Gran Sasso Science Institute

Doctoral Thesis

ESSAYS ON THE ECONOMIC GEOGRAPHY OF OPPRESSIVE VIOLENT DEVIANT BEHAVIOURS

Author:

Daria Denti

Supervisor:

Prof. Alessandra Faggian

Prof. Simona Iammarino

A thesis submitted in the fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Social Sciences

10th November 2019

Declaration of Authorship

I, Daria Denti, declare that this thesis titled, "Essays in the Economic Geography of Oppressive Violent Deviant Behaviours" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself

SIGNED: _____

DATE: 27-01-2020

ABSTRACT

This dissertation consists of four self-contained essays on the economic geography of oppressive violent deviant behaviours, a definition identifying violent violations of social norms driven by oppressive urges against minorities. The common thread running through the essays is the quantitative exploration of how the local socioeconomic characteristics influence the occurrence and the scale of violence directed towards disempowered groups.

The first chapter investigates the relationship between hate events and the socioeconomic/cultural characteristics of Italian Local Labour Market Areas, through a unique database with georeferenced hate manifestations. The estimation of a hurdle model identifies the local conditions acting as risk factors for hate occurrence and for hate frequency. The geography of refugees' hosting structures appears as a predictor of both the occurrence and the frequency of hate, whereas foreign resident population does not display any significant influence. At the same time, trust works as a protective factor reducing the occurrence of hatred at the local level. Moreover, once the hurdle of experiencing at least one hate manifestation is crossed, fewer conditions are needed to fuel further hate in the same place.

The second chapter analyses the influence of real-world socioeconomic features on online hate in Italy, exploiting a novel dataset on geotagged hate tweets. Results show a strong empirical association between the local economic dimension and cyberhate. Economic insecurity is robust risk factors for online hate, as well as economic inequality. The latter influences online hate along two channels: the local outlook of income inequality and the relative importance assigned to individualistic values by the established family type in the area.

In the third chapter I investigate whether school bullying is affected by a cultural shock from migration. The analysis exploits the natural shock from migration which occurred in the UK after the 2004 European Union enlargement to empirically estimate how a sudden and sizeable migration inflow influences school bullying. The findings -robust to endogeneity of immigrants' location choice- highlight that the cultural shock from

migration determines an increase in school violence. The paper also suggests that existing language barriers act as a moderator for the migration shock by increasing its effect.

The final chapter explores whether women' propensity to report sexual crimes to the police is influenced by the local availability of specialized services for victims of sexual offences. Applying the synthetic control method to the UK, the empirical investigation shows that the local availability of dedicated services, as refuges and professional help, increases women's willingness to seek justice. The positive effect of the local provision of specialized services holds even after the occurrence a of countrywide and prominent media campaign about sexual offences. This last finding further supports the importance of the availability of nearby services, since it suggests that nation-scale initiatives do not work as substitutes.

AKNOWLEDGEMENTS

Firstly, I would like to express my sincere gratitude to my advisors Prof. Alessandra Faggian and Prof. Simona Iammarino for having endorsed and continuously supported my PhD research with their motivation, advice and guidance. Their supervision helped me in all the time of research and writing of this thesis. I could not have imagined having better advisors and mentors for my PhD study. Alessandra embraced the ideas that have been developed in this thesis at a time when the study of resentment and intolerance in applied economics and regional science was at the very seminal stage and, from that moment onwards, she has been supportive and enthusiastic about every stage of this PhD thesis. Simona welcomed me at the Department of Geography and the Environment of the London School of Economics and since the very beginning of my visiting period she has provided me with insightful suggestions about my research ideas as well as valuable comments on my work.

I am also grateful to both the GSSI and the London School of Economics for the great opportunity that they offered me through the Erasmus+ exchange programme which allowed me to spend extremely valuable months in the UK working on my papers on bullying and gender-based violence.

I thank Marco Di Cataldo, Davide Luca and Paola Proietti for the sound discussions and comments on my work.

I am also extremely grateful to my GSSI colleagues for the wonderful time that we spent together in open-spaces and everywhere we went to, building wonderful friendships. And to Giulio and Federico for three amazing years of constantly thinking outside the box.

I also thank my friends for being always there for me and for making me handle everything with a smile. From being my strongest supporters and my safe harbour to the never-ending lifts to/from airport to the last-minute places to stay (particularly when I lock myself outside!). Thanks Beppe, Federico, Loris, Lucia, Mattia and Simone. And thanks Andrea for our long discussions on what should drive us doing research (and quite often also on the meaning of life!).

I would also like thank my family for always boosting me to follow my dreams.

List of Main Tables

"Bigotry and racism are among
the deadliest social ills
plaguing the world today. […]
The only way to destroy them
is to expose them, to reveal them
for the insidious evils
they really are"
S. Lee

"Luke don't give in to hate.
That leads to the dark side."
O.W. Kenobi

Introduction

The past few years have witnessed a mounting attention towards deviant behaviours such as hate events, online hate speech, school bullying and gender-based violence. The increasing concern is motivated by their growing records, which indicate that these negative events are disseminating fast alongside their high societal costs.

This subset of deviant behaviours refers to violent violations of social norms against a target with the aim to harass the victim's group. Violence is driven by oppressive urges against minorities, disempowered, stigmatized or different groups and it manifests itself by targeting victims who match the identifying characteristics of the aimed group (Gerstenfeld, 2017; Perry, 2001). Criminologists, sociologists as well as social psychologists acknowledge that these behaviours, although different along several dimensions, are all enacted by inter-group strains triggered by prejudices (Gerstenfeld, 2017; Hall, 2013; Hogg, 2006; Arthur G Miller, 2004; Staub, 2004). The central role of prejudices in activating oppressive violent behaviours has prompted scholars to introduce the hypothesis that places also matter, being that prejudices are cultural features rooted in the local social milieus (Gerstenfeld, 2017; Hall, 2013; Perry, 2001).

The role of the local dimension finds further support also in the existing evidence showing that the observed geography of violence against minorities cannot be explained by means of the personality traits of perpetrators (Green, McFalls, & Smith, 2001; Hall, 2013). Hence, the factors driving the escalation of violence have to reside elsewhere than

personality and places represent a sound dimension to explore, as human beings do not live in a social vacuum (Castells, 2001).

The conceptualized nexus between oppressive violent deviant behaviours and places opens up to the empirical investigation of the geography of the socioeconomic risk factors associated to the occurrence and the size of these violent events. This thesis aims at contributing to this investigation through the quantitative analysis of the spatial elements which either promote or counter the proliferation of violence against minorities. More into details, it explores the local geographies of both the risk and the protective factors associated to four types of oppressive violent deviant behaviours considering two different countries and it presents evidence supporting a non-spatially blind policy approach to counter the proliferation of these behaviours. In addition, the thesis focuses on scrutinizing the relative role of local economic drivers with respect to sociocultural drivers in shaping violent intolerance against minorities, hence contributing to the ongoing debate opposing the perception of sociocultural threats to the perception of economic threats as the causes of people's resentment (Card, Dustmann, & Preston, 2012; Hainmueller & Hopkins, 2014; McCann, 2019; Rodríguez-Pose, 2018). Finally, by focusing on actual violent behaviours, this thesis adopts an approach that looks at revealed preferences instead of at attitudes' declaration as most of the existing literature does (Hainmueller & Hopkins, 2014).

The remainder of this introduction is organised as follows. Section 1 provides background information on the phenomenon. Section 2 introduces the literature and the research gaps. Sections 3 provides the aim of the thesis as well as a description of its structure.

1.Background

The count and the diffusion of oppressive violent deviant behaviours are sharply growing. In the US, both hate crimes and hate groups have increased by 30% between 2014 and 2017 (FBI, 2017; Southern Poverty Law, 2019). In a panel of European countries, hate crimes have nearly doubled between 2013 and 2016 (OSCE-ODHIR, 2019). Enlarging the perspective to behaviours which are not generally listed as crimes, such as school bullying and online hate speech, numbers remain striking. The trend on online hate speech is marked: in a recent special Eurobarometer, 75% of respondents

said they had seen online hate speeches on social platforms (Eurobarometer, 2016) and in the US the share is 66% (Duggan, 2017). The data also show large shares of cyberhates produced by people not belonging to any organised hate group (Hall, 2013). Also, available data on social networks contribute to grasp the size of this phenomenon: in 2018 Facebook has removed about 7.9 million hate speech contents and YouTube has deleted more than 15,000 channels per quarter for the same reasons[1]. With respect to school bullying, the recent update of the PISA survey has inserted a specific question on school bullying, allowing to measure an average 18% of students reporting that they are frequently experiencing bullying across OECD countries (OECD 2017). At the global level, the share rises to 30% (UNESCO 2018). Finally, with respect to gender-based violence, considering the European Union scale, one in three women says someone has hurt them physically and one in two has experienced sexual harassment (Fundamental Rights Agency, 2014). In the US, the rate of rape or sexual assault has increased from 1,4 victimizations per 1000 persons age 12 or older in 2017 to 2,7 per 1000 in 2018 (US Department of Justice, 2019).

While it is straightforward to list hate as a violent behaviour aimed at oppressing the targeted minorities, more details are provided with respect to school bullying and gender-based violence. School bullying identifies aggressive behaviour among school-aged children that involves a real or perceived power imbalance (OECD, 2017). Bullying is a deliberate abuse of physical or psychological force or power, threatened or actual, against a person that is a member of minority groups, less protected, or perceived to transgress traditional boundaries. The power imbalance -real or perceived- is a pivotal characteristic of bullying distinguishing it from conflict (Juvonen & Graham, 2014). The role of power alongside the target that generally belongs to a minority are relevant lines to embed bullying within violent oppressive behaviours. In some countries, such as the UK, anti-bullying programs are part of broader policy to counter hate. Gender-based violence can be identified as violent deviant behaviours where the oppressive urges aim at reinforcing the subordinate status of women in society (Hodge, 2011; Pain, 2000; Perry, 2001). Notably, some hate-countering legislation contemplate also gender (Gerstenfeld, 2017).

[1] https://transparency.facebook.com/community-standards-enforcement#hate-speech ; https://transparencyreport.google.com/youtube-policy/removals?hl=en.

Alongside the worrying trends displayed by this type of behaviours, criminologists have detailed how oppressive deviant behaviours hurt the victims more than other violent behaviours, as they pivot on blaming victims on their identifying characteristics (gender, ethnicity, sexual orientation, religion). Moreover, oppressive violent deviant behaviours also impact the wider community, acting as "*message*" events (Iganski, 2008), since they create an *in terrorem* effect that outspreads the individual victim to reach all community members creating a sense of group vulnerability, social tension and fear (Hall, 2013; OSCE-ODIHR, 2009), overall determining social and economic costs (Glaeser, 2005). Hence, due to both their growing prominence and their persistent negative impact on society, more and more States and international institutions are legislating in order to stop the spread of these phenomena[2].

Acts deriving from oppressive violent deviant behaviours remain largely unreported mainly due to the victims' fear of reprisal and/or sense of social marginalisation (Fundamental Rights Agency, 2014; Gagliardone *et al.*, 2015; OECD, 2015). This limited willingness to report means that policies to counter these occurrences which are based only on the crime dimension will fail to be effective, since many perpetrators would not be reported. Hence, prevention policy initiatives appear a more suitable tool (Anderberg, Rainer, Wadsworth, & Wilson, 2016; Gagliardone et al., 2015; Wilkinson & Pickett, 2017). The key role of prevention calls for the identification of the elements which either promote or oppose the manifestation of oppressive violent deviant behaviours, hence further supporting a spatial exploration of both risk and protective factors relating to violence against minorities and disempowered groups. The next section details the state of the literature investigating risk and protective factors.

[2]European Commission #SayNoStopVAW initiative; EU High Level Group on Combating Intolerance; UN Sustainable Development Goals #5; UNESCO initiative to counter hate speech; OSCE Hate Crime Monitor; UK Action Plan against Hate; OECD "How much of a problem is bullying at school?"; OSCE Office for Democratic Institutions and Human Right (ODIHR) Hate Crime Reporting; European Union Fundamental Right Agency (FRA), Germany Network Enforcement Act (NetzDG), United States Matthew Shepard and James Byrd Jr. Hate Crimes Prevention Act; England and Wales Crown Prosecution Office and Association of Chief Police Officers Hate Crime Manual and Tactical Guidance; Council of Europe Strategy on the Rights of the Child 2012-2015; ABC Europe Anti-Bullying Campaign.

2.Literature and research gaps

For many decades, the prominent perspective in the analysis of oppressive violent deviant behaviours was pivoting on the role of personality, in the attempt of determining which constellation of individual characteristics make up a "*prejudiced personality*" to act against the out-group members. This methodology, labelled "*personality approach"*, is not particularly consistent with empirical evidence on perpetrators of violent oppressive acts (Green et al., 2001; Hall, 2013). Indeed, it suffers from the "*fundamental attribution bias*", meaning that it overestimates the role of personality traits while neglecting both the role of situations and the volatility patterns of oppressive violent acts along time (Brown, 1995 p.34; Craig, 2002; Zimbardo, 2004). More into details, the "*personality approach"* fails to account for the fact that evil actions are widely performed by people who do not display wicked and malevolent personality traits (Baumeister & Vohs, 2004). Figures highlight that no more than 5% of hate crimes in the US are committed by members of hate groups and most hate crimes are committed by ordinary people with no sadistic or extremely biased personality traits (Gerstenfeld, 2017, p.97). The same patterns emerge with respect to cyberhate (Hall, 2013), gender-based violence (Keller, Mendes, & Ringrose, 2018) and school bullying (Álvarez-García, García, & Núñez, 2015; Espelage, Van Ryzin, & Holt, 2018). Another element that cannot be explained by the "*personality approach*" resides in the displayed volatility patterns of oppressive violent acts along time; to this respect, these events are characterized by temporal trends that do not fit with the temporal trends of the share of population with wicked and malevolent personality traits (Brown, 1995; Craig, 2002; Zimbardo, 2004).

As a consequence, another perspective has emerged, focusing on "*the power of situations*" (Miller, 2004). The key point of this second approach, labelled "*the situationist approach*", is that the characteristics of the socioeconomic milieu in which people live exert a strong role in the occurrence of oppressive violent deviant behaviours. More precisely, the characteristics of places act as triggers, escalating existing prejudices into oppressive violation of social norms.

Prejudices are the pivoting element of the "*situationist approach*". They serve the self-reinforcement of in-group boundaries and the protection of the in-group social identity and they are generated from the internalization of stereotypes (Gerstenfeld, 2017 p. 106). Notably, stereotypes are intrinsic by-products of the process through which human

beings get to know the world. They are judgements, or "*illusory correlations*" (Brown, 1995), generated as drawbacks by the innate cognitive process of social categorization. Through the idiosyncratic social categorization process, human beings build their knowledge of the world, identifying which social group they belong as individuals -that is their social identity or in-group membership (Allport, 1988; Ashforth & Mael, 1989; Brown, 2014; Putnam, 2007). By constructing the social identity through the characteristics of the in-group in which they fit, individuals clearly draw also the boundaries of the out-group(s). However, this parting is biased, since it is complemented by a perceptive distortion against the out-groups caused by two tendencies. First, the out-group's members are seen as more similar to each other than the members of the in-group; second, there is a tendency to perceive negative things as endogenous and idiosyncratic to the out-group, while exogenous to the in-group (Hewstone & Jaspars, 1984; Pettigrew, 1979; Whitley JR & Kite, 2010). These judgments are based only on the individual's group affiliation rather than on information actually learned about the specific individual and they are defined as stereotypes. Stereotypes are characterized by being persistent features of the sociocultural milieu (Perry, 2001 p.51) and by being self-perpetuating -since we recall things better when they are consistent with our stereotypes (Hall, 2013 p.86). Prejudices are the attitudes generated from these judgements (Whitley JR & Kite, 2010). Prejudices then act as stimuli, among in-group members, for actual behaviours aimed at preserving a given socio-economic and cultural status quo whenever there is real or perceived strain between in-group and out-group (Hogg, 2006; Perry, 2001; Whitley JR and Kite, 2010). These behaviours may often be violent and oppressive deviances, by infringing social norms and sometimes also laws.

Having identified why prejudices are a constituting feature of oppressive violent deviant behaviours and being that "*prejudice is a too prevalent a phenomenon to be simply consigned to the province of individual pathology*" (Brown, 1995, p.10), it follows that what has to be investigated is which local elements influence the escalation of existing prejudices in oppressive violent deviant behaviours. Frameworks such as "*the strain theory*", "th*e real conflict theory*", "*the integrated threat theory*" and "*cultural persistence*" are amongst the conceptual set-ups used to investigate the role of role of socio-economic, cultural and demographic variables in influencing oppressive violent deviant behaviours.

The "*strain theory*" is among the most widespread theories used to analyse the association between social factors and oppressive violent deviant behaviours and it states that, whenever members of a community experience a gap between culturally prescribed goals and their means and opportunities of attaining these legitimately, these people will experience a strain which, in turn, opens to deviant behaviours, including bias-based violence (Perry, 2009). Traditional "*strain theory*" has been challenged by many scholars, due to the fact that it does not accommodate for the evidence that many perpetrators do not experience any limitation in the means they need to acquire a desirable social status. Indeed, many of them belongs to economically and powerful groups in the society (Abrams, 2010; Hall, 2013). One of the most influential frameworks capable of accounting for this critique is the Perry (2001) "*structured action theory of doing difference*", which identifies the causes for social frustration and violent oppression of minorities in the fear that "*different people*" threaten and challenge established social hierarchical structures and values. In this account, bias manifestations are triggered by threats to the social status quo and to the social identity. The "*real conflict theory*" focuses on the economic dimension as trigger for oppressive violence, considering the role of inter-group actual conflicts of interest on valued commodities (such as jobs, land, power) or opportunities (Gerstenfeld, 2017; Green et al., 2001). Still within the economic dimension, the "*integrated threat theory*" stresses the role of symbolic and perceived threat together with actual threats, an element which appears to be particularly important in economic downturns (Stephan & Stephan, 2000). These theories relate to the influence of a sense of unfulfillment of needs in escalating prejudices into violent behaviours (Staub, 1999). Finally, the cultural dimension is relevant, since culture is the main repository of stereotypes and prejudice (Brown, 1995; Gerstenfeld, 2017; Perry, 2001).

Notably, elements such as the social identity, social values, employment and culture have a strong place-specific dimension. Thus, geography emerges as a relevant issue to consider. To this regard, another relevant spatial feature is given by the distinction between marginalized and central areas, where the former are endowed with geographic elements of seclusion, disconnectedness and backwardness that may, in turn, promote resentment and frustration (Rodriguez-Pose, 2018).

Hence, there is a foundation for a regional science perspective to the investigation of the determinants of violent acts against minorities and disempowered groups. This

rationale has just started to be exploited; recent seminal contributions have explored the relationships between some spatial features and specific oppressive deviant behaviours, outlining relevant associations (Voigtlander and Voth, 2012; Alesina, Brioschi and Ferrara, 2016; Anderberg, Mantovan and Sauer, 2018; Medina et al., 2018; Tur-Prats, 2018). Besides, growing empirical evidence on recent electoral outcomes provide further support to the relevance of the geographical patterns in shaping resentful attitudes (i.a. Guiso, Herrera, Morelli, & Sonno, 2017; McCann, 2019; Rodríguez-Pose, 2018; Rodrik, 2018). Overall, results outline strong and meaningful influences of the geography dimension on intolerance and resentment.

Existing contributions focus mainly on attitude declaration through surveys and electoral outcomes to gauge a measure for intolerance (i.a. Alesina, Brioschi and Ferrara, 2016; Barone, D'Ignazio, de Blasio, & Naticchioni, 2016; Colussi & Pestel, 2017; Halla, Wagner, & Zweimüller, 2017; Hangartner, Dinas, Marbach, Matakos, & Xefteris, 2018; Tur-Prats, 2018). With regards to attitude declaration, mainly collected through surveys, scholars are warning about a social desirability bias which may push respondents to provide answers reflecting what is considered as socially acceptable rather than their true attitude (Hainmueller & Hopkins, 2014). At the same time, electoral outcomes fail to account for the fact that intolerance is a pre-political feature, hence it may be at work also among people that abstain from voting. Therefore, contributions focusing on measures of actual manifestations are called for since they contribute to a fine-tuned assessment.

To conclude, different disciplines in the social sciences have been providing both conceptual frameworks and introductory evidence prompting for a quantitative spatial exploration of the role of places in relationship with oppressive violent deviant behaviours. This thesis develops a contribution in this perspective by presenting empirical evidence identifying both risk and protective factors nested into the local dimension with respect to Italy and the UK. Due to the seminal stage of the literature investigating from a quantitative perspective the socioeconomic geography of violent behaviours against minorities, the research questions presented in the thesis mainly add to building the knowledge base.

3.Aim and structure of the thesis

This thesis aims at providing empirical evidence on how local socioeconomic characteristics influence violence against disempowered groups. This aim is pursued by identifying suitable measures for actual manifestations of these behaviours, going beyond existing measures constructed out of attitude's declaration. To this regard, a relevant effort has been devoted to collecting spatially fine-grained data on the phenomena of interest which are, thus far, available from different and scattered sources.

Given the actual germinal stage of the literature analysing the geography of these behaviours, the scope of the empirical analysis is comprehensive. Therefore, notwithstanding the main relationship scrutinised by each paper, the broad potential pool of local features which may exert some influence is considered. Moreover, in paper 1,2 and 3 the results from the baseline empirical estimations are gauged through several robustness checks, including the estimation of competing modelling strategies.

The thesis is organised into four chapters, written in the form of academic papers. Paper 1 and 2 focus on Italy, investigating which risk and protective factors relate to real-world hate events and cyberhate respectively. Paper 3 target the UK and it assesses the relevance of a cultural threat generated from a migration shock in generating oppressive violence at school. Finally, paper 4 takes a policy-oriented perspective by evaluating the effect of local policies in favouring the countering of oppressive violence against women, again in the UK.

More into details, Paper 1 focuses on hate manifestations -vandalism, physical attacks, verbal attacks- happening in Italy to empirically detect which characteristics of places are related with the occurrence and the frequency of hatred. The results show that the local sociocultural dimension displays meaningful associations with hate manifestations, whereas the economic dimension does not appear as exerting relevant influences.

Paper 2 addresses the rising phenomenon of cyberhate, again focusing on Italy, to detect whether there are real-world features capable of influencing this online behaviour. The findings support a meaningful association between the economic dimension and online hate speech and they also outline that online hate speech and real-world hate manifestations are associated with different risk and protective factors.

Paper 3 target the UK, to investigate the effect of a cultural shock due to migration on the local level of school bullying. The evidence presented in the paper strongly

supports the role of the cultural shock in determining oppressive violence at the school level.

The evidence presented in the Paper 1, 2 and 3 supports the need for non-spatially blind policies to counter oppressive violent deviant behaviours. Hence, paper 4 adopts a policy perspective to analyse the effect of a local policy to counter violence against women on the victims' propensity to report sexual offences in the UK. The comparative case study analysis performed in the paper confirms that locally grounded policies are effective in countering oppressive violence.

Collaborations

I state that Paper 1 and 2 are jointly co-authored with Alessandra Faggian (Gran Sasso Science Institute). Both papers were conceived by all of the authors. I carried out the data collection and the econometric estimations. Extracts from Paper 1 and 2 have been summarised in two technical reports I have been engaged with within the C.O.N.T.R.O. project (Counter Narratives AgainsT Racism Online). The project is coordinated by UNAR, the anti-racial discrimination national office and financed through the Rights, Equality and Citizenship Programme of the European Commission. Paper 3 and 4 were realised benefitting from the Erasmus+ exchange programme at the Department of Geography and Environment of the London School of Economics and of subsequent field work in the UK to gather further data. Paper 3 will be soon submitted to a peer-review journal.

References

Abrams, D. (2010). Processes of Prejudice: Theory, Evidence and Intervention. Manchester: EHRC.

Aizer, A. (2010). The Gender Wage Gap and Domestic Violence Anna. American Economic Review, 100(4), 1847–1859. https://doi.org/10.1002/hbm.21045.Effects

Alesina, A., Brioschi, B., & Ferrara, E. La. (2016). Violence Against Women: A Cross-cultural Analysis for Africa. National Bureau of Economic Research Working Paper Series, No. 21901. https://doi.org/10.3386/w21901

Allport, G. W. (1988). The Nature Of Prejudice. Addison-Wesley.

Álvarez-García, D., García, T., & Núñez, J. C. (2015). Predictors of school bullying perpetration in adolescence: A systematic review. Aggression and Violent Behavior, 23,

126–136. https://doi.org/10.1016/j.avb.2015.05.007

Anderberg, D., Mantovan, N., & Sauer, R. M. (2018). The dynamics of domestic violence: Learning about the Match, (April). https://doi.org/10.2307/2117958

Anderberg, D., Rainer, H., Wadsworth, J., & Wilson, T. (2016). Unemployment and Domestic Violence: Theory and Evidence. The Economic Journal, (1230).

Ashforth, B. E., & Mael, F. (1989). Social identity theory and the organization. The Academy of Management Review, 14(1), 20–39. https://doi.org/10.2307/258189

Bansak, K., Hainmueller, J., & Hangartner, D. (2016). How economic, humanitarian, and religious concerns shape European attitudes toward asylum seekers Article (Accepted version) (Refereed) How Economic, Humanitarian, and Religious Concerns Shape European Attitudes toward Asylum-Seekers. Science, 354(6309), 217–223. https://doi.org/10.1126/science.aag2147

Barone, G., D'Ignazio, A., de Blasio, G., & Naticchioni, P. (2016). Mr. Rossi, Mr. Hu and politics. The role of immigration in shaping natives' voting behavior. Journal of Public Economics, 136(8228), 1–13. https://doi.org/10.1016/j.jpubeco.2016.03.002

Baumeister, R., & Vohs, K. (2004). Four Roots of Evil. In A.G Miller (Ed.), The Socio-Psychology of Good and Evil. Guilford.

Brown, R. (1995). Prejudice: Its social psychology. Prejudice: Its social psychology. Malden: Blackwell Publishing.

Brown, R. (2014). Prejudice. Its social psychology. Journal of Multilingual and Multicultural Development (Vol. 35). https://doi.org/10.1080/01434632.2013.803723

Card, D., Dustmann, C., & Preston, I. (2012). Immigration, wages, and compositional amenities. Journal of the European Economic Association, 10(1), 78–119. https://doi.org/10.1111/j.1542-4774.2011.01051.x

Castells, M. (2001). The Internet Galaxy: Reflections on the Internet, Business, and Society. New York, NY, USA: Oxford University Press, Inc.

Colussi, T., & Pestel, N. (2017). Minority Salience and Political Extremism ∗.

Craig, K. M. (2002). Examining hate-motivated aggression: A review of the social psychological literature on hate crimes as a distinct form of aggression. Aggression and Violent Behavior, 7(1), 85–101. https://doi.org/10.1016/S1359-1789(00)00039-2

Duggan, M. (2017). Online harassment 2017. Pew Research, 1–85. https://doi.org/202.419.4372

Dustmann C, Vasiljeva K., P. D. A. (2016). Refugee Migration and Electoral Outcomes. CReAM Discussion Paper, (19/16). https://doi.org/10.1016/S1043-2760(97)84344-5

Espelage, D. L., Van Ryzin, M. J., & Holt, M. K. (2018). Trajectories of bully perpetration across early adolescence: Static risk factors, dynamic covariates, and longitudinal outcomes. Psychology of Violence, 8(2), 141–150. https://doi.org/10.1037/vio0000095

Eurobarometer. (2016). Media pluralism and democracy – Special Eurobarometer 452. Digital Single Market. Retrieved from https://ec.europa.eu/digital-single-market/en/news/media-pluralism-and-democracy-special-eurobarometer-452

FBI. (2017). FBI — Hate Crime- Victims. Retrieved from https://ucr.fbi.gov/hate-crime/2017/topic-pages/victimsFundamental Rights Agency. (2014). Violence against women : an EU - wide survey.

Gagliardone, I., Gal, D., Alves, T., & Martinez, G. (2015). Countering Online Hate Speech. UNESCO Series on Internet Freedom. https://doi.org/978-92-3-100105-5

Gerstenfeld, P. B. (2017). Hate Crimes: Causes, Controls, and Controversies. SAGE Publications. Retrieved from https://books.google.it/books?id=v35ZDwAAQBAJ

Glaeser, E. L. (2005). The political economy of hatred. Quarterly Journal of Economics, (February), 45–86.

Green, D. P., McFalls, L. H., & Smith, J. K. (2001). Hate Crime: An Emergent Research Agenda. Annual Review of Sociology, 27(1), 479–504. https://doi.org/10.1146/annurev.soc.27.1.479

Guiso, L., Herrera, H., Morelli, M., & Sonno, T. (2017). Populism: Demand and Supply *, 1–63. Retrieved from http://www.heliosherrera.com/populism.pdf

Hainmueller, J., & Hopkins, D. J. (2014). Public Attitudes Toward Immigration. Annual Review of Political Science, 17, 225–249. https://doi.org/10.1146/annurev-polisci-102512-194818

Hall, N. (2013). Hate crimes. 2nd Edition. Routledge.

Halla, M., Wagner, A. F., & Zweimüller, J. (2017). Immigration and Voting for the Far Right. Journal of the European Economic Association, Volume 15, Issue 6, 1 December 2017, Pages 1341–1385, (6), 1341–1385.Hangartner, D., Dinas, E., Marbach, M., Matakos, K., & Xefteris, D. (2018). Does Exposure to the Refugee Crisis Make Natives More Hostile? American Political Science Review, (17). https://doi.org/10.2139/ssrn.3042936

Hewstone, M., & Jaspars, J. (1984). Social dimensions of attribution. In H. Tajfel (Ed.), The social dimension: European developments in social psychology. Cambridge University Press. Retrieved from https://books.google.it/books?id=ilIIeD5SCKUC

Hodge, J. . (2011). Gendered Hate: Exploring Gender in Hate Crime Law. Boston, MA: Northeastern University Press.

Hogg, M. A. (2006). Social Identity Theory. In Contemporary social psychological theories. (pp. 111–136). Stanford University Press.

Iganski, P. (2008). "Hate crime" and the city. https://doi.org/10.1332/policypress/9781861349408.001.0001

Juvonen, J., & Graham, S. (2014). Bullying in Schools: The Power of Bullies and the Plight of Victims. Ssrn. https://doi.org/10.1146/annurev-psych-010213-115030

Keller, J., Mendes, K., & Ringrose, J. (2018). Speaking 'unspeakable things': documenting digital feminist responses to rape culture. Journal of Gender Studies, 27(1), 22–36. https://doi.org/10.1080/09589236.2016.1211511

McCann, P. (2019). Perceptions of regional inequality and the geography of discontent: insights from the UK. Regional Studies, 1–12. https://doi.org/10.1080/00343404.2019.1619928

Medina, R. M., Nicolosi, E., Brewer, S., & Linke, A. M. (2018). Geographies of organized hate in America: A regional analysis. Annals of the American Association of Geographers, 108(4), 1006–1021. https://doi.org/10.1080/24694452.2017.1411247

Miller, A. G (Ed.). (2004). The social psychology of good and evil. The social psychology of good and evil. New York, NY, US: The Guilford Press.

OECD. (2015). Students' Well-Being. PISA 2015 Results, III. https://doi.org/http://dx.doi.org/10.1787/9789264273856-en IS

OSCE-ODHIR. (2019). Hate Crime Monitor. Retrieved from http://hatecrime.osce.org/incidents

OSCE-ODIHR. (2009). Preventing and responding to hate crimes - A resource guide for NGOs in the OSCE region. Retrieved from http://www.osce.org/odihr/39821?download=true

Pain, R. (2000). Place, social relations and the fear of crime: a review. Progress in Human Geography, 24(3), 365–387. https://doi.org/10.1191/030913200701540474

Perry, B. (2001). In the Name of Hate: Understanding Hate Crimes. https://doi.org/10.1017/CBO9781107415324.004

Perry, B. (2009). The Sociology of Hate: Theoretical Approaches. In B. Perry (Ed.), Hate Crimes. Westport: Praeger.

Pettigrew, T. F. (1979). The Ultimate Attribution Error: Extending Allport's Cognitive Analysis of Prejudice. Personality and Social Psychology Bulletin, 5(4), 461–476. https://doi.org/10.1177/014616727900500407

Putnam, R. (2007). Diversity and community in the twenty-first century the 2006 Johan Skytte prize lecture". Housing Policy Debate (Vol. 19). https://doi.org/10.1080/10511482.2008.9521631

Rodríguez-Pose, A. (2018). The revenge of the places that don't matter (and what to do about it). Cambridge Journal of Regions, Economy and Society, 11(1), 189–209. https://doi.org/10.1093/cjres/rsx024

Rodrik, D. (2018). Populism and the economics of globalization. Journal of International Business Policy, 1(1), 12–33. https://doi.org/10.1057/s42214-018-0001-4

Staub, E. (1999). The Roots of Evil: Social Conditions, Culture, Personality, and Basic Human Needs. Personality and Social Psychology Review, 3(3), 179–192. https://doi.org/10.1207/s15327957pspr0303

Staub, E. (2004). Basic human needs, altruism, and aggression. In The social psychology of good and evil. (pp. 51–84). New York, NY, US: The Guilford Press.

Stephan, W. G., & Stephan, C. W. (2000). An integrated threat theory of prejudice. In Reducing prejudice and discrimination. (pp. 23–45). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Tur-Prats, A. (2018). Family Types and Intimate-Partner Violence: A Historical Perspective. The Review of Economics and Statistics, 1–45. https://doi.org/10.1162/rest_a_00784

US Department of Justice. (2019). Criminal Victimization, 2018.

Voigtlander, N., & Voth, H. (2012). Persecution Perpetuated: the Medieval Origins of anti-Semitic Violence in Nazi Germany. Quarterly Journal of Economics, 127(2), 1339–1392. https://doi.org/10.1093/qje/qjs019.Advance

Whitley JR, B. E., & Kite, M. E. (2010). The Psychology of Prejudice and Discrimination (2nd edition). Journal of Chemical Information and Modeling (Vol. 53). Wadsworth.

Wilkinson, R. G., & Pickett, K. E. (2017). The enemy between us: The psychological and social costs of inequality. European Journal of Social Psychology, 47(1), 11–24. https://doi.org/10.1002/ejsp.2275

Zimbardo, P. (2004). A situationist perspective on the psychology of evil: Understanding how good people are transformed into perpetrators. The Social Psychology of Good and Evil: Understanding Our Capacity for Kindness and Cruelty., 21–50. https://doi.org/10.1097/01.nmd.0000207370.95835.47

CHAPTER 1

Mapping the Italian Geography of Hatred.

Abstract. Where does hatred happen in Italy? The paper investigates the relationship between hate and the socioeconomic/cultural characteristics of the 611 Italian Local Labour Market Areas, through a unique database with georeferenced hate manifestations. Through a hurdle model estimation, the conditions relating to hate occurrence and the conditions relating to hate frequency are identified. The geography of refugees' hosting structures is a robust predictor of both occurrence and frequency of hate, whereas foreign resident population does not display any significant influence. Results support the sociocultural dimension in promoting hatred at local level rather than the economic one. Moreover, once the "*hurdle*" of experiencing at least one hate manifestation is crossed, fewer conditions are needed to fuel further hate in the same place.

Keywords: refugees, migrants, hate, Local Labour Market Area

1.1 Introduction

Hate manifestations have been increasing in their count in recent years: considering a sample of 14 European countries, figures show that xenophobic hate crimes have doubled from 2013 to 2016 (European Commission, 2017). The diffusion of hate impacts heavily society, reducing freedom perception for the objects of hatred, propagating negative stereotypes, promoting discriminative behaviours, increasing the cost of correct information (Glaeser, 2005). Overall, this exerts negative consequences such as increasing interaction costs, false signalling, inefficient matching, segregation and lock-in in inefficient allocations. As a consequence of the increasing evidence on the magnitude and the pervasiveness of hate manifestations, national and international institutions are

legislating and promoting plans to counter the phenomenon[3]. Furthermore, hate within Europe is deeply intertwined with the migrant/refugees' question, being migrants and refugees the rising target of hate manifestations (Assimakopoulos, Baider, & Millar, 2017). The growing bulk of evidence highlights that the traditional perspective of classifying hate manifestations as associated to isolated "*bad seeds*" within a society fails to account for a vast number of events (Gerstenfeld, 2017; Perry, 2001). Hence, hate-related studies have been increasingly targeting the influence of elements which characterize the local milieu. Nonetheless, few efforts have thus far been devoted to the empirical investigation of the relationship between hate manifestations and the socio-economic, political and cultural geography of places (Jefferson & Pryor, 1999; Medina, Nicolosi, Brewer, & Linke, 2018; Voigtlander & Voth, 2012).

Therefore, there is room for analysing whether it is possible to identify risk factors associated to hate, contributing to build a sound understanding of the growing phenomenon. Within this area of investigation, it is also interesting to comprise the potential role of the unprecedented inflow of refugees: remarkably, it has still to be assessed how much of hate events is actually related to tensions in interactions with resident minorities and/or with refugees rather than to other characteristics of the local environment, albeit that hate behaviour narratives currently strongly pivot on social minorities.

This research addresses these issues focusing on Italy, whose hate events growth rate (49% between 2013 and 2016) exceeds the European average rate (16,8% between 2013 and 2016) (ENAR, 2015; OSCE-ODHIR, 2017)[4]. Hate is analysed across 611 Local Labour Market Areas (LLMAs), by exploiting an original dataset where novel geolocated data on hate events are merged with administrative and historical data.

The analysis supports the role of local features in influencing hatred. Overall, results point to the fact that the sociocultural outlook of places plays a determinant role in fostering hatred, consistently with extant contributions showing a robust association between local conditions and people's life satisfaction and consequent behaviour (Veneri

[3]Some examples: OSCE ODIHR Hate Crime Reporting; US Matthew Shepard and James Byrd Jr. Hate Crimes Prevention Act; England and Wales Crown Prosecution Office and Association of Chief Police Officers Hate Crime Manual and Tactical Guidance.
[4] See Annex A1 for more details

& Murtin, 2019). At the same time, the characteristics of the local economic outlook do not appear as relevant risk factors. In this regard, findings align with scholars arguing that social anxiety against minorities is mainly driven by sociocultural drivers (Hainmueller & Hopkins, 2014; Newman, 2013). Moreover, consistently with recent works (Fratesi, Percoco, & Proietti, 2019; Hangartner et al., 2018), there is evidence of an association between local social tension and refugees'/asylum seekers inflows. The perception of a social threat, such as the presence of asylum seekers centres, is positively associated with hatred. Conversely, elements reducing uncertainty, like trust and social interactions, are negatively associated with it.

This work relates also to the recent literature on geography of resentment (Rodríguez-Pose, 2018), by targeting associations between local features and frustration.

1.2 The determinants of hate: relevant literature and research questions

Criminologists define hate events as violent oppressive behaviours targeting victims on the basis of the perpetrator's prejudice against the actual or perceived status of the target (Craig, 2002). Hate also affects the wider community, acting as "*message*" event (Iganski, 2008), i.e. creating an *in terrorem* effect that outspreads the individual victim to reach all community members, prompting social tension and fear (Hall, 2013; OSCE-ODIHR, 2009) along with socioeconomic costs (Glaeser, 2005). Data increasingly support the role of places (Baumeister & Vohs, 2004; Green et al., 2001) rather than the one of personality traits in influencing hate (Miller, 2004). Indeed, in the US no more than 5% of hate crimes are committed by members of hate groups and most hate crimes are committed by ordinary people with no sadistic or extremely biased personalities (Gerstenfeld, 2017; Hall, 2013). Therefore, hate is "*not abnormal; rather it is a normal (albeit extreme) expression of the biases that are diffused throughout the culture and history in which it is embedded*" (Perry, 2001 p.37)

According to this perspective, local factors act by escalating prejudice into hate events (Hall, 2013, p.9), being prejudice "*an aversive or hostile attitude toward a person who belongs to a group, simply because he belongs to that group, and is therefore presumed to have the objectionable qualities ascribed to that group*" (Allport, 1988). Indeed, hate manifestations often hinge on the perpetrator identifying the target due to the victim's group, with no actual knowledge of

the victim herself (Gerstenfeld, 2017 p. 61; Glaeser, 2005). As outlined by Perry (2001 p.29): "*The victim simply represents the Other in generic terms.*" This perspective stems from criminologists and social psychologists to institutions (OSCE-ODHIR, 2017; UK Crown Prosecution Service, 2018; US Hate Crime Statistics Act, 2010). Prejudices are persistent and pervasive components of the cultural milieu, being by-products of the "*social categorization*" process used by human beings to get to know the multifaceted world (Allport, 1988). This classification suffers from perception distortions due to the tendency to view out-group members more similar to each other than the in-group members and the tendency to perceive negative things as endogenous and idiosyncratic to the out-group, while exogenous to the in-group (Hewstone & Jaspars, 1984; Pettigrew, 1979). These perception biases open for stereotypes: judgments or "*illusory correlations*" (Brown, 1995, p.10) about individuals based only on their group affiliation rather than on information actually learned about them. Stereotypes are, in turn, self-perpetuating - since we recall things better when they are consistent with our stereotypes (Hall, 2013 p.86)- and persistent features of the sociocultural al milieu (Perry, 2001 p.51). Prejudice are the attitudes stemming from stereotypes (Gerstenfeld, 2017 p. 106). Finally, since "*prejudice is a too prevalent a phenomenon to be simply consigned to the province of individual pathology*" (Brown, 1995, p.10), the factors driving the escalation of prejudice towards violent behaviours do not reside in personality traits, but elsewhere. And the local milieu appears an interesting dimension to explore since it influences how people live and interact.

Social psychologists and criminologists have detailed the role of a wide array of local factors in fostering the escalation of prejudices towards bias manifestations, spamming from the socioeconomic to the cultural dimension (Gerstenfeld, 2017; Glaeser, 2005; Hall, 2013; Staub, 2004; Waller, 2014). Being that the literature highlights that there is no single factor relating to people committing hatred, this paper pursues a systematic analysis across an exhaustive range of socio-economic, cultural and spatial factors. In this way, it is possible to address the following research question:

1) *Which are the local factors identified by extant literature displaying an empirical association with hate event in Italy?*

By no mean exhaustive, this analysis allows to scrutinize the association between hatred and local confounding factors, and it can also be helpful in directing future research efforts aimed at carefully identifying specific mechanisms. The analysis starts by

detailing the potential local factors relating to hatred as identified in the literature, alongside with the proxies to gauge them.

The social structure of a community is associated to hate according to different strands of literature. The in-group/out-group dichotomy (Hainmueller & Hopkins, 2014; Hall, 2013), the "*strain theory*" and the "*difference theory*" (Perry, 2001b) identify actual changes in the social structure as key elements triggering hate through threats to established social identity (Gerstenfeld, 2017). To this regard, several variables are considered. First, the share of foreign resident population, displaying positive association with intolerance among incumbent population in some studies (Newman, 2013; Putnam, 2007), but a negative one in others since it enables frequent and less costly interactions between different groups, thus weakening the in-group/out-group dichotomy according to the "*contact theory*" (Kemeny & Cooke, 2017; Pettigrew & Tropp, 2006). The overall impact on hatred is not obvious and needs empirical investigation. Alongside foreign residents, the following measures are included: a measure for the breakdown of the traditional family through the number of divorced/separated couples (Halla, Wagner, & Zweimüller, 2017; Jefferson & Pryor, 1999), a measure for the demographic dynamism through the aging index -measured as the ratio of population over 65 on the population below 15 (Barone et al., 2016), population density, as it can favour interaction (Glaeser, 2014) as well as it captures non-linear urban effects. Also crime is contemplated, since it is associated to uncertainty and social stress (Bianchi, Buonanno, Pinotti, 2012; Dustmann & Fasani, 2016; Otto & Steinhardt, 2014), although literature highlights that hatred is fuelled by stories about crimes of the hated group more than actual figures (Glaeser, 2005). Again, empirical results are needed to determine the final outcome. Finally, the analysis also controls for the local population size.

The economic dimension may matter, through competition for scarce resources -like jobs and welfare protection-, and scapegoating due economic insecurity (Craig, 2002; Garland & Treadwell, 2012; Gerstenfeld, 2017). The proxy for economic insecurity is unemployment share (Barone et al., 2016; Bratti, Deiana, Havari, Mazzarella, & Meroni, 2017; Dustmann, Vasiljeva, & Piil Damm, 2016; Guiso, Sapienza, & Zingales, 2016). Then, also human capital is introduced, identified by some scholars as negatively relating to hate, since education eases overcoming of prejudices (Glaeser, 2005). However, an opposite view claims that self-interest concerns may push affluent people to become

intolerant by internalizing that minority protection costs will be mainly rebated on them (Dustmann et al., 2016) Finally, economic inequality is gauged by means of the Gini index for income inequality.

More than 13 million people live in the portion of Italy identified as "*Inner Areas*", amounting to around the 60% of Italian territory and characterized by an enduring process of geographic marginalization entailing social frustration (Barca, Casavola, & Lucatelli, 2014). Since the latter may exert some influence on hate, the Italian territory is consolidated depending on its belonging to either inner or central areas. Alongside this main classification, also the presence of hub for services and business is considered: on the one hand, the catchment dimension of hubs can promote social interactions and collaboration between diverse people; on the other hand, it can elect hubs as the preferred destination for marginalized people in need for social protection and jobs, determining competition for scarce economic resources.

By countering uncertainty, trust in institutions can reduce perception of threat and fear (Guiso et al., 2017; Glaeser, 2005; Perry, 2001). The proxy for trust is given by the voting turnout at the 2014 European Parliament Election, whose voting attendance has approached nearly 60%. The choice of European elections allows to account for the relevance of European agenda within national discourses on migration inflows (Becker, Fetzer, & Novy, 2017). In terms of collaboration and solidarity, both a measure for participation in associations and a measure for civic capital are introduced. The former conveys information on the actual collaborative outlook of places while the latter identifies values and beliefs which help cooperation (Guiso et al., 2016). Civic capital is proxied by the engagement of Italian municipalities in the Resistance against the Nazi-Fascist during the Liberation War in 1940s, a measure also used by Guiso et al. (2016).

Prejudices are persistent and consistent elements of the local cultural milieu (Perry, 2001), hence calling for the exploration of the potential effects of the local cultural outlook. To this respect, Italy represents a stimulating case: until quite recently it comprised several states, whose cultural outlook variety is recognized as being decisive for current regional identities (Broers, 2003). More into details, the cultural identities of the pre-Unitarian states are acknowledged as a relevant source of sociocultural spatial heterogeneity in Italy (Felice, 2015), due to the non-negligible and established differences in terms of social norms and institutions. Notably the geographical plot of the local

constituencies on which the territorial order of the Kingdom of Italy was constructed after Unification was given by the geography of the pre-unification states, due to their strong sociocultural identity (Cerreti, 2011). Therefore, the actual Italian regional geography is consolidated on pre-Unitarian states after the Aix-La-Chapelle Treaty (1748) and after the Wien Congress (1815). Political preferences are included to account for the potential role of anti-immigrant and anti-elite rhetoric in influencing hate (Glaeser, 2005b; Guiso et al., 2016; Waller, 2014) and grouped following Mudde (2007).

Alongside these local elements identified by extant hate-related literature, the soaring and unprecedented phenomenon of refugees' inflow towards Europe may represent another relevant factor to analyse. Refugees entail out-group features with respect to the incumbent population in receiving destinations (Fratesi et al., 2019) and they are increasingly targeted within the political debate as absorbing tight social benefit budget (Bratti et al., 2017; Dustmann et al., 2016; Steinmayr, 2016). The Italian strategy to deal with refugees' inflow is the Protection System for Asylum Seekers and Refugees (SPRAR), created by Law No 189/2002. The System consists on a decentralized network of local institutions implementing reception projects for forced migrants. SPRARs have been one of the rising targets of hate events (OSCE-ODHIR, 2017): noteworthy, in several cases, rumours and speculations about upcoming SPRAR openings have been sufficient to fuel hate-related vandalism, protests and speeches (Lunaria, 2017). Thus, the geography of SPRAR is included as a spatial element potentially related to hate, allowing to empirical investigation of another research question:

2) *Is the unprecedented phenomenon of refugees' inflow related with the geography of hate?*

1.3 Data description

The starting point is the build-up of a novel dataset merging new data on hate manifestations along with other data on administrative and cultural features. Hate data are taken from the database "*Cronache di ordinario razzismo*" realized by Lunaria[5], an Italian non-for-profit organizations collaborating with OSCE in the realisation of the Hate Crime Data Reporting[6] and with the Fundamental Rights Agency of the European

[5] www.lunaria.org
[6] http://hatecrime.osce.org

Union[7]. Each episode listed in the Lunaria database is categorised by the nature of the event: damages to buildings or properties, discriminations, verbal violences, physical assault. The database provides information about the time and place of the event, people involved and legal consequences. The same dataset is used in the Italian White Paper on Racism. Measuring hate through data collected by journalists, NGOs and civil society groups is increasingly pursued by institutions in many countries, to overcome the limitations inherent to structural under-reporting from hate crime victims (Gerstenfeld, 2017; OSCE-ODHIR, 2017; Pro-Publica, 2019). Data spam from the beginning of 2016 to the beginning of 2017. Overall, 253 out of 611 LLMAs experienced at least one hate manifestation in the considered time-span, with a mean value of 3.4. Table 1 summarizes the descriptive statistics for the considered hate events, both in absolute counts and in number of hate events per 10000 inhabitants.

Table 2: Summary statistics for hate manifestatons

Variable	Number obs	Mean	Standard Deviation	Min	Max
Hate events	611	1.421	4.03	0	50
Hate events >0	253	3.431	5.69	1	50
Hate events per 100000 inhabitants	611	1.421	3.263	0	28.232
Hate events per 100000 inhabitants	253	3.420	4.346	0.256	28.232

Empirical estimation refers to the 611 Italian Local Labour Market Areas (LLMAs) as units of observation: sub-regional geographical areas, identified by ISTAT, where the bulk of the labour force lives and works, and where establishments can find the largest amount of the labour force necessary to occupy the offered jobs. This spatial unit of observation appears to be suitable to comply with hate manifestations as the latter often involve population of neighbouring municipalities[9]. Moreover, LLMAs allow to account for possible spillover effects due to the fact that immigration flows and/or a refugee-hosting structure in a single municipality could also affect the surrounding municipalities through the mobility decisions of agent and/or information flows (Barone et al., 2016).

The dataset merges administrative, cultural and hate data, mapping them on LLMAs geography.

[7] https://fra.europa.eu/en/databases/anti-muslim-hatred/node/2031

[9] i.e. hate events against refugees' hosting centres often gather population from neighbouring towns as well as hate behaviours related to sport events. Other examples refer to hate against students happening in schools serving different municipalities (Lunaria, 2017)

Figure 1 Geography of hate, foreign resident population, SPRAR and pre-Unitarian states in the Italian LLMAs

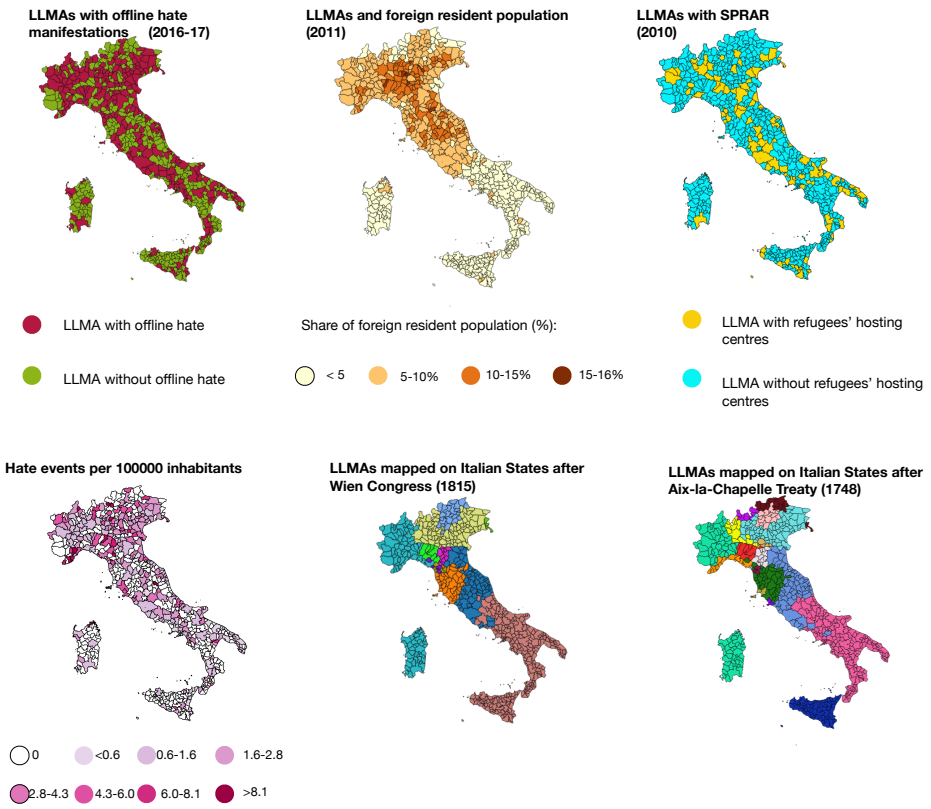


Figure 1 portrays occurrence of hate, foreign resident populations and the geography of refugees hosting centres on the 611 LLMAs. With respect to refugees' hosting structures (SPRAR), the distance between each LLMA and the closest municipality with a SPRAR is computed, using the latitude and the longitude of the centroid of each LLMAs. The lower part of Figure 1 outlines the geography of Italian pre-Unitarian States whose cultural outlooks is acknowledged to be still relevant nowadays.

Table 2 contains the detailed description of the variables considered in the analysis. The classification of LLMAs in either inner or central areas depends on the categorization of the majority of each LLMA territory. Coming to the presence of hub, the pursued classification divides LLMAs containing at least one hub from LLMAs without any hub[10]. Lagged explanatory variables are used.

[10] Data sources and descriptive statistics are detailed in Annex A2, Table A2.1 and Table A2.2

Table 2: Data description

Variable	Description
HATE (YN)	Dummy variable taking value 1 if hate event in a LLMA; 0 if otherwise
Hate manifestations	Count variable measuring the number of hate events in each LLMA
LLMAs sociodemographic outlook	
Foreign resident population	Share of non-Italian resident population as % of total resident population
Split population	Share of divorced and separated couples as % of total resident population
Aging index	Ratio of LLMA resident population over 65 on resident population over 15
Population density	resident population per hectare (ha)
Crime	Number of recorded crimes in LLMA every 1000 inhabitants
Population	Resident population in millions
Non-Western foreign resident population	Share of resident population with either African or Asian citizenship as % of total resident population
Diversity Index	Measure of ethnic diversity in % computed using Alesina *et al.* (2003) index for ethnic fractionalization
LLMAs economic features	
Human capital	Share of population with a college degree as % of total resident population older than 20
Unemployment	Share of unemployed as % of labour supply
Gini index	Gini index for income inequality in % computed on tax declarations of resident households
Trust and collaboration	
Voting turnout	Share of voting turnout in % at the 2014 European Parliament Elections
Civic capital	Dummy variable taking value 1 if at least one municipality was awarded the valour medal for outstanding behaviour in the Liberation War by the President of the Italian Republic; 0 otherwise
Social networks	Number of local units of non-for-profit organizations active every 1000 inhabitants
LLMAs geographic characteristics	
Distance from closest refugees' hosting facility	LLMA centroid distance from the closest refugees' hosting structure measured in decimal degrees
Hub in the LLMA	Dummy variable taking value 1 if LLMA contains at least a municipality or a cluster of municipalities acting as catchment areas; 0 otherwise
Central Area	Dummy variable taking value 1 if the majority of the LLMA is classified as central rather than inner; 0 otherwise
LLMAs political preferences	
Right parties' votes	Share of vote in % given for rightist parties (Northern League, FdI, Forza Italia) at the 2014 European Parliament Elections
Left parties' votes	Share of vote in % given for leftist parties (Democratic Party, Green Party, L'Altra Europa IdV) at the 2014 European Parliament Elections
5 Stars votes	Share of vote in % given for 5 Stars Movement at the 2014 European Parliament Elections
Centre parties' votes	Share of vote in % given for centre parties (Scelta Europea, Io Cambio, NCD) at the 2014 European Parliament Elections
LLMAs cultural outlook	
Part of a pre-Unitarian state in 1748	Categorical variable mapping LLMAs on Italian pre-Unitarian states outlook after Aix-La-Chapelle Treaty

1.4 Empirical methodology

The outcome variable is a count variable displaying many zero: 358 out of 611 LLMAs have zero hate manifestations[11.] Moreover, it cannot be assumed that there exists any LLMA a priori immune from experiencing hatred: every Italian region displays at least one hate manifestation, as well as 99 provinces out of 110. These features of the dependent variable suggest estimation based on hurdle model (Heilbron, 1994; Mullahy, 1986)[12]. The basic idea of hurdle model is that the data-generating process is driven by two different arrays of parameters and that there can be a systematic difference in the statistical process governing the "*hurdle*" with respect to the statistical process governing the subsequent count for positive manifestations (Cameron & Trivedi, 1998). In the case considered in the paper, the statistical process referring to occurrence of hate and the statistical process governing the expected frequency of hate manifestation, given that at least one manifestation has occurred, are assumed to be two separate processes. Estimates of the parameter vectors can be obtained by separate maximization of the two log-likelihood functions. The probability that a hate manifestation will happen in a LLMA *i* is specified through a logit model as follows,

$$P(y_i = 0|X_i) = \frac{\exp(X_i'\gamma)}{1+\exp(X_i'\gamma)} \tag{1}$$

A truncated-at-zero negative binomial (TNB) model is then used to account for the number of hate manifestations, defined over the sample of LLMAs with positive count, as follows

$$P(y_i = j|y_i > 0, W_i) = \frac{P(y_i = j|W_i)}{1-P(y_i = 0|W_i)} = \frac{\Gamma(\theta+y_i)}{\Gamma(y_i+1)\Gamma(\theta)} r_i^{y_i}(1-r_i)^{\theta}, \tag{2}$$

$$r_i = \frac{\lambda_i}{\lambda_i+\theta} \; and \; \lambda_i = \exp(W_i'\beta) \tag{3), (4}$$

[11] See also Figure A2.1 in Annex A2.

[12] The chosen modelling strategy is assessed against other model specifications: Poisson, ZIP, ZINB, NBREG, Hurdle (NBREG), Hurdle (Poisson) Specification tests show that the hurdle model is preferred to alternative model specifications (see Annex A5).

By separating the occurrence of any hate manifestation from the expected frequency, it is possible to assess whether any variable exerts its effect mainly through the occurrence or the frequency of events. The approach adopted in the paper aims to account for a wide range of features potentially influencing hate; thus, several model specifications are estimated, progressively augmenting the set of covariates in X_i and W_i. In all specifications regional-level fixed effects and the share of foreign resident population are included; the latter deserves a throughout exploration due to its high relevance in recent works on voting decisions (i.a. Barone et al., 2016; Colussi & Pestel, 2017; Crescenzi, Di Cataldo, & Faggian, 2018; Halla et al., 2017). Regional fixed effects allow to capture all the features of the region that may affect the occurrence of hate manifestations, such as the regional degree of exposure to globalization and regional policies governing migration, refugees and ethnic inclusiveness. In all regressions, standard errors are clustered at regional level. Then, through postestimation diagnostics and robustness check the fitted model specification is determined, containing only robust significant covariates. The same steps are pursued for both stages of the hurdle model.

1.5 Results

Hurdle model stage 1: binary logit for the occurrence of hate manifestation. Table 3 reports the fitted model specification estimation results[13], after goodness-of-fit and robustness check for the selection of core-covariates. Standard errors are clustered at regional level in column 1 and at province level in column 2. With respect to non-nationals, the association of the share of foreign resident population with the occurrence of hate is not significant. At the same time, however, the effect of refugees' hosting structure location displays a mild significance also when controlling for socioeconomic, geographic and cultural confounding factors. To this respect, the farther away from the LLMA the refugees' hosting centre is located, the lower the occurrence of a hate manifestation.

Notably, also anecdotal evidence on hate manifestations confirms an important turnout of refugees-related events: in many cases rumours on hypothetical opening of refugees' hosting structures were sufficient to trigger hate events. This is consistent with

[13] Detailed results for the considered model specifications are available in the Annex A.3, Table A3.1

established literature on hatred showing that intergroup hatreds are generally based on stories about the hated group that do not need to be supported by facts (Glaeser, 2005; Hainmueller & Hopkins, 2014). The refugees hosting centres may work as a source for uncertainty at the local level and, consequently, the shock generates fear linked to a sociocultural threat. This result also relates to literature on refugees' effects on labour market, where it has been established that refugees' arrival acts as an exogenous shock (Borjas and Monras, 2017).

Table 3: 1st stage hurdle model logit estimates with regional and province FE for fitted specification

HATEYN	(1) Odd ratios	(2) Odd ratios Province FE
split pop	1.564**	1.564**
	(1.003 - 2.440)	(1.084 - 2.257)
pop density (ln)	0.694**	0.694**
	(0.492 - 0.978)	(0.510 - 0.945)
population (ln)	4.043***	4.043***
	(2.831 - 5.772)	(2.636 - 6.199)
educated people	0.588	0.588
	(0.288 - 1.201)	(0.306 - 1.130)
educated pop (sq)	1.028*	1.028*
	(0.997 - 1.059)	(0.998 - 1.058)
trust	0.965**	0.965**
	(0.936 - 0.996)	(0.935 - 0.997)
distance SPRAR (ln)	0.784*	0.784*
	(0.596 - 1.033)	(0.597 - 1.031)
Right parties' votes	1.075***	1.075***
	(1.033 - 1.118)	(1.031 - 1.120)
5-star votes	1.089***	1.089**
	(1.029 - 1.152)	(1.017 - 1.166)
Observations	611	611
Regional FE	YES	NO
Province FE	NO	YES

Confidence interval in parentheses (errors clustered at regional level in column 1 and at province level in column 2)
**** p<0.01, ** p<0.05, * p<0.1*

The remaining variables behave consistently with the literature. Social stress, proxied by the rate of divorced/separated couples, is a significant risk factor for the occurrence of hate events, analogously to existing findings on the determinants of voting for far-right parties (Halla et al., 2017) and on the determinants of hate groups in the US (Jefferson & Pryor, 1999). Population density displays a negative association, supporting contact theory claims that frequent interaction between different groups reduces intolerance. Crime rate is never significant[14]. This result allows to argue that there is consistency with previous findings showing that relationship between hatred and the

[14] Subsets of crime -homicides, robberies, thefts and usury- have been considered, but again there is non-significance of any of the different subsets.

criminality of the hated group is often minimal, as the "*crime*" topic serves mainly as rhetoric tool to sustain self-consistency of hate (Glaeser, 2005). Population size is always positively related with the occurrence of hate, suggesting that bigger LLMAs favour the appearance of hate events. With respect to economic dimension, we have a slightly positive relationship between high-educated population and hatred occurrence. Trust in institutions displays a negative influence, consistently with the positive role of trust in reducing uncertainty. Controls for political preferences work accordingly to established findings on voting: anti-elite discourse and anti-minorities act pandering fear (Glaeser, 2005; Guiso, Herrera, Morelli, & Sonno, 2017). The fact that the LLMA is located in either a central or inner area and the presence of major catchment areas appears not to be significant. Overall, for the first stage of the hurdle model, the drivers of hate occurrence belong to the sociocultural sphere, rather than to the economic one.

Postestimation Diagnostic and Robustness Check. The fitted model specification results presented in Table 3 have been identified applying goodness-of-fit and robustness check to the whole set of potential covariates. Wald test, likelihood ratio test and selection criteria based on the Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC) allow to identify the following potential core-covariates: divorced/separated couples, population density, population size, human capital, trust, distance from the closest SPRAR, rightist parties votes and 5 stars votes. Subsequently 2048 regressions have been performed where the dependent binary variable is regressed on core variables - which are included in all regressions - and all possible combinations of other non-core variables. Robustness check confirms the findings for all core covariates. Assessments for goodness-of-fit, specification errors, multicollinearity, sensitivity and specificity do not outline any concern[15.]. The role of big LLMAs (Milan and Rome) as leverage points is assessed by estimating the fitted model with and without the two LMMAs: again, results do not change as shown in columns (1) in Table 4. The empirical investigation has outlined that the geography of refugees' hosting facilities depicts some level of significance in predicting hate occurrence. Since this geography has been changing from 2010, following the increasing refugees' inflow, it is further analysed whether the result holds when the geography of refugees' hosting centres changes.

[15] In Appendix A3, for completeness, results for postestimation and robustness check are presented.

Hence, the SPRAR geography in 2013 is introduced, since this year displays another upward twist in asylum seekers inflows, driving a consequent increase in the refugees' hosting centres. In 2010 the SPRAR hosting capacity amounted to 3146 places (SPRAR, 2010), in 2013 the hosting capacity rose up to 10000 places[16] (SPRAR, 2014). Column (2) in Table 4 shows estimation results for the fitted model when distance from the closest SPRAR in 2013 is used. Results still hold.

Table 4: 1st-stage hurdle model logit estimates with regional FE for fitted model with different robustness checks

HATEYN	(1) Odds Ratio *no Rome&Milan*	(2) Odds Ratio *SPRAR 2013*
split pop	1.564**	1.561**
	(1.003 - 2.440)	(1.006 - 2.424)
pop density (ln)	0.694**	0.688**
	(0.492 - 0.978)	(0.488 - 0.971)
population (ln)	4.042***	4.054***
	(2.831 - 5.772)	(2.834 - 5.800)
educated people	0.588	0.592
	(0.288 - 1.202)	(0.283 - 1.238)
educated pop (sq)	1.028*	1.027*
	(0.997 - 1.059)	(0.996 - 1.060)
Trust	0.965**	0.965**
	(0.936 - 0.996)	(0.935 - 0.995)
Distance SPRAR (ln)	0.784*	0.744*
	(0.596 - 1.033)	(0.532 - 1.039)
Right parties' votes (EU14)	1.075***	1.073***
	(1.033 - 1.118)	(1.032 - 1.116)
5-star votes (EU14)	1.089***	1.091***
	(1.029 - 1.152)	(1.031 - 1.154)
Observations	609	611
Regional FE	YES	YES

Confidence interval in parentheses (errors clustered at regional level)
**** p<0.01, ** p<0.05, * p<0.1*

The influence of non-Italian residents is further scrutinised, by comparing results under the following specifications: *(i)* the baseline fitted specification with the share of foreign resident population (column 1 -Table 5) *(ii)* the increase in ethnic diversity nested with the augmenting migration flows, to check whether ethnic varieties may relate to hate through the so called "*diversity fatigue*" (column 2 -Table 5)[17]; *(iii)* non-western foreign resident population (column 3- Table 5). None specification alters results, further supporting the insight that legal migration inflow adjusts within the local economy.

[16] See also figure A3.3 in Annex A3

[17] 3 indexes for ethnic diversity have been considered: the fractionalization index, the Theil index and the fractionalization index developed by Alesina, Devleeschauwer, Easterly, Kurlat, & Wacziarg, (2003). For the sake of brevity, here only results for the Alesina et al. (2003) fractionalization index are showed, since findings do not change for any of the indexes.

Finally, to address the role of persistency of the geography of culture, pre-Unitarian states are introduced. They are an enduring cultural feature of Italy, since the current outlook of Italian regions spurs directly from pre-Unitarian states' identities (Broers, 2003). Thus, the fitted model specification is estimated consolidating existing LLMAs geography on two pre-Unitarian states configurations: 1748 (Aix-La-Chapelle Treaty) and 1815 (Wien Congress).

Table 5: 1st-stage hurdle model logit estimates with regional FE for fitted model with different measures for foreign resident population

HATEYN	(1) Odds Ratio *foreign res pop*	(2) Odds Ratio *ethnic diversity*	(3) Odds Ratio *non-western res*
foreign resident pop	1.025		
	(0.925 - 1.137)		
ethnic diversity (%)		0.991	
		(0.825 - 1.192)	
sh_nonwest			58.85
			(4.55e-07 - 7.608e+09)
split pop	1.573**	1.562**	1.589**
	(1.016 - 2.435)	(1.011 - 2.411)	(1.041 - 2.425)
pop density (ln)	0.679**	0.698**	0.671**
	(0.492 - 0.938)	(0.503 - 0.968)	(0.481 - 0.936)
population (ln)	4.060***	4.047***	4.035***
	(2.874 - 5.737)	(2.812 - 5.826)	(2.837 - 5.738)
educated people	0.582	0.589	0.592
	(0.281 - 1.209)	(0.291 - 1.193)	(0.281 - 1.245)
educated pop (sq)	1.028*	1.027*	1.028*
	(0.997 - 1.060)	(0.997 - 1.059)	(0.996 - 1.060)
trust	0.964**	0.966**	0.964**
	(0.934 - 0.996)	(0.936 - 0.997)	(0.934 - 0.996)
Distance SPRAR (ln)	0.780*	0.785*	0.779*
	(0.588 - 1.034)	(0.595 - 1.037)	(0.586 - 1.034)
Right parties' votes (EU14)	1.075***	1.075***	1.076***
	(1.032 - 1.119)	(1.033 - 1.118)	(1.031 - 1.124)
5-star votes (EU14)	1.088***	1.089***	1.089***
	1.573**	(1.029 - 1.152)	(1.031 - 1.151)
Observations	611	611	611
Regional FE	YES	YES	YES

*Confidence interval in parentheses (errors clustered at regional level) *** p<0.01, ** p<0.05, * p<0.1*

More into details, the model is estimated considering pre-Unitarian fixed effects instead of regional fixed effects. Results are outlined columns 1 and 2 in Table 6 below. Notably, allowing for the cultural geography rather than for the administrative geography remarkably improves the significance of the effect of the refugees' hosting facilities. This finding appears to support the fact that the relationship between refugees' hosting centres and hate events is enacted through the dimension of a cultural threat.

Table 6: 1st-stage hurdle model logit estimates with regional FE for fitted model with pre-unitarian geography

HATEYN	(1) Odds Ratio *Aix-La-Chapelle Treaty (1748)*	(2) Odds Ratio *Wien Congress (1815)*
split pop	1.467***	1.560***
	(1.132 - 1.901)	(1.211 - 2.010)
pop density (ln)	0.773***	0.748***
	(0.639 - 0.935)	(0.630 - 0.889)
population (ln)	3.690***	3.488***
	(2.715 - 5.015)	(2.749 - 4.425)
educated people	0.649	0.680
	(0.361 - 1.166)	(0.376 - 1.230)
educated pop (sq)	1.024*	1.023
	(0.999 - 1.050)	(0.994 - 1.053)
Trust	0.967**	0.975*
	(0.942 - 0.993)	(0.950 - 1.001)
Distance SPRAR (ln)	**0.740*** **	**0.792** **
	(0.632 - 0.866)	**(0.656 - 0.957)**
Right parties' votes (EU14)	1.066***	1.075***
	(1.031 - 1.103)	(1.028 - 1.124)
5-star votes (EU14)	1.044*	1.063***
	(0.996 - 1.093)	(1.040 - 1.087)
Observations	611	611
Regional FE	NO	NO
Pre-unitarian States FE	YES	YES

*Confidence interval in parentheses (errors clustered at regional level) *** p<0.01, ** p<0.05, * p<0.1*

Hurdle model stage 2: TNB model for the frequency of hate manifestation. Stage 1 analysis has identified the local factors associated with hate occurrence. The second stage estimates which local features are associated with expected frequency of hate, once the "*hurdle*" of experiencing at least one hate manifestation has occurred. As in stage 1, several specifications for a TNB are estimated, progressively augmenting the set of controls (results for the different model specifications in Annex A4 -Table A4.1). Then, through goodness-of-fit and robustness checks the fitted model specification is identified, whose estimation delivers the results summarized in Table 7, with robust standard errors clustered at regional level in column 1 and robust standard errors clustered at province level in column 2.

Distance from the closest refugees hosting centre still displays a meaningful association. In this second stage, once that at least one hate event has occurred, the further away the refugees hosting centre is located, the more the frequency of hate events gets reduced. This finding provides additional support for refugees' hosting structures acting as a local risk factor.

Table 7: 2nd stage hurdle model TNB estimates for fitted model specification

hate frequency	(1) Fitted model -IRR- cluster Region	(2) Fitted model -IRR- cluster Province
pop density (ln)	0.809*	0.809
	(0.641 - 1.021)	(0.611 - 1.072)
population (ln)	3.019***	3.019***
	(2.379 - 3.832)	(2.321 - 3.927)
distance SPRAR_10 (ln)	**0.880****	**0.880***
	(0.786 - 0.985)	**(0.758 - 1.022)**
Hub in the LLMA	1.420**	1.420**
	(1.020 - 1.978)	(1.005 - 2.007)
Centre parties' votes	1.023*	1.023
	(1.000 - 1.048)	(0.994 - 1.054)
Right parties' votes	1.108**	1.108***
	(1.005 - 1.221)	(1.025 - 1.197)
Observations	253	253
Regional FE	YES	NO
Province FE	NO	YES

Confidence interval in parentheses (errors clustered at regional level in column 1 and at province level in column 2)
**** p<0.01, ** p<0.05, * p<0.1*

Looking at confounding features, population density displays significant negative association with the expected rate of hate manifestation, again relating to the "*contact theory*"; the size of the LLMA is still positively associated with expected frequency of hate. Results about political preferences suggest that non-populist parties adapt to populist concerns once intolerance has gained visibility within the public discourse (Glaeser, 2005; Guiso et al., 2017; Halla et al., 2017). Finally, the presence of a hub in the LLMA acts as a risk factor. This finding suggests that hubs are the preferred destination for marginalized people in need for social services, determining competition for scarce public resources and, consequently, social tensions.

Remarkably, fewer conditions are at work once the hurdle of experiencing at least one hate event has been crossed.

Goodness-of-fit and Robustness Check. The fitted model specification in Table 6 has been identified through goodness-of-fit tests (Wald test, likelihood ratio test, AIC and BIC) to identify the following set of core covariates: population density, population size, distance from closest SPRAR, presence of a hub in the LLMA, rightist parties vote and centre parties vote. To check for robustness of the fitted model specification, it is assessed whether the identified core-covariates are sensitive to the inclusion of additional controls. Robustness check, together with goodness-of-fit tests, confirm that the fitted model specification delivers robust identification of predictors. It is also confirmed the

non-significance of the following testing covariates: share of foreign resident population, crime rate, aging index, share of divorced/separated couples, the LLMA being a central area, voting either leftist parties or 5 stars, non-for-profit organizations local units and civic capital. There are no concerns for multicollinearity[18.] Again, the role of big LLMAs is gauged, by estimating the fitted model specification with and without Rome and Milan: results do not change (columns 1 in Table 8). Since the geographical allocation of refugees' hosting structures holds a significant association with expected frequency of hate manifestation, the TNB fitted model is estimated changing the refugees' hosting centres geography as in stage 1. That is, considering SPRAR geography in 2013. Results are summarized in Table 8, columns 2.

Table 8: 2nd stage hurdle model TNB estimates without influential points and with a different geography for refugees' hosting centres

hate frequency	(1) -IRR- *no Rome&Milan*	(2) -IRR- *SPRAR_13*
pop density (ln)	0.804*	0.802*
	(0.636 - 1.016)	(0.634 - 1.013)
population (ln)	3.075***	3.005***
	(2.343 - 4.035)	(2.337 - 3.865)
distance SPRAR (ln)	**0.870****	**0.836*****
	(0.774 - 0.977)	**(0.740 - 0.945)**
Hub in the LLMA	1.413**	1.389*
	(1.010 - 1.976)	(0.993 - 1.943)
Right parties' votes	1.024**	1.024*
	(1.000 - 1.048)	(1.000 - 1.049)
Centre parties' votes	1.112**	1.108**
	(1.006 - 1.230)	(1.008 - 1.217)
Observations	251	235
Regional FE	YES	YES

Confidence interval in parentheses (errors clustered at regional level)
**** p<0.01, ** p<0.05, * p<0.1*

Then, it is controlled for different ways to consider ethnic variety, again considering ethnic diversity and the share of non-western population. Results presented in columns 1-3 in Table 9 confirm the non-significance of foreign resident population on expected frequency of hate. Finally, due to the reduced number of units of observations, estimation convergence with pre-Unitarian State fixed effect cannot be reached.

[18] see Table A4.2 in appendix A4 for details

Table 9: 2nd stage hurdle model TNB estimates controlling for different measures for foreign resident population

hate frequency	(1) -IRR- *foreign res pop*	(2) -IRR- *ethnic diversity*	(3) -IRR- *non-western res*
foreign resident pop	1.016		
	(0.956 - 1.080)		
ethnic diversity (%)		1.030	
		(0.939 - 1.130)	
sh_nonwest			129.7
			(0.306 - 55,012)
pop density (ln)	0.803*	0.802*	0.790*
	(0.628 - 1.027)	(0.626 - 1.028)	(0.614 - 1.017)
population (ln)	2.996***	2.972***	3.006***
	(2.381 - 3.769)	(2.375 - 3.718)	(2.383 - 3.791)
distance SPRAR (ln)	**0.875****	**0.876****	**0.875****
	(0.778 - 0.984)	**(0.781 - 0.983)**	**(0.785 - 0.976)**
Hub in the LLMA	1.424**	1.426**	1.443**
	(1.019 - 1.990)	(1.016 - 2.000)	(1.029 - 2.024)
Right parties' votes	1.022*	1.022*	1.022*
	(0.998 - 1.047)	(0.997 - 1.047)	(0.997 - 1.048)
Centre parties' votes	1.107**	1.108**	1.110**
	(1.003 - 1.221)	(1.004 - 1.222)	(1.005 - 1.226)
Observations	253	253	253
Regional FE	YES	YES	YES

Confidence interval in parentheses (errors clustered at regional level)
**** p<0.01, ** p<0.05, * p<0.1*

1.6 Conclusions

Local factors have been increasingly acknowledged as associated to hate by many scholars in criminology and social psychology. This paper provides an empirical investigation about whether this is the case with respect to the Italian context, where hate events have been increasing at a fast rate in the past years, overcoming the European average trend. Evidence supports the association between several local features and hate events, both with respect to the occurrence and the expected frequency. Hence, there is evidence for a relevant role of geography on the manifestation of hatred.

More into details, the effect of changes in the local outlook of non-nationals appears to depend on the characteristics of migration itself. On the one hand, foreign resident population does not exert any influence on hate manifestations, even considering only the more distant cultures. On the other hand, refugees' hosting centres appear to be a predictor of both hate occurrence and expected hate frequency. One possible explanation, relating also to literature on refugees' effects on labour market (Borjas and Monras, 2017) and to the literature on sociocultural determinants of intolerance, implies that asylum seekers hosting facilities are relevant sources of perceived sociocultural

threats, since they are perceived as abrupt changes, or shocks, to the local sociocultural structure.

Besides migrants, hate occurrence is also associated to social frustration, driven by the gap between traditional societal ambitions/hierarchies and social changes. Remarkably, actual crime figures do not exert any significant effect, consistently with theoretical arguments arguing that crime is mainly a rhetoric tool to sustain self-consistency of hate (Glaeser, 2005).

Trust in institution represents an important channel to counter hate, as well as densely populated places where interaction among different people is eased, thus allowing for more frequent information flows and uncertainty reduction, which, in turn, reduce the negative effects of prejudices.

Since the analysis has identified sociocultural elements rather than economic features capable of influencing the occurrence and the frequency of hatred, overall the Italian case aligns to existing contributions showing that the sociocultural sphere is more relevant than the economic one in triggering hate (Card et al., 2012; Hainmueller & Hopkins, 2014; Perry, 2001).

Another interesting finding of the paper refers to the fact that once hatred has been experienced by a community, fewer conditions are needed for triggering further repetitions of hate. This result suggests that the overcoming of a social desirability barrier may unleash an escalation of hate driven by existing social frustration.

From a policy perspective, findings support both the need for investigation of situational factors to disentangle hate determinants and channels of transmission and the importance of local features.

References

Alesina, A., Devleeschauwer, A., Easterly, W., Kurlat, S., & Wacziarg, R. (2003). Fractionalization. *Journal of Economic Growth*, *8*(2), 155–194. https://doi.org/10.1023/A:1024471506938

Allport, G. W. (1988). *The Nature Of Prejudice.* Addison-Wesley.

Assimakopoulos, S., Baider, F. H., & Millar, S. (2017). *Online Hate Speech in the European Union: A Discourse- Analytic Perspective.* https://doi.org/10.1007/978-3-319-72604-5

Barca, F., Casavola, P., & Lucatelli, S. (2014). *A Strategy for Inner Areas in Italy: Definitions, Objectives, Tools and Governance.*

Barone, G., D'Ignazio, A., de Blasio, G., & Naticchioni, P. (2016). Mr. Rossi, Mr. Hu and politics. The role of immigration in shaping natives' voting behavior. *Journal of Public Economics*, *136*(8228), 1–13. https://doi.org/10.1016/j.jpubeco.2016.03.002

Baumeister, R., & Vohs, K. (2004). Four Roots of Evil. In A.G Miller (Ed.), *The Socio-Psychology of Good and Evil.* Guilford.

Becker, S. O., Fetzer, T., & Novy, D. (2017). Who voted for Brexit? A comprehensive district-level analysis. *Economic Policy*, *32*(92), 601–650. https://doi.org/10.1093/epolic/eix012

Bianchi, M., Buonanno, P., Pinotti, P. (2012). Do Immigrants Cause Crime ? To cite this version : Do Immigrants Cause Crime ? ∗ Universit e. *Journal of the European Economic Association*, *10*(6), 1318–1347.

Borjas, G. J. (2017). The Labor Market Consequences of Refugee Supply Shocks. *Economic Policy*, *32*(91), 361–413.

Bratti, M., Deiana, C., Havari, E., Mazzarella, G., & Meroni, E. C. (2017). What are You Voting for? Proximity to Refugee Reception Centres and Voting in the 2016 Italian Constitutional Referendum. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3045757

Broers, M. (2003). The Myth and Reality of Italian Regionalism :A Historical Geography of Napoleonic Italy, 1801–1814. *The American Historical Review*, *18*(3), 688–709.

Cameron, A. C., & Trivedi, P. K. (1998). *Regression Analysis of Count Data(2nd Edition).* Cambridge University Press.

Card, D., Dustmann, C., & Preston, I. (2012). Immigration, wages, and compositional amenities. *Journal of the European Economic Association*, *10*(1), 78–119. https://doi.org/10.1111/j.1542-4774.2011.01051.x

Cerreti C. (2011). La Rappresentazione del Territorio. In *L'Unificazione Italiana.* Treccani

Colussi, T., & Pestel, N. (2017). Minority Salience and Political Extremism ∗.

Craig, K. M. (2002). Examining hate-motivated aggression: A review of the social psychological literature on hate crimes as a distinct form of aggression. *Aggression and Violent Behavior*, *7*(1), 85–101. https://doi.org/10.1016/S1359-1789(00)00039-2

Crescenzi, R., Di Cataldo, M., & Faggian, A. (2018). Internationalized at work and localistic at home: The 'split' Europeanization behind Brexit. *Papers in Regional Science*, *97*(1), 117–132. https://doi.org/10.1111/pirs.12350

Dustmann, C, Vasiljeva, K., & Piil Damm, A. (2016). Refugee Migration and Electoral Outcomes. *CReAM Discussion Paper*, (19/16). https://doi.org/10.1016/S1043-2760(97)84344-5

Dustmann, C., & Fasani, F. (2016). The Effect of Local Area Crime on Mental Health. *Economic Journal*, *126*(593), 978–1017. https://doi.org/10.1111/ecoj.12205

ENAR. (2015). *Racist Crime in Europe. ENAR Shadow Report 2013-2014.*

European Commission. (2017). *Annual report on ECRI's activities 2016*. Retrieved from http://www.coe.int/t/dghl/monitoring/ecri/activities/Annual_Reports/Annual report 2016.pdf

Felice E. (2015). On the reconstruction and interpretation of regional inequality in Italy in the long-run: main results and future lines of research. MPRA Working Paper 66310

Fratesi, U., Percoco, M., & Proietti, P. (2019). Who welcomes them? Determinants of the spatial distribution of asylum seekers in Italy. *Regional Studies*, *53*(1), 118–130. https://doi.org/10.1080/00343404.2018.1525487

Garland, J., & Treadwell, J. (2012). The New Politics of Hate ? An Assessment of the Appeal of the English Defence League Amongst Disadvantaged White Working Class Communities in England 1. *England Journal of Hate Studies*, *10*(1), 123-141.

Gerstenfeld, P. B. (2017). *Hate Crimes: Causes, Controls, and Controversies*. SAGE Publications. Retrieved from https://books.google.it/books?id=v35ZDwAAQBAJ

Glaeser, E. (2014). *Triumph of the City*. Penguin

Glaeser, E. L. (2005). The Political Economy of Hatred*. *Quarterly Journal of Economics*, *120*(1), 45–86. https://doi.org/10.1162/qjec.2005.120.1.45

Green, D. P., McFalls, L. H., & Smith, J. K. (2001). Hate Crime: An Emergent Research Agenda. *Annual Review of Sociology*, *27*(1), 479–504. https://doi.org/10.1146/annurev.soc.27.1.479

Guiso, L, Herrera, H., Morelli, M., & Sonno, T. (2017). Populism: Demand and Supply *, 1–63. Retrieved from http://www.heliosherrera.com/populism.pdf

Guiso, L, Sapienza, P., & Zingales, L. (2016). Long-term persistence. *Journal of the European Economic Association*, *14*(6), 1401–1436. https://doi.org/10.1111/jeea.12177

Hainmueller, J., & Hopkins, D. J. (2014). Public Attitudes Toward Immigration. *Annual Review of Political Science*, *17*, 225–249. https://doi.org/10.1146/annurev-polisci-102512-194818

Hall, N. (2013). *Hate crimes. 2nd Edition*. Routledge.

Halla, M., Wagner, A. F., & Zweimüller, J. (2017). Immigration and Voting for the Far Right. *Journal of the European Economic Association, Volume 15, Issue 6, 1 December 2017, Pages 1341–1385*, (6), 1341–1385.

Hangartner, D., Dinas, E., Marbach, M., Matakos, K., & Xefteris, D. (2018). Does Exposure to the Refugee Crisis Make Natives More Hostile? *American Political Science Review*, (17). https://doi.org/10.2139/ssrn.3042936

Heilbron, D. C. (1994). Zero-Altered and other Regression Models for Count Data with Added Zeros. *Biometrical Journal*, *36*(5), 531–547. https://doi.org/10.1002/bimj.4710360505

Hewstone, M., & Jaspars, J. (1984). Social dimensions of attribution. In H. Tajfel (Ed.), *The social dimension: European developments in social psychology*. Cambridge University Press. Retrieved from https://books.google.it/books?id=ilIIeD5SCKUC

Iganski, P. (2008). *"Hate crime" and the city*. https://doi.org/10.1332/policypress/9781861349408.001.0001

Jefferson P.N. and Pryor F. (1999). On the geography of hate. *Economic Letters*, *65*(5), 389–395. https://doi.org/10.3917/reco.566.1249

Kemeny, T., & Cooke, A. (2017). Urban Immigrant Diversity and Inclusive Institutions. *Economic Geography*, *00*(00), 1–25. https://doi.org/10.1080/00130095.2017.1300056

Lunaria. (2017). *Cronache di Ordinario Razzismo. Quarto Libro Bianco sul Razzismo in Italia.*

Medina, R. M., Nicolosi, E., Brewer, S., & Linke, A. M. (2018). Geographies of organized hate in America: A regional analysis. *Annals of the American Association of Geographers*, *108*(4), 1006–1021. https://doi.org/10.1080/24694452.2017.1411247

Miller, A.G (Ed.). (2004). *The social psychology of good and evil. The social psychology of good and evil.* New York, NY, US: The Guilford Press.

Mudde, C. (2007). *populist radical right parties in Europe*. Cambridge University Press.

Mullahy, J. (1986). Specification and testing of some modified count data models. *Journal of Econometrics*, *33*(3), 341–365. https://doi.org/10.1016/0304-4076(86)90002-3

Newman, B. J. (2013). Acculturating Contexts and Anglo Opposition to Immigration in the United States. *American Journal of Political Science*, *57*(2), 374–390. https://doi.org/10.1111/j.1540-5907.2012.00632.x

OSCE-ODHIR. (2017). *Hate Crime Reporting.*

OSCE-ODIHR. (2009). *Preventing and responding to hate crimes - A resource guide for NGOs in the OSCE region*. Retrieved from http://www.osce.org/odihr/39821?download=true

Otto, A. H., & Steinhardt, M. F. (2014). Immigration and election outcomes: Evidence from city districts in Hamburg. *Regional Science and Urban Economics 45, 67-79*.

Perry, B. (2001). *In the Name of Hate: Understanding Hate Crimes*. Routledge. Retrieved from https://books.google.it/books?id=doDHZXvq19YC

Pettigrew, T. F. (1979). The Ultimate Attribution Error: Extending Allport's Cognitive Analysis of Prejudice. *Personality and Social Psychology Bulletin*, *5*(4), 461–476. https://doi.org/10.1177/014616727900500407

Pettigrew, T. F., & Tropp, L. R. (2006). A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology*, *90*(5), 751–783. https://doi.org/10.1037/0022-3514.90.5.751

Pro-Publica. (2019). Hate News Index. Retrieved from https://projects.propublica.org/hate-news-index/

Putnam, R. (2007). *Diversity and community in the twenty-first century the 2006 Johan Skytte prize lecture". Housing Policy Debate* (Vol. 19). https://doi.org/10.1080/10511482.2008.9521631

Rodríguez-Pose, A. (2018). The revenge of the places that don't matter (and what to do about it). *Cambridge Journal of Regions, Economy and Society*, *11*(1), 189–209. https://doi.org/10.1093/cjres/rsx024

SPRAR. (2010). *Al di là dell'Emergenza. Compendio Statistico del Sistema di Protezione per Richiedenti Asilo e Rifugiati - Anno 2010.*

SPRAR. (2014). *Rapporto Annuale SPRAR. Sistema di Protezione per Richiedenti Asilo e Rifugiati.*

Staub, E. (2004). Basic human needs, altruism, and aggression. In *The social psychology of good and evil.* (pp. 51–84). New York, NY, US: The Guilford Press.

Steinmayr, A. (2016). Exposure to Refugees and Voting for the Far-Right: (Unexpected) Results from Austria. *IZA Discussion Paper*, *9790*(9790), 63.

UK Crown Prosecution Service. (2018). Hate Crime Strategy 2017-2020.

US Hate Crime Statistics Act. (2010). Retrieved from https://ucr.fbi.gov/hate-crime/2010/resources/hate-crime-2010-hate-crime-statistics-act.pdf

Veneri, P., & Murtin, F. (2019). Where are the highest living standards? Measuring well-being and inclusiveness in OECD regions. Regional Studies, 53(5), 657–666. https://doi.org/10.1080/00343404.2018.1463091

Voigtlander, N., & Voth, H. (2012). Persecution Perpetuated: the Medieval Origins of anti-Semitic Violence in Nazi Germany. Quarterly Journal of Economics, 127(2), 1339–1392. https://doi.org/10.1093/qje/qjs019.Advance

Waller, J. (2014). Becoming Evil. Genocide: A Reader.

A.1 Facts and Figures on Hate trends

Relevant trend of a subset of hate events: xenophobic hate crimes recorded by the police in Italy (Figure A1.1) and in a panel of 14 European countries (including Italy) (Figure A1.2). Clearly, Italy is experiencing a faster growth rate with respect to the European panel.

Figure A1.1 Italy and 14 EU countries: trend in xenophobic hate crimes 2013-2016

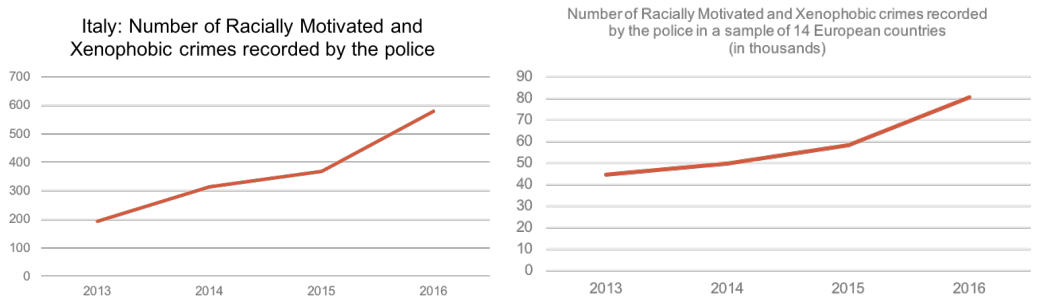
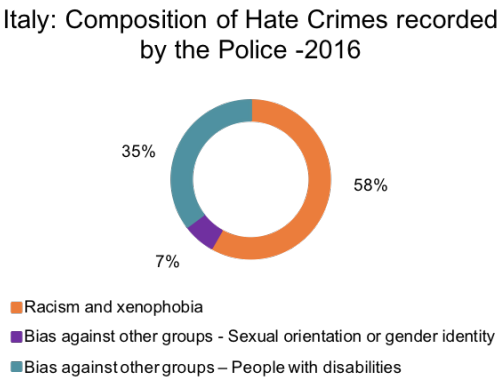


Figure A1.2 Italy: Hate crimes shares by target in 2016 (OSCE-ODHIR 2016)



A.2. Descriptive statistics

Table A2. 3: Data source

Variable	Source
HATE (YN)	Lunaria database on hate manifestation (2016/17)
Hate manifestations	Lunaria database on hate manifestation (2016/2017)
LLMAs sociodemographic outlook	
Foreign resident population	ISTAT -15th Population Census (2011)
Split population	ISTAT -15th Population Census (2011)
Aging index	ISTAT -15th Population Census (2011)
Population density	ISTAT -15th Population Census (2011)
Crime	ISTAT -Growth Policy Indicators: Rule of Law and Order (2011)
Population	ISTAT -15th Population Census (2011)
Non-Western foreign resident population	ISTAT -15th Population Census (2011)
Diversity Index	ISTAT -15th Population Census (2011)
LLMAs economic features	
Human capital	ISTAT -15th Population Census (2011)
Unemployment	ISTAT -15th Population Census (2011)
Gini index	Italian Ministry of Economy and Finance (fiscal year 2011)
Trust and collaboration	
Voting turnout	Ministry of the Interior Elections database (2014)
Civic capital	The President of the Italian Republic database on honours and awards
Social networks	ISTAT -9th Industry and Services Census (2011)
LLMAs geographic characteristics	
Distance from closest refugees' hosting facility	Ministry of the Interior and National Association of Italian Municipalities (ANCI) SPRAR website (2007, 2010, 2013)
Hub in the LLMA	Agency for Territorial Cohesion, Classification of Italian Municipalities according to Inner Areas methodology (2014)

Central Area	Agency for Territorial Cohesion, Classification of Italian Municipalities according to Inner Areas methodology (2014)
LLMAs political preferences	
Right parties' votes	Ministry of the Interior Elections database (2014)
Left parties' votes	Ministry of the Interior Elections database (2014)
5 Stars votes	Ministry of the Interior Elections database (2014)
Centre parties' votes	Ministry of the Interior Elections database (2014)
LLMAs cultural outlook	
Part of a pre-Unitarian state in 1748	Treccani Encyclopedia
Part of a pre-Unitarian state in 1815	Treccani Encyclopedia

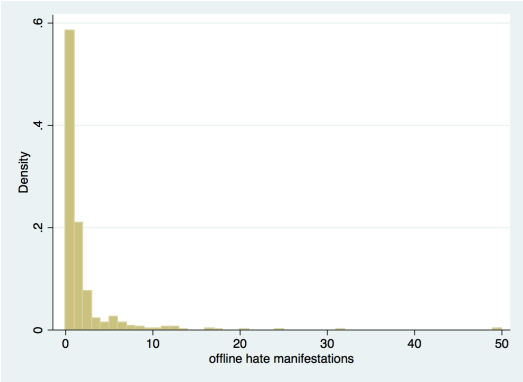
Table A2.4: Descriptive statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Geography of hate (2016-17)					
HATE (YN)	611	0.4124386	0.4926766	0	1
HATE COUNTS	611	1.420622	4.02999	0	50
Socio-demographic features of LLMAs (2011)					
Foreign resident population	611	5.799721	3.777264	0.2714932	16.39703
Split population	611	3.770977	1.509572	0.7702183	8.319397
Aging index[19]	611	16.69316	4.646057	6.93617	34.9384
Population density	611	2.056063	2.915807	0.1041765	31.07485
Crime	611	19.93512	7.503294	7.467119	55.05622
Population	611	0.9727	2.5829	0.03138	36.85101
Economic characteristics of the LLMA (2011)					
Human capital	611	10.28096	2.737488	3.578903	21.07972
Unemployment	611	11.94874	6.136332	1.486878	27.40376
Gini index	611	41.135	2.866571	33.47302	54.87072
Trust endowment of LLMA (2014)					
Voting turnout (trust)	611	56.41075	12.89531	25.47039	85.69295
Civic Capital	611	0.2684124	0.4434965	0	1
Social networks	611	5.528334	2.32694	.9486006	14.95564
Geographic characteristics of LLMA					
Distance from closest refugees' hosting facility in 2010	611	3.6711	2.8916	0. 0777543	19.23096

[19]Scaling of variables may lead to convergence issues in numerical methods estimation. A general rule of thumb to control for convergence issues arising from variables scale is the following: whenever the ratio between the biggest and the smallest standard deviation is more than 10, re-scaling should be considered to reduce the ratio below 10 (Scott-Jong, 1997 p. 60). Following this approach, the aging index and distance form closest refugees' center have been rescaled -multiplying both by a factor 10. Scaling has been applied also to population: expressed in millions of inhabitants, again to reach convergence.

Hub in the LLMA	611	0.3698854	0.483169	0	1
Central Area	611	0.3927987	0.4887728	0	1
Political preferences in LLMAs (2014)					
Right parties' votes	611	27.23696	8.619548	5.359276	55.28928
Left parties' votes	611	44.17219	9.47298	9.82249	73.67038
5 Stars votes	611	20.27569	5.679561	1.990223	38.42243
Center parties' votes	611	6.647648	5.036799	.2094972	48.85226
Cultural milieu of LLMAs					
Part of a pre-Unitarian state in 1748	611	10.37971	5.99452	1	18
Part of a pre-Unitarian state in 1815	611	7.216039	4.618119	1	12

Figure A2.1: histogram of the density of repetitions of hate manifestations



A.3 Hurdle model 1st stage.

Estimation results for model specifications with a progressively augmented set of controls. Starting from the baseline specification in which the share of foreign resident population is considered as the unique predictor for hate, confounding variables are subsequently added, to control for: socio-demographic characteristics of the LLMAs, economic outlook, trust, geographic characteristics, political preferences, social networks, civic capital. Estimation results presented in Table A3.1 below.

Table A3.1: 1ststage hurdle model logit estimates with regional FE for several model specifications

HATEYN	(1) Odds Ratio	(2) Odds Ratio	(3) Odds Ratio	(4) Odds Ratio	(5) Odds Ratio	(6) Odds Ratio	(7) Odds Ratio	(8) Odds Ratio
foreign resident pop	1.058	0.992	1.005	1.017	1.021	1.018	1.026	1.026
	(0.977 - 1.145)	(0.913 - 1.078)	(0.928 - 1.088)	(0.940 - 1.099)	(0.939 - 1.109)	(0.915 - 1.133)	(0.918 - 1.146)	(0.918 - 1.148)
splitted pop		1.781***	1.866***	1.804***	1.855***	1.657**	1.636**	1.637**
		(1.187 - 2.674)	(1.250 - 2.784)	(1.218 - 2.673)	(1.239 - 2.778)	(1.078 - 2.549)	(1.063 - 2.518)	(1.063 - 2.521)
aging index		0.960	0.960	0.960	0.952	0.965	0.966	0.966
		(0.898 - 1.026)	(0.896 - 1.029)	(0.897 - 1.027)	(0.892 - 1.016)	(0.906 - 1.029)	(0.904 - 1.031)	(0.905 - 1.031)
pop density (ln)		0.785	0.791	0.781	0.775	0.693**	0.725*	0.729*
		(0.558 - 1.106)	(0.551 - 1.136)	(0.533 - 1.144)	(0.549 - 1.093)	(0.488 - 0.983)	(0.504 - 1.045)	(0.511 - 1.039)
crime		0.988	0.985	0.989	0.986	0.986	0.988	0.988
		(0.950 - 1.027)	(0.947 - 1.026)	(0.949 - 1.031)	(0.946 - 1.029)	(0.944 - 1.030)	(0.944 - 1.034)	(0.944 - 1.034)
population (ln)		3.504***	3.357***	3.521***	3.096***	3.437***	3.488***	3.467***
		(2.695 - 4.555)	(2.525 - 4.462)	(2.633 - 4.709)	(2.223 - 4.311)	(2.442 - 4.838)	(2.497 - 4.872)	(2.457 - 4.891)
educated people			0.554	0.578	0.578	0.636	0.633	0.629
			(0.258 - 1.188)	(0.276 - 1.209)	(0.276 - 1.213)	(0.303 - 1.338)	(0.298 - 1.344)	(0.292 - 1.355)
educated pop (sq)			1.028*	1.026	1.026	1.025	1.024	1.025
			(0.995 - 1.061)	(0.995 - 1.058)	(0.995 - 1.057)	(0.994 - 1.057)	(0.993 - 1.057)	(0.993 - 1.058)
unemployment			1.039	1.032	1.033	1.001	1.007	1.007
			(0.964 - 1.120)	(0.963 - 1.106)	(0.962 - 1.108)	(0.921 - 1.088)	(0.921 - 1.100)	(0.921 - 1.100)
gini index			0.989	0.975	0.976	0.928	0.927	0.928
			(0.886 - 1.105)	(0.875 - 1.087)	(0.876 - 1.087)	(0.823 - 1.048)	(0.821 - 1.048)	(0.821 - 1.048)
trust				0.969*	0.969*	0.963**	0.962**	0.962**
				(0.938 - 1.001)	(0.937 - 1.002)	(0.933 - 0.993)	(0.933 - 0.992)	(0.932 - 0.993)
distance SPRAR (ln)					0.755*	0.777*	0.780	0.779
					(0.568 - 1.003)	(0.577 - 1.047)	(0.573 - 1.061)	(0.573 - 1.061)
hub in the LLMA					1.266	1.157	1.110	1.106
					(0.701 - 2.286)	(0.674 - 1.987)	(0.650 - 1.897)	(0.649 - 1.885)
central area					0.871	1.020	1.051	1.053
					(0.520 - 1.458)	(0.601 - 1.730)	(0.620 - 1.781)	(0.624 - 1.777)
left parties votes						1.000	1.000	1.000
						(0.955 - 1.046)	(0.956 - 1.047)	(0.955 - 1.047)
center parties votes						0.989	0.992	0.992
						(0.943 - 1.037)	(0.946 - 1.040)	(0.947 - 1.040)
right parties votes						1.078***	1.078***	1.077***
						(1.031 - 1.127)	(1.031 - 1.128)	(1.030 - 1.127)
5 star votes						1.090***	1.094***	1.093***
						(1.022 - 1.162)	(1.025 - 1.167)	(1.024 - 1.167)
social networks							1.082	1.082
							(0.921 - 1.271)	(0.921 - 1.271)
civic capital								1.054
								(0.663 - 1.677)
Observations	611	611	611	611	611	611	611	611
Regional FE	YES	YES	YES	YES	YES	YES	YES	YES

Confidence interval in parentheses (errors clustered at regional level)
**** p<0.01, ** p<0.05, * p<0.1*

Robustness of core covariates. Wald test, likelihood ratio test AIC and BIC diagnostics on the whole set of potential covariates allow to identify potential core-covariates. Then, 2048 regressions are estimated where the dependent binary variable is regressed on core variables - which are included in all regressions - and all possible combinations of other non-core variables. Below the summary of significance levels for every variable are outlined, by showing the minimum level reached by the p-value for testing variables and the maximum level reached by the p-value for core covariates. First, testing covariates are never significant. (Table A3.2). Then, considering core covariates, all core covariates confirm themselves as robust significant predictors in more than 75% of the 2048 model specifications. Population size, the share of separated/divorced couples, political preferences and trust in institutions are significant across the whole 2048 regressions.

Table A3.2: minimum p-value for every testing-covariate of the logit model with regional FE after 2048 regressions estimation

Testing Covariates (2048 regressions)	**p-value min (robust se)**
Foreign resident population	0.487
Aging index	0.213
Crime rate	0.520
Social networks	0.228
Center parties votes	0.638
Left parties votes	0.665
Unemployment	0.814
Gini index for income inequality	0.220
Educated people	0.131
Central Area	0.543
Hub in the LLMA	0.424
Civic capital	0.678

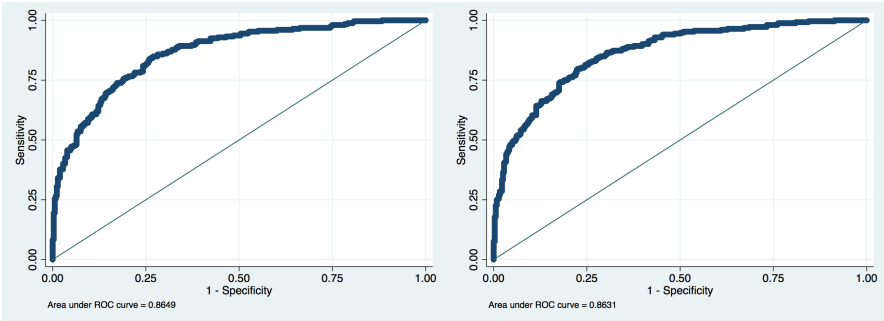
Table A3.3maximum p-value for every core-covariate of the logit model with regional FE after 2048 regressions estimation

Core Covariates (2048 regressions)	**p-value max (robust se)**
Educated people (sq)	0.145
Pop density	0.243
Population size	0.000
Split population	0.071
Right parties votes	0.002
5 stars parties votes	0.011
Trust	0.037
Distance from closest SPRAR	0.139

Goodness-of-fit is assessed through the Hosmer-Lemeshow test. Specification errors are gauged using STATA command `linktest`. The fitted model specification is robust

to both[20]. With regards to specificity and sensitivity, the area below the ROC curve for the fitted model specification equals 0,865. Figures A3.1 and A3.2 show the ROC curve for full model specification (column 8 in Table A3.1) and fitted model specification (column 1 in Table 2)

Figure A3.1: ROC curve full model *Figure A3.2: ROC curve fitted model*



The fitted model displays a VIF below 2,5 general tolerance above 0,4 and the condition number is way below 15, thus multicollinearity is not a source for concern (Table A3.4)

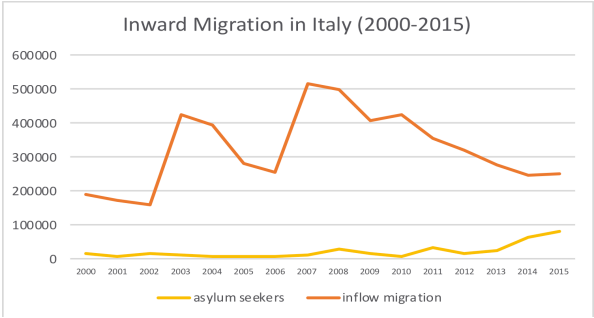
Table A3.4: Multicollinearity diagnostics for the fitted logit model

Variable	VIF	SQRT VIF	Tolerance	R-Squared		Eigenval	Cond Index
Split population	1.56	1.25	0.6415	0.3585	1	3.0008	1.0000
Pop density	2.59	1.61	0.3862	0.6138	2	1.5567	1.3884
Population size	3.00	1.73	0.3338	0.6662	3	1.2062	1.5773
Educated people	1.97	1.41	0.5065	0.4935	4	0.9282	1.7981
Educated pop (sq)	1.28	1.13	0.7828	0.2172	5	0.7366	2.0184
Trust	1.58	1.26	0.6327	0.3673	6	0.5979	2.2402
Distance SPRAR (ln)	1.42	1.19	0.7054	0.2946	7	0.3918	2.7675
Rightist parties votes	1.22	1.10	0.8223	0.1777	8	0.3627	2.8763
5 stars votes	1.18	1.08	0.8503	0.1497	9	0.2191	3.7009
Mean VIF	**1.75**					Condition Number	3.7009
						Det(correlation matrix)	0.0717

To perform robustness check for the association between the refugees hosting centre and hate occurrence, the geography of refugees hosting centres is changed from 2010 to 2013, since the latter has been characterized by an upward twist in refugees' inflow (see Figure A3.3) and refugees' hosting capacity.

[20] Hosmer_Lemeshow test: estat gof, group (10) gives a p-value = 0.6674; following Allison P. (2013), we also check for estat gof, group (9) and estat gof, group (11) and again, we have results supporting good fitness; linktest gives a p value for hatsq = 0.723

Figure A3.3: Asylum seekers and inflow migration in Italy 2000-2015 (OECD data)



A.4. Hurdle model 2nd stage

Table A4.1: 2nd stage hurdle model TNB estimates with regional FE for several model specifications

hate frequency	(1) -IRR-	(2) -IRR-	(3) -IRR-	(4) -IRR-	(5) -IRR-	(6) -IRR-	(7) -IRR-	(8) -IRR-
foreign resident pop	1.185	1.004	1.006	1.008	1.019	0.917	1.009	1.009
	(0.919 - 1.529)	(0.909 - 1.109)	(0.901 - 1.124)	(0.906 - 1.122)	(0.917 - 1.132)	(0.822 - 1.022)	(0.915 - 1.112)	(0.919 - 1.109)
aging index		0.981	0.969	0.967	0.961	0.936	0.954	0.954
		(0.878 - 1.096)	(0.864 - 1.086)	(0.862 - 1.084)	(0.852 - 1.083)	(0.771 - 1.135)	(0.845 - 1.077)	(0.846 - 1.076)
crime		0.994	0.996	0.996	0.998	1.018	1.001	1.001
		(0.971 - 1.017)	(0.974 - 1.018)	(0.974 - 1.019)	(0.978 - 1.019)	(0.976 - 1.061)	(0.977 - 1.025)	(0.976 - 1.027)
split pop		1.118	1.057	1.050	1.081	1.159	1.172	1.171
		(0.925 - 1.352)	(0.814 - 1.372)	(0.806 - 1.368)	(0.816 - 1.431)	(0.759 - 1.770)	(0.870 - 1.579)	(0.869 - 1.577)
pop density (ln)		0.819	0.789	0.790	0.733	0.533**	0.731*	0.734*
		(0.576 - 1.166)	(0.554 - 1.126)	(0.554 - 1.126)	(0.490 - 1.098)	(0.314 - 0.905)	(0.531 - 1.007)	(0.533 - 1.011)
population (ln)		3.080***	2.812***	2.824***	2.629***	5.071***	3.160***	3.133***
		(2.433 - 3.900)	(2.254 - 3.508)	(2.256 - 3.534)	(2.031 - 3.404)	(2.808 - 9.159)	(2.463 - 4.055)	(2.346 - 4.182)
educated people			1.067**	1.068*	1.038	1.076	1.002	1.002
			(1.001 - 1.138)	(1.000 - 1.140)	(0.964 - 1.118)	(0.946 - 1.224)	(0.932 - 1.076)	(0.930 - 1.080)
unemployment			1.038	1.035	1.032	1.091	1.012	1.011
			(0.906 - 1.190)	(0.906 - 1.183)	(0.897 - 1.188)	(0.889 - 1.340)	(0.893 - 1.146)	(0.891 - 1.147)
gini index			0.977	0.975	0.991	0.991	0.975	0.974
			(0.917 - 1.040)	(0.920 - 1.035)	(0.917 - 1.071)	(0.846 - 1.160)	(0.920 - 1.033)	(0.915 - 1.036)
trust				0.991	0.992	0.980	0.998	0.998
				(0.959 - 1.025)	(0.963 - 1.021)	(0.947 - 1.014)	(0.979 - 1.017)	(0.978 - 1.018)
hub in the LLMA					1.304	1.420	1.196	1.199
					(0.879 - 1.933)	(0.836 - 2.413)	(0.791 - 1.806)	(0.795 - 1.809)
central area					1.056	1.316	1.141	1.136
					(0.687 - 1.623)	(0.624 - 2.776)	(0.754 - 1.727)	(0.728 - 1.772)
distance SPRAR (ln)					0.861*	0.817	0.876*	0.878*
					(0.740 - 1.002)	(0.621 - 1.075)	(0.765 - 1.002)	(0.769 - 1.002)
right parties votes						1.074**	1.028	1.028
						(1.013 - 1.139)	(0.977 - 1.082)	(0.977 - 1.082)
5 star votes						0.991	1.012	1.013
						(0.890 - 1.103)	(0.933 - 1.098)	(0.938 - 1.094)
left parties votes						1.016	0.998	0.997
						(0.961 - 1.075)	(0.949 - 1.048)	(0.948 - 1.050)
center parties votes						1.159**	1.116**	1.116**
						(1.022 - 1.313)	(1.016 - 1.225)	(1.017 - 1.225)
social networks							1.106	1.107
							(0.954 - 1.283)	(0.955 - 1.283)
civic capital								1.038
								(0.658 - 1.637)
Observations	253	253	253	253	253	253	253	253
Regional FE	YES	YES	YES	YES	YES	YES	YES	YES

Confidence interval in parentheses (errors clustered at regional level)
**** p<0.01, ** p<0.05, * p<0.1*

The fitted model displays a VIF below 2,5, tolerance generally above 0,4 and the condition number is way below 15, therefore multicollinearity does not appear to be a concern (see table A4.2 below).

Table A4.2: Multicollinearity diagnostics for the fitted zero-truncated negative binomial model

Variable	VIF	SQRT VIF	Tolerance	R-Squared		Eigenval	Cond. Index
Pop density	2.05	1.43	0.4887	0.5113	1	2.1138	1.000
Pop size	2.33	1.53	0.4290	0.5710	2	1.0228	1.4376
SPRAR distance	1.30	1.14	0.7711	0.2289	3	0.9443	1.4962
Right parties' votes	1.10	1.05	0.9077	0.0923	4	0.6513	1.8015
Center parties' votes	1.00	1.00	0.9961	0.0039	5	0.2678	2.8096
Mean VIF	**1.56**					Condition number	2.8096
						Det. Correlation matrix	0.3561

A.5 Model comparison statistics: log likelihood, AIC and BIC

By comparing log-likelihood, AIC and BIC of the chosen model specification (Hurdle TNB) with alternative model specifications, the hurdle model with TNB as second stage outperforms the other specifications.

Table A5.1: Assessment of hurdle model with TNB against alternative model specifications

Model	Log likelihood	AIC	BIC
Hurdle (TNB)	**-630.1602**	**1290.320**	**1356.547**
Hurdle (Poisson)	-667.3519	1358.704	1411.685
ZINB	-646.7366	1331.473	1415.36
ZIP	-682.5391	1403.078	1486.965
Nbreg	-658.1091	1330.218	1361.124
Poisson	-718.9134	1443.827	1457.072

CHAPTER 2

Where Do Angry Birds Tweet?

Abstract. This paper analyses the influence of real-world socioeconomic and institutional features on cyberhate. Exploiting a novel dataset on Italian geo-tagged tweets, empirical estimation finds strong evidence of association between self-interest and cyberhate: *(i)* the more unequal the place, the more online intolerant behaviours increases; *(ii)* self-interest family values are positively associated with online hate. Finally, online hate is associated to situational features that differ from the ones related to offline hate. Results are robust to different model specifications.

Keywords: hate, online behaviour, inequality, self-interest, informal institutions, Heckman sample selection

2.1 Introduction

Online hate speech is defined as: "words or symbols diffused through the Internet, that are derogatory and/or intimidating on the basis of race, religion, sexual orientation, and so on" (McGonagle, 2013). Cyberhate is surging in size and diffusion and widely recognized as one of the most relevant challenges posed by social media platforms on the Internet (Brocato, 2016; Council of Europe, 2004; ENAR, 2016; European Commission, 2018.; Mitts, 2018; OSCE-ODIHR, 2010). Its prominence is supported also by figures: in 2018, Facebook has removed around 7,9 million pieces of content related to hate speech worldwide21; YouTube has cancelled more than 15.000 channels from July to September 2018[22]; similarly, Twitter has received more than 250.000 reports of hate-related contents between January and June just referring to a single country[23]. The relevance of cyberhate is confirmed also by the ongoing trend towards law enforcement to counter it (Assimakopoulos, Baider, & Millar, 2017). Due to these salient and worrying trends, online hate occupies a prominent position in numerous research fields, from computer science and criminology to economics and psychology (Gagliardone, Gal, Alves, & Martinez, 2015; Müller & Schwarz, 2018; Musto, Semeraro, de Gemmis, & Lops, 2016; Silva, Mondal, Correa, Benevenuto, & Weber, 2016). The main focus of these studies is on the impact of online hate speeches and on the strategies to counter it.

[21] https://transparency.facebook.com/community-standards-enforcement#hate-speech
[22] https://transparencyreport.google.com/youtube-policy/removals?hl=en
[23]http://blogs.lse.ac.uk/mediapolicyproject/2018/08/16/removals-of-online-hate-speech-in-numbers/

Despite the increasing efforts in detecting and removing cyberhate, little is known about its drivers, although alleviating the causes of cyberhate is recognized as one of the less divisive policy approaches since it avoids interfering with the fundamental right to freedom of speech13. In this work, the primary objective is contributing to the investigation of the determinants of online hate speech through the empirical estimation of the effects exerted by the spatial economic outlook on the proliferation of cyberhate.

Online hate does not happen in a situational vacuum: technological systems and contents are socially produced and culturally influenced (Castells, 2001). Online hatemongers are grounded in an offline environment whose influence extends also towards online behaviours. At the same time, online hate structurally differs from offline hate, as confirmed by recent evidence scrutinizing the relationship between the two (ElSherief, Kulkarni, Nguyen, Wang, & Belding, 2018; Hine et al., 2016; Mitts, 2019; Müller & Schwarz, 2018). Hence, it may be that online hate is associated to idiosyncratic situational risk factors. Up to now and to the best of my knowledge, there is still a lack of systematic empirical evidence specifically targeting the links between real-world features and online hate speech, while increasingly called for by hate scholars (Gagliardone et al., 2015; Hall, 2013; Silva et al., 2016).

The focus of this work is on the effects of the spatial economic outlook on cyberhate. Online hate serves the hatemongers' needs of protecting the extant status quo against some kind of perceived threats posed by out-group members, where the status quo relates to the places in which hatemongers live and work. The spatial economic outlook is a relevant source for threats perception since the out-group members may be considered as competitors for the distribution of scarce resources such as jobs and welfare (Guiso et al., 2017). Notably, the fear of competitors increases in places affected by economic inequality, because inequality determines settled unequal distributions of achievements and endowments. Since there is evidence showing that economic inequality affects real-world resentful behaviours (Wilkinson and Pickett, 2017), it is interesting to assess whether its role reaches also the digital sphere.

In this paper, the role of offline contexts with respect to cyberhate is scrutinised targeting the Italian Twitter dimension, being Twitter designed as a public space where both re-broadcasting to large audience and echo-chambers creation are particularly easy to achieve. Findings highlight a strong association between economic interests and online

hate. Results show a significant role exerted by economic inequality and self-interest: people living in unequal places positively relate to online hate, referring also to existing evidence on anti-immigrant voting (Dustmann, Vasiljeva, & Piil Damm, 2016) and on pro/anti-sociality behaviours (Côté, House, & Willer, 2015); analogously, family types characterized by the transmission of self-centeredness and individualistic values exert a strong and robust positive boost to online hate. Finally, also economic insecurity displays a significant association. Findings hold to several robustness checks as well as to estimation of competing modelling strategies.

By identifying risk and protective effects of contextual features on online hate speech, the paper contributes also to highlighting continuity and discontinuity between the online and offline hate events. Lastly, results contribute to the literature on the effects of long-term determinants and social norms on current geography of violent behaviours (Alesina et al., 2016; Tur-Prats, 2018).

The next section reviews prior literature allowing to formulate the research questions. Then, data alongside empirical strategy are described. Afterwards findings are discussed, and conclusions provided.

2.2 Online hate speech: Background literature and research questions

Online hate speech is intended to harm, harass, intimidate and humiliate targeted groups, promoting violence and insensitivity (Perry & Olsson, 2009). It represents a rising form of oppressive deviant behaviour, increasingly generated and shared through social media platforms exploiting fast and inexpensive extensive reaching (Silva et al., 2016). In a Special Eurobarometer survey, 75% of respondents has witnessed online hate speech on social platforms (Eurobarometer, 2016); in US, the share amounts to 66% (Duggan, 2017). Data also show large shares of cyberhate produced by people unrelated to any organized hate group (Hall, 2013).

Online hate speech is a deviant behaviour characterized by: *(i)* being oppressive against minorities, *(ii)* working as a "*message*" act against a whole minority group (Cohen-Almagor, 2014), *(iii)* serving to reinforce the in-group/out-group dichotomy (Hall, 2013). Being both dangerous and harmful and by exploiting Internet fast and global reach, it contributes to foster an atmosphere in which bias-motivated violence is encouraged, subtly or explicitly (Gagliardone et al., 2015; Müller & Schwarz, 2017).

Undoubtedly similar to real-world hate in many ways - oppressive aim, prejudice-based, serving in-group identity against out-group challenges-, online hate is a distinct form of oppressive deviant behaviour. Its impact on victims can be extremely durable, although being purely verbal: hyperlinking, online searchability, content shared by users make contents remain traceable and retrievable (Sunstein, 2017). Moreover, it gets public and available to a potentially global audience without any mediation (McGonagle, 2013). Cyberhate is characterized by the perpetrators' perceived sense of anonymity or deindividuation while they are online (Perry & Olsson, 2009): the apparent feeling of not being accountable for online speeches and the lack of social norms mediation determine lessening the sense for accountability and moderation (Banks, 2010; Citron & Norton, 2011). This "*online disinhibition effect*" (Suler, 2004) lowers the effect of real-world social barriers in countering aggressive and radical behaviours (ElSherief et al., 2018; Sunstein, 2017). Aside anonymity, online hate is subject to less restrictive legal constraints (Perry & Olsson, 2009), being not listed as crime in many countries, again reducing perception of social constraints.

Cyberhate is generated in an environment where it is easier to foster in-group/out-group dichotomy as well as social identity: growing evidence confirms that high shares of social media platforms users connect and bond mainly with users sharing the same views (Himelboim, Mccreery, & Smith, 2013), reinforcing social group identity (McNamee, Peterson, & Peña, 2010). Virtual in-groups are created along with "*echo chambers*": closed systems where stereotypes and prejudices are amplified and reinforced (Sunstein, 2017). Moreover, since many users exploiting Internet as cheap and fast source of information are not equipped with the necessary critical skills to evaluate the legitimacy of the information presented to them and social mediation is reduced, uncritical acceptance may work towards further reinforcement of prejudices (Hall, 2013; Perry, 2001). Overall, promoting intergroup online contacts to offset prejudices can be particularly difficult and poorly effective (McGonagle, 2013).

Having outlined the main features characterizing online hate has a distinct oppressive deviant behaviour, real-world elements that can potentially related to cyberhate are introduced. To this respect, the analysis refers to literature addressing oppressive deviant behaviours in general, being cyberhate such a novel phenomenon that literature

identifying its risk factors is at seminal and exploratory stage (Gerstenfeld, 2017; McNamee et al., 2010; Perry & Olsson, 2009).

The focus of investigation is on the effect of the economic geographies of places on online hate. To this regard, threat perception can work also along the economic dimension, again being real and/or perceived and boosted by economic downturns (Guiso et al., 2017). Economic inequality may influence violent oppressive behaviours, by increasing the level of social anxiety against potential threats to established achievements and endowments (Elgar, Craig, Boyce, Morgan, & Vella-Zarb, 2009; Wilkinson & Pickett, 2017). Existing empirical evidence on the correlation between income inequality and offline social tension is mixed and addressing different contexts (Côté et al., 2015; Jefferson & Pryor, 1999; Layte & Whelan, 2014; Medina, Nicolosi, Brewer, & Linke, 2018); moreover, few contributions have considered a fine-grained territorial scale. Moreover, extant literature shows that people with a high educational attainment may pursue anti-social behaviours due to a feeling of entitlement and self-orientation (Cote et al., 2015; Piffs and Robinson, 2017), which may be further fuelled when there is social intolerance due to the fact that high-social status people will internalize that they will have to bear the cost of protecting the victims of social intolerance (Dustmann et al., 2019). The existing evidence on the association between educated people and anti-social behaviours is not conclusive (Piffs and Robinson, 2017), as some places are characterized by a positive association whereas other display a negative one (Kraus & Callaghan, 2016). Hence, it appears interesting to assess the significance and the direction of the association in the Italian context. It also relevant to assess from an empirical point of view the moderating role of income inequality and educated people. This interaction is supported by existing contributions showing that inequality plays a role in increasing the anti-social behaviours enacted by people with a high social status (Piffs and Robinson, 2017 Cote et al., 2015). Being the focus on inequality, it is also appealing to analyse the role of family given its focal imprinting on the spatial geographies of social norms, including the transmission of equality/non-equality value (Bertocchi & Bozzano, 2016; Duranton & Rodriguez-Pose, 2010) as well as to several oppressive violent behaviours (Brown, 2014; Tur-Prats, 2018). Todds (1990) has identified an organising principle for the classification of family types based on the relationship between siblings in the family. On the one hand, siblings are considered equal, whereas, at the other extreme, parents may favour one

particular child (often the eldest) at the expense of the others. On this basis, families are labelled "*equal*" or "*unequal*". Notably, these family values have already been acknowledged as capable of influencing relevant socioeconomic outcomes (Duranton & Rodrıguez-Pose, 2010). The equality indicator identified by Todds is given by what happens to family property after the death of the parents. Equality is said to be strongest where family property is divided most evenly between siblings, or (more usually) between brothers. Areas in which equal familial systems are operating are identified, therefore, by inheritance laws and practices. The classification of family types following from Todds (1990) identifies two egalitarian family types -communitarian and egalitarian nuclear- and two non-egalitarian family types -stem and incomplete stem-. Non-egalitarian family types are identified by the self-interested dimension enforced through individualistic standards (Duranton & Rodrıguez-Pose, 2010). The same individualistic standards channelled by the "*non-egalitarian*" family types are also recognised within social psychology as booster for social anxiety and violence, since they increase the relevance of preserving the existing social status from potential threats (Wilkinson & Pickett, 2017). Thus, it appears engaging to bridge these two strands of literature by assessing whether self-interest transmitted through non-egalitarian family types influence online hate.

Alongside the main dimension under investigation, it is important to account for competing spatial features which may as well influence the production of online hate. To this respect, online hate can serve users' needs of backing existing social norms, social identities and/or hierarchies that they perceive to be under threat (Suler, 2004; Sunstein, 2017). Threats may be sociocultural (Hainmueller & Hopkins, 2014), given by real or perceived agents of change with regards to societal identities and status quo (Perry, 2001). Examples are migrants, refugees, women, LBGT people, alternative family forms. Whether and how geography may influence online hate is another appealing issue to address; although ICT technology is capable of enabling communication overcoming physical distance, it has not yet ruled out the role of proximity, serving as complement rather than a substitute in many circumstances (Glaeser, 2014). Hence, the classification of the Italian territory in central or marginalised areas as well as the presence of catchment areas in the considered spatial unit of observation are included among potential cofounders. Social relationships are influenced by social structures. Noteworthy, the

deindividuation process inherent to online behaviours weakens their role (Suler, 2004; Sunstein, 2017), therefore the reach towards the digital sphere of different types of social structure is accounted for among confounding features, considering trust and collaboration which have an established association with pro/anti-sociality (Guiso et al., 2017; Wilkinson & Pickett, 2017). Finally, also several real-world deviant behaviours may relate to social strain and tension. Crime is one of the main sources for social tension (Pinotti, 2015) and distress (Dustmann & Fasani, 2016); then real-world hate events are releases of tensions arising from threat perception (Perry, 2001); gambling relates to compensation needs arising from feelings of material/social deprivation (Nyman, Welte, & Dowd, 2008; Welte, Wieczorek, Barnes, & Tidwell, 2006); and finally, anti-vax movements hinge on mistrust in institutions and individualistic urges over collaborative behaviours (Kennedy, 2019; WHO, 2018; Wolfe, 2002).

Therefore, different strands of literature pinpoint several real-world dimensions potentially influencing online hate, allowing to formulate the first research question for this study:

1) *To what extent real-world economic geographies of insecurity and inequality may influence online hatred behaviours in Italy, after controlling for relevant potential confounding factors?*

The previous chapter has scrutinised which local features relate to real-world hate in Italy. By identifying risk and protective situational features associated to cyberhate in the same context, it is also possible to address the second research question:

2) *Which patterns of continuity and discontinuity can be detected between situational features influencing online and offline hate in Italy?*

2.3 Data

Measuring online hate in Italy. The measure for online hate is given by the corpus of Twitter geo-tagged data created by Musto et al. (2016) and used to create the Italian Hate Map[25], in turn inspired by the Humboldt University Hate Map targeting US[26]. The database contains more than 75.000 georeferenced tweets generated in Italy throughout 2017 and identified through a pipeline of algorithms for data extraction, semantic

[25] http://www.voxdiritti.it/ecco-le-mappe-di-vox-contro-lintolleranza/
[26] https://users.humboldt.edu/mstephens/hate/hate_map.html

processing, sentiment analysis and content classification (Musto, Semeraro, Lops, & Gemmis, 2015). Although capturing only a part of cyberhate, Twitter is a valuable source for measuring digital hate speech, being widely used to propagate cyberhate[27] as well as characterized for having public content (ElSherief et al., 2018; Himelboim et al., 2013). Tweets are aggregated at Local Labour Market Area (LLMA) to account for the impossibility to detect whether the tweets are posted during working time, commuting time or leisure time. In fact, the 611 Italian LLMAs divide Italy in sub-regional geographical areas where the bulk of the labour force lives and works, hence appearing suitable to alleviate for the unfeasibility to assign Tweets to either workplace, third places or home.

Measures for the local economic outlook. The main dimension of interest is given by the economic outlook, with specific focus on economic insecurity and inequalities, due to the increasing interest on their effect on resentment, social anxiety and intolerance. Economic insecurity is assessed through the unemployment rate and the rate of company failures and overdue payments. Economic inequality is captured by the Gini index on income. Also the size of educated population in each LLMA is considered. The latter measure also relates to the level of affluence in the LLMA, being the higher wages payed by employers to workers with tertiary education (OECD, 2012).

Measures for self-interest values. Todds (1990) geography of family types for Italy is used to discriminate between egalitarian and non-egalitarian families. Both egalitarian and non-egalitarian family types are present in Italy. The most frequent family type is the egalitarian nuclear, being the dominant family type in the 60% of LLMAs; then the communitarian family type which is the prominent family type in the 25% of LLMAs. The incomplete stem family type covers the 12.6% of LLMAs and the stem family type the 1.96%. Vaccine coverage proxies the attitude to put collective needs before individual deeds.

Potential cofounders. Alongside, other dimensions are considered, since they represent potential cofounding features at the same time contributing to deliver a more comprehensive picture of the risk and protective factors associated to online hate. Sociocultural potential risk factors are proxied using the share of resident migrants and

[27] Amnesty International "Toxic Twitter"; Demos "Misogyny on Twitter"

share of disrupted families. The geographic outlook is conveyed through the classification of each LLMA as either central or peripherical following the Italian Agency for Territorial Cohesion (ATC) classification[28]. The presence of major catchment areas within each LLMA is considered, again following classification from ATC. Since one of the aims of the paper is analysing the patterns of continuity/discontinuity between contextual features associated with real-world hate and cyberhate, also the geography of refugees' hosting centres (SPRAR) is introduced, since they display robust associations with real-world hate as shown in the previous chapter. To capture different kinds of social structures the following proxies are considered: voting turnout at the 2014 European Parliament elections to measure trust in formal institutions and the number of non-for-profit local units to account for collaborative network. Finally, to convey a measure for the real-world deviant behaviours which will be assessed in their relationship with cyberhate, the following measures are introduced: the crime rate and real-world hate manifestations, already used to measure real-world hate in chapter 1. The empirical investigation also controls for population size, population density and political preferences.

A novel dataset is compiled by merging these different data and consolidating them at LLMA geography. Figure 1 depicts the geography of online hate, offline hate, family types and vaccine hesitancy; the latter conveyed by decrease in vax coverages. More details are outlined in Table 1[29].

[28] www.agenziacoesione.gov.it

[29] Data sources and descriptive statistics are in Table A1.1 and A1.2 in Annex A1.

Figure 1: LLMA geography of hate tweets, offline hate, egalitarian/non-egalitarian family types, vax hesitancy

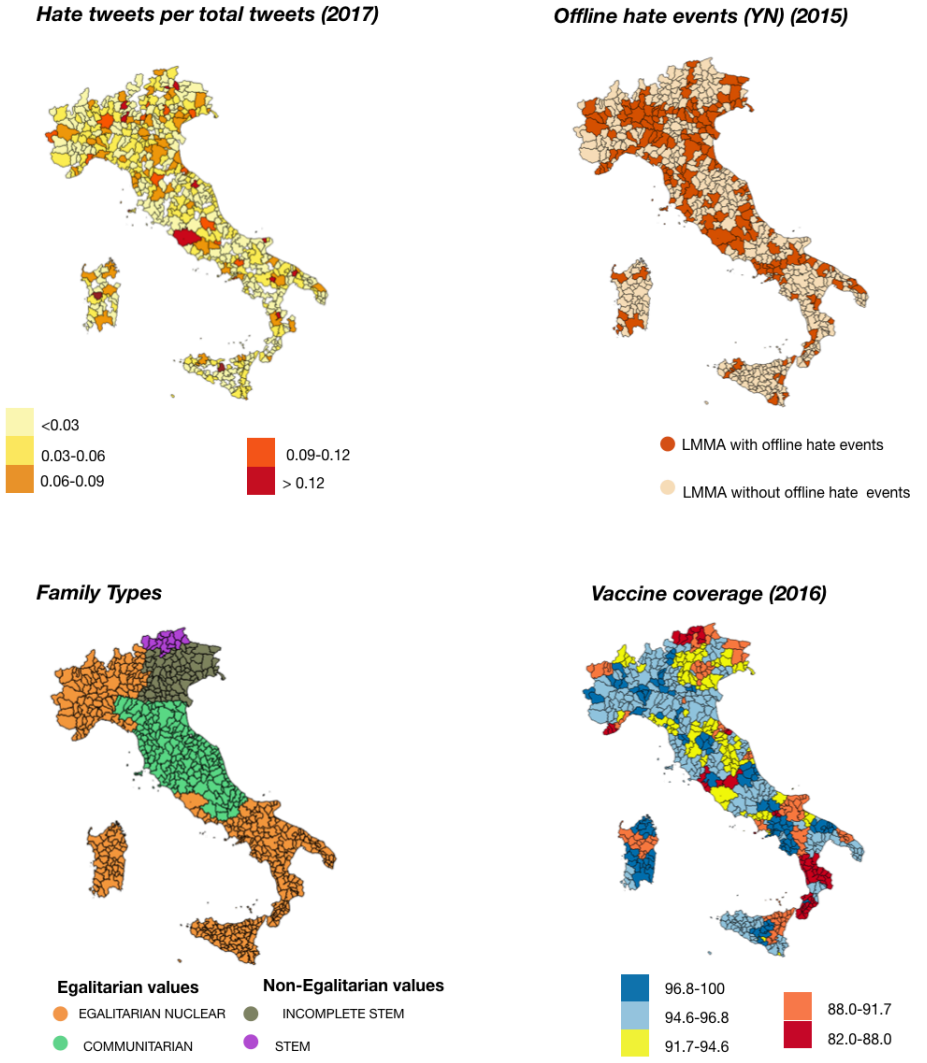


Table 1: Data description and sources

Variable	Description
	Geography of hate tweets
onlinehate	number of hate tweets on total tweets in 2017
z	Dummy variable taking value 1 if there is at least one hate tweet in 2017; 0 if otherwise
	Sociocultural features of LLMAs
Foreign resident population	Share of non-Italian resident population as % of total resident population
Split population	Share of divorced and separated couples as % of total resident population
Aging index	Ratio of resident population over 65 on resident population over 15
Population density	resident population per hectare (ha)
Population	Resident population in hundred thousand

Economic characteristics of LLMAs	
Human capital	Share of population with college degree as % of resident population > 20
Unemployment	Share of unemployed as % of labour supply
Gini index	Gini index for income inequality in % computed on tax declarations of resident households
Company failures	Number of failed companies every 10000 companies in 2007
Overdue payments	Number of overdue payments every 1000 inhabitants in 2009
Perception of job security	Employees feeling likely to lose their current job within the next 6 months and highly unlikely to find a similar job as share of total employees
Geographic characteristics of LLMAs	
Distance from closest SPRAR in 2010	LLMA centroid distance from the closest refugees' hosting structure measured in decimal degrees
Hub in the LLMA	Dummy variable taking value 1 if LLMA contains at least a municipality or a cluster of municipalities acting as catchment areas; 0 otherwise
Central Area	Dummy variable taking value 1 if the majority of the LLMA is classified as central rather than "inner"; 0 otherwise
Social Structures of LLMAs	
Social networks	Number of local units of non-for-profit organizations every 1000 inhab
Trust	Share of voting turnout in % at the 2014 European Parliament Elections
Family types	Categorical variable mapping the geography of family types
Behavioural features of LLMAs	
Crime	Number of recorded crimes every 1000 inhabitants
Offline hate (2015)	Dummy variable taking value 1 if there had been at least one offline hate event in 2015; 1 if otherwise
Offline hate (2013)	Dummy variable taking value 1 if there had been at least one offline hate event in 2013; 1 if otherwise
Tweets	Number of total geotagged tweets (irrespective of the topic)
Gambling availability	Number of gambling Number of gambling facilities per inhabitant
Measles vaccine coverage	Share of measles vaccination coverage among children aged 24 months
Diphtheria vaccine coverage	Share of diphtheria vaccination coverage among children aged 24 months
Political preferences in LLMAs	
Right parties' votes	Share of vote in % given for rightist parties (Northern League, FdI, Forza Italia) at the 2014 European Parliament Elections
Left parties' votes	Share of vote in % given for leftist parties (Democratic Party, Green Party, L'Altra Europa IdV) at the 2014 European Parliament Elections
5 Stars votes	Share of vote in % given for 5 Stars Movement at the 2014 European Parliament Elections
Center parties' votes	Share of vote in % given for center parties (Scelta Europea, Io Cambio, NCD) at the 2014 European Parliament Elections
Exclusion restrictions (Sample Selection Model)	
No internet age index	Share of population over 70-year-old on population below 15-year-old
No internet age index robustness check	Share of population over 74-year-old on population below 15-year-old

2.4 Empirical Strategy and Estimation

The dependent variable is continuous with a non-negligible fraction of zeros[30], suggesting three competing estimation strategies: single-stage OLS, the two-part model (2PM) and the sample selection model[31] (SSM) (i.a. Cameron & Trivedi, 2010; Santos-Silva, Tenreyro, & Windmeijer, 2015). The single-stage OLS assumes that the share of zero values in the outcome variable does not determine any bias in estimation, whereas the 2PM and the SSM specifications imply that these zero must be accounted for to get unbiased results. Moreover, although 2PM and SSM are both assuming that the zeros in the dependent variable must be modelled, they nonetheless differ on the nature of these assumptions. The 2PM assumes that the zero in the dependent variable are true zeros, hence the outcome variable is fully observed and there is no selection bias to address. Conversely, the SSM assumes that the zero values for the outcome variable indicate observations for which the potential outcome is missing. Whenever this is the case, 2PM estimation suffers from selection bias and the SSM is a more suitable option (Dow & Norton, 2003). Following the ongoing vigorous debate opposing 2PM and SSM advocate, a balanced approach is pursued. Data do not entail evident features indicating one model over the other, thus 2PM, SSM and the single-stage OLS are estimated to identify the most appropriate through postestimation diagnostics as detailed in the literature (Cameron & Trivedi, 2010; Dow & Norton, 2003; Santos-Silva et al., 2015; J. Wooldridge, 2002). Since 2PM displays the strongest predictive power[32], its results are presented as main findings, whereas SSM outcomes and single-stage OLS outcomes are outlined in robustness checks. In all model specifications regional fixed effects are included and errors are clustered at regional level.

2PM assumes that we are observing actual outcome with only true zero observations, allowing independent mechanisms for the participation decision ($onlinehate = 0$ versus $onlinehate > 0$) and the amount decision (the magnitude of online hate, when it is positive (Wooldridge, 2002). The two parts are estimated separately, considering a binary outcome model (probit) for the first part $Pr(onlinehate > 0)$ and linear regression to model the second part $E(\ln onlinehate \mid onlinehate > 0)$, where log-normality accounts for correcting

[30] Zeros for the dependent variable amount to 14,24%
[31] The latter is also referred as Type-II-Model or Heckman Selection Model.
[32] See Annex A4.

the right-skewness of the dependent variable, once the condition $onlinehate > 0$ is applied (Cameron & Trivedi, 2010). Formally, let z be a binary indicator of positive online hate events such that $z = 1$ *if* $onlinehate > 0$ and $z = 0$ if $onlinehate = 0$. For $onlinehate > 0$, $f(ln\, onlinehate / z = 1)$ is the conditional density of $ln\, onlinehate$. The 2PM can be summarised as

$$f(\ln onlinehate\,|X) = f(x) = \begin{cases} \Pr(z = 0|X), & if\ onlinehate = 0 \\ \Pr(z = 1|X)\, f(\ln onlinehate\,|z = 1, X), & if\ onlinehate > 0 \end{cases} \quad (1)$$

The same regressors are assessed in both parts of the 2PM, since there are no obvious ex-ante exclusion restrictions. The dependent variable is $lnonlinehate$, measured by logs of the share of hate tweets on total tweets in each LLMA in 2017. Given that hate-related tweets are collected by focusing on the identification and the extraction of Twitter contents characterized by specific hate-related keywords, hate-related tweets extraction does not include the general flow of tweets. Therefore, to measure the total amount of tweets produced in a given area a different Twitter corpus is used, collected by Cheng, Caverlee, Lee, & Sui (2011) and up to now the largest and finest-grained geo-tagged Twitter databases. In the robustness check, results are assessed using resident population as normalizing factor for the dependent variable. Results from this alternative normalization confirm the main findings of the paper. A more detailed discussion is presented in the robustness checks.

<u>2PM Stage 1: Probit.</u> The fitted probit specification, whose results are summarized in Table 2, column 1, is identified through robustness tests[33] applied to different model specifications, with progressively increasing set of covariates to account for different class of contextual features[34]. The egalitarian nuclear family is the family type of reference, and therefore is not included in the regression analysis. All family variables coefficients can be interpreted as relative to the family type of reference. *NE (E)* indicates non-egalitarian (egalitarian) family types.

[33] Wald test, contrast test, log likelihood, AIC and BIC are used to discriminate between potential core covariates and testing covariates. Robustness of core covariates is assessed by running 1024 regressions, keeping the potential core covariates fixed while turning all the testing covariates in every possible combination.

[34] For estimation results of these preliminary model specifications, see Annex A2, Table A2.1.

Table 2: 2PM 1st stage: fitted probit specification and robustness checks

Variables	(1) coef *fitted model*	(2) coef *job insecurity '15*	(3) coef *job insecurity '13*	(4) coef *vax coverages '14*	(5) coef *no big cities*	(6) coef *se cluster Province*
Self-interest dimension						
educated people	0.148***	0.133***	0.129**	0.159***	0.148***	0.129**
	(0.0559)	(0.0512)	(0.0521)	(0.0567)	(0.0559)	(0.0588)
incomplete stem family (NE)	**4.951*** (0.741)**	**4.980*** (0.777)**	**4.963*** (0.777)**	**5.548*** (0.763)**	**4.949*** (0.741)**	**4.963*** (0.851)**
stem family (NE)	**4.677*** (1.120)**	**4.482*** (1.170)**	**4.466*** (1.151)**	**5.127*** (1.200)**	**4.674*** (1.120)**	**4.466*** (1.219)**
communitarian family (E)	**-0.827** (0.412)**	**-0.867** (0.433)**	**-0.850* (0.437)**	**-0.716* (0.390)**	**-0.827** (0.412)**	**-0.850** (0.372)**
vax coverage '16	-0.0493***	-0.0447**	-0.0432**		-0.0493***	-0.0432
	(0.0179)	(0.0190)	(0.0198)		(0.0179)	(0.0298)
vax coverage '14				-0.0817***		
				(0.0305)		
Economic insecurity dimension						
unemployment	-0.0421**			-0.0440**	-0.0421**	
	(0.0199)			(0.0215)	(0.0199)	
job insecurity '15		**0.780*** (0.279)**				
job insecurity '13			**0.471*** (0.152)**			**0.471** (0.200)**
offline hate '15	0.518	0.559*	0.549*	0.508	0.518	0.549*
	(0.325)	(0.328)	(0.329)	(0.327)	(0.325)	(0.316)
Controls	*YES*	*YES*	*YES*	*YES*	*YES*	*YES*
Observations	611	611	611	611	607	611
Regional FE	*YES*	*YES*	*YES*	*YES*	*YES*	*YES*
Control: LLMA population size and political preferences						

Robust standard errors in parentheses: columns 1-5 cluster errors at regional level; column 6 cluster errors at province level
**** p<0.01, ** p<0.05, * p<0.1*

From Table 2 - column 1, risk factors for predicted probability of online hate are mainly related to the self-interest dimension. A significant risk factor is given by non-egalitarian family types, supporting the persistent channeling of self-interested values along families; also vax hesitancy behaves as meaningful risk factor, suggesting the prevalence of individualistic needs against collective deeds; finally, the share of educated people positively relates to online hate, a finding which deserves further investigation. To the one hand, it can relate to self-interest theory being that more educated people are also the more affluent (OECD, 2012). To this regard, recent evidence shows how affluent people develop resentment and intolerance against minorities by internalising the welfare burden they have to bear for their protection (Côté et al., 2015; Dustmann et al., 2016); nonetheless, it may also relate to the fact that Twitter users are generally endowed with a high level of human capital. With respect to the labour market, the results in column 1 of Table 2 show an increase in employment is positively related to online hate. This result, apparently counterintuitive, could be driven by some features characterizing the Italian

job market, namely the ongoing trend of high level of job insecurity (OECD, 2018). Thus, even if new jobs are created, they do not fully counter economic insecurity perception. To further explore this insight, a measure for job insecurity perception by labour force is considered, using the ISTAT Well Being and Sustainability Indicator on job insecurity for 2013 and 2015 (ISTAT, 2014, 2016). Results (columns 2 and 3, Table 2) support the insight that the perception of job insecurity by workers is positively related to online hate. Results stay robust controlling for cofounding factors like population size and political preferences, which behave consistently with the literature (Barone, D'Ignazio, de Blasio, & Naticchioni, 2016; Glaeser, 2005).

Postestimation and robustness checks. The fitted model is assessed considering vax hesitancy geography in another point in time. Due to data availability at fine-grained spatial scale, diphtheria vax coverage in 2014[35] is introduced, which confirms vax hesitancy as robust predictor of online hate occurrence (Table 2: column 4). The fitted model specification is also estimated after removing the biggest LLMAs -Rome, Milan and Naples- to get consistency of findings (Table 2: column 5). Finally, errors at clustered at province level, to see that results hold aside vax hesitancy (column 6 in Table 2). Postestimation diagnostics with respect to the fitted model specification do not outline severe concerns: multicollinearity does not appear as an issue; the model performs well in terms of goodness-of-fit and specification; it also displays outstanding discrimination in terms of sensitivity and specificity[36].

2PM Stage 2: OLS. Potential core covariates are identified as in stage 1; then, by running 131072 regressions with all potential core covariates fixed while turning testing covariates in every possible combination, the fitted model specification is defined, whose results detailed in Table 3 column 1.

In the second stage of the 2PM, the self-interest dimension still displays relevant associations with online hate.

[35] Measles and diphtheria are both established diseased in vaccine programs, historically characterized by similar coverage shares (Italian Ministry of Health, 2018)

[36] See Annex A2 for detailed results

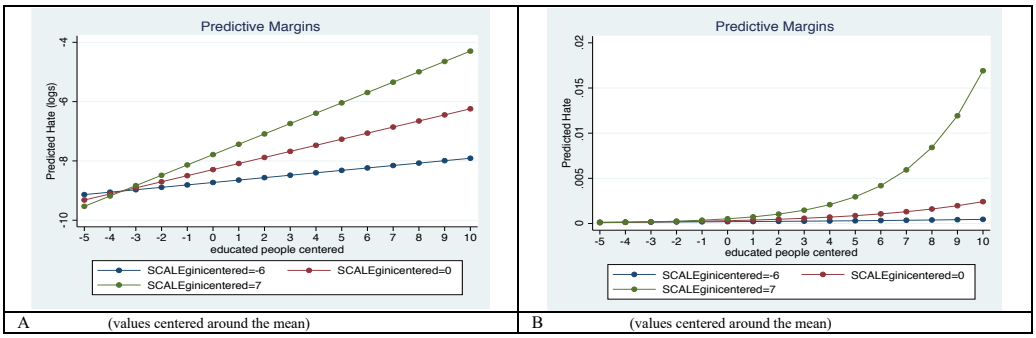
Table 3: 2PM 2nd stage: fitted OLS specification and robustness checks

Variables	(1) 2PM fitted OLS coef *cluster region*	(2) 2PM fitted OLS coef *cluster province*	(3) 2PM fitted OLS coef *no outlier*	(4) 2PM fitted OLS coef *hate '13*
Self-interest dimension				
educated people	0.205***	0.213***	0.210***	0.213***
	(0.0307)	(0.0363)	(0.0362)	(0.0369)
Gini index	**0.0723****	**0.0814****	**0.0896****	**0.0838****
	(0.0309)	**(0.0382)**	**0.210*****	**(0.0384)**
inequality*educated people	**0.0206*****	**0.0180****	**0.0185****	**0.0173***
	(0.00791)	**(0.00888)**	**(0.00885)**	**(0.00882)**
incomplete stem family (NE)	0.541	1.595**	1.606**	1.840**
	(0.720)	(0.652)	(0.652)	(0.845)
stem family (NE)	0.954	2.723***	2.733***	2.905***
	(0.633)	(0.680)	(0.679)	(0.876)
communitarian family (E)	0.618	-	0.530	0.504
	(0.853)		(0.381)	(0.358)
offline hate '15	0.715***	0.548***	0.548***	
	(0.126)	(0.142)	(0.142)	
offline hate '13				0.480***
				(0.139)
hub in the LLMA	0.788***	0.909***	0.911***	0.995***
	(0.157)	(0.167)	(0.167)	(0.166)
central area	0.0417	0.0205	0.0420	-0.0447
	(0.143)	(0.160)	(0.157)	(0.162)
Observations	524	524	523	524
R-squared	0.504	590	0.601	0.588
Regional FE	YES	NO	YES	YES
Province FE	NO	YES	NO	NO

*Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1*

The level of income inequality is positively associated with the share of hate tweets on total tweets. Moreover, alongside this direct effect, there is also evidence of a significant association for the interaction between the size of human capital and income inequality[37], further detailed in Figure 2.

Figure 2: Moderating role of income inequality on self-interest



A (values centered around the mean) B (values centered around the mean)

[37] The same interaction term has been estimated in the first part of the 2PM to find non-significance. Values are centred around their mean. Other interactions with the remaining core predictors were considered, namely with the presence of a hub in the LLMA, the LLMA being a central area and real-world hate events. None of these interactions prove itself to be significant.

As the Gini index increases, the marginal semi-elasticity of the size of educated people on predicted online hate events grows[38]. Hence, *ceteris paribus*, places that are more economically unequal increase influence on online hate. Previous occurrence of offline hate events displays a positive association, too. Finally, there is evidence that the presence of catchment areas in the LLMA represents a risk factor for cyberhate, providing at the same time further support to the fact that notwithstanding the digital nature of cyberhate, real-world geography still matters.

Postestimation and robustness checks[39]. Analogously to robustness checks performed for the first part of the 2PM, the fitted model is estimated considering province fixed effects, to get that main results are confirmed (Column 2 -Table 3). There are hints for a mild outlier, therefore its relevance is analysed: estimation summarized in column 3 of Table 3 shows that the mild outlier does not affect the main results and that, by accounting for it, overall estimation improves. Finally, we change the geography for deviant behaviour, considering occurrence of real-world hate in 2013, to see that results hold (Column 2 -Table 4). With respect to postestimation diagnostics[40], both multicollinearity and omitted variable tests do not rise any concern.

2.5 2PM discussions of findings and robustness check

In both stages of the 2PM the economic sphere exerts the most relevant role in association to cyberhate. Influence on hate tweets occurs through the geography of economic insecurity perception, income inequality and local persistency of self-interest values as transmitted by non-egalitarian family types. Moreover, the self-interest dimension exerts a significant influence in both stages of the 2PM, influencing both the occurrence and the size of hate tweets. Self-interest plays a role in actual local outlook, as shown by the effects of income inequality, the size of affluent people and the size of vaccine hesitant population. Furthermore, self-interest plays a role also along the social value outlook of places, as shown by the effects of non-egalitarian family types. This second element is entailed with more persistent characteristics, being that value

[38] The dependent variable is in logarithmic term; thus, regression coefficients provide semi-elasticity measures.
[39] Results have been double-checked the results from both stage estimations with results from the user-written command -tpm- (Bellotti, Deb, Manning, & Norton, 2015), to get the same findings.
[40] See Annex A2.

transmission is more enduring than current economic conditions and that family types display really long persistence (Duranton & Rodrıguez-Pose, 2010).

Real world hate manifestations are positively related to cyberhate, being that previous hate events contribute to nurture the perception of social tension and anxiety. Moreover, they may also work as signal for increased protection needs from disempowered group, therefore increasing the perception of further welfare burden by affluent people.

The results also outline robust evidence on the non-significant role exerted either by crime or non-national residents, supporting extant hate-related literature argument that hatred does not rely on actual crime figures, rather it hinges on stereotypes (Brown, 2014; Glaeser, 2005).

The considered outcome variable does not allow to exclude a priori that observed zero online hate events might be observations for which the potential outcome is latent (Dow and Norton, 2003). Following the literature, it is estimated a model capable of accounting for this potential selection bias, relaxing the assumption that decision to tweet about hate and amount of hate tweets are independent. Then, it is assessed which model has the strongest predictive power. The first competing model that is estimated is a Sample Selection Model (SSM) with two interdependent processes -a selection equation (probit) for the probability of having a positive outcome and a conditional equation (OLS) for the sub-set with $onlinehate > 0$[41]. Potential collinearity and identification issues that may arise in SSM estimates are alleviated by identifying an exclusion restriction in the selection equation and by constraining the set of covariates in the conditional equation to be a strict subset of the covariates in the selection equation[42] (Wooldridge, 2002). The paper introduces an exclusion restriction which is given by the share of

[41] Formally,

$$Pr(onlinehate > 0|X_1) = \Phi(X_1\beta_1, \varepsilon_1)$$

$$E(\ln onlinehate|onlinehate > 0, X_2) = X_2\beta_2 + E(\varepsilon_2|onlinehate > 0, X_1)$$
$$= X_2\beta_2 + \rho\sigma_2\lambda(X_1\beta_1)$$

Where $X_1$are the covariates for the condition equation and X_2 are the covariates for the selection equation. $\lambda(X_1\beta_1)$ is the inverse Mills ratio $\lambda(X_1\beta_1) = \varphi(X_1\beta_1)/\Phi(X_1\beta_1)$, a non-selection hazard linking the two equations.

[42] X_2 is a sub-set of X_1 and X_1 contains the exclusion restriction.

population above 70-years old over digital natives. National statistics on internet usage by population (ISTAT, 2018) outline that nearly the 91% of the population over 75 years-old had not accessed internet in 2017[43]. Furthermore, Twitter statistics about the age profile of user in Italy highlights that the 96% of Twitter users is above 70-years old (Global Web Index, 2015). This variable satisfies the condition for being an exclusion restriction (Cameron & Trivedi, 2010)[44].

Through two-step procedure it is possible to get consistent and robust estimates[45] (Greene, 2003; Leung & Yu, 1996; Wooldridge, 2002) with error clustered at regional level. Fitted SSM specification is robust to several checks: clustering error at province level, testing different specifications for the exclusion restriction, removing influential points and outliers. Postestimation diagnostics do not highlight any concern for multicollinearity, goodness-of-fit or omitted variables. Outliers and influential points do not alter estimation results. Overall, we conclude that estimation results from SSM are quite robust. In order to discriminate among the competing frameworks, a discrimination test based on predictive power with respect to our data is performed following established literature (Dow & Norton, 2003; Madden, 2008; Santos-Silva et al., 2015). Results support 2PM as the preferred specification. Anyhow there is continuity in many of the findings between 2PM and SSM.

Alongside the nature of zeros in the outcome variable, also their magnitude represents a feature worth considering in choosing the proper modelling strategy. To this regard, results from modelling strategies accounting for zero values are compared with results from a model specification where the zeros are assumed not to be a cause for concern. More into details, a single-stage OLS on the entire population is estimated and its predictive power is gauged against the 2PM and the SSM. Results, detailed in the Appendix A3, show that the single-stage OLS displays the lowest predictive power among competing model specifications. In any case, findings from single-stage OLS confirm main findings.

[43] Same percentage applies to 2016.
[44] Other potential exclusion restrictions have been tested, such as broadband connectivity and 3G connectivity, but they do not satisfy the necessary conditions
[45] For a detailed description of the SSM estimation and postestimation, see Annex A3.

Results presented thus far refer to the share of hate tweet on total tweets in each LLM as dependent variable. Another factor for normalization is considered, given by resident population in each LLMA. Hence, in this case, the dependent variable is given by hate tweets per inhabitant. Again, the three different competing model specifications are estimated: 2PM, SSM and single-stage OLS. Results confirms the findings of the paper[46].

2.6 Conclusions

This paper presents the first work which empirically investigates which situational features relates to digital hate in Italy, targeting fine space granularity. By exploiting a novel dataset merging geotagged hate tweet with administrative and historical data, empirical estimation has detected a major role played by the economic dimension, even when controlling for many different potential confounding factors. Economic insecurity, inequality and non-egalitarian social norms appear as significant and robust risk factors for online hate. With respect to economic insecurity, the results relate to recent findings identifying a relationship between economic insecurity and populism (Guiso et al., 2017), providing further support to the association between economic anxiety and behaviours against targeted groups. Regarding inequality, the findings presented in this paper relate to growing contributions in different fields: recent evidence has highlighted that people are less collaborative towards society in more unequal places (Côté et al., 2015); self-interest displays correlation also in anti-migrant voting behaviours due to welfare burden concerns (Dustmann et al., 2016). Further steps will address potential endogeneity of income inequality through an instrumental variable approach. Also, vaccine hesitancy, a deviant behaviour which is consistent with self-interest, emerges as risk factors for cyberhate.

Notably, neither migrants/refugees or crime appear significantly related to cyberhate, whereas previous real-world hate manifestation are significant risk factors. Further research is needed to address each channel of transmission in details as well as longitudinal Twitter data to scrutinize causal relationships.

The findings outline that cyberhate is boosted and countered by situational features that are distinct from the one relating to offline hate, providing empirical support to

[46] See Annex A3.

growing literature analysing cyberhate as a distinctive oppressive behaviour. In addition, the results contribute to the ongoing debate opposing the perception of sociocultural threats to the perception of economic threats as the causes of people's resentment (Card, Dustmann, & Preston, 2012; Hainmueller & Hopkins, 2014; McCann, 2019; Rodríguez-Pose, 2018), showing that for cyberhate the economic dimension is the more relevant one. Results from chapter 1 show, instead, that the sociocultural dimension is the one associated with real-world hate events. Hence, being that the two resentful behaviours are associated to different sphere, at least in Italy, they should be addressed by different types of policy.

At the same time, the meaningful association between real-world characteristics and cyberhate suggests broadening current policy to counter cyberhate. Currently, a relevant bulk of policy initiatives countering online hate is devoted to improving commitment and accountability of ISP providers, social media platforms and digital users, supported also by the creation of technical tool automatedly detecting online hate speech (ElSherief et al., 2018). However, this type of policy suffers from acknowledged shortcomings due to prosecutions vagaries that dampen the deterrence effect as well as the cheap and fast relocation of hate contents in more favourable jurisdictions (McGonagle, 2013). Results discussed in this paper add another important point to these acknowledged limitations, stressing that situational characteristics of places represent another relevant dimension to address to reduce risk factors associated to cyberhate.

References

Alesina, A., Brioschi, B., & Ferrara, E. La. (2016). Violence Against Women: A Cross-cultural Analysis for Africa. *National Bureau of Economic Research Working Paper Series*, *No. 21901*. https://doi.org/10.3386/w21901

Assimakopoulos, S., Baider, F., & Millar, S. (2017). *Hate Speech in the European Union: A Discourse- Analytic Perspective*. SpringerBriefs.

Banks, J. (2010). Regulating hate speech online. *International Review of Law, Computers & Technology*, *24*(3), 233–239. https://doi.org/10.1080/13600869.2010.522323

Barone, G., D'Ignazio, A., de Blasio, G., & Naticchioni, P. (2016). Mr. Rossi, Mr. Hu and politics. The role of immigration in shaping natives' voting behavior. *Journal of Public Economics*, *136*(8228), 1–13. https://doi.org/10.1016/j.jpubeco.2016.03.002

Bellotti, F., Deb, P., Manning, W., & Norton, E. (2015). twopm: Two-part Model. *The Stata Journal*, *15*(1), 3–20.

Bertocchi, G., & Bozzano, M. (2016). Origins and Implications of Family Structure across Italian Provinces in Historical Perspective, (10327), 1–23.

Brocato, R. (2016). *Hate Speech Online: Assessing Europe's Capacity to Tackle an Emerging Threat.* (Freedom from Fear No. Issue 12).

Brown, R. (2014). *Prejudice. Its social psychology. Journal of Multilingual and Multicultural Development* (Vol. 35). https://doi.org/10.1080/01434632.2013.803723

Cameron, A. C., & Trivedi, P. K. (2010). *Microeconometrics Using Stata: Revised Edition*. StataPress.

Castells, M. (2001). *The Internet Galaxy: Reflections on the Internet, Business, and Society*. New York, NY, USA: Oxford University Press, Inc.

Cheng, Z., Caverlee, J., Lee, K., & Sui, D. Z. (2011). Exploring millions of footprints in location sharing services. In *Fifth International AAAI Conference on Weblogs and Social Media*.

Citron, D. K., & Norton, H. (2011). Intermediaries and hate speech: Fostering digital citizenship for our information age. *Boston University Law Review*, *91*, 1435–2131. https://doi.org/10.2139/ssrn.1764004

Cohen-Almagor, R. (2014). Countering Hate on the Internet. *Annual Review of Law and Ethics*, *Vol. 22*, 431–443.

Côté, S., House, J., & Willer, R. (2015). High economic inequality leads higher-income individuals to be less generous. *Proceedings of the National Academy of Sciences*, *112*(52), 15838–15843. https://doi.org/10.1073/pnas.1511536112

Council of Europe. Convention on Cybercrime: Protocol on Xenophobia and Racism. CETS no.185 (2004).

Dow, W. H., & Norton, E. C. (2003). Choosing Between and Interpreting the Heckit and Two-Part Models for Corner Solutions. *Health Services & Outcomes Research Methodology*, *4*(5). https://doi.org/10.1023/A:1025827426320

Duggan, M. (2017). Online harassment 2017. *Pew Research*, 1–85. https://doi.org/202.419.4372

Duranton, G., & Rodrıguez-Pose, A. (2010). Family Types and the Persistence of. *Regional Studies*, (November).

Dustmann, C, Vasiljeva, K., & Piil Damm, A. (2016). Refugee Migration and Electoral Outcomes. *CReAM Discussion Paper*, (19/16). https://doi.org/10.1016/S1043-2760(97)84344-5

Dustmann, C., & Fasani, F. (2016). The Effect of Local Area Crime on Mental Health. *Economic Journal*, *126*(593), 978–1017. https://doi.org/10.1111/ecoj.12205

Elgar, F. J., Craig, W., Boyce, W., Morgan, A., & Vella-Zarb, R. (2009). Income Inequality and School Bullying: Multilevel Study of Adolescents in 37 Countries. *Journal of Adolescent Health*, *45*(4), 351–359. https://doi.org/10.1016/j.jadohealth.2009.04.004

ElSherief, M., Kulkarni, V., Nguyen, D., Wang, W. Y., & Belding, E. (2018). Hate Lingo: A Target-based Linguistic Analysis of Hate Speech in Social Media, (Wolfson). Retrieved from http://arxiv.org/abs/1804.04257

ENAR. (2016). *Racism and discrimination in the context of migration in Europe. ENAR Shadow Report.*

Eurobarometer. (2016). *Media pluralism and democracy – Special Eurobarometer 452. Digital Single Market.* Retrieved from https://ec.europa.eu/digital-single-market/en/news/media-pluralism-and-democracy-special-eurobarometer-452

European Commission. (2018.). *Third Evaluation of the Code of Conduit on Countering Illegal Online Hate Speech.*

Gagliardone, I., Gal, D., Alves, T., & Martinez, G. (2015). *Countering Online Hate Speech. UNESCO Series on Internet Freedom*. https://doi.org/978-92-3-100105-5

Gerstenfeld, P. B. (2017). *Hate Crimes: Causes, Controls, and Controversies*. SAGE Publications. Retrieved from https://books.google.it/books?id=v35ZDwAAQBAJ

Glaeser, E. (2014). *The Triumh of the City*. Penguin

Glaeser, E. L. (2005). The Political Economy of Hatred*. *Quarterly Journal of Economics*, *120*(1), 45–86. https://doi.org/10.1162/qjec.2005.120.1.45

Global Web Index. (2015). *GWI Social Q1*.

Greene, W. (2003). *Econometric Analysis (5th Edition)*. Prentice Hall.

Guiso, L., Herrera, H., Morelli, M., & Sonno, T. (2017). Populism: Demand and Supply *, 1–63. Retrieved from http://www.heliosherrera.com/populism.pdf

Hainmueller, J., & Hopkins, D. J. (2014). Public Attitudes Toward Immigration. *Annual Review of Political Science*, *17*, 225–249. https://doi.org/10.1146/annurev-polisci-102512-194818

Hall, N. (2013). *Hate crimes. 2nd Edition*. Routledge.

Himelboim, I., Mccreery, S., & Smith, M. (2013). Birds of a Feather Tweet Together: Integrating Network and Content Analyses to Examine Cross-Ideology Exposure on Twitter. *Journal of Computer-Mediated Communication*, *18*(2), 40–60. https://doi.org/10.1111/jcc4.12001

Hine, G. E., Onaolapo, J., De Cristofaro, E., Kourtellis, N., Leontiadis, I., Samaras, R., … Blackburn, J. (2016). Kek, Cucks, and God Emperor Trump: A Measurement Study of 4chan's Politically Incorrect Forum and Its Effects on the Web, (Icwsm), 92–101. https://doi.org/10.1007/s00213-003-1442-x

Hosmer D., & Lemeshow S. (2000*). Applied Logistic (2nd edition)*. Wiley International

ISTAT. (2014.). *Rapporto BES*.

ISTAT. (2016). *Rapporto BES*.

ISTAT. (2018). Aspects of Daily Life: Internet - Age Detail. In Multipurpose Survey on Households: Aspects of Daily Life -General Part. Retrieved from http://dati.istat.it/Index.aspx?lang=en&SubSessionId=31db8ed9-fc80-4a53-914a-98bba42146bc

Jefferson P.N. and Pryor F. (1999). On the geography of hate. *Economic Letters*, *65*(5), 389–395. https://doi.org/10.3917/reco.566.1249

Kennedy, J. (2019). Populist politics and vaccine hesitancy in Western Europe: an analysis of national-level data. *European Journal of Public Health*, *29*(3), 512–516. https://doi.org/10.1093/eurpub/ckz004

Layte, R., & Whelan, C. T. (2014). Who feels inferior? A test of the status anxiety hypothesis of social inequalities in health. *European Sociological Review*, *30*(4), 525–535. https://doi.org/10.1093/esr/jcu057

Leung, S. F., & Yu, S. (1996). On the choice between sample selection and two-part models. *Journal of Econometrics*, *72*(1–2), 197–229. https://doi.org/10.1016/0304-4076(94)01720-4

Madden, D. (2008). Sample selection versus two-part models revisited: The case of female smoking and drinking. *Journal of Health Economics*, *27*(2), 300–307. https://doi.org/10.1016/j.jhealeco.2007.07.001

Manning, W. G., Duan, N., & Rogers, W. H. (1987). Monte Carlo evidence on the choice between sample selection and two-part models. *Journal of Econometrics*, *35*(1), 59–82. https://doi.org/10.1016/0304-4076(87)90081-9

McGonagle, T. (2013). *The Council of Europe against online hate speech: Conundrums and challenges. Council of Europe Conference Expert Paper.* https://doi.org/10.1007/978-3-319-31238-5_6

McNamee, L. G., Peterson, B. L., & Peña, J. (2010). A call to educate, participate, invoke and indict: Understanding the communication of online hate groups. *Communication Monographs*, *77*(2), 257–280. https://doi.org/10.1080/03637751003758227

Medina, R. M., Nicolosi, E., Brewer, S., & Linke, A. M. (2018). Geographies of organized hate in America: A regional analysis. *Annals of the American Association of Geographers*, *108*(4), 1006–1021. https://doi.org/10.1080/24694452.2017.1411247

Mitts, T. (2018). *From Isolation to Radicalization: Anti_Muslim Hostility and Support for ISIS in the West.*

Müller, K., & Schwarz, C. (2017). Fanning the Flames of Hate: Social Media and Hate Crime. *Ssrn*. https://doi.org/10.2139/ssrn.3082972

Müller, K., & Schwarz, C. (2018). Making America Hate Again? Twitter and Hate Crime Under Trump. *Ssrn*. https://doi.org/10.2139/ssrn.3149103

Musto, C., Semeraro, G., de Gemmis, M., & Lops, P. (2016). Modeling community behavior through semantic analysis of social data: The italian hate map experience. In *Proceedings of the 2016 Conference on User Modeling Adaptation and Personalization* (pp. 307–308). ACM.

Musto, C., Semeraro, G., Lops, P., & Gemmis, M. de. (2015). CrowdPulse: A framework for real-time semantic analysis of social streams. *Information Systems*, *54*, 127–146. https://doi.org/10.1016/j.is.2015.06.007

Nyman, J. A., Welte, J. W., & Dowd, B. E. (2008). Something for nothing: A model of gambling behavior. *Journal of Socio-Economics*, *37*(6), 2492–2504. https://doi.org/10.1016/j.socec.2008.02.011

OECD. (2012). *Education at a glance: Highlights.* OECD Publishing, Paris

OECD. (2018). *Employment Outlook 2018*. OECD Publishing, Paris

OSCE-ODIHR. (2010). *Incitement to Hatred vs. Freedom of Expression: Challenges of combating hate crimes motivated by hate on the Internet.* Retrieved from https://www.osce.org/odihr/68750

Perry, B. (2001). *In the Name of Hate: Understanding Hate Crimes*. https://doi.org/10.1017/CBO9781107415324.004

Perry, B., & Olsson, P. (2009). Cyberhate: the globalization of hate. *Information & Communications Technology Law*, *18*(2), 185–199. https://doi.org/10.1080/13600830902814984

Pinotti, P. (2015). Immigration Enforcement and Crime. *American Economic Review*, *105*(5), 205–209. https://doi.org/10.1257/aer.p20151040

Santos-Silva, J. M. C., Tenreyro, S., & Windmeijer, F. (2015). Testing Competing Models for Non-negative Data with Many Zeros. *Journal of Econometric Methods*, *4*(1). https://doi.org/10.1515/jem-2013-0005

Silva, L., Mondal, M., Correa, D., Benevenuto, F., & Weber, I. (2016). Analyzing the Targets of Hate in Online Social Media, (March). Retrieved from http://arxiv.org/abs/1603.07709

Suler, J. (2004). The Online Disinhibition Effect. *CyberPsychology & Behavior*. https://doi.org/10.1089/1094931041291295

Sunstein, C. (2017). *#Republic: Divided Democracy in the Age of Social Media*. Princeton University Press.

Todds, E. (1990). *L'invention de l'Europe*. Seuil.

Tur-Prats, A. (2018). Family Types and Intimate-Partner Violence: A Historical Perspective. *The Review of Economics and Statistics*, 1–45. https://doi.org/10.1162/rest_a_00784

Welte, J. W., Wieczorek, W. F., Barnes, G. M., & Tidwell, M.-C. O. (2006). Multiple Risk Factors for Frequent and Problem Gambling: Individual, Social, and Ecological. *Journal of Applied Social Psychology*, *36*(6), 1548–1568. https://doi.org/10.1111/j.0021-9029.2006.00071.x

WHO. (2018). *Assessment Report of the Global Vaccine Action Plan. Strategic Advisory Group of Experts on Immunization.*

Wilkinson, R. G., & Pickett, K. E. (2017). The enemy between us: The psychological and social costs of inequality. *European Journal of Social Psychology*, *47*(1), 11–24. https://doi.org/10.1002/ejsp.2275

Wolfe, R. M. (2002). Anti-vaccinationists past and present. *Bmj*, *325*(7361), 430–432. https://doi.org/10.1136/bmj.325.7361.430

Wooldridge, J. (2002). *Econometric Analysis of Cross-Section and Panel Data (2nd ed)*. The MIT Press.

Annex A1. Descriptive statistics

Table A1.1 Data sources

Variable	Source
Geography of hate tweets	
onlinehate	Musto *et al.,*(2016, 2015)
Sociocultural features	
Foreign resident population	ISTAT -15th Population Census (2011)
Split population	ISTAT -15th Population Census (2011)
Aging index	ISTAT -15th Population Census (2011)
Population density	ISTAT -15th Population Census (2011)
Population	ISTAT -15th Population Census (2011)
Economic characteristics	
Human capital	ISTAT -15th Population Census (2011)
Unemployment	ISTAT -15th Population Census (2011)
Gini index	Ministry of Economy and Finance (fiscal year 2011)
Company failures	ISTAT -Justice and Security- bankruptcies (2007)
Overdue payments	ISTAT -Justice and Security- protests (2009)
Perception of job security	ISTAT -Well Being Indicators- (2013, 2015)
Geographic characteristics	
Distance from closest SPRAR in 2010	Ministry of the Interior and National Association of Italian Municipalities (ANCI) SPRAR website (2007, 2010, 2013)
Hub in the LLMA	Agency for Territorial Cohesion, Classification of Italian Municipalities according to Inner Areas (2014)
Central Area	Agency for Territorial Cohesion, Classification of Italian Municipalities according to Inner Areas (2014)
Social Structures	
Social networks	ISTAT -9th Industry and Services Census (2011)
Trust	Ministry of the Interior Elections database (2014)
Family types	Duranton *et al.*, (2010); Todds (1990)
Behavioural features	
Crime	ISTAT -Growth Policy Indicators: Rule of Law and Order (2011)
Offline hate (2015)	Lunaria (2018)
Offline hate (2013)	Lunaria (2018)
Tweets	Cheng *et al* (2011)
Gambling availability	Customs Agency
Measles vaccine coverage (2016)	Wired (2018) – Ministry of Health (2018)
Diphtheria vaccine coverage (2014)	Wired (2018) – Ministry of Health (2018)
Political preferences	
Right parties' votes	Ministry of the Interior Elections database (2014)
Left parties' votes	Ministry of the Interior Elections database (2014)
5 Stars votes	Ministry of the Interior Elections database (2014)
Center parties' votes	Ministry of the Interior Elections database (2014)

Exclusion restrictions (Sample Selection Model)	
No internet age index	ISTAT -15th Population Census (2011)

Figure A1.1. Box plots of hate tweets per total tweets in LLMA and of logs of hate tweets per total tweets in LLMA to correct for right skewness

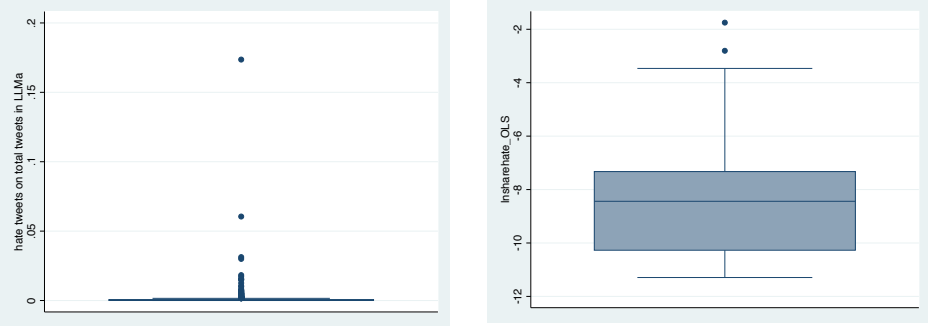


Figure A1.2: Box plots of hate tweets per LLMA inhabitants and of logs of hate tweets per inhabitants to correct for right skewness

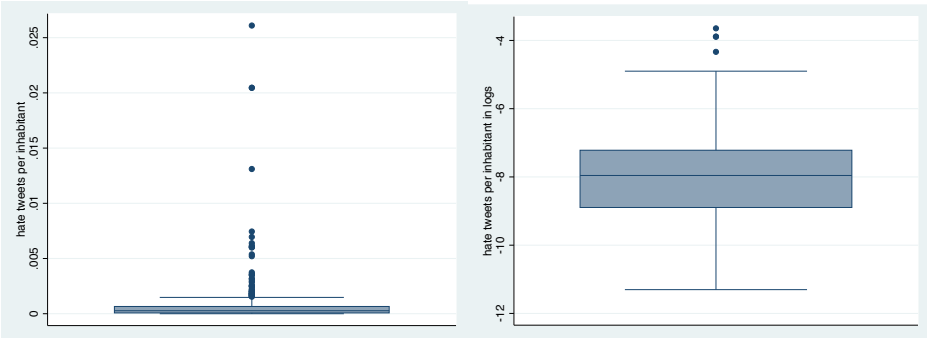


Table A1.2: Descriptive statistics

Variable[47]	Obs	Mean	Std. Dev.	Min	Max
Geography of hate tweets (2017-18)					
$onlinehate$	611	0.000663	0.00185	0	0.0261
z	611	0.858	0.350	0	1
Sociocultural features (2011)					
Foreign resident population	611	5.799721	3.777264	0.2714932	16.39703
Split population	611	3.770977	1.509572	0.7702183	8.319397
Aging index	611	16.69316	4.646057	6.93617	34.9384
Population density	611	2.056063	2.915807	0.1041765	31.07485
Population	611	0.09727	0.25829	0.314	3.6851
Economic characteristics (2011)					
Human capital	611	10.28096	2.737488	3.578903	21.07972
Unemployment	611	11.94874	6.136332	1.486878	27.40376
Gini index	611	41.135	2.866571	33.47302	54.87072
Company failures	611	12.43	5.680	1	30.10
Overdue payments	611	23.45	11.47	4.400	48
Perception of job security 2013	611	5.473344	0.7065942	3.571887	7.629317
Perception of job security 2015	611	3.159242	0.3876799	2.227504	4.645005
Geographic characteristics					
Distance from closest refugees' hosting facility in 2010	611	3.6711	2.8916	0. 0777543	19.23096
Hub in the LLMA	611	0.3698854	0.483169	0	1
Central Area	611	0.3927987	0.4887728	0	1
Social Structures					
Social networks	611	5.528334	2.32694	.9486006	14.95564
Trust	611	56.41075	12.89531	25.47039	85.69295
Family types	611	1.92144	1.275384	1	4
Deviant behaviours					
Crime	611	19.93512	7.503294	7.467119	55.05622
Offline hate (2015)	611	0.270	0.444	0	1
Gambling availability	611	2.242603	0.6932691	0. 340329	5.066981
Offline hate (2013)	611	0.260	0.439	0	1
Measles vaccine coverage (2016)	611	93.96	3.659	82	100
Diphtheria vaccine coverage (2014)	611	94.99364	2.617369	85.9	100

[47] Scaling of variables may lead to convergence issues in numerical methods estimation. A general rule of thumb to control for convergence issues arising from variables scale is the following: whenever the ratio between the biggest and the smallest standard deviation is more than 10, re-scaling should be considered to reduce the ratio below 10 (Scott-Jong, 1997 p. 60). Following this approach, population is expressed in hundred thousand, the aging index and distance form closest refugees' center have been rescaled -multiplying both by a factor 10.

Political preferences in LLMAs (2014)					
Right parties votes	611	27.23696	8.619548	5.359276	55.28928
Left parties votes	611	44.17219	9.47298	9.82249	73.67038
5 Stars votes	611	20.27569	5.679561	1.990223	38.42243
Center parties votes	611	6.647648	5.036799	.2094972	48.85226
Exclusion restrictions (SSM)					
No internet age index	611	3.716	1.134	1.372	8.266
No internet age index robustness check	611	2.528	0.828	0.865	5.961

Some covariates that have not proved to be significant are omitted for readability issues: robberies, homicides, thefts, usury and gambling machines per inhabitants.

Annex A2. 2PM model

Table A2.1: 2PM stage 1 -Estimation results by increasing the considered covariates

Variables	(1) 2PM probit coeff	(2) 2PM probit coeff	(3) 2PM probit coeff	(4) 2PM probit coeff	(5) 2PM probit coeff	(6) 2PM probit coeff
foreign resident pop	-0.0272	-0.0456	-0.0379	-0.0667*	-0.0657	-0.0596
	(0.0349)	(0.0387)	(0.0373)	(0.0400)	(0.0403)	(0.0414)
split pop	0.251*	0.196	0.196	0.219	0.199	0.146
	(0.150)	(0.145)	(0.144)	(0.147)	(0.142)	(0.146)
aging index	-0.0347*	-0.0473**	-0.0463**	-0.0487*	-0.0466*	-0.0415*
	(0.0193)	(0.0211)	(0.0224)	(0.0248)	(0.0250)	(0.0249)
population density	0.155	0.150	0.213	0.184	0.160	0.134
	(0.151)	(0.137)	(0.163)	(0.160)	(0.163)	(0.164)
population (ln)	0.897***	0.903***	0.896***	0.871***	0.878***	0.913***
	(0.113)	(0.138)	(0.151)	(0.160)	(0.157)	(0.153)
unemployment		-0.0457**	-0.0478**	-0.0631***	-0.0573**	-0.0547**
		(0.0228)	(0.0235)	(0.0242)	(0.0273)	(0.0257)
educated people		0.140*	0.141**	0.132*	0.138*	0.176**
		(0.0731)	(0.0692)	(0.0747)	(0.0732)	(0.0803)
gini index		-0.0514**	-0.0508**	-0.0361	-0.0320	-0.0440*
		(0.0250)	(0.0255)	(0.0243)	(0.0260)	(0.0226)
protests		0.0174	0.0152	0.0136	0.0152	0.00757
		(0.0247)	(0.0248)	(0.0255)	(0.0254)	(0.0266)
company failures		-0.0120	-0.0113	-0.0101	-0.00748	-0.00203
		(0.0246)	(0.0241)	(0.0264)	(0.0282)	(0.0299)
central area			-0.461	-0.510	-0.504	-0.513
			(0.301)	(0.312)	(0.333)	(0.330)
hub in the LLMA			0.290	0.350	0.311	0.241
			(0.394)	(0.408)	(0.420)	(0.412)
distance SPRAR (ln)			-0.0427	-0.0352	-0.0298	-0.0143
			(0.132)	(0.117)	(0.125)	(0.136)
collaborative outlook				-0.0680	-0.0647	-0.0573
				(0.0457)	(0.0500)	(0.0481)
trust				0.0119	0.0110	0.00930
				(0.0117)	(0.0117)	(0.0111)
incompl stem family (NE)				4.247***	4.353***	4.681***
				(0.388)	(0.439)	(0.571)
stem family (NE)				3.787***	3.564***	4.220***
				(0.353)	(0.428)	(0.882)
communitarian family (E)				-1.017*	-1.026*	-0.967*
				(0.564)	(0.567)	(0.525)
crime rate					0.00385	0.00602
					(0.0267)	(0.0290)
vax coverage					-0.0353	-0.0379*
					(0.0220)	(0.0214)
offline hate (Y/N)					0.567**	0.573*
					(0.283)	(0.302)
right parties votes						0.0202
						(0.0173)
left parties votes						-0.00601
						(0.0130)
5 star votes						0.0119
						(0.0197)
center parties votes						-0.0202
						(0.0184)
Observations	611	611	611	611	611	611
Regional FE	YES	YES	YES	YES	YES	YES

*Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1*

The first stage of the 2PM considers a binary outcome variable Pr($onlinehate$ >0). Therefore, it is independent from the normalizing factor used to weight hate tweets in the second part.

To assess robustness of the potential core covariates identified through postestimation diagnostics 4096 regressions are estimated, keeping the potential core covariates fixed while turning all the other covariates in every possible combination. Table A2.2 reports the minimum p-value for every testing covariate, showing that they never display robust significant correlation with the dependent variable.

Table A2.2: minimum p-value for testing covariates in 1024 regressions

Testing variables	p value min	
foreign resident pop	0.289	
split pop	0.251	
crime	0.611	
population density	0.214	
protests	0.550	
company failures	0.173	
hub in the LLMA	0.366	
distance SPRAR (ln)	0.741	
left parties votes	0.173	
social networks	0.409	
center voting	0.073	*p-value*≤ 0.100 *in 4.68% of regressions*
trust	0.2603	

Then, the same exercise is done considering potential core covariates only, results are reported in Table A2.3: population size, educated people, unemployment, political preferences and deviant behaviours are robust core covariate.

Table A2.3: maximum p-value for potential core covariates

Core variables	p value max	
aging index	0.242	*p-value*≤ 0.100 *in 25.78% of regressions*
population (ln)	0.010	
unemployment	0.174	*p-value*≤ 0.100 *in 80.37% of regressions*
educated people	0.024	
gini index	0.583	*p-value*≤ 0.100 *in 9.37% of regressions*
central area	0.252	*p-value*≤ 0.100 *in 0% of regressions*
right parties' votes	0.122	*p-value*≤ 0.100 *in 92.97% of regressions*
5-star votes	0.094	
vax coverage	0.131	*p-value*≤ 0.100 *in 91.41% of regressions*
offline hate	0.133	*p-value*≤ 0.100 *in 80.86% of regressions*

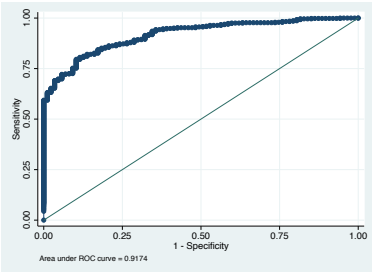
Family types are then added checking for their significance through `contrast` test and `contrast r.` in Stata to reach the fitted model specification. Table A2.4 shows that multicollinearity does not appear to be a cause for concern, given a weak Cramer association between categorical variables (Cramer's V = 0.1216); mean VIF way below 10; tolerance above 0.4 and a small value for the condition number.

Table A2.4: Multicollinearity check for stage 1 of 2PM

Variable	VIF	SQRT VIF	Tolerance	R-Squared		Eigenval	Cond Index
Population size	1.63	1.28	0.6121	0.3879	1	1.7995	1.0000
Educated people	1.67	1.29	0.5995	0.4005	2	1.3948	1.1358
Unemployment	1.32	1.15	0.7596	0.2404	3	1.1234	1.2657
Rightist parties votes	1.13	1.06	0.8856	0.1144	4	0.8428	1.4612
5 stars votes	1.39	1.18	0.7175	0.2825	5	0.4775	1.9415
Measles vax coverage '16	1.06	1.03	0.9397	0.0603	6	0.3621	2.2292
Mean VIF	**1.37**					Condition Number	2.2292
						Det(correlation matrix)	0.4108

The ROC curve shows that the fitted model specification is characterized by outstanding discrimination, given an area below the curve greater than 0.9 (Figure A2.1)

Figure A2.1: ROC curve for fitted probit specification



STATA command `-linktest-` assesses the fitted model specification: we get a p-value for _hatsq of 0.608 and a p-value for _hat of 0.000. Thus, the linktest is not significant, meaning that the model is well fitted. Finally, the Hosmer-Lemeshow test for goodness-of-fit test of the fitted model confirms that the model fits data well: `estat gof, group (10)` gives a p-value = 0.8709; following Allison P. (2013), it is also checked for `estat gof, group (9)` and `estat gof, group (11)` and again, results support good fitness.

Proceeding as in stage 1, first more than 131072 regressions are estimated, keeping all core covariates fixed while making testing covariates turn in all possible combinations. From this assessment, the confirmed robust core covariates are listed in Table A2.5. Table A2.6 summarizes the testing variables which are not robust enough to be considered in the fitted model specification.

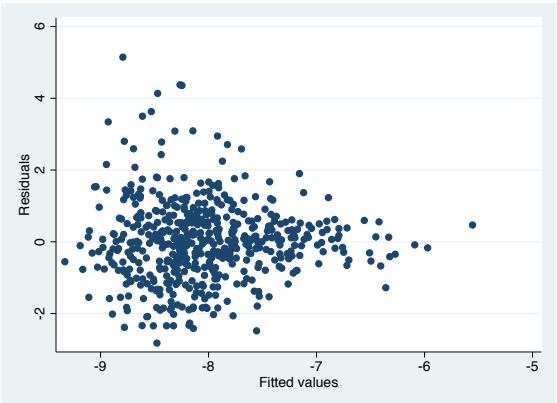
Table A2.5: Thresholds significance level of core covariates after robustness check

Core variables	p value max	
educated people	0.001	
hub in the LLMA	0.018	*p-value*≤ 0.100 *in 98.84% of regressions*
central area	0.031	*p-value*≤ 0.100 *in 71.67% of regressions*
offline hate '15	0.153	*p-value*≤ 0.100 *in 60.53% of regressions*

Table A2.6: Thresholds significance level of testing covariates after robustness check

Testing variables	p value min	
crime rate	0.070	*p-value*≤ 0.100 *in 14.11% of regressions*
aging index	0.035	*p-value*≤ 0.100 *in 14.11% of regressions*
foreign resident pop	0.315	
split pop	0.026	*p-value*≤ 0.100 *in 23.87% of regressions*
population density	0.040	*p-value*≤ 0.100 *in 20.53% of regressions*
population (ln)	0.055	*p-value*≤ 0.100 *in 1.67% of regressions*
unemployment	0.155	
Gini index	0.204	
protests	0.148	
company failures	0.202	
distance SPRAR (ln)	0.245	
right parties votes	0.670	
left parties votes	0.245	
5S votes	0.295	
center voting	0.067	*p-value*≤ 0.100 *in 0.69% of regressions*
social networks	0.131	
trust	0.249	
vax coverage	0.210	

Checking for multicollinearity, the mean VIF equals 6.79 that is below the threshold level of 10. Cramer's associations among categorical variables are weak, aside association between hub and central Areas which is moderate. Therefore, there are no serious issues for collinearity. The Ramsey test for omitted variables -estat ovtest- returns a p-value = 0.1597, therefore there are no omitted variables. Checking for outliers, the plot of residuals over fitted values highlights a potential mild outlier, for residual values above 5.



The outlier corresponds to the LLM "*Petralia Sottana*", located in Sicily. Noteworthy, the geo-located dataset of online hate tweets "The Italian Map of Hate" allocates a surprisingly high number of tweets to this LLM. Thus, OLS estimation of the fitted model is done removing the outlier, and results do not change (Column 4 in Table 3, main text).

Annex A3. SSM model

Fitted model specification and robustness check. To alleviate potential collinearity and identification issues that may arise in Heckman estimates, an exclusion restriction is introduced in the selection equation, and the set of covariates in the conditional equation is constrained to be a strict subset of the covariates in the selection equation (Wooldridge, 2002). Therefore, X_2 is a sub-set of X_1 and X_1 contains at least one covariate -the exclusion restriction- that is significantly negatively related to the dependent variable of the selection equation, while being unrelated to the dependent variable of the conditional equation. The first stage implies identifying the exclusion restriction. To this respect, a valid exclusion restriction must be such that: *(i)* it is significantly and negatively associated with the dependent variable of the selection equation; *(ii)* it is not a relevant predictor of outcome in the conditional equation (Cameron and Trivedi, 20010). The considered exclusion restriction is given by the share of population above 70-years old over digital natives, which satisfies both conditions. Second, consistent and robust estimates are achieved by performing two-step procedure (Cameron and Trivedi, 2010; Wooldridge, 2002). Third, errors are clustered at regional level to account for potential heteroskedasticity. To account for these three relevant conditions to get robust estimates, the SSM is estimated following the two-step procedure as follows (Greene, 2003; Wooldridge, 2002; Leung and Yu, 1996): the probit model with exclusion restriction is estimated first to find also the value for the inverse Mills ratio; the OLS regression model is estimated adding the inverse Mills ratio to the regressors[48].The regressors in the OLS model are constrained to be a subset of the probit regressors. Results are presented in

[48] Use of STATA command -heckman- is ruled out since it is not possible to use the twostep option together with clustered standard error.

columns 1-2 in Table A3.1. It is also checked the robustness of the fitted model by clustering error at province level (columns 3-4 in Table A3.1), to see that results hold.

Table A3.1: SSM fitted specification, errors clustered at regional and province level

Variables	(1) SSM-probit cluster region	(2) SSM-OLS cluster region	(3) SSM-probit cluster province	(4) SSM-OLS cluster province
share of + 70 over -15	-0.167*		-0.167*	
	(0.0994)		(0.0939)	
population (ln)	0.987***	0.121	0.987***	0.121
	(0.146)	(0.103)	(0.159)	(0.0839)
unemployment	-0.0469**	-0.0281	-0.0469*	-0.0281
	(0.0192)	(0.0260)	(0.0262)	(0.0307)
educated people	0.153***	0.170***	0.153***	0.170***
	(0.0586)	(0.0352)	(0.0572)	(0.0275)
hub in the LLMA	0.141	0.317*	0.141	0.317**
	(0.377)	(0.162)	(0.388)	(0.144)
central area	-0.390	-0.170*	-0.390	-0.170
	(0.306)	(0.0966)	(0.297)	(0.131)
right parties votes	0.0289*	0.00726	0.0289*	0.00726
	(0.0171)	(0.0131)	(0.0170)	(0.0115)
5 star votes	0.0265*	0.0247*	0.0265	0.0247
	(0.0146)	(0.0140)	(0.0178)	(0.0171)
vax coverage	-0.0444**	-0.00354	-0.0444*	-0.00354
	(0.0176)	(0.0157)	(0.0268)	(0.0163)
offline hate	0.566*	0.213*	0.566*	0.213**
	(0.298)	(0.123)	(0.304)	(0.108)
incomplete stem family (NE)	5.387***	0.667**	5.387***	0.667**
	(0.697)	(0.239)	(0.822)	(0.259)
stem family (NE)	5.021***	-0.210	5.021***	-0.210
	(1.083)	(0.401)	(1.183)	(0.473)
communitarian family (E)	-0.843*	-0.426	-0.843**	-0.426
	(0.492)	(0.279)	(0.421)	(0.418)
Inverse Mills ratio		0.972***		0.972***
		(0.229)		(0.317)
Observations	611	524	611	524
R-squared		0.233		0.233
Regional FE	YES	YES	YES	YES

*Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1*

Then, it is verified the robustness of the exclusion restriction, by considering over 74-year-old instead of over 70-year-old, since the share of internet users decreases even more as age rises (Table A3.2). Results stay consistent, moreover, by increasing the age threshold for the elders does not alter much the effect of the exclusion restriction. Therefore, the exclusion restriction performs well.

Table A3.2: different specifications for the exclusion restriction

Variables	(1) SSM probit (74)	(2) SSM OLS (74)	(3) SSM probit (70)	(4) SSM OLS (70)
Pop over 74/pop under 15	-0.257*			
	(0.137)			
pop over 70/ pop under 15			-0.188*	
			(0.0994)	
population (ln)	0.979***	0.0892	0.981***	0.0878
	(0.146)	(0.102)	(0.145)	(0.102)
unemployment	-0.0435***	-0.0272	-0.0424**	-0.0273
	(0.0164)	(0.0259)	(0.0166)	(0.0259)
educated people	0.183***	0.145***	0.182***	0.144***
	(0.0515)	(0.0375)	(0.0519)	(0.0375)
Gini index	-0.0425	0.0383	-0.0396	0.0385
	(0.0535)	(0.0267)	(0.0527)	(0.0267)
inequality*educated people	0.000154	0.00884	0.000408	0.00887
	(0.0164)	(0.00791)	(0.0164)	(0.00790)
Hub in the LLMA	0.129	0.348**	0.134	0.348**
	(0.375)	(0.159)	(0.378)	(0.159)
Central area	-0.404	-0.167	-0.403	-0.167
	(0.307)	(0.0990)	(0.309)	(0.0993)
Right parties votes	0.0320**	0.00269	0.0321**	0.00261
	(0.0162)	(0.0148)	(0.0164)	(0.0148)
5 star votes	0.0287**	0.0173	0.0296**	0.0171
	(0.0146)	(0.0153)	(0.0145)	(0.0153)
vax coverage	-0.0425**	-0.00239	-0.0425**	-0.00238
	(0.0193)	(0.0174)	(0.0192)	(0.0174)
offline hate '15	0.571*	0.216*	0.578*	0.216*
	(0.310)	(0.118)	(0.310)	(0.118)
incomplete stem family	5.484***	0.636**	5.460***	0.633**
	(0.686)	(0.249)	(0.683)	(0.249)
stem family	5.438***	-0.466	5.404***	-0.472
	(1.023)	(0.486)	(1.025)	(0.486)
communitarian family	-0.850*	-0.285	-0.842*	-0.284
	(0.494)	(0.345)	(0.492)	(0.344)
Inverse Mills ratio		0.825***		0.818***
		(0.264)		(0.267)
Observations	611	524	611	524
R-squared		0.237		0.237
Regional FE	YES	YES	YES	YES

Robust standard errors in parentheses
**** p<0.01, ** p<0.05, * p<0.1*

One of the most acknowledged among critical issues with regards to SSM is multicollinearity (*i.a.* Cameron and Trivedi, 2010). The two-step estimation with exclusion restriction allows to alleviate for this, which is further evaluated by performing the following postestimation tests. First, multicollinearity for the condition equation is measured to get a mean VIF of 4.46 that does not suggest concern for multicollinearity. However, in SSM, multicollinearity represent a cause for warning also with respect to the inverse Mills ratio (Leung and Yu, 1996). To account for this issue, the inverse Mills ratio is regressed on the condition equation regressors to get a mean VIF = 4.43, suggesting

that collinearity between the inverse Mills ratio and the condition equation regressors is not a severe issue. Then, multicollinearity for the selection equation, is considered, to get a mean VIF equal to 1.91, again suggesting that collinearity does not represent a cause for alarm in both stages of SSM. Regarding goodness-of-fit and omitted variables, the `linktest` for the selection equation and the Ramsey test for the condition equation support the fitted model specification[49]. Then, outliers are considered, since the data are characterized by a surprisingly number of hate tweets allocated to the LLM "Petralia Sottana". Therefore, the SSM specification is estimated removing the outlier and results are presented in Table A3.3.

Table A3.3: effects of mild outlier on SSM condition equation

Variables	(1) SSM-OLS	(2) SSM-OLS no outlier
population (ln)	0.121	0.101
	(0.103)	(0.0912)
unemployment	-0.0281	-0.0204
	(0.0260)	(0.0211)
educated people	0.170***	0.174***
	(0.0352)	(0.0341)
Hub in the LLMA	0.317*	0.315*
	(0.162)	(0.162)
Central area	-0.170*	-0.164
	(0.0966)	(0.0961)
Right parties votes	0.00726	0.00844
	(0.0131)	(0.0135)
5 star votes	0.0247*	0.0258*
	(0.0140)	(0.0139)
vax coverage	-0.00354	-0.00533
	(0.0157)	(0.0171)
offline hate	0.213*	0.210
	(0.123)	(0.122)
incomplete stem family (NE)	0.667**	0.656**
	(0.239)	(0.232)
stem family (NE)	-0.210	-0.163
	(0.401)	(0.405)
communitarian family (E)	-0.426	-0.400
	(0.279)	(0.282)
Inverse Mills ratio	0.972***	0.856***
	(0.229)	(0.243)
Observations	524	523
R-squared	0.233	0.244
Regional FE	YES	YES

Robust standard errors in parentheses
**** p<0.01, ** p<0.05, * p<0.1*

[49]The linktest for the selection equation gives a p-value for _hatsq = 0.713 and a p-value for _hat=0.000. The Ramsey test for the condition equation gives a p-value of 0.8989 implying that we cannot refuse the assumption of no omitted variables.

To see that, differently from 2PM, removing the outlier slightly changes results, by making the feature of LLMA as central area and offline hate occurrence no more significant predictor for the amount of hate tweets. The role of influential points in the selection equation is scrutinised to get that they do no drive results (Table A3.4).

Table A3.4: effects of mild outlier on SSM selection equation

Variables	(1) SSM-probit	(2) SSM- probit no influentials
share of + 70 over -15	-0.167*	-0.167*
	(0.0994)	(0.0994)
population (ln)	0.987***	0.987***
	(0.146)	(0.146)
unemployment	-0.0469**	-0.0469**
	(0.0192)	(0.0192)
educated people	0.153***	0.153***
	(0.0586)	(0.0586)
Hub in the LLMA	0.141	0.141
	(0.377)	(0.377)
Central area	-0.390	-0.390
	(0.306)	(0.306)
Right parties votes	0.0289*	0.0289*
	(0.0171)	(0.0171)
5 star votes	0.0265*	0.0265*
	(0.0146)	(0.0146)
vax coverage	-0.0444**	-0.0444**
	(0.0176)	(0.0176)
offline hate (Y/N)	0.566*	0.566*
	(0.298)	(0.298)
incomplete stem family (NE)	5.387***	5.887***
	(0.697)	(0.694)
stem family (NE)	5.021***	5.521***
	(1.083)	(1.083)
communitarian family (E)	-0.843*	-0.843*
	(0.492)	(0.492)
Observations	611	607
Regional FE	YES	YES

Robust standard errors in parentheses
**** p<0.01, ** p<0.05, * p<0.1*

Assessing SSM results against the 2PM allows to conclude that, once the selection bias due to age stratification in the LLMA is controlled for, all meaningful predictors identified in the probit specification in 2PM are consistent predictors in SSM too; moreover, in the SSM set-up, offline hate occurrence becomes a robust predictor for the occurrence of online hate (columns 1 and 3 – Table A3.2). Finally, a single-stage OLS is estimated, alongside an alternative normalization for the dependent variable, given by hate tweets divided by resident population. Results for these alternative specifications are summarized in Table A3.5 below, to confirm the main results.

Table A3.5: Estimation results for competing model specifications (2PM and single-stage OLS) on two alternative normalizations for the dependent variable

	lny_a = *ln(share of hate tweets on total tweets)*			lny_b = *ln(hate tweets on population)*		
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	single stage OLS coef	2PM OLS coef	2PM OLS coef	single stage OLS coef	2PM OLS coef	2PM OLS coef
	cluster region	*cluster region*	*cluster province*	*cluster region*	*cluster region*	*cluster province*
educated people	0.245***	0.205***	0.213***	0.198***	0.130***	0.129***
	(0.0296)	(0.0307)	(0.0363)	(0.0324)	(0.0331)	(0.0296)
Gini index	**0.0738*** **	**0.0723** **	**0.0814** **	**0.0493* **	**0.0429* **	**0.0596* **
	(0.0279)	**(0.0309)**	**(0.0382)**	**(0.0236)**	**(0.0222)**	**(0.0331)**
educated people#	**0.0136* **	**0.0206*** **	**0.0180** **	**0.00239**	**0.0133* **	**0.0139** **
Gini index	**(0.00702)**	**(0.00791)**	**(0.00888)**	**(0.00829)**	**(0.00741)**	**(0.00691)**
Hub in the LLMA	0.905***	0.788***	0.909***	0.490***	0.349**	0.445***
	(0.167)	(0.157)	(0.167)	(0.131)	(0.146)	(0.150)
Central area	0.0639	0.0417	0.0205	-0.117	-0.167*	-0.193
	(0.154)	(0.143)	(0.160)	(0.114)	(0.0923)	(0.144)
incomplete stem family (NE)	0.158	0.541	1.595**	0.257***	0.671***	2.217***
	(0.650)	(0.720)	(0.652)	(0.0871)	(0.0759)	(0.836)
stem family (NE)	0.910*	0.954	2.723***	1.159***	1.139***	3.577***
	(0.546)	(0.633)	(0.680)	(0.135)	(0.0949)	(0.858)
communitarian family (E)	-0.0610	0.618	-	-0.169	0.511	-
	(0.818)	(0.853)		(0.317)	(0.376)	
offline hate '15 (Y/N)	0.887***	0.715***	0.548***	0.389***	0.179*	0.0989
	(0.132)	(0.126)	(0.142)	(0.118)	(0.0971)	(0.129)
Observations	611	524	524	611	524	524
R-squared	0.504	0.459	0.590	0.301	0.223	0.363
Regional FE	YES	YES	NO	YES	YES	NO
Province FE	NO	NO	YES	NO	NO	YES

Annex A4. Discriminating between 2PM, SSM and single-stage OLS

There is a well-established and enduring debate between Heckman sample selection and two-part models in econometrics. Issues such as the validity of the exclusion restriction and the ability of detecting multicollinearity are the most challenged features with respect to SSM. Whereas considering the 2PM, the quite stringent assumption of independence between occurrence of one event and its expected amount may not hold in many scenarios. As a consequence, many contributions have been produced and challenged to identify formal statistic procedures capable of discriminating between the two models, especially when it is not possible to choose ex-ante on theoretical arguments (Dow & Norton, 2003; Madden, 2008; Santos-Silva et al., 2015).

The starting point are the baseline results with respect to the SSM: in all model specifications, the coefficient for the inverse Mills ratio is significant; to this respect, I have also performed the preliminary test for the existence of selection bias as suggested by Wooldridge (2003): a standard t test on the estimated coefficient of the inverse Mills ratio under the null of no selection bias, $H_0: \beta_{IMR} = 0$. There is preliminary support for the existence of a selection bias[50]. However, discriminating between 2PM and SSM relying only of this test cannot be sufficient, since the power of the test will be limited by the presence of collinearity issues (Dow and Norton, 2003; Leung and Yu, 1996; Madden, 2008). Indeed, if a standard t test is performed as specification test, then also the degree of multicollinearity between the inverse Mills ration and the condition equation regressors should be reported. Therefore, the next check implies assessing the degree of collinearity between the inverse Mills ratio and X_2. This can be done by regressing the inverse Mills ratio on X_2: for this regression we get a VIF = 4.43, suggesting that collinearity does not appear to be an issue (Madden, 2008). However, a high VIF for the inverse Mills ratio is a sufficient but not a necessary condition for collinearity, therefore finding a low value for the VIF between the inverse Mills ratio and X_2 is not enough to rule out collinearity (Madden, 2008). An even more robust assessment evaluates the value of the condition number for the regression of the inverse Mills ratio on X_2: if the value of the condition number is below 20, then there is no strong collinearity weakening the

[50] The *t* test gives a p = 0.0004

SSM (Leung and Yu, 1996). Also this procedure is pursued to get more evidence ruling out severe collinearity[51]. Another statistical method entails determining the value for correlation coefficient of the error terms from the selection and the condition equation, ρ, using the procedure detailed by Manning, Duan, & Rogers (1987); in the case of this paper $\rho = 0.833$, further supporting the absence of strong collinearity issues affecting our SSM specification .

Having assessed that model specifications appear quite robust, it is important to gauge which modelling strategy fits data better. Therefore, the discrimination test based on predictive power of opposing models with respect to the data is used (Cameron and Trivedi, 2010; Dow and Norton, 2003; Madden 2009). In order to get prediction from the 2PM and the single-stage OLS retransformation bias has to be accounted for (Cameron and Trivedi, 2010). To this respect, it is possible either to assume that error terms in the OLS equation are normally distributed with mean 0 and variance σ , implying that we can get the following expression for unconditional means

$$E[onlinehate|X] = \Phi(X_1\alpha_1)\exp\left(X_2\alpha_2 + \frac{\sigma_2^2}{2}\right)$$

or this strong assumption can be relaxed by claiming that error terms in the OLS equation are *i.i.d*, allowing to apply the Duan's smearing factor such that retransformation works as follows,

$$E[onlinehate|X] = \Phi(X_1\alpha_1)\exp\left(X_2\alpha_2 + \frac{\sum_{j=1}^{N}\varepsilon_{2j}}{N}\right)$$

Both approaches are considered, first assuming that residuals are normally distributed; second, considering the weaker assumption that residuals are independent and identically distributed, applying the Duan smearing retransformation.

Then, predictions form the SSM are determined, through the following expression for unconditional mean (Dow and Norton, 2003; Santos-Silva *et al.*, 2015)

$$E[onlinehate|X] = \Phi(X_1\alpha_1 + \rho\sigma_2)\exp\left(X_2\alpha_2 + \frac{\sigma_2^2}{2}\right)$$

[51] Leung and Yu (1996) detail also how to assess multicollinearity between the inverse Mills ratio and the condition equation regressors through the condition number. We follow the procedure detailed by them to get a condition number equal to 10.1368 for our condition equation. Notably the value for the condition number is way below the threshold level of 20 that they indicate as the value that, once crossed, signals collinearity problems.

Prediction form the three competing specifications are detailed in the following tables.

Table A4.1a: Predictive power of competing modelling strategies

a. y = *hate tweets on population*	*Obs*	*Mean*	*Std. Dev.*	*Min*	*Max*
hate tweets on resident population	611	**0,000663**	0,001851	0	0,026099
yhat_2PM_normal	611	**0,000652**	0,000645	3.87e-06	0,007366
yhat_2PM_Duan	611	0,000864	0,000855	5.13e-06	0,009766
yhat_SSM_Normal	611	0,000738	0,000678	0,000086	0,006916
yhat_OLS_Normal	611	0,000832	0,001183	0,0000926	0,011189
yhat_OLS_Duan	611	0,001169	0,001546	0,0001302	0,015722

Assessing predictions from SSM against predictions from 2PM and single-stage OLS -under normality and *iid* of errors- with respect to observed data, allows to conclude that 2PM under normality assumption outperform the competing specifications in terms of mean. Same conclusions apply when the outcome variable is given by hate tweets on total tweets, as summarized in the table below.

Table A41.b: Predictive power of competing modelling strategies

b. y = *hate tweets on total tweets*	*Obs*	*Mean*	*Std. Dev.*	*Min*	*Max*
hate tweets on total tweets	611	**0,001301**	0,0078335	0	0,173595
yhat_2PM_normal	611	**0,001293**	0,0037686	5.81e-06	0,073603
yhat_2PM_Duan	611	0,001629	0,0047469	7.32e-06	0,092711
yhat_OLS_Normal	611	0,001443	0,0041946	0,000047	0,076929
yhat_OLS_Duan	611	0,002009	0,0058409	0,000064	0,107123

CHAPTER 3

Looking Ahead in Anger. The effects of foreign migration on youth resentment in England

Abstract. This paper investigates whether school violence is affected by a migration shock capable of determining sudden and sizeable changes in the local composition of cultural-ethnic groups. The analysis exploits the natural shock from migration which occurred in the UK after the 2004 European Union enlargement to empirically estimate how the migrants' inflow has influenced bullying among 15-years-old across the 150 English Local Authorities. Findings highlight a relevant effect of the shock from migration, which is robust to the accounting for potential endogeneity with respect to immigrants' location choice. The role of existing language barriers as moderator for the effect of the migration shock is also scrutinised, to find that they increase its effect.

Keywords: resentment, bullying, migration, conflict, cultural threat, inequality, England, contact theory

3.1 Introduction

The present paper will examine the effect of foreign immigration on youth resentment at the local level, focusing on teenagers. This topic emerges as relevant, in the light of the thriving research detailing the meaningful effects which foreign migration exerts on many local socioeconomic elements, such as the labour market and some forms of local discontent. These research efforts are focused on the adult population only, hence, little is known about the effects that foreign migration might exert on youth resentment. The latter appears an important element to investigate since youth cohorts will grow older, hence understanding whether immigration contributes to shape their attitudes is relevant also for a better comprehension of adult cohort behaviours, especially considering that

adolescent cohorts are few years away from voting and entering the job market. Moreover, recent evidence shows that local conflicts between ethnic/cultural groups are leading to ethnic/cultural-based victimization at school (Álvarez-García, García, & Núñez, 2015; OECD, 2017). In the US and in the UK, for instance, ethnic/cultural features are the motivations behind most bullying incidents at school happening lately (Childline, 2019; Marsh & Mohdin, 2018; Southern Poverty Law, 2019; Weale, 2019).

This paper investigates the effect of the fast and large migration influx which occurred after the EU 2004 enlargement on school bullying in England. This analysis is possible given new available data on school-bullying among 15-years-old covering all local authorities in England (Health and Social Care Information Centre of the UK Government, 2016). 15-years-old appears as a valuable young cohort to assess the relationship between school bullying and sociocultural threat from migrations, since they are experiencing a local milieu overcoming the borders of family and school (Southern Poverty Law, 2019) to reach the more complex local socioeconomic structures. Addressing the English context is appealing for two main reasons. First, the United Kingdom displays a remarkably high share of bullying victimization, ranking 4th among OECD countries in terms of highest bullying victimization at school and 2nd if we restrict the subset to OECD European countries (OECD, 2017). Second, the EU 2004 enlargement has generated a sizeable shock from migration for the UK, up to the point that in few years Polish has become the most common non-British nationality/country-of-birth (Office for National Statistics - ONS, 2017)[52]. The EU 2004 Accession migration to the UK pinpoints itself as a potential trigger for sociocultural threat (i.a. Goodwin & Milazzo, 2017), because the spatial distribution of migrants from Eastern Europe after the Accession follows quite distinct patterns compared to the locational preference of Eastern European migrants that have arrived prior to 2004 (Becker & Fetzer, 2017). In fact, Accession migrants have often settled in places with little or low familiarity with Eastern European culture (Pollard, Latorre, & Sriskandarajah, 2008).

Criminology and social psychology provide a sound background to support the relevant nexus between foreign migration and youth resentment, having identified bullying as the specific hate-related behaviour enacted by young cohorts. According to

[52]Polish represents by far the largest nationality of Accession Countries migrants.

these contributions, school bullying happens to protect existing sociocultural hierarchies against perceived threats posed by minority groups (Gerstenfeld, 2017). Both the perceived social threats and the sociocultural hierarchies that bullying is meant to protect are taken from the broader environment (Vertovec & Coen, 2002), since schools are not "*hermetically sealed institutions*" (Southern Poverty Law, 2019) immunized from the local context. Anecdotal evidence on England is highlighting how bullying at school is mainly influenced by what is happening in the wider community in terms of migration history and the consequent blurring of social identities driven by the changing ethnic/cultural outlooks (Reynolds, 2008; Sime et al., 2017). Alongside criminology and social psychology, the relationship between local social anxiety and migration-driven changes is gaining momentum also in the socioeconomic literature (i.a. Goodwin & Milazzo, 2017; Halla, Wagner, & Zweimüller, 2017; Rodrik, 2018), although with respect to adult cohorts only. Evidence shows that the sociocultural outlook brought along by newcomers may be perceived as a menace to group-related attitudes and to symbols shared by the local incumbent population. Therefore, the sociocultural dimension[53] is recognized as the channel through which migration influences social tensions and hostility (Card et al., 2012; Guiso et al., 2017; Hainmueller & Hopkins, 2014; Hangartner et al., 2018). On this bulk of evidence, Newman (2013) has developed a multicultural "*defended neighbourhood*" hypothesis stating that large influx of migration will be most culturally threatening for citizens residing in contexts with minimal pre-existing experience of the population that is moving to their place. Pre-existing experience of incoming migrants eases the mitigation of the cultural shock, since the receiving place is already familiar with some levels of their cultural outlook.

Up to now and to the best of my knowledge, the investigation on the relationship between immigration and social tension has focused on the adult population, notwithstanding that figures are showing how bullying is increasingly characterized by a cultural/ethnic polarization (Rogers et al., 2017; Southern Poverty Law, 2019) and how bullying magnitude is growing in many countries (OECD, 2017; UNESCO, 2017). Moreover, current research on school bullying is calling for the exploration of the geography of sociocultural triggers (Pells, Ogando Portela, & Espinoza, 2016; Tippett &

[53]Defined also as sociotropic by many contributions (Hainmueller & Hopkins, 2014)

Wolke, 2014; Wilkinson & Pickett, 2017), stimulated by the non-negligible spatial heterogeneity in bullying rates (Health and Social Care Information Centre of the UK Government, 2016; OECD, 2017).

Hence, on the one side bullying literature identifies local sociocultural threats as risk factors for bullying and on the other side social sciences recognize that fast and sudden migration inflows can act as meaningful sources for sociocultural threats at local level. Therefore, the main goal of the present paper is to build upon these developments, by bringing them together in the form of empirically assessing whether school violence is influenced by sudden and sizeable inflows of an immigrant group in places where the immigrant group has largely been absent. In other terms, the paper empirically assesses whether a "*defended schools*" hypothesis may apply to young cohorts, analogously to what has been demonstrated with respect to the "*defended neighbourhood*" hypothesis for adult cohorts.

In the paper, bullying is measured using new data from the What About YOUth? Survey (Health and Social Care Information Centre of the UK Government, 2016), which are characterized by being robust at the fine spatial granularity of Upper-Tier Local Authority (UTLA)[54]. According to this survey, more than the 50% of English 15-years-old report that they had been bullied in some form in the previous couple of months (Health and Social Care Information Centre of the UK Government, 2015b). These data are combined in a novel dataset with the Becker, Fetzer & Novy (2017) measure of exposure of UTLAs to migration occurring after the 2004 EU Enlargement and with administrative data on the local socio-economic outlook to control for cofounding features. The resulting novel database contains individual-level data on the 15-years-old participants to the survey which are merged with spatial-level data on the local exposure to the migration shock. The database allows to analyse, through estimation of a multilevel logit, whether the bullying victimization is influenced by the migration shock controlling also for individual characteristics of the target -such as ethnicity, gender, deprivation.

Estimation results presented in the paper entails several remarkable insights supporting the "*defended school*" hypothesis. First, local exposure to sharp and fast

[54]England consists of 152 Upper-Tier Local Authority (UTLA): in the WAY-2014 two UTLAs are merged to the nearest neighbour due to small population (City of London with Hackney and Isles of Scilly with Cornwall). Therefore, survey results of the WAY-2014 refer to 150 UTLAs and the same holds for the analysis in this paper.

migration inflows of unfamiliar cultural groups emerges as a robust risk factor for bullying, even when individual-level features and other potential confoundings are controlled for. Noteworthy, this result is confirmed also when the potential endogeneity of the scrutinized regressors is accounted for. This finding provides strong evidence that a local cultural shock from migration propagates also inside schools. This evidence aligns with existing contributions supporting the role of perceived sociocultural threat in boosting social tension and violence among adult cohorts, at the same time broadening the range of engaged actors to include 15-years-old.

The outcomes outline robust evidence of non-negligible associations between the local economic geography and bullying victimization, alongside the influence of the migration shock. To this respect, findings show that local poverty per se does not constitute a risk factor for bullying victimization, while places characterized by higher spatial polarization in terms of income are associated to more bullying.

Finally, the paper also presents and measures a potential moderator for the effect of migration on school bullying, namely language barriers, measured through the school level exposure to pupils whose first language is not English. The results provide empirical support to this point, showing that a higher exposure to non-native English speakers at school level corresponds to a stronger effect of the migration shock on school violence. These findings relate to the existing literature on the role of local language barriers in preventing the assimilation of cultural diversity even where different cultural groups reside in the same place (Newman, 2013). The existence of language barriers prevents the integration of the different cultural groups settled in the area; hence the arrival of a new cultural group cannot benefit from the place being capable of dealing with cultural diversity.

Overall, the results depict robust patterns of local differences in bullying victimization, strongly supporting a non-spatially blind approach with respect to school violence, with implications also in terms of policy design. The empirical evidence presented here clearly suggest including within bullying-prevention programs both measures to favor the local assimilation of unassimilated cultural groups and measure to moderate the local outlook economic inequality.

The remainder of the paper is organized as follows. Initially, a review of relevant research on bullying and sociocultural threat is outlined. Then, data and variables used in

the subsequent empirical estimation are introduced and described. The subsequent section details the empirical strategies and the different models that are assessed in the paper. The results are then presented and discussed. Finally, concluding comments are presented.

3.2 Review of Related Literature and Research Questions

The present paper cuts across different disciplines, being related to the literature on the risk factors of school bullying, the literature on the effects of foreign migration and research on the influences of places on health and well-being ("*situationist approach*").

Bullying is both a societal challenge and a public health concern: its numbers are growing[55] as well as its costs in term of health, human capital and social fabric (Ammermueller, 2012; Brown & Taylor, 2008; Tippett & Wolke, 2014). The negative effects caused by bullying, alongside its pervasiveness, have pushed it in the top of in many institutional agendas[56]. Research targeting the identification of risk factors associated with school bullying has been thriving, too. Seminal approaches have focused on the link between bullying and individual/family characteristics, though producing mixed empirical evidence up to the point that scholars have asserted the impossibility of identifying robust individual types for either bullying perpetrators or victims (Álvarez-García et al., 2015; Espelage et al., 2018). Also the evaluation of bullying-prevention programs based on individuals' personality traits has found that there might be factors other than individual and family characteristics acting as triggers (Juvonen & Graham, 2014; OECD, 2017). The identification of this "*personalized bias*" has pushed research to broaden the scope of investigation to include the effects of places[57] (Cook, Williams, Guerra, Kim, & Sadek, 2010; Espelage, 2014; Migliaccio & Raskauskas, 2016; Tippett &

[55] The recently included questions about bullying in the OECD PISA survey returned figures showing that, on average, the 18% of interviewed students experiences being bullied at least few times per month (OECD, 2017). The UNESCO survey at global level depicts that around 30% of students aged 13-15 experience bullying globally (UNESCO, 2018).

[56] Council of Europe Strategy on the Rights of the Child 2012-2015; ABC Europe Anti-Bullying Campaign; ENABLE European Network Against Bullying in Learning and Leisure Environment; UN SDG-4

[57] The approach of including the characteristics of places is labelled "*ecological perspective*" and it has been developed hinging on the seminal work of Bronfenbrenner (1979). The "*ecological perspective*" embeds violent behaviours within an environment organized in individual and local structures, with the aim of understanding how characteristics at different levels can either promote or moderate school bullying (Espelage, 2014; Hong & Espelage, 2012).

Wolke, 2014). School characteristics, neighbourhood characteristics and macro socioeconomic determinants have been consequently inserted in the analysis of school-bullying risk factors. Empirical evidence is still far from being conclusive (Álvarez-García et al., 2015; Espelage et al., 2018), although meta-analysis shows that examining the role of individual characteristics without any local element conveys a limited understanding of bullying and that community factors display important effect size (Cook et al., 2010).

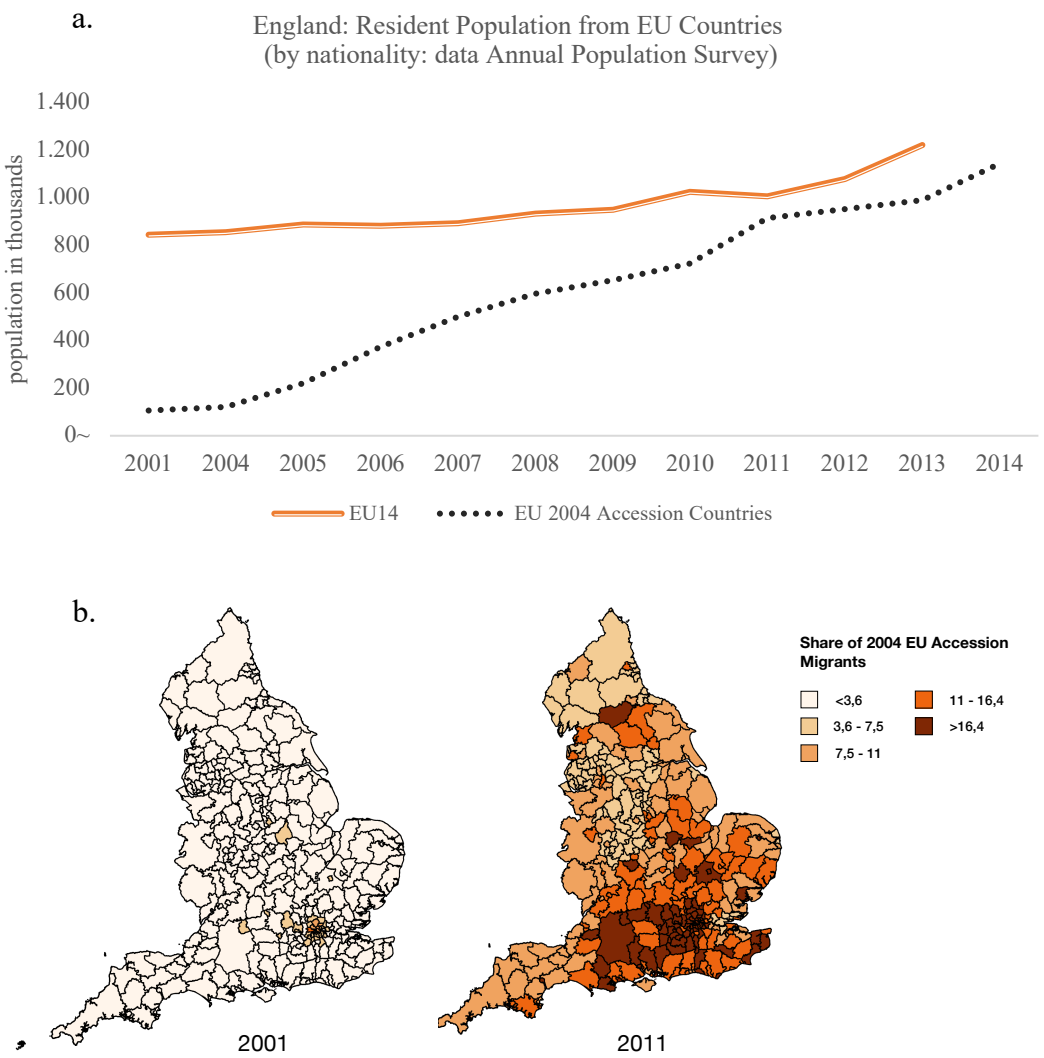
The growing awareness of the influence of places in shaping school bullying opens up to explore whether it is meaningful to bridge the literature on oppressive violence happening at school and the literature on the role of foreign migration in shaping social anxiety. To this respect, empirical research on the relationship between migration and places is increasing, mainly investigating along the economic dimension (i.a. Peri, 2016), the crime dimension (i.a. Pinotti, 2015) and electoral outcomes (i.a. Barone, D'Ignazio, de Blasio, & Naticchioni, 2016; Halla et al., 2017). Evidence is showing that the incumbent population develops hostility and resentment mostly triggered by threat to the sociocultural identity (Card et al., 2012; Enos, 2016; Hainmueller & Hopkins, 2014; Tabellini, 2017). Moreover, the relevance of sociocultural threats calls for a strong focus on local geographies, rather than on the nation as a whole, since many elements determining the identity of a group have a local dimension (Hopkins, 2010), starting from social norms and informal institutions. Within this field of research, increasing emphasis is placed on local demographic changes caused by migration inflows (Hopkins, 2010; Newman, 2013), as these changes threaten the status quo by affecting extant expectations of the incumbent population about the composition of the local community. Thus, the focus of investigation is the size of the exposure to newcomers considering the prior socio-ethnic structure. The same strand of literature pinpoints that the analysis on the effects of perceived sociocultural threats from migration should focus on outcomes capturing revealed preferences, given the growing evidence showing how individuals' response to migration is better captured through behaviours (such as electoral outcomes and violence) rather than through expressed opinions, the latter often suffering from social desirability biases (Hainmueller & Hopkins, 2014) Therefore, being that youth cohorts do not vote, school violence may represent an interesting revealed outcome to address.

Within this set up, foreign migration in England as a consequence of the EU 2004 Accession represents an appealing feature to assess. In the aftermath of the 2004 European Union Enlargement, eight Eastern European countries, alongside Cyprus and Malta, joined the European Union. The UK did not enforce any transition rules with respect to movement of people, experiencing a resulting mass migration significantly larger than anticipated. Figure 1 shows the sharp increase in Eastern Europeans arriving to England compared to the trend of EU 14 citizens, after the 2004 Accession (Figure 1.a). Figure 1.b details that the spatial distribution of these migrants is quite heterogeneous and distinct compared to Western Europeans (Becker et al., 2017). Eastern Europeans settling to England after the Accession have been spreading throughout the place being driven by job opportunities , with relevant settlements in small and rural areas which have previously attracted very few migrants[58] (Pollard et al., 2008).

Moreover, Eastern Europeans arriving after the EU04 Accession generally have scarce familiarity with either England or the English language and a low experience of cultural diversity in general, being that their home societies are predominantly ethnic-white (Sales, Ryan, Rodriguez, & Angelo, 2008). This fast and relevant arrival of unfamiliar cultural groups in many English areas after the 2004 EU enlargement represents a potential sociocultural threat fitting within the "*defended neighbourhood*" hypothesis (Becker & Fetzer, 2017). Indeed, figures show that hate crimes against Eastern Europeans started sharply rising after the Accession (Human Rights First, 2008; Rzepnikowska, 2019). In 2013, data collected from 26 Police Force Area in England display that one person every 14 hours have been arrested for hate crimes against Polish people (Mcdevitt, 2014). Following the Accession, British voters became increasingly concerned about the effects of immigration (Goodwin & Milazzo, 2017; Meleady, Seger, & Vermue, 2017). New anecdotal evidence suggest that the perceived threats to social identity posed by Eastern Europeans in the wider environment are at work also in schools, due to young cohorts' urges to protect the same social identity (Sime et al., 2017).

[58]only a small part of EU 2004 Accession migrants moved to London (28%), which was instead the preferred destination for Eastern Europeans migrants before 2004 (46% of them moved to London) (Becker et al., 2017)

Figure 1: Migration trends towards England by EU 2004 Accession Countries and the Spatial Distribution of Migrants from Accession Countries before and after the Accession



According to the "*defended neighbourhood*" hypothesis and its recent updates, the fact that Accession migrants have settled in areas which were not previously familiar with Eastern European culture ignites the perception of cultural threat in the local population, triggering tension and disorder. Then, according to bullying related literature, pupils might mimic this social stress in the school environment, by enacting cultural and ethnic violence to protect their social identity. Hence, there is a rationale for empirically verify whether the sociocultural threats arising from sudden inflow of migrants enact school violence in line with the "*defended neighbourhood*" hypothesis. To deal with established concerns about endogeneity of immigrant's location choices, a Control Function estimator approach and LIML estimation are performed among robustness checks,

instrumenting the current outlook of EU Accession migrants through shift-share with respect to the geography of EU 2004 Accession Countries migrants in 2000.

This paper refers also to the literature investigating the geography of health and well-being (i.a. Bilger & Carrieri, 2013; Diez Roux & Mair, 2010). The social determinants of health approach (SDH) focuses on analysing the social structures that influence people's chances to be healthy (Viner et al., 2012). Social determinants are defined by the World Health Organization as the conditions in which people are born, grow, live, work and age; conditions that are indeed shaped by socio-economic forces and their geographic distribution, therefore being local-scaled (Commission on Social Determinants of Health of the World Health Organization, 2008). Within this strand of literature, local communities, neighbourhood and school represents proximal factors that channel crucial structural factors towards adolescents (Viner et al., 2012). Noteworthy, there is still little systematic geographic work on the effects of socioeconomic determinants on adolescent health. Preliminary empirical works have already identified some influences of strain on school bullying (Patchin & Hinduja, 2011).

A final element considered in the paper refers to the role of pre-existing language barriers at local level in moderating the effects of a cultural shock from migration. This issue investigates the relationship between the effect of the cultural shock from migration and the pre-existing degree of assimilated cultural diversity in a given area. The degree of cultural diversity in a place is related to the "*defended neighbourhood*" hypothesis, since the prejudice activated in response to the initial appearance of a new cultural group may be moderated by pre-existing assimilation of cultural diversity, achieved through the prolonged presence of extant different cultural groups (Newman, 2013). However, for cultural diversity to be assimilated, it is important that the different cultural groups residing in the place experience repeated and intimate contact to attenuate prejudice, as argued by the "*contact theory*" (Hainmueller & Hopkins, 2014). For this contact to take place, the different groups need to communicate. The presence of language barriers inhibits contact among diverse groups, resulting in increased perception of obstacles to the incumbent dominant cultural group posed by the outgroups and increased uneasiness toward them overall (Hainmueller & Hopkins, 2014). Therefore, for a given level of pre-existing local cultural heterogeneity, language barriers may determine different levels of cultural diversity assimilation, which in turn may generate different effects of a given

cultural shock from inflows of new migrants. Evidence about England shows that speaking the language of the hosting country is a fundamental social norm for the incumbent cultural group, implying that cultural minorities have to abide to it to reduce the cultural distance from the dominant culture[59]. Therefore, the paper provides a preliminary answer on whether the level of pre-existing language distance between the dominant cultural group and the other cultural minorities moderates the effect of the cultural shock generated by inflows of a new cultural group.

3.3 Data and Descriptive Statistics

Previous contributions on bullying victimization have mainly considered national and/or cross-country level surveys targeting limited samples of schools (Álvarez-García et al., 2015; OECD, 2017; Tippett & Wolke, 2014). This paper specifically addresses the salience of the local dimension through a comprehensive geographical approach across the 150 Upper Tier Local Authorities (UTLAs), by assembling an original dataset that matches novel data on 15-years-old experience of being bullied resulting from the What About YOUth? 2014 survey (WAY-2014) with geographically disaggregated socio-economic data. Hence, the paper targets a geographic dimension that has been largely under scrutinized in bullying-related research, notwithstanding existing evidence shows non-negligible spatial heterogeneity in bullying rates at different geographic scales (Health and Social Care Information Centre of the UK Government, 2016; OECD, 2017).

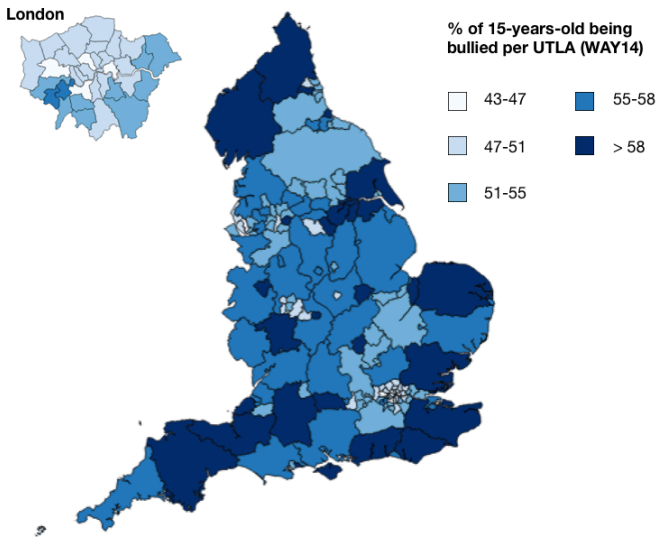
Measuring the experience of being bullied at local level. The WAY-2014 survey represents a landmark as it covers the entire English context, providing robust data at a fine-grained geography[60]. Up to now, the WAY-2014 survey is also the largest-scale survey addressing school bullying, since more than 110.000 15-years-old answered questions about bullying. It has been carried on between Sept. 2014 and Jan. 2015, targeting pupils that have been identified using the UK Department for Education National Pupil Database (NPD) and it contains questions addressing bullying

[59]95% of respondents to the British Social Attitude Survey in 2014 thinks that to be "*truly British*" people have to speak English (NatCen Social Research, 2014).

[60]The WAY-2014 is characterized by sufficient observations collected in each UTLA to achieve a +/- 3% margin error at a 95% CI for youth in England aged 15 in each UTLA (Health and Social Care Information Centre of the UK Government, 2015a). A random sampling methodology was used in order to ensure that no bias was introduced into the sample selection and that the survey sample reflected the population.

specifically[61]. Out of the WAY-2014 survey, Przybylski and Bowes (2017) have designed an anonymized database on individual experience of bullying containing 110,788 observations, corresponding to the number survey participants having answered bullying-related questions. They amount to the 92% of total respondents, showing good internal consistency (α=0.90) and accounting for a weighted population of 520,221 pupils. To this respect, almost all previous studies have sampled less than 50,000 pupils, often across countries (Álvarez-García et al., 2015).

Figure 2: The geography of bullied 15-years old across England UTLA



Therefore, this dataset entails suitable characteristics to provide a locally robust proxy for bullying in England at a fine spatial granularity, allowing to develop an empirical estimation capable of sound inclusion of the local-level covariates. The geography of the experience of being bullied from WAY-2014 shows non-negligible spatial heterogeneity across the UTLAs, as outlined in Figure 2.

Measuring the exposure to a migration shock at local level. The fine-grained WAY-2014 data on the 15-years-old experience of being bullied are merged with administrative data, starting from the most recent shock from migration inflows that has occurred to England, using the Becker, Fetzer & Novy (2017) database, which contains a measure

[61]Questions 46-47 (Health and Social Care Information Centre of the UK Government, 2015a). More details on WAY-2014 survey and bullying-related questions are provided in Annex A1. Cyberbullying is not considered in this analysis.

for the EU 2004 Accession migration. Using data from the 2001 and 2011 census, Becker and Fetzer (2016) create the following measure of exposure to migration from EU accession shock,

$$AccessionShock_{UTLA_j} = \frac{Accession\ Migrants_{UTLA_{j,2011}} - Accession\ Migrants_{UTLA_{j,2001}}}{EU\ Migrants_{UTLA_{j,2001}}} \quad (1)$$

where the numerator for the *j*-th UTLA is given by the 2011-2001 difference in the size of UTLA *j* residents coming from EU Accession countries. The denominator is conveyed by the size of UTLA *j* resident population coming from EU countries that have been members of the European Union before 2004. The measure accounts for both the magnitude of immigration from EU accession countries and its effect relative to migration from the Western European countries (Becker & Fetzer, 2017). Finally, this measure is particularly relevant because it takes into account that a great influx of any immigrant group will be perceived as more of a threat among incumbents in places where the same immigrant group had previously been largely absent (Becker & Fetzer, 2017). Therefore, it is a measure consistent with the "*defended neighbourhood*" hypothesis, since places which are less experienced of migrants receive a stronger cultural shock from the inflow of migrants belonging to a new cultural group[62].

Control variables. To include potential confounding issues that are relevant with respect to bullying and oppressive deviant behaviours, several variables are considered. The economic dimension is increasingly acknowledged as significant element associated with bullying. Recent research has been outlining significant influence exerted by income inequality on school bullying at cross-national level (Due et al., 2009; Elgar et al., 2013; Wilkinson & Pickett, 2017) and at neighbourhood level, although the latter has been mainly limited to subsets of neighbourhoods located in cities (Juvonen & Graham, 2014; van der Ende et al., 2012). This paper considers both income deprivation and income inequality. The proxy for the former is conveyed through the Index of Income Deprivation, the official measure for relative income deprivation for local areas in England, provided by the UK Department for Communities and Local Governments (Department for Communities and Local Governments of the UK Government, 2015).

[62] To assess the relevance of demographic changes rather than static demographic outlooks, I have also estimated the effect of the local share of Eastern Europeans after the Enlargement, to get that there is no significant effect when levels rather than changes are considered.

The Income Deprivation Average Score gives a population-weighted summary of the average level of income deprivation in the UTLA, through the proportion of UTLA population experiencing deprivation relating to low income; if a UTLA scores 0,52, it means that 52% of the population is income deprived in that area. Moreover, since it relates to a proportion of the relevant population experiencing income deprivation, the Income Deprivation Average Score can be used to compare UTLAs on an absolute scale (Department for Communities and Local Governments of the UK Government, 2015). Notably, it is not a measure for inequality, because it measures only the share of income-deprived population. To grasp the inequality dimension, a measure for spatial income polarization is introduced, calculating the variation coefficient for the distribution of income across neighbourhoods within the same UTLA using data on income level at neighbourhood-scale from the Department for Communities and Local Governments.

Social frustration is measured through the number of children in need and the share of split population. Conversely, the share of same sex couples conveys a measure for openness. Also, UTLA population size and its ethnic composition are considered. Given the focus on the effect of a migration shock, local exposure to migration from other international contexts is assessed both as a cofounder and as a competitive treatment, again using data from Becker, Fetzer & Novy (2017). The school ethnic composition at UTLA level is contemplated, due to existing empirical evidence supporting its potential influence on bullying as well as on in-group/out-group relationships (Burgess & Platt, 2018). Since a recent study on US has highlighted that particularly predominantly white schools are more hostile environments for racial minorities (Rogers et al., 2017), the share of British white pupils is used to proxy established incumbent ethnicity. Other cofounding factors are given by crime, identified by the literature as a local element that may exert influence on school violence (Bowes et al., 2009) and unemployment share, another proxy for socioeconomic hardship (Guiso et al., 2017). Geography is also considered in terms of rural/urban dichotomy since extant literature has identified that this distinction may play a role (Juvonen & Graham, 2014; Smokowski, Cotter, Robertson, & Guo, 2013)[63].

[63]Other control variables that have been considered are juvenile crime, stock of human capital, religious outlook of places, adopted children, population density. However, they do not display significance at the same time not improving results fit.

Table 1: Descriptive statistics

Variable	number of observations	mean	sd	min	max
Level 1 (individual) variables					
Survey Weight	110788	4.696	2.859	1.079	33.85
Traditional bullying	110788	0.297	0.457	0	1
Belongs to non-white minority	110788	0.23131	0.42167	0	1
Male	110788	0.477037	0.499475	0	1
Lives deprived neighbourhood	110788	0.42957	0.495017	0	1
Level 2 (Local Authority) variables					
Split population (share)	110788	2.595	0.443	0.706	3.778
Same sex couples (share)	110788	0.0966	0.071	0.0203	0.549
Ethnic diversity Index (Simpson Inverse)	110788	6.478	9.731	0.332	54.61
IMD average income score	110788	0.153	0.0492	0.052	0.276
IMD crime score	110788	0.067	0.434	-0.804	1.019
Unemployment (share)	110788	5.528	3.455	0.037	12.002
Income variation coefficient	110788	0. 519	0.158	0.128	1.067
Children looked after (per 100k children)	110788	1.399	1.170	0.0315	5.378
Population size (ln)	110788	1.172	0.601	-0.984	2.684
2004 EU Accession Migration Shock	110788	0.171	0.116	0.0113	0.668
Exposure to migration from ROW	110788	0.062	0.066	-0.0005	0.3993
Rural	110788	0.3022	0.4592	0	1
Secondary school: British white (share)	110788	0.713	0.252	0.0475	0.966
Secondary school: percentage change EAL pupils 2012-2007	110788	1.353	0.368	0.296	2.753
School: Language diversity (Simpson Index)	110788	0.867	0.102	0	1

Finally, the Przybylski and Bowes (2017) database on individual experience of bullying contains also individual information on pupil's gender, ethnicity, neighbourhood deprivation. Gender and ethnicity are among the most scrutinized individual-level features in bullying-related literature (i.a. Álvarez-García et al., 2015; Espelage, 2014),

therefore they represent relevant control variables to include into the analysis. Neighbourhood deprivation is increasingly investigated, also as alternative to socio-economic status, being that the latter has failed to produce a meaningful and robust association with bullying (i.a. Hong & Espelage, 2012; Sykes, Piquero, & Gioviano, 2017; van der Ende et al., 2012). UTLA codes are used to match individual-level data and local level data. Data description is summarized in Table 1[64].

3.4 Estimation Strategy: 2-Level Hierarchical Model

Pearson chi-square test of bullying victimization provides supports for significant differences between UTLAs[65], as already suggested by Figure 2. The database allows to proper account for the salience of the spatial dimension, combining the predictive power of individual-level and spatial-level data. More into details, the variables contained in the dataset allow to analyse whether the bullying victimization is influenced by local features -as a migration shock- rather than individual characteristics of the target -such as ethnicity, gender, deprivation-, or both.

Being the outcome variable binary, multilevel (ML) logit estimation will be performed on two nested levels: each *i*-pupil (level 1) is nested in the *j*-UTLA (level 2). Formally letting y_{ij} denote the binary response variable of being bullied on the *i*-th pupil within the *j*-th UTLA,

$$\Pr(y_{ij} = 1) = \alpha_0 + \alpha_{0j} + \delta EU04_j + \alpha_1 x_{1ij} + \cdots + \alpha_k x_{kij} + \cdots + \beta_1 z_{1j} + \cdots + \beta_m z_{mj} \quad (2)$$

where $EU04_j$ is the migration shock in UTLA *j* after the 2004 European Union Accession measured through the Becker & Fetzer (2016) index of exposure. x_{1ij} to x_{kij} are the explanatory control variables at the individual level: *ith* pupil's gender (male or female/unknown), deprivation of neighbourhood of residency identified by postcode (yes/no)[66], belonging to a minority (yes/no); z_{1j} to z_{mj} indicate the explanatory control variables at the UTLA level *jth*: UTLA rurality (yes/no), share of split population, children in need, share of same sex couples, income deprivation score, share of British

[64]Data sources are presented in Annex A1, Table A1.1
[65]*p-value = 0.0000*
[66]The postcode identifies a small area with an average of 15 properties, which falls within a Lower-Layer Super Output Area (LSOA), which, in turn falls within a local authority.

white pupils at secondary school, population size, ethnic composition, crime and unemployment. α_{0j} is the random intercept (RI) component, capturing the UTLA effect for each *j*-th UTLA; it is assumed to be independent of the model covariates and independent and identically distributed as follows: $\alpha_{0j} \sim N(0, \tau^2)$. Finally, errors are clustered at UTLA level in all model specifications.

Preliminary stages are undertaken to further gauge suitability of estimation methodology[67], following established contributions in the literature. First, survey weights are scaled to reduce bias estimation (Asparouhov, 2006; Carle, 2009; Rabe-Hesketh & Skrondal, 2006a)[68]. Then, the pertinence of the hierarchical model is verified through postestimation diagnostics of the unconditional model with RI (Hamilton, 2013; Maas & Hox, 2005; Snjiders & Bosker, 2012)[69].

Figure 3: Ranking plot of estimated residuals for the 150 UTLA

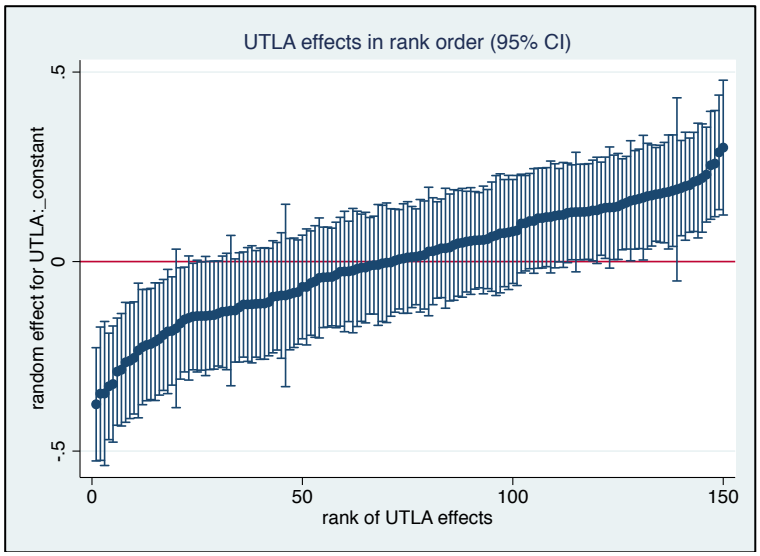


Figure 3 plots the estimates of the UTLAs effects from the unconditional model specification in rank order. Clearly, the probability of being bullied for 15-years old is significantly above (below) average in a substantial amount of UTLAs, as outlined by the mass of confidence intervals which do not overlap the horizontal line at zero. Since the

[67] The proposed ML logit is estimated by maximum likelihood using Stata melogit. Sample size is large (110,788 observations) and the number of clusters exceeds 100 (150 UTLAs), hence maximum likelihood estimation delivers accurate standard errors (Bryan & Jenkins, 2016; Hox, 2010).

[68] A detailed discussion is provided in Annex A2.

[69] lr test unconditional logit vs unconditional ML logit: *p-value = 0.0000; DEFF = 7.187*; Wald statistic for estimated UTLA level variance against a chi-square distribution with 1 d.f. (taking p-value/2 since the distribution is one-sided). H_0: *UTLA level variance = 0* is rejected.

hierarchical model is supported as estimation strategy, its fitted model specification is identified following the bottom-up approach (Hox, 2010). After unconditional model estimation, individual-level variables are added to the random intercept specification, testing their significance and goodness-of-fit of the resulting model. Then, local level variables are added in a further model specification, again testing for significance and overall improvement to model specification. Afterward, random slopes are introduced and finally the specification with cross-level interactions is assessed. The different model specifications are evaluated through likelihood ratio tests[70]. Variables are centred around grand mean. It is therefore possible to get the fitted model specification which is given by the 2-level logit with random intercept (RI), whose estimation results are presented in Table 2 column (1)[71].

Results in Table 2 - column 1 indicate that the rapid growth of EU 2004 accession countries migrants is a significant risk factor for bullying victimization, as shown by the value of the odds ratio that is greater than 1. Hence, after controlling for individual and local-level cofounder, the places where the size of Eastern European residents has increased considerably after the EU04 Accession are the ones with a higher likelihood of pupils experiencing school bullying[72]. Therefore, *ceteris paribus*, sudden and sizeable upturns of migration inflows in places which are unfamiliar with the incoming group display a positive association with school violence. This finding can be related to the "*defended neighbourhood*" hypothesis since the migration shock happening at local level triggers a cultural threat perception in the incumbent population (Hangartner et al., 2018; Newman, 2013).

Literature addressing adult cohorts has already highlighted that social tension is often the consequence of a perceived challenge to the social status quo by out-groups (Card et al., 2012; Hainmueller & Hopkins, 2014). Results from the empirical estimation presented in Table 2 – column 1 confirm that young cohorts are not immune from these tensions and that cultural shock transmits to social disorders also in the school environment, as

[70]RI individual-level-predictors model vs RI unconditional model lr test: *p-value = 0.0000*; RI 2-level-predictors model vs RI individual-level-predictors model lr test: *p-value = 0.0000*.

[71] Table 2 reports results for the unweighted model specification. Results for model specifications with scaled weights are presented in Annex A2.

[72]For the Becker and Fetzer (2016) measure for the migration shock to increase, a place needs a large share of Eastern European resident population in 2011 with respect to 2001, relative to the overall European population living in the area.

speculated in the literature (Pells et al., 2016; Vertovec & Coen, 2002) and suggested by anecdotal evidence (Southern Poverty Law, 2019). Therefore, there is preliminary evidence favouring a "*defended school*" hypothesis.

Table 2: Estimated Odds ratios for two-level random intercept logit

	(1)	(2)	(3)	(4)	(5)
	2-level RI logit *robust se*	2-level RI logit *robust se no London*	2-level RI logit *robust se no 6 biggest MBs*[73]	2-level RI logit *robust se no 14 biggest MBs*[74]	2-level RI logit *robust se interactions*
Measures of association (Odds ratio)					
2004 EU Accession Migration shock	**1.237**** **[1.061,1.441]**	**1.261**** **[1.045,1.522]**	**1.264**** **[1.045,1.527]**	**1.254**** **[1.033,1.522]**	**1.225**** **[1.047,1.432]**
Belongs to Minority					0.755*** [0.721,0.790]
British White in School (sh)					0.897 [0.726,1.108]
Belongs to Minority*Brit White in School (sh)					1.230** [1.050,1.441]
Lives in Deprived Neighbourhood					1.153*** [1.116,1.191]
UTLA Income Deprivation					0.175*** [0.063,0.486]
Deprived Neighbourhood*UTLA Income Deprivation					0.555* [0.293,1.053]
Individual level controls	*YES*	*YES*	*YES*	*YES*	*YES*
UTLA level controls	*YES*	*YES*	*YES*	*YES*	*YES*
Log likelihood	-66120.259	-54567.593	-52041.291	-48191.669	-66114.119
Observations	110788	89959	85554	79062	110788
Cluster	150	118	113	105	150

Individual level controls: gender, ethnicity, deprivation of neighbourhood of residency.
UTLA level controls: looked-after children, split population, same-sex couples, crime rate, ethnic composition, population size, unemployment, rural/urban area

95% confidence intervals in brackets; p < 0.10, ** p < 0.05, *** p < 0.001; errors clustered at UTLA level*

To gauge the influence of leverage points, the fitted model specification is estimated also removing London and the other big Metropolitan Boroughs (Table 2, columns (2)-(4)). Finally, two meaningful cross-level interactions are considered[75].

[73]London, Birmingham, Leeds, Sheffield, Bradford, Manchester

[74]London, Birmingham, Leeds, Sheffield, Bradford, Manchester, Liverpool, Bristol, Kirklees, Coventry, Leicester, Wakefield, Wigan, Wirral

[75]The current debate on cross-level interactions in multilevel analysis confronts scholars arguing that multilevel estimation allows consistent and unbiased cross-level interactions when there is significant random slope for the level-1covariate included in the interaction term (*i.a.* Hox, 2010) with scholars arguing that the random slope does not need to be included (i.a. Rabe-Hesketh & Skrondal, 2012). In the paper, random slopes are never significant, therefore cross-section interactions are estimated without including the random slopes and robustness is assessed by estimating a Generalized Estimation Equation (GEE) logit model.

First, it appears reasonable that the likelihood of being bullied for a 15-year-old living in a deprived neighbourhood is not neutral to the overall level of deprivation at the local authority level. Therefore, the UTLA-level of income deprivation is considered as moderator for the effect of the level of deprivation of the neighbourhood in which the pupil resides. Second, the likelihood of being bullied for a 15-year-old belonging to a minority might not be neutral to the share of the incumbent ethnicity in the local school population, especially in the light of the emerging figures on school bullying outlining that predominantly white schools are more hostile environments for ethnic/cultural minorities (Rogers et al., 2017). Hence, the share of British white pupils at school level is considered as a moderator for the individual characteristic of belonging to a minority[76]. Through the inclusion of these interaction terms in the analysis, it is possible to check whether there are explanatory variables at the local level serving as moderators for individual level relationships, given a level of the cultural shock from migration. Estimation with these interactions is summarized in column 5 of Table 2. The findings show that both interactions are significant and that, at the same time, the effect exerted by the migration shock remains stable in magnitude, direction and significance. *Ceteris paribus*, the more income-deprived the UTLA, the more the likelihood of bullying victimization of pupils living in deprived neighbourhoods decreases, as shown by the value of the odds ratio for the interaction term which is smaller than 1. This result suggests that local poverty promotes a solidarity effect among deprived pupils when coping with a shock. Then, estimates also depict that places where the share of the British-white school population is higher than the grand-mean have higher odds of minority pupils being bullied, for a given level of the cultural shock. This finding pinpoints another channel through which the local sociocultural dimension influence bullying, alongside the "*defended school*" hypothesis and without offsetting its impact. More into details, high shares of the incumbent cultural group relative to any other cultural groups imply that the place is more culturally homogenous, hence it is more likely that undifferentiated prejudices are produced (Hopkins, 2010). Hence, for any given level of

[76] Other interactions with the migration shock have been tested, all showing non-significance. A first set of considered interactions refers to the local level and are: the interaction between the migration shock and crime; the interaction between the migration shock and local economic deprivation; the interaction between the migration shock and local ethnic diversity. A second set considers a cross level interaction between belonging to a minority and the migration shock. See Annex A2, Table A2.4

the cultural shock from migration, the lack of pre-existing acclimation with cultural diversity triggers increased hostility against everyone not aligning to the cultural dominant outlook. Results presented in Table 2 -column 6 supports this point also with respect to 15-years pupils.

Following the literature on multilevel estimation with survey data, the procedure to identify the fitted model as well as the estimations summarized in table 2 are performed for the unweighted model and for models with scaled survey weights (Rabe-Hesketh & Skrondal, 2006b). Results align, therefore only results for the unweighted model are presented in the paper[77]. Results from this first part also show that, by focusing the investigation on fine-grained spatial dimensions, it becomes possible to account for the existing geographic heterogeneity displayed by data on bullying. In fact, adding local-level covariates substantially reduces the estimate of between-UTLA variance of school bullying[78]. Hence, a large fraction of bullying spatial variance is explained by local features proxied by significant covariates. Overall, estimation results support non-spatially blind strategy where the individual and the local levels are jointly considered.

A further step in the analysis is considering estimates of the fitted model specification with interactions once leverage points are removed. Results are outlined in Table 3, columns 1-3. The effect of the interaction between the share of British white pupils and the ethnicity of each individual pupil doubles when Metropolitan Boroughs are removed.

Results in Table 3 -columns 1-4- show how odds ratio rises from 1.232 to more than 2 for the coefficient referring to the interaction[79]. This finding again aligns to existing evidence supporting the role of sociocultural features in influencing resentment. More into details, it shows that places which are more socially static and homogeneous and which are not endowed with established local policies to deal with multiculturality are characterized by stronger hostility against every culturally diverse group (Hopkins, 2010).

[77]Convergence of results among weighted and unweighted model specifications is presented in Annex A2, Tables A2.1-A2.2.

[78]The fitted model specification explains nearly the 90% of the observed spatial variance of school bullying (PVC = 89.3%). The Median Odds Ratio (MOR) is reduced by 8.56%. See Annex A2, table A2.3 for measures of variation of clustering for the fitted model specification. Best Linear Unbiased Predictors (BLUPS) for each UTLA Random Intercept in the fitted ML logit specification are specified in Annex A3, Table A3.1.

[79]More details are provided in Annex A3, Figure A3.1

Finally, it is also analysed the effect of spatial economic polarization as interaction term. The geography of spatial income polarization is assessed by the coefficient of variation of income across spatial subunits[80] for each UTLA, calculated using data on the local geography of income (Department for Communities and Local Governments of the UK Government, 2015).

Table 3: Estimated Odds ratios and variance components for two-level random intercept logit with interactions

	(1)	(2)	(3)	(4)
	2-level RI logit *robust se* *interactions* *no London*	2-level RI logit *robust se* *interactions* *no 6 biggest MBs*	2-level RI logit *robust se* *interactions* *no 14 biggest MBs*	2-level RI logit *robust se* *interactions* *income polarization*
Measures of association (Odds ratio)				
2004 EU Accession Migration shock	**1.240** [1.033,1.487]**	**1.255** [1.044,1.508]**	**1.255** [1.039,1.516]**	**1.224** [1.022,1.466]**
Belongs to Minority	0.720*** [0.685,0.757]	0.714*** [0.679,0.751]	0.726*** [0.687,0.766]	0.756*** [0.722,0.791]
British White in School (sh)	0.877 [0.624,1.234]	0.945 [0.662,1.348]	1.016 [0.694,1.486]	0.891 [0.6974,1.137]
Belongs to Minority*Brit White in School (sh)	**2.118*** [1.635,2.743]**	**2.282*** [1.759,2.959]**	**2.056*** [1.568,2.695]**	**1.254** [1.074,1.464]**
Lives in Deprived Neighbourhood	1.182*** [1.141,1,223]	1.183*** [1.142,1,226]	1.177*** [1.133,1,223]	1.1457*** [1.109,1,182]
UTLA Income Deprivation	0.130*** [0.0496,0.339]	0.131*** [0.0491,0.351]	0.149*** [0.0525,0.424]	
Deprived Neighbourhood*UTLA Income Deprivation	**0.505** [0.256,0.997]**	**0.536* [0.262,1.097]**	**0.507* [0.231,1.112]**	
UTLA Spatial Income Polarization				1.081 [0.890,1.313]
Deprived Neighbourhood*UTLA Spatial Income Polarization				**1.304** [1.064,1.598]**
Individual level controls	*YES*	*YES*	*YES*	*YES*
UTLA level controls	*YES*	*YES*	*YES*	*YES*
Log likelihood	-54547.291	-52022.415	-48179.544	-52041.291
Observations	89959	85554	79062	110788
Cluster	118	113	105	150

Individual level controls: gender, ethnicity, deprivation of neighbourhood of residency.
UTLA level controls: looked-after children, split population, same-sex couples, crime rate, ethnic composition, population size, unemployment, rural/urban area

*95% confidence intervals in brackets; * p < 0.10, ** p < 0.05, *** p < 0.001; errors clustered at UTLA level*

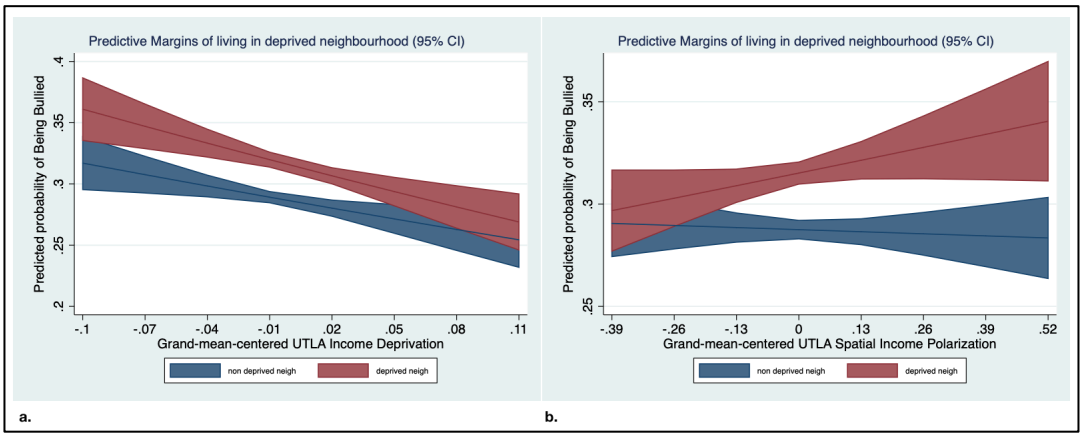
The lower the coefficient, the more neighbourhoods are homogenous in terms of income, whereas a high coefficient implies that there is a relevant difference in income level across neighbourhoods in a given UTLA. Considering the inherent high correlation

[80]Subunits are given by Lower-Layer Super Output Areas (LSOAs): small areas designed to be of a similar population size, with an average of approximately 1,500 residents or 650 households. There are 32,844 LSOAs in England (Department for Communities and Local Governments of the UK Government, 2015)

between deprivation and inequality, it is estimated a 2-level logit with RI where income spatial polarization is introduced instead of spatial income deprivation. Result are summarized in column 4 of Table 3: the more spatially unequal the UTLA in terms of income, the more school victimization increases for pupils living in deprived neighbourhoods, as outlined also in Figure 4.

Figure 4 summarizes the opposite effects exerted by income deprivation (Figure 4.a) and income inequality (Figure 4.b) on school bullying. This finding provides the first spatially robust evidence at a fine-grained scale supporting inequality over poverty as positively correlated to school violence, affecting young people from deprived neighbourhoods.

Figure 4: Moderating effect of economic deprivation and economic inequality on school bullying



3.5 Postestimation diagnostics and robustness checks.

Likelihood ratio test of estimated ML logit specifications over logit specification (Hamilton, 2013) always support ML, as well as Wald test. Support for the role of UTLA-level features in influencing bullying victimization comes from Median Odd Ratios (MOR), above the threshold value of 1 (Austin & Merlo, 2017). Wald test confirm significance of covariates (Rabe-Hesketh, 2008). The model appears not to suffer from severe multicollinearity (mean VIF = 3.85, EU 2004 migration accession VIF = 1.81)[81]. Following the bottom-up approach (Hox, 2010), other UTLA-level covariates have been assessed (juvenile crime, exposure to migration from other contexts, stock of human

[81]To further account for collinearity among control variables, estimation has been performed removing controls with VIF above 4.5 (unemployment and crime rate) and results do not change.

capital, religious outlook, adopted children, population density) showing both non-significance and no improvement in model fit. Given the central role of migration shock in the analysis, exposure to migration from other contexts is estimated both as a control variable and as alternative main variable of interest[82]. It displays non significance in either case. To control for potential concern about non-normally distributed residual errors, mis-specification and outliers, the 2-level logit with RI is estimated with White robust standard errors (Hox, 2010) and model estimations is performed removing big Metropolitan Boroughs (see Table 2, columns 2-4; Table 3, columns 2-3). Robustness is assessed also by comparing the 2-level logit with RI with a population-averaged (or marginal) logit model estimated through Generalized Estimating Equation (GEE) approach. The latter represent a competing model specification for two-level data with assumed within-cluster correlation, requiring fewer assumptions to be satisfied to deliver robust and unbiased results (Hubbard et al., 2010). In particular, the GEE marginal logit delivers consistent estimates of population-averaged effects even when dependency among individuals in clusters is not properly modelled, thus assumptions on the structure of spatial heterogeneity are less stringent. (Rabe-Hesketh & Skrondal, 2012). Results, detailed in the appendix[83], confirm the findings from the 2-level logit with RI. Noteworthy, even when relaxing the assumption on the distribution of random effects, estimated odds ratios are extremely close to the ones estimated in the 2-level logit with RI. Consistency on estimation results comparing 2-level logit with RI and GEE logit allows to conclude that the identified significant associations between risk factors, protective factors and the likelihood of bullying victimization across English UTLAs are robust to underlying assumption of dependencies of pupils within UTLA. Also, interactions have been further assessed through the GEE approach, again confirming results from the 2-level logit with RI.

[82]See Table A2.5 in Annex A2. Further checks with regards to other potentially competing features referring to the ethnic dimensions have been considered. First the local endowment of ethnic diversity is considered as a control variable in all model specifications. Results show that it behaves as a mild protective factor whose significance disappears when big metropolitan contexts are removed at the same time it does not affect the strong influence of the migration shock. Second, results do not change also when the analysis focuses on the UTLAs where the impact of the migration shock has been stronger.

[83]Estimation results for the GEE logit are presented in the Annex A3, Tables A3.2-A3.3. Model is estimated using Stata command xtgee with robust standard errors and allowing observations to be uncorrelated between UTLAs, while correlated within UTLAs.

The GEE logit specification is used also to address the issue of potential endogeneity and sorting effects of EU 04 Accession immigrants' location choice. Given the hierarchical structure of our dataset and the binary nature of the outcome variable implying model non-linearity, a Control Function approach as discussed by Petrin and Train (2010) is used. Following a two-step estimation procedure, first the potentially endogenous variable is regressed against an extra regressor in the reduced form equation alongside local-level control variables from the structural equation. Then, the reduced form residuals are plugged into the structural equation together with the endogenous explanatory variable and the other control variables (Petrin & Train, 2010; Wooldridge, 2010, 2014)[84]. Since the endogeneity analysis is pursued on EU 2004 Accession migrants, the extra regressor is framed following the well-established shift-share approach, considering the interaction between national inflows by country of origin with immigrants' past geographic distribution (Card, 2001). This instrument is consistent with the exclusion restriction, since it is more than plausible that neither the national size of Eastern Europeans migrating to England after the 2004 Accession nor the local cultural outlook more than a decade before the time in which bullying is observed have some direct effect on current school bullying. Thus, considering that the number of immigrants from EU accession countries prior to 2004 is essentially flat, their distribution at the 2001 census is considered as the baseline date (Becker & Fetzer, 2017). Formally, the extra regressor which allow the structural equation to be identified is given by

$$m_j = \sum_C \frac{M_{Cjt^0}}{M_{Ct^0}} \frac{\Delta M_C}{POP_{jt-1}} \tag{3}$$

where $\frac{M_{Cjt^0}}{M_{Ct^0}}$ identifies the share of immigrants from country of origin *C* in place *j* at reference date t^0, which precedes the date of the measurement of the endogenous regressor. ΔM_C is the country level number of new arrivals from country *C* at the date of measurement of the endogenous regressor and POP_j is the population of place *j* in the previous period. The shift-share regressor is a weighted average of the national inflow rates from each country of origin with weights given by settled distribution of immigrants. Thus, stage 1 is the OLS regression of the following linear reduced form

[84]Bootstrapping is applied to improve efficiency of estimation (Wooldridge, 2010).

$$EU04 = \pi_1 m + \pi_2 \Omega + v_2 \quad (4)$$

where $EU04$ is the size of EU 2004 Accession migration shock in 2011, is the potentially endogenous regressor, m is the shift-share regressor described above and Ω includes all the local-level control variables used in the baseline model specification. OLS regression is performed with robust standard errors, generating the estimated value for stage 1 residuals, $\hat{v}_2$. The second stage is the GLM estimation of the logit model, where $\hat{v}_2$ is included among regressors in $Pr(y_1 = 1|EU04, \Omega, v_2)$ together with cluster robust standard errors (Wooldridge, 2010). Both stages are embedded in a bootstrapping program to account for improved efficiency and robustness of standard errors (Petrin & Train, 2010). Results are summarized in Table 4, considering, respectively, the entire UTLAs set and the subset without the leverage points.

Table 4: Control Function approach: two-step estimation (bootstrapping 1000 replications)

	(1)	(2)	(3)	(4)
	CF approach: 2nd stage GLM logit bootstrap		CF approach: 2nd stage GLM logit bootstrap *without London and 13 biggest MBs*	
	OR	coef	OR	coef
2004 EU Accession Migration shock	**1.941*** **	**0.663*** **	**1.616** **	**0.480** **
		(0.187)		**(0.187)**
Belongs to Minority	0.755***	-0.281***	0.726***	-0.320***
		(0.021)		(0.026)
British white share (school)	0.888	-0.119	0.897	-0.109
		(0.096)		(0.173)
Belongs to Minority # Brit white share in school	1.220**	0.199**	2.059***	0.722***
		(0.069)		(0.163)
Lives in Deprived Neighbourhood	1.151***	0.141***	1.175***	0.161***
		(0.015)		(0.018)
Income Deprivation	0.188***	-1.674***	0.177***	-1.733***
		(0.354)		(0.423)
Deprived Neighbourhood # Income Deprivation	0.568*	-0.567*	0.527*	-0.641*
		(0.306)		(0.363)
Stage 1 residuals	0.570**	-0.561**	0.696*	-0.362*
		(0.209)		(0.219)
Individual level controls	*YES*	*YES*	*YES*	*YES*
UTLA level controls	*YES*	*YES*	*YES*	*YES*
DWH test (p-value)	0.007		0.0980	
Observations	110788	110788	79062	79062
Clusters	150	150	105	105

	CF approach: 1st stage OLS bootstrap	CF approach: 1st stage OLS bootstrap *without London and 13 biggest MBs*
	coef	coef
Shift share (Card, 2001)	-0.826***	-1.042***
	(0.254)	(0.298)
Controls (all UTLA-level covariates)	YES	YES
F-statistic	10.57***	12.24***

** p < 0.10, ** p < 0.05, *** p < 0.001; UTLA-level cluster robust standard errors in parentheses*

Results from stage 1 show that the shift-share regressor is a negative significant predictor of Eastern European spatial preference in England after 2004. This result is consistent with figures showing that the spatial distribution of EU 2004 Accession countries migrants has not followed the established settlements of Eastern European migrants reaching England before 2004 (i.a. Becker & Fetzer, 2017); in fact, the shift-share regressor exerts a negative significant influence on EU 2004 Accession migrants' location choices. The Control Function approach, while adjusting for endogeneity of migrants' location choices, confirms results from the baseline GEE specification: all covariates in Table 4 - column 1 display both sign and significance in line with the baseline GEE. The same holds when leverage points are excluded from estimation, as displayed in Table 4 -column 3. The residuals enter significantly in the second stage, as confirmed also by the Durbin-Wu-Hausman (DWH) test, highlighting that Eastern European migrants' location choices after the EU 2004 Accession are endogenous. The instrument fits well the actual distribution of Eastern Europeans across UTLAs: the first stage coefficient has a negative sign and it is highly significant, confirming the absence of an "*enclave effect*" in the Eastern European locational preferences after the 2004 Accession, as suggested in the literature (Becker & Fetzer, 2017; Pollard et al., 2008).

Furthermore, the F-statistic is above 10, meaning that our estimates do not suffer from the issue of a weak instrument[85]. The impact of the migration shock on the likelihood of bullying happening at school is now much larger and highly significant, suggesting that, among the sources of bias those delivering attenuation, measurement error may play a role, although it appears also plausible that the introduction of the instrumental variable allows to recover a Local Average Treatment Effect (LATE), which was not visible before, as estimation was focused on Average Treatment Effect (ETA) (Becker, 2016).

[85]Results are assessed also through a conditional mixed-process model with Limited Information Maximum Likelihood using the -cmp- command, to get that findings hold estimation results from the conditional mixed-process model with LIML are presented in Annex A3, Table A3.4

3.6 Pre-existing language barriers as moderator for the migration-driven cultural shock

Finally, this paper presents a preliminary investigation on the potential role of pre-existing language barriers in moderating the cultural shock from the inflow on unfamiliar migrants. Language barriers are one of the most relevant features hindering contact and the build-up of familiarity among different cultures, therefore reinforcing each group's social identity and increasing inter-group distance (Newman, Hartman, & Taber, 2012). Being that language barriers may prevent cultural integration to happen in a place, they negatively affect the assimilation of cultural diversity at local level (Newman et al., 2012).

The language barriers are considered with respect to young cohorts, hence the measure of exposure to language barriers at school level is designed using the English School Census data on secondary school pupils not speaking English as first language. The School Census figures rely on the Department of Education's definition of EAL as "*a pupil whose first language is known or believed to be other than English*". First language is further defined as "*the language in which the child was initially exposed during early development and continues to be exposed to this language in the home or in the community*" (Department for Education of the UK Government, 2011). Notably, data on the share of EAL pupils do not convey any information on the actual EAL pupils' level of English proficiency. However, the 2018 School Census has released data on the level of EAL pupils' English proficiency at UTLA level (Office for National Statistics, 2019), showing that more than 60% of EAL pupils are not English-fluent. Therefore, figures on the school population not having English as first language can still provide with a broad indication of school level exposure to language barriers. The paper introduces a measure of exposure to language barriers at school level following the relevant role assigned to changes rather than levels within sociocultural issues related to migration (*i.a.* Hainmuller & Hopkins, 2014; Newman, 2013; Hopkins, 2010). The proposed measure of school-life exposure to language barriers at school level for pupils aged 15 in 2014 is given by the percentage change of non-native English speaker pupils from 2007 to 2012[86]. Formally, the following 2-level logit model is estimated

[86]Although 15-years-old have enrolled in the school system in 2004, the School Census started collecting data on non-native English speakers in 2007. Alternatives specifications for exposure to language barriers are discussed later in the session and detailed in Annex A4.

$$Pr(y_{ij} = 1) = \alpha_0 + \alpha_{0j} + \gamma LB_j \times EU04_j + \alpha_1 x_{1ij} + \cdots + \alpha_k x_{kij} + \cdots + \beta_1 z_{1j} + \cdots + \beta_m z_{mj} \quad (5)$$

where LB_j is the proxy for exposure to language barriers at school level for UTLA *j*, and the remaining variables are the same as in eq. (1). Eq. (5) considers the interaction between exposure to school language barriers and exposure to cultural shock from migration. Being that robustness checks have highlighted the importance of accounting for the endogeneity of the migration shock, eq. (5) contains an interaction term where one of the variables is endogenous. Results from the estimation of eq. (5) through the Control Function approach (Wooldridge, 2010, ch.6 and ch.9)[87]are summarized in Table 5 -column 1. There is preliminary evidence suggesting that language barriers act as moderator for a cultural shock triggered by migration.

Table 5: Estimated Odds ratios for two-level random intercept logit specification with language barriers as moderator for cultural shock

	CF approach: 2nd stage GLM logit	
	(1)	(2)
Measures of association (Odds ratio)		
Language Barriers Exposure	0.963*	0.966*
	[0.925,1.002]	[0.928,1.006]
2004 EU Accession Migration shock	**1.988*** **	**1.873*** **
	[1.420,2.784]	**[1.300,2.700]**
Language Barriers Exposure * 2004 EU Accession Migration shock	**1.262***	**1.272***
	[0.981,1.623]	**[0.981,1.651]**
Stage 1 residuals	0.633**	0.664**
	[0.447,0.896]	[0.451,0.979]
Individual level controls	*YES*	*YES*
UTLA level controls	*YES*	*YES*
Cross level interactions	*NO*	*YES*
Log likelihood	-66125.39	-66120.75
Observations	110788	110788
Cluster	150	150

95% confidence intervals in brackets; $p < 0.10$, ** $p < 0.05$, *** $p < 0.001$;*

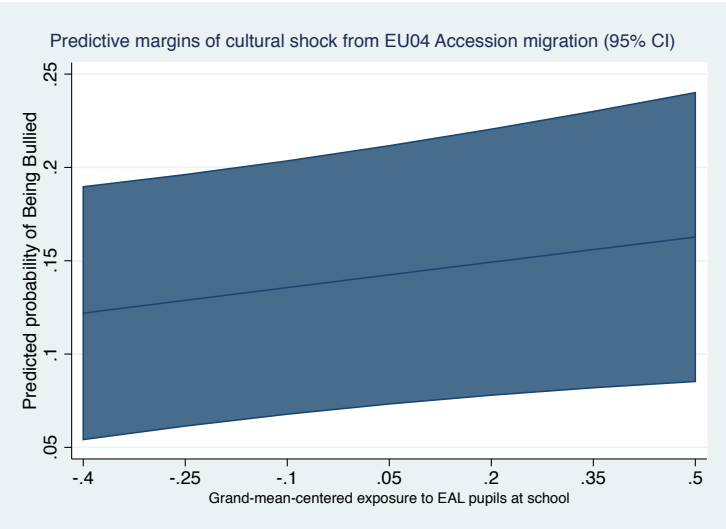
errors clustered at UTLA level

As summarized in Figure 5, the wider the percentage change in language barriers experienced at school level, the stronger the effect of the cultural shock from Eastern European migration. Findings are robust to the inclusion of the other interactions described in the paper (Table 5, column 2). Moreover, the results hold also with alternative specifications for the measure of exposure to language barriers and when

[87]Stage 1 is again the OLS regression of the linear reduced form outlined in eq.(4); therefore results from Stage 1 are the same as the ones summarized in Table 4. Residuals from estimation of eq. (4) are plugged into eq. (5) to get the results presented in Table 5.

UTLAs are restricted to a subset where Eastern Europeans pupils are not leveraging EAL students[88]. Although preliminary, this evidence supports the relationship between extant exposure to low level of assimilation of cultural diversity in places and the effect of a cultural shock determined by the arrival of a new and unfamiliar cultural group. Areas with relevant size of language barriers at school suffer from a lower assimilation of cultural diversity, which widens the effect of a cultural shock from migration. Distances among extant cultural groups due to linguistically unassimilated people hinder the local acculturation with respect to diversity. Since extant acculturation to cultural diversity eases the mitigation of a cultural shock due to newcomers, wherever this acculturation has not taken places mitigation is prevented.

Figure 5: Moderating effect of language barriers on school bullying



3.7 Conclusions

To conclude, this paper has studied the effect of a sudden and fast migration inflow on the likelihood of school violence for 15-year-old pupils, a school population segment that is more exposed to the external environment than younger pupils. Through hierarchical logit estimation, coupled with Generalized Equilibrium Estimation and Control Function approach to account for endogeneity issues, findings suggest that

[88]Alternative specifications have considered: *(i)* different time span for the exposure to language barriers during school-life; *(ii)* an index for language barriers complexity where the size of language diversity at school level is considered along with exposure to language barriers. See Annex A4 for details.

perceived threats posed by a sudden and sizeable arrival of migrants, seen as outsiders with respect to the established social identity, act as a risk factor for school violence.

The "*defended school*" hypothesis appears relevant. Young cohorts experience higher level of school violence in places where there is a cultural shock due to the arrival of migrants whose culture is not known by the incumbent population. The evidence presented in the paper confirms that sudden inflows of unfamiliar migrants represent a robust risk factor even after controlling for local socioeconomic characteristics, individual qualities and potential endogeneity of the migration shock.

The results of this paper support the idea that school violence is driven by cultural threats which are given by the broader local context (Vertovec and Coen, 2002). The findings presented here relates to the fact that cultural identity and social hierarchies do not intrude in people's life abruptly when reaching adulthood, rather they influence beliefs and behaviours through our being and acting in a given social environment. Therefore, as growing up implies being more exposed and active in the local social environment, it also determines increased influence of the environment itself on actions and perceptions.

The triggering effect exerted by migration inflows on school bullying outlined in this paper can be related to the thriving literature scrutinizing the geography of resentments (McCann, 2019; Rodríguez-Pose, 2018). The existing contributions have uncovered various channels through which local socio-economic features boost frustration and tension within and across communities: evidence supports relevant association between migration and social tension, due to the incumbent population's perceived threats to their in-group identity (Hainmueller & Hopkins, 2014), also channelled through voting decisions (i.a. Halla et al., 2017; Otto & Steinhardt, 2014). Results from this paper adds another channel, identifying school bullying as influenced by migration shocks, too.

Findings from this paper also aligns with extant contributions showing that the effects of migration on social tension are channelled by dynamic changes rather than by static levels (Becker & Fetzer, 2017; Hainmueller & Hopkins, 2014; Newman, 2013).

Another interesting finding refers to the economic dimension, since evidence from the paper depicts that income inequality rather than deprivation behaves a risk factor for bullying. This result adds a fine-grained geography evidence to existing cross-country data on the positive relationship between income inequality and bullying.

In the paper it is also assessed that language barriers among established cultural groups influence the effect of the arrival of newcomers in places. The evidence presented suggest that language barriers increase the negative effect of cultural shock from migration, hence confirming that they play a role in preventing a place to become familiar with cultural diversity per se. By harming this process, language barriers limit the local endowment of cultural diversity assimilation, thus reducing the local capacity to ease the absorption of a cultural shock brought by new migrants. This issue deserves further investigation.

Overall, the analysis presented in the paper strongly suggests that that bullying-prevention programs should consider the local context. This conclusion entails non-negligible policy implication, since it supports place-based policies as most appropriate to targeting bullying than place-neutral policies. More into details, bullying-prevention programs should consider local demographic changes when related to immigration of new cultural groups, alongside language barriers at school level. With respect to the latter, policy should aim at reducing them, so to favour the local assimilation of different cultures and improve the place's fit to cope with new incoming culture. Finally, also the level of local spatial income polarization should be addressed.

The exploratory investigation of the local risk factors associated to the first wave of the WAY survey has contributed to shed some light on the role of spatial features with respect to school bullying. At the same time highlighting the relevance of bullying-related survey designed to grasp the local level figures. Obviously, this study considers only England, therefore there are questions about generalization of results to other contexts. Moreover, due to data availability, school-level data are considered at the local level, therefore not allowing to refer to the specific school attended by pupils within the local authority. Also, results on the effect of spatial income polarization open up to further investigation on the role of spatial segregation, which is not analysed here being beyond the scope of the present work. Still, the local authority represents an interesting spatial unit, since it allows to account for spillovers from conterminous neighbourhoods within a local authority and it encompasses both school and residency. The latter feature is non-negligible when dealing with this specific age cohorts, since certain tier of schools -like secondary schools- are not located in every neighbourhood, implying that the residency neighbourhood may often not coincide with the school location. Thus, local authorities as spatial units of observation may help to alleviate some of the disadvantages of focusing

on small areas when analysing the effects of exposure to local features (Burgess & Platt, 2018).

Also, with respect to cultural variety of pupils in English schools, information is time-fragmented: the School Census data on languages spoken in schools are available for 2012 only and data about pupils' nationality/country of birth for 2017-2018 only and suspended afterwards. Thus, the measure for language barriers at school level amounts to a broad estimate of existing obstacles in communication among different cultures. Nonetheless, many research efforts are providing support for the role of language barriers in fostering cultural distance (*i.a.* Hainmueller & Hopkins, 2014), therefore further research is needed for a more refined scrutiny of their moderating role with respect to a cultural shock from migration.

References

Alesina, A., Brioschi, B., & Ferrara, E. La. (2016). Violence Against Women: A Cross-cultural Analysis for Africa. *National Bureau of Economic Research Working Paper Series*, *No. 21901*. https://doi.org/10.3386/w21901

Álvarez-García, D., García, T., & Núñez, J. C. (2015). Predictors of school bullying perpetration in adolescence: A systematic review. *Aggression and Violent Behavior*, *23*, 126–136. https://doi.org/10.1016/j.avb.2015.05.007

Ammermueller, A. (2012). Violence in European schools: A widespread phenomenon that matters for educational production. *Labour Economics*, *19*(6), 908–922. https://doi.org/10.1016/j.labeco.2012.08.010

Asparouhov, T. (2006). General Multi-Level Modeling with Sampling Weights. *Communications in Statistics - Theory and Methods*, *35*(3), 439–460. https://doi.org/10.1080/03610920500476598

Austin, P. C., & Merlo, J. (2017). Intermediate and advanced topics in multilevel logistic regression analysis. *Statistics in Medicine*, *36*(20), 3257–3277. https://doi.org/10.1002/sim.7336

Barone, G., D'Ignazio, A., de Blasio, G., & Naticchioni, P. (2016). Mr. Rossi, Mr. Hu and politics. The role of immigration in shaping natives' voting behavior. *Journal of Public Economics*, *136*(8228), 1–13. https://doi.org/10.1016/j.jpubeco.2016.03.002

Becker, S. O. (2016). Using instrumental variables to establish causality. *IZA World of Labor*, (April), 1–10. https://doi.org/10.15185/izawol.250

Becker, S. O., & Fetzer, T. (2017). Does Migration Cause Extreme Voting? *Warwick Working Paper Series*, (306).

Becker, S. O., Fetzer, T., & Novy, D. (2017). Who voted for Brexit? A comprehensive district-level analysis. *Economic Policy*, *32*(92), 601–650.

https://doi.org/10.1093/epolic/eix012

Bilger, M., & Carrieri, V. (2013). Health in the cities: When the neighborhood matters more than income. *Journal of Health Economics*, *32*(1), 1–11. https://doi.org/10.1016/j.jhealeco.2012.09.010

Bowes, L., Arseneault, L., Maughan, B., Taylor, A., Caspi, A., & Moffitt, T. E. (2009). School, neighborhood, and family factors are associated with children's bullying involvement: A nationally representative longitudinal study. *Journal of the American Academy of Child and Adolescent Psychiatry*, *48*(5), 545–553. https://doi.org/10.1097/CHI.0b013e31819cb017

Bronfenbrenner, U. (1979) *The Ecology of Human Development*. Harvard University Press

Brown, S., & Taylor, K. (2008). Bullying, education and earnings: Evidence from the National Child Development Study. *Economics of Education Review*, *27*(4), 387–401. https://doi.org/10.1016/j.econedurev.2007.03.003

Bryan, M. L., & Jenkins, S. P. (2016). Multilevel modelling of country effects: A cautionary tale. *European Sociological Review*, *32*(1), 3–22. https://doi.org/10.1093/esr/jcv059

Burgess, S., & Platt, L. (2018). *Inter-ethnic Relations of Teenagers in England's Schools: the Role of School and Neighbourhood Ethnic Composition* (Discussion Paper Series No. CPD 07/18).

Card, D. (2001). Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration. *Journal of Labor Economics*, *19*(1), 22–64. https://doi.org/10.1086/209979

Card, D., Dustmann, C., & Preston, I. (2012). Immigration, wages, and compositional amenities. *Journal of the European Economic Association*, *10*(1), 78–119. https://doi.org/10.1111/j.1542-4774.2011.01051.x

Carle, A. C. (2009). Fitting multilevel models in complex survey data with design weights: Recommendations. *BMC Medical Research Methodology*, *9*(1), 49. https://doi.org/10.1186/1471-2288-9-49

Childline. (2019). Race hate crimes against children reach 3 year high.

Commission on Social Determinants of Health of the World Health Organization. (2008). *Closing the gap in a generation: health equity through action on the social determinants of health. Final Report of the Commission on Social Determinants of Health.*

Cook, C. R., Williams, K. R., Guerra, N. G., Kim, T. E., & Sadek, S. (2010). Predictors of bullying and victimization in childhood and adolescence: A meta-analytic investigation. *School Psychology Quarterly*, *25*(2), 65–83. https://doi.org/10.1037/a0020149

Department for Communities and Local Governments of the UK Government. (2015). *The English Index for Multiple Deprivation. Guidance.* Retrieved from https://www.gov.uk/government/statistics/english-indices-of-deprivation-2015

Department for Education of the UK Government. (2011). *School Census.*

Diez Roux, A. V., & Mair, C. (2010). Neighborhoods and health. *Annals of the New York Academy of Sciences*, *1186*(1), 125–145. https://doi.org/10.1111/j.1749-6632.2009.05333.x

Due, P., Merlo, J., Harel-Fisch, Y., Damsgaard, M. T., Holstein, B. E., Hetland, J., … Lynch, J. (2009). Socioeconomic inequality in exposure to bullying during adolescence: A comparative, cross-sectional, multilevel study in 35 countries. *American Journal of Public Health*, *99*(5), 907–914. https://doi.org/10.2105/AJPH.2008.139303

Elgar, F. J., Pickett, K. E., Pickett, W., Craig, W., Molcho, M., Hurrelmann, K., & Lenzi, M. (2013). School bullying, homicide and income inequality: A cross-national pooled time series analysis. *International Journal of Public Health*, *58*(2), 237–245. https://doi.org/10.1007/s00038-012-0380-y

Enos, R. D. (2016). What the Demolition of Public Housing Teaches Us about the Impact of Racial Threat on Political Behavior. *American Journal of Political Science*, *60*(1), 123–142. https://doi.org/10.1111/ajps.12156

Espelage, D. L. (2014). Ecological Theory: Preventing Youth Bullying, Aggression, and Victimization. *Theory into Practice*, *53*(4), 257–264. https://doi.org/10.1080/00405841.2014.947216

Espelage, D. L., Van Ryzin, M. J., & Holt, M. K. (2018). Trajectories of bully perpetration across early adolescence: Static risk factors, dynamic covariates, and longitudinal outcomes. *Psychology of Violence*, *8*(2), 141–150. https://doi.org/10.1037/vio0000095

Gerstenfeld, P. B. (2017). *Hate Crimes: Causes, Controls, and Controversies*. SAGE Publications. Retrieved from https://books.google.it/books?id=v35ZDwAAQBAJ

Goodwin, M., & Milazzo, C. (2017). Taking back control? Investigating the role of immigration in the 2016 vote for Brexit. *British Journal of Politics and International Relations*, *19*(3), 450–464. https://doi.org/10.1177/1369148117710799

Guiso, L., Herrera, H., Morelli, M., & Sonno, T. (2017). Populism: Demand and Supply *, 1–63. Retrieved from http://www.heliosherrera.com/populism.pdf

Hainmueller, J., & Hopkins, D. J. (2014). Public Attitudes Toward Immigration. *Annual Review of Political Science*, *17*, 225–249. https://doi.org/10.1146/annurev-polisci-102512-194818

Halla, M., Wagner, A. F., & Zweimüller, J. (2017). Immigration and Voting for the Far Right. *Journal of the European Economic Association, Volume 15, Issue 6, 1 December 2017, Pages 1341–1385*, (6), 1341–1385.

Hamilton, L. (2013). *Statistics with Stata (8th Ed)*. Cengage.

Hangartner, D., Dinas, E., Marbach, M., Matakos, K., & Xefteris, D. (2018). Does Exposure to the Refugee Crisis Make Natives More Hostile? *American Political Science Review*, (17). https://doi.org/10.2139/ssrn.3042936

Hawkins, O., & Moses, A. (2016). Polish population of the United Kingdom, House of Commons Library Briefing paper N. CBP7660

Health and Social Care Information Centre of the UK Government. (2015a). *Health and Well Being of 15 years old in England. What About YOUth? Survey. Technical Report. UK Data Service. SN: 7894*. https://doi.org/http://doi.org/10.5255/UKDA-SN-7894-3

Health and Social Care Information Centre of the UK Government. (2015b). *Health and Well Being of 15 years old in England. What Avout YOUth? Survey. Main Findings. UK Data*

Service. SN:7894. https://doi.org/http://doi.org/10.5255/UKDA-SN-7894-3

Health and Social Care Information Centre of the UK Government. (2016). What About YOUth? Survey, 2014. [data collection]. 2nd Edition. UK Data Service. SN: 7894.

Hong, J. S., & Espelage, D. L. (2012). A review of research on bullying and peer victimization in school: An ecological system analysis. *Aggression and Violent Behavior*, *17*(4), 311–322. https://doi.org/10.1016/j.avb.2012.03.003

Hopkins, D. J. (2010). Politicized Places: Explaining Where and When Immigrants Provoke Local Opposition. *American Political Science Review*, *104*(1), 40–60. https://doi.org/10.1017/S0003055409990360

Hox, J. J. (2010). *Multilevel Analysis (2nd ed.)*. New York: Routledge.

Hubbard, A. E., Ahern, J., Fleischer, N. L., Van der Laan, M., Satariano, S. A., Jewell, N., … Satariano, W. A. (2010). To GEE or not to GEE: comparing population average and mixed models for estimating the associations between neighborhood risk factors and health. *Epidemiology*, 467–474.

Human Rights First. (2008). 2008 Hate Crime Survey.

Juvonen, J., & Graham, S. (2014). Bullying in Schools: The Power of Bullies and the Plight of Victims. *Annual Review of Psychology*, *65*(1), 159–185. https://doi.org/10.1146/annurev-psych-010213-115030

Maas, C. J., & Hox, J. (2005). Sufficient Sample Sizes for Multilevel Modeling. [References]. *J. Res. Methods Behav. Soc. Sci.*, *1*(3), 86–92. https://doi.org/10.1027/1614-1881.1.3.86

Marsh, S., & Mohdin, A. (2018). Record number of UK children excluded for racist bullying. *The Guardian*, 9–12. Retrieved from https://www.theguardian.com/education/2018/nov/30/record-number-of-uk-children-excluded-for-racist-bullying

McCann, P. (2019). Perceptions of regional inequality and the geography of discontent: insights from the UK. *Regional Studies*, 1–12. https://doi.org/10.1080/00343404.2019.1619928

Mcdevitt, J. (2014). New figures reveal dramatic increase in hate crimes against Polish people | Johnny McDevitt | Society | The Guardian. Retrieved from http://www.theguardian.com/society/2014/jun/11/polish-people-rise-in-attacks-blame-recession-politicians-media

Medina, R. M., Nicolosi, E., Brewer, S., & Linke, A. M. (2018). Geographies of organized hate in America: A regional analysis. *Annals of the American Association of Geographers*, *108*(4), 1006–1021. https://doi.org/10.1080/24694452.2017.1411247

Meleady, R., Seger, C. R., & Vermue, M. (2017). Examining the role of positive and negative intergroup contact and anti-immigrant prejudice in Brexit. *British Journal of Social Psychology*, *56*(4), 799–808. https://doi.org/10.1111/bjso.12203

Migliaccio, T., & Raskauskas, J. (2016). *Bullying as a Social Experience*. Routledge.

Müller, K., & Schwarz, C. (2018). Making America Hate Again? Twitter and Hate Crime Under Trump. *Ssrn*. https://doi.org/10.2139/ssrn.3149103

NatCen Social Research. (2014). *British Social Attitudes: the 31st Report*. Retrieved from www.bsa-3.natcen.ac.uk

National Association for Language Development in the Curriculum. (2015). *Languages in schools: More about the languages of bilingual pupils, England.*

Newman, B. J. (2013). Acculturating Contexts and Anglo Opposition to Immigration in the United States. *American Journal of Political Science*, *57*(2), 374–390. https://doi.org/10.1111/j.1540-5907.2012.00632.x

Newman, B. J., Hartman, T. K., & Taber, C. S. (2012). Foreign language exposure, cultural threat, and opposition to immigration. *Political Psychology*, *33*(5), 635–657. https://doi.org/10.1111/j.1467-9221.2012.00904.x

OECD. (2017). Students' Well-Being. *PISA 2015 Results*, *III*. https://doi.org/http://dx.doi.org/10.1787/9789264273856-en IS

Office for National Statistics. Population of the UK by Country of Birth and Nationality: 2016. Statistical Bullettin (2017).

Office for National Statistics - ONS. (2019). *International Migration and the Education Sector – What Does the Current Evidence Show?*

Otto, A. H., & Steinhardt, M. F. (2014). Immigration and election outcomes: Evidence from city districts in Hamburg. *Regional Science and Urban Economics 45, 67-79*.

Patchin, J. W., & Hinduja, S. (2011). Traditional and nontraditional bullying among youth: A test of general strain theory. *Youth and Society*, *43*(2), 727–751. https://doi.org/10.1177/0044118X10366951

Pells, K., Ogando Portela, M., & Espinoza, P. (2016). Poverty and Inequity: Multi-Country Evidence on the Structural Drivers of Bullying. In *Ending the Torment: Tackling Bullying from the Schoolyard to Cyberspace*. UN Office of the Special Representative of the Secretary-General on Violence against Children.

Peri, G. (2016). Immigrants, Productivity, and Labor Markets. *Journal of Economic Perspectives*, *30*(4), 3–30. https://doi.org/10.1257/jep.30.4.3

Petrin, A., & Train, K. (2010). A Control Function Approach to Endogeneity in Consumer Choice Models. *Journal of Marketing Research*, *47*(1), 3–13. https://doi.org/10.1509/jmkr.47.1.3

Pinotti, P. (2015). Immigration Enforcement and Crime. *American Economic Review*, *105*(5), 205–209. https://doi.org/10.1257/aer.p20151040

Pollard, N., Latorre, M., & Sriskandarajah, D. (2008). Floodgates or turnstiles. … *Flows to (and from) the UK*, (April). Retrieved from http://www.wsmp.org.uk/documents/wsmp/Migration (general) research and reports/IPPR Floodgates or Turnstiles.pdf

Przybylski, A. K., & Bowes, L. (2017). Cyberbullying and adolescent well-being in England: a population-based cross-sectional study. *The Lancet Child and Adolescent Health*, *1*(1), 19–26. https://doi.org/10.1016/S2352-4642(17)30011-1

Rabe-hesketh, S. (2008). *Multilevel and Longitudinal Modelling Using Stata*. Statacorp. Retrieved from http://books.google.com/books?id=woi7AheOWSkC

Rabe-Hesketh, S., & Skrondal, A. (2006a). *Multilevel and longitudinal modeling using Stata Third Edition: Volume I: Continuous Responses. American Statistician* (Vol. 60). https://doi.org/10.1198/tas.2006.s56

Rabe-Hesketh, S., & Skrondal, A. (2006b). Multilevel modelling of complex survey data. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *169*(4), 805–827. https://doi.org/10.1111/j.1467-985X.2006.00426.x

Rabe-Hesketh, S., & Skrondal, A. (2012). Dichotomous or binary responses. *Multilevel and Longitudinal Modeling Using Stata. Volume II: Categorical Responses, Counts, and Survival*, 501–574.

Reynolds, G. (2008). The impacts and experiences of migrant children in UK secondary schools, (47), 34.

Rodríguez-Pose, A. (2018). The revenge of the places that don't matter (and what to do about it). *Cambridge Journal of Regions, Economy and Society*, *11*(1), 189–209. https://doi.org/10.1093/cjres/rsx024

Rodrik, D. (2018). Populism and the economics of globalization. *Journal of International Business Policy*, *1*(1), 12–33. https://doi.org/10.1057/s42214-018-0001-4

Rogers, J., Franke, M., Yun, J., Ishimoto, M., Diera, C., Cooper Geller, R., … Brenes, T. (2017). *Teaching and Learning in the Age of Trump: Increasing Stress and Hostility in America's High Schools*. Retrieved from https://idea.gseis.ucla.edu/publications/teaching-and-learning-in-age-of-trump

Roodman, D. (2011). Fitting Fully Observed Recursive Mixed-process Models with cmp. *The Stata Journal: Promoting Communications on Statistics and Stata*, *11*(2), 159–206. https://doi.org/10.1177/1536867X1101100202

Rzepnikowska, A. (2019). Racism and xenophobia experienced by Polish migrants in the UK before and after Brexit vote. *Journal of Ethnic and Migration Studies*, *45*(1), 61–77. https://doi.org/10.1080/1369183X.2018.1451308

Sales, R., Ryan, L., Rodriguez, M., & Angelo, A. D. (2008). *Polish Pupils in London Schools: opportunities and challenges*.

Sime, D., Kakela, E., Corson, S., Tyrell, N., McMellon, C., Kelly, C., & Moskal, M. (2017). Eastern European Young People in Brexit Britain: Racism, Anxiety and a Precarious Future [Research and Policy Briefing No.1], (1).

Smokowski, P. R., Cotter, K. L., Robertson, C., & Guo, S. (2013). Demographic, Psychological, and School Environment Correlates of Bullying Victimization and School Hassles in Rural Youth. *Journal of Criminology*, *2013*, 1–13. https://doi.org/10.1155/2013/137583

Snjiders, T., & Bosker, R. (2012). *Multilevel Analysis: an Introduction to Basic and Applied Multilevel Analysis (2nd ed.)*. SAGE Publications Inc.

Southern Poverty Law. (2019). *Hate at School*.

Sykes, B. L., Piquero, A. R., & Gioviano, J. P. (2017). Code of the Classroom? Social Disadvantage and Bullying Among American Adolescents, U.S. 2011-2012. *Crime and Delinquency*, *63*(14), 1883–1922. https://doi.org/10.1177/0011128716641431

Tabellini, M. (2017). Gifts of the Immigrants,Woes of the Natives: Lessons from the Age of Mass Migration, *Harvard Business School BGIE Unit Working Paper No. 19-005.*

Tippett, N., & Wolke, D. (2014). Socioeconomic status and bullying: A meta-analysis. *American Journal of Public Health*, *104*(6), 48–59. https://doi.org/10.2105/AJPH.2014.301960

Tur-Prats, A. (2018). Family Types and Intimate-Partner Violence: A Historical Perspective. *The Review of Economics and Statistics*, 1–45. https://doi.org/10.1162/rest_a_00784

UNESCO. (2017). *School Violence and Bullying Global Status Report Sustainable Development Goals United Nations Educational, Scientific and Cultural Organization.* https://doi.org/10.2105/AJPH.2011.300430

UNESCO. (2018). *School and Bullying.*

van der Ende, J., Jansen, W., Jansen, P. W., Verhulst, F. C., Mieloo, C., Berkel, A. D., … Veenstra, R. (2012). Prevalence of bullying and victimization among children in early elementary school: Do family and school neighbourhood socioeconomic status matter? *BMC Public Health*, *12*(1), 1–10. https://doi.org/10.1186/1471-2458-12-494

Vertovec, S., & Coen, R. (2002). Introduction. In S. Vertovec & R. Coen (Eds.), *Conceiving Cosmopolitanism.* Oxford University Press, Inc.

Viner, R. M., Ozer, E. M., Denny, S., Marmot, M., Resnick, M., Fatusi, A., & Currie, C. (2012). Adolescence and the social determinants of health. *The Lancet*, *379*(9826), 1641–1652. https://doi.org/10.1016/S0140-6736(12)60149-4

Weale, S. (2019). Xenophobic Bullying Souring Lives of Eastern European Pupils in UK. *The Guardian.* Retrieved from https://www.theguardian.com/society/2019/aug/22/xenophobic-bullying-souring-lives-of-east-european-pupils-in-uk

Wilkinson, R. G., & Pickett, K. E. (2017). The enemy between us: The psychological and social costs of inequality. *European Journal of Social Psychology*, *47*(1), 11–24. https://doi.org/10.1002/ejsp.2275

Wooldridge, J. M. (2010). *Econometric Analysis of Cross-Section and Panel Data (2nd ed).* The MIT Press.

Wooldridge, J. M. (2014). Quasi-maximum likelihood estimation and testing for nonlinear models with endogenous explanatory variables. *Journal of Econometrics*, *182*(1), 226–234. https://doi.org/10.1016/j.jeconom.2014.04.020

Annex A1: Data description

WAY 2014 bullying-related questions. Within the WAY 2014 survey, being bullied was defined as *"…when another person, or a group of people, say or do nasty and unpleasant things to him or her. It is also bullying when a person is teased repeatedly in a way he or she does not like or when he or she is deliberately left out of things. Bullying may happen over the Internet or by text or phone messages. It is not bullying when a person is teased in a friendly and playful way*" (Health and Social Care Information Centre of the UK Government, 2015a). Given this definition of bullying, respondents were asked to rate six statements referring to traditional bullying using a 5-points response scale ranging from 0 = "*I haven't been bullied this way in the past couple of months*" 1 = "*It has happened once or twice*", 2 = "*2 or 3 times a month*", 3 = "*2 or 3 times a week*", to 4 = "*Several times a week.*" The six statements referred to traditional bullying (verbal, physical and relational) worked as follows: "*I was called mean names, was made fun of, or teased in a hurtful way*"; "I *was hit, kicked, pushed, shoved around, or locked indoors*"; "*Other people left me out of things on purpose, excluded me from their group of friends, or completely ignored me*"; "*Other people told lies or spread false rumours about me and tried to make other people dislike me*"; "*Other people made fun of me because of my body weight"*; "*Other people made sexual jokes, comments, or gestures to me*" (Health and Social Care Information Centre of the UK Government, 2015a). To obtain a measure of experience of bullying, Przybylski & Bowes (2017) use the following indicator: if the sum of the number of traditional bullying items endorsed by each participant at greater than "*2 or 3 times a month*", then the pupils is a bully-victim.

Data sources and correlation matrix

Table A1.1: Data sources

Variable	Source
Level 1 (individual) Variables	
Survey Weight	What About Youth (2016) UK National Health Service Digital (data collected in 2014)
Traditional bullying	What About Youth (2016) UK National Health Service Digital (data collected in 2014)
Belongs to non-white minority	What About Youth (2016) UK National Health Service Digital (data collected in 2014)
Male	What About Youth (2016) UK National Health Service Digital (data collected in 2014)
Lives deprived neighbourhood	What About Youth (2016) UK National Health Service Digital (data collected in 2014)
Level 2 (Local Authority) Variables	
Split population (share)	ONS 2011 Census
Same sex couples (share)	ONS 2011 Census
Ethnic diversity Index (Simpson Inverse)	ONS 2011 Census
IMD average income score	UK Department of Communities and Local Governments. English Indices of Deprivation 2011
IMD crime score	UK Department of Communities and Local Governments. English Indices of Deprivation 2011
Unemployment (share)	ONS 2011 Census
Income variation coefficient	UK Department of Communities and Local Governments. English Indices of Deprivation 2011
Children looked after (per 100k children)	2010-11 Children in Need Census (CIN). UK Department of Education
Population size (ln)	ONS 2011 Census
2004 EU Accession Migration Shock	Becker S, Fetzer T and Novy D (2017)
Exposure to migration from ROW	Becker S, Fetzer T and Novy D (2017)
Rural	2011 ONS Rural-Urban Classification (RUC11)
Secondary school: British white (share)	UK Department of Education. School Census 2011
Secondary school: percentage change EAL pupils 2012-2007	UK Department of Education. School Census 2012-2007
School: Language diversity (Simpson Index)	UK Department of Education. School Census 2012-2007

	2004 EU migration accession shock	British white share (school)	Ethnic diversity	Crime score	Pop size (ln)	Split population	Same sex couples	UTLA looked-after children	Unemployment	Income deprivation	Spatial Income polarization	EAL exposure (school)	Language diversity (school)
2004 EU migration accession shock	1												
British white share (school)	-0,4046	1											
Ethnic diversity	0,4410	-0,9161	1										
Crime score	0,4703	-0,6991	0,6171	1									
Pop size (ln)	-0,1137	0,1481	-0,1191	-0,3025	1								
Split population	-0,1649	0,6791	-0,641	-0,2766	0,0422	1							
Same sex couples	-0,1606	-0,3917	0,3236	0,2168	-0,0919	-0,0445	1						
UTLA looked-after children	-0,0851	0,4279	-0,3831	-0,1829	-0,0338	0,3752	-0,1296	1					
Unemployment	0,4049	-0,4198	0,3597	0,6911	-0,6291	-0,1405	0,0607	-0,1871	1				
Income deprivation	0,4340	-0,3558	0,3056	0,692	-0,3000	0,0283	0,1129	-0,1018	0,823	1			
Spatial Income polarization	-0,4895	0,4907	-0,5063	-0,6568	0,2539	0,0654	-0,1976	0,1063	-0,6381	-0,8075	1		
EAL exposure (school)	0,0512	0,3849	-0,3408	-0,1570	-0,0703	0,2830	-0,1750	0,4055	-0,1019	-0,1268	0,0998	1	
Language diversity (school)	-0,0554	-0.1971	0.1875	0,0756	0,1132	-0,2590	0,1747	-0,1069	--0,1133	-0,2255	0,1268	-0,0309	1

Annex A2: Rescaling survey weights and adequate integration point identification

Established literature on multilevel modelling has shown that survey weights have to be rescaled to alleviate for a series of issues leading to biased estimation (Carle, 2009; Asparouhov, 2006; Rabe-Hesketh & Skrondal, 2006). Therefore, weights provided in the WAY 2014 survey are rescaled according to strategies suggested in the literature. Formally,

$$w_{ij}^{A} = w_{ij}\left({n_j}/{\sum_i w_{ij}}\right)$$

$$w_{ij}^{B} = w_{ij}\left({\sum_i w_{ij}}/{\sum_i w_{ij}^2}\right)$$

Weights A are obtained scaling original weights so that the new weights sum to the cluster sample size; Weights B are obtained scaling original weights so that the new weights sum to the effective cluster size. Following Rabe-Hesketh, and Skrondal (2010), multilevel estimation is performed using Weight A, Weight B and no weight, respectively. Results are compared and as long as there is convergence, there is no warning for biased results. Furthermore, given the large-enough number of clusters (greater than 50), convergence of estimation results would allow to focus on unweighted multilevel specification for postestimation analysis (Carle, 2009). Another relevant estimation design issue refers to the identification of the adequate number of integration points to be used in the maximum likelihood estimation with adaptive quadrature process inherent in STATA melogit. To address this second issue, parameter estimates for the unconditional model with different numbers of integration points are compared for each weighting strategy, to check whether they remarkably change or not. Results are summarized in Table A2.1: both intercept and between-UTLA variance do not change while integration points change. Increasing the number of integration points does not deliver relevant improvements in estimates, therefore 3 integration points are chosen as default value, to speed up convergence. Main results will be nonetheless assessed against 15 integration points specification for robustness check (Table A2.2), confirming that estimation performed with 3 integration points is an appropriate strategy. Both Table A2.1 and A2.2 show that the three weighting strategies (Weight A, Weight B and

unweighted) deliver converging results. Therefore, there is reduced concern with respect to biased estimation (Carle, 2009; Rabe-Hesketh & Skrondal, 2006). Finally, postestimation diagnostics can be performed on the unweighted specification, reducing existing computational limitation with regards to weighted Multilevel Logit estimation.

Table A2.1 Estimates for different numbers of integration points for the unconditional models with different weighting strategies

a) Weight A	(1)	(2)	(3)	(4)
	coeff 3 int points	coeff 5 int points	coeff 7 int points	coeff 15 int points
Intercept	-0.878***	-0.878***	-0.878***	-0.878***
	(0.0159)	(0.0161)	(0.0161)	(0.0161)
UTLA level intercept variance	0.0309***	0.0309***	0.0309***	0.0309***
	(0.00404)	(0.00411)	(0.00411)	(0.00411)
Observations	110788	110788	110788	110788
b) Weight B				
	coeff 3 int points	coeff 5 int points	coeff 7 int points	coeff 15 int points
Intercept	-0.878***	-0.878***	-0.878***	-0.878***
	(0.0159)	(0.0161)	(0.0161)	(0.0161)
UTLA level intercept variance	0.0304***	0.0304***	0.0304***	0.0304***
	(0.00398)	(0.00404)	(0.00404)	(0.00404)
Observations	110788	110788	110788	110788
c) Unweighted				
	coeff 3 int points	coeff 5 int points	coeff 7 int points	coeff 15 int points
Intercept	-0.870***	-0.870***	-0.870***	-0.870***
	(0.0152)	(0.0152)	(0.0152)	(0.0152)
UTLA level intercept variance	0.0278***	0.0278***	0.0278***	0.0278***
	(0.00410)	(0.00410)	(0.00410)	(0.00410)
Observations	110788	110788	110788	110788

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.001$; errors are clustered at UTLA level

Table A2.2: Comparison of estimation results for the fitted model specification with 3 and 15 integration points for each weighting strategy

	Unweighted		*Weight A*		*Weight B*	
	3 int points	15 int points	3 int points	15 int points	3 int points	15 int points
Measures of association (Odds ratio)						
Individual level variables						
Belongs to non-white minority	0.737***	0.737***	0.740***	0.740***	0.740***	0.740***
	[0.710,0.765]	[0.710,0.765]	[0.707,0.776]	[0.707,0.776]	[0.706,0.775]	[0.706,0.775]
Male	0.570***	0.570***	0.570***	0.570***	0.570***	0.570***
	[0.555,0.585]	[0.555,0.585]	[0.554,0.585]	[0.554,0.585]	[0.554,0.586]	[0.554,0.586]
Lives in Deprived Neighbourhood	1.150***	1.150***	1.162***	1.162***	1.162***	1.162***
	[1.116,1.186]	[1.116,1.186]	[1.124,1.202]	[1.124,1.202]	[1.124,1.202]	[1.124,1.202]
UTLA level variables						
Income Deprivation	0.153***	0.153***	0.166***	0.166***	0.167***	0.167***
	[0.0721,0.326]	[0.0721,0.326]	[0.0608,0.451]	[0.0608,0.451]	[0.0617,0.454]	[0.0617,0.454]
2004 EU Accession Migration shock	**1.237****	**1.237****	**1.237****	**1.237****	**1.236****	**1.236****
	[1.028,1.487]	**[1.028,1.487]**	**[1.057,1.446]**	**[1.057,1.446]**	**[1.057,1.446]**	**[1.057,1.446]**
Rural area	1.051*	1.051*	1.046	1.046	1.045	1.045
	[0.998,1.106]	[0.998,1.106]	[0.987,1.108]	[0.987,1.108]	[0.987,1.108]	[0.987,1.108]
Split population	1.142***	1.142***	1.137***	1.137***	1.136***	1.136***
	[1.075,1.212]	[1.075,1.212]	[1.062,1.218]	[1.062,1.218]	[1.061,1.216]	[1.061,1.216]
Same sex couples	0.752*	0.752*	0.734**	0.734**	0.741**	0.741**
	[0.558,1.015]	[0.558,1.015]	[0.565,0.955]	[0.565,0.955]	[0.573,0.958]	[0.573,0.958]
UTLA looked after children	1.005***	1.005***	1.004***	1.004***	1.004***	1.004***
	[1.003,1.006]	[1.003,1.006]	[1.003,1.006]	[1.003,1.006]	[1.003,1.006]	[1.003,1.006]
Unemployment	1.004	1.004	1.002	1.002	1.002	1.002
	[0.991,1.018]	[0.991,1.018]	[0.985,1.019]	[0.985,1.019]	[0.985,1.019]	[0.985,1.019]
Crime score	1.008	1.008	0.994	0.994	0.995	0.995
	[0.933,1.089]	[0.933,1.089]	[0.920,1.074]	[0.920,1.074]	[0.921,1.074]	[0.921,1.074]
Ethnic diversity	0.996*	0.996*	0.996**	0.996**	0.996**	0.996**
	[0.992,1.001]	[0.992,1.001]	[0.993,1.000]	[0.993,1.000]	[0.993,1.000]	[0.993,1.000]
Pop size (ln)	1.007	1.007	1.000	1.000	0.999	0.999
	[0.965,1.050]	[0.965,1.050]	[0.950,1.052]	[0.950,1.052]	[0.950,1.052]	[0.950,1.052]
Brit white share (school)	0.871	0.871	0.895	0.895	0.896	0.896
	[0.699,1.084]	[0.699,1.084]	[0.725,1.105]	[0.725,1.105]	[0.725,1.106]	[0.725,1.106]
Measures of variation of clustering						
Intercept variance (se)	0.00300**	0.00300**	0.00345**	0.00345**	0.00316**	0.00316**
	0.0011	0.0011	0.0011	0.0011	0.0011	0.0011
Observations	110788	110788	110788	110788	110788	110788
Cluster	150	150	150	150	150	150

*95% confidence intervals in brackets * p < 0.10, ** p < 0.05, *** p < 0.00; errors are clustered at UTLA level*

Table A2.3: ML logit fitted model specifications estimation results

	(1)	(2)	(3)	(4)	(5)
Measures of association (Odds ratio)	2 level logit RI	2 level logit RI *robust se*	2 level logit RI *no London*	2 level logit RI *no 6 biggest MBs*	2 level logit RI *no14 biggest MBs*
Individual level variables					
Belongs to non-white minority	0.737***	0.737***	0.730***	0.738***	0.754***
	[0.710,0.765]	[0.704,0.772]	[0.698,0.763]	[0.704,0.773]	[0.717,0.791]
Male	0.570***	0.570***	0.569***	0.571***	0.574***
	[0.555,0.585]	[0.554,0.586]	[0.553,0.586]	[0.554,0.588]	[0.556,0.592]
Lives in Deprived Neighbourhood	1.150***	1.150***	1.177***	1.180***	1.175***
	[1.116,1.186]	[1.113,1.189]	[1.139,1.216]	[1.141,1.220]	[1.135,1.217]
UTLA level variables					
Income Deprivation	0.153***	0.153***	0.117***	0.120***	0.141***
	[0.0721,0.326]	[0.0578,0.406]	[0.051,0.273]	[0.0501,0.285]	[0.05681,0.285]
2004 EU Accession Migration shock	**1.237****	**1.237****	**1.261****	**1.264****	**1.254***
	[1.028,1.487]	**[1.061,1.441]**	**[1.011,1.572]**	**[1.008,1.584]**	**[0.988,1.591]**
Rural area	1.051*	1.051*	1.055**	1.052*	1.052*
	[0.998,1.106]	[0.993,1.112]	[1.001,1.112]	[0.998,1.110]	[0.996,1.111]
Split population	1.142***	1.142***	1.183***	1.188***	1.179***
	[1.075,1.212]	[1.070,1.218]	[1.105,1.266]	[1.110,1.273]	[1.097,1.267]
Same sex couples	0.752*	0.752**	0.856	0.830	0.824
	[0.558,1.015]	[0.584,0.969]	[0.567,1.291]	[0.547,1.261]	[0.539,1.258]
UTLA looked-after children	1.005***	1.005***	1.005***	1.005***	1.005***
	[1.003,1.006]	[1.003,1.006]	[1.003,1.007]	[1.003,1.007]	[1.003,1.007]
Unemployment	1.004	1.004	1.007	1.007	1.005
	[0.991, 1.018]	[0.998,1.021]	[0.993,1.021]	[0.992,1.023]	[0.989,1.021]
Crime score	1.008	1.008	1.004	0.987	0.984
	[0.933,1.089]	[0.933,1.089]	[0.919,1,098]	[0.901,1.081]	[0.896,1,082]
Ethnic diversity	0.996*	0.996**	0.989*	0.991	0.993
	[0.992,1.001]	[0.993,0.999]	[0.976,1.002]	[0.978,1.004]	[0.979,1.007]
Pop size (ln)	1.007	1.007	1.013	1.013	1.006
	[0.965,1.050]	[0.957,1.058]	[0.969,1.060]	[0.963,1.060]	[0.953,1.061]
British white share (school)	0.871	0.871	0.730*	0.763	0.844
	[0.699,1.084]	[0.711,1.067]	[0.504,1.057]	[0.520,1.1120]	[0.561,1.268]
Measures of variation of clustering (UTLA)					
intercept variance (se)	0.0030**	0.0030**	0.0027**	0.0027**	0.0029**
Proportional Change of Variance (PVC)	89.3%				
Likelihood ratio test against logit with FE	11.42***		8.33**	8.24 **	8.50**
ICC	0.091	0.091			
MOR	1.0533	1.0533			
Log likelihood	-66120.259	-66120.259	-54567.593	--52041.291	-48191.669
Observations	110788	110788	89959	85554	79062
Cluster	150	150	118	113	105

*95% confidence intervals in brackets; * p < 0.10, ** p < 0.05, *** p < 0.001; column 2 errors clustered at UTLA level*

Table A2.4: ML logit fitted model specifications estimation results with different interactions

	(1)	(2)	(3)	(4)
Measures of association (Odds ratio)	2 level logit RI	2 level logit RI	2 level logit RI	2 level logit RI *no 6 biggest MBs*
2004 EU Accession Migration shock	**1.216***	**1.198***	**1.231****	**1.253****
	[0.985,1.503]	**[0.989,1.451]**	**[1.040,1.458]**	**[1.051,1.494]**
Income Deprivation	0.169***			
	[0.0617,0.463]			
Income Deprivation*2004 EU Accession Migration shock	1.042			
	[0.0685,15.86]			
Crime score		1.004		
		[0.928,1.086]		
Crime score*2004 EU Accession Migration shock		1.071		
		[0.798,1.436]		
Ethnic diversity			0.998	
			[0.993,1.003]	
Ethnic diversity*2004 EU Accession Migration shock			0.999	
			[0.990,1.008]	
Pupils belonging to minority				1.145
				[0.927,1.415]
Pupils belonging to minority* 2004 EU Accession Migration shock				0.939
				[0.558,1.582]
Individual level controls	YES	YES	YES	YES
UTLA level controls	YES	YES	YES	YES
Log likelihood	-66115.776	-66114.052	-66114.105	-66118.407
Observations	110788	110788	89959	85554
Cluster	150	150	150	150

*95% confidence intervals in brackets; * p < 0.10, ** p < 0.05, *** p < 0.001; errors clustered at UTLA level*

The effect of exposure to migration from other contexts is assessed using the Becker *et al.* (2017) data on the 2011-2001 growth of immigrants from every country except the EU 04 Accession countries. This variable is considered both as control variables an as main independent variable of interest. The estimation results, presented in table A2.5, show no significant association between exposure to migration from other contexts and bullying.

Table A2.5: ML logit fitted model specifications estimation results with exposure to migrations from other contexts (ROW exposure)

	(1)	(2)
	Level 1-2 predictors *ROW exposure as control*	Level 1-2 predictors *ROW exposure as treatment*
Measures of association (Odds ratio)		
2004 EU Accession Migration shock	**1.237****	
	[1.028,1.487]	
Exposure to ROW migration 2011-2001	0.956	0.958
	[0.670,1.364]	[0.667,1.376]
Individual level controls	YES	YES
UTLA level controls	YES	YES
Log likelihood	-66120.228	-66122.732
Observations	110788	110788
Cluster	150	150

*95% confidence intervals in brackets; * p < 0.10, ** p < 0.05, *** p < 0.001; error clustered at UTLA level*

Annex A3. Postestimation diagnostics

Table A3.1: Best Linear Unbiased Predictions for each UTLA Random Intercept in the fitted ML logit specification

UTLA name	predicted RI	UTLA name	predicted RI	UTLA name	predicted RI	UTLA name	predicted RI	UTLA name	predicted RI
North West		**South West**		**London**		**Yorkshire & the Humber**		**West Midlands**	
Tameside	0,034	Devon	0,045	Richmond u.Thames	0,060	Kirklees	0,049	Telford and Wrekin	0,058
Manchester	0,029	Dorset	0,025	Tower Hamlets	0,055	E. Riding of Yorkshire	0,042	Coventry	0,032
Bolton	-0,000	Bournemouth	0,024	Harrow	0,047	Wakefield	0,035	Worcestershire	0,015
Oldham	-0,004	Wiltshire	0,023	Hounslow	0,035	Doncaster	0,032	Solihull	0,010
Wirral	-0,007	Plymouth	0,022	Redbridge	0,0315	Leeds	0,032	Walsall	0,010
Blackpool	-0,008	Torbay	0,021	Lewisham	0,031	Bradford	0,027	Shropshire	0,000
Stockport	-0,012	Gloucestershire	0,017	Hackney	0,022	Rotherham	0,026	Wolverhampton	-0,002
Blackburn & Darwen	-0,0123	North Somerset	0,015	Havering	0,022	North Lincolnshire	0,022	Staffordshire	-0,010
Trafford	-0,016	South Gloucestershire	0,012	Haringey	0,021	Kingston upon Hull	0,020	Warwickshire	-0,012
Salford	-0,016	Bristol	0,012	Croydon	0,0172	NE Lincolnshire	0,016	Herefordshire	-0,017
Wigan	-0,018	Bath & NE Somerset	0,007	Barnet	0,012	Sheffield	0,013	Sandwell	-0,019
Lancashire	-0,019	Somerset	0,005	Barking & Dagenham	0,001	Calderdale	0,013	Dudley	-0,028
Cheshire W & Chester	-0,0221	Cornwall	0,005	Newham	0,001	Barnsley	-0,008	Stoke-on-Trent	-0,029
Liverpool	-0,024	Swindon	-0,032	Westminster	0,0056	York	-0,034	Birmingham	-0,048
St. Helens	-0,027	Poole	-0,038	Merton	0,002	North Yorkshire	-0,063	**East Midlands**	
Warrington	-0,027	**South East**		Kensington & Chelsea	0,001	**East of England**		Derbyshire	0,032
Bury	-0,037	Milton Keynes	0,061	Brent	-0,001	Norfolk	0,051	Derby	0,024
Cumbria	-0,043	West Berkshire	0,057	Southwark	-0,001	Southend-on-Sea	0,048	Lincolnshire	0,023
Halton	-0,046	West Sussex	0,049	Bexley	-0,007	Essex	0,004	Rutland	0,021
Knowsley	-0,051	Portsmouth	0,034	Hammersmith &Fulham	-0,010	Thurrock	-0,001	Nottingham	0,018
Rochdale	-0,055	East Sussex	0,024	Islington	-0,010	Cambridgeshire	-0,007	Leicestershire	0,014
Sefton	-0,056	Hampshire	0,014	Wandsworth	-0,011	Suffolk	-0,008	Nottinghamshire	0,004
Cheshire East	-0,058	Brighton & Hove	0,006	Lambeth	-0,015	Luton	-0,012	Leicester	-0,013
North East		Kent	0,000	Hillingdon	-0,016	Peterborough	-0,018	Northamptonshire	-0,065
Hartlepool	0,063	Reading	-0,004	Bromley	-0,017	Hertfordshire	-0,024	Telford & Wrekin	0,058
aaaaaaaaaSunderland	0,043	Medway	-0,007	Waltham Forest	-0,024	Bedford	-0,035		
Northumberland	0,032	Oxfordshire	-0,009	Kingston upon Thames	-0,025	Central Bedfordshire	-0,055		
South Tyneside	0,023	Surrey	-0,015	Sutton	-0,026				
Gateshead	0,012	Wokingham	-0,017	Enfield	-0,033				
Redcar & Cleveland	0,003	Windsor & Maid	-0,017	Ealing	-0,038				
Stockton-on-Tees	-0,003	Slough	-0,017	Greenwich	-0,051				
Newcastle upon Tyne	-0,016	Bracknell Forest	-0,020	Camden	-0,057				
Middlesbrough	-0,022	Buckinghamshire	-0,020						
Darlington	-0,024								
North Tyneside	-0,034								
County Durham	-0,084								

Figure A3.1: Moderating effect of the share of British white 15-year-old at school

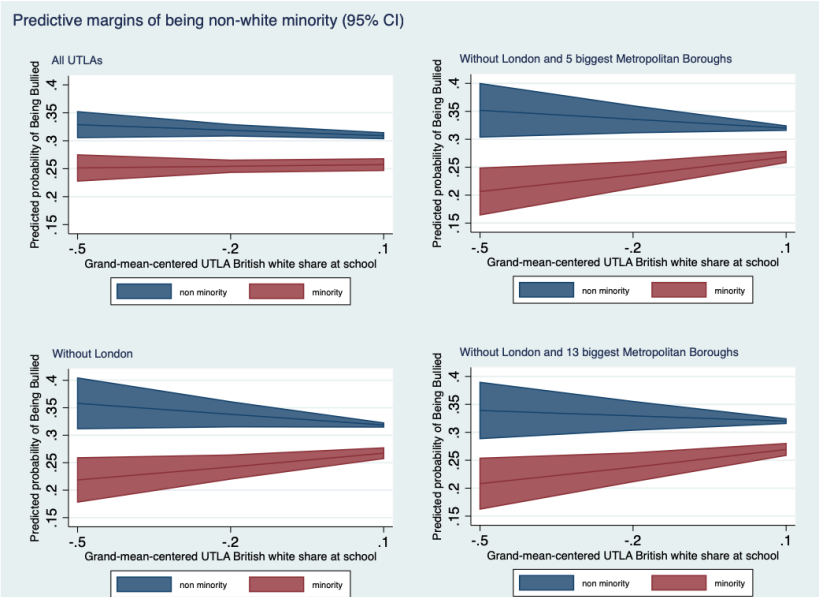


Table A3.2: GEE estimates for marginal logit for bullying victimization

	(1)	(2)	(3)	(4)
	Marginal logit (GEE)	Marginal logit (GEE) *without London*	Marginal logit (GEE) *without 6 biggest MBs*	Marginal logit (GEE) *without14 biggest MBs*
2004 EU Accession Migration shock	**1.237****	**1.259****	**1.260****	**1.251****
	[1.061,1.442]	**[1.043,1.519]**	**[1.042,1.525]**	**[1.030,1.520]**
Belongs to Minority	0.737***	0.730***	0.738***	0.754***
	[0.704,0.772]	[0.689,0.774]	[0.695,0.783]	[0.710,0.800]
Male	0.570***	0.569***	0.571***	0.574***
	[0.554,0.586]	[0.552,0.587]	[0.553,0.590]	[0.554,0.594]
Lives in Deprived Neighbourhood	1.150***	1.176***	1.180***	1.175***
	[1.113,1.189]	[1.136,1.218]	[1.138,1.223]	[1.132,1.221]
Income Deprivation	0.153***	0.120***	0.122***	0.144***
	[0.058,0.402]	[0.045,0.317]	[0.045,0.330]	[0.050,0.413]
Rural area	1.050*	1.054*	1.052*	1.052*
	[0.991,1.112]	[0.993,1.119]	[0.991,1.116]	[0.991,1.117]
Split population	1.143***	1.183***	1.188***	1.178***
	[1.073,1.218]	[1.110,1.260]	[1.113,1.268]	[1.100,1.262]
Same sex couples	0.757**	0.855	0.831*	0.825**
	[0.590,0.971]	[0.687,1.065]	[0.683,1.012]	[0.686,0.991]
UTLA looked-after children	1.005***	1.005***	1.005***	1.005***
	[1.003,1.006]	[1.003,1.007]	[1.003,1.007]	[1.003,1.007]
Unemployment	1.004	1.007	1.007	1.005
	[0.988,1.021]	[0.990,1.024]	[0.989,1.026]	[0.986,1.024]
Crime score	1.008	1.004	0.986	0.984
	[0.933,1.089]	[0.913,1.103]	[0.896,1.086]	[0.891,1.087]
Ethnic diversity	0.996**	0.989**	0.991*	0.993
	[0.993,0.999]	[0.980,0,998]	[0.982,1.000]	[0.984,1.003]
Pop size (ln)	1.008	1.014	1.013	1.005
	[0.959,1.059]	[0.963,1.068]	[0.956,1.074]	[0.984,1.071]
British white share (school)	0.871	0.732*	0.765	0.846
	[0.711,1.065]	[0.522,1.025]	[0.539,1.086]	[0.584,1.225]
Wald test (p-value)	0.0000	0.0000	0.0000	0.0000
Observations	110788	89959	85554	79062
Clusters	150	118	113	105

Exponentiated coefficients; 95% confidence intervals in brackets; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.001$; error clustered at UTLA level

Table A3.3: GEE estimates for marginal logit for bullying victimization with interaction terms

	(1)	(2)	(3)	(4)	(5)
	Marginal logit (GEE)	Marginal logit (GEE) *no London*	Marginal logit (GEE) *no 6 biggest MBs*	Marginal logit (GEE) *no 14 biggest MBs*	Marginal logit (GEE) *Spatial income polarization*
2004 EU Accession Migration shock	**1.226**** **[1.049,1.433]**	**1.239**** **[1.032,1.487]**	**1.252**** **[1.041,1.507]**	**1.253**** **[1.036,1.515]**	**1.222**** **[1.020, 1.463]**
Belongs to Minority	0.755*** [0.721,0.790]	0.720*** [0.685,0.757]	0.714*** [0.679,0.751]	0.726*** [0.687,0.766]	0.756*** [0.722,0.791]
British white share (school)	0.856 [0.697,1.052]	0.741* [0.527,1.042]	0.784 [0.552,1.113]	0.864 [0.595,1.253]	0.845 [0.664,1.075]
Belongs to Minority*Brit white share in school	1.232** [1.052,1.443]	2.117*** [1.634,2.743]	2.280*** [1.758,2.958]	2.055*** [1.567,2.694]	1.255** [1.075,1.466]
Lives in Deprived Neighbourhood	1.153*** [1.116,1.191]	1.181*** [1.141,1.223]	1.183*** [1.141,1.226]	1.176*** [1.132,1.222]	1.144*** [1.108,1.181]
UTLA Income Deprivation	0.224** [0.1079,0.634]	0.177*** [0.064,0.484]	0.175*** [0.063,0.485]	0.204** [0.699,0.595]	
Deprived Neighbourhood*UTLA Income Deprivation	0.553* [0.292,1.045]	0.503** [0.255,0.991]	0.535* [0.262,1.094]	0.507* [0.231,1.111]	
UTLA Spatial income polarization					0.962 [0.794,1.165]
Deprived Neighbourhood*UTLA Spatial Polarization					1.305** [1.066,1.598]
Individual level controls	*YES*	*YES*	*YES*	*YES*	*YES*
UTLA level controls	*YES*	*YES*	*YES*	*YES*	*YES*
Wald test (p-value)	0.0000	0.0000	0.0000	0.0000	0.0000
Observations	110788	89959	85554	79062	110788
Clusters	150	118	113	105	150

*Exponentiated coefficients; 95% confidence intervals in brackets; * p < 0.10, ** p < 0.05, *** p < 0.001; error clustered at UTLA level*

The Conditional-Mixed Process approach allows individual equations to take various forms, including binary and continuous; moreover, each equation in the system may use a different estimation technique (Roodman, 2011). It allows for endogenous variable testing through LIML estimation. Multiple endogenous variables are allowed. Therefore, endogeneity assessments performed through the Control Function approach and results are displayed in Table A3.4 below.

Table A3.4: Conditional-Mixed Process Approach: LIML

CMP LIML estimation			
Structural equation (logit)	**Odds ratio**	**Reduced form equation (OLS)**	**coef**
Belongs to Minority	0.848***	Shift Share (Card, 2001)	-0.633**
	[0.825,0.871]		(-2.35)
Male	0.714***	Income Deoprivation	0.704*
	[0.702,0.726]		(1.66)
Lives in Deprived Neighbourhood	1.088***	Rural	-0.0132
	[1.067,1.110]		(-0.79)
Income Deprivation	0.371**	Split population	0.0330
	[0.190,0.726]		(1.02)
EU 2004 migration accession	1.484**	Same sex couples	-0.539***
	[1.049,2.098]		(-3.99)
Rural area	1.030*	UTLA looked-after children	0.000394
	[0.996,1.067]		(0.52)
Split population	1.060**	Unemployment	-0.00407
	[1.005,1.119]		(-0.67)
Same sex couples	1.028	Crime score	0.0379
	[0.776,1.360]		(1.12)
UTLA looked-after children	1.003***	Ethnic diversity	0.00734**
	[1.002,1.004]		(2.66)
Unemployment	1.002	Pop size (ln)	0.00135
	[0.992,1.012]		(0.08)
Crime score	0.984	British white share (school)	0.0570
	[0.933,1.039]		(0.55)
Ethnic diversity	0.997		
	[0.994,1.001]		
Pop size (ln)	1.002		
	[0.973,1.032]		
British white share (school)	0.933		
	[0.817,1.067]		
Belongs to Minority * Brit white share in school	1.118**		
	[1.016,1.230]		
Deprived Neighbourhood * Income Deprivation	0.712*		
	[0.487,1.040]		
Observations	110788		

*95% confidence intervals in brackets; * p < 0.10, ** p < 0.05, *** p < 0.001; error clustered at UTLA level*

Table A3.5: ML logit estimates for UTLAs experiencing the strongest impact of the migration shock

	(1)
	ML logit
2004 EU Accession Migration shock	**1.334****
	[1.064,1.672]
Ethnic diversity	0.996
	[0.992,1.002]
Individual level controls	*YES*
UTLA level controls	*YES*
Observations	37801
Clusters	50

*Exponentiated coefficients; 95% confidence intervals in brackets; * p < 0.10, ** p < 0.05, *** p < 0.001; error clustered at UTLA level*

Table A3.6: Effects of the size of Eastern Europeans population on school bullying

	(1)	(2)
Measures of association (Odds ratio)	2 level logit RI	2 level logit RI
2004 EU Accession Migration shock	**1.239****	
	[1.030,1.491]	
Local size of Eastern European pop	0.00188	0.00086
	[0.000,11.74]	[1.69e-07, 4.3677]
Local size of Eastern European pop *2004 EU Accession Migration shock	2.38470e+13	
	[5.60e-28,1.02e+54]	
Individual level controls	YES	YES
UTLA level controls	YES	YES
Log likelihood	-65126.533	-65128.733
Observations	110788	110788
Cluster	150	150

*Exponentiated coefficients; 95% confidence intervals in brackets; * p < 0.10, ** p < 0.05, *** p < 0.001; error clustered at UTLA level*

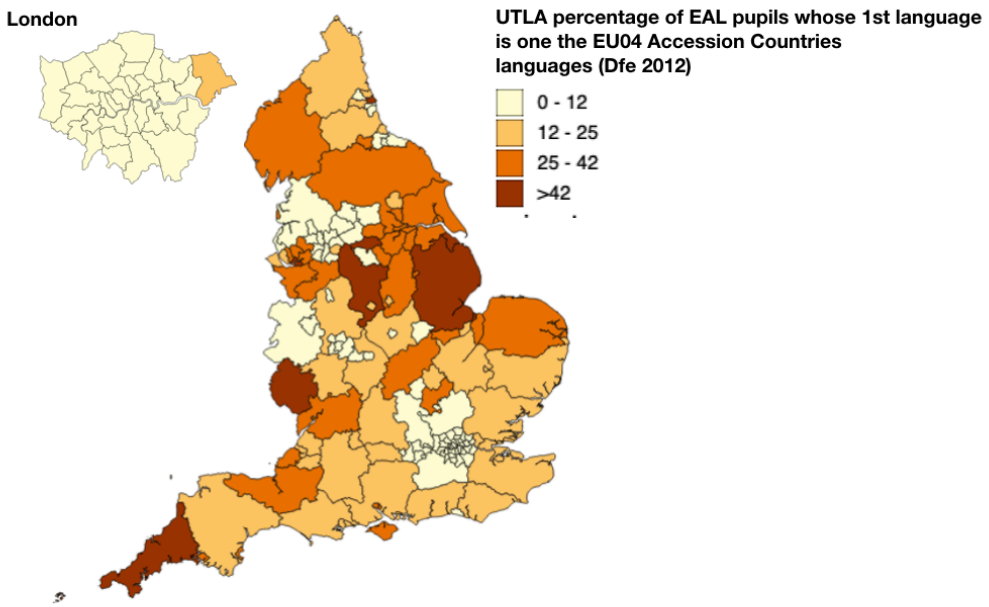
Annex A4. Language barriers as moderating variable

This section describes several robustness checks with respect to the role of language barriers as moderators for the effects of the cultural shock from migration. More into details, the role of existing language barriers in moderating the effect of the shock is assessed considering the potential leverage effect exerted by places where the language barriers may be heavily determined by Eastern European students and the measure for language barriers. Given that there is endogeneity with respect to the shock from migration, all estimations are performed using the Control Function approach, as detailed in the main paper. Hence, estimation follows a two-stage process where the endogenous regressor is estimated against the shift-share regressor given by eq. (3) in stage 1 and the resulting residuals are plugged in the second stage estimation of eq. (5), following Wooldridge (2010, ch. 6 and 9).

Since the object of this investigation is to assess whether existing language barriers among cultural groups act as moderator for the cultural shock triggered by the arrival of a new cultural group, language barriers have to be analysed also considering in which areas they are not leveraged by EAL students belonging to EU 04 Accession Countries nationalities. The 2011 School Census has detailed the first language spoken by EAL pupils at UTLA level (DfE school census data, reported by National Association for Language Development in the Curriculum, 2015), allowing the identification of UTLAs where the EU 04 young population amounts to the highest share of pupils not proficient in English. These UTLAs are summarized in Figure A4.1 and in Table A4.1. Therefore, in testing the robustness of the moderating effect of language barriers on the cultural shock from migration on school violence, here eq. (4) and (5) are estimated both on the

entire set of UTLAs and on the subset of 91 UTLAs where EAL students are not leveraged by Eastern European[89]. In these 91 UTLAs, the large majority of EAL students speaks Panjabi, Urdu, Bengali, Gurajati which have been the most common non-English home languages also before the EU Enlargement (National Association for Language Development in the Curriculum, 2015). The subset of 91 UTLAs is still characterized by a consistent geography of the EU 04 Accession migration shock, having a mean value for the Accession shock equal to 0,179, whereas the mean on the entire set of UTLAs is 0,171. This feature is due to the fact that a non-negligible part of EU 04 Accession migration is without dependant children, being high shares of migrants single and unmarried while others have left their children in the home country (Hawkins & Moses, 2016; Pollard et al., 2008).

Figure A4.1: Geography of England UTLAs with respect to Eastern European pupils as share of pupils not having English as 1st language



[89]For robustness, estimation has been also performed by including and excluding Barnet, Oxfordshire, Havering, Peterborough and South Tyneside in the subset of 91 UTLAs, since although they have Eastern Europeans as the bigger share among EAL students, their share is really close to the second largest. Results do not change.

Table A4.1: UTLAs where the share of Eastern European pupils having English as additional language is the largest fraction among non-native English speakers at school (DfE school census data, reported by NALDIC, 2015b)

UTLACODE	UTLA name	share of EE04 on EAL	2nd nationality on EAL	Share of 2nd nationality on EAL
E09000003	Barnet	7,80431728	*Gujarati*	7,73406
E10000025	Oxfordshire	13,0552392	*Punjabi*	11,77391
E09000016	Havering	13,377193	*Yoruba*	10,41666
E10000012	Essex	14,7019153	*Bengali*	6,51470
E10000032	West Sussex	16,2463037	*Urdu*	8,97547
E10000011	East Sussex	16,6584767	*Portuguese*	8,64864
E06000057	Northumberland	17,8502879	*Punjabi*	12,47601
E06000025	South Gloucestershire	18,4659091	*Punjabi*	7,8125
E06000037	West Berkshire	18,8405797	*Portuguese*	8,36120
E06000033	Southend-on-Sea	19,4244604	*Urdu*	8,00351
E10000016	Kent	20,1303523	*Nepali*	9,1659
E06000014	York	20,3773585	*Turkish*	7,54717
E06000022	Bath and North East Somerset	20,5752212	*Malayalam*	7,07964
E06000054	Wiltshire	21,3550136	*Nepali*	8,02168
E08000037	Gateshead	21,5053763	*Urdu*	11,02151
E10000003	Cambridgeshire	21,6176471	*Bengali*	7,37763
E10000009	Dorset	23,7687366	*Bengali*	11,1349
E06000029	Poole	24,0963855	*Malayalam*	14,60843
E06000035	Medway	24,204947	*Punjabi*	9,2226
E06000028	Bournemouth	24,2491657	*Portuguese*	14,12680
E06000047	Durham	24,6355685	*Punjabi*	6,26822
E10000034	Worcestershire	24,7793858	*Urdu*	18,67278
E10000029	Suffolk	25,1387347	*Portuguese*	11,15427
E10000008	Devon	25,529661	*Arabic*	5,98516
E06000045	Southampton	26,5675801	*Punjabi*	13,63214
E10000021	Northamptonshire	26,8835176	*Bengali*	8,92260
E06000027	Torbay	27,1708683	*German*	11,48459
E06000056	Central Bedfordshire	27,6884602	*Chinese*	6,47059
E06000031	Peterborough	28,2472505	*Punjabi*	24,44661
E08000011	Knowsley	28,5714286	*Yoruba*	23,80952
E06000046	Isle of Wight	29,3617021	*Bengali*	9,36170
E08000036	Wakefield	29,7770004	*Punjabi*	24,1364
E10000013	Gloucestershire	30,4153793	*Gurajati*	10,23000
E06000026	Plymouth	30,9387755	*Arabic*	7,91836
E10000024	Nottinghamshire	30,9876933	*Urdu*	11,42316
E06000050	Cheshire West and Chester	31,0810811	*Bengali*	6,75676
E08000010	Wigan	32,9761905	*Chinese*	8,09524
E06000007	Warrington	33,8129496	*Urdu*	12,8297
E08000013	St. Helens	34,7222222	*Chinese*	13,8889
E10000020	Norfolk	34,9309912	*Portuguese*	19,64868
E06000024	North Somerset	35,1973684	*Bengali*	9,53947
E10000023	North Yorkshire	35,8431861	*Punjabi*	9,52084
E06000010	Kingston Upon Hull, City of	36,3054778	*Arabic*	7,75690
E08000017	Doncaster	37,1414914	*Urdu*	12,57170
E10000006	Cumbria	37,3853211	*Chinese*	9,63302
E10000027	Somerset	39,6663079	*Portuguese*	11,19483
E06000009	Blackpool	40,1769912	*Bengali*	9,38053
E06000013	North Lincolnshire	40,4135338	*Bengali*	23,1203
E06000049	Cheshire East	41,6405761	*Urdu*	5,82341
E06000011	East Riding of Yorkshire	41,9047619	*Russian*	6,34920
E06000005	Darlington	42,081448	*Punjabi*	10, 40724
E06000052	Cornwall	45,8144796	*Portuguese*	11,19909

E08000016	Barnsley	45,8745875	*Shona*	5,94059
E06000012	North East Lincolnshire	47,2906404	*Bengali*	17,2413
E06000006	Halton	51,6853933	*Chinese*	11,2359
E10000019	Lincolnshire	60,4275144	*Portuguese*	9,07097
E10000007	Derbyshire	61,5384615	*Punjabi*	12,08791
E06000019	Herefordshire	64,3835616	*Malayalam*	8,41487
E08000023	South Tyneside	100	-	-

Table A4.2 summarizes results from the robustness check estimations.

Table A4.2: Estimated Odds ratios for the 2nd stage of the Control Function Approach with language barriers as moderator for cultural shock

	Control Function Stage 2: GLM logit				
	(1)	(2)	(3)	(4)	(5)
	Language barriers 2012-2007	*Language barriers 2013-2007*	*Language barriers 2013-2007*	*Language diversity*	*Language diversity*
	subset UTLAs	*all UTLAs*	*subset UTLAs*	*all UTLAs*	*subset UTLAs*
Measures of association (Odds ratio)					
Language Barriers Exposure (2012-2007)	0.959 [0.893,1.030]				
Language Barriers Exposure (2013-2007)		0.967* [0.929,1.005]	0.950* [0.883,1.022]		
Language Barriers Complexity				0.962 [0.919,1.008]	0.953 [0.77,1.036]
2004 EU Accession Migration shock	**1.516** [1.036,2.219]**	**1.887*** [1.305,2.728]**	**1.324* [0.952,1.841]**	**1.762*** [1.305,2.378]**	**1.643** [1.218,2.217]**
Lang. Barriers Exposure ('12-'07) * 2004 EU Accession Migration shock	**1.378** [1.004,1.892]**				
Lang. Barriers Exposure ('13-'07) * 2004 EU Accession Migration shock		**1.252* [0.993,1.579]**	**1.512** [1.139,2.008]**		
Language Barriers Complexity * 2004 EU Accession Migration shock				**1.258** [1.006,1.666]**	**1.336* [0.953,1.871]**
Stage 1 Residuals	0.850 [0.605,1.195]	0.663** [0.449,0.980]	0.980 [0.741,1.212]		0.772 [0.581,1.077]
Individual level controls	*YES*	*YES*	*YES*	*YES*	*YES*
UTLA level controls	*YES*	*YES*	*YES*	*YES*	*YES*
Cross level interactions	*NO*	*NO*	*NO*	*NO*	*NO*
Log Likelihood	-38109.97	-66120.45	-38109.91	-66120.15	-38109.68
Observations	64949	110788	64949	110788	64949
Cluster	91	150	91	150	91

Individual level controls: gender, ethnicity, deprivation of neighbourhood of residency.
UTLA level controls: looked-after children, split population, same-sex couples, crime rate, ethnic composition, population size, unemployment, rural/urban area

*95% confidence intervals in brackets; * p < 0.10, ** p < 0.05, *** p < 0.001; errors clustered at UTLA level*

Column 1 shows the estimation results from the baseline model specification presented in the main paper when the set of UTLAs is restricted to the subset where

Eastern European students do not amount to the largest share of EAL pupils. Then, to further assess the robustness of language barriers at school level in moderating the effect of a cultural shock from migration an alternative time span for exposure to language barriers is considered, given by the percentage change of non-native English speaker pupils from 2007 to 2013. Results are presented in columns 2 and 3 with respect to the entire set of UTLAs and the subset which is not leveraged by Eastern Europeans EAL students.

Then, an index for language barriers complexity at school level is introduced. This index considers both the exposure to language barriers (given by the percentage change during school-life) and the size of language diversity. The latter is measured calculating the Simpson's index of Diversity with respect to the mother-tongue of pupils at school level in each UTLA on the 2012 School Census Data (DfE school census data, reported by National Association for Language Development in the Curriculum, 2015); formally the index for language barriers complexity for UTLA *j*-th, LC_j, is given by

$$LC_j = (1 - D)_j \times (\frac{EAL_{12}}{EAL_7})_j \qquad \text{(A4.1)}$$

where $(1 - D)_j$ is the Simpson's index of Diversity for the mother-tongue languages different from English spoken at school level in UTLA *j*-th which is multiplied by the percentage change in the size of pupils not having English as 1st language in secondary school in UTLA *j*-th. Eq (A4.1) shows that language complexity will increase due to an increase in language diversity for a given level of exposure to language barriers as well as due to an increase in the level of exposure to language barriers, given a level of language diversity. Results are presented in columns 4 and 5 of table A4.2.

The direct effect of the EU04 migration shock is still significant in all model specifications and the moderating role of language barriers hold, overall increasing the effect of the shock on the likelihood of bullying. Results do not change by adding the cross-level interaction between individual deprivation and UTLA-level deprivation and the one between individual ethnic-group and UTLA school-level share of British white pupils.

The endogeneity of Eastern Europeans location choices does not appear to be significant when the UTLAs characterized by a high share of Eastern European students are removed, as shown by the non-significance of the residuals from stage 1 in columns

1, 3 and 5. Hence, estimation of these three specifications is performed again removing residuals from stage 2 (Wooldridge, 2010, ch. 9) to get that increased significance of results. A possible explanation for the fact that endogeneity does not appear to be an issue when UTLAs with high shares of Eastern European pupils are removed could reside either in the choice of the instrumental variable or in some socioeconomic explanation. To this respect, the instrument does not appear to be weak, as displayed by the F-statistics greater than 10, and it also largely acknowledged as particularly suitable in addressing migration. From the socioeconomic perspective, data show that the restricted set of UTLAs is still characterized by a consistent geography of the EU 04 Accession migration shock, having a mean value for the Accession shock equal to 0,179, whereas the mean on the entire set of UTLAs is 0,171. Hence, even in these UTLAs there is still a sizeable inflow of EU04 migrants. However, these migrants are mainly without dependent children, since the share of Eastern Europeans at school is way lower than the share of other cultural minorities. Empirical evidence from Table A4.1 is showing that location choices for these types of Eastern Europeans is no affected by the geography of those Eastern Europeans who settled in England before the Accession. Therefore, it may be the case than endogeneity of Eastern Europeans' settling preferences after the EU04 Accession is relevant only with respect to some characteristics of the migrants themselves. This issue deserves further investigation, which goes beyond the scope of the present paper.

CHAPTER 4

Coming Out of the Woods. Do nearby support services influence the propensity to report sexual violence?

Abstract. This paper presents an empirical investigation of the effect of a local policy for the provision of specialized support services for victims of sexual crimes on women' propensity to report sexual offences to the police. Applying the synthetic control method to the UK, the investigation outlined in the paper provides evidence supporting a positive effect of the local policy. The positive effect of the policy holds even in the occurrence a of high-profile media campaign about sexual offences, which is anecdotally related to a countrywide increase in the propensity to report. Hence, prominent media coverages of sexual offences do not appear to work as substitute for the local provision of VAWG support services with respect to the UK.

Keywords: women, gender violence, austerity, policy evaluation, synthetic control, local policy, comparative case study

4.1 Introduction

Nowadays sexual assaults remain extremely frequent hence Violence Against Women and Girls (VAWG) support services[90] are needed to help victims to find dedicated support, to break the feelings of isolation and embarrassment and to make the first move towards seeking justice (Council of Europe, 2011).

Within European countries, one in 10 women has experienced some form of sexual violence since the age of 15 and one in 20 women has been raped. (Fundamental Right Agency, 2014).The scarcity of nearby VAWG support services affects propensity to report because it prevents many victims from any reachable safe place capable of

[90] VAWG support services include rape crisis centers, sexual assault referral centers, refuges, independent domestic violence advisors, independent sexual violence advisors, domestic violence outreach projects, services for ethnic minority women, support for trafficked women and women in prostitution.

providing with tailored help to cope with the harm, name violence and seek justice (Bonnewit & DeSantis, 2016; Fundamental Rights Agency, 2014). In the absence of these services the feeling of isolation arising as a consequence of embarrassment and/or fear of reprisals may represent a barrier too high to overcome for victims, reducing also reporting to the police (Bonnewit & DeSantis, 2016; European Commission -DG Justice and Consumers, 2017; Imkaan, 2013). Given their role in reducing victims' perception of isolation, VAWG services should be easily reachable, hence they need to be geographically even distributed (All-Party Parliamentary Group on Domestic and Sexual Violence Inquiry, 2015; Council of Europe, 2011).

The acknowledged role of nearby VAWG services in supporting victims is not preventing them from disappearing in many places due to the budget cuts imposed by austerity. Since VAWG services heavily rely on public spending, the budget reductions imposed by the austerity programmes have hit heavily their spatial distribution (Towers & Walby, 2012; Women's Aid, 2015; Women's Budget Group, 2019). Within this widespread trend, it is nonetheless possible to detect exceptions. In fact, there are places that have decided to prioritize their VAWG support services within their local budgets even in times of austerity, actually realizing local interventions aimed at providing VAWG support services.

Therefore, it is interesting to empirically analyse whether the places that have decided to foster the provision of local VAWG support services during austerity have actually influenced the victims' willingness to report compared to places where local VAWG support services have been decommissioned.

This paper provides an examination of these effects by evaluating a place where a local programme supporting VAWG services has been realized during austerity against a comparable spatial unit where the same policy has not been implemented. This comparative case study aims at detecting whether the introduction of the local policy is related to meaningful differences in the victims' propensity to report. This objective is pursued by comparing the propensity to report sexual crimes between the place in which the local programme has been adopted and the control unit where the local programme has not been adopted. The comparison is performed between places which were similar before the introduction of the local policy (Abadie, Diamond, & Hainmueller, 2015).

Similarity has to be fulfilled in terms of the propensity to report sexual offences to the police before the introduction of the local policy and in terms the same endowment of local VAWG support services. The comparative case study analysis in the paper follows the synthetic control method (Abadie, Diamond, & Hainmueller, 2010a). Hence, the place which has implemented a local policy for the provision of VAWG support services is assessed against a synthetic counterfactual which simulates what the path for women's willingness to report sexual crimes would be if the place did not undergo the local policy.

The paper considers the local policies supporting VAWG services in England and Wales, due to two relevant reasons. First, in England and Wales there is no national-level specification of the mandatory requirements for the local provision of VAWG support services. As a consequence, the spatial distribution of VAWG support services depends on local choices(Coy, Kelly, & Foord, 2009). Second, the austerity programme in 2010 has determined a considerable change in the policies covering VAWG services, altering the way they are funded and commissioned. The austerity programme has introduced a structural shift from funding through grant aid from the national level to the local commissioning, where the scope, scale and nature of the service are specified by the local funder. The devolution from the central to the local government of both the decision making about resource allocation and the commissioning of VAWG services has made them entirely dependent on the implementation of a local policy of support[91]. While stressing the local commissioning, the austerity programme has also imposed relevant general cuts to public budget spending (Hastings, Bailey, Bramley, Gannon, & Watkins, 2015). Therefore, local authorities had to decide whether or not to realize local policies for VAWG support services considering their tight budget decisions about which services to either prioritize or cut (Hirst & Rinne, 2012). Several reports have highlighted the sizeable impact of the severe budgets cuts on the local availability of VAWG support services (All-Party Parliamentary Group on Domestic and Sexual Violence Inquiry, 2018; Towers & Walby, 2012; Women's Budget Group, 2019). The critical outlook characterizing England and Wales is also stated in several recommendations by the

[91] Before the introduction of the austerity progamme and the devolution to local authorities, the majority of funding for VAWG support services was provided through the national level, whereas after the introduction of the austerity programme the majority of funding is provided by the local level (Coy et al., 2009; Hirst & Rinne, 2012)

United Nations Committee on the Elimination of Discrimination against Women (CEDAW), which has recommended the UK to mitigate the impact of austerity measures on VAWG specialized services and to adopt legislative and comprehensive policy measures to protect women from violence across the whole country (United Nations, 2013b, 2019).

The precarious situation of the VAWG support service sector after the beginning of the austerity programme has also prompted several analyses aimed at collecting local evidence about the effect of austerity on the supply of specialized services for female victims of sexual violence. Academic researchers, journalists, activists and representatives of women organizations have collected data about the changing landscape of VAWG service provision across England and Wales (Grierson, 2018; The Bureau of Investigative Journalism, 2017; Towers & Walby, 2012). The investigation carried on in this paper starts from gathering this scattered evidence to allow the build-up of a comprehensive and sound database on the geography of the programmes on VAWG services carried on by local authorities before and after the beginning of the austerity programme. This outlook conveys a snapshot of the distribution of local policies supporting the provision of nearby VAWG services, allowing to gauge their spatial distribution before austerity and the geography of local interventions following the beginning of austerity.

Out of this comprehensive mapping, Brighton and Hove emerges as the place where a locally funded programme for VAWG support services has been pursued during austerity. Hence, the comparative case study analyses the effects of the local policy pursued by Brighton and Hove on the local propensity to report sexual crimes to the police, comparing it with a synthetic comparable unit where the same policy has not been applied.

The results from the synthetic control method suggest that the introduction of the local policy supporting VAWG services has exerted a positive effect of the propensity to report sexual assaults in Brighton and Hove compared to the synthetic control unit. Noteworthy, the positive effect of the local policy holds also after the high-profile media coverage of sexual offences given by "*Operation Yewtree*"[92] happened. The latter is

[92]"*Operation Yewtree*" is a police investigation on sexual abuse allegations which occurred in the UK at the end of 2012 and which has received a countrywide prominent media coverage. More details are provided in the next section.

considered as a trigger in women's propensity to report sexual offences countrywide (ONS, 2018b). Moreover, the results still hold when the estimation includes socioeconomic local level features which may contribute to influencing the propensity to report by victims of sexual assaults.

The remaining of the paper is organized as follows. First a background description of the effect of VAWG support services in supporting victims is provided, together with the recent history of the provision of VAWG support services in England and Wales. Then, the proposed empirical approach is explained with respect to the research question and the characteristics of the database design are specified. Afterwards, results are presented and discussed and, finally, preliminary conclusions are detailed alongside further research steps to improve the strength of findings.

4.2 Background

VAWG support services provide confidential and dedicated support including safe locations to stay, legal advice and psychological help. They do not require the victims to report offences to the police; nonetheless, they exert a positive effect in this regard by helping victims understanding what they have been through and by dismantling possible self-inflicted blames (Boba & Lilley, 2009; Imkaan, 2013; Towers & Walby, 2012; World Bank, 2014). Qualitative evidence with respect to the UK highlights that they exert a positive effect in helping women to name violence (Krishnan & Bewley, 2014) and the same insight is further supported by evidence at the European level showing that the availability of VAWG support services increases the propensity of victims to leave a violent partner by 9% (Bettio & Ticci, 2017; European Commission -DG Justice and Consumers, 2017). Hence, these services help victims to build the awareness that they deserve justice for what they have experienced, lowering cultural and actual barriers to police reporting. VAWG support services need to be evenly distributed across places, as their spatial proximity to victims is a fundamental aspect for their capability of reducing barriers (Bonnewit & DeSantis, 2016; Coy et al., 2009). The absence of nearby support services imposes to victims the additional cost of searching and travelling to find support, hence increasing their sense of isolation (Imkaan, 2013). More practically, VAWG support services like refuges provide viable answer to the housing problem which represent a relevant barriers for many women who lack the sufficient level of economic

independency to leave a violent relationship (Bettio & Ticci, 2017). Extant literature shows that female victims, especially poor ones, needs a set of reachable services capable of meeting their needs in terms of food, shelter and psychological assistance. The absence of the nearby provision of these tangible services is likely to prevent the victim from both escaping and reporting the violent situation (Baker, Billhardt, Warren, Rollins, & Glass, 2010). Evidence on the UK confirms the relevance of nearby services capable of responding to these needs in influencing the victims' propensity to escape a violent situation and to start a path to seek justice (Fahmy, Williamson, & Pantazis, 2016). A recent survey on the UK has highlighted that 28% of female victims encountering barriers to the fulfilment of the need for a safe accommodation give up the attempt of leaving the violent situation (Miles & Smith, 2019).

As a consequence, the geographic distribution on VAWG support services is identified as a pivotal element in determining women's propensity to come forward (Council of Europe, 2011; European Parliament, 2014). The local availability of VAWG support services also implies a community-level effort to counter violence against women, therefore signalling to victims a local commitment to a culture of respect of women rights, which can contribute to reducing the perception of socio-cultural prejudices (Ellsberg et al., 2015; World Bank, 2014). The fundamental role of VAWG specialized support services and the need for them to be evenly geographically distributed has also been advocated by a conspicuous corpus of political and legal initiatives by several international and national bodies, like the United Nations, the European Union and the Council of Europe (Council of Europe, 2011; European Parliament, 2014; United Nations, 2012).

With respect to VAWG support services, England and Wales are characterized by two main elements: *(i)* a structural shift in the policy framework to provide VAWG support services due to the introduction of the austerity programme; *(ii)* an enduring history of unequal geography of VAWG support services, dating before the kick-off of the austerity programme in 2010.

By looking at the 2009 Map of Gaps Report, realized by the Equality and Human Rights Commission together with End Violence Against Women, it appears that one over four local authorities did not have any specialized service before the start of the austerity (Coy et al., 2009). The report labelled the situation of victims of sexual violence

as "*a postcode lottery*", due to the critical impact of the place of residency on the availability of support. The spatial distribution of VAWG support service described in the 2009 Map of Gaps Report was linked to a policy framework in which the funding of VAWG support services was mainly realized through national grant aid benefitting both the statutory services and the non-statutory services (Heady, Kail, & Yeowart, 2011; Hirst & Rinne, 2012). Local bodies were not subject to any mandatory commitment with respect to the provision of VAWG support services[93] (Coy et al., 2009).

The role of local bodies in the provision of VAWG support services received a strong emphasis in the 2010-2015 Government strategy to counter VAWG: "*A Call to End Violence Against Women and Girls. 2010-2015*"[94], which followed the austerity program (All-Party Parliamentary Group on Domestic and Sexual Violence Inquiry, 2015; Home Office, 2011). The new prominent role assigned to local authorities implied also the devolution of the funding responsibilities of the VAWG specialized services to local budgets. Notably, the strong emphasis in the local authority commissioning of VAWG support services was again not matched with any mandatory duty for local authorities with respect to either provision and/or minimum funding for VAWG support service and it was also intertwined with the severe cuts which were imposed on local authorities' budgets (All-Party Parliamentary Group on Domestic and Sexual Violence Inquiry, 2018; United Nations, 2013a; Women's Budget Group, 2019). Overall, both the assessment of needs and the policy responses were devolved to the local level, while cutting the national budget previously dedicated to VAWG services. The Government strategy implied a national-level commitment to almost £40 millions of ring-fenced funding between 2011 and 2015 from the Home Office and Ministry of Justice. However, these ring-fenced resources served different relevant issues relating VAWG, prioritising national-level programmes and those services that are activated only after a victim actually reports to the police[95] (All-Party Parliamentary Group on Domestic and Sexual Violence Inquiry,

[93] The requisite for local governments to invest in sexual violence support services was identified by Public Service Agreement 23 -Make Communities Safer-, although without any mandatory requirements of service provision for local bodies

[94] The strategy was followed by an action plan and a progress review in 2011. The action plan was first updated in 2012 and again in 2013 and 2014.

[95] More than half of the ring-fenced resources have been devoted to the Police and the Criminal Justice system, for services dedicated to police-reporting victims (All-Party Parliamentary Group on Domestic and Sexual Violence Inquiry, 2015).

2015). Support to local commissioning had to share the remaining national resources with other purposes, such as national-level campaigns, international schemes, projects with universities and schools, health-related issues[96], children abuse and youth gang (Home Office, 2014). Consequently, national funds for the local level of VAWG support services were heavily reduced[97], shifting the funding dimension on the local level. Devolution implied that local authorities were attributed with the discretionary power to identify whether they wanted to provide local VAWG support services. Whenever this was the case, the same local authority had to support the policy with its own budget, in a context where local budgets were experiencing cuts of resources due to austerity. As a consequence, the overall geography of VAWG support services experienced a severe reduction in service provision (All-Party Parliamentary Group on Domestic and Sexual Violence Inquiry, 2018). Across England, specialist refuges have being decommissioned and replaced by general homelessness units where users with different needs apply for the same bed (Imkaan, 2013; Women's Aid, 2013). In some cases, there has been no replacement for decommissioned refuges (All-Party Parliamentary Group on Domestic and Sexual Violence Inquiry, 2015; Imkaan, 2013). Just one year after the take-off of the 2011-2015 government strategy to counter VAWG, the sexual violence and domestic violence sector was suffering from cuts in funding from local authorities (Towers & Walby, 2012). Looking at the general picture, figures define a severe contraction in services and increasing numbers of victims unable to access the required help:

- more than 75% of England's local authorities slashed their spending on VAWG services between 2011 and 2017 (Women's Budget Group, 2019);
- 17% of women's refuges were forced to close between 2010 and 2014 (Women's Budget Group, 2019);

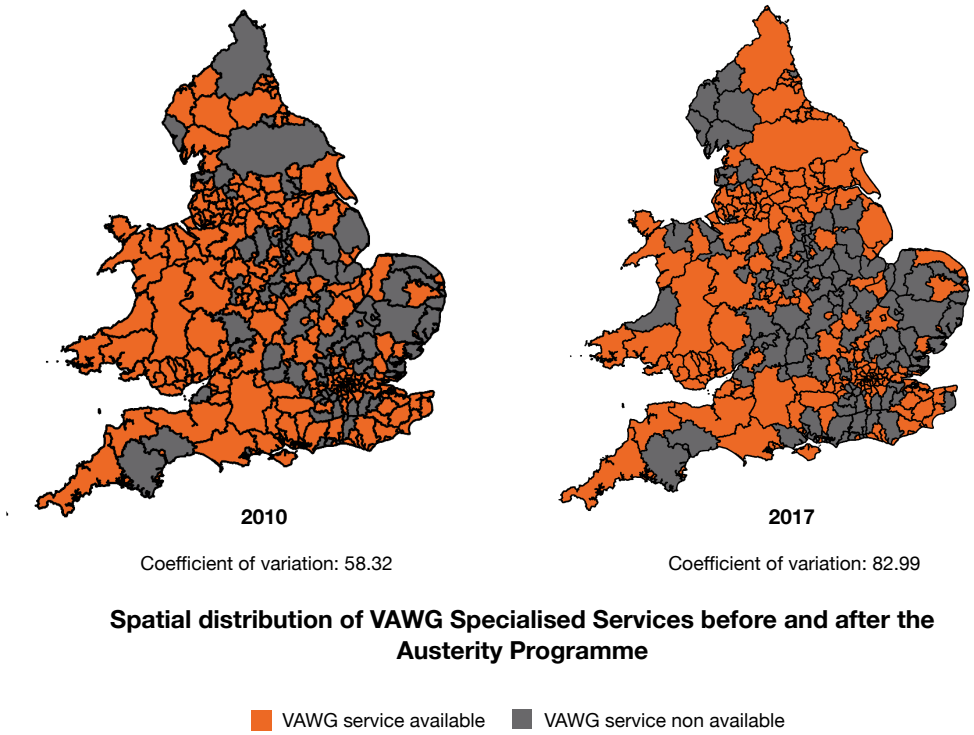
[96] A significant share of funding is dedicated to Sexual Assault Referral Centers (SARC), whose focus is getting forensic evidence, creating a case for prosecution and helping victims to cooperate with the criminal justice system. Hence, SARCs are extremely helpful in supporting victims, but only after that they have reported to police offence. So, they do not exert a relevant role in lowering barriers to police reporting (All-Party Parliamentary Group on Domestic and Sexual Violence Inquiry, 2018)

[97] Even though the UK Government provided around £4 million year between 2011 and 2015 to fund around 80 rape support centres across England and Wales (UK Government, 2015), there has been an overall the reduction in VAWG service provision determined by cuts to other public budgets (All-Party Parliamentary Group on Domestic and Sexual Violence Inquiry, 2015).

- nearly 50% of VAWG support services run by the voluntary sector were operating at least one activity without any funding or on a reduced budget already after two year from the beginning of the austerity program (Women's Aid, 2013);
- an increasing number of VAWG support services has substituted trained staff with voluntary unpaid work, resulting in growing service discontinuity and in refusal of victims whose needs require professional skills (Turgoose, 2016; Women's Aid, 2013, 2014, 2015);
- around 25%-30% of women seeking escape in refuges had been turned away between 2010 and 2016 (Women's Aid, 2013, 2014, 2015; Women's Budget Group, 2019);
- 47% of specialist Black and Minority Ethnic-led organisations reporting a significant loss of funding in after the kick-off of the austerity programme (All-Party Parliamentary Group on Domestic and Sexual Violence Inquiry, 2015).

Austerity cuts have increased the unevenness of the spatial distribution of VAWG support services (All-Party Parliamentary Group on Domestic and Sexual Violence Inquiry, 2018; Women's Budget Group, 2019), making it strongly dependent on each local body's decision on its budget spending (Towers & Walby, 2012). So, the "*postcode lottery*" of VAWG support services is still at work notwithstanding the implementation of the cross-governmental strategy to tackle violence against women and girls "*A Call to End Violence Against Women and Girls. 2011-2015*". This spatial heterogeneity has been acknowledged also by the First Report of Session 2017-2019 of the Joint Committee on the Draft of the Domestic Abuse Bill (House of Lords House of Commons Joint Committee on the Draft Domestic Abuse, 2019). Figure 1 shows the spatial distribution of VAWG support services in 2009, before the beginning of the austerity programme, and in 2017. The data underlying the maps (Coy et al., 2009; Holly, 2017; Women's Aid, 2019) confirm a 18% net reduction in the number of local VAWG support services between 2009 and 2017. Moreover, the coefficient of variation of the spatial distribution of VAWG services has increased by nearly the 25% after 10 years of austerity, supporting a sizeable increase in the spatial dispersion of local VAWG support sevices.

Figure 1: The Spatial Distribution of VAWG Support Services before and after the Austerity Programme



The enduring "*postcode lottery*" of VAWG support services is affected by the intertwined nexus between the austerity cuts to public funding and the subsequent level of commitment of local authorities in prioritising VAWG services among their local policies, given the severe reduction on the overall stream of Government grants.

Upper-Tier Local Authorities (UTLAs) are the public bodies in charge of assessing local needs of VAWG support services and of providing the necessary funding to implement them, in case they decide to realise them[98]. Funding of this type of local policy is mainly determined by the allocation of resources from the UTLA controllable income[99]. Controllable income is mainly funded by Government grants, council tax, sales, fees and charges, trading and investment income. Noteworthy, Government grants

[98]UTLAs are shire county councils, containing shire districts, 'single tier' London boroughs, metropolitan districts and unitary authorities, which are statutory responsible for social care and education services, alongside all other local services (National Audit Office, 2018). There are 174 UTLAs in England and Wales.

[99] Reserves have proven to be particularly relevant in the austerity period, since local authorities can use their reserve to provide a resource to cushion the impact of unexpected events or emergencies in prioritised areas (National Audit Office, 2018). However, differently from controllable income, they do not constitute an enduring source of funding for services.

constitute a relevant share of the funding (National Audit Office, 2018). Hence, the reduction in Government grants entailed a policy decision by each UTLA on which local services had to bear the consequences of reduced resources. At the same time, there was no mandatory requirement on a general budget cut across the whole menu of local services. Thus, each UTLA was still capable of discretionary choosing the local policies which were considered as a priority and to allocate resources to back them. In the case of VAWG services, prioritisation implied assessing a local need for increased/improved service provision through the implementation of a local policy. To this regard, UTLA councils across England and Wales made different decisions, with the majority of UTLAs deciding not to prioritise the implementation of local policies for VAWG support services (Women's Budget Group, 2019). Notably, it is not only the local statutory VAWG support services depending on the implementation of a dedicated local policy, but also the local non-statutory services, whose activities mainly rely on local authority funding[100] (Heady et al., 2011; Women's Aid, 2015).

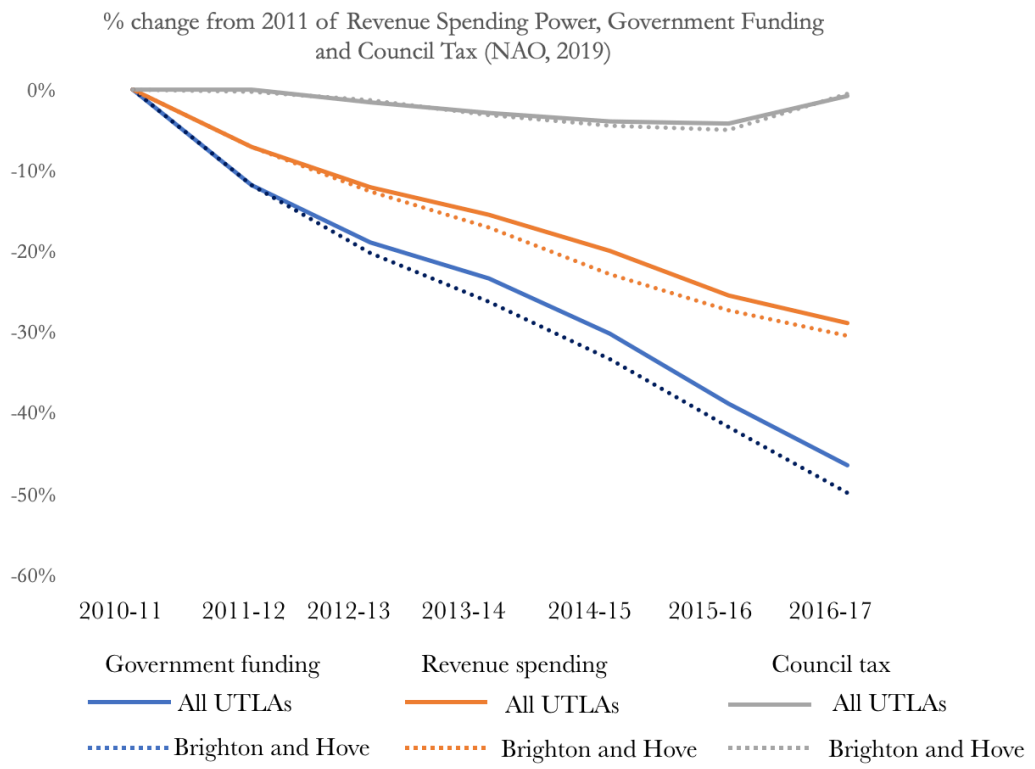
Brighton and Hove is a Unitary Authority of Sussex, hence it is appointed for the policy choice of implementing VAWG support services. Evidence on local authorities choices on the realization of local interventions to provide VAWG support services in time of austerity shows that Brighton and Hove pursued the implementation of these service (The Bureau of Investigative Journalism, 2017). At the same time, figures also show that Brighton and Hove belongs to the share of UTLAs which has been more severely hit by the austerity programme (National Audit Office, 2019), due to a higher reduction in government funding compared to the remaining UTLAs, as summarized also in Figure 2 below. Hence, the policy choice of providing VAWG support services was not done benefitting from a milder impact of the austerity programme with respect to the UTLAs that decided to decommission. By looking more into details in the Brighton and Hove Council budget in the years of the austerity programme, figures outline that the discretionary policy choices made by the Council reduced some services more than the average across UTLAs -such as highway and transportations and culture (National Audit Office, 2019). Hence, the Council rebated the reduction in revenue

[100] On average, more than 50% of funding for non-statutory organizations providing VAWG support services comes from the local authority (Women's Aid, 2015)

spending power on some services, at the same time deciding to devote spending for the implementation of a policy providing VAWG support services.

Given the policy choice of the Council of Brighton and Hove, the majority of the others UTLAs decided to decommission VAWG support services. This spatial heterogeneity in the discretionary choice of implementing a local policy for the provision of VAWG support services across England and Wales gives room for the analysis of the impact of the actual policy adoption on the propensity to report sexual crimes to the police by female victims. Besides, it also allows to gauge the impact of this type of local policy with respect to other emerging interventions which are anecdotally related to an increase in the willingness to report by victims, namely communication and media events.

Figure 2. Trends in Brighton and Hove and the other UTLAs Revenue Spending Power, Government Funding and Council Tax from the beginning of the austerity programme



It is currently conjectured that high-profile media coverage of sexual offences exerts a positive influence on women's willingness to report sexual offences to the police (Mendes, Ringrose, & Keller, 2018; ONS, 2018a). Communication and media campaigns facilitate victims in understanding that their own history of sexual violence is part of a broader structural social problem, rather than an individual experience that arose from

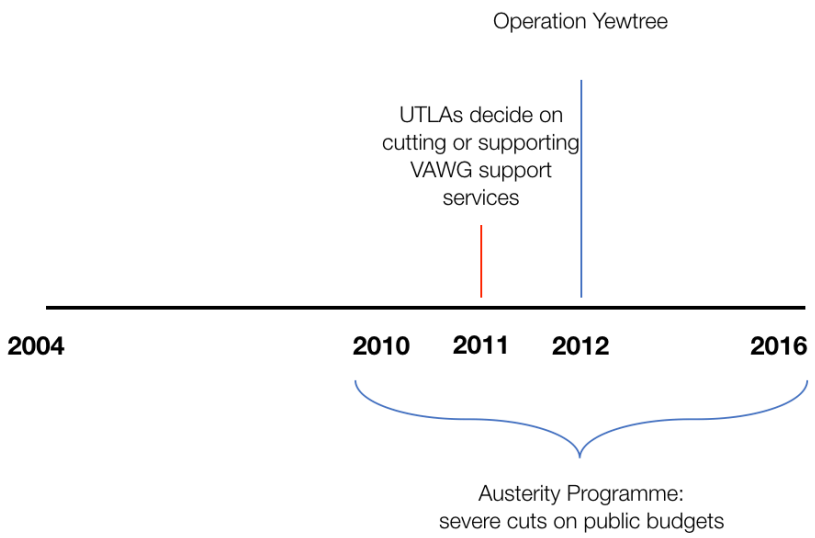
encounters with '*bad men*'. Thus, high-profile reportage by the media may contribute to reducing the feeling of isolation perceived by victims, helping them to come forward and seek justice. Moreover, their potential effect is not spatially limited, due to the national and/or global reach of the media coverage. The influence of high-impact media exposure of sexual offences is particularly important with respect to the analysis performed in this paper, since a relevant media campaign happened in the UK at the end of 2012, more less after two years from the beginning of the austerity programme. Starting from an ITV programme which featured five women recounting being abuse by the late television presenter and charity fund-raiser Jimmy Savile during the 1970s, the Metropolitan Police Service (MPS) was requested by the Association of Chief Police Officers to take the lead in assessing and scoping the claims made in the programme. The MPS investigation - labelled "*Operation Yewtree*" started by collating all the investigation against Savile to subsequently broad to include other perpetrators due to hundreds of victims coming forward to report sexual assaults perpetrated by other people as a results of the media coverage of "*Operation Yewtree*" (NSPCC, 2013). So, the high-profile coverage of "*Operation Yewtree*" influenced victims' willingness to come forward to report both recent and non-recent offences countrywide (ONS, 2018a). The latter feature has been labelled "*Yewtree effect*" (Ford, 2013) and it relates to the analogous expression "*Weinstein effect*" (Maddaus, 2018), introduced in the public discourse after the global media reach of the #*Metoo* movement.

Nonetheless, although contributing to reducing barriers from embarrassment and shame, communication events do not provide victims with reachable safe place to escape and nearby tailored support. Consequently, whether safe places and support are locally available or not may still remain an element capable of affecting the propensity to report. To this respect, the synthetic control analysis allows to gauge some evidence on whether high-media coverage of sexual offences acts as a substitute for the provision of nearby VAWG support services in influencing victims' propensity to come forward or whether the actual need for specialized protection still represents a pivotal element notwithstanding any positive media effect.

Figure 3 outlines the timeline of the events which have been described thus far. In April 2010 the UK Government announced the austerity programme, determining cuts in public grants to UTLAs starting from the 2011 budget. In November 2010, the 2010-

2015 Government strategy to counter VAWG - "*A Call to End Violence Against Women and Girls. 2010-2015*"- was applied increasing the role of discretionary commissioning of UTLAs on VAWG services. In 2011, UTLAs identified which local policies to pursue given the reduction of available resources. A big share of UTLAs practiced decommissioning of VAWG support services (Women's Budget Group, 2019). Brighton and Hove implemented a local policy for the provision of VAWG support services. Between the end of 2012 and the beginning of 2013, "*Operation Yewtree*" received a prominent media coverage, which is supposed to have stimulated many victims of sexual offences to come forward and report to the police both current and past offences countrywide (ONS, 2018c).

Figure 3: Timeline of the relevant event for the evaluation of the effect of local policies providing nearby VAWG support service



4.3 Methodology

The paper aims at measuring the effects of the policy intervention undertaken by the Council of Brighton and Hove to provide VAWG support services in time of austerity. The effects of this policy are estimated by means of a comparative case study, where Brighton and Hove is the unit exposed to the policy intervention. This exposed (treated) unit will be compared to an unexposed unit (untreated) designed through the synthetic control method (Abadie et al., 2010a). The unexposed unit is a synthetic version of Brighton and Hove, designed to be as similar to Brighton and Hove as possible with the

exception of the introduction of the local policy for the provision of VAWG specialized services. Hence, the evolution of silence breaking with respect to sexual assault for Brighton and Hove is estimated and compared with the evolution of silence breaking with respect to sexual assault for a synthetic counterfactual which simulates Brighton and Hove in the absence of the policy intervention.

The synthetic control method exploits a panel regression of outcomes on covariates (excluding treatment), and a binary variable indicating the treatment status of individual observations is specified. Then, an optimization process gives weights to individual control observations, such that the trends in covariates and outcomes of the synthetic control match those of the treated unit prior to treatment as closely as possible. Application of these weights to the control observations during the treatment period allows for a synthetic control, or counterfactual, that can be compared with the actual trend of treatment (Peri & Yasenov, 2016).

The synthetic control method hinges on the premise that a combination of comparison units (the synthetic control) often does a better job of reproducing the characteristics of the unit of interest than any single comparison unit alone (Abadie et al., 2010a). Motivated by this consideration, the comparison unit in the synthetic control method is selected as the weighted average of all potential comparison units that best resembles the characteristics of the case of interest (Abadie et al., 2015). These units compose the donor pool of comparison units.

The design of the donor pool must consider several steps detailed in the literature (Abadie et al., 2015). First, it is important to restrict the donor pool to units with characteristics similar to the treated unit. Then, units affected by the intervention of interest or by events of a similar nature should be excluded from the donor pool. Finally, units that may have suffered large idiosyncratic shocks to the outcome of interest during the study period should also be excluded if such shocks would have not affected the treated unit in the absence of the treatment.

4.4 Data description

Measuring the propensity to report sexual offences at local level. In England and Wales sexual offences are measured through two main sources: the Crime Survey for England and Wales (CSEW) and sexual offences reported to and recorded by the police (ONS,

2018b). The CSEW is particularly reliable in providing estimates on the prevalence of the volume of sexual assaults at country level, since the survey methodology allows to grasp offences that are not reported to the police, although with data that are not robust at fine-grained spatial scale. Instead, police recorded crime figures are suitable to gauge victims' demands on the police in relation to sexual offences and they are collected at the local level (ONS, 2018b). Hence, police recorded crime figures are a suitable proxy to convey a measure for the local propensity to report sexual crimes. Community Safety Partnerships (CSP) [101] is the smallest spatial units on which data on records of sexual crimes[102] on women above 16-year-old are available. There are 315 CSPs in England and Wales and they constitute the spatial units for the empirical investigation in the present paper. The policy decision about the implementation of local VAWG specialized services is attributed to the 174 Upper-Tier Local Authorities, which often coincide with the CSPs[103].

Control variables. Through the review of the relevant literature, it is assessed which local level features may contribute in influencing the propensity to report by victims of sexual assaults. These features are considered since they may be potential cofounding factors. The geography of unemployment and of the wage of women are considered, due to their role in determining women's socioeconomic independence and the consequent level of socioeconomic barriers they face when deciding to report a sexual assault (Aizer, 2010; Anderberg et al., 2016; European Commission -DG Justice and Consumers, 2017). Alongside, the share of female population is considered. Data sources are detailed in Annex A1.

4.5 Sample description

The synthetic control method needs a proper selection of the sample, especially with regard to the potential comparison units which are used to design the synthetic counterfactual.

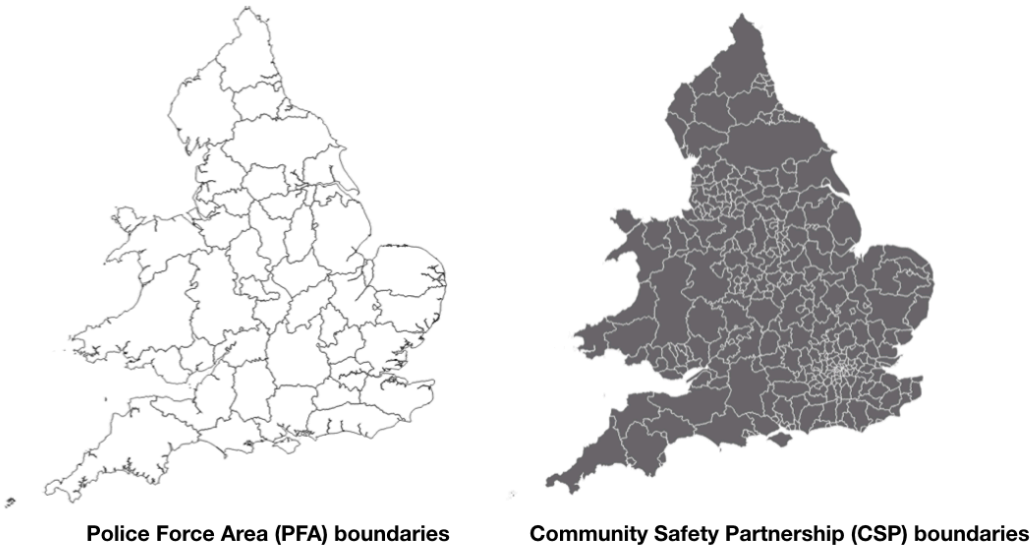
[101]CSPs are statutory bodies set up under Section 5-7 of the Crime& Disorder Act 1988.

[102]For this investigation, data on sexual crimes refers to rape and sexual assault against a female victim older that 16.

[103] For Unitary Authorities, Metropolitan Boroughs and London Boroughs, the CSP coincide with the UTLA. Only for Non-Metropolitan Districts, the corresponding UTLA (which is the County) includes several CSPs. In these cases, the policy decision taken at the UTLA level is applied to every CSPs belonging to the UTLA.

The first step is to curb the donor pool to units with characteristics similar to the treated unit. A first assessment on the suitable sample of comparable units for the synthetic control method is prompted by the variable used to proxy the propensity to report sexual offences. Recall that the outcome of interest is measured using figures on police records of sexual crimes on women. Hence, it is needed for these police records to accurately match reports by victims. Whenever this is not the case, it means that the local police force has inaccurately classified and/or dismissed a relevant share of reports of sexual offences. Areas where there is evidence of this inaccuracy by the police force have to be discarded, since they are not endowed with a reliable proxy to measure the willingness to report by female victims. To gauge the accuracy of the local police records of sexual offences against women, evidence from official audits on the integrity of crime data are used. These data are available at local Police Force Areas (PFAs) level. PFAs are the local units in which policing for England and Wales is subdivided; there are 43 local Police Force Areas (PFAs) in England and Wales and each CSP belongs to a PFA, as summarized by Figure 4.

Figure 4: The geography of Police Force Areas and Community Safety Partnerships in England and Wales



In 2013, Her Majesty's Inspectorate of Constabulary and Fire & Rescue Services (HMICFRS) started an independent inspection program to evaluate the accuracy of crime data recording by the 43 local PFAs in England and Wales, including a focus on sexual assault and rape (Her Majesty's Inspectorate of Constabulary, 2014). HMICFRS

examined the integrity of crime data for each police force, releasing force-specific report, which have been subsequently updated with follows-up between 2016 and 2019. Reviewing these reports, it is possible to identify 14 PFAs which displays high level of accuracy in recording (Her Majesty's Inspectorate of Constabulary, 2019, 2018, 2017, 2016, 2014). They are outlined in Figure 5a[104] and they include Sussex, the PFA in which Brighton and Hove falls. The analysis will discard from the sample all CSPs not belonging to these 14 PFAs, giving a remaining sample of 125 CSPs.

Another important element to consider in the construction of the donor pool is given by the local endowment of VAWG support services before the beginning of the austerity programme. Recall that VAWG support services before austerity were mainly financed through national grants rather than with local resources, hence a local decision of decommissioning after the devolution meant cutting their funding and, consequently, making them partially or entirely close. This further screening allows to identify places sharing similar characteristics before the structural change in the public policy framework due to the austerity programme (Abadie et al., 2015). More into details, it is assumed that the opportunity cost of starting a local policy is different between local authorities which did not have to bear the start-up costs for the build-up of new facilities and local authorities which had to evaluate also to bear the cost of facilities build-up. The Map of Gaps Report outlines that Brighton and Hove was endowed with VAWG support services before austerity (Coy et al., 2009). The same report also maps which other CSPs had a local endowment of VAWG support services even before 2010, as shown by Figure 5b. The CSPs not endowed with VAWG support services before the beginning of austerity are discarded[105].

Finally, it is important to exclude from the donor pool those spatial units where the same or a similar policy has been implemented. Hence, a map of local policy for the provision of VAWG support services is needed. The discretionary choice of implementing and funding a local policy for the provision of specialized VAWG support services started in 2011, after the introduction of the austerity programme in 2010. Up

[104]The PFAs are: Cambridgeshire, Cumbria, Dorset, Dyfed-Powys, Durham, Gwent, Lancashire, Lincolnshire, Merseyside, Metropolitan Police, Norfolk, Staffordshire, South Wales and Sussex. The City of London PFA is consolidated with the Metropolitan Police PFA.

[105]Notably, none of the discarded UTLAs implemented a local policy for the provision of VAWG support services during austerity.

to now and to the best of my knowledge, a comprehensive map of the local decisions on whether or not to implement a local policy supporting VAWG services after the kick-off of austerity is not available. However, several actors have collected data on this respect through interviews to VAWG service practitioners and FOIA requests[106] to the local authorities with respect to their local policy initiatives on the provision of VAWG support services from 2011 to 2016 (Grierson, 2018; The Bureau of Investigative Journalism, 2017; Towers & Walby, 2012). Through the collection and the analysis of these different data sources, it is possible to create an initial outlook of the discretionary policy choices made at the local level with respect to the local provision of VAWG services between 2011 and 2016.

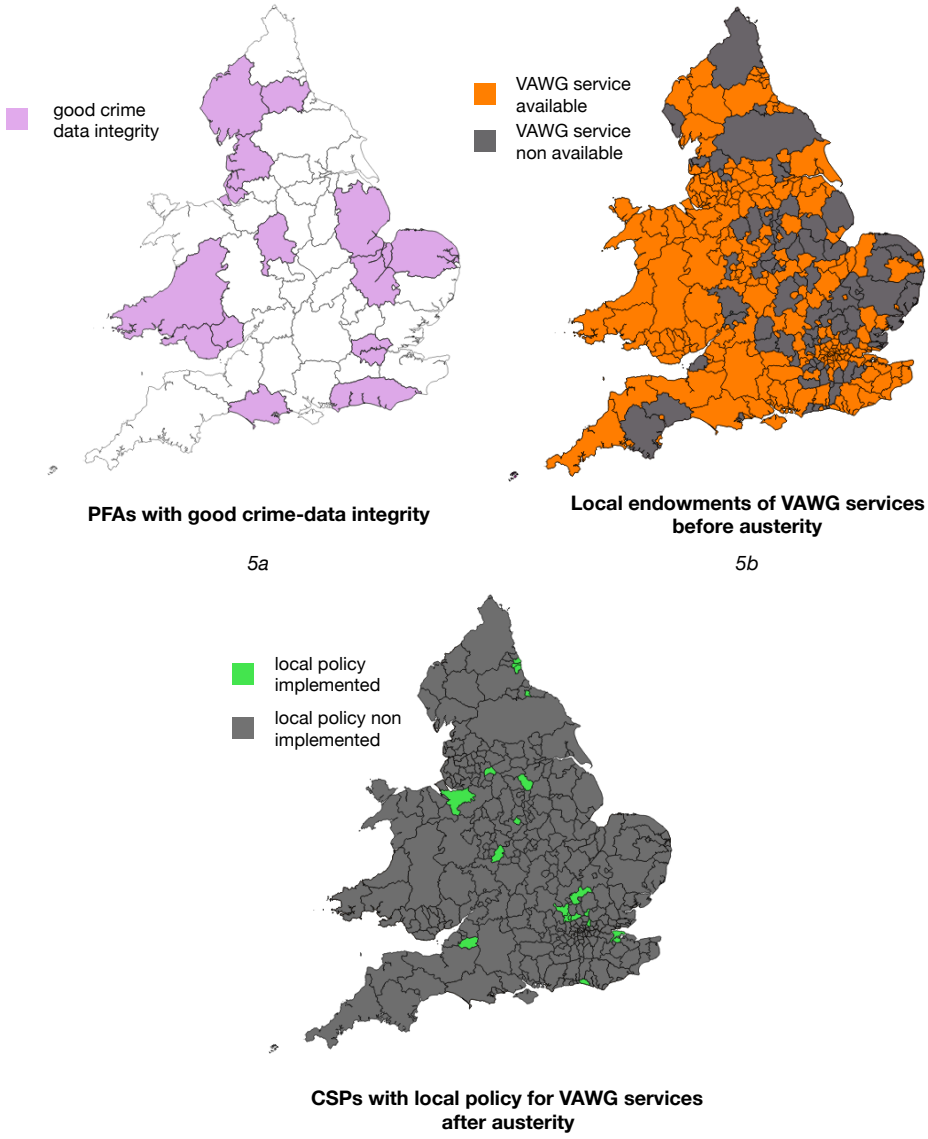
By mapping the local policy choices on the corresponding CSPs across England and Wales, the evidence gathered and scrutinized thus far allow to identify 16 CSPs pursuing local realisation of VAWG support services, alongside Brighton and Hove, which are summarized in Figure 5c. Among the 17 CSPs, 15 refer to PFAs with inaccurate police recording, thus they were already dropped from the sample. Within the remaining 2 CSPs, besides the treated unit -Brighton and Hove- 1 CSP concerns the London Metropolitan Police[107]. This CSP is dropped from the sample, since it cannot contribute to the donor pool of CSPs which is used to define the synthetic counterfactual (Abadie et al., 2010a). Recall that the synthetic counterfactual is meant to reproduce the willingness to report sexual crimes that would have been observed for Brighton and Hove in the absence of the local policy, hence it has to be designed out of a donor pool of CSPs where a local policy for VAWG support service has been absent before and after the austerity.

Finally, by crossing the map of CSPs where the police force is accurate in handling sexual crime reports with the map of CSPs where VAWG support services were provided even before 2010, it is possible to outline the sample of CSPs which will be used to perform the comparative case study. This sample is outlined in Figure 6. Alongside the "*treated*" unit, which is highlighted in green, there are other 85 CSPs in the sample which

[106]The Freedom of Information Act gives to citizens the right to access recorded information held by public sector organizations.
[107]This CSP is Waltham Forest. The comparative case study has chosen Brighton and Hove rather than Waltham Forest since the latter is likely to be affected by spatial spillovers due to its belonging to the London area.

can be exploited to design the synthetic counterfactual. Notably, the CSP of Westminster has been dropped from the sample due to the relevant night economy characterizing the place with consequences also on the size of sexual crimes (ONS, 2018a). This element represents an idiosyncratic feature characterizing that unit, hence it may introduce biases in the design of the synthetic control (Abadie et al., 2010a).

Figure 5: Relevant maps for the definition of the donor pool of CSPs for the synthetic control method



PFAs with good crime-data integrity

5a

Local endowments of VAWG services before austerity

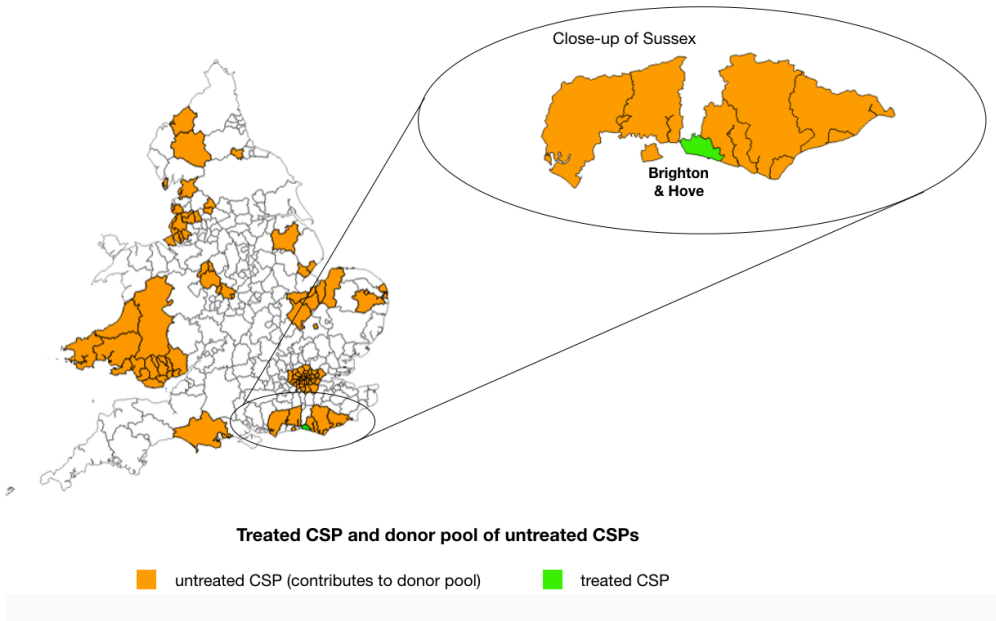
5b

CSPs with local policy for VAWG services after austerity

5c

With respect to the time dimension, the present investigation considers CSP-level police recorded crime figures for the period 2004 – 2016[108], giving 7 years of pre-intervention data.

Figure 6: map of the sample of CSPs considered for synthetic control method (85 untreated CSPs and 1 treated CSP)



4.6 Results

Having defined the sample of CSPs to exploit for the comparative case study analysis, as well as the treated unit and the suitable donor pool of CSPs which can be used to design the synthetic untreated unit, the synthetic control method is applied. Through this approach, the effects of local provision of VAWG services on women's propensity to repost sexual in Brighton and Hove are evaluated by analysing how this propensity would have evolved in Brighton and Hove after 2011 in the absence of a local policy of provision.

Following the literature (Abadie et al., 2010a, 2015), the synthetic Brighton and Hove is designed as the weighted combination of the CSPs in the donor pool which most closely resembled Brighton and Hove before the devolution of VAWG support services determined by the austerity programme.

[108] Police records after and before 2003 are not directly comparable due to the introduction of the National Crime Recording Standard in 2002.

To define the synthetic control model specification, the pre-intervention periods is set between 2004 and 2010, since the devolution to local bodies of deciding whether or not to realize and fund VAWG support services became operative in 2011. Then, a vector of pre-intervention characteristics and outcomes for the exposed unit is defined. Following Dustmann, Schönberg, & Stuhler (2017) this vector includes of the value of the outcome variable in each pre-intervention period and the average over the entire pre-intervention periods of control variables. In the case of the present paper, these variables are (1) the unemployment rate, (2) the real wage of women, (3) the share of female population between 16 and 44 ages. Similarly, a matrix containing the same variables for the unaffected unit is defined. By including the lagged values of the outcome variable, the model specification accounts for the issue of omitted variables, since the lagged outcomes include the effects of any predictor whether or not it is included in the analysis (Athey & Imbens, 2006). The findings from this estimation[111] are summarized in Table 1, which compares the pre-treatment characteristics of the actual Brighton and Hove with that of the synthetic Brighton and Hove.

Table 1: Predictor means for the propensity to report sexual offences by women before the introduction of the local policy in Brighton and Hove

Variable	**Brighton and Hove**	
	Real	**Synthetic**
Reporting sexual crimes on women in 2010	5,114522	5,129826
Reporting sexual crimes on women in 2009	4,988298	5,00303
Reporting sexual crimes on women in 2008	4,935929	4,950125
Reporting sexual crimes on women in 2007	4,827552	4,842507
Reporting sexual crimes on women in 2006	4,871129	4,886971
Reporting sexual crimes on women in 2005	4,975691	4,990995
Reporting sexual crimes on women in 2004	4,705814	4,722017
Share of Unemployment	6,731772	7,863447
Real wage of female workers	6,284676	6,272486
Share of female population between the ages of 16 and 44	0,4848319	0, 4805767

All variables except lagged reporting of sexual crimes are averaged for the 2004-2010 period. Reporting of sexual crimes is measured for the count of sexual crimes records on women per 100.000 women in the CSP and it is transformed in logs. Real wage of female workers is measured in 2001 CPI and it is transformed in logs

[111] Estimation is performed using the STATA module `synth` (Abadie, Diamond, & Hainmueller, 2010b)

From table 1 it appears that the synthetic Brighton and Hove accurately reproduces the values that reporting sexual crime and reporting sexual crimes predictor variables had in Brighton and Hove a prior to the introduction of the local policy. The value of the share of unemployment does not convey an accurate reproduction of the same value for the real Brighton and Hove. This feature does not represent a cause for concern since unemployment does not have substantial power predicting the propensity to report sexual crime in Brighton and Hove before 2011, as shown by the estimated weight given to the predictor[112]. For robustness, the estimation is repeated removing unemployment and findings do not change. Overall, estimated predictor weights have a balanced distribution, implying that the selection of the CSPs from the donor pool to fit the propensity to report sexual crimes in the synthetic Brighton and Hove does not depend mostly from a single predictor.

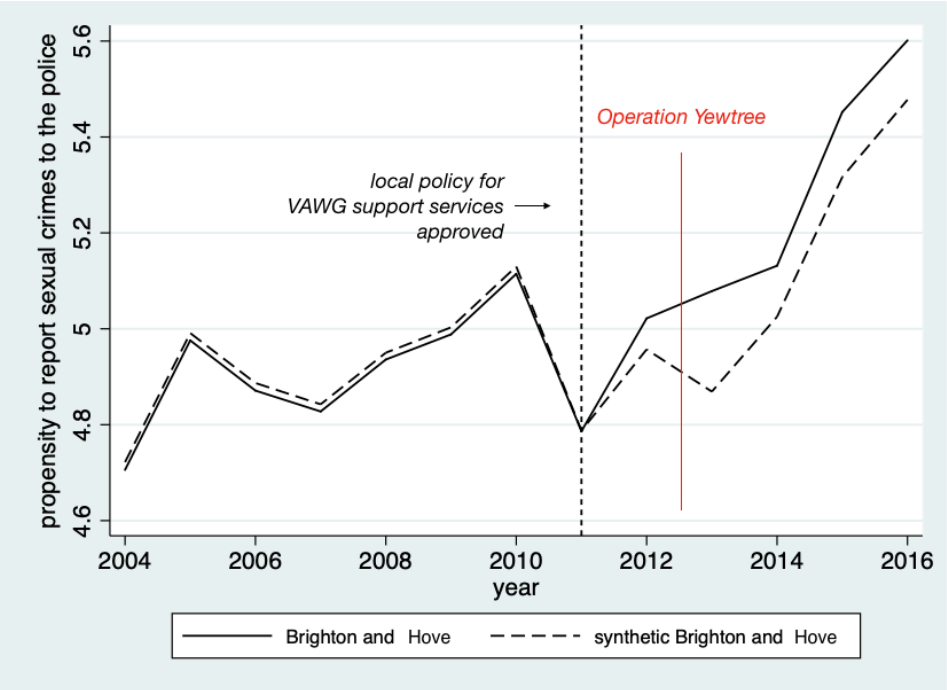
The synthetic Brighton and Hove is reproduced by a combination of 78 CSPs from the donor pool, since 7 CPSs are assigned zero weights. Among the CSPs which are assigned positive weights, the following 7 CSPs contribute the most to the synthetic control unit: Barrow-in-Furness, Burnemouth, Burnely, Darlington, Haringey, Lambeth, Newham, Southwark[113].

Figure 7 displays the reporting of sexual crimes by women for Brighton and Hove and its synthetic counterpart during the period 2004–2016.

[112] The information regarding variable weights tell how much importance is given to each predictor in the model based on their predictive power. The detailed values for the weights assigned to each predictor in the pre-treatment path for Brighton and Hove see Table A2.4 in Annex 2.

[113] The complete outlook of the weights assigned to the CSPs belonging to the donor pool is available in Appendix A2, Table A2.1

Figure 7: Trends in the propensity to report sexual crimes to the police by women: Brighton and Hove vs. synthetic Brighton and Hove



Notice that the reporting of sexual crimes in the synthetic Brighton and Hove very closely track the trajectory of this variable in Brighton and Hove for the entire pre-treatment period. Combined with the high degree of balance on all predictors as showed by Table 1, this suggests that the synthetic Brighton and Hove provides a sensible approximation to the propensity to report sexual crimes to the police that would have been happening in Brighton and Hove in 2011–2016 in the absence of the local policy for the provision of VAWG support services.

The estimate of the effect of the local policy providing VAWG support services on women's propensity to report in Brighton and Hove is the difference between police records of sexual crimes on women in Brighton and Hove and its synthetic version after the implementation of the local policy. The two lines begin to diverge after several months from the passage of the policy, a feature that is probably due to the fact the actual implementation of the policy happened in several months. Staring from 2012, after one year from the approval of the local policy, the two lines start diverging noticeably. The gap between the two lines suggest a positive effect of the local provision of nearby VAWG support services on the propensity to report to the police by women. Between the end of 2012 and the beginning of 2013, the high-profile media coverage of "*Operation*

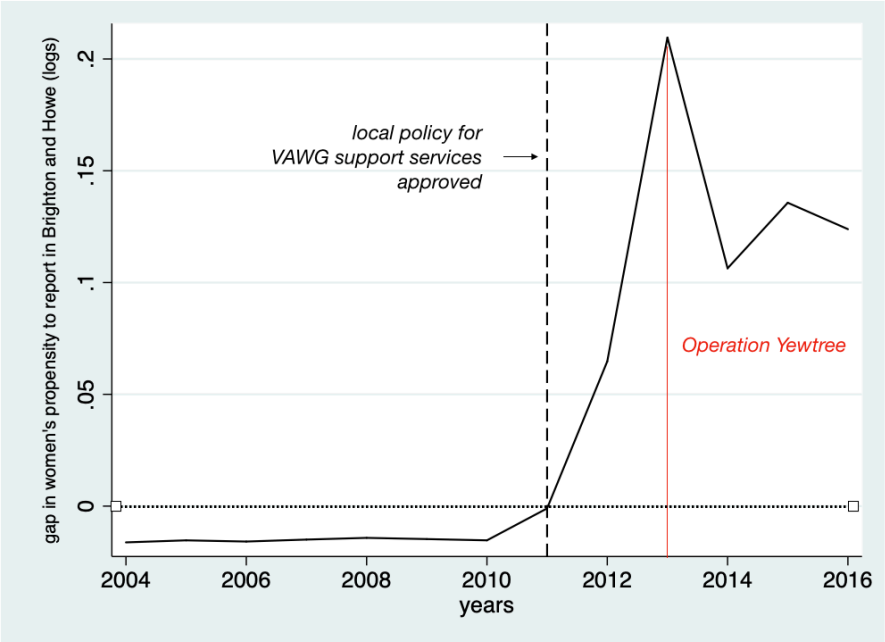
Yewtree" started, disseminating an emotional wave of support and solidarity for victims of sexual crimes countrywide. This wave has been labelled "*Yewtree effect*" and it is associated to a countrywide increase in reporting of sexual offences. The paths of both Brighton and Hove and of its synthetic counterfactual are consistent with the "*Yewtree effect*. Noteworthy, the upward path of Brighton and Hove remains higher than the one of its synthetic counterpart, suggesting that the positive effect exerted by the nation-wide media campaign does not substitute for the role of nearby VAWG support services in stimulating women to come forward and seek justice.

The size of the difference in the propensity of report sexual crimes to the police in Brighton and Hove compared with its synthetic counterpart where no local policy has been adopted is non negligible. Figures show that, from 2012 to 2013, Brighton and Hove had around 40% more reports. Then, with "*Operation Yewtree*" beginning its country-wide media coverage, the gap started reducing, although remaining above 10%.

The upward trend in the pre-treatment period between 2008 and 2010 is concurrent with a commitment by the UK Government to improve VAWG support services, which mainly happened through national-funded programmes that were later reduced as a consequence of the economic recession (Coy et al., 2009; Hirst & Rinne, 2012).

Figure 8 plots the yearly estimates of the impacts of the local policy for the provision of VAWG support services in Brighton and Hove, that is, the yearly gaps in women's propensity to report sexual crimes between Brighton and Hove and its synthetic counterpart. Figure 8 suggests that the local policy had a large effect on women's propensity to report, and that this effect remained positive even after the high-profile media campaign related to "*Operation Yewtree*", supporting that large-scale communication campaigns do not work as substitute for the local provision of specialized services..

Figure 8: Dynamic effects for the Brighton and Hove post-2011 local policy to support VAWG support services on women's propensity to report sexual crimes to the police (in logs)



In order to assess the robustness of these results, estimation is repeated reducing the number of lags for police records of sexual offences by women and results hold[114].

4.7 Inference

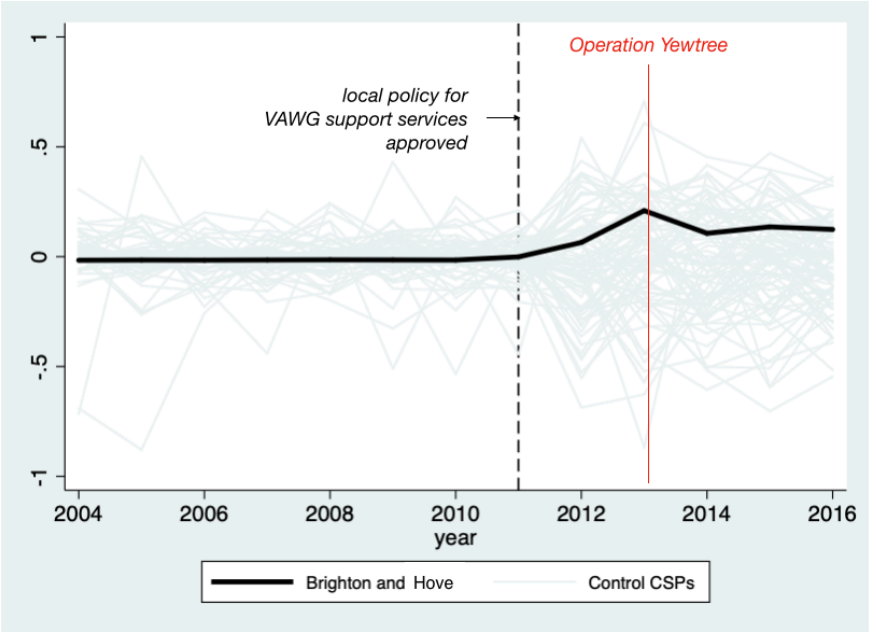
To evaluate the significance of estimates, it is assessed whether the results could be driven entirely by chance. Hence, following the literature (Abadie et al., 2010a; Dustmann et al., 2017) placebo tests are used to check how often results of this magnitude could be obtained by choosing a CSP at random for the study instead of Brighton and Hove. Following Abadie et al. (2010a), the placebo studies are realized by applying the synthetic control method to the CSPs that did not implement the local policy for the provision of VAWG support services during the sample period of the present study. If the placebo studies create gaps of magnitude similar to the one estimated for Brighton and Hove, then there is no significant evidence of a positive effect of the local policy on the propensity to report to the police by female victims. If, on the other hand, the placebo studies show that the gap estimated for Brighton and Hove is unusually large relative to the gaps for the CSPs that did not implement the local policy, then there is significant

[114] See Annex A2, Tables A2.2, A2.3 and A2.4 and Figures A2.1 and A2.2.

evidence of a positive effect of the local policy on the propensity to repost sexual crimes to the police in Brighton and Hove. The estimation process for the placebo tests consists of iteratively applying the synthetic control method used to estimate the effect of the local policy for Brighton and Hove to every other CSP in the donor pool. In each iteration the local policy intervention is reassigned to one of the 85 control CSPs, shifting Brighton and Hove to the donor pool. That is, in each iteration one of the CSPs in the donor pool is assumed to have passed a local policy for the provision of VAWG support services in 2011, instead of Brighton and Hove. We then compute the estimated effect associated with each placebo run. This iterative procedure depicts the distribution of estimated gaps for the CSPs where no intervention took place. Figure 8 summarizes the results from the placebo tests. The superimposed black line denotes the gap estimated for Brighton and Hove, while the light blue lines represent the gap associated with each of the 85 runs of the test. That is, the light blue lines show the difference in women's propensity to report sexual crimes between each CSP in the donor pool and its respective synthetic version. From figure 8, it appears that the estimated gap for Brighton and Hove during the 2011-2016 period is unusually large relative to the distribution of the large majority of CSPs in the donor pool. In other terms, after 2011, the gap between propensity to report sexual crimes by women of the synthetic and actual Brighton and Hove is larger in magnitude than the gap for most placebo CSPs. The results demonstrate also that before 2011, the gap between Brighton and Hove and its synthetic control is extremely small and also smaller than a similar gap for most placebo. The small gap before 2011 and the large and growing gap after 2011 suggest that Brighton and Hove's gap is because of the introduction of the local policy.

At the same time, figure 9 outlines that the propensity to report sexual crime by women during the pre-treatment period is not well reproduced for some CSPs by their synthetic controls. Graphically, this feature is outlined by the lines far below and far above in the pre-treatment period. The CSPs with a bad fit in the pre-treatment period have to be accounted for, since they are characterized by extreme values in the pre-treatment period which do not allow to identify a combination of CSPs in the sample capable of reproducing their path before 2011. To account for this, forthcoming steps not included in the present paper will include repeating the analysis without those CSPs which display a bad fit in the pre-intervention period.

Figure 9: Per-capita gaps in the propensity to report sexual crimes to the police (in logs) in Brighton and Hove and placebo gaps in all 85 control CSPs



4.8 Conclusion and further research

This paper has explored the effects of the introduction of a local policy for the provision of nearby VAWG support services on the propensity to report sexual offences to the police. By analysing a comparative case study with respect to England and Wales, where local authorities have the discretionary policy power to choose whether to realize local VAWG support service or not, the paper presents evidence suggesting a positive effect of the local policy when implemented.

The findings outline that the provision of VAWG support services implemented by the Council of Brighton and Hove had a sizeable impact in increasing the willingness to report to the police by female victims. Furthermore, the positive influence on women's propensity to report related to the provision of nearby VAWG support services is not offset when nation-wide campaign occurred.

These findings support the need for an even spatial distribution of VAWG support service for the UK, an issue that had been previously pinpointed by qualitative works (All-Party Parliamentary Group on Domestic and Sexual Violence Inquiry, 2018; Imkaan, 2013; Women's Aid, 2015).

The investigation carried on in the paper is introductory. Several elements need to be included to fine-tune the proper sample for the comparative case study. To this regard, an important element to consider is given by the effectiveness of the local justice to prosecute identified perpetrators. This feature is relevant in influencing women's propensity to report since it is perceived as a signal of the commitment of the local justice system in tackling sexual crimes (Spohn & Tellis, 2012). To this respect, the audit reports on the efficiency and efficacy of Criminal Courts at the local level have been gathered (RMP Rape Monitoring Group, 2013) and they will be processed to further fine-tune the selection of the donor pool of local authorities used to design the synthetic Brighton and Hove.

By the same token, another element which may contribute to victims' perception of fear is represented by low level detention rate related to sexual offences. Low detention levels mean that the local police force has discarded records of sexual assault after investigation. Hence, the synthetic control method will be performed discarding form the donor pool all CSPs belonging to a PFA with a low detention rate. To re-define the sample of CSPs accordingly, figures from the Inspectorate of Constabulary on detection rate per PFA will be used (RMP Rape Monitoring Group, 2013).

Moreover, even if local authorities' budget is the main source of funding for VAWG support services after 2011, data on other funding streams for VAWG support services will be gathered, to fine-tune the map of local VAWG support services in England and Wales during austerity.

Another potential robustness check is given by a Difference-in-Difference estimation, which needs a throughout analysis of pre-treatment trends to assess whether the assumption of parallel trends hold.

Another step which will be pursued consists of a detailed review of other elements which can potentially influence women's propensity to report. Data from the Crime Survey for England and Wales and from the Survey on Gender-based Violence against Women carried on by the Fundamental Rights Agency of the European Commission will be analysed to identify other relevant features to include in the empirical investigation. Information from these surveys are particularly relevant since they are derived by interviewing victims and alleviating the issue of under-reporting. Although they cannot provide robust spatially fine-grained data, these surveys allow to identify general elements

preventing women from reporting which will be empirically assessed as potential cofounders in the analysis.

References

Abadie, A., Diamond, A., & Hainmueller, A. J. (2010a). Synthetic control methods for comparative case studies: Estimating the effect of California's Tobacco control program. *Journal of the American Statistical Association*, *105*(490), 493–505. https://doi.org/10.1198/jasa.2009.ap08746

Abadie, A., Diamond, A., & Hainmueller, J. (2010b). SYNTH: Stata module to implement Synthetic Control Methods for Comparative Case Studies. *Statistical Software Components*, 457334. Retrieved from http://ideas.repec.org/c/boc/bocode/s457334.html

Abadie, A., Diamond, A., & Hainmueller, J. (2015). Comparative Politics and the Synthetic Control Method. *American Journal of Political Science*, *59*(2), 495–510. https://doi.org/10.1111/ajps.12116

Aizer, A. (2010). The Gender Wage Gap and Domestic Violence Anna. *American Economic Review*, *100*(4), 1847–1859. https://doi.org/10.1002/hbm.21045.Effects

All-Party Parliamentary Group on Domestic and Sexual Violence Inquiry. (2015). *The Changing Landscape of Domestic and Sexual Violence Services*. Retrieved from https://rapecrisis.org.uk/pdfs/2308_appg-changing-landscape-report-2015.pdf

All-Party Parliamentary Group on Domestic and Sexual Violence Inquiry. (2018). *Report into the Funding and Commissioning of Sexual Violence and Abuse Services 2018*.

Anderberg, D., Rainer, H., Wadsworth, J., & Wilson, T. (2016). Unemployment and Domestic Violence: Theory and Evidence. *The Economic Journal*, (1230).

Athey, S., & Imbens, G. . (2006). Identification and Inference in Nonlinear Difference-in-Difference Models. *Econometrica*, *74*(2), 431–497. https://doi.org/10.1111/j.1468-0262.2006.00668.x

Boba, R., & Lilley, D. (2009). Violence Against Women Act (VAWA) Funding: A nationwide assessment of effects on rape and assault. *Violence Against Women*, *15*(2), 168–185. https://doi.org/10.1177/1077801208329146

Bonnewit, A., & DeSantis, E. (2016). The Issue of Violence Against Women in the European Union. Directorate General for Internal Policies Policy Deparment C: Citizens' Rights and Constitutional Affairs.

Council of Europe. Council of Europe Convention on preventing and combating violence against women and domestic violence - CETS 210 - (Istanbul Convention) (2011). Retrieved from https://rm.coe.int/168008482e

Coy, M., Kelly, L., & Foord, J. (2009). *Map of Gaps 2*.

Dustmann, C., Schönberg, U., & Stuhler, J. (2017). Labor supply shocks, native wages, and the adjustment of local employment. *Quarterly Journal of Economics*, *132*(1), 435–483.

https://doi.org/10.1093/qje/qjw032

Ellsberg, M., Arango, D. J., Morton, M., Gennari, F., Kiplesund, S., Contreras, M., & Watts, C. (2015). Prevention of violence against women and girls: What does the evidence say? *The Lancet*, *385*(9977), 1555–1566. https://doi.org/10.1016/S0140-6736(14)61703-7

European Commission -DG Justice and Consumers. (2017). *Women's Economic Independence, a Way out of Violence: Policies and practices to promote.*

European Parliament. (2014). *Report with recommendations to the Commission on combating Violence Against Women (2013/2004(INL)) Plenary sitting 2009 - 2014.*

Ford, R. (2013). Yewtree effect leads to rise in reports of sex abuse. *The Times.*

Fundamental Rights Agency. (2014). *Violence against women: an EU - wide survey.*

Grierson, J. (2018). Council Funding For Women's Refuges Cut By Nearly £7m Since 2010. *The Guardian.* Retrieved from https://www.theguardian.com/society/2018/mar/23/council-funding-womens-refuges-cut-since-2010-england-wales-scotland

Hastings, A., Bailey, N., Bramley, G., Gannon, M., & Watkins, D. (2015). The Cost of the Cuts: The Impact on Local Government and Poorer Comunities. *Joseph Rowntree Foundation Report*, (March), 5–127.

Heady, L., Kail, A., & Yeowart, C. (2011). *Understanding the stability and sustainability of the violence against women voluntary sector.*

Her Majesty's Inspectorate of Constabulary. (2014). *Crime-recording: making the victim count.*

Her Majesty's Inspectorate of Constabulary. (2016). *Crime data integrity inspection.*

Her Majesty's Inspectorate of Constabulary. (2017). *Crime data integrity inspection*

Her Majesty's Inspectorate of Constabulary. (2018). *Crime data integrity inspection*

Her Majesty's Inspectorate of Constabulary. (2019). *Crime data integrity inspection*

Hirst, A., & Rinne, S. (2012). *The impact of changes in commissioning and funding on women-only services.*

Home Office. (2011). *Call to End Violence Against Women and Girls: Action Plan. HM Government.*

Home Office. (2014). *2014 Violence Against Women and Girls Action_Plan.*

House of Lords House of Commons Joint Committee on the Draft Domestic Abuse. (2019). *First Report of Session 2017–19 Report.*

Imkaan. (2013). *Beyond the Labels: women and girls' views on the 2013 mayoral strategy on violence agaist women and girls.*

Maddaus, G. (2018). Weinstein Effect: Sexual Harassment Claims Up 12% Nationwide.

Mcclenaghan, M., & Andersson, J. (2017). *Revealed : Thousands of vulnerable women turned away as refuge funding is cut.* Retrieved from

https://www.thebureauinvestigates.com/stories/2017-10-16/a-system-at-breaking-point

Mendes, K., Ringrose, J., & Keller, J. (2018). #MeToo and the promise and pitfalls of challenging rape culture through digital feminist activism. *European Journal of Women's Studies*, *25*(2), 236–246. https://doi.org/10.1177/1350506818765318

National Audit Office. (2018). *Local authorities' budgets: Departmental Overview*. Retrieved from https://fullfact.org/economy/local-authorities-budgets/

National Audit Office. (2019). Financial sustainability of local authorities. Retrieved from http://www.nao.org.uk/wp-content/uploads/2013/03/Local-Authority-Full-Report.pdf

NSPCC. (2013). *"Giving Victims a Voice" A joint MPS and NSPCC report into allegations of sexual abuse made against Jimmy Savile under Operation Yewtree.*

ONS. (2018a). *Sexual offences in England and Wales: year ending March 2017.*

ONS. (2018b). Sexual offences in England and Wales: Year Ending March 2017. *Office for National Statistics*. Retrieved from https://www.ons.gov.uk/peoplepopulationandcommunity/crimeandjustice/articles/sexualoffencesinenglandandwales/yearendingmarch2017

ONS. (2018c). *Sexual offending : victimisation and the path through the criminal justice system.*

Peri, G., & Yasenov, V. (2016). *The Labor Market Effects of a Refugee Wave: Synthetic Control Method meets the Mariel Boatlift* (National Bureau of Economic Research No. No. w21801). Retrieved from http://giovanniperi.ucdavis.edu/uploads/5/6/8/2/56826033/mariel_may_10_2016.pdf

RMP Rape Monitoring Group. (2013). *Local Area Data.*

Spohn, C., & Tellis, K. (2012). The Criminal Justice System's Response to Sexual Violence. *Violence Against Women*, *18*(2), 169–192. https://doi.org/10.1177/1077801212440020

Taylor, K. (2013). *A Growing Crisis of Unmet Needs: what the figures alone don 't tell.*

The Bureau of Investigative Journalism. (2017). No Refuge Investigation : Refuge Cuts. Retrieved from https://www.thebureauinvestigates.com/stories/2017-10-16/a-system-at-breaking-point

Towers, B. J., & Walby, S. (2012). Measuring the impact of cuts in public expenditure on the provision of services to prevent violence against women and girls, (January), 1–58. Retrieved from http://www.nr-foundation.org.uk/wp-content/uploads/2012/03/Measuring-the-impact-of-cuts-in-public-expenditure-on-the-provision-of-services-to-prevent-violence-against-women-and-girls-Full-report-3.pdf

Turgoose, D. (2016). Domestic violence: Interdisciplinary perspectives on protection, prevention and intervention. In S. Hilder & V. Bettinson (Eds.), *Domestic Violence: Interdisciplinary Perspectives on Protection, Prevention and Intervention* (pp. 1–332). https://doi.org/10.1057/978-1-137-52452-2

UK Government. (2015). *A Call to End Violence against Women and Girls: Progress Report 2010-15*. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/409510/VAWG_Progress_Report_2010-2015.pdf

United Nations. (2012). *Handbook for Legislation on Violence against Women. Handbook for Legislation on Violence against Women*. https://doi.org/10.18356/5e37558d-en

United Nations. (2013a). *Concluding Observations of the Committee on the Elimination of Discrimination Against Women. New York UN*.

United Nations. (2013b). Concluding observations on the seventh periodic report of the United Kingdom of Great Britain and Northern Ireland.

United Nations. (2019). *Concluding observations on the eighth periodic report of the United Kingdom of Great Britain and Northern Ireland*.

Walby, S., Olive, P., Towers, J., Francis, B., Strid, S., Krizsán, A., … Agarwal, B. (2013). *Overview of the worldwide best practices for rape prevention and for assisting women victims of rape. Gender Equality*. https://doi.org/10.2861/37950

Women's Aid. (2013). *Annual Survey 2013*.

Women's Aid. (2014). *Annual Survey 2014*.

Women's Aid. (2015). *Annual Survey 2015*.

Women's Budget Group. (2019). *Triple whammy : The impact of local government cuts on women*.

World Bank. (2014). *Voice and Agency: Empowering Women and Girls for Shared Prosperity*.

Annex A1. Data Sources

Variable	Source
Police records of sexual offences to a female older than 16	The UK Home Office
Female population	ONS
Female population aged between 16 and 44	ONS
Unemployment	The UK Labour Force Survey
Wage differential	The UK Labour Force Survey

Annex A2. Estimation Results

Table A2.1: State weights in the synthetic Brighton and Hove

CSP	Weight	CSP	Weight	CSP	Weight

Barking&Dagenham	0,001	Enfield	0,003	Newcastle un. Lyme	0,001
Barnet	0,001	Fenland	0,001	**Newham**	**0,037**
Barrow-in-Furness	**0,051**	Fylde	0	Newport	0,001
Bexley	0,001	Great Yarmouth	0,002	Norwich	0,002
Blackburn&Darwen	0,001	Greenwich	0,003	Pembrokeshire	0,001
Blackpool	0,002	Hackney	0,006	Pendle	0
Blaenau Gwent	0,001	Hammersmith&Fulham	0,002	Peterborough	0,003
Boston	0,005	**Haringey**	0,029	Powys	0
Bournemouth	**0,198**	Harrow	0,001	Redbridge	0,001
Brent	0,002	Hastings	0,003	Rhondda Cynon Taf	0,001
Bridgend	0,001	Havering	0,002	Richmond up Thames	0,001
Bromley	0,001	Hillingdon	0,003	Rother	0,001
Burnley	**0,064**	Horsham	0,001	South Norfolk	0
Caerphilly	0,001	Hounslow	0,002	South Ribble	0,001
Cambridge	0,001	Huntingdonshire	0,001	**Southwark**	**0,192**
Cannock Chase	0,001	Islington	0,008	St. Helens	0,001
Cardiff	0,005	Kensington&Chelsea	0,002	Stafford	0,001
Carlisle	0,001	King's Lynn W Norfolk	0,001	Stoke-on-Trent	0,001
Carmarthenshire	0,001	Kingston up Thames	0,007	Sutton	0,001
Ceredigion	0,001	**Lambeth**	0,22	Swansea	0,001
Chichester	0,001	Lancaster	0,001	Tamworth	0,001
Chorley	0,001	Lewes	0,001	Torfaen	0,002
Croydon	0,002	Lewisham	0,002	Tower Hamlets	0,02
Darlington	**0,073**	Lichfield	0,001	Vale of Glamorgan	0
Dorset	0,001	Merthyr Tydfil	0,001	Wandsworth	0,001
Ealing	0,002	Merton	0,001	Wealden	0
Eastbourne	0,002	Monmouthshire	0,001	West Lindsey	0,001
Eden	0,001	Neath Port Talbot	0	Wirral	0,002
				Worthing	0,002

Table A2.2: Propensity to report sexual offences by women predictor means before the introduction of the local policy in Brighton and Hove. Three lags for the outcome variable.

Variable	**Brighton and Hove**	
	Real	**Synthetic**
Reporting sexual crimes on women in 2010	5,114522	5,120042
Reporting sexual crimes on women in 2008	4,935929	4,941746
Reporting sexual crimes on women in 2005	4,975691	4,982148
Share of Unemployment	6,731772	6,746063
Real wage of female workers	6,284676	6,291065
Share of female population between the ages of 16 and 44	0,4848319	0,4854871

Figure A2.1: Trends in the propensity to report sexual crimes to the police by women: Brighton and Hove vs. synthetic Brighton and Hove, Four lags for the outcome variable.

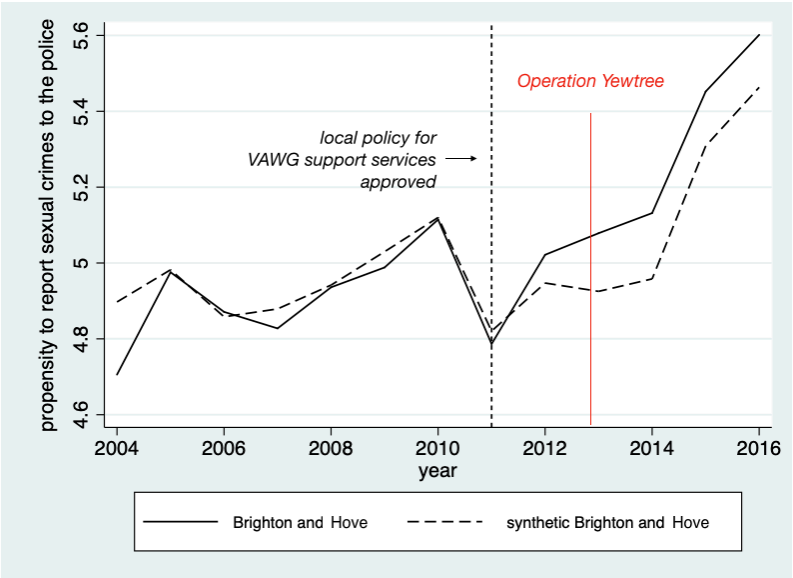


Table A2.3: Propensity to report sexual offences by women predictor means before the introduction of the local policy in Brighton and Hove. Four lags for the outcome variable.

Variable	**Brighton and Hove**	
	Real	**Synthetic**
Reporting sexual crimes on women in 2010	5,114522	5,131205
Reporting sexual crimes on women in 2008	4,935929	4,95183
Reporting sexual crimes on women in 2007	4,827552	4,844358
Reporting sexual crimes on women in 2005	4,975691	4,993105
Share of Unemployment	6,731772	6,75561
Real wage of female workers	6,284676	6,308414
Share of female population between the ages of 16 and 44	0,4848319	0,4862817

Figure A2.2: Trends in the propensity to report sexual crimes to the police by women: Brighton and Hove vs. synthetic Brighton and Hove, Four lags for the outcome variable.

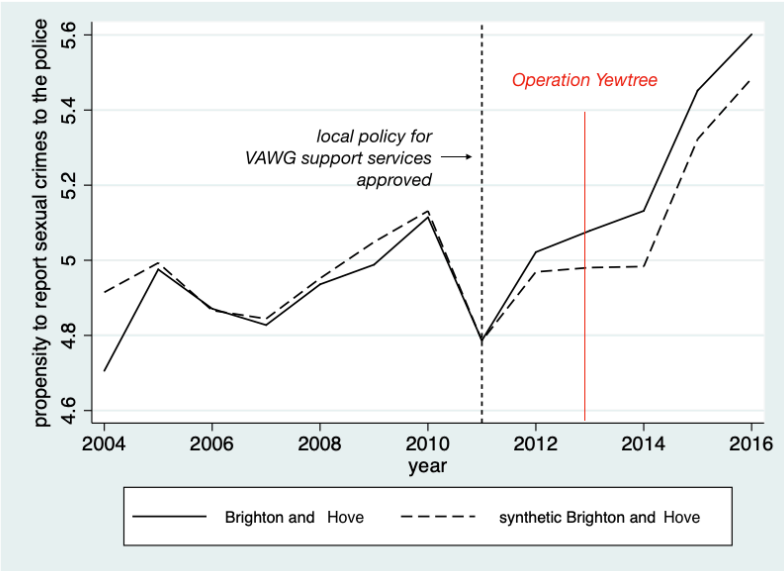


Table A2.4: Synthetic Brighton and Hove Predictor Weights for all specification

Variable	Weights		
	All outcome lags	**4 outcome lags**	**3 outcome lags**
Reporting sexual crimes on women in 2010	0,1479	0,2519	0,3076
Reporting sexual crimes on women in 2009	0,1480	-	-
Reporting sexual crimes on women in 2008	0,1311	0,2115	-
Reporting sexual crimes on women in 2007	0,1141	0,1986	0,2892
Reporting sexual crimes on women in 2006	0,1025	-	-
Reporting sexual crimes on women in 2005	0,1312	0,3266	0,3842
Reporting sexual crimes on women in 2004	0,2253	-	-
Share of Unemployment	4,34e-29	0,0021	0,0022
Real wage of female workers	1,94e-26	0,0040	0,0061
Share of female population between the ages of 16 and 44	3,88e-26	0,0052	0,0107

Table A2.5: RMSPE for all specification

RMSPE	**Model specification**		
	All outcome lags	**4 outcome lags**	**3 outcome lags**
	6,47e-11	0,0748	0,0743

Conclusions.

To conclude, the results presented throughout the four papers comprised in the thesis contributes to the build-up of the knowledge base about the role of places with respect to violent behaviours against minorities and disempowered groups.

Analysing four different types of behaviours and considering two different countries, the evidence which emerges from this thesis strongly identifies economic and sociocultural elements as non-negligible risk factors for hate, cyberhate, school bullying.

Economic inequality emerges as a relevant issue both in Italy and in the UK, with respect to cyberhate and school bullying respectively. The relationship between economic inequality and social uneasiness has also been pinpointed by scholars focusing on electoral outcomes. The evidence presented in this thesis outlines an even more severe effect exerted by inequality, since the findings of chapter 2 and 3 outline how economic inequality can actually drive people towards oppressive violence.

Threats to the established sociocultural identity are another key driver for oppressive violence emerging from the analysis detailed in the thesis. It relates both to hatred manifestation in Italy and to school bullying in the UK. In the latter case, the evidence supports a strong causal relationship between a cultural threat due to a migration shock and oppressive violence happening at school. This result is particularly relevant since it shows that young cohorts are not immune from enacting violence against minorities.

By collecting evidence on the non-negligible role of local geographies on the occurrence and the size of oppressive violent deviant behaviours, this thesis provides support for a non-spatially-blind policy approach to counter these rising phenomena. Each chapter has shown that the characteristics of places strongly matter, hence neglecting them to focus on personality-based policy programme may be ineffective. The limited effectiveness of personality-based policies has already been pinpointed in the literature (Gerstenfeld, 2017; Hall, 2013; Perry, 2001), advocating an approach more focused on the role of situations.

At the same time, up to now, the scope of analysis of the socioeconomic factors relating to oppressive violent deviant behaviours has mainly considered the national and the cross-national level. The investigations presented in this thesis specifically addressed fine-space granularity to assess whether local heterogeneities are relevant, to get that they indeed strongly matter. Hence, these findings do not only add to existing contributions advocating to target socioeconomic issues for an effective countering of oppressive violence, but they also show that these issues have to be considered at the local level.

The relevance of the local dimension is also stressed in the final chapter, where it is showed how a local-level policy can indeed help countering violence against women, even when a relevant national-level campaign is introduced.

Overall, the results presented in the four chapters supports the argument that oppressive violence is not enacted by "*bad seeds*" who are unavoidably born with specific personality traits, but it is enacted in places which become resentful due to socioeconomic issues which can be modified and/or prevented through policy implementation.

The evidence outlined in this thesis wants to contribute to an enduring effort, which finds its contemporary roots in the philosophical thinking in the aftermath of the World War II and that can be epitomised through the words of Hannah Arendt (1964): "*I changed my mind and do no longer speak of "radical evil." … It is indeed my opinion now that evil is never "radical," that it is only extreme, and that it possesses neither depth nor any demonic dimension. It can overgrow and lay waste the whole world precisely because it spreads like a fungus on the surface.*"

Evil is not extraordinary, rather it is banal, and we need to know what makes it proliferate in order to defy it.

References

Arendt H. (1964). Correspondence with Scholem Gershom Gerhard. Adolf Eichmann File Series 1938-1968. The US Library of the Congress

Gerstenfeld, P. B. (2017). Hate Crimes: Causes, Controls, and Controversies. SAGE

Hall, N. (2013). *Hate crimes. 2nd Edition*. Routledge.

Perry, B. (2001). *In the Name of Hate: Understanding Hate Crimes*. https://doi.org/10.1017/CBO9781107415324.004