negligible and non-significant effect on the populist vote in Trentino and South Tyrol during the Italian 2018 general elections. Our results indicate that exposure to fake news is entirely dictated by self-selection in misinformation bubbles, meaning that the causal channel between voting and fake news goes on a single direction, with individuals being exposed to misinformation because of their political presences. Also, these findings suggest that, unless the two effects cancelled each other, there is neither an extensive nor intensive margin to the effect of fake news: indeed, the non-significance of the exposure coefficient indicates that individuals who followed these disseminators on social media did not modify their final voting preference more than their unexposed peers.

In a simple two party model, if all voters were exposed to the same pieces of information, and if each voter had varying ability in recognising true facts from false ones, then misinformation could play a role in shifting median voter preferences. However, when social media filter bubbles are able to produce personalised information feeds, it is difficult to believe any given voter is exposed to the same information as the next one. Our results provide empirical evidence indicating either that the average ability to recognise true from false facts is underestimated, or that preferences are not dictated by the absolute proportion of facts perceived as true. This interpretation would support the hypothesis that the social media information bubbles can already shift the perceived proportion of supporting facts independently of the presence of misinformation.

This does not mean that the fake news is less problematic, but only that the causes of the populist shift in voting have to be found elsewhere, and that misinformation still thrives within these filter bubbles. The persistence of very similar differences in voting behaviour conditional on linguistic grouping and broadband penetration indicates that echo chambers most likely had a role in determining final preferences. In this sense, fake news would rather stand as the embodiment of shared narrations within groups of voters which are further reinforced by confirmation bias and the increasing personalisation of social media echo chambers. It may be possible that, once entered a misinformation bubble, partisan opinions are being reinforced by the presence of fake news, which might then ensure continued support for partisan beliefs as reputable sources of information are progressively removed and discredited in favour of 'alternative facts'. In the presence of personalised filter bubbles, preference dictates facts and not the other way around; social media plays a role in as much as it provides individuals with the information they want to believe.

Our final notes address the validity of our results and suggest future pathways for research, in the hope that our work spurs further empirical and theoretical contributions on the role misinformation has on voting.

The exploitation of a natural experiment Trentino-Alto Adige/Südtirol imposes a constraint on the external validity of our results, as the relationship between misinformation and voting might differ in other regional and national contexts. A similar methodology to ours could be applied to different contexts and with different units of observation, as language groups and broadband penetration have proven as good predictors for access to fake news. Survey data might also shed more light on individual preferences and social media behaviours.

Our estimates are also robust to a single definition of populism. Due to the liquid nature of this phenomenon, it is possible for our figures to change under different conceptions of populism, such as its more ethno-nationalistic departures. Our methodology could then be replicated using different text-bags as shown in Appendix B. In this sense, further work is certainly needed to address competition between party platforms sharing contiguous filter bubbles and investigate whether misinformation can affect voting preferences between populist platforms, favouring certain versions of populism over other ones.

Finally, as argued earlier, we believe that future research should focus on the relationship between misinformation and echo chambers. As filter bubble access is already determined by prior preferences and individual characteristics, the attention of researchers and policy makers should rather focus on the socio-economic determinants for access into misinformation bubbles. If changes in voting behaviour are unaffected by increases in exposure to fake news, then, we believe, the personalisation of information streams in social media is ultimately the most important and socially-poignant factor that should be addressed when studying misinformation and the rise of populism.

Acknowledgements

The authors would like to thank Thomas Bassetti, Markus Cappello, Paloma Díaz Topete, Federico Maria Ferrara, Stefano Gagliarducci, Matilde Giaccherini, Stephen Hansen, Joanna Kopinska, Alessandro Martinello, Tim Munday, Francesco Pierri, Alessandro Pizzigolotto, Alberto Franco Pozzolo, Chris Redl, Giancarlo Spagnolo, Arthur Turrell, Davide Viviano, Daniela Vuri, Anzelika Zaiceva for their invaluable suggestions and support. We are grateful to participants to the Bank of England-King's College conference 'Modelling with Big Data and Machine Learning: Interpretability and Model Uncertainty', the PhD Seminar in Economics and Finance at the University of Rome Tor Vergata, the FBK IRVAPP internal seminar, the University of Modena and Reggio Emilia seminar for their comments. All remaining errors are ours.

Supplementary material

Datasets and codes used in our estimation can be found in the online data archive, available at the url address: https://sites.google.com/site/michelecantarella1992/data-archive-by-paper.

Bibliography

Abel, A., Vettori, C., and Forer, D. (2012). Learning the neighbour's language: The many challenges in achieving a real multilingual society. the case of second language acquisition in the minority-majority context of south tyrol. In for Minority Issues & European Academy Bolzano/Bozen, E. C., editor, *European Yearbook of Minority Issues, vol. 9*, pages 271–304. Brill Academic Publishers, Leiden.

Aitchison, J. (1982). The statistical analysis of compositional data. *Journal of the Royal Statistical Society. Series B (Methodological)*, 44(2):139–177.

Albertazzi, D. and McDonnell, D. (2008). *Twenty-First Century Populism. The Spectre of Western European Democracy*. Palgrave Macmillan.

Allcott, H. and Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, 31(2):211–236.

Allcott, H., Gentzkow, M., and Yu, C. (2019). Trends in the diffusion of misinformation on social media. *Research & Politics*, 6(2).

Astat (2011). Censimento della popolazione 2011: Determinazione della consistenza dei tre gruppi linguistici della provincia autonoma di bolzano-alto adige. Astat Informazioni.

Astat (2017). Annuario statistico della provincia di bolzano. Provincia Autonoma di Bolzano/Alto Adige, Istituto provinciale di statistica - ASTAT.

Autorità per le Garanzie nelle Comunicazioni (2018). News vs fake nel sistema dell'informazione, interim report nell'ambito dell'indagine conoscitiva di cui alla delibera n. 309/16/cons. Interim report.

Azzimonti, M. and Fernandes, M. (2018). Social media networks, fake news, and polarization. Working Paper 24462, National Bureau of Economic Research.

Barrera Rodriguez, O., Guriev, S. M., Henry, E., and Zhuravskaya, E. (2017). Facts, alternative facts, and fact checking in times of post-truth politics. Cepr discussion paper no. dp12220.

Bischof, D. and Senninger, R. (2018). Simple politics for the people? complexity in campaign messages and political knowledge. *European Journal of Political Research*, 57(2):473–495.

Bound, J., Jaeger, D. A., and Baker, R. M. (1995). Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. *Journal of the American Statistical Association*, 90(430):443–450.

Boutyline, A. and Willer, R. (2016). The social structure of political echo chambers: Variation in ideological homophily in online networks. *Political Psychology*, 38(3):551–569.

Bovet, A. and Makse, H. A. (2019). Influence of fake news in twitter during the 2016 US presidential election. *Nature Communications*, 10(1).

Caiani, M. and Graziani, P. R. (2016). Varieties of populism: insights from the italian case. *Italian Political Science Review*, 46(2):243–267.

Campante, F., Durante, R., and Sobbrio, F. (2017). Politics 2.0: The Multifaceted Effect of Broadband Internet on Political Participation. *Journal of the European Economic Association*, 16(4):1094–1136.

de Vreese, C. H., Esser, F., Aalberg, T., Reinemann, C., and Stanyer, J. (2018). Populism as an expression of political communication content and style: A new perspective. *International Journal of Press/Politics*, 23(4):423–438.

Del Vicario, M., Vivaldo, G., Bessi, A., Zollo, F., Scala, A., Caldarelli, G., and Quattrociocchi, W. (2016). Echo chambers: Emotional contagion and group polarization on facebook. *Scientific reports*, 6:37825.

Durante, R., Pinotti, P., and Tesei, A. (2019). The political legacy of entertainment TV. *American Economic Review*, forthcoming.

Ebner, C. V. (2016). The long way to bilingualism: the peculiar case of multilingual south tyrol. *International Journal for 21st Century Education*, 3(2):25.

Galbraith, J. K. and Hale, J. T. (2008). State income inequality and presidential election turnout and outcomes. *Social Science Quarterly*, 89(4):887–901.

Giglietto, F., Iannelli, L., Rossi, L., Valeriani, A., Righetti, N., Carabini, F., Marino, G., Usai, S., and Zurovac, E. (2018). Mapping italian news media political coverage in the lead-up of 2018 general election. *SSRN Electronic Journal*.

Guess, A., Nyhan, B., and Reifler, J. (2018). Selective exposure to misinformation: Evidence from the consumption of fake news during the 2016 us presidential campaign. *European Research Council*.

Guiso, L., Herrera, H., Morelli, M., and Sonno, T. (2017). Populism: Demand and supply. *Center for Economic Policy Research Discussion Paper*, 11871.

Gunther, R., Beck, P. A., and Nisbet, E. C. (2019). “Fake news” and the defection of 2012 obama voters in the 2016 presidential election. *Electoral Studies*.

Hersh, E. D. and Nall, C. (2015). The primacy of race in the geography of income-based voting: New evidence from public voting records. *American Journal of Political Science*, 60(2):289–303.

Kaltwasser, C. R., Taggart, P. A., Espejo, P. O., and Ostiguy, P. (2017). *The Oxford Handbook of Populism.* Oxford University Press.

Kennedy, P. J. and Prat, A. (2019). Where do people get their news? *Economic Policy*, 34(97):5–47.

Kumar, A. and Sebastian, T. M. (2012). Sentiment analysis on twitter. *International Journal of Computer Science Issues*, 9(4):372–378.

Lewis-Beck, M. S. and Nadeau, R. (2011). Economic voting theory: Testing new dimensions. *Electoral Studies*, 30(2):288–294.

Liberini, F., Redoano, M., Russo, A., Cuevas, A., and Cuevas, R. (2018). Politics in the Facebook Era Evidence from the 2016 US Presidential Elections. The Warwick Economics Research Paper Series (TWERPS) 1181, University of Warwick, Department of Economics.

Lord, C. G., Ross, L., and Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology*, 37(11):2098–2109.

Madestam, A., Shoag, D., Veuger, S., and Yanagizawa-Drott, D. (2013). Do political protests matter? evidence from the tea party movement. *The Quarterly Journal of Economics*, 128(4):1633–1685.

Martin, G. J. and Yurukoglu, A. (2017). Bias in cable news: Persuasion and polarization. *American Economic Review*, 107(9):2565–2599.

Nyhan, B. and Reifler, J. (2010). When corrections fail: The persistence of political misperceptions. *Political Behavior*, 32(2):303–330.

Pariser, E. (2011). *The Filter Bubble.* Penguin Books Ltd.

Pauwels, T. (2011). Measuring populism: A quantitative text analysis of party literature in belgium. *Journal of Elections, Public Opinion and Parties*, 21(1):97–119.

Polk, J., Rovny, J., Bakker, R., Edwards, E., Hooghe, L., Jolly, S., Koedam, J., Kostelka, F., Marks, G., Schumacher, G., et al. (2017). Explaining the salience of anti-elitism and reducing political corruption for political parties in europe with the 2014 chapel hill expert survey data. *Research & Politics*, 4(1):2053168016686915.

Rooduijn, M. (2014). The nucleus of populism: In search of the lowest common denominator. *Government and Opposition*, 49(4):572–598.

Rooduijn, M. and Pauwels, T. (2011). Measuring populism: Comparing two methods of content analysis. *West European Politics*, 34(6):1272–1283.

Roozenbeek, J. and van der Linden, S. (2018). The fake news game: actively inoculating against the risk of misinformation. *Journal of Risk Research*, pages 1–11.

Schkade, D., Sunstein, C. R., and Hastie, R. (2007). What happened on deliberation day. *California Law Review*, 95:915.

Shin, J., Jian, L., Driscoll, K., and Bar, F. (2018). The diffusion of misinformation on social media: Temporal pattern, message, and source. *Computers in Human Behavior*, 83:278–287.

Shu, K., Sliva, A., Wang, S., Tang, J., and Liu, H. (2017). Fake news detection on social media. *ACM SIGKDD Explorations Newsletter*, 19(1):22–36.

Sunstein, C. R. (2002). *Republic.com*. Princeton University Press.

Sunstein, C. R. (2018). *#Republic*. Princeton University Press.

Taggart, P. A. (2000). *Populism*. Open University Press.

Törnberg, P. (2018). Echo chambers and viral misinformation: Modeling fake news as complex contagion. *PLOS ONE*, 13(9):e0203958.

Appendix

2.A Measuring exposure to fake news

The problem of measuring exposure to fake news is far from a trivial one. Other studies have relied on different methods, such as individual level survey data or the total social media shares for each piece of misinformation. The structure of our data, however, imposes to obtain a proxy for exposure that can be broken down at municipality level and that is available for all units of analysis. There is no way to reconstruct how each fake news has been shared across social media, and who has been exposed to it. It is, however, possible to construct an approximate measure of the following each Facebook page which disseminate fake news has in a given municipality.

This information can only be acquired through a single source: the Facebook Audience Insight Tools (see www.facebook.com/ads/manager/audiences). This tool, generally intended for advertising purposes, allows to access information on active Facebook users, who can be filtered by their place of residence, age, gender, language and many other interests. The tool yields demographic (broken down by age, gender, relationship status, education, and employment) and online activity (page likes and social media use) information on all active individuals in the targeted audience. While it is not possible to filter an audience by their appreciation for a particular Facebook page, it is possible to target an audience through 'interests', and then be presented with a list of pages which correlate with these interests amongst the selected audience, where the number of audience-specific likes for each of these pages is also displayed.

We found a number of interest to be correlated with appreciation of fake news disseminators in Italy. Keyworks such as 'Immigrazione', 'Stato sovrano', 'Scienza di confine', 'Illuminati', 'Medicina alternativa', 'Casta', 'Teoria del complotto', 'Notizie Incredibili', 'Popolo', 'News24', 'Sovranità', 'La casta', 'Massoneria', and 'Notizie.it', all returned information on many disseminators included in 'Black lists' compiled by debunking websites.

While the demographic information of the audience tool is biased towards Facebook users, we are confident that page likes which is information obtained from these tools presents an unbiased figure of exposure to misinformation, as it is expected for the the utmost majority of fake news to travel through social media.

We then used the Facebook Audience Insight Tools to extract a sample of pages likes on Facebook, using the 'News24' interest. We collected this information for each Municipality in the Trentino-Alto Adige/Südtirol region, for all users aged 24+, in order to ensure that these individuals had the opportunity to vote at both 2013 and 2018 elections.

The collection process required a few precautions to be taken. First, as the tool will not display figure on a location if the number of users in that area is below a certain (unspecified) threshold, we collected all exposure information for pairs of municipalities $Trento + Municipality_i$, using Trento (capital of the Trentino-Alto Adige/Südtirol) as the baseline to be subtracted at the end of the collection. This method ensured that: (1) information could be collected for all locations, and (2) that the tool always returned the same list of pages.

A second issue is generated by the 'interest' targeting options, as the tool only returns information on the number of users belonging to a given 'interest' audience who also liked that page. This means that the number of observed likes through interest targeting will be a fraction – approaching unity, for certain interests – of the total audience, and that exposure figures will be under-estimated. To adjust for this issue, we make two further assumptions: (1) that the total audience of these pages is composed solely of people residing in Italy and (2) that this fraction is the same for both Trentino Alto Adige/Südtirol and Italy. More formally, we assume that $\mathit{obslikes}_{IT}/\mathit{totlikes}_{IT} = \mathit{obslikes}_{TA}/\mathit{totlikes}_{TA}$, where the subscript IT and refer to users located in Italy, TA to those in Trentino Alto Adige/Südtirol. This means that the ratio between observed likes and total likes is equal across the country. Under this assumption, we divide for each page the total number of likes by the audience figures obtained in the Italian territory. We then adjust the shares collected in Trentino Alto Adige/Südtirol by this scalar.

The inability to target specific pages through the audience tool implies that not all fake news disseminators will be captured by this methodology, as some are – perhaps purposely – unreachable through interest targeting.

To account for this issue, we develop a simple yet powerful predictive model that we use to impute the social media following for each of these missing pages, following from the intuition that the number of individuals liking a page in a given municipality will be proportional to the total Facebook likes of said page, holding specific effects from each location, such as its size and the language group, as fixed. If we assume that all page likes follow a similar functional form, we can then estimate the parameters from this function for observed pages, and then use the model to impute the following of unobserved pages. In the model:

$$y_i = \alpha + totlikes_i\beta + municip_i'\gamma + totlikes_i{\times}municip_i'\delta + collegio_i'\zeta + e_i$$

y_i indicates the total number of Facebook likes each observed page has in each municipality, for $i = \{1, 2, ..., M{\times}P\}$, where M indicates the total number of municipalities, and P the total number of observed pages. $Totlikes$, instead, indicates the total number of likes each of these pages has on Facebook, while $Municip$ and $Collegio$ are column vectors of dummies for municipality and constituency, respectively. The pages used in our estimation are presented in Table 2.A.1, where observed and modelled pages are labelled as 'donor' and 'recipient' respectively.[38]

The model allows to predict the number of likes for an observed page in a given location as a function of the total number of likes of said page, allowing for different slopes and intercepts between municipalities.[39] With 3504 observations and a multiple R-squared of 0.9237, the model achieves a satisfying fit and its parameters are extracted and used to impute, for each municipality, the number of people liking the other pages appearing in 'black lists' using the total number of likes of these pages only. After these values have been imputed, they are summed with the observed values in each location to obtain an estimate of the total number of social media likes to fake news disseminators in each municipality.

While it is reasonable to argue that the fake news phenomenon has risen to mainstream attention only in the eve of the 2017 referendum and 2018 elections, it should be noted that few of these pages already existed before 2013. To account for this issue, we adopt a 'conservative' methodology and, for 2013, exposure to fake news is still computed using pages which existed in the date of the 2013 elections, keeping the total number of likes for each of these pages unaffected. While the results shown in Section 7 make use of these adjusted figures, our estimates are nearly unaffected by the use of a less conservative indicator where exposure to fake news is set at zero for 2013.

As a final note, it should be remarked that this imputation method is not stochas-

[38] A very similar imputation model was also used to correct figures for the 'Chedonna.it' page. As the exposure figures are rounded by the nearest hundred after a page reaches 1,000 likes in a given location, figures for this have been adjusted accordingly using the variation in the other observed pages.

[39] The coefficient vector γ will produce fixed effects specific to the municipality, while the vector of the interaction coefficients δ will allow for different slopes.

tic, but deterministic. However, as the final variable will consist in the sum of these estimates, we are generally uninterested in correctly simulating within-municipality variation, while we feel that between-municipality variation is a second-order problem considering the good fit of the model. Also, we remark that the purpose of this imputation is to improve the figures of exposure so that the effect of each additional like to a disinformation disseminator can be quantified with more precision. In any case, even in the presence of over or under-estimation of our exposure figures, the sign, and the statistical presence of an effect of fake news on voting should remain unchanged, as the imputation mostly scales the total number of likes in a municipality upwards. Indeed, we constructed another alternative variable for exposure to misinformation using only information from the pages we managed to observe with the Facebook Audience Insight Tool, and our final figures are again unaffected by these changes.

Last but not least, our results are only partially reproducible, as the Facebook API suffers from severe transparency issues that affect the possibility to perfectly replicate this method. It is well possible, then, that exposure figures may vary slightly when collected in a different moment in time. Also, some pages may have changed their name or have been unpublished in the meanwhile, complicating, again, the reproduction of our results. Indeed, many of the pages reported in Table 2.A.1 were closed by Facebook in May 2019 following a flagging campaing from the non-governmental organisation AVAAZ.[40] In any case, the original data set used in our calculations is made available in the online data archive.

[40] Ibid. 17

Table 2.A.1: Fake news disseminators

Page	likes	followers	founded in	type	Page	likes	followers	founded in	type
Adesso Basta	508971	522664	06-Apr-16	donor	Notizie live	11309	11246	03-Nov-14	recipient
Chedonna.it	1996969	1905609	03-Dec-13	donor	Questa è l'Italia	3982	3987	26-Jan-12	recipient
Citazioni che ispirano	322977	324125	02-Nov-13	donor	Questa è l'Italia di oggi	19674	20961	01-Aug-16	recipient
DiariodelWeb.it	506535	512444	19-Dec-14	donor	Rimani informato	44758	53156	03-Jul-17	recipient
Dimissioni e tutti a casa	784505	749631	19-Aug-13	donor	Rothschild: la Bestia che domina il mondo	19172	19163	21-Jul-11	recipient
Giornale Interattivo	525468	525468	03-Mar-14	donor	Scienza di confine	81788	81060	11-Feb-12	recipient
Italia Patria Mia	375094	369117	29-May-15	donor	Segreto di Stato	6918	6934	07-Dec-16	recipient
Mister Link	436253	404946	11-Apr-11	donor	Se ti fai un'altra birra la prima non s'incazza	205564	204537	11-Nov-12	recipient
Quello che i TG non dicono	369617	377979	23-Oct-16	donor	SocialTv Network (websocialtv)	99181	107729	05-Mar-16	recipient
Silenzi e falsità della stampa italiana	853552	861987	06-Nov-13	donor	Sono senza parole	938033	905287	29-Jun-14	recipient
24H Italia News	171561	173703	07-Dec-16	recipient	Sovranità Monetaria & Debito Pubblico	2128	2158	16-Nov-11	recipient
Aprite gli occhi	148648	149245	08-Feb-11	recipient	Stop alle scie chimiche	32729	32471	19-Oct-09	recipient
Autismo e Vaccini	22391	23016	22-Jul-13	recipient	Ultimissime	2889	2876	29-Apr-14	recipient
Avvistamenti di Creature Mitologiche	181711	179820	20-Mar-12	recipient	Un male tutto italiano, questo regime è una vera piaga sociale	10912	10686	10-Jun-11	recipient
Banda Bassotti	32666	32776	29-Jan-13	recipient	Vaccini Basta	27314	28314	21-Jan-12	recipient
Catena Umana	34464	34372	16-May-14	recipient	Zona grigia	66149	65974	29-Apr-16	recipient
Come i treni a vapore	132131	136164	07-Feb-14	recipient	Stop Invasione	18197	18718	23-Oct-15	recipient
Contro i poteri forti	11441	11591	23-May-15	recipient	Un caffe al giorno	235717	236577	21-Jul-16	recipient
Cose che nessuno ti dirà di nocensura.com	1666222	1612016	16-Aug-10	recipient	Non cielo dicono	153986	160074	25-Jun-16	recipient
CrimeNews (mafiacapitale.info) (Catena Umana 3)	11804	11813	13-Oct-15	recipient	Roby	1425198	1414607	06-Apr-11	recipient
CSSC - Cieli Senza Scie Chimiche	10531	10619	08-Dec-11	recipient	Sardegna Today	5428	5550	04-Sep-18	recipient
Eco(R)esistenza	85473	84418	02-Jun-11	recipient	Dangerous News	5385	5275	07-Jul-13	recipient
Informati (informatitalia)	211461	207818	07-Apr-13	recipient	Il Matto Quotidiano	12313	11952	02-Apr-15	recipient
IO VI Spengo	481267	550151	12-Apr-15	recipient	Imola Oggi	9188	9403	16-Sep-18	recipient
Italia Malata	96781	96935	02-Sep-15	recipient	Italiani Uniti per la Patria	96382	98380	28-Apr-16	recipient
Killuminati Soldiers	2406	2434	23-Jan-16	recipient	Scenari Economici	39224	42555	06-Mar-13	recipient
John Koenig Mb Koenig	2585	2670	16-Feb-14	recipient	Informare X Resistere	1086691	1020283	16-Jul-09	recipient
L'alimentazione e gli illuminati	10032	10095	10-Aug-11	recipient	Piovegovernoladro	55839	54235	19-Dec-13	recipient
L'antipolitica	14517	14519	29-Apr-12	recipient	Jeda News	372364	357423	02-Jan-12	recipient
La Vera Italia	3742	3722	16-May-15	recipient	Breaknotizie	81656	80762	10-Sep-12	recipient
La Verità ci Rende Liberi	68857	72846	20-Jun-18	recipient	Controinformo	27095	26917	13-Jan-16	recipient
Le notizie che non ti aspetti	133697	133185	26-Jun-14	recipient	DirettaNews24	130449	130453	09-Oct-16	recipient
Mafia Capitales (whatsappecultura)	99487	98732	30-Dec-08	recipient	FascinAzione	1866	1910	03-Apr-16	recipient
Movimento anti NWO	12324	12274	20-Feb-12	recipient	TG 5 Stelle	16934	16216	26-Jun-13	recipient
Neovitruvian	6742	6763	09-Dec-10	recipient	CVDiariodelPollineo	1569	1572	14-Oct-16	recipient
NO alla dittatoriale Unione Europea e al Nuovo Ordine Mondiale	2629	2714	11-Nov-11	recipient	Notizie in Movimento	116951	118686	29-Sep-14	recipient
Notiziario (notiziario.face)	22882	22943	15-Nov-12	recipient					

*N*otes: Donor and recipient pages used in the estimation of total page likes by municipality. Previous page names are shown in parentheses.

2.B Electoral gains for specific parties

Table 2.B.1: OLS & 2SLS estimates of the effect of misinformation on voting

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Votes to LEGA & M5S				Votes to LEGA		Votes to PD	
VARIABLES	OLS	2SLS	OLS	2SLS	OLS	DiD	OLS	2SLS
Exposed to fake news	-0.098	1.311***	-0.692***	-0.092	0.108*	0.498	0.402***	2.536
	(0.201)	(0.483)	(0.071)	(0.859)	(0.057)	(0.413)	(0.088)	(2.361)
Exposed to fake news × Year of election	0.361**	-0.599**	0.700***	0.336	-0.042	-0.368	-0.331***	-1.773
	(0.163)	(0.281)	(0.059)	(0.542)	(0.049)	(0.370)	(0.056)	(1.584)
Year of election	256.464***	123.155***	138.995***	143.795***	18.794*	-22.750	14.033	-30.587
	(37.831)	(30.074)	(15.169)	(19.820)	(9.756)	(56.376)	(11.913)	(46.276)
Votes to SVP			-0.701***	-0.469*	0.220***	0.460*	-0.309***	0.360
			(0.033)	(0.263)	(0.034)	(0.262)	(0.042)	(0.596)
votes LEGA & M5S					0.797***	1.050***		
					(0.063)	(0.335)		
Broadband connections	0.167***	0.006	-0.037*	-0.023	-0.034**	-0.022	-0.051**	-0.020
	(0.051)	(0.099)	(0.022)	(0.020)	(0.014)	(0.017)	(0.023)	(0.091)
Electorate size	0.118*	-0.057	0.384***	0.260	-0.130***	-0.277*	0.264***	-0.217
	(0.060)	(0.044)	(0.030)	(0.169)	(0.020)	(0.164)	(0.038)	(0.405)
Income per capita (natural log)	-376.316***	304.781	94.346**	168.232	-25.723	-27.482	74.868	305.217
	(139.694)	(219.144)	(39.916)	(119.900)	(32.217)	(35.055)	(50.886)	(385.097)
Abstentions and invalid votes	-0.351***	0.202	-0.072	-0.053	0.082***	0.157	-0.250***	0.202
	(0.135)	(0.159)	(0.056)	(0.148)	(0.021)	(0.104)	(0.043)	(0.465)
Observations	584	584	584	584	584	584	584	584
Adjusted R-squared	0.960	0.898	0.992	0.982	0.989	0.981	0.992	0.803
F-test (Exposition)		45.03		71.82		45.91		71.82
F-test (Interaction)		59.19		93.37		53.50		93.37

*N*otes: OLS and IV estimates for the effect of misinformation on vote to specific parties by municipality: Lega and M5S, Lega only, and Partito Democratico. Standard errors robust to clustering by municipality in parentheses.

*p<.05; **p<.01; ***p<.001

2.C correlation between text-based populist scores and CHES data

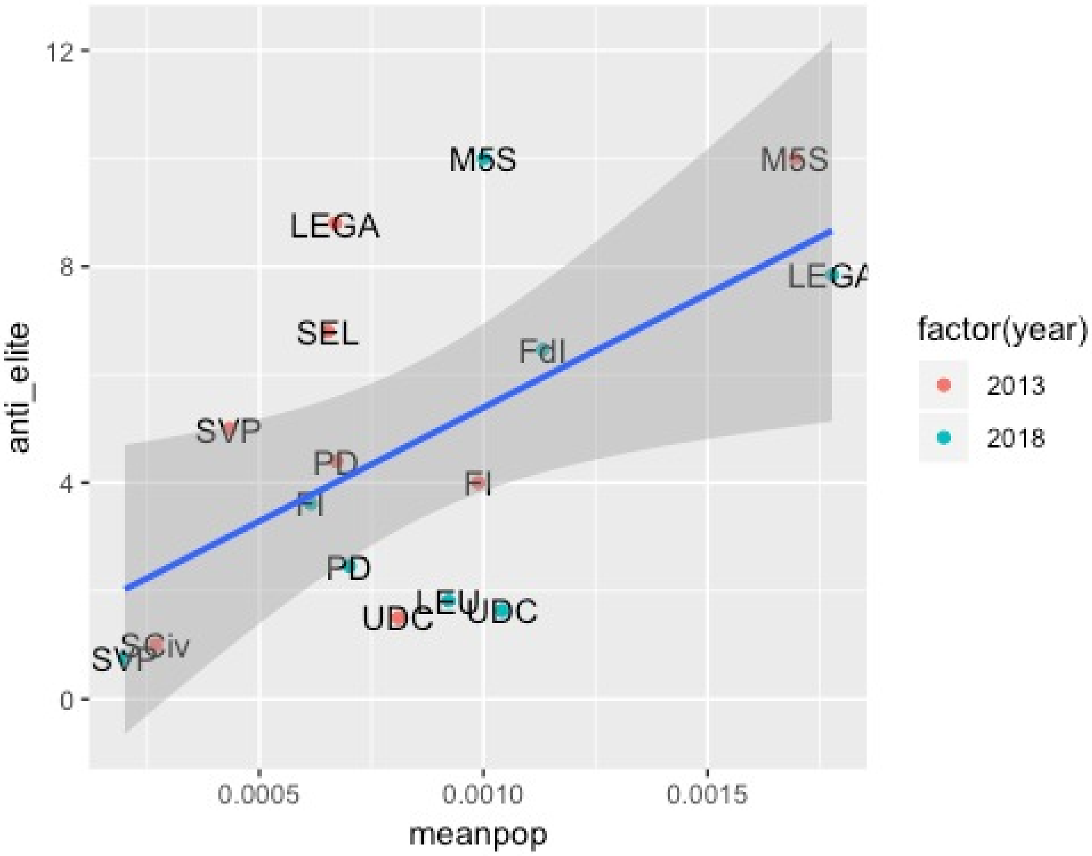


Figure 2.C.1: Text analysis scores of social media posts from parties and their leaders during the 2013 and 2018 elections campaigns (x-axis) and their relationship with the scores on the variable 'People vs the Elites' from the Chapel Hill Expert Survey for year 2014 (y-axis). Higher values on the y-axis correspond to higher salience of anti-establishment and anti-elite rhetoric on a scale from 0 to 10. Scores are computed using the 'Anti-establishment/ aggressive' text bag.

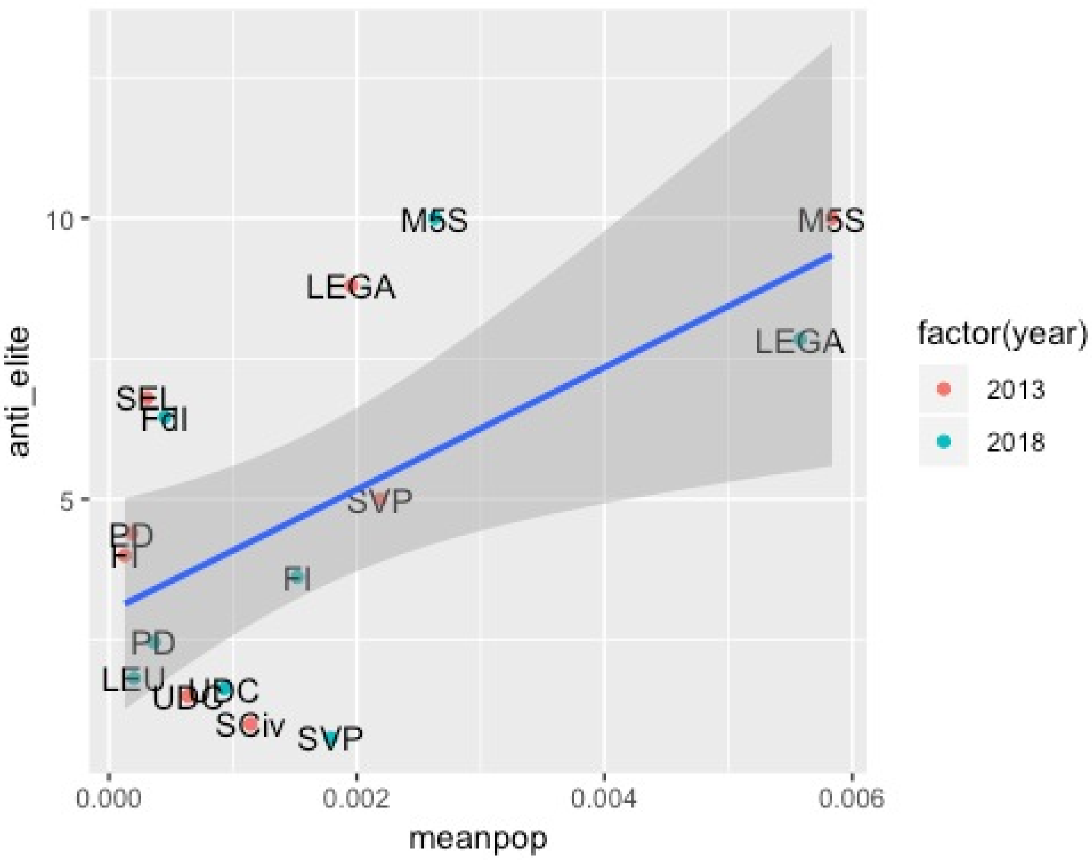


Figure 2.C.2: Text analysis scores of social media posts from parties and their leaders during the 2013 and 2018 elections campaigns (x-axis) and their relationship with the scores on the variable 'People vs the Elites' from the Chapel Hill Expert Survey for year 2014 (y-axis). Higher values on the y-axis correspond to higher salience of anti-establishment and anti-elite rhetoric on a scale from 0 to 10. Scores are computed using the 'Assertive' text bag.

Chapter 3

#Portichiusi: the human costs of migrant deterrence in the Mediterranean

Abstract

Using daily data on forced migration from the IOM, I compare trends in flows and mortality across three major migration routes in the Mediterranean, analysing the effects of the introduction of rescue-deterrence policies in Italy. Controlling for exogenous shocks which affect push and pull factors in mobility, along with sea state conditions and route-day fixed effects, I find that the reduction in refugee migration flows in the Central Mediterranean has been modest, at best. At the same time, these policies have generated a permanent increase in daily mortality rates in the Central Mediterranean, having grown by more than 4 deaths per day. Finally, I investigate whether variations in mortality are sufficient to offset migration flows. Increases in mortality rates, however, are only accompanied by a short-term negative displacement effect, as migration attempts are delayed by increases in absolute mortality, rather than being prevented.

Keywords: *costs of migration, forced migration, EU refugee crisis, deterrence policies*

JEL codes: F22, J15, J61, J68

3.1 Introduction

Forced migration flows to EU member states through the Mediterranean Sea reached their peak in 2015 and have continued to persist in the last several years. Turmoil in the Middle East and Africa has generated an unprecedented displacement of individuals toward the EU, many of whom have attempted to reach the shores of Europe by crossing the Mediterranean sea.

This migration route is far from safe, and the death tool has been steadily increasing over the course of time. Indeed, the evolution of the refugee crisis into a humanitarian disaster has dominated political discourse and created friction between EU member states. In response to the crisis, patrol operations in the Central Mediterranean (such as operations *Mare Nostrum*, *Triton* and *Themis*) have intensified, but the effort has been far from concerted, with these operations relying, mostly, on voluntary contributions from EU member states, and the focus of these operations gradually shifting from search and rescue (S&R, onwards) to border patrol. Furthermore, uncertainties in the implementation and interpretation of the Common European Asylum System (Dustmann et al., 2017a), including major weaknesses in the asylum procedures envisioned under the Dublin regulation, where the duty of examination of an asylum application falls to the country of first arrival, have only led to further discord and discontent.

In response to these developments, some private initiatives have arisen and NGOs, such as Open Arms, Sea-Watch, Mediterranea, Mèdecins Sans Frontières and Save the Children, amongst many others, have been conducting S&R operations in the Central (and in rarer instances, Eastern) Mediterranean area.[1] These activities have not gone unnoticed.

While some regulation efforts have been made by national authorities, their presence has been, more or less, tolerated. In Italy, the first of these attempts, the so-called "Minniti Code" was enacted during the former Gentiloni government in July 2017, and required NGOs conducting S&R operations in the Central Mediterranean to comply to a code of conduct. The adoption of said code was not met without opposition from NGOs.[2] However, the code still allowed NGOs to perform their operations, even if within specified boundaries.

[1] European Union Agency for Fundamental Rights (2019); 2019 update - NGO ships involved in search and rescue in the Mediterranean and criminal investigations; available at: https://fra.europa.eu/en/publication/2019/2019-update-ngos-sar-activities; last accessed: August 27th, 2019.

[2] La Repubblica (2017); Migranti, codice Ong: Msf non firma. Minniti: "Chi non sottoscrive regolamento è fuori"; available at: https://www.repubblica.it/cronaca/2017/07/31/news/migranti_msf_non_firma_codice_ong-172058967/; last accessed: August 27th, 2019.

This code was not the only measure enacted during the Gentiloni government, as this initiative was accompanied by a much larger change in migrant policy consisting of strengthening the role of the Libyan coast guard in border patrol operations and in concluding agreements with Libyan tribal chiefs to break their links with migrant smugglers.[3] Critics of this "desert diplomacy" approach argued that these measures only led to the further exploitation of refugees and severe violations of human rights.[4]

While the NGOs have tried to make up for these national and intergovernmental deficiencies with their own means, the Mediterranean Member States have found themselves facing, mostly alone, a problem to which different lines of thought have given conflicting answers. On the other hand, we have the idealistic approach, in defense of a moral obligation to rescue; on the other, an approach, which, for the lack of a better term, could be defined as realist, supporting thesis that the reduction of rescue efforts would reduce migration attempts and, consequently, even deaths.

Indeed, in the midst of this debate, non-withstanding the policies of the Gentiloni government, populist voices have campaigned in favour of curbing all rescue operations in the Mediterranean, arguing that the presence of S&R operation, even if regulated, encourages migration attempts, implying that the interruption of these activities would deter migrants from crossing the sea.[5] As the hostility of natives towards refugees grew, some of those voices provided an electoral platform for their sentiments.[6] In Italy, the electoral success of the Lega Nord under the new guise of a "national-populist" party platform, and the nomination of its leader Matteo Salvini as the Minister of Interior of the M5S-Lega coalition government, has marked the beginning of a new era in migrant policy for Italy, an era that is in fulfillment

[3]The Guardian (2017); Italian minister defends methods that led to 87% drop in migrants from Libya; https://www.theguardian.com/world/2017/sep/07/italian-minister-migrants-libya-marco-minniti; last accessed: August 28th, 2019

[4]Politico.eu (2019); Italy's 'minister of fear'; https://www.politico.eu/article/marco-minniti-italy-minister-of-fear/; last accessed: August 28th, 2019.

[5]As remarked, amongst the others, by declarations of Matteo Salvini – La Repubblica, January 19th, 2019; Centinata annegati senza aiuto, ma per Salvini è colpa delle ong – Luigi Di Maio – Il Fatto Quotidiano, April 23th, 2017; Ong 'taxi del Mediterraneo'? Di Maio fa insinuazioni senza dare soluzioni; available at: https://www.ilfattoquotidiano.it/2017/04/23/ong-taxi-del-mediterraneo-di-maio-fa-insinuazioni-senza-dare-soluzioni/3538861/; last accessed: August 27th, 2019 – and, over the continent, Nigel Farage – The Telegraph, September 1st, 2015; Nigel Farage: EU has opened doors to migration exodus of biblical proportions; available at: https://www.telegraph.co.uk/news/politics/nigel-farage/11836131/Nigel-Farage-EU-has-opened-doors-to-migration-exodus-of-biblical-proportions.html; last accessed: August 27th, 2019.

[6]While reverse causation between anti-migrant sentiment and the success of these parties cannot be excluded, evidence from Hangartner et al. (2019) suggests that inflows of refugees can change natives' attitudes and policy preferences, exploiting a natural experiment conditioning inflows of refugees to Greek islands. Whether these changes are further fuelled by preexisting institutional weaknesses leading poor integration of migrants in local communities and labour markets, is, however, still unclear.

of this narrative. The Twitter hashtag "#Portichiusi" embodied this new stance, which championed, quite literally, the closure of ports to all vessels carrying rescued migrants.

Indeed, since the appointment of the M5S-Lega government in June 2018, led by prime minister Giuseppe Conte, the Italian migration policy has been consistently characterised by a rescue-deterrence stance: S&R activities of NGO vessels operating in the Central Mediterranean have been actively hindered, while the Italian coast guard has also been prevented from rescuing shipwreck victims outside Italian territorial waters (IOM, 2018). In multiple occasions, the Ministry of Interior prevented vessels carrying rescued migrants from reaching Italian shores: these actions went as far as minister Salvini risking his own prosecution for kidnapping after ordering the Italian Coast Guard ship U. Diciotti not to let rescued refugees disembark. His prosecution was only deterred by the Parliament voting in favour of his immunity from prosecution.[7] These deterrence efforts also culminated in two decree laws (known as *Decreto Sicurezza* and *Decreto Sicurezza bis*), enacted in October 4th, 2018 and June 15th, 2019, respectively, which toughened immigration laws, deprived asylum applicants of many legal guarantees and *de facto* criminalised migrant rescue operations.

It is still unclear whether the rescue-deterrence policies promoted by the government have been successful in their aims. Does signalling an abdication of responsibility to rescue stranded people at sea, increasing the perceived cost of crossing, actually reduces migration flows? And what is the actual cost of such a policy in terms of human lives being lost? In this paper, I aim to answer these questions, estimating the impact of deterrence policies such as the one promoted in Italy in terms of variations in migration flows and their human costs.

At the same time, the question whether increases in mortality rates alone affect decisions to migrate remains unanswered, as the introduction of the policy is not sufficient to determine whether changes in flows can be directly attributed by variations in mortality, or rather by other factors. There are specific economic assumptions concerning risk-adversity and learning processes in refugee migration decisions and smugglers activity which underlie rescue-deterrence policies. My aim is to test whether permanent and short-term changes in 'environmental' risk factors can affect (or change) migratory flows in this context.

The paper is organised as follows: Section 3.2 discusses the literature on the refugees crisis and migration decisions under risk; section 3.3 describes the data

[7] Reuters (2019); Italian parliament saves Salvini from migrant kidnapping probe; available at: https://www.reuters.com/article/us-italy-politics-salvini/italian-parliament-saves-salvini-from-migrant-kidnapping-probe-idUSKCN1R11Y4; last accessed: August 28th, 2019.

sources used in this paper, while section 3.4 elaborates and discusses the econometric models used to produce the final estimates. Then, the following three sections present the main results; showing the impact of rescue-deterrence on migration flows (section 3.5) and their effect on mortality rates (section 3.6), while estimates for the impact of variations in mortality on migration attempts are contained in section 3.7. Finally, section 3.8 presents the conclusion.

3.2 Overview of the literature

An influential work from Dustmann et al. (2017a) first focused on the European refugee crisis from both a policy and labour integration perspective. The authors found that coordination across the EU has been lacking, and that member states have interpreted their responsibility within the Geneva Convention for Refugees with much liberty, which also underlines how asylum applications have been far from equally distributed across countries. Notably, the study also finds that integration into labour markets has been more difficult for refugees than for economic migrants. Weaknesses in asylum policies across EU member states, along with their limited effectiveness in controlling migration flows, were already underlined in Hatton et al. (2004), who called for more coordination across the EU, and in Facchini et al. (2006), who argued that these inefficiencies are the systemic product of strategic delegation in this policy field.

All these studies place emphasis on the distinction between forced migrants (or refugees) and economic migrants which is, indeed, very important. Economic migrants have long been studied in the literature and their migration decisions have long been framed within economic theory.[8] Risk can, indeed, act both as a pull and push factor both for economic and forced migrants.

Forced migration, however, responds to different causes, such as violent conflicts, as studied in Schmeidl (1997), Vogler and Rotte (2000), Neumayer (2005), and Melander and Öberg (2007), amongst others. Direct changes in exposure to risk factors in the origin country[9] can then affect decisions to leave. However, 'pull' factors in forced migration, influencing the decision on which country to migrate to, including how to reach this destination, are much less unambiguous. Forced migrants are assumed to account for all these risk factors when considering their expected utility for staying or leaving.

Recent forced migration flows to Europe and their determinants were studied

[8]See Lee (1966), and Borjas (1989)

[9]See also Rodriguez and Villa (2012) for how displacement responds to changes in kidnapping risk, and Deschenes and Moretti (2009) for changes in mortality.

in Brück et al. (2018), who focused on ‘push’ and ‘pull’ factors influencing refugee migration decisions. In regards to the flows following from the Arab Spring, violence in the origin country and unemployment in the host country appear as significant predictor of arrivals.

Some attempts have been made to quantify the impact of rescue-deterrence policies in Italy and their influence as (negative) pull factors. In this regard, the work from Cusumano and Villa (2019) is certainly worth mentioning. This work similarly focuses on the effect of “desert diplomacy” and “#Portichiusi” policies on migration flows, estimating a monthly reduction of 12,119 (SE: 2,417; ≈408 daily) and 6,539 (SE: 2,717; ≈220 daily) arrivals by the Sea, respectively. Most notably, no significant effect from the presence of NGOs S&R vessels is found. Descriptive evidence in the relationship between official search and rescue operations and mortality was also provided in Deschenes and Moretti (2009).

I expand on this work by providing a more robust analysis of mortality and migration flows, using high frequency data and controlling for exogenous time-dependent migration shocks by adopting a comparative approach with other migration routes. In any case, the value of the present work resides beyond the analysis of flow-deterrence as this is, to the best of the author’s knowledge, the first attempt to provide a comprehensive analysis of the impact of these policies on mortality rates, and on the impact of mortality on short-term decisions to migrate, with a focus on reverse causality issues.

It is straightforward that rescue-deterrence policies comply to an underlying economic rationale: individual costs of migration are increased as agents update their information on the perceived riskiness of a migration attempt. For a deterrence policy to be successful, this perception can be modified either directly, through factual increases in mortality (an "extensive" margin), or indirectly, through updates in expectations generated by the announcement of the policy itself (the "intensive" margin).

Previously cited literature focused precisely on this mechanism, where variations in migration flows are attributed to environmental changes in risk. The perception of risk among migrants and refugees, however, might diverge significantly from the general population. The literature on migration choice is not unfamiliar to risk attitudes, and can instruct the design of this study.

The relatively lower risk-aversion of migrants is documented by a number of studies. Looking at economic migrants, Jaeger et al. (2010) first found that lower risk-aversity makes individuals more likely to migrate. These findings were later reconfirmed, in a different context, by Dustmann et al. (2017b).

Risk attitudes were first studied in the context of the EU refugee crisis in Bocquého et al. (2018), which studied the risk preferences of refugees in an experimental setting, and framed the results within cumulative prospect theory. The authors conclude that forced migrants are characterised by low loss aversion and a preference for gains, implying that deterrence measures might be less effective on this population.

This literature is valuable to us, but does not provide much insight into the possibility of changes in risk aversion. Under this line of thought, flows might change due to increases in mortality, in accordance with individual risk aversions, which would remain the same. Another strand of literature, instead, investigated whether external events might affect individual preferences for risk. In the context of our research question, these studies might suggest that increases in the mortality rates of the Mediterranean could also generate permanent changes in migration behaviours.

A paper from Callen et al. (2014) focused on direct exposure to violence and changes in risk attitudes, arguing that traumatic experiences lead to individuals developing a preference for certainty. However, these results were reassessed and criticized by Vieider (2018), while Akgüç et al. (2016) also provided evidence that individual risk attitudes are unaffected by substantial changes in the environment.

3.3 Data

I use data from the International Organization for Migration (IOM) and Frontex, integrating these sources with data from Eurostat, the Italian Air Force (ITAF – *Aeronautica Militare*, in Italian), the Armed Conflict Location and Event Data Project (ACLED) and information on NGOs S&R vessels activity from a number of news sources.

The IOM reports and disseminates data on both migration flows[10] and missing migrants[11] in the Mediterranean Sea. IOM data on registered arrivals is collected 'through consultations with the ministry of the interior, coast guards, police forces and other relevant national authorities IOM (2018). This data includes arrivals both by land and by sea, and is reported for all major countries of first arrival in Europe.

Information on missing migrants, instead, is reported for each of the three main migration routes in the Mediterranean: namely, Western (possible destination: Spain), Central (Italy or Malta) and Eastern (Greece and, more rarely, Cyprus). This data is collected from multiple sources, which are detailed in the data-set:

[10] Available at: https://migration.iom.int/europe?type=arrivals
[11] Available at: https://missingmigrants.iom.int/downloads

these could be, for example, government authorities, the press, alarm phones or search and rescue vessels. The estimated date of the accident is also based on these sources, which are classified by their quality: therefore, I drop all observations on incidents based on information from only one media source (level 1), so as to minimise measurement error.[12]

Both figures on arrivals and missing migrants are generally reported daily, with a few relevant exceptions. While for Italy, Malta, Greece and Cyprus daily flows by sea have been available since 2014, Spanish authorities have only began to do so recently, as accesses have reported monthly up until the end of 2017. Daily arrivals to Spain for 2017 are then imputed by dividing total monthly flows across each day in a month.

Frontex also supplies monthly information on the origin of migrants crossing the sea.[13] I have used these figures to reconstruct the nationality of migrants attempting to cross the Mediterranean. Monthly unemployment figures from Eurostat for Spain, Italy and Greece were also collected, so that they can be used as 'pull' factor controls in my econometric model, following from the findings of Brück et al. (2018).

The Italian Air Force releases its *Meteomar* bulletin every six hours, announcing forecasts and present conditions for the sea state in the Mediterranean. The bulletin offers notices and forecasts for 22 areas in the Mediterranean. Comparing information on migration routes from the IOM and the coordinates from these areas, I match three "seas" where most migration attempts occur: the Libyan sea for the Central Mediterranean Route, the Alboran sea for the Western Route, and the Aegean sea for the Eastern Route. Running a simple text mining technique, I then reconstruct daily sea state conditions for each of the three routes, generating indicators for occurrences (and their forecasts) of seastorms (meaning a ≥ 7 score on the Douglas scale) and thunderstorms.

Information on political instability at exit points is obtained through ACLED (Raleigh et al., 2010). I proxy political instability through the daily number of fatal accidents linked to violent political unrest registered in the largest exit point countries (Libya and Tunisia for the Central Mediterranean route, Morocco and Algeria for the Western Mediterranean, Turkey for the Eastern Mediterranean). Instability in origin countries is, instead, not relevant to this analysis, as I study variations in flows comparatively across routes, meaning that these shocks are absorbed by design once controlling for shocks affecting a single route.

An overview of non-governmental S&R operations is provided by the European

[12]More information on the methodology of the Missing Migrants project from the IOM is available at: https://missingmigrants.iom.int/methodology.

[13]Available at: https://frontex.europa.eu/along-eu-borders/migratory-map/

Union Agency for Fundamental Rights.[14] Using this information, I reconstruct the time-frame of activities for each NGO vessel, and create an indicator for the number of active private S&R vessels on each day for every given route.

Finally, I harmonise and aggregate this information to construct a final dataset, which includes daily information from January 2017 to July 2019. As, starting from July 2017, the introduction of the "desert diplomacy" policy in the Central Mediterranean route significantly affected the number of accesses, the empirical model will also allow for the introduction of multiple policies. At the same time, this time-frame is convenient as, during these years, no other significant unilateral changes in migration policies in the Eastern and Western Mediterranean exit and entry points have taken place.[15]

All flows to the five Mediterranean states are aggregated into the three main routes, for 911 days. This simplification is motivated by two distinct necessities: firstly, this is the only way to reduce all data sources to a common number of dimensions and, secondly, the final destination of missing migrants can never be fully ascertained. Nevertheless, the absence of significant unilateral changes in migration policies in Malta (meaning that flows to Malta will be as well affected by policy changes in Italy), along with the negligible volume of migrants who eventually land on this island, allows the aggregation of all flows into the three distinct routes without compromising the estimates.

3.4 Empirical Model

As a model for migration flows, I propose the following specification, using a group-level difference in difference estimator:

$$\begin{aligned} Attempts_{rt} = \beta_0 + Policy^{'}\beta_1 + Vessels_{rt}\beta_2 + Deaths^{'}\beta_3 + Seastate^{'}\beta_4 + \\ Unrest^{'}\beta_5 + PullShocks^{'}\beta_6 + PushShocks^{'}\beta_8 + \gamma_r + \delta_t + \epsilon_{rt} \end{aligned} \tag{3.1}$$

where β_0 is a constant and $Policy$ is a vector of migration policies ($Policy_{1rt}, Policy_{2rt}, ...,$ $Policy_{Prt}$) at time t in each migration route r, which can either be western, central, or eastern Mediterranean. The outcome variable $Attempts_{rt}$ indicates the number of daily migration attempts – that is, the sum of deaths and successful arrivals by

[14] Ibidem, 1

[15] For reference, see https://ec.europa.eu/home-affairs/sites/homeaffairs/files/what-we-do/policies/european-agenda-migration/20190306_managing-migration-factsheet-step-change-migration-management-border-security-timeline_en.pdf.

sea. As it takes less than a day to reach Italy from Libya,[16] I assume that missing or dead individuals would have reached their destination on the same day of their disappearance, if they survived. $Vessels_{rt}$, instead, captures the number of active search and rescue vessels operated by NGOs, while $Deaths$ is a vector of lags and moving averages for reported dead or missing migrants.

Route specific variation in sea conditions is captured by the $Seastate$ vector, including lagged values for sea state conditions in each route, while γ_r and δ_t are, respectively, fixed effect specific to the migration route, week and month. This specification essentially establishes a group-level difference-in-differences setting, where daily fixed effects are captured by the δ_t term, while γ_r covers the baseline effect of each migration route. As the $Policy$ vector already captures the interaction between time and the 'treated' route, the coefficients contained in β_1 will yield the impact of each policy on the dependent variable.

Finally, $Unrest$ specifies lags and moving averages for political unrest in the exit regions, while $PushShocks$ and $PullShocks$ are vectors of push and pull factors affecting the relative supply of migrants in a given route, in a given day. These factors should include those economic, cultural and geographic determinants which might affect the number of attempts in a given route and which are not already captured by route and time fixed affects.

$PullShocks$ includes lagged values for the monthly level of unemployment in the main destination of each of the three routes – so either Spain, Italy or Greece.

Given that modelling socio-economic shocks from each country of origin would generate an excessive loss of degrees of freedom, possibly leading to over-specification bias, $PushShocks$ controls for exogenous shocks specific to populations which only cross a single route. This term is obtained by first identifying, for each route, the monthly share of the nationalities which, during the full study window, have attempted crossing on this specific route.

This monthly ratio is then multiplied by the daily attempts, and then de-trended, to obtain an approximation of the daily variation in migrants whose preference for a single route is infinitely elastic with regards to the pull-factors conditioning access through the other routes. In other words, this term controls for variations in route-specific push-factors, which might pose as a source of bias for total amount of accesses which cannot already be controlled by the fixed effects terms in the model. While I could still assume route-specific migration shocks to be random and uncorrelated

[16] The Economist (2015); Everything you want to know about migration across the Mediterranean; available at: https://www.economist.com/the-economist-explains/2015/04/21/everything-you-want-to-know-about-migration-across-the-mediterranean; last accessed: September 2nd, 2019

with the policy adoption, the very small number of groups in my analysis (n = 3) still requires shocks to be shared across the three routes for them not to affect the results.

The model proposed in equation (3.1) is then re-adapted to study daily mortality in each route, as in equation (3.2):

$$\begin{aligned} Deaths_{rt} = \beta_0 + Policy^{'}\beta_1 + Vessels_{rt}\beta_2 + Attempts_{rt}\beta_3 + Seastate^{'}\beta_4 + \\ Unrest^{'}\beta_5 + PullShocks^{'}\beta_6 + PushShocks^{'}\beta_8 + \gamma_r + \delta_t + \epsilon_{rt} \end{aligned} \tag{3.2}$$

where $Deaths_{rt}$ indicates the daily number of dead or missing migrants per route, while the fourth term in the right side of the equation refers to migration attempts in the same day. Compared to equation (3.1), this term is not a vector of lags and moving averages, but refers to the number of migration attempts on the same day.

Indeed, I assume that daily attempts might be affected by previous variations in mortality because of changes in the perceived risk of migration, but daily fatal accidents on the sea would be solely dependent on the number of attempts on the same day, as I assume the mortality rate $Deaths_{rt}/Attempts_{rt}$ to be exogenous to the number of attempts.[17]

Models (3.1) and (3.2) are more than sufficient for testing the impact of rescue-deterrence policies in terms of flow reduction and mortality. In these cases, potential reverse causality issues between flows and mortality are tempered by the assumption that these policies generate a permanent and exogenous change in perceived mortality rates.

I am, however, also interested in whether mortality itself affects flows as a deterrent, independently from policy actions. However, estimating the effect of variations of perceived mortality in the sea over migration attempts is a much more complicated issue.

First of all, how actors – being migrant or smugglers – update their information on mortality should be discussed. A perfectly informed and rational actor would update his/her information on migration risk using the expectation $E[Deaths_{rt}/Attempts_{rt}]$ for the mortality coefficient. These expectation could be based on previous values for this coefficient, such as $Deaths_{rt-1}/Attempts_{rt-1}$ or a ratio of moving averages.

But do individuals really account for the denominator in this coefficient? Increases in absolute mortality in the sea might be caused by increases in flows, with relative mortality remaining fixed; but it seems difficult to believe that individuals

[17] Also, on a slightly more technical note, as the IOM already adjusts the date of fatal accidents to the most plausible day, there is no need to lag the co-variate group to account for delays in the reporting of deaths.

would prioritize relative over absolute information on mortality.

The behavioural literature has presented evidence in support of this assumption. Availability heuristics (as first studied in Tversky and Kahneman, 1974) have been shown to influence individual expectations on the likelihood of an event based on 'the ease with which instances or occurrences can be brought to mind.' These heuristics can lead to biases where other relevant information is not as easily available. Large tragedies in the Mediterranean sea usually receive large media coverage and, due to their salience, tragic events might also be easier to recall. In contrast, information on flows might not be as easily available, or flows could be mistakenly held as fixed. This could lead to biases in decision making, where individuals would value absolute over relative mortality.

Also, even if the denominator were known, it is unlikely that individuals are immediately exposed to the mortality ratio, rather than having to compute that for themselves from absolute figures. Indeed, studies in affect heuristics (see Slovic et al., 2000) have shown how perceived risk is greater when information is updated through frequency-based scales rather than communicated through probabilities.

It follows that, in order to measure the effect of increases in mortality and its effects on the perceived risk of migration, absolute increases in mortality should be used as explanatory variables, rather than changes in relative mortality.

However, perceived mortality – if informed by the aforementioned processes – cannot be considered exogenous to the number of migration attempts, as total deaths are clearly a function of the latter. Indeed, if individuals – either smugglers or migrants – do not take into account the total number of attempts to inform the decision to leave the exit point, and if perceived mortality is only affected by the absolute death tool, then perceived mortality is not exogenous, and the effects of reported deaths on attempts will be upwardly biased.

Removing this bias is then paramount. Further, lagging deaths is not a sufficient solution to the problem. As discussed in Bellemare et al. (2017), the use of lagged explanatory variables cannot appease endogeneity concerns when there are reasons to believe lagged values are still suffering from endogeneity. In this case, a different empirical strategy is needed.

For increases in absolute mortality to be really exogenous, these increases should be assigned independently from migration flows. The solution to this problem will come from an instrumental variable approach and, more precisely, by exploiting the random variation generated by relative mortality itself.

Indeed, as my estimates will later confirm, the relative mortality rate can be exogenous to the number of migration attempts, while still retaining relevant predictive power on absolute mortality. This holds under a specific condition: being relative mortality measured from the ratio between the daily number of deaths and attempts, this ratio needs to be orthogonal to the attempts.

In other words, the mortality rate needs to be influenced by random chance only, pertaining for example by unexpected weather conditions or hazards outside of the control of smugglers and/or migrants. If the number of migration attempts further affected mortality, then the effect of attempts on the ratio between deaths and attempts would be positive. More formally, mortality can then be modelled as:

$$Mortality_{rt} = \frac{Deaths_{rt}}{Attempts_{rt}^{\eta}} \tag{3.3}$$

with the η exponent capturing any non-linear effect of attempts on the mortality rate, given the number of deaths. For η equal to -1, mortality is an increasing linear function of the attempts, convex for $-1 < \eta < 0$, concave for $\eta < -1$. For η close to zero, instead, mortality is independent from the number of attempts given the deaths. If this condition holds, then the mortality rate will only be endogenous if influenced by factors which actors can control for. Controlling for these factors, using for example weather forecasts on sea state conditions, the remaining variation in relative mortality is then as good as random.

This orthogonality condition can be tested linearly simply by regressing mortality on the attempts rate: if the attempts coefficient is zero, then attempts are surely independent from mortality. However, η can be estimated empirically as, taking logs on both sides, and rearranging, equation 3.3 turns into the regression equation:

$$log(Mortality_{rt}) = C_{rt} - \eta log(Attempts_{rt}) + \epsilon_{rt} \tag{3.4}$$

setting the constant term to equal the time and route fixed effects, as $C_{rt} = \beta_0 + \gamma_r + \delta t$, then a fixed effect model can provide an estimate for η.

As will be shown later, the estimates for η which I obtain are not statistically different from zero, confirming that relative mortality is indeed exogenous to the number of migration attempts. The other conditions that the instrumental variable have to satisfy are much easier to discuss: higher mortality rates clearly affect absolute mortality, as the first stage regressions will later show, satisfying the relevance condition; while the exclusion restriction, which usually cannot be tested, holds under the rather reasonable assumption that, after controlling for time and route fixed effects in migration flows, the relative mortality ratio does not affect the number of migration attempts other than through the absolute mortality channel.

Should these conditions hold, then instrumenting absolute mortality through its relative component will yield an unbiased estimate for the effect of each additional death in the Mediterranean on migration attempts. Filtered by this instrument, reported deaths become as good as random, as increases in absolute mortality are randomised through increases in relative mortality, controlling, in this way, for the stocks of flows already affecting the total number of dead or missing.

This being considered, the resulting first stage regressions will be:

$$\begin{aligned} Deaths^{'} = \beta_0 + Policy^{'}\beta_1 + Vessels_{rt}\beta_2 + Mortality^{'}\beta_3 + Seastate^{'}\beta_4 + \\ Unrest^{'}\beta_5 + PullShocks^{'}\beta_6 + PushShocks^{'}\beta_8 + \gamma_r + \delta_t + \epsilon_{rt} \end{aligned} \quad (3.5)$$

where $Deaths$ and $Mortality$ are vectors of lags and moving averages for absolute and relative mortality. Model (3.5), indeed, implies as many first stages as the number of lags in the model. For clarity purposes, I will use a one-day lag and a varying moving average for the past days, but more combinations are certainly possible.

Finally, these values from model (3.5) are plugged in the second stage equation (3.6):

$$\begin{aligned} Attempts_{rt} = \beta_0 + Policy^{'}\beta_1 + Vessels_{rt}\beta_2 + \widehat{Deaths}'\beta_3 + Seastate^{'}\beta_4 + \\ Unrest^{'}\beta_5 + PullShocks^{'}\beta_6 + PushShocks^{'}\beta_8 + \gamma_r + \delta_t + \epsilon_{rt} \end{aligned} \quad (3.6)$$

where $\widehat{Deaths}$ is a vector of predicted values from the first stages in equation (3.5).

Finally, the nature of the model and data used requires a few words to be spent on the correct calculation of standard errors.

In panel difference-in-difference designs, it is often suggested that standard errors should be clustered by group and time. However, in instances, such as ours, when data is already aggregated at the group level, clustering standard errors by group and time is not any different from not clustering at all (Cameron and Miller, 2015). In these cases, it is suggested to simply cluster by group, and leave the time dimension unclustered (Bertrand et al., 2004; Cameron and Miller, 2015).

Collapsing individual data on the group level of aggregation in order to obtain more consistent standard errors is, after all, already a standard practice in the applied literature (see Bertrand et al., 2004).

However, in panels with small n and large t, we also need to account for serial correlation and time-wise heteroscedasticity. In these instances, an asymptotically efficient Panel Corrected Standard Errors (PCSE) estimator, as proposed by Beck and Katz (1995), is often considered appropriate. The estimator assumes distur-

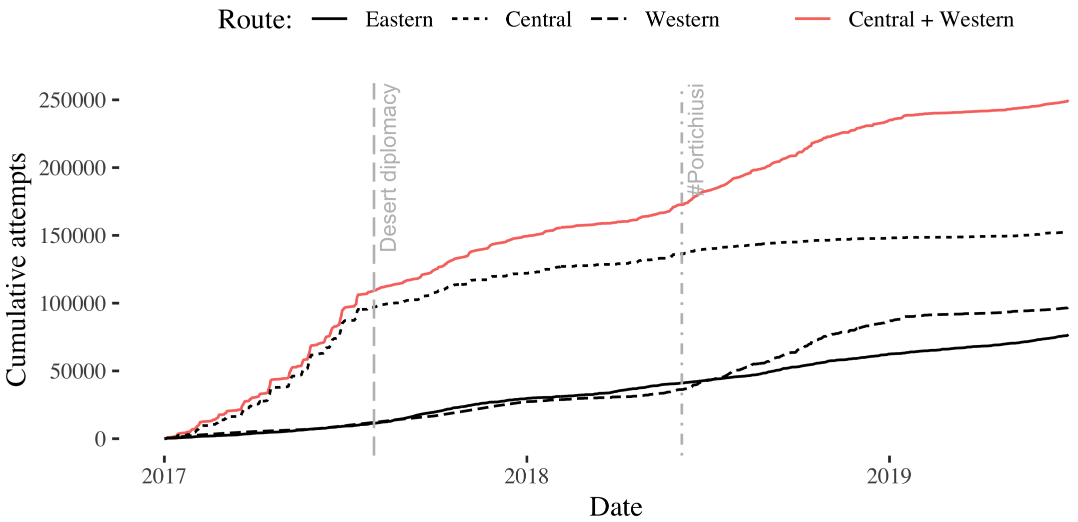


Figure 3.5.1: Cumulative migration attempts in the Mediterranean

bances to be serially correlated and heteroskedastic, and will then be used to produce robust standard errors for the remainder of the paper.

3.5 Deterrence and migration flows

Cumulative migration flows across the three routes, starting from January 2017, are plotted in figure 3.5.1.

While, naturally, the figure only focuses on the time-driven variation in flows, omitting the influence of other important predictors, preliminary visual evidence already suggests that the most significant impact on migration flows is linked with the introduction of the "desert diplomacy" approach under the Gentiloni government. Comparing trends in the Central Mediterranean with the other routes, parallel trends in growth between the three routes have been disrupted by the introduction of the policy, after which the slope of the Central Mediterranean route has changed.

Rescue-deterrence policies (#Portichiusi) enacted during the Conte government, instead, do not seem to have had a great impact on flow reduction: indeed, since the introduction of Minniti's policies ("desert diplomacy"), cumulative attempts in the Central Mediterranean have grown linearly, suggesting that no significant change in trends has occurred. However, while trends in growth between the Central and Eastern Mediterranean have remained parallel, the Western route has witnessed a peculiar increase in accesses, suggesting that a relocation effect might have taken place.

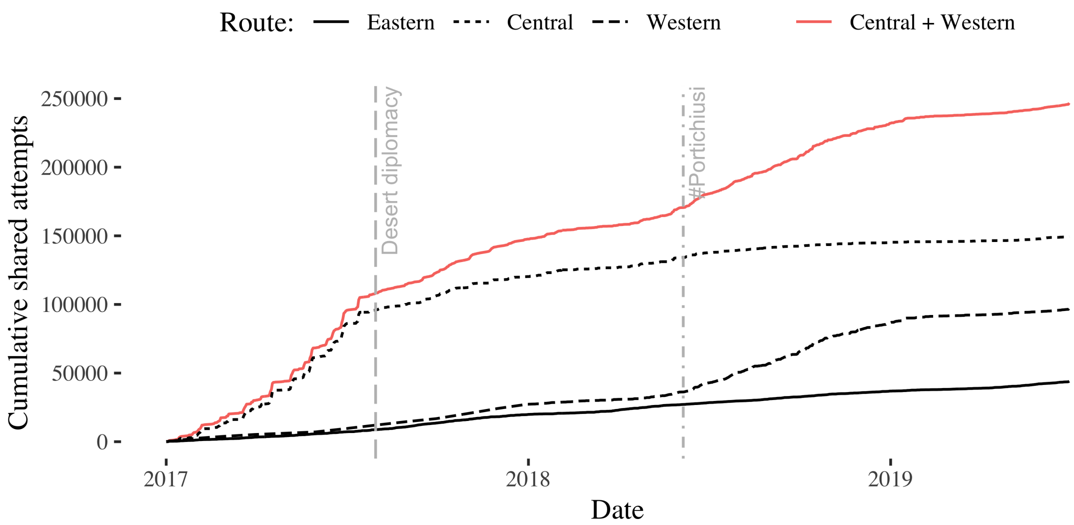


Figure 3.5.2: Cumulative migration attempts in the Mediterranean ($A \cap B \cap C$ set)

Results from figure 3.5.1 might, however, be misleading: shocks unique to a specific route might lead to a bias in our estimates, especially if we suspect the presence of route-specific and time-dependent migration shocks from specific countries. In these cases, the only solution is to suppress the 'unique' components from flows, filtering out all entries from countries of origin where migrants do not have access to all three routes.

More formally, suppose that there are three sets of nationalities A, B and C. Nationalities in A have had at least one migrant attempting to cross, say, the Western Mediterranean route across the estimation window, and the same goes for the other sets and routes. Focusing on the $A \cap B \cap C$ subset enables us to focus only on those migrants who had the option to choose between the three routes, controlling for route-specific shocks.

I do so in figure 3.5.2, by obtaining the monthly proportion of shared flows from Frontex data s_{rt}, and applying this ratio to the daily number of arrivals. This method is valid as long as we assume that the daily nationality composition to reflect monthly figures, a decently reasonable simplification.

After adjusting flow figures for this component, previous trends appear to persist, including the relocation effect, which even experiences a higher relative increase.

Regression results are plotted in table 3.5.1. Column (1) estimates the baseline effect of the two policies, omitting all controls except for day and route fixed effects. I find a statistically significant reduction of 392 migrants per day after the introduction of Minniti's policies, and a statistically significant -122 reduction for Salvini's

Table 3.5.1: Effect of deterrence policies on daily migration attempts

	Daily migration attempts							
	Total Attempts						Shared Attempts	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
#Portichiusi	−122.033***	−126.731*	−159.237***	−158.160***	−131.227**	−56.944**	−156.669**	−45.571*
	(40.755)	(73.842)	(54.424)	(54.542)	(59.785)	(22.726)	(66.928)	(27.044)
#Portichiusi (W. Route)						92.375***		102.841***
						(10.110)		(9.030)
Desert diplomacy	−391.969***	−360.457***	−383.621***	−383.135***	−369.846***	−361.071***	−372.303***	−332.384***
	(10.434)	(20.443)	(16.688)	(12.924)	(6.887)	(23.310)	(18.346)	(25.097)
Desert diplomacy (W. Route)						−16.479		11.747
						(10.813)		(9.207)
No of S&R Vessels		−2.389	−9.770*	−9.711*	−7.024	−0.312	−11.164	0.752
		(7.998)	(5.521)	(5.157)	(4.969)	(4.876)	(7.035)	(5.564)
Reported Deaths (lag 1)		2.940***	2.925***	2.925***	2.923***	2.927***	2.899***	2.898***
		(0.223)	(0.203)	(0.203)	(0.215)	(0.217)	(0.195)	(0.214)
Reported Deaths (MA14)		5.283***	5.059***	5.058***	4.159***	3.506***	5.113***	3.256***
		(0.376)	(0.456)	(0.450)	(1.106)	(1.027)	(0.594)	(0.862)
Wind Storm 00h00			−38.981***	−39.020***	−38.629***	−36.130***	−35.823***	−32.322***
			(4.090)	(3.823)	(6.377)	(6.659)	(5.108)	(8.051)
Wind Storm (lag 1)			−24.090*	−24.138*	−22.038	−20.433	−26.196**	−20.573
			(14.272)	(13.915)	(13.590)	(14.076)	(12.519)	(13.474)
Thunder Storm 00h00			−16.953	−16.829	−19.011	−18.882	−17.983	−19.255
			(21.405)	(20.960)	(20.108)	(20.074)	(20.210)	(19.255)
Thunder Storm (lag 1)			−17.581**	−17.490**	−19.423**	−18.636**	−17.261*	−17.477*
			(8.662)	(8.871)	(8.895)	(9.177)	(8.849)	(9.353)
Instability at exit point (lag 1)			−9.609	−9.618	−8.498	−8.269	−9.057	−7.596
			(8.454)	(8.447)	(9.208)	(9.269)	(8.226)	(9.052)
Instability at exit point (MA14)			−13.289	−12.584	−13.916***	−28.127***	−10.747	−33.531***
			(12.834)	(7.752)	(4.717)	(3.796)	(7.985)	(5.206)
Youth unemp. in host country (log, lag 5 months)				−36.009	−90.520	219.977	−135.419	312.283
				(278.925)	(253.148)	(329.335)	(285.290)	(302.867)
Country of origin controls	*No*	*No*	*No*	*No*	*Yes*	*Yes*	*No*	*No*
Route fixed effects	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Time fixed effects	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
R^2	0.43	0.45	0.455	0.455	0.475	0.478	0.459	0.482
Adjusted R^2	0.143	0.172	0.177	0.176	0.206	0.208	0.182	0.216
Observations	2,733	2,691	2,691	2,691	2,691	2,691	2,691	2,691

Note: *p<0.1; **p<0.05; ***p<0.01

Beck and Katz (1995) Panel Corrected Standard Errors, clustered by route.

Table 3.5.2: Effect of deterrence policies on daily migration attempts (log)

	Daily migration attempts (log)							
	Total Attempts						Shared Attempts	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
#Portichiusi	−0.790***	−0.049	−0.604	−0.801**	−0.612	−0.733	−0.769	−0.448
	(0.168)	(0.690)	(0.485)	(0.377)	(0.508)	(0.549)	(0.476)	(0.585)
#Portichiusi (W. Route)						−0.060		0.227*
						(0.122)		(0.127)
Desert diplomacy	−0.862*	−0.286	−0.622**	−0.723***	−0.651***	−1.066**	−0.694***	−0.853*
	(0.441)	(0.300)	(0.309)	(0.144)	(0.175)	(0.433)	(0.152)	(0.459)
Desert diplomacy (W. Route)						−0.716***		−0.496***
						(0.141)		(0.144)
No of S&R Vessels		0.141	0.015	0.006	0.003	−0.027	−0.018	−0.017
		(0.159)	(0.115)	(0.077)	(0.092)	(0.102)	(0.079)	(0.108)
Reported Deaths (lag 1) (log)		0.224***	0.207**	0.212**	0.221***	0.221***	0.207**	0.215***
		(0.082)	(0.085)	(0.083)	(0.084)	(0.084)	(0.082)	(0.083)
Reported Deaths (MA14) (log)		0.021	0.043	0.030	0.002	0.034	0.050	0.028
		(0.049)	(0.039)	(0.042)	(0.062)	(0.056)	(0.033)	(0.052)
Wind Storm 00h00			−0.715***	−0.707***	−0.689***	−0.714***	−0.700***	−0.692***
			(0.147)	(0.143)	(0.151)	(0.152)	(0.143)	(0.150)
Wind Storm (lag 1)			−0.346**	−0.337**	−0.349**	−0.369***	−0.373***	−0.388***
			(0.157)	(0.158)	(0.149)	(0.140)	(0.142)	(0.131)
Thunder Storm 00h00			−0.035	−0.060	−0.087	−0.091	−0.064	−0.090
			(0.124)	(0.120)	(0.101)	(0.100)	(0.117)	(0.098)
Thunder Storm (lag 1)			−0.157*	−0.176*	−0.208**	−0.217***	−0.166*	−0.203**
			(0.088)	(0.090)	(0.081)	(0.077)	(0.097)	(0.086)
Instability at exit point (lag 1)			0.035***	0.037***	0.042***	0.042***	0.031***	0.038***
			(0.006)	(0.007)	(0.011)	(0.010)	(0.007)	(0.011)
Instability at exit point (MA14)			−0.479***	−0.610***	−0.626***	−0.465***	−0.547***	−0.495***
			(0.128)	(0.124)	(0.095)	(0.116)	(0.110)	(0.120)
Youth unemp. in host country (log, lag 5 months)				6.743***	4.959***	2.176	5.413***	2.953
				(1.899)	(1.877)	(2.456)	(1.536)	(2.375)
Country of origin controls	*No*	*No*	*No*	*No*	*Yes*	*Yes*	*No*	*No*
Route fixed effects	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Time fixed effects	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
R^2	0.449	0.458	0.48	0.485	0.507	0.509	0.485	0.508
Adjusted R^2	0.172	0.185	0.214	0.222	0.254	0.256	0.221	0.254
Observations	2,733	2,691	2,691	2,691	2,691	2,691	2,691	2,691

Note: *p<0.1; **p<0.05; ***p<0.01

Beck and Katz (1995) Panel Corrected Standard Errors, clustered by route.

policies. Other regressors are then reintroduced in the next specification, until the full model is presented in column (5).

Column (2) introduces controls for absolute mortality – including a one-day lag and a moving average for the previous two weeks – and for the number of S&R vessels operating in the sea. Relative mortality controls are, instead, shown in the appendix. The policy coefficients are not particularly affected, and, as the next specification will confirm, it is worth considering that the presence of vessels does not appear to be connected to increases in migration activity, in contrast with the accusations that their presence has acted as a positive pull factor to refugee migration. The interpretation of the mortality coefficients is, instead, less straightforward, as these estimates for mortality are not robust to reverse causality: hence the positive coefficient for both regressors.[18]

Columns (3) and (4) introduce controls for sea state conditions, political instability at exit points and unemployment in the first arrival country. Weather conditions appear as a significant deterrent for migration behaviour, increasing the precision of the model.

Column (5) completes the model, adding controls for route-specific migration shocks. These controls reduce the effect of both the #Portichiusi (moving to -131.227, SE: 59.785) and desert diplomacy policy (-369.846, SE: 6.887). Other regressors retrieve their signs and magnitudes, with a few exceptions; notably, the 14 days moving average for political instability at exit point now becomes a statistically significant predictor of migration attempts. Column (7) retrieves the same specification, but changes the dependent variable so that only the $A \cap B \cap C$ subset is represented, as done for figure 3.5.2. The policy coefficients maintain a comparable magnitude.

Finally, as I suspect relocation effects to the Spanish route, in column (6) I add dummies for policy spill-over effects in the other routes, using the Eastern Mediterranean route as a baseline, to the specification from column (5). I repeat the same exercise in column (8), using the model from column (7).

These results suggest that, undoubtedly, a dislocation effect has been generated. However, given the limitations of the data at hand, it is not currently possible to find which of the two policies ultimately caused the dislocation effect in 2018. At best, the rescue-deterrence policies have generated a ≈157 individuals reduction in daily flows, at worse, this reduction is of only ≈46 refugees per day – and barely significant from a statistical standpoint as well.

Until further evidence is available, discerning whether rescue-deterrence policies

[18] In any case, as will be discussed later, vessel present is also not immune to this issue.

in the Central Mediterranean have had an immediate (as exemplified in figures 3.5.1 and 3.5.2) effect on migration flows towards Spain, or whether the relocation effect has been generated by previous "desert diplomacy" attempts, disrupting migration routes to the coast from within Libya, is left to the judgement of the reader.

In any case, it is evident – both from visual evidence and from coefficient analysis – that total flows in the Mediterranean have not changed much since the introduction of the policy, even if countries like Italy or Malta were comparatively less affected. Holding the Eastern Mediterranean route as the baseline, differences in the rescue-deterrence coefficients from column (6) point at an increase in 35.431 migrants per day across the Western and Central routes, with a 24.873 standard error: after the introduction of rescue deterrence policies, total refugee migration flows have barely changed at all. Coming back to figures 3.5.1 and 3.5.2, and looking at the cumulative flows towards both the central and Western Mediterranean routes, it is evident that, apart for seasonal variations, the increase in flows has been constant over the last years.

Table 3.5.2 shows the same specifications, but with logged dependent variable. These results are similar to the ones obtained from the previous model. Moreover, across all final specifications (columns 5, 6, 7 and 8) the effect of #Portichiusi remains statistically non-significant.

As a word of caution, it should be reminded that the estimates of the policy effect on migratory flows relies on the efficacy of the use of other routes, restricted to the common nationalities of origin, as a counter-factual. The efficacy of this strategy might, then, be limited to the assumption that other unobserved factors influencing daily migration attempts are properly controlled for. However, as said, the marginal contribution from this study goes beyond the estimation of the effect of these policies in terms of daily migration attempts. Indeed, the estimates for the effects of these policies in terms of their effects on mortality does not rely on this assumption, as it is already possible to control for these factors simply because the these models will already control for the number of migration attempts. At the same time, the model estimating the effect of mortality on migration flows relies on the deterrence effects of mortality across all routes, making the need of a counterfactual superfluous.

3.6 Deterrence and human costs

Trends in absolute mortality in the Mediterranean are reported in figure 3.6.1 where, again, cumulative sums for each of the three routes are plotted, starting from January

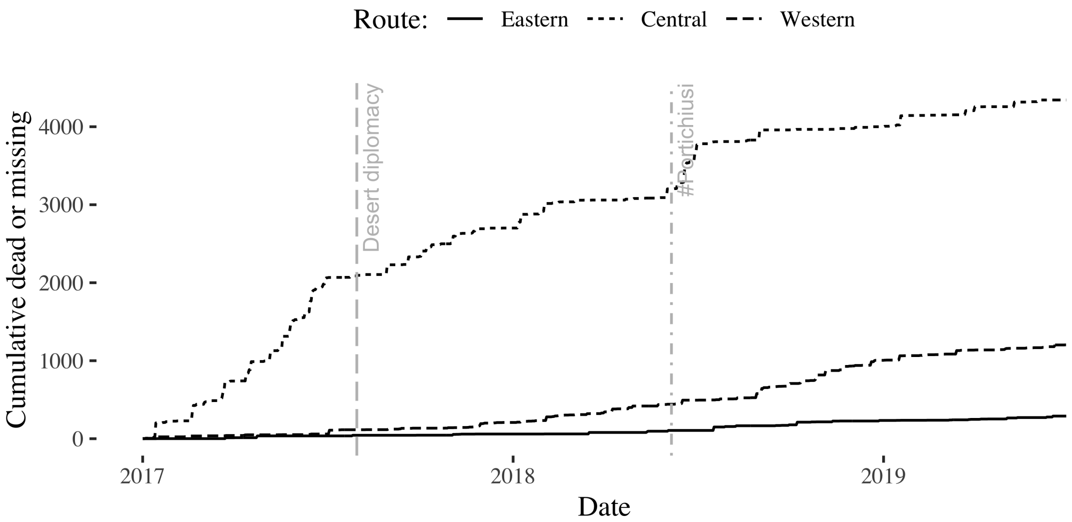


Figure 3.6.1: Cumulative reported dead-or-missing in the Mediterranean

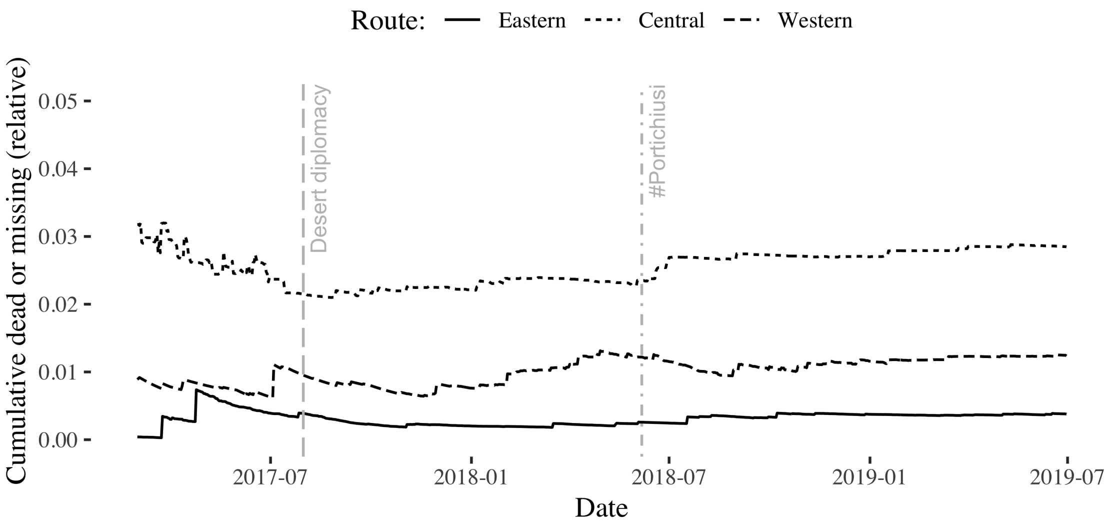


Figure 3.6.2: Cumulative dead-or-missing over cumulative arrivals in the Mediterranean

2017. As mentioned earlier, I also filter by source using the quality criteria set by the IOM, so that only reported deaths than can be fully attributed to migration attempts are estimated. The true mortality rate could then, be, understated.

On the one hand, the figure suggests that trends in mortality have followed a linear path over the entire estimation window in both the Eastern and Western Mediterranean routes. On the other hand, the Central Mediterranean route has experienced two discontinuities in regard to the occasion of the introduction of the two policies. Indeed, since the introduction of "desert diplomacy", mortality in the sea has seen a significant reduction,[19] while "#Portichiusi" rescue-deterrence policies seems to have generated an opposite effect, increasing mortality by a significant margin.

What is disconcerting, at least from this initial visual evidence, is that, after the introduction of rescue-deterrence policies, trends in absolute mortality in the Eastern and Western routes have been proportional to the flows in migration while, in the Central route, mortality has increased, even if migration flows, as seen earlier, have diminished.

Figure 3.6.2 provides some further checks to these findings, showing changes in the ratio between cumulative dead-or-missing accidents and cumulative migration attempts in the Mediterranean.[20] As the ratio in the Central route increases after the introduction of rescue-deterrence policies, the previous interpretation remains fundamentally unchanged.

Table 3.6.1 presents regression output for the effect of deterrence policy on absolute mortality. As migration attempts, being included among the covariates, are already controlled for in the model, there is no reason to worry about the relocation effect which affected previous estimates.

Column (1) introduces a basic specification including the policy vector and migration attempts as controls. These first results might falsely suggest that rescue-deterrence policies have had no effect on mortality rates. Further controls are introduced until the full model is achieved in column (5). Column (2) introduces the number of active NGO S&R vessels as a control, overturning previous results, and suggesting that reverse correlation affected estimates from column (1): this time, the effect of "desert diplomacy" appears not statistically different from zero, while an increase in 4.426 deaths per day is revealed after the appointment of Giuseppe Conte's government. It is important to note that the presence of S&R vessels is not exogenous to the number of deaths, as these vessels are already operated in response to increases in total mortality in the Mediterranean: the estimate for the vessels coefficient is, then, not robust to reverse causality.

Column (3) adds controls for the sea state, while column (4) adds further push and pull factor controls. Column (5) controls for migration shock. In conformity with the estimates from table 3.5.1, sea state controls suggest a negative effect of adverse conditions (one day earlier) on mortality, given the reduction of migration attempts. The other controls add more noise to the model, as expected, with the adjusted r-squared being reduced from their inclusion. The coefficient for the policy

[19] Increases in mortality on land are not to be excluded, but their study goes beyond the aims of this research.

[20] The first months are censored so that the ratio can stabilise.

Table 3.6.1: Effect of deterrence policies on daily mortality

	Daily reported dead or missing						
	Total dead or missing					Minus S&R	Mortality Rate
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
#Portichiusi	0.154	4.426***	4.333***	4.424***	4.444***	3.938***	0.075***
	(0.113)	(0.623)	(0.701)	(0.563)	(0.639)	(0.609)	(0.026)
Desert diplomacy	−2.942***	0.069	−0.043	0.444	0.434	−0.648	−0.055***
	(0.374)	(0.552)	(0.539)	(0.624)	(0.660)	(0.631)	(0.017)
Migration Attempts	0.011***	0.011***	0.011***	0.010***	0.010***	0.010***	−0.0001***
	(0.0005)	(0.0005)	(0.0004)	(0.0005)	(0.0004)	(0.0003)	(0.00001)
No of S&R Vessels		0.865***	0.842***	0.891***	0.881***	0.745***	0.015**
		(0.138)	(0.155)	(0.139)	(0.149)	(0.143)	(0.007)
Wind Storm 00h00			0.116	0.111	0.120	0.188	0.011***
			(0.277)	(0.201)	(0.215)	(0.235)	(0.004)
Wind Storm (lag 1)			−0.338	−0.086	−0.101	−0.041	−0.010
			(0.281)	(0.458)	(0.439)	(0.458)	(0.010)
Thunder Storm 00h00			−0.435	−0.465	−0.473	−0.631	0.003
			(0.584)	(0.581)	(0.597)	(0.626)	(0.013)
Thunder Storm (lag 1)			−0.155*	−0.213***	−0.219***	−0.276	−0.009
			(0.089)	(0.080)	(0.084)	(0.250)	(0.009)
Instability at exit point (lag 1)				0.801*	0.799*	0.844*	0.007
				(0.425)	(0.427)	(0.447)	(0.004)
Instability at exit point (MA14)				−0.823	−0.835	−0.993	0.023
				(0.837)	(0.861)	(0.913)	(0.016)
Youth unemp. in host country (log, lag 5 months)				3.501	2.796	2.787	−0.269**
				(3.688)	(3.693)	(3.135)	(0.107)
Country of origin controls	*No*	*No*	*No*	*No*	*Yes*	*Yes*	*Yes*
Route fixed effects	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Time fixed effects	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
R^2	0.087	0.1	0.101	0.4	0.4	0.4	0.382
Adjusted R^2	0.086	0.099	0.099	0.094	0.092	0.092	0.065
Observations	2,733	2,733	2,733	2,691	2,691	2,691	2,691

Note: *p<0.1; **p<0.05; ***p<0.01
Beck and Katz (1995) Panel Corrected Standard Errors, clustered by route.

Table 3.6.2: Effect of deterrence policies on daily mortality (log)

	Daily reported dead or missing (log)						
	Total dead or missing					Minus S&R	Mortality Rate
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
#Portichiusi	−0.068**	0.251*	0.240*	0.302***	0.305***	0.308***	0.059***
	(0.032)	(0.131)	(0.142)	(0.098)	(0.095)	(0.089)	(0.020)
Desert diplomacy	−0.527***	−0.302***	−0.310***	−0.269***	−0.264***	−0.339***	−0.023**
	(0.091)	(0.096)	(0.101)	(0.069)	(0.067)	(0.063)	(0.011)
Migration Attempts (log)	0.102***	0.100***	0.099***	0.100***	0.100***	0.092***	−0.003
	(0.022)	(0.022)	(0.023)	(0.022)	(0.022)	(0.020)	(0.002)
No of S&R Vessels		0.065**	0.063**	0.073***	0.074***	0.068***	0.012**
		(0.029)	(0.032)	(0.022)	(0.022)	(0.020)	(0.005)
Wind Storm 00h00			0.022***	0.026**	0.025**	0.030**	0.007***
			(0.007)	(0.012)	(0.012)	(0.013)	(0.002)
Wind Storm (lag 1)			−0.067	−0.058	−0.056	−0.047	−0.008
			(0.041)	(0.040)	(0.039)	(0.040)	(0.006)
Thunder Storm 00h00			−0.001	0.004	0.005	0.003	0.002
			(0.060)	(0.061)	(0.062)	(0.054)	(0.008)
Thunder Storm (lag 1)			0.007	0.011	0.011	0.013	−0.005
			(0.013)	(0.007)	(0.008)	(0.025)	(0.007)
Instability at exit point (lag 1)				0.019**	0.020**	0.019*	0.006*
				(0.008)	(0.009)	(0.011)	(0.003)
Instability at exit point (MA14)				0.043	0.045	0.039	0.014
				(0.075)	(0.076)	(0.069)	(0.010)
Youth unemp. in host country (log, lag 5 months)				−0.784**	−0.707*	−0.631*	−0.166***
				(0.334)	(0.362)	(0.345)	(0.060)
Country of origin controls	*No*	*No*	*No*	*No*	*Yes*	*Yes*	*Yes*
Route fixed effects	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Time fixed effects	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
R^2	0.058	0.109	0.111	0.418	0.418	0.412	0.379
Adjusted R^2	0.057	0.108	0.108	0.121	0.12	0.11	0.061
Observations	2,733	2,733	2,733	2,691	2,691	2,691	2,691

Note: *p<0.1; **p<0.05; ***p<0.01

Beck and Katz (1995) Panel Corrected Standard Errors, clustered by route.

vector confirm a 4.444 increase in absolute mortality caused by the introduction of rescue-deterrence policies, while the previous policy attempt still does not reach the 5% significance level.

As discussed earlier, data on missing migrants in the Mediterranean is collected from multiple sources. Since some of these sources are, indeed, rescue vessels, the reduction of search and rescue activities in the Mediterranean route caused by the deterrence policy itself might stand as a potential source of bias. Indeed, it is straightforward that, for each additional NGO ship involved in search and rescue ceasing operations, the probability of deaths in the sea being reported by one of these vessel lowers considerably.

This could pose as a significant source of downward bias, as the number of S&R vessels is correlated with the introduction of the policy, meaning that our estimates will underestimate the effect of the deterrence policy in terms of mortality rates.

These hurdles, however, can be easily overcome. Indeed, I discussed how IOM missing migrants data details the original source of information for each record. In this way, I were able to filter out all casualties reported by S&R vessels, producing an alternate variable $Deaths^{*}_{tr}$ where the post-policy variation in deaths is not caused by the reduction of operational S&R vessels. While total mortality will still be lower, differential rates in mortality before and after the introduction of the policy should not be biased by the introduction of the policy.

Figures for the effect of the effect of deterrence policies on $Deaths^{*}_{tr}$ are presented in column (6). Interestingly, the difference in the policy coefficients between columns (5) and (6) is positive, suggesting that S&R vessels still play an important role in reporting dead or missing individuals, and that most bias concerns can be disregarded.

Finally, a new specification is presented in column (7), this time using relative mortality ($Deaths_{tr}/Attempts_{tr}$) as the dependent variable. Estimates reveal a change in the relative mortality of 0.075% and -0.055% for the policies of Salvini and Minniti, respectively. Also, the magnitude of the attempts coefficients is so small it can effectively be argued that the mortality rate is independent from attempts, as discussed earlier, providing evidence in support of the exogeneity of the relative mortality rate. Reverse causality concerns for the effect of vessels presence on mortality rates persist.

Table 3.6.2 shows results from a logarithmic specification, providing further evidence in support of previous findings. Estimates from table 3.6.2 point at an approximate 36% increase in mortality since the introduction of rescue deterrence policies, which essentially cancels out the reduction in sea mortality obtained during the

Minniti era.[21]

Most importantly, the logarithmic specification allows for the estimation of the η parameter indicating the elasticity of mortality to the number of migration attempts in column (7), as discussed in section 3.4. The non-significance of the η coefficient suggests that the ratio between daily deaths and migration attempts is infinitely elastic to the attempts, keeping absolute mortality fixed – as it is absorbed by the fixed effects – and controlling for all other conditions which might affect migrants' or smugglers' decision to leave a shore.

3.7 Deterrence of mortality

I now turn to the final question addressed in this paper. As discussed, rescue-deterrence policies rest on an a 'realist' stance, which can be broken down to a simple mechanism, according to which: (i) deaths by the sea are announced by the press; (ii) smugglers and/or refugees update their information accordingly; (iii) depending on average risk adversity of these actors, migration decisions change and flows should be affected.

Previous results did not provide any evidence on the channel through which rescue-deterrence can affect migratory patterns. Given that rescue-deterrence implies an increase in the risk of crossing the sea as a deterrent to migration, it should be questioned how changes in mortality can affect decisions to interrupt ongoing migration attempts.

This question is addressed in table 3.7.1, where instrumental variable estimates for the effect of absolute mortality on migration attempts are shown. As discussed in section 3.4, daily reported deaths – the endogenous variable – are instrumented through relative mortality, referring to ratio between deaths and migration. Reduced form estimates showing the direct effect of relative mortality on migration attempts are shown in table 3.A.1 in the appendix.

The identification of the effects of mortality now rests on a different strategy, as the effect of the announcement of deaths on migratory flows does not rely on the counterfactual from the two routes unaffected by the policy. Essentially, the use of relative mortality as an instrumental variable allows to study the effects of absolute mortality on a 'complier' group only, where increases in mortality are caused by random shocks and not by endogenous variation in the number of migration attempts.

[21] Recall, also, that we cannot assess whether the Desert Diplomacy approach led to variations in land mortality as migratory land routes in Africa relocated towards Spain.

Table 3.7.1: Effect of mortality on migration attempts

	Daily migration attempts			
	Total Attempts			
	(1)	(2)	(3)	(4)
#Portichiusi	−20.367	−19.416	−37.940	−42.662
	(23.313)	(22.361)	(32.151)	(28.823)
#Portichiusi (W. Route)	107.224***	107.698***	101.186***	98.499***
	(11.356)	(11.335)	(12.131)	(11.107)
Desert diplomacy	−390.612***	−392.273***	−376.570***	−373.058***
	(16.579)	(18.669)	(13.576)	(15.993)
Desert diplomacy (W. Route)	−2.954	−2.528	−8.196	−11.520
	(10.381)	(10.387)	(11.227)	(10.364)
No of S&R Vessels	7.422	7.655	3.871	3.095
	(5.336)	(5.197)	(6.764)	(6.065)
Reported Deaths (lag 1)	−1.450***	−1.268***	−2.141***	−2.032***
	(0.371)	(0.300)	(0.515)	(0.410)
Reported Deaths (MA3)		−0.522		
		(1.135)		
Reported Deaths (MA7)			4.244**	
			(1.837)	
Reported Deaths (MA14)				5.042***
				(0.774)
Wind Storm 00h00	−33.256***	−33.133***	−34.315***	−34.234***
	(7.442)	(7.445)	(7.301)	(7.080)
Wind Storm (lag 1)	−21.415	−21.264	−22.844	−22.567
	(15.491)	(15.530)	(15.252)	(15.337)
Thunder Storm 00h00	−22.452	−22.743	−19.647	−20.208
	(19.455)	(19.637)	(18.852)	(18.235)
Thunder Storm (lag 1)	−22.961*	−23.069*	−19.947*	−20.536*
	(12.870)	(12.810)	(11.714)	(11.441)
Instability at exit point (lag 1)	−8.814	−8.919	−8.309	−9.111
	(8.610)	(8.676)	(8.010)	(8.523)
Instability at exit point (MA14)	−36.499***	−36.655***	−33.719***	−30.299***
	(4.905)	(5.158)	(2.893)	(3.712)
Youth unemp. in host country (log, lag 5 months)	333.480	336.601	298.914	268.669
	(363.611)	(365.567)	(350.728)	(329.016)
Country of origin controls	*Yes*	*Yes*	*Yes*	*Yes*
Route fixed effects	*Yes*	*Yes*	*Yes*	*Yes*
Time fixed effects	*Yes*	*Yes*	*Yes*	*Yes*
Adjusted R^2	0.172	0.169	0.168	0.169
First Stage (1): Adj. R^2	0.141	0.142	0.144	0.152
First Stage (1): F Test	81.385	86.595	90.746	1925.109
First Stage (2): Adj. R^2		0.205	0.338	0.581
First Stage (2): F Test		203.446	287.186	260.202
Observations	2,691	2,691	2,691	2,691

Note: *p<0.1; **p<0.05; ***p<0.01
Beck and Katz (1995) Panel Corrected SEs, clustered by route.
(1) First stage for endogenous variable: Reported Deaths (lag 1)
(2) First stage for endogenous variable: Reported Deaths (MA3-14)

Table 3.7.2: Effect of mortality on migration attempts (log)

	Daily migration attempts (log)			
	Total Attempts			
	(1)	(2)	(3)	(4)
#Portichiusi	-0.606	-0.600	-0.709	-0.728
	(0.533)	(0.524)	(0.602)	(0.584)
#Portichiusi (W. Route)	-0.005	-0.001	-0.049	-0.080
	(0.131)	(0.121)	(0.140)	(0.141)
Desert diplomacy	-1.152^{***}	-1.159^{***}	-1.125^{***}	-1.108^{***}
	(0.405)	(0.414)	(0.409)	(0.418)
Desert diplomacy (W. Route)	-0.630^{***}	-0.623^{***}	-0.690^{***}	-0.723^{***}
	(0.147)	(0.131)	(0.147)	(0.152)
No of S&R Vessels	0.007	0.010	-0.022	-0.027
	(0.101)	(0.098)	(0.119)	(0.113)
Reported Deaths (lag 1) (log)	-0.103	-0.082	-0.165	-0.152
	(0.127)	(0.122)	(0.157)	(0.128)
Reported Deaths (MA3) (log)		-0.037		
		(0.087)		
Reported Deaths (MA7) (log)			0.175	
			(0.113)	
Reported Deaths (MA14) (log)				0.187^{***}
				(0.059)
Wind Storm 00h00	-0.713^{***}	-0.712^{***}	-0.726^{***}	-0.728^{***}
	(0.158)	(0.157)	(0.161)	(0.156)
Wind Storm (lag 1)	-0.387^{***}	-0.386^{***}	-0.401^{***}	-0.401^{***}
	(0.142)	(0.143)	(0.146)	(0.146)
Thunder Storm 00h00	-0.058	-0.060	-0.040	-0.037
	(0.092)	(0.093)	(0.085)	(0.083)
Thunder Storm (lag 1)	-0.226^{***}	-0.224^{***}	-0.222^{***}	-0.211^{***}
	(0.078)	(0.082)	(0.076)	(0.070)
Instability at exit point (lag 1)	0.033^{**}	0.033^{**}	0.031^{*}	0.028^{*}
	(0.016)	(0.016)	(0.017)	(0.016)
Instability at exit point (MA14)	-0.488^{***}	-0.491^{***}	-0.476^{***}	-0.463^{***}
	(0.096)	(0.090)	(0.104)	(0.105)
Youth unemp. in host country (log, lag 5 months)	2.710	2.762	2.273	1.891
	(2.693)	(2.611)	(2.436)	(2.410)
Country of origin controls	*Yes*	*Yes*	*Yes*	*Yes*
Route fixed effects	*Yes*	*Yes*	*Yes*	*Yes*
Time fixed effects	*Yes*	*Yes*	*Yes*	*Yes*
Adjusted R^2	0.242	0.240	0.242	0.239
First Stage (1): Adj. R^2	0.343	0.344	0.347	0.351
First Stage (1): F Test	614.596	356.021	462.355	4566.619
First Stage (2): Adj. R^2		0.317	0.403	0.65
First Stage (2): F Test		3251.559	122.568	466.95
Observations	2,691	2,691	2,691	2,691

Note: *p<0.1; **p<0.05; ***p<0.01
Beck and Katz (1995) Panel Corrected Standard Errors, clustered by route.
(1) First stage for first endogenous variable - Reported Deaths (lag 1)
(2) First stage for second endogenous variable - Reported Deaths (MA3-14)

The models from table 3.7.1 retain the same co-variates from table 3.5.1, column (6), with only the endogenous variable vector *Deaths* being replaced. Column (1) contains a single one-day lag in said vector. Columns (2) to (4) switch the endogenous variables vector between pairings of one-day lags and varying moving averages for the days preceding t – moving from 3, to 7 and finally 14 days before each attempt. In all cases, all first stage specifications successfully pass the F-tests for excluded instruments, suggesting that relative mortality is indeed a relevant predictor for absolute mortality.

Looking at the main estimates for the effect of mortality, results are now overturned, evidencing that endogeneity concerns for absolute mortality were, indeed, valid. The effect of increases in mortality one day before the migration attempt is turned negative, ranging between a 1.450 and 2.141 reduction in migrants attempting crossing per day. Most importantly, analysis of the moving averages coefficients reveals this to reduction to be explained as temporary near-term displacement. Indeed, no effect is detected within the 3 days window, and the reduction of migration attempts is completely re-absorbed after 7 days, suggesting that refugees do not reverse their migration decision as a consequence of increases in the risk of crossing, but only postpone their departure by a few days.

While these results are only valid for the first 14 days, and are supposed to affect migration decisions already taken (and not initial migration decisions), the fact that the "#Portichiusi" policy coefficient loses all its statistical significance indicates that its remaining components, including, as found earlier, the permanent increase in mortality, are also ineffective in reducing migration flows.

The insignificance of the policy coefficient in the Central Mediterranean indicates that the remaining variation in flows linked with the policy introduction is mostly driven by random shocks, and that the only significant reduction in flows is to be attributed to relocation effects, which might have been, in turn, generated by the previous "desert diplomacy" approach.

The use of a logarithmic specification in table 3.7.2 confirms the estimates from the previous specifications. However, while the sign and magnitude of the coefficients corroborate the relocation effect hypothesis, this time, only the 14 days moving average yields a statistically significant coefficient.

Overall, these results reinforce the view that forced migration is qualitatively different from economic migration. These results can either support previous findings in the literature suggesting that forced migrants are characterised by extremely low levels of risk-aversion, or imply that these individuals often have little to no choice when it comes to decide whether to pursue or not their migration attempt.

3.8 Conclusions

It is widely considered that coordination in refugee migration policies and asylum procedures in the EU has been lacking, leaving the EU destination countries to deal with the crisis by themselves. Persistence of flows, dissatisfaction with the EU framework and political opportunism from national-populist forces have led to an unprecedented strengthening of anti-migrant sentiment in the countries involved. In Italy, the appointment of a new government in June 2018 heralded the beginning of a novel rescue-deterrence stance in migration, which set to reduce migration flows by effectively making the crossing of the Mediterranean riskier.

Based on high-frequency data on migration flows and migrant disappearances in the Mediterranean across three major migration routes, I provide empirical evidence on the effect of rescue-deterrence policies on both flows and mortality rates, and investigate the effect of increases in absolute mortality on decisions to continue ongoing migration attempts.

I find that both absolute and relative mortality in the route towards Italy and Malta have increased by at least 4 deaths per day since the introduction of the policy, and that this increase in mortality has been accompanied by a very modest reduction in flows. Finally, this work shows that forced migration flows do respond to short-term changes in mortality, but that this effect is only temporary, as migration attempts are delayed and reabsorbed within few weeks. Notably, there is also evidence to suggest that permanent increases in mortality are also ineffective. These findings suggest that risk aversion among refugees might be particularly low, or that forced migrants are unable to change their ongoing migration decisions.

In the view of the author, the evidence uncovered by this paper provides a strong point in support of the repeal – or a large revision – of rescue-deterrence practices in migration policy. Signalling the abdication of a responsibility to rescue refugees in distress, while hindering the humanitarian initiatives engaged in search and rescue operation, is not only relatively ineffective as a strategy for flow disruption, but most importantly carries an unacceptable cost in terms of human lives.

Acknowledgements

The author would like to thank Nicolò Fraccaroli, Anzelika Zaiceva, Matteo Villa and Roberto Volpe for their invaluable advice and support. The author also wishes to thank the participants of the BISA Postgraduate Research Workshop on International Migration Politics, and Teodora Tsankova in particular, for their useful comments and suggestions. All remaining errors are mine.

Bibliography

Akgüç, M., Liu, X., Tani, M., and Zimmermann, K. F. (2016). Risk attitudes and migration. *China Economic Review*, 37:166–176.

Beck, N. and Katz, J. N. (1995). What to do (and not to do) with Time-Series Cross-Section Data. *The American Political Science Review*, 89(3):634–647.

Bellemare, M. F., Masaki, T., and Pepinsky, T. B. (2017). Lagged Explanatory Variables and the Estimation of Causal Effect. *The Journal of Politics*, 79(3):949–963.

Bertrand, M., Duflo, E., and Mullainathan, S. (2004). How Much Should We Trust Differences-In-Differences Estimates? *The Quarterly Journal of Economics*, 119(1):249–275.

Bocquého, G., Deschamps, M., Helstroffer, J., and Joxhe, M. (2018). Risk and Refugee Migration. CREA Discussion Paper Series 18-08, Center for Research in Economic Analysis, University of Luxembourg.

Borjas, G. J. (1989). Economic Theory and International Migration. *International Migration Review*, 23(3):457–485. PMID: 12282789.

Brück, T., Dunker, K., Ferguson, N., Meysonnat, A., and Nillesen, E. E. (2018). Determinants and Dynamics of Forced Migration to Europe: Evidence from a 3-D Model of Flows and Stocks. IZA Discussion paper series No. 11834, IZA - Institute of Labor Economics.

Callen, M., Isaqzadeh, M., Long, J. D., and Sprenger, C. (2014). Violence and Risk Preference: Experimental Evidence from Afghanistan. *American Economic Review*, 104(1):123–148.

Cameron, A. C. and Miller, D. L. (2015). A Practitioner's Guide to Cluster-Robust Inference. *Journal of Human Resources*, 50(2):317–372.

Cusumano, E. and Villa, M. (2019). Sea rescue NGOs: a pull factor of irregular migration? Robert Schuman Centre for Advanced Studies Policy Briefs 22, European University Institute.

Deschenes, O. and Moretti, E. (2009). Extreme weather events, mortality, and migration. *The Review of Economics and Statistics*, 91(4):659–681.

Dustmann, C., Fasani, F., Frattini, T., Minale, L., and Schönberg, U. (2017a). On the economics and politics of refugee migration. *Economic Policy*, 32(91):497–550.

Dustmann, C., Fasani, F., Meng, X., and Minale, L. (2017b). Risk Attitudes and Household Migration Decisions. IZA Discussion paper series No. 10603, IZA - Institute of Labor Economics.

Facchini, G., Lorz, O., and Willmann, G. (2006). Asylum seekers in europe: the warm glow of a hot potato. *Journal of Population Economics*, 19(2):411–430.

Hangartner, D., Dinas, E., Marbach, M., Matakos, K., and Xefteris, D. (2019). Does Exposure to the Refugee Crisis Make Natives More Hostile? *American Political Science Review*, 113(2):442–455.

Hatton, T. J., Richter, W. F., and Faini, R. (2004). Seeking asylum in europe. *Economic Policy*, 19(38):5–62.

IOM (2018). Mixed migration flows in the Mediterranean, Compilation of Available Data and Information, June 2018. Report, International Organization for Migration.

Jaeger, D. A., Dohmen, T., Falk, A., Huffman, D., Sunde, U., and Bonin, H. (2010). Direct Evidence on Risk Attitudes and Migration. *The Review of Economics and Statistics*, 92(3):684–689.

Lee, E. S. (1966). A Theory of Migration. *Demography*, 3(1):47–57.

Melander, E. and Öberg, M. (2007). The Threat of Violence and Forced Migration: Geographical Scope Trumps Intensity of Fighting. *Civil Wars*, 9(2):156–173.

Neumayer, E. (2005). Bogus Refugees? The Determinants of Asylum Migration to Western Europe. *International Studies Quarterly*, 49(3):389–409.

Raleigh, C., Linke, A., Hegre, H., and Karlsen, J. (2010). Introducing ACLED: An Armed Conflict Location and Event Dataset. *Journal of Peace Research*, 47(5):651–660.

Rodriguez, C. and Villa, E. (2012). Kidnap risks and migration: evidence from Colombia. *Journal of Population Economics*, 25(3):1139–1164.

Schmeidl, S. (1997). Exploring the Causes of Forced Migration: A Pooled Time-Series Analysis, 1971-1990. *Social Science Quarterly*, 78(2):284–308.

Slovic, P., Monahan, J., and MacGregor, D. G. (2000). Violence risk assessment and risk communication: The effects of using actual cases, providing instruction, and employing probability versus frequency formats. *Law and Human Behavior*, 24(3):271–296.

Tversky, A. and Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157):1124–1131.

Vieider, F. M. (2018). Violence and Risk Preference: Experimental Evidence from Afghanistan: Comment. *American Economic Review*, 108(8):2366–2382.

Vogler, M. and Rotte, R. (2000). The effects of development on migration: Theoretical issues and new empirical evidence. *Journal of Population Economics*, 13(3):485–508.

Appendix

3.A Reduced form regressions

Table 3.A.1: Reduced form effect of relative mortality on migration attempts

	Daily migration attempts			
	Total Attempts			
	(1)	(2)	(3)	(4)
#Portichiusi	−23.522	−23.238	−25.124	−31.007
	(22.498)	(22.275)	(23.631)	(25.142)
#Portichiusi (W. Route)	105.398***	105.406***	105.627***	105.844***
	(11.137)	(11.089)	(11.181)	(11.364)
Desert diplomacy	−385.188***	−385.144***	−385.853***	−386.332***
	(16.802)	(16.727)	(17.297)	(17.899)
Desert diplomacy (W. Route)	−3.731	−3.553	−4.323	−5.444
	(10.651)	(10.627)	(10.602)	(10.860)
No of S&R Vessels	6.556	6.566	6.549	6.260
	(5.122)	(5.104)	(5.181)	(5.405)
Relative Mortality (lag 1)	−30.667***	−27.508***	−33.265***	−35.643***
	(5.872)	(5.034)	(5.654)	(5.724)
Relative Mortality (MA3)		−9.401		
		(9.331)		
Relative Mortality (MA7)			30.886*	
			(17.516)	
Relative Mortality (MA14)				117.955***
				(25.021)
Wind Storm 00h00	−34.068***	−33.955***	−34.100***	−34.148***
	(7.341)	(7.232)	(7.304)	(7.138)
Wind Storm (lag 1)	−20.498	−20.452	−20.816	−20.559
	(14.965)	(14.913)	(15.101)	(14.796)
Thunder Storm 00h00	−20.730	−20.557	−20.680	−20.929
	(20.270)	(20.265)	(20.162)	(20.108)
Thunder Storm (lag 1)	−22.297*	−22.104*	−22.864*	−22.951*
	(12.017)	(11.990)	(12.249)	(12.048)
Instability at exit point (lag 1)	−8.578	−8.558	−8.668	−8.692
	(8.792)	(8.783)	(8.871)	(8.823)
Instability at exit point (MA14)	−34.886***	−34.703***	−35.380***	−35.390***
	(4.021)	(4.020)	(4.352)	(4.377)
Youth unemp. in host country (log, lag 5 months)	313.451	313.874	310.137	310.875
	(358.085)	(357.342)	(359.457)	(360.388)
Country of origin controls	*Yes*	*Yes*	*Yes*	*Yes*
Route fixed effects	*Yes*	*Yes*	*Yes*	*Yes*
Time fixed effects	*Yes*	*Yes*	*Yes*	*Yes*
R^2	0.466	0.466	0.466	0.466
Adjusted R^2	0.191	0.191	0.191	0.191
Observations	2,691	2,691	2,691	2,691

Note: *p<0.1; **p<0.05; ***p<0.01

Beck and Katz (1995) Panel Corrected SEs, clustered by route.

Chapter 4

Mind the wealth gap: a new allocation method to match micro and macro statistics for household wealth

Abstract

The financial and economic crisis recently experienced by many European countries has increased demand for timely, coherent and consistent distributional information for the household sector. In the Euro area, most of the NCBs collect such information through wealth surveys (Eurosystem Household Finance and Consumption Survey, HFCS), which are often used to provide distributional financial accounts and inform monetary policy. These surveys, however, can often suffer from a number of biases, usually caused by differential non-response and under-reporting behaviour, leading to a mismatch with macroeconomic aggregates. In this paper, we develop a novel allocation method which combines distributional information from a power law (Pareto) model with re-weighting and imputation procedures so to address these issues simultaneously, filling the gap with macro statistics and retaining the micro data structure of the survey. Finally, we produce updated distributional indicators for four Euro-Area countries, documenting significant improvements in the measurement of wealth inequality.

Keywords: *Wealth distribution, Non-response, Pareto distributions, Survey*

calibration, Household Finance and Consumption Survey.

JEL codes: D31

4.1 Introduction

The financial and economic crisis recently experienced by many European countries has increased demand for timely, coherent and consistent distributional information for the household sector. The G20 data gap initiative has encouraged the production and dissemination of distributional information on income, consumption, saving, and wealth for the household sector. Eurostat and the European Statistical System have agreed in the "Vienna Memorandum" in 2016 to work towards the same objective as far as consumption, and income are concerned. In recent years, also the measurement of household wealth is becoming increasingly important for policy-making. This is especially the case in societies where job insecurity is growing and where the welfare state is no longer able to ensures acceptable living standards to all individuals. In these circumstances, household wealth becomes an important buffer to guarantee an economic well-being.

The measurement of household wealth is receiving high priority especially in the agenda of national central banks (NCBs). In the Euro area, most of the NCBs collect such information through wealth surveys (Eurosystem Household Finance and Consumption Survey, HFCS). The main advantage of surveys is that they provide distributional information.

Distributional information on wealth is used for instance for financial stability purposes. Overall, it helps have a better understanding of the distributional effects of monetary and fiscal policies, but also allows to study the distribution of risk and debt and to evaluate how much households would be affected by exogenous shocks (see, for instance, Ampudia et al., 2016; Michelangeli and Rampazzi, 2016). Moreover, they help NCBs to have a better understanding of the distributional effects of monetary policy. In the aftermath of the global financial crisis, and following the advent of nonstandard monetary policy measures, a debate started about how monetary policy might affect inequality (Coibion et al., 2017; Colciago et al., 2019). Indeed, a highly accommodating monetary policy, such as quantitative easing (QE), may favor richer households disproportionately, thereby contributing to a more unequal income and wealth distribution.

The use of wealth surveys is not straightforward since they are usually affected by quality issues that need to be addressed. The most glaring one is caused by *differential unit non-response* (Vermeulen, 2018), which happens when the probability of the selected households to participate in the survey is lowered conditionally on other observed or latent factors. This is usually the case for very rich households: as these households tend to concentrate a large share of wealth in their hands, such missing wealth will not show up in survey data, leading to truncation in the survey data, and degrading quality of survey weights as wealth approaches the truncation point.

Second, since wealth surveys include both complex and sensitive items, respondents are not always able or even willing to report the correct value of an item. This measurement error problem can be caused by both item comparability issues – where the survey item is not recording the exact same item used in the corresponding macroeconomic aggregate – or, more commonly, by mis-reporting behavior, which generally differs across the groups of the population or across wealth components. For instance, rich households may be prone to under-report their wealth because of social desirability bias, preference for privacy or, more simply, tax avoidance. Moreover, differential non-response and under-reporting are more common for financial assets (Essig and Winter, 2009), and rich households tend to have larger portfolio shares of these assets.

The ideal solution to overcome these problems would be to match survey data with administrative records (such as tax records or credit registers). Unfortunately, when such data exists and are not limited in scope, they are not usually available to researchers for confidentiality reasons. Because of that, recent literature is developing trying to adjust survey data with the limited external information publicly available such as aggregate totals from national accounts or rich list. Vermeulen (2016, 2018) uses Forbes World Billionaires list in combination with some wealth surveys to estimate a Pareto model. He finds out that the use of the rich list increases the quality of the results (compared to estimating a Pareto model from survey data alone). Building on this approach, Chakraborty et al. (2019); Chakraborty and Waltl (2018); Waltl (2018) extend the analysis by benchmarking survey results to the National Accounts.

This paper belongs to this recent literature which tries to combine survey data with national accounts in order to produce more reliable indicators about the distribution of household wealth. We contribute to the literature in three ways. First, while the existing papers only focus on nonresponse at the tail of the distribution, we present a method to adjust also for underreporting. Dealing with both aspects simultaneously is important even when the interest lies in estimating the wealth held at the top of the distribution. The papers previously mentioned replace households with a net wealth above a given threshold with synthetics data drawn from a Pareto distribution. In addition, they use the information provided by households who have been classified as rich in the survey to estimate the portfolio composition at the top of the distribution. It is clear the failing to correctly select rich households may cause biased results. One difficulty to select rich households is the underreporting behaviour. However, the main advantage of our approach is that it enables us to compute distributional indicators than refers also to "norich" households, such as those relating the financial vulnerability. The second contribution of our method is that it aims at producing a modified database in which survey data are adjusted for the quality issues above mentioned. The existing papers are mainly focused on methods specially tailored to a specific issue, that is the estimation of total wealth held at the top. The final product of our method is to create an adjusted survey that can be used for many alternative indicators. Our third contribution is methodological. Even if we apply methods that are well-established such as imputation and calibration, we show that the two approaches are closely related and we present a new way how they can be jointly used.

The paper is structured as follows: section 4.2 describes the data sources used in our estimation; section 4.3 discusses methodologies for the correction of differential non-response, presenting a novel method for the reallocation of survey weights after fitting the data to a Pareto probability function; section 4.4 addresses the measurement error problem; while section 4.5 shows distributional results for four Euro Area countries, and section 4.6 concludes.

4.2 Data

This paper makes use of one primary source of information – the second round of the Household Finance and Consumption Survey (HFCS) – and two sources of auxiliary information – that is, Financial Accounts and Rich List data.

The HFCS collects household-level information on financial and non-financial portfolios, including assets, liabilities, income and consumption. So far, the survey has been conducted twice, with the first wave having taken place between 2010 and 2011 (the fieldwork reference period varies across states) and the second one between 2013 and 2015.

We focus on the second round of HFCS, but our methods can as well be applied to the cross-section from the first one. We also restrict our analysis to four countries only: Italy, France, Germany and Finland. This choice is motivated by a number of considerations, related to the differences across surveys.

Focusing on these four countries not only allows us to direct our attention to a smaller number of cases from which rich lists are also available, but also to assess the robustness of our method comparatively, focusing on a group of surveys characterised by deep differences both methodological and behavioural in nature. These difference have arisen, for example, from the use of over-sampling both at household (as in the French and Finnish survey) and regional level (as in German one), the use of matched administrative data (as in Finnish one) or by a large truncation in survey data (as in Italian one).

As mentioned, survey weights and reported values from the HFCS are adjusted through the use of auxiliary information. We use Financial Accounts as a prior benchmark, ensuring that the final information contained in the survey correctly sums up to the total wealth in a given economy. Wealth figures extracted from financial accounts refer to the financial instruments which most resemble the wealth items contained in the surveyed portfolios. We use gross wealth as the main variable of interest, but our results can easily be replicated using net wealth wealth instead. To maximise comparability between wealth in the HFCS and in the Financial Accounts, we construct our gross wealth variable as the sum of deposits, bonds, shares, mutual funds, money owed to the household, value of insurances and pension funds, business wealth and housing wealth.

Rich lists have already seen use in the literature, finding use in Vermeulen (2018) and Chakraborty and Waltl (2018) with the aim of finding appropriate estimates of the Pareto tail in the HFCS.

The integration of rich lists in household surveys, however, has generated a number of concerns regarding the reliability of such a source of information. The methodology behind the compilation of said lists is often obscure, and usually only figures for net worth are provided, with no instrument breakdown being given. In contrast, adjustments based on Type II Pareto (recommended by Blanchet et al., 2017, and implemented in Waltl, 2018) provide the advantage of relying on percentile totals obtained from external sources (administrative records such as tax data, as in the case of France), rather than relying on rich lists.

However, this is only possible in those – rare – cases when such information exists and is accessible. When these sources are not available, rich lists might remain a reliable substitute to this kind of auxiliary information, and evidence from Waltl (2018) indicates that, after the integration of rich lists, there might be little difference between the wealth estimated by Pareto I and Pareto II distributions.

In our case, we use wealthy household data from *Forbes' Billionaires List*, substituting this information from larger region-specific lists [1] when available. We also

[1] such as *Challenges' "Les 500 plus grandes fortunes de France"* for France, and the Manager

adjust our rich list data, given that rich lists often report wealth in net terms rather than gross, and usually do not offer detailed information on portfolio composition. This adjustment can be performed in a number of ways: in our case, we first obtain gross wealth by obtaining the liabilities to net wealth ratio from top wealth observations from the HFCS, and then multiplying the inverse of the ratio by the observed net wealth in the rich list. We also sample a number of rich lists observations to have zero liabilities (and adjust gross wealth accordingly) based on the share of top survey observations with no liabilities. Once gross wealth is obtained, a similar procedure is followed to reconstruct instrument specific shares.[2] In this way, estimates for portfolio compositions among top fortunes can be achieved, and rich list data can be fully integrated with the HFCS for estimation purposes.

4.3 The missing wealth

4.3.1 Pareto tail estimation

A relatively small number of missing billionaires can have non-negligible effect on the distributional features of a wealth survey. Wealth distributions have always been characterised by their extreme right skewness, with few individuals at the top of the distribution holding most of the fortunes of an economy.

Most often, the sampling process is unable to capture many of these households. In other instances, the household may be unwilling to respond. What matters is that, when a survey is unable to record these observations in the tail, the entire distribution is affected: as not only the survey will suffer from truncation after a certain level of wealth, but the quality of survey weights will deteriorate as wealth increases (as discussed in: Blanchet et al., 2018; Vermeulen, 2018). The HFCS, as many other wealth surveys, is not immune to this malaise: even in the presence of over-sampling, some of top fortunes might still be missing.

It is well known that these issues emerge when computing the Horvitz Thompson estimator $\sum_{i=1}^{S} p_i w_i$ (the summation of the product of weight p_i and wealth w_i for each individual household i) from the HFCS: indeed, the estimator will only cover a fraction of the total reflected in the Financial Accounts. These issues are usually corrected by proportionally allocating the remaining wealth to all households in the survey. Any attempt, however, in correcting for these mismatches between the FAs

Magazin list for Germany.

[2]We also experimented with Aitchison's 1982 log-ratio transformation to model liabilities and instrument shares in the HFCS survey as a function of wealth, estimate the functional parameters with an estimator of choice (MML or OLS), and then use them to impute shares of households in the rich list.

and survey data without accounting for the missing top wealth is bound to generate biased figures.

Many researchers have then focused on finding the "missing wealthy" and, to do so, most have relied on fitting a Pareto Distribution using external information. The Pareto Distribution is a probability distribution based on a power law, first created by the Italian economist Vilfredo Pareto in his studies on wealth distributions. Its use on HFCS wealth data is well-established, as shown, among the others, in Bach et al. (2019), Chakraborty and Waltl (2018), Eckerstorfer et al. (2016), Vermeulen (2016; 2018), and Waltl (2018).

Essentially, these approaches assume that, over a certain wealth threshold (w_0), the complementary cumulative distribution (CCDF) of wealth (w_i) is approximated by a power law, which (for $w_i \geq w_0$) can be expressed as:

$$P(W \leq w) = 1 - (w_0/w_i)^{\alpha} \tag{4.1}$$

where the parameter $\alpha \in \mathbb{R}_{\geq 0}$ indicates the shape of the tail. The rationale for this approximation is that wealth accumulation is thought to follow a random multiplicative process (RMP), leading to a "fat-tailed" distribution of wealth.[3] With regards to the HFCS survey, the Pareto distribution has notably been used to produce figures for the total wealth above the threshold (w_0), replacing the empirical tail.

A first complication, however, arises when estimating the shape of the tail. If the survey is suffering from truncation, the Pareto tail parameter cannot be estimated from survey data only. Vermeulen (2018) tested a number of methods for estimating the shape parameter α using truncated data, finding linear regression estimates (vis-à-vis maximum likelihood and other methods known in the literature) to yield the most consistent results, if an upper bound was defined. In this case, through the integration of information from the Forbes "rich list" – or similar lists – to "anchor down" the very top of the distribution, the truncation in the CCDF – caused by what the author defines the *differential non-response problem* – can be controlled for, obtaining a closer estimation of the true shape parameter.

Ordering observations by rank i so that $w_1 \geq w_2 ... \geq w_S$, linear estimates for α,

[3]The process of random growth leading to exponential accumulation of wealth has been discussed extensively in the literature: see, for reference, Wold and Whittle (1957), Steindl (1965) and Jones (2015).

after imputing the rich list, can then be obtained through the following specification:[4]

$$ln((i-1/2)\bar{P}_{I_{[j\leq i]}i}/\bar{P}) = C - \alpha ln(w_i) \tag{4.2}$$

Where $\bar{P}_{I_{[j\leq i]}i}/\bar{P}$ adjusts the rank of each individual observation by weighting it by the ratio between the average weight of the first i observations $\bar{P}_{I_{[j\leq i]}i}$ and the average weight $\bar{P}$ in the survey.

As discussed earlier, Chakraborty and Waltl (2018) and Waltl (2018) confirmed this estimator to produce precise and consistent estimates when information on top tail observations is provided, proving as an adequate substitute for Generalised (or Type II) Pareto applications when these cannot be implemented. Rich lists are also typically more easily available than administrative and registry data, which are instead used in the method proposed in Blanchet et al. (2018). For these reasons, thanks to the its ease of use and its parsimonious requirements in terms of auxiliary information, the regression method will be adopted as our preferred approach for Pareto shape parameter estimation.

Other issues relate to the selection of the threshold (w_0). So far, prior literature has adopted the rather arbitrary threshold of 1 million EUR. This is an unnecessary, if not dangerous, assumption. If the arbitrary threshold is above the real one, estimates for the α parameter will suffer from a deterioration in precision, especially if we believe that under-sampling increases with wealth. If the threshold is below the real one, the error in estimation will be even greater, depending on how many non-Pareto distributed observations are now used in the estimation of the shape parameter. Also, different economies are characterized by different distributions of wealth, therefore the draconian adoption of a single threshold for all countries is far from an optimal solution.

As for our estimation, we relax the 1 million EUR threshold assumption by proposing a different – and more objective – method for threshold selection. The proposed method has been used extensively in the literature for threshold estimation in Type II Pareto. As the Type I Pareto – which we use – is still a special case of the Type II Pareto, this method is still valid and can be applied to our model. Equation 4.3 illustrates our approach:

$$E[W - w_0 | W > w_0] = \frac{\sum_{j=1}^{i_{w_i=w_0}} p_j(w_j - w_0)}{\sum_{j=1}^{i} p_j} \tag{4.3}$$

[4]After rearranging and taking logs from the empirical counterpart of equation (4.1): $i/n \cong (w_0/w_i)^\alpha$, and correcting for bias by subtracting 0.5 from the rank of the observation (see Gabaix and Ibragimov, 2011).

Indeed, following from Yang (1978), an useful property of the mean excess function[5] $E[W-w_0|W>w_0]$ is its linearity in w_0 if the distribution is Pareto. Following from this property, Davison and Smith (1990) first proposed a graphical method for threshold detection, meaning that the threshold can be set once the mean excess function becomes approximately linear. This process can be solved numerically by estimating $E[W-w_{0i}|W>w_{0i}]$ for each possible value for w_0, and then selecting the value w_0^* where goodness-of-fit is maximised.

Langousis et al. (2016) confirmed this one to be the most robust method for threshold estimation, provided that the increase in estimation variance is accounted for. In our case, we could assign, for each observed value of w_{0i}, the weight $q_i = i$, meaning that each observation will be weighted by the size of the sample on the right side of the corresponding threshold. As long as the thresholds are ordered by their rank i, choosing the value for w_{0i} where the r-squared of regression is maximised will yield the optimal threshold.

We then adopt a variation of this method, which can be generalised to account for survey weights by assigning the weight: $q_i = (\sum_{j=1}^{i} p_j)$. This is a convenient approach as, assuming that differential non-response bias to increase with wealth, this assumption will allow us to find the threshold by assigning greater weights to observations where this bias is lower.

A similar approach for finding the threshold could entail the property studied in equation 4.2, estimating the α parameter with increasing thresholds, and again selecting the optimal threshold which maximised the model fit. While this method would, essentially, replicate the same process described earlier, we nonetheless rely on the mean excess approach for two reasons: first, its affirmed usage in the literature and, second, the fact that the mean excess approach relies on fitting the conditional mean of reported values, rather than the empirical CCDF, meaning that bias from weights will be reduced. We then believe that, in a survey setting, using the mean excess method will yield a more robust estimate of the Pareto threshold.

Figure 4.3.1 provides a graphical intuition of the output of this process for the four selected countries, showing the estimated threshold and showing, given this threshold, linear fits for the mean excess conditional on wealth. As it appears, this approach provides immediate and tangible benefits over an arbitrary threshold selection: in all cases, the new threshold is found to be lower than 1 million EUR, meaning that subsequent estimates on tail behaviour will significantly benefit in precision.

After the threshold and Pareto shape have been determined, a final choice con-

[5] Also known as the mean residual life function.

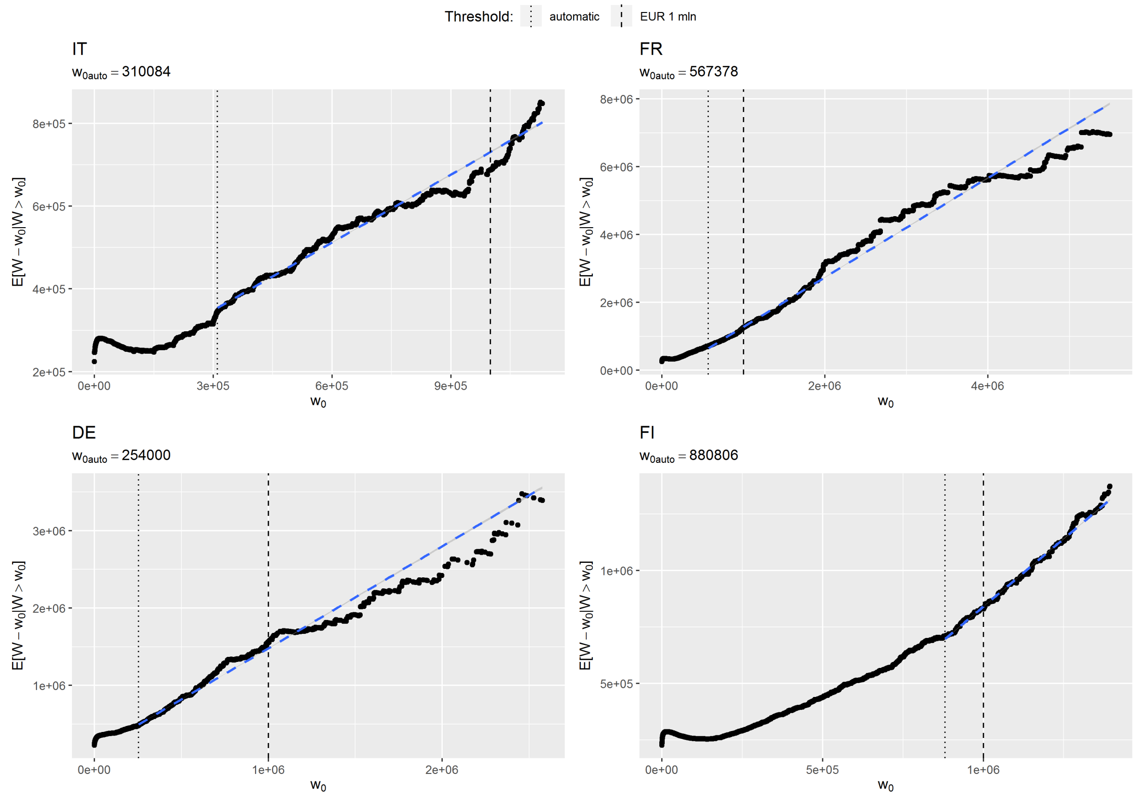


Figure 4.3.1: Mean excess plots for gross recorded wealth in the HFCS. Predicted Pareto thresholds and linear fits estimated using the proposed methodology.

cerns how to use this information.

After a reliable approximation of the shape parameter α has been obtained, total wealth in the top tail of the survey can be quickly re-estimated: this is, undoubtedly, the most common approach. This estimate is usually obtained by multiplying the sum of survey weights after the threshold w_0 by the mean of the Pareto distribution:[6] the resulting figure will yield the total amount of wealth for all households whose wealth exceeds w_0, and the new survey total is obtained by summing this value with the wealth at the bottom of the survey.

While increasing the coverage between the HFCS survey and financial accounts, this approach alone is unable to fill the remaining gap, as this method relies on the implicit assumption that the HFCS adequately samples household wealth for the range of wealth before the distribution starts following the Pareto power law. It may be recalled that, if observations are ranked in order of wealth, an unbiased estimate of the wealth of the lower part of the distribution will be provided by the

[6] The Pareto first moment, which can be found in $\alpha w_0/(\alpha - 1)$ for $\alpha > 1$

Horvitz Thompson estimator, for $i = (1, 2, ..., s_{w_i=w_0}, ..., S)$:

$$t^{HT}_{W_{bot}} = \sum_{i=s+1}^{S} p_i w_i \tag{4.4}$$

However, this also assumes that the sum of the weights of the top tail adequately estimates the number of households in the top tail. Should this be the case, top tail adjustments will only require re-estimating the total amount of wealth at the top of the Pareto distribution, as previous studies have done. There are, however, reasons to move beyond this assumption, as survey weights are rarely calibrated on the basis of household wealth, and these approaches still leave a sizeable gap in the coverage between the HFCS and financial accounts. Also, if we assume a truncated distribution to be following a power law, the re-estimation of the Pareto tail parameter implicitly assumes that the empirical cumulative distribution is suffering from bias precisely because of the truncation, meaning that a number of observations is inevitably missing. For these reasons, top tail adjustments should focus not only at estimating the wealth in the Pareto tail, but also the number of households in it, and should be integrated with adjustments for measurement error.

In this case, then, information on the Pareto tail can be used for the adjustment of individual survey weights. This step raises a two-sided problem, as not only we need to provide consistent estimates for the number of households in the Pareto tail, but we also need to use this information to adjust survey weights – for all survey observations – accordingly.

Indeed, re-weighting can be achieved in a number of ways. In its most simple approach, weights could simply be re-scaled to the new adjusted level. However, this approach would generate one major difficulty which should be dealt with, as the final distribution of demographic characteristics will also be biased towards the features of richest households. This issue can be solved in a single adjustment, through survey calibration, as described in the next subsection.

4.3.2 Methods for adjusting survey data

Let t_W be the population total for variable of interest w to be estimated using survey data. A classical estimator is the Horvitz–Thompson one:

$$t^{HT}_{W} = \sum_{i=1}^{S} p_i w_i \tag{4.5}$$

where p_i is the sampling weight for the i–th household and S is the selected sample.

As discussed, if both micro and macro data were perfect, t_W^{HT} should be equal to t_W, the corresponding macro aggregate. In practice, however, this is seldom the case. In fact, because of nonresponse and underreporting, the value of the aggregated micro data is generally below the macro aggregates. The methods that can be used to adjust survey data to fill such a gap can be grouped into two broad categories.

The first one is the design-based approach, which aims at correcting survey data by modifying the sampling weights p_i through re-weighting methods while keeping the individual responses w_i unchanged (Deville and Särndal, 1992; Särndal, 2007). In the literature, this approach is mainly used: (i) to force consistency of certain survey estimates to known population quantities; (ii) to reduce non-sampling errors such as nonresponse errors and coverage errors; (iii) to improve the precision of estimates (Haziza et al., 2017).

The second, model-based approach aims at adjusting the individual responses collected through the survey w_i while sampling weights p_i are left unchanged. It requires a model for the distribution of the measurement error and auxiliary information to estimate the parameters of the model. Among the several models available in the literature, those most suitable for our purposes are imputation methods. For a general description, see the seminal work of Rubin (1976, 1987). These methods can be used both for the problem of underreporting and for the problem of nonresponse.

The two approaches have some shared traits, so that the distinction is not always clear-cut. For example, the weighting adjustment can also be seen as a method of imputation consisting in compensating for the missing responses by using those of the respondents with the most similar characteristics; in the same way, the imputation of plausible estimates in lieu of respondents' claimed values can be thought as a re-weighting method.

Yet it should be borne in mind that the choice of the method of adjustment is basically driven by three factors. First, it depends on the estimator of interest. For example, if the interest is to estimate the share of total wealth held by rich households, the use of the Pareto method could be sufficient. Second, the choice depends on the magnitude of the gap to fill. If it is considerable, one may want to combine different methods. Finally, the choice depends on the information that is available. If, for example, the only available auxiliary information is in the form of population totals, the design-based approach might be the only feasible one; but if auxiliary data are available at the individual level, then the model-based methods may represent the most effective solution.

In our case, the design-based approach provides a standard approach to the solution of our problem through survey calibration as, for now, we intend to adjust the density of individual observations to account both for the truncation and the decay in quality with increasing wealth.

In the design-based approach, the calibration method for estimating the population total is solved through the following optimisation problem:

$$min \sum_{i=1}^{S} \frac{(p_i^* - p_i)^2}{p_i} \quad s.t. \quad t_Z = \sum_{i=1}^{S} p_i^* z_i \tag{4.6}$$

Where $p_i^* = p_i a_i$. The adjustment factor a_i is a function of the variables used in calibration $z_i = (z_{i1}, z_{i2}, ..., z_{ik})$ and it is computed so that final weights meet benchmark constraints, $\sum_{i=1}^{S} p_i^* z_i = t_Z$, while being, at the same time, kept as close as possible to the initial ones. Benchmark constraints are defined with respect to $t_Z = (t_1, t_2, ..., t_k)$, that is the known vector of population totals or counts of the calibration variables.

The final output is a single new set of weights to be used for all variables. The magnitude of the adjustment factors and, therefore, the variability of the final set of weights is a function of the number of constraints (dimension of the vector t_Z) and the unbalance (difference between the Horvitz-Thompson estimate and the population total). Very variable weights hinder the quality of final estimates for subpopulations and for variables not involved in the calibration procedure. For these reasons, weights are usually required to meet range restrictions such as to be positive and/or within a chosen range.

The method was originally developed to improve the efficiency of the estimators and to ensure coherence with population information, but then it was also largely applied to adjust for nonresponse (Särndal and Lundström, 2005). For example, Little and Vartivarian (2005) show that if the variables used to construct the weights are associated both with non-participation and with the variable of interest, the bias and the variance of the estimator are reduced.

Calibration is a model-assisted procedure: the model implicitly assumed by the procedure is essentially a linear regression between z and t. It does not require an explicit modelling of nonresponse or measurement error: this is implied by the choice of the calibration variables and the functional form chosen for the adjustment factor a_i. The latter can be interpreted as an estimate of the inverse of the response probability of unit i. The vector of auxiliary variables used in calibration usually refers to socio-demographic characteristics but it could also include variables relating households' income or wealth such as the amount of deposits or housing wealth.

The main problem with the use of household balance sheet data in re-weighting methods is that wealth is generally very skewed and concentrated in the hands of a small group of the population that has both low propensity to participate in the survey and different socio-demographic characteristics from the average population. As a consequence, calibration procedures, in order to be used with wealth data, will need to be enhanced by using information on power law behaviour in the tail, as discussed in the next subsection.

4.3.3 Tail households estimation

The Pareto re-weighting adjustment can be paired with calibration methods so as to correct for differential non-response bias.

This can be achieved by including, in the auxiliary variables vector t_z used for calibration, the estimated number of households in and out of the Pareto tail, as well as the estimated wealth held in these two parts of the distribution, along with Horvitz–Thompson estimators for selected demographic indicators.[7] We will refer to this method as 'Pareto-calibration' from now on.

The main intuition is that restraining the calibration with the estimated number of households in the top survey plus their estimated wealth ensures that observations in the top are still Pareto distributed and that the weight decay is accounted for. At the same time, distortions from the original weights are minimised, while statistical representativeness for demographic factors is retained. This procedure only needs to be performed once.

This method, however, requires the number of households, as the wealth these households hold, to be estimated. These estimates be achieved using the methodologies used earlier for the estimation of power laws with truncated data.

Should there be no reasonable doubt to suspect measurement error or misreporting, then the difference between Financial Accounts and survey data can entirely be attributed to the missing wealth at the top of the distribution, and one could subtract the Horvitz Thompson estimator for wealth in the lower part of the survey distribution from the Financial Accounts total, and then divide the resulting value by the Pareto first moment[8] to obtain the number of households to which calibrate the observations in the tail.

Such an estimate, however, would assume no presence of under-reporting and, most importantly, no truncation.

We then pursue a less restrictive method which only provides re-weighting for tail observations, without assuming perfect coverage. In this way, we can preserve the truncation in survey data, but still account for the unobservable rich and adjust

[7] Provided survey data has already been calibrated to these demographic characteristics.
[8] $\alpha w_m/(\alpha - 1)$.

the decay in weight quality amongst the observed points.

Indeed, the larger the gap, the more the decay in the ECDF will be noticeable, as will be seen later in figure 4.3.2. Tail observations can still follow a Pareto distribution which, as discussed, can be retrieved quite consistently using the method from Vermeulen (2018). If the missing households were to be imputed, the logarithm of the complementary cumulative distribution function would be uniformly distributed, and the decay would disappear. While we cannot impute these households as they are, by definition, unobserved, we can still estimate them, and use this information to readjust only the Pareto tail households present in the survey.

A rough approximation of true amount of rich household can easily be obtained by dividing the sum of adjusted survey weights at in the Pareto tail by the value of the Pareto CDF at the maximum recorded survey value:

$$P_T \approx \frac{\sum_{i=s}^{m} p_i}{1 - (w_0/w_m)^\alpha} \tag{4.7}$$

where m indicates the observation with the maximum recorded value, while the Pareto distribution starts at s. The main intuition behind this approach is that, provided the survey data does not suffer from truncation, containing all households, the Glivenko-Cantelli theorem would hold, and the empirical cumulative distribution of wealth would converge to the theoretical one with increasing m.

But, in our case, the following is instead true:

$$\lim_{n\to\infty} \frac{\sum_{i=s}^{m} p_i}{\sum_{i=s}^{N} p_i} - (1 - \frac{w_0}{w_m})^\alpha \geq 0 \tag{4.8}$$

meaning that the empirical CDF will always suffer from a bias equal or larger than zero, the larger the gap. This happens because the denominator $\sum_{i=s}^{N} p_i$, in our case, suffers from truncation, as observations between m and N are unknown. This issue is usually not solved by simply imputing the rich list, as a gap can still persist, as the left column of figure 4.3.2 clearly shows.

However, replacing the $\sum_{i=s}^{N} p_i$ term with P_T, the equality holds, and the equation can be quickly rearranged into equation 4.7. In this case, then, the theoretical PDF will not integrate to one at the maximum wealth recorded in the survey, as the empirical PDF instead does, and estimates for the total number of households in the Pareto tail can be drawn.

Analytically, the estimate from equation 4.7 should be the same for each n-th observation in the tail, replacing w_m with w_n. In practice, with empirical data, variability in survey weights will produce variations. This becomes an especially important problem when we recall that weight quality can deteriorate the closer observed wealth gets to the truncation point. The estimate can then be improved by estimating P_T for each value of wealth over a range of top tail observations, then estimating the mean $\hat{P}_T$ – and producing confidence intervals.

$$\hat{P}_T = \frac{1}{m}\sum_{j=s}^{m} \frac{\sum_{i=s}^{n} p_i}{1-(w_0/w_n)^{\alpha}} \tag{4.9}$$

Improvements can be achieved after weighting each cumulative sum by the inverse rank, or inverse weighted rank (also, recall that observations are ranked in descending order), of the highest observation, so that each weight equals the number of data points, or sample items, it comprises. Other improvements, which we will not explore for the time being, can be achieved using bounded Pareto models, so that the probability density function will integrate to one at the 'true' maximum wealth.

Finally, an estimator for the number of missing, unobserved, households at the right of the tail can be computed as $\hat{P}_{miss} = \hat{P}_T(w_0/w_m)^{\alpha}$. As differential non-response should increase with wealth, the difference between $\hat{P}_T$ and $\hat{P}_{miss}$ can be also larger than the sum of weights in the tail: in this case, weights in the observed tail can be made to conform to this estimate, so that $\hat{P}_{obs} = \hat{P}_T(1-(w_0/w_m)^{\alpha})$, with one single calibration adjustment, as discussed earlier.

In this case, the calibration vectors should then include, other than the demographic totals, the adjusted number of total households in the top and in the bottom part of the distribution separately. It is straightforward that the adjusted number of households in the bottom part of the survey can be obtained by subtracting from the sum of weights of observations below the threshold the difference between the sum of weights in the tail and $\hat{P}_T$.

As we do not to yet intend to adjust for under-reporting and item comparability, but we also want to retain representativeness of the survey in terms of household wealth below the Pareto threshold, calibration benchmarks for these households can also be set to the initial Horvitz-Thompson estimator for total wealth across these households.

The calculation of calibration benchmarks for observable Pareto tail wealth is, instead, less trivial. To obtain an estimate, we first compute expected wealth in the tail given the shape parameter α and the total number of households in the tail $\hat{P}_T$. This estimate is obtained, as in Vermeulen (2018) and Chakraborty and Waltl (2018), by multiplying $\hat{P}_T$ by the mean of the Pareto distribution $\alpha w_0/(\alpha-1)$. From this product, we then subtract Pareto tail wealth after the truncation point can, then, be grossed up again by multiplying $\hat{P}_{miss}$ by the Pareto first moment

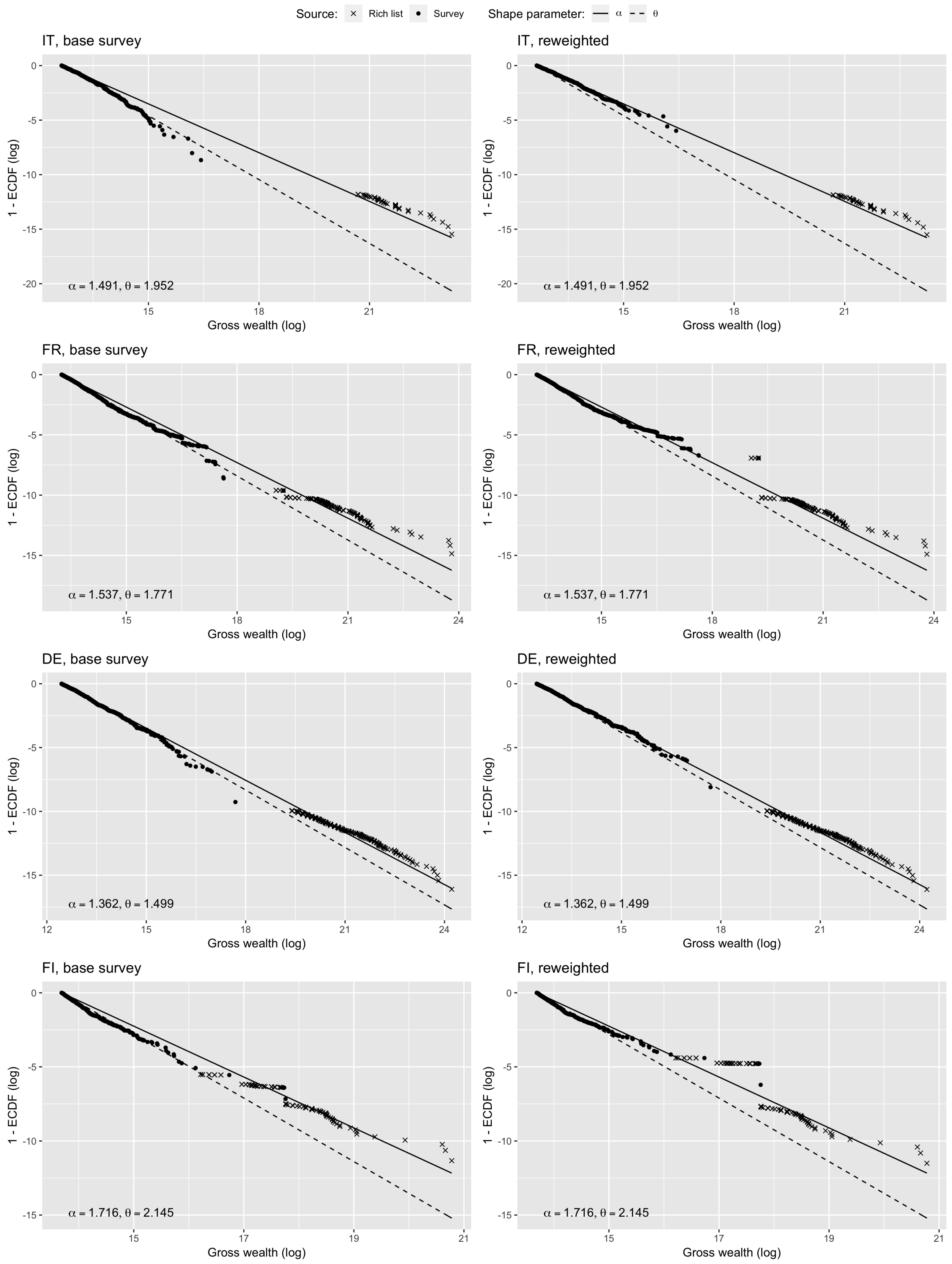


Figure 4.3.2: Empirical cumulative distribution functions (log scale) for survey wealth distributions in the Pareto Tail. Re-weighting achieved by using the Pareto calibration method, using the calibration vectors from equations 4.11 and 4.12.

$\alpha w_m/(\alpha-1)$, setting the new threshold at the maximum recorded wealth in the survey. This is possible because the Pareto shape parameter does not change along the Pareto distribution. Equation 4.10 simplifies this process, producing an estimator for observed tail wealth to which calibrate the survey:

$$\hat{t}_{obs} = \frac{\alpha \hat{P}_T(w_0 - w_m(w_0/w_m)^\alpha)}{(\alpha - 1)} \tag{4.10}$$

As mentioned earlier, calibrating initial survey weights to the adjusted total observed wealth and the adjusted observed weights means that we can force survey weights to conform to a Pareto distribution without also allocating the wealth held by unobserved households to the observed households.

The final vectors calibration benchmarks vector will then be:

$$t_Z = (\hat{t}_{obs}, \hat{P}_{obs}, t_B^{HT}, \hat{P}_B, t_X^{HT}) \tag{4.11}$$

$$z_i = (w_{iI_{[w_i \geq w_0]}}, I_{[w_i \geq w_0]}, w_{iI_{[w_i < w_0]}}, I_{[w_i < w_0]}, x_i) \tag{4.12}$$

where $\hat{P}_{obs}$ is the estimated number of observed households in the Pareto tail, and $I_{[w_i \geq w_0]}$ is an indicator variable for observations in the tail, as $\hat{P}_B$ and $I_{[w_i < w_0]}$ are for the observations not in the tail; $\hat{t}_{obs}$ is the estimated observable wealth in the Pareto tail, and $w_{iI_{[w_i \geq w_0]}}$ are individual reported values for observed wealth in the Pareto tail; t_B^{HT} is a vector of Horvitz-Thompson estimators decomposing the initial wealth of observations below the threshold into their corresponding portfolio items, and $w_{iI_{[w_i < w_0]}}$ is a vector of these observed portfolio items; and similarly t_X^{HT} is a vector of Horvitz-Thompson estimators for demographic totals, while individual demographic values are contained in the vector x_i.

After calibration, total 'implied' wealth is then obtained by summing the estimate for total wealth after truncation to the Horvitz-Thompson estimator of wealth for all observations in the survey, after calibration:

$$\hat{t}_W = \frac{\alpha w_m \hat{P}_{miss}}{(\alpha - 1)} + \sum_{i=1}^{S} p_i^* w_i \tag{4.13}$$

Provided that the we have a good approximation of wealth distribution in the tail, remaining differences in coverage between the estimate obtained in equation 4.13 and the Financial Accounts will then be left to under-reporting or comparability issues. In this case, under-reporting adjustments should only focus on reallocating this remaining difference.

Figure 4.3.2 illustrates the outcome of this process, showing the empirical CCDF on a log-log scale before, and after, the adjustment. The graphs, essentially, provide a graphical intuition for the linear relationship described earlier in equation 4.2. Re-weighted figures are produced by using the proposed Pareto-calibration method. α indicates the Pareto shape parameter estimated by imputing the rich list, while θ shows these estimation results with survey data only.

The benefits from this process are immediately clear: survey weights are adjusted to account for the distribution of wealth, so that non-response bias is corrected, but their variance is also preserved. At the same time, consistency with other demographic parameters is maintained too.

Table 4.3.1 shows coverage ratios between survey wealth estimates and Financial Accounts. Column (1) shows initial coverage ratios, while column (2) displays coverage ratio for adjusted data, and column (3) grosses up survey wealth by first estimating the number of rich households in the tail, and then by estimating total implied wealth in the tail, as described. Columns (4) and (5) show the estimated number of households in the Pareto tail, along with the number of "missing rich".

Table 4.3.1: The missing gap: Pareto adjustments for gross wealth

	Coverage Ratios			Estimated tail households		
	Base	Adjusted	Implied	Total	95% C.I.	Missing
Country	(1)	(2)	(3)	(4)	(5)	(6)
IT	0.553	0.647	0.727	5,483,837	± 243.406	19366.110
FR	0.673	0.762	0.779	3,003,389	± 60.255	289.339
DE	0.827	0.908	1.039	9,889,923	±463.185	7913.568
FI	0.917	1.019	1.059	107,881	±31.047	92.065

Notes: Coverage ratios and estimated number of households in the tail. Re-weighting achieved by using the Pareto calibration method, using the calibration vectors from equations 4.11 and 4.12.

Results from Table 4.3.1 show how the Pareto-calibration procedure can significantly improve the coverage between surveys and Financial Accounts.

In the case of Finland, where not only respondents' assets are matched with registry data – so that no under-reporting can be expected – but oversampling has been employed, the gap is filled immediately by Pareto-reweighting. Missing rich are minimised, and the increase between Pareto-adjusted and implied survey wealth is negligible, and should only be attributed to the superimposition between survey and rich list data.

We also see significant improvements when differential non-reporting becomes a particularly large issue. This is the case of Italy and Germany, where the absence of over-sampling, in the former case, and the use of regional over-sampling, in the latter one, lead to a particularly large number of missing households in the very top. Interestingly, implied wealth figures for Germany point again at a complete coverage between survey and Financial Accounts, meaning that there are few reasons to suspect misreporting, even in the lack of matching with registry data.

The same cannot be said for Italy, where estimates for implied wealth only explain less than half of the remaining gap – even if coverage has significantly improved from the paltry 55% from the base survey. In this case, the remaining gap is most likely explained by under-reporting behaviour.

Finally, in the French survey, the presence of over-sampling points at an extremely small number of missing observations; however, the persistence of a mismatch between survey data and Financial Accounts points again at the presence of under-reporting.

Overall, these figures suggest that the proposed Pareto-calibration approach can produce substantial improvements in survey coverage. While the need for other adjustments for the correction of under-reporting and the improvement of item comparability will vary on a case-by-case basis, these adjustments can now be performed without needing to worry about non-response bias. These adjustments will be discussed in the next sections.

4.4 Allocating the remaining wealth

Survey calibration, when combined with information from the Pareto distribution, can then stand as a powerful tool for correctly reallocating the household wealth distribution in surveys suffering from truncation. However, as discussed, the Pareto-calibration methodology will not always successfully fill the remaining gap between survey data and macroeconomic aggregates. In these cases, survey data can be further adjusted using imputation methods in conjunction with information from the Pareto distribution.

This section provides an overview of the methods that could be used to adjust survey data for measument errors caused by under-reporting behavior and, to a lesser degree, item comparability, while ensuring that the adjustments obtained through Pareto-calibration are correctly taken into account.

The proposed allocation methods will produce estimates that are in line with the totals from National Accounts, adjusting for under-reporting only across the observed distribution, maintaining consistency with the Pareto estimates. Our aim is to show under which conditions these methods will retain this consistency, and not to provide a conclusive answer on how to solve the measurement error problem.

Indeed, if most of these differences can be attributed to under-reporting behaviour, imputation models should be justified by behavioural theory and evidence, and adjustment of such nature could be implemented before implementing the Pareto-calibration adjustment. Such adjustments are, unfortunately, not possible until further auxiliary sources of information have been made available.

This being considered, we can make an educate guess, and assume that (1) measurement error between rich list and survey data is comparable and (2) that relative error is indifferent to wealth, at least among the very rich. In this case, we can assume the relative error to converge in probability to a constant, which we will denote $(w^* - w)/w \xrightarrow{p} \zeta$, so that, on average, the unobserved 'true' total wealth will be given by $\hat{w}_i^* = \zeta w_i$, provided that $\zeta \perp w$.

Thanks to Slutsky's theorem, survey wealth would still be Pareto distributed after adjusting for measurement error. As it follows, total wealth in the survey would scale up to $\zeta t_W^{HT} = \sum_{i=1}^{S} \zeta p_i w_i$, and the Pareto CDF would turn into $F_\alpha(\zeta w_i) = 1 - (\zeta w_0/\zeta w_i)^\alpha$. Simplifying this last formula and updating equation 4.13 for measurement error, we obtain an the following estimate for total wealth:

$$\zeta \hat{t}_W = \zeta(\frac{\alpha w_m \hat{P}_{miss}}{(\alpha - 1)} + \sum_{i=1}^{S} p_i w_i) \tag{4.14}$$

Meaning that our Pareto-adjusted estimate $\hat{t}_W$ is indifferent to the scaling of the variables. Should this be the case, the coefficient for Pareto-adjusted coverage ratio, given the Financial Accounts total wealth, as in $\zeta = t_Z/\hat{t}_W$, will yield the scalar to which re-allocate reported survey wealth.

A breakdown of adjusted coverage ratios by portfolio item can also be provided, but requires portfolio compositions in the unobserved Pareto tail to be inferred from observed, and Pareto-adjusted, survey data. This is not a trivial exercise but, for simplicity, we will assume that the portfolio item composition over total wealth in the observed Pareto tail provides a satisfying approximation of the unobserved instrument composition in the unobserved part.

In this case, equation 4.14 changes into:

$$t_Z = \theta(\frac{\alpha w_m \hat{P}_{miss}}{(\alpha - 1)} s_{obs} + \sum_{i=1}^{S} p_i z_i) \tag{4.15}$$

where $z_i = (z_{i1}, z_{i2}, ..., z_{ik})$ is a vector of portfolio items, $t_Z = (t_1, t_2, ..., t_k)$ refers to the macroeconomic aggregate of these items, $\theta = (\theta_1, \theta_2, ..., \theta_k)$ is a vector of coverage coefficients for each item, and $s_{obs} = (s_{obs,1}, s_{obs,2}, ..., s_{obs,k})$ is a vector indicating the composition of portfolio items as a share of total wealth in the observed Pareto tail.

Reallocation of wealth can then be achieved in two ways, as long as we impose the orthogonality restraint between wealth and relative error, and we take into account these adjusted coverage ratios.

Proportional allocation is the standard approach. In this case, we would allocate the remaining gap by multiplying each z_i by the respective θ. This process assumes (i) measurement error relative to wealth to be constant across all households, and (ii) this same error to differ across all portfolio items. In this way, the adjustment factor will be item-specific, but each household would receive the same adjustment for each portfolio item, relative to its initial value. As a result, the gap between financial accounts and survey data will be filled, and each portfolio item in the survey would be allocated so to match the corresponding total from the financial accounts.

However, not all households under-report by the same degree, and, most importantly, there is no reason to assume item-by-item relative differences in mis-reporting behaviour from each household. Also, reallocating each instrument to match its macro-economic aggregate is bound to create distortions in the wealth distribution once again.

The aforementioned methods used in survey calibration can then prove useful in providing an alternative to proportional allocation, which would instead provide a single household-specific adjustment for all portfolio items.

Recall that survey calibration methods find the optimal adjustment factor a_i which minimises the quadratic distortion of new weights relative to prior ones. Nothing prevents us from changing the minimization problem so that the adjustment factor a_i is minimised with respect to a quadratic loss function for reported wealth values.

$$min \sum_{i=1}^{S} \frac{(a_i w_i - w_i)^2}{w_i} \quad s.t. \quad \zeta \sum_{i=1}^{S} p_i w_i = \sum_{i=1}^{S} p_i a_i w_i \tag{4.16}$$

This guarantees the adjustment factor a_i to average to ζ while also remaining independent from wealth, as the function will attempt to minimise distortions relative to the initial reported value. Within-household portfolio compositions is also retained and, at the same time, only the amount of missing wealth unexplained by the Pareto-calibration allocation, and its implied unobserved wealth, is allocated.

The main drawback from this approach is the reduced dimensionality of the calibration vector, as portfolio items have to be summed together for the optimisation to succeed. This is a negligible issue, at this stage, and we leave to future research the task to further explore the possibility of adapting this calibration allocation to the multivariate case.

As a result, the difference between financial accounts and the adjusted survey data, corrected for the missing tail[9], will approximate zero, providing a rough correction for most measurement error issues.

As this method will reallocate all portfolio items by the same household-specific factor, it is still possible, after the adjustments, to find divergences between the Horvitz-Thompson estimators of each portfolio item and their corresponding macroeconomic aggregate. In this case, these difference will indicate the presence of more profound issues, arising either from the sampling process or from the low comparability between certain portfolio items and their macro counterparts.

It is well possible for the survey to be representative in terms of distribution of wealth, but see this quality deteriorate once we the coverage of specific portfolio items. This is possible when considering that differential non-response might also affect the sampling of holders of specific portfolio items: some households might be, for example, more or less inclined to reveal information on their wealth depending on their asset composition. These issues are highlighted in those instances where linkage with registry data ensures that measurement error is essentially non-existent, as it is the case for Finland.

In these cases, as we have seen, it is not uncommon for the coverage ratio to approach unity once we control for the unobserved Pareto tail, while the coverage for individual portfolio items might differ significantly, sometimes being lower, sometimes higher, than one. The persistence of differences in coverage between portfolio items after the Pareto adjustment, when the gap between total survey wealth and the macroeconomic aggregate approaches zero, and especially when individual responses have no misreporting error, implies that differential non-response can also be conditioned by participation into financial markets.

In these instances, then, these remaining differences should be attributed to the sampling process. A 'last-resort' allocation adjustment can then be performed by calibrating the weights once more, using the same procedure described in section 4.3, ensuring that the portfolio items sum up to $t_Z(\theta^T)^{-1}$, and that the wealth that has been allocated through Pareto-calibration – along with the usual demographic characteristics – remains fixed. This procedure requires us to assume that the estimate for the wealth in the Pareto tail we have obtained is correct, and most

[9] An adjusted that, we recall, can easily be obtained using equation 4.13, and replacing w_m with ζw_m and w_i with $w_i a_i$.

importantly that this estimate has not been influenced by the sampling error for individual portfolio items.

As a result, the survey estimate for individual portfolio items will approach unity once the unobserved Pareto tail is taken into account, but the estimated distribution of wealth will not be affected by the reallocation of these items.[10]

Conceptual comparability issues across Financial Account instruments and portfolio items in the HFCS may also require this calibration method to be performed separately on separate instrument groupings. In our case, we grouped instruments based on their comparability (using the comparability scale provided by Chakraborty and Waltl, 2018), and performed this item allocation on financial assets only, as business and housing wealth recorded in the survey are not considered to be perfectly comparable with their macroeconomic aggregates from the Financial Accounts. Then, assuming that measurement error is constant within the financial assets grouping, this expedient ensures that the adjustment will not be biased by the low comparability of specific items.

4.5 Results

Table 4.5.1 shows distributional results indicating the share of gross wealth held by the top 1, 5, 10, and 20 weighted percentiles, along with the bottom 50%. Weighted Gini inequality indices are also presented in column (6), while column (7) provides the estimated Pareto tail parameter α given the data. These figures are reproduced under each allocation method.

The first set of rows ('Base Survey') presents distributional figures from the un-adjusted HFCS data. As it is well known, truncation in top wealth distribution can cause survey estimates to understate the true level of wealth inequality, and the figures presented in the table provide support for this possibility. Indeed, estimates from the un-adjusted HFCS would suggest wealth inequality in Italy, whose survey suffers from the largest truncation, to be very close to the inequality level in Finland, where the truncation is minimised.

Column (7) displays the Pareto tail coefficients. In the first set of rows, the α parameter is estimated using survey data only, meaning that this is the Pareto estimate that survey data yields when truncation is not corrected through the imputation of a rich list.[11]

[10]Recall that, by design, the calibration equation is enforcing total wealth and the sum of weights in the Pareto tail to remain fixed to the prior estimates obtained in section 4.3, ensuring that survey data remains Pareto distributed, with tail shape parameter α.

[11]These are the same tail parameters denoted as θ in Figure 4.3.2.

Table 4.5.1: The missing gap: distributional wealth indicators

	Wealth Shares						
	Top 1%	Top 5%	Top 10%	Top 20%	Bot 50%	Gini	Tail α
Country	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				Base Survey			
IT	0.114	0.293	0.424	0.598	0.103	0.597	1.952
FR	0.178	0.361	0.494	0.664	0.069	0.663	1.771
DE	0.226	0.443	0.575	0.744	0.030	0.741	1.499
FI	0.120	0.286	0.416	0.595	0.101	0.596	2.145
				Survey & missing tail			
IT	0.221	0.379	0.495	0.648	0.091	0.647	1.491
FR	0.198	0.377	0.507	0.672	0.067	0.671	1.537
DE	0.329	0.520	0.634	0.779	0.026	0.776	1.362
FI	0.137	0.299	0.427	0.602	0.099	0.604	1.717
				Pareto-calibration & missing tail			
IT	0.266	0.431	0.540	0.682	0.084	0.677	1.456
FR	0.274	0.438	0.556	0.705	0.061	0.704	1.500
DE	0.360	0.548	0.657	0.793	0.025	0.789	1.365
FI	0.192	0.351	0.472	0.635	0.091	0.634	1.602
				Par-cal, Proportional allocation & missing tail			
IT	0.299	0.470	0.581	0.715	0.077	0.716	1.427
FR	0.239	0.408	0.531	0.687	0.070	0.699	1.558
DE	0.337	0.517	0.632	0.774	0.033	0.734	1.376
FI	0.270	0.435	0.546	0.689	0.081	0.867	1.541
				Par-cal, Wealth calibration & missing tail			
IT	0.300	0.465	0.569	0.702	0.078	0.674	1.442
FR	0.264	0.428	0.548	0.698	0.063	0.714	1.534
DE	0.357	0.549	0.658	0.795	0.024	0.761	1.365
FI	0.189	0.349	0.471	0.636	0.090	0.635	1.584
				Par-cal, Wealth/portfolio calibration, & missing tail			
IT	0.303	0.470	0.580	0.712	0.073	0.707	1.432
FR	0.244	0.422	0.543	0.695	0.064	0.693	1.511
DE	0.362	0.549	0.658	0.794	0.025	0.789	1.365
FI	0.143	0.312	0.440	0.613	0.100	0.610	1.781

Notes: Wealth share by percentile, gini inequality coefficients and Pareto tail parameters for Italy, France, Germany and Finland, estimated using different adjustments for the HFCS data, and accounting for the unobserved part of the Pareto tail.

For all following sets of rows, the unobserved part of the Pareto tail is included when computing all distributional indicators. To do so, these missing households are imputed as a single observation whose weight and wealth equals, respectively, the estimated number of unobserved households and the estimated average wealth in the unobserved Pareto tail, both obtained using the methods presented in section 4.3.

The second second set of rows ('Survey & missing tail') displays estimates produced using the un-adjusted survey data, plus the missing tail households. Depending on the size of the truncation in the Pareto tail, inequality estimates can be affected considerably. For surveys, such as the Italian and German ones, where truncation bias is particularly pronounced, the sole inclusion of these unobserved households increases the share of wealth held by the top 1% households by at least 10.7 and 10.3 percentage points, respectively. This increase is much less pronounced for the French and Finnish survey, where the truncation is also much more modest.

The inclusion of the unobserved tail raises inequality levels for all surveys considered, but again these increases are proportional to the size of the truncation. Finally, estimates for the Pareto tail parameter are now corrected for the truncation by imputing the rich list and using the estimation procedure described in section 4.3. These are the same parameters earlier shown in Figure 4.3.2.

Survey weights are then adjusted using the proposed Pareto-calibration method to produce the figures shown in the third set of rows ('Pareto-calibration & missing tail'). After this adjustment, between-country differences across distributional indicators start to reduce. This time, an increase in inequality, while less remarkable than in the previous step, can still be noted across all surveys. This increase suggests, should there be a reason to suspect survey weights to degrade due to differential non-response, that the proposed adjustment can play an important contribution in the measurement of inequality through the adjustment of existing survey data points.

Notably, the estimated α parameters are not particularly affected by the adjustment, indicating that the combination of survey calibration with Pareto estimates still distributes survey data in accordance with the estimated Pareto distribution.

However, this parameter changes the most after adjusting the Finnish survey. This phenomenon is easily explained. In this case, the superimposition between survey and rich list data can cause some problems because some rich list observations appear twice, creating a small loss in consistency in the estimation of the Pareto tail parameter. A solution, which we did not pursue for the time being, would be to cut the rich list before the truncation point, ensuring that no observations are repeated twice. In any case, the Finnish survey is only marginally affected by the reallocation of weights: the estimated inequality remains mostly stable across all iterations of the adjustments, indicating that, when sampling and mis-reporting error are minimised, the proposed method reallocates wealth parsimoniously.

The row sets from fourth to sixth adjust the survey for measurement error, applying the procedures described in section 4.4. For countries like Finland and Germany, where mis-reporting can be considered to be negligible issue, these adjustments might not be needed, and remaining divergences in portfolio item coverage against macroeconomic aggregates should be treated as sampling issues and adjusted through weight calibration, as detailed in section 4.4, and shown in the last set of rows in Table 4.5.1.

In the fourth set of rows ('Par-cal, proportional allocation & missing tail'), portfolio items are scaled proportionally to the Financial Accounts aggregates. Proportional allocation, however, seems like an inadequate solution to the measurement error problem. While proportionally allocated items do not generate severe distortions in the estimated Pareto distribution, proportional allocation will most likely affect the portfolio allocation within each household. In cases such as the Finnish one, where individual survey responses are matched with registry data, and where the low coverage of certain items can only be attributed to sampling issues, proportional allocation generates excessive and unnecessary variation in the estimates.

Then, a much more conservative approach would require to use the calibration methodologies described in section 4.4 to reallocate total wealth, leaving portfolio compositions intact, and then dealing with sampling issues separately. The fourth ('Par-cal, Wealth calibration, & missing tail') and fifth ('Par-cal, Wealth/portfolio calibration, & missing tail') sets of rows show how distributional figures are affected by this approach.

In both cases, significant differences over proportional allocation can be noted. First of all, the Pareto tail parameter is always closer to the initial estimate, meaning that the reallocation process, this time, leaves the distributional features of the survey intact. Secondly, inequality figures appear much more similar to the estimates produces in the previous steps. Indeed, the final output shows comparable results across all surveys, where the increases in inequality, compared to the initial survey data, are proportional to the severity of both truncation and mis-reporting problems. Most importantly, the α parameter is still close enough to the one estimated initially, suggesting, once again, that both adjustment do not give rise to unnecessary distortions in the tail wealth distribution.

While wealth calibration should not be treated as a substitute for proper models for adjusting for measurement error, especially when this error is linked to socio-economic or behavioural factors, these calibration-based methods can still assist in the production of distributional figures without exposing the researcher to the risk of misrepresenting the distribution of household wealth and individual asset compositions.

Also, the use of portfolio calibration (as in the final set of rows) can definitely help when measurement error is supposed to be null (Finland, and Germany to a lesser degree), and when models have been developed to correctly address mis-reporting behavior. In these cases, the wealth calibration step can be skipped entirely, while the portfolio calibration can be paired with Pareto-calibration within the same step, so that the weighted sum of each portfolio item is kept consistent with the corresponding macro-economic aggregate, producing consistent and correct distributional figures.

4.6 Conclusions

In this paper, we have shown how a combination of well-established methodologies for the fitting of a Pareto distribution and the calibration of survey data can be used in order to adjust wealth data from household surveys for the correction of differential non-response and misreporting, fully retaining the micro data structure of the survey, and filling an important gap in the literature.

We have applied these methods to the HFCS data, using the 2014 Finnish, French, German, and Italian surveys, and employing, as auxiliary sources of information, rich list data from Forbes or national press sources, along with household sector aggregates from the Financial Accounts.

We have shown how these adjustments can play a particularly important role in the production of correct distributional financial accounts for the household sector, suggesting that inequality estimates from the original survey data can widely understate the population parameters, depending on the severity of both differential non-response problem and measurement error.

Based on evidence from prior literature, we believe our Pareto estimates to yield an acceptable approximation of the true wealth distribution, given the quality of the auxiliary sources of information. Our method ultimately relies on the assumption that wealth is correctly sampled for the most part of the distribution, but that sampling quality degrades as the wealth approaches a truncation point.

While we have provided a much needed template for the use of Pareto tail estimation methodologies in conjunction with survey calibration methodologies, further work is needed for the refinement of this approach. In particular, the estimation process for the number of households in the tail can be further validated and improved, enhancing its robustness to the degradation in quality of survey weights in the top tail of the Pareto distribution. Also, alternatives to rich lists should be employed whenever possible. The reallocation of survey weights through calibration can also be further improved by enriching the auxiliary variables vector with more information.

The correction of measurement error requires a set of assumptions that are, eventually, more challenging to defend. Most importantly, we assume relative measurement error to be constant and thus independent from wealth. Relaxing this assumption has major implications for the wealth distribution, and thus further work should be pursued for the development robust models for misreporting based on economic and behavioural theory.

Nonetheless, we have provided a general framework for the reallocation of weights that is – potentially – fully compatible with these models. Should the measurement error problem be corrected separately, then differential non-response – for both total wealth and specific portfolio items – can be dealt with using the methodologies we have presented.

Other than calling for further work on measurement error, our contribution opens the opportunity for pursuing further analytical work on household finance using the adjusted micro data-set. As a matter of fact, most studies on survey data corrections based on Pareto models have focused on providing aggregate estimates for the missing wealth, without looking at their implications for statistical inference.

Indeed, without these adjustments, not only distributional financial accounts can be miscalculated, but econometric models may also end up providing faulty estimates when frequency weights used in the estimation are affected by differential non-response bias. There is no guarantee that the covariates used in multivariate models are able to adequately control for these distributional issues, especially when the explanatory variable is wealth itself. Thanks to our contribution, it now possible to reassess whether prior findings from applied research are robust to differential non-response bias, and provide corrections in case they are not.

Acknowledgements

The authors would like to thank all members of the EG-LMM and EG-DFA working groups for their advice and suggestions. All remaining errors are ours.

Bibliography

Aitchison, J. (1982). The statistical analysis of compositional data. *Journal of the Royal Statistical Society. Series B (Methodological)*, 44(2):139–177.

Ampudia, M., van Vlokhoven, H., and Żochowski, D. (2016). Financial fragility of euro area households. *Journal of Financial Stability*, 27:250 – 262.

Bach, S., Thiemann, A., and Zucco, A. (2019). Looking for the missing rich: tracing the top tail of the wealth distribution. *International Tax and Public Finance*, 26(6):1234–1258.

Blanchet, T., Flores, I., and Morgan, M. (2018). The weight of the rich: Improving surveys using tax data. *WID.world WORKING PAPER SERIES N° 2018/12*.

Blanchet, T., Fournier, J., and Piketty, T. (2017). Generalized pareto curves: theory and applications. *WID.world WORKING PAPER SERIES N° 2017/3*.

Chakraborty, R., Kavonius, I. K., Pérez-Duarte, S., and Vermeulen, P. (2019). Is the top tail of the wealth distribution the missing link between the household finance and consumption survey and national accounts? *Journal of Official Statistics*, 35(1):31 – 65.

Chakraborty, R. and Waltl, S. R. (2018). Missing the wealthy in the HFCS: micro problems with macro implications. ECB Working Paper Series No 2163, European Central Bank.

Coibion, O., Gorodnichenko, Y., Kueng, L., and Silvia, J. (2017). Innocent bystanders? monetary policy and inequality. *Journal of Monetary Economics*, 88:70 – 89.

Colciago, A., Samarina, A., and de Haan, J. (2019). Central bank policies and income and wealth inequality: A survey. *Journal of Economic Surveys*, 0(0).

Davison, A. C. and Smith, R. L. (1990). Models for exceedances over high thresholds. *Journal of the Royal Statistical Society. Series B (Methodological)*, 52(3):393–442.

Deville, J.-C. and Särndal, C.-E. (1992). Calibration estimators in survey sampling. *Journal of the American Statistical Association*, 87(418):376–382.

Eckerstorfer, P., Halak, J., Kapeller, J., Schütz, B., Springholz, F., and Wildauer, R. (2016). Correcting for the missing rich: An application to wealth survey data. *Review of Income and Wealth*, 62(4):605–627.

Essig, L. and Winter, J. K. (2009). Item non-response to financial questions in household surveys: An experimental study of interviewer and mode effects. *Fiscal Studies*, 30(3-4):367–390.

Gabaix, X. and Ibragimov, R. (2011). Rank - 1 / 2: A simple way to improve the ols estimation of tail exponents. *Journal of Business & Economic Statistics*, 29(1):24–39.

Haziza, D., Beaumont, J.-F., et al. (2017). Construction of weights in surveys: A review. *Statistical Science*, 32(2):206–226.

Jones, C. I. (2015). Pareto and piketty: The macroeconomics of top income and wealth inequality. *Journal of Economic Perspectives*, 29(1):29–46.

Langousis, A., Mamalakis, A., Puliga, M., and Deidda, R. (2016). Threshold detection for the generalized pareto distribution: Review of representative methods and application to the NOAA NCDC daily rainfall database. *Water Resources Research*, 52(4):2659–2681.

Little, R. J. and Vartivarian, S. (2005). Does weighting for nonresponse increase the variance of survey means? *Survey Methodology*.

Michelangeli, V. and Rampazzi, C. (2016). Indicators of financial vulnerability: a household level study. Questioni di Economia e Finanza (Occasional Papers) 369, Bank of Italy, Economic Research and International Relations Area.

Rubin, D. B. (1976). Inference and missing data. *Biometrika*, 63(3):581–592.

Rubin, D. B., editor (1987). *Multiple Imputation for Nonresponse in Surveys*. John Wiley & Sons, Inc.

Särndal, C.-E. (2007). The calibration approach in survey theory and practice. *Survey Methodology*, page 99.

Särndal, C.-E. and Lundström, S. (2005). *Estimation in Surveys with Nonresponse*. John Wiley & Sons, Ltd.

Steindl, J. (1965). *Random processes and the growth of firms: a study of the Pareto law*. Economic theory and applied statistics. Hafner Pub. Co.

Vermeulen, P. (2016). Estimating the top tail of the wealth distribution. *American Economic Review*, 106(5):646–50.

Vermeulen, P. (2018). How fat is the top tail of the wealth distribution? *Review of Income and Wealth*, 64(2):357–387.

Waltl, S. (2018). Multidimensional Wealth Inequality: A Hybrid Approach toward Distributional National Accounts in Europe. In *Proc. 35th IARIW General Conference (IARIW 2018)*.

Wold, H. O. A. and Whittle, P. (1957). A model explaining the pareto distribution of wealth. *Econometrica*, 25(4):591–595.

Yang, G. L. (1978). Estimation of a biometric function. *The Annals of Statistics*, 6(1):112–116.