UNIVERSITA' COMMERCIALE "LUIGI BOCCONI"
PhD SCHOOL

PhD program in Public Policy and Administration

Cycle: XXXIII

Disciplinary Field: SECS-P/03

Essays in Wellbeing, Health and Human Capital Development

Advisor: Simone GHISLANDI

PhD Thesis by
Nirosha Elsem Varghese
ID number: 3056044

Year 2022

Thesis Abstract

The foremost policy challenges confronting the European Union include (i) improving human capital development and economic growth in a society characterized by low fertility and increasing ageing population (Carone, 2005; Serban, 2012) and (ii) ensuring sustainability of the health care systems to ensure better, timely and affordable health care (Rebba, 2014).

The challenges associated with human capital accumulation have been a central point of debate concerning EUs future economic growth. The negative demographic trend in EU if not accompanied by a good quality of education and technological innovation might not be able to compensate for a sizeable drop in skilled workforce (Saini and Keswani Mehra, 2017). Given the advantages of improving human capital outcomes such as increasing labor market productivity, lowering crime, improving health and overall economic growth, human capital investment is of utmost importance (Bleakley, 2010; Izushi and Huggins, 2004). In EU, the ageing population may also induce stress on the fiscal viability of social and health care systems due to increasing comorbidities and chronic diseases among the elderly, reduced income from taxes via longer schooling, contracting workforce, early retirement and crises including recession and pandemics. Additionally, inefficient, and injudicious government policies in the face of crises may result in an unsustainable health care system which may not be able to provide optimal support to the population. In this regard, human capital formation and population health are two outcomes that are subject to policy interventions as they have important implications for quality of life and wellbeing.

Building on these observations, my thesis focuses broadly on two instruments that may significantly influence human capital development and population health. In policy discussions aimed at improving human capital outcomes, a neglected area of research is sleep. Although a much researched topic in medicine, the economic and social penalty of sleep disorders is poorly understood. With the aim of contributing to the scarce literature on sleep and its role in human capital accumulation, chapter 1 presents the findings from four research papers that study the correlates, prevalence estimates and consequences of sleep problems at different points across the lifecycle. Chapter 2 and 3 of this thesis focuses on the next instrument, which are policies, in particular austerity measures and behavioral interventions targeting health systems and public health. To summarize, this thesis is comprised of three main chapters that investigates how physiological processes and government policies independently influence human capital, health and behavioral outcomes across life-cycle.

In chapter 1, I study how sleep, a basic physiological need and an important part of daily routine affect human capital development. We provide evidence suggesting that sleep matters for a range of

outcomes over the lifespan. Our findings call for minimizing disparities in sleep health by developing and implementing effective interventions to reduce overall inequalities in health and wellbeing.

In chapter 1.1, we estimate the prevalence of insufficient sleep and poor sleep in the Italian adult population. Given the importance of sleep, it is vital to monitor the prevalence of indicators of sleep quality and quantity at a population level and compare these estimates and trends within and between countries. Using a cross-sectional study conducted in 2019 on a sample of 3120 subjects, representative of the general Italian adult population, our study shows that the prevalence of sleep dissatisfaction and insufficient sleep in the Italian adult population is likely increasing and relatively high as compared to other European countries. Additional findings identify being older, being a female, poor socio-economic status as correlates of sleep problems. Our findings call for tailored interventions to improve sleep quality and quantity among population with a disadvantaged socioeconomic background. If confirmed by longitudinal data, our results have important implications from a public health perspective.

In chapter 1.2, we study the use of social media and technology among adolescents and if they are associated to sleep onset difficulties. To explore this relationship, we use cross sectional data from the 2013-2014 wave of the Health Behavior in School-aged Children (HBSC) survey conducted on 3172 adolescents aged 11 to 15 years in Northern Italy. Our results show sleep-onset difficulties to be higher among adolescents reporting a higher use of electronic device, social media, and video streaming platforms such as YouTube. Our findings claim for interventions and guidelines to minimize excess use of technology among adolescents. Interventions should include contributions from health professionals and educators to create a well-coordinated parent centered strategy to increase awareness on evolving technologies, social media platforms and the deleterious effects on sleep among adolescents.

In chapter 1.3, we investigate the role of infant night wakings in explaining parity progression. Infant night wakings may influence parents' experience of physical and socio-psychological aspects of childrearing (via disrupted parental sleep), and we hypothesize this experience to play an important role in decisions related to further family planning. Using data from ALSPAC (Avon Longitudinal Study of Parents and Children), we find that an increase in the night wakings of the first child predicts a decreased likelihood of intention to have a second child and this relationship to be significant only until 3 years of age of first child. Our findings call for policy makers to consider the need to put in place, some form of support to improve the parenting experiences of first time parents so that such experiences do not contribute to inhibit further fertility intentions. Such support could include helping new parents "towards managing their postpartum sleep expectations" (Richter et al., 2019) and ways to reduce sleep deprivation after birth.

In chapter 1.4, I study the effects of sleep disruption on child cognitive and non-cognitive outcomes. Sleep plays a vital role in brain maturation which is required for development of memory consolidation, learning and behavioral functions. Prior literature provides evidence linking sleep disruption in childhood to poor human capital outcomes. However, there is limited empirical (and causal) evidence on the dynamics of sleep problems in early childhood and how they influence later outcomes. Using a longitudinal cohort study from UK, this study, for the first time provides evidence on timing of exposure to sleep disruption and human capital outcomes.

This chapter contributes to the literature on sleep and human capital formation in two ways. First, using ALSPAC data, I adapt examples of life course research that investigates the relationship between exposure to shocks in early life such as poverty, maltreatment, instability of family structure and cognitive outcomes, health behaviors and subjective wellbeing to childhood sleep problems (Duncan et al., 1998; Green et al., 2018). This approach can best explain the pattern of relationship between exposure to sleep deprivation across childhood (for example, presence of early sensitive period or duration effects) and later cognitive and non-cognitive outcomes. Our results show that the length of time exposed to sleep disruption across childhood increases the risk for adverse cognitive and non-cognitive outcomes and this may be driven by exposure to sleep disruption in early/mid childhood. On balance, our findings indicate the need to intervene during early childhood. Second, findings from scarring effects strengthen our conclusion that sleep in the past matters for later outcomes irrespective of sleep in the present.

Third, we causally identify the effect of sleep disruption in childhood on later cognitive and school outcomes using exposure to neighbour noise as an instrument. We isolate the causal impact of sleep disruption based on the identifying assumption that exposure to neighbour noise is exogenous to our set of human capital outcomes and affects the outcomes only through sleep disruption (after controlling for a range of socio-economic, dwelling and neighbourhood characteristics). We find evidence for causal effect of sleep disruption in early childhood on outcomes including IQ and total SDQ scores measured during 8 to 15 years of age.

In chapter 2, I discuss the effects of hospital closures in Italy on patient health outcomes. Austerity policies are a frequent response to economic crises and fiscal deficits in both developed and developing countries (Reeves et al., 2014). While the policies may vary in regard to their context, intensity and implementation, such models recommend reducing public expenses and social investments. In Europe, the austerity policies have not spared the health system and budgetary restrictions may very often include closure of hospital facilities. Hospital closures are said to be advantageous if the closed hospitals are inefficient or underutilized. Such closures may actually improve patient outcomes. On the other hand, closures could potentially harm population health through congestion in nearby hospitals and increased travel time in accessing care. In our study, using patient discharge data of AMI patients in Italy, we find that individuals belonging to municipalities exposed to a home hospital closure (closure of a hospital receiving most admission from a municipality in a year) compared to those who were not, had an increased likelihood of experiencing an in-hospital death, increase in cardiac and circulatory disease related 30, 90 and 365-day readmissions and length of stay following an initial AMI admission. We also find that hospital closures are associated with an increase in travel time and bed utilization rates, which in the second stage worsens patient outcomes. Most importantly, the effects of closure are persistent indicating the adaption to this negative shock to be very slow.

The negative effects of hospital closures are most significant during emergency crises leaving the hospital system to be overburdened with less capacity and huge demand. One such crisis is pandemics. In chapter 3, I focus on the COVID-19 pandemic which put a huge stress on the public health system in Europe and all over the world. The chapter focuses on the commitment and acceptability to COVID-19 preventive measures, and in general a snapshot of the public perception of the European population during the pandemic. When health care capacity is limited and saving lives depends solely on non-pharmaceutical interventions, understanding public behavior during the crises is of utmost importance in targeting policies aimed at the flattening the curve.

BIBLIOGRAPHY

Bleakley H. Health, human capital, and development. Annu. Rev. Econ. 2010;2; 283-310.

Carone G. Long-term labour force projections for the 25 EU Member States: A set of data for assessing the economic impact of ageing. European Commission Economic Paper 2005.

Duncan GJ, Yeung WJ, Brooks-Gunn J, Smith JR. How much does childhood poverty affect the life chances of children? American sociological review 1998; 406-423.

Green MJ, Stritzel H, Smith C, Popham F, Crosnoe R. Timing of poverty in childhood and adolescent health: evidence from the US and UK. Social Science & Medicine 2018;197; 136-143.

Izushi H, Huggins R. Empirical analysis of human capital development and economic growth in European regions. Third report on vocational training research in Europe: background report. Luxembourg: Office for Official Publications of the European Communities 2004.

Rebba V. 2014. The long-term sustainability of European health care systems. (Ed)^(Eds). Dipartimento di Scienze Economiche" Marco Fanno"; 2014.

Reeves A, McKee M, Basu S, Stuckler D. The political economy of austerity and healthcare: Cross-national analysis of expenditure changes in 27 European nations 1995–2011. Health policy 2014;115; 1-8.

Richter D, Krämer MD, Tang NK, Montgomery-Downs HE, Lemola S. Long-term effects of pregnancy and childbirth on sleep satisfaction and duration of first-time and experienced mothers and fathers. Sleep 2019;42.

Saini S, Keswani Mehra M. Quality of Schooling: Child Quantity-Quality Tradeoff, Technological Progress and Economic Growth. 2017.

Serban AC. Aging population and effects on labour market. Procedia Economics and Finance 2012;1; 356-364.

Acknowledgements

There are not enough words to convey my deep gratitude and respect for my thesis advisor, Professor Simone Ghislandi. Simone has been a perfect mentor to me and has contributed to the thesis with a major impact. Simone believed in my potential, my work and always spoke for me in ways I would not have for myself. Thank you for always carving out time to listen to my ideas and my concerns, encouraging me to think independently and for providing me a wonderful environment the last four years for doing research that I am now proud of.

I acknowledge the members of my thesis committee: Prof. Aleksandra Torbica and Prof. Giovanni Fattore for your invaluable suggestions and feedback which has tremendously improved the thesis. My discussions with Aleksandra regarding my research, career interest and other aspects of graduate student life has been inspiring and very helpful. My time with Giovanni as a student of Health Policy has immensely shaped my thinking of health economics discipline and policy making. Thank you for your encouraging words and kindness throughout the PhD journey.

I am indebted to Silvano Gallus, Alessandra Lugo and Eugenio Santoro from Mario Negri for their guidance on the research on sleep in the Italian population, which has been instrumental in writing the first chapter of my thesis. I was lucky to collaborate with some amazing colleagues without whom the thesis would not be what it is now. Finally, I acknowledge the administrative staff of the PhD school, Silvia Acquati and Chiara Peverelli and other friends from the PhD cohort who has made my PhD experience a happier one.

A very special word of thanks goes for my parents, Varghese and Mariamma. Thank you, dad, for pushing me when I needed that extra push and your endless support in every way possible. Thank you, Mom, for your unconditional sacrifice in raising me and for the faith you have instilled in me. I am forever indebted to both of you.

My love and gratitude go out to my siblings Nimisha and Nikhil. Thank you, both for looking out for me, for your invaluable support, humor, and fun times over the years. Thanks to my amazing brother in law Roshan for your guidance, love, and encouragement. Thanks to my niece and nephew, Serah and Nathaniel for making our lives happier, and brighter. I hope you read my thesis one day.

A special thanks to my wonderful partner Anandu for your love, patience, and motivation over the years. I am glad I chose to walk with you in this journey of life.

Contents

CHAPTER 1

Sleep problems across the lifespan: Prevalence, causes and consequences

This chapter is comprised of four sets of research findings that brings together a comprehensive picture on sleep across life course, correlates, and potential effects on human capital outcomes.

Chapter 1.1

Sleep dissatisfaction and Insufficient sleep duration in the Italian population[1]

ABSTRACT

To investigate the prevalence and possible determinants of sleep quality and quantity, we used data from a cross-sectional study conducted in 2019 on a sample of 3120 subjects, representative of the general Italian adult population. Sleep dissatisfaction was reported by 14.2% and insufficient sleep (duration) by 29.5% of adults. Sleep dissatisfaction and insufficient sleep were directly related with age (p for trend <0.001), and inversely related with socioeconomic class (p for trend <0.001) and income (p for trend<0.001). Sleep dissatisfaction was higher among women (odds ratio, OR=1.30; 95% confidence interval, CI: 1.05-1.60). Insufficient sleep was inversely related to education (p for trend<0.001) and more frequent in current compared to never smokers (OR: 1.32; 95% CI: 1.08-1.61). Sleep dissatisfaction was higher among divorced/separated compared with married subjects (OR=1.75; 95% CI: 1.20-2.58) and lower among subjects living with children aged 0-14 years (OR=0.48, 95% CI: 0.33-0.70). Pet owners more frequently had sleep dissatisfaction (OR=1.35, 95% CI: 1.08-1.68) and insufficient sleep (OR=1.46, 95% CI: 1.23-1.73). In Italy, self-perceived sleep problems appear to be increasing. Sleep problems can contribute to aggravating health disparities in the society. The unfavourable relationship with pets (and the favourable ones with children) should be confirmed by longitudinal studies.

INTRODUCTION

An adequate and healthy sleep is required for a good quality of life, better productivity and wellbeing (Moore, 2012). Insufficient sleep has been found to be associated with physical health outcomes, including poor cardiovascular health, high blood pressure, insulin resistance and obesity (Cappuccio et al., 2010; Cespedes Feliciano et al., 2018; Gangwisch et al., 2010; Mitchell et al., 2013), mental health consequences, such as depression, suicidal and other self-harm behavior (Gangwisch et al., 2010; Liu et al., 2017; Roberts and Duong, 2014), as well as overall mortality (Åkerstedt et al., 2017; Cappuccio et al.,

---

[1] This was a joint project with Alessandra Lugo[a], Simone Ghislandi[b], Paolo Colombo[c], Roberta Pacifici[d] and Silvano Gallus[a] and published in *Scientific Reports.*

a) Istituto di Ricerche Farmacologiche Mario Negri IRCCS, Italy
b) Bocconi University, Italy
c) Istituto DOXA, Italy
d) Istituto Superiore di Sanità, Italy

2010). An estimated 7 to 9 hours of total sleep is recommended for the adult population for healthy sleep (Hirshkowitz et al., 2015) whereas the US National Sleep Foundation (NSF) has put forward the four key indicators for good sleep quality to include i) less difficulty in initiating sleep, ii) sleeping more time while in bed, iii) less than two instances of night waking and iv) ability to go back to sleep within 20 minutes of a night awakening. More recently, sleep satisfaction was added as a further key indicator for good sleep quality (Ohayon et al., 2017; Ohayon et al., 2018).

Given the importance of sleep on all aspects of health (Cappuccio et al., 2010), it is vital to monitor the prevalence of indicators of sleep quality and quantity at a population level and compare these estimates and trends within and between countries. Sleep quality can be measured either as a subjective rating regarding one's satisfaction with sleep and/or as an index computed from several components related to sleep quality. Among others (Knutson et al., 2017), one widely used instrument for measuring sleep quality is the 24-item Pittsburgh Sleep Quality Index (PSQI), (Buysse et al., 1989) with translations and validations in several languages (Curcio et al., 2013; Manzar et al., 2018). Efforts made to shorten the instrument using factor analyses showed that subjective sleep quality (i.e., sleep satisfaction) and sleep duration explained a significant portion of the variance of the PSQI score (Manzar et al., 2018).

To ensure objectivity of the measures, the self-completed questionnaires on sleep, when possible, are supplemented by objective measurements of sleep quality such as polysomnographic, actigraphy or movements captured using other phone apps or body wearables (Marino et al., 2013). However, the obvious trade-off between higher accuracy of device supported measures versus the inexpensive and ease of large sample collection via self-reported information has led to population based epidemiological studies to use self-reported measures on sleep (Manzar et al., 2018).

In Italy, there are very few studies based on representative samples of the general population on self-reported sleep problems (Boffi et al., 2015; Leger et al., 2008; Ohayon and Smirne, 2002). In particular, a cross-sectional study based on a sample of 3970 subjects, representative of the Italian general population found that 10.1% of Italian adults are dissatisfied with sleep (Ohayon and Smirne, 2002). The data used in this study dates back to 1997-1998. To our knowledge, no data after 2005 are available on sleep quality in Italy. Information on time spent sleeping is included in various surveys conducted by the Italian National Institute of Statistics. Using 2008-2009 time use data from ISTAT, Boffi and colleagues showed that in Italy, an average of 8 hours and 17 minutes is spent for sleeping. Overall, they show that time spent sleeping has reduced in Italy, when compared to the past and to other European countries.

Given the paucity of data on the issue in the Italian population, we take advantage of a representative survey on Italian adults conducted in 2019 to assess the prevalence and determinants of sleep dissatisfaction and insufficient sleep (duration). To ensure comparability, we chose two questions on a subjective rating of overall sleep quality and actual sleep duration from the validated Italian version of the PSQI (Curcio et al., 2013).

METHODS

We used data from a survey, conducted by DOXA - the Italian branch of the Worldwide Independent Network/Gallup International Association (WIN/GIA) - and coordinated by the Italian National Institute of Health and by Mario Negri Institute (Liu et al., 2019). Our analysis employs the data collected during February-April 2019 on a sample of 3120 subjects aged 15 years and over, representative of the general adult Italian population (52.4 million inhabitants in 2019), in terms of sex, age, area of residence and socioeconomic characteristics (Lugo et al., 2017).

Representative multistage sampling was applied to select participants. Briefly, the first stage involved the selection of 114 municipalities in all the 20 Italian regions, on the basis of region and size of municipalities. The second stage involved a random selection of an adequate number of electoral wards in each municipality, so that more or less affluent areas of each municipality were represented in the right proportions. In the third stage, individuals were randomly sampled from each electoral ward within strata of sex and age. A 'quota' method based on sex and age was used to select adolescents aged 15-17 years, since their names were not included in the electoral rolls. Finally, the data processing also involved generating statistical weights for each subject to ensure representativeness of the Italian population aged 15 years and older.

The interviews were conducted in Italian by *ad hoc* trained interviewers using a structured questionnaire in the context of a computer-assisted personal in-house (CAPI) interview.

Besides demographic (e.g., age, sex and marital status) and socioeconomic characteristics (e.g., geographic area, level of education, perceived socioeconomic level and self-reported family income), we collected information on smoking status, presence of children with various ages in the household and on pet ownership to indicate if the participants own one or more cats, one or more dogs, or one or more cats or dogs. Consistent with the validated Italian version of PSQI (Curcio et al., 2013), subjective evaluation of sleep quality was asked through the question: 'During the past month, how would you rate your sleep quality overall?'. Possible answers were: (1) very good; (2) quite good; (3) quite bad (4) very bad. We defined individuals to be dissatisfied with sleep if participants evaluated their overall quality of sleep to be quite bad or bad (Grandner et al., 2006). Sleep quantity was reported in hours of sleep as a continuous measure and was assessed using the question: 'During the past month, how many hours of actual sleep did you get at night?’. For our analysis, we defined ‘insufficient sleep’ as the ‘sleep duration that is likely too brief to meet physiologic needs’ (Grandner, 2019). We used the cut-off of 6 hours per night of sleep duration to disentangle those with insufficient sleep (≤6 hours per night), as recommended by several panels and sleep societies (Watson et al., 2015). According to the NSF recommendations on sleep time duration, sleeping 5 to 6 hours may be appropriate (not recommended) for those who are 65 and above as the need for sleep decreases with age (Hirshkowitz et al., 2015). Therefore, we conduct a sensitivity check with a cutoff of ≤5 hours of sleep per night defined as insufficient sleep duration for this age group.

To investigate determinants of sleep dissatisfaction and insufficient sleep, odds ratios (OR) and the corresponding 95% confidence intervals (CI) were estimated using multiple logistic regression models adjusted for sex, age, level of education, and geographic area. As a robustness check, we also repeat the analysis stratified by age groups (15-44, 45-64 and 65 and above) to examine if the determinants for sleep dissatisfaction and insufficient sleep duration vary by age. Our study is of exploratory nature and a significance level of 5% was chosen. The results were reported without any correction for multiplicity(Bender and Lange, 2001). Statistical weights were used in all analyses to ensure representativeness of our sample. Statistical analysis was performed using SAS V.9.4 and Stata V.15 statistical software.

RESULTS

Of the 3120 Italian participants, 23.6% described their overall sleep quality as very good, 62.2% as quite good, 12.1% as quite bad and 2.1% as very bad. Mean sleep duration was 7.00 hours per night (SD: 1.16; data not shown in tables).

Table 1 shows prevalence estimates for sleep dissatisfaction and insufficient sleep duration according to selected socioeconomic characteristics and lifestyle habits. Overall, 14.2% (95% CI: 13.0%-15.5%) of subjects reported having sleep dissatisfaction. The prevalence of subjects with sleep dissatisfaction was higher among women than men (OR=1.30; 95% CI: 1.05-1.60), whereas no relationship with sex was observed for insufficient sleep. Both sleep dissatisfaction and insufficient sleep increased with age: compared to <45 year olds, the ORs for sleep dissatisfaction were 2.81 (95% CI: 2.12-3.71) for 45-64 years and 4.17 (95% CI: 3.08-5.64; p for trend<0.001) for ≥65 years; the corresponding ORs for insufficient sleep were 2.39 (95% CI: 1.96-2.91) and 3.25 (95% CI: 2.60-4.07; p for trend<0.001), respectively. Whereas no statistically significant relation was observed between level of education and sleep dissatisfaction, insufficient sleep decreased with increasing of level of education (p for trend<0.001). Both sleep dissatisfaction and insufficient sleep decreased with increasing socioeconomic class (p for trend<0.001 for both measures of sleep) and income (p for trend <0.001). No relation was observed between smoking status and sleep duration, but compared to never, current smoker had more frequently insufficient sleep (OR= 1.32; 95% CI: 1.08-1.61).

[Table 1 about here]

Table 2 shows prevalence estimates, and corresponding ORs, for sleep dissatisfaction and insufficient sleep by marital status, presence of children in household and ownership of pets. Compared to married, divorced or separated subjects had more frequently sleep dissatisfaction (OR=1.75; 95% CI: 1.20-2.58), but not insufficient sleep. Subjects living with children aged 0-14 years (OR=0.48; 95% CI: 0.33-0.70) had less frequently sleep dissatisfaction compared to participants without children. Results were similar in an additional specification when comparing participants living with no children aged 0-14 to living with children below 5 but not between 6-14 years (OR=0.28; 95% CI= 0.11-0.70) and living with children between 6-14 years but not below 5 (OR=0.49; 95% CI = 0.32-0.75). Overall, 32.1% had a pet (either a cat or a dog), 14.7% a cat and 23.2% a dog. Owning a pet (cat or dog) and owning both cats and dogs was directly related to both sleep dissatisfaction (OR=1.35; 95% CI: 1.08-1.68 for a pet and OR=1.61; 95% CI: 1.07-2.42 for both cats and dogs) and insufficient sleep (OR=1.46; 95% CI: 1.23-1.73 for a pet and OR=2.12; 95% CI:1.52-2.95 for both cats and dogs). Additionally, owning dogs but not cats was directly related to insufficient sleep duration (OR=1.43, 95% CI: 1.15-1.78).

[Table 2 about here]

Table S1 and **S2** replicates the results in Table 1 but stratified by age groups 15-44, 45-64 and ≥ 65 years, respectively. Higher sleep dissatisfaction and insufficient sleep among women were statistically significant only among the age group 45-64 years. Results for relationship between income and both sleep dissatisfaction and insufficient sleep was observed in all age groups whereas the results for education and perceived socioeconomic class were significant only among those aged 65 and older. A higher prevalence of sleep dissatisfaction and insufficient sleep was reported among current smokers aged 15-44 years whereas being a former smoker among those aged 65 and older was associated with both lower sleep dissatisfaction and insufficient sleep.

Table S3 and S4 replicates the results in Table 2 but stratified by age groups. The relationship between being divorced or separated and both sleep dissatisfaction and insufficient sleep was significant in all age groups except among those aged 45-64 years. Among those aged 15-44, there was a lower prevalence of insufficient sleep duration. The negative relationship between living with children in various age groups and both sleep dissatisfaction and insufficient sleep was statistically significant only among those aged

45-64 years. The positive relationship between pet ownership and sleep dissatisfaction was observed only among those aged 45-64 whereas insufficient sleep duration had a higher prevalence among pet owners belonging to both age groups 45-64 and 65 and older.

Table S5 reports findings for participants aged 65 and above when the cutoff of insufficient sleep duration is changed to ≤5 hours. Results for participants aged 65 and above do not change when the cutoff for insufficient duration is changed from ≤6 hours to ≤5 hours.

DISCUSSION

Our exploratory study shows that the prevalence estimates of sleep dissatisfaction and insufficient sleep in the Italian general adult population are 14% and 30%, respectively. Age and socio-economic characteristics are major determinants of sleep quality and quantity. Living with children is directly related with sleep quality whereas having pets is inversely related with both sleep quality and quantity.

Compared with the only representative study using our definition of sleep dissatisfaction, the sleep quality of Italian adults appears to be worsening from 10% in 1996-1997 (Ohayon and Smirne, 2002) to 14% in 2019. Our estimates on the proportion of adults with insufficient sleep in Italy were higher than those observed in selected nationally representative surveys when comparable definitions were used, (i.e., actual sleep <7 or ≤6 hours per day), including Finland (14%) and Australia (17%) (Hublin et al., 2007; Magee et al., 2009), and similar to those observed in the USA (28% to 40%), with the higher estimates coming from studies using the Behavioral Risk Factors Surveillance System (BRFSS) and the National Health and Nutrition Examination Survey (NHANES) surveys respectively (Grandner, 2017; Grandner et al., 2014; Krueger and Friedman, 2009; Liu et al., 2016; Singh et al., 2005).

We confirm the results of most, but not all (Grandner et al., 2015; Soldatos et al., 2005; Zilli et al., 2009), previous studies in several European populations, including Italy (Ohayon and Smirne, 2002), showing higher subjective sleep dissatisfaction to be increasing with age (Madrid-Valero et al., 2017; Ohayon and Paiva, 2005; Ohayon and Zulley, 2001). Our analysis also showed insufficient sleep to be higher among older age groups, in agreement with current evidence (Akerstedt et al., 2017; Boffi et al., 2015; Grandner, 2017; Grandner et al., 2015). In fact, previous studies have shown that sleep becomes more disturbed and lighter with age and this could be attributed to physiologic changes or decrease in time of melatonin production (Grandner, 2017).

Gender differences in sleep quality has been previously reported (Nowakowski et al., 2013). Our findings are consistent with past literature that shows women to report more sleep dissatisfaction compared to men

(Ohayon and Paiva, 2005; Ohayon and Partinen, 2002; Ohayon and Zulley, 2001; Ursin et al., 2005). These differences in sleep quality have been attributed to biological events in the life course, including hormonal influences or as a manifestation of sex differences in mental health which is more prevalent in women (van den Berg et al., 2009; Zhang and Wing, 2006). Moreover, women have been shown to express emotional concerns and bodily symptoms more readily in general compared to men (Groeger et al., 2004). However, we do not find significant results for the relationship between insufficient sleep and sex, in agreement with some (Groeger et al., 2004; Hume et al., 1998) but not all previous studies (Basner et al., 2014; Dalmases et al., 2018; Whinnery et al., 2014).

In our analysis, the dimensions of socio-economic status (education, income, and perceived socio-economic status) had a negative relationship with sleep dissatisfaction and insufficient sleep. Previous studies conducted in the US, European and Australian populations have consistently shown that lower socio-economic background, lower income families and low levels of education are risk factors for insufficient sleep and poor subjective sleep quality (Grandner, 2019; Magee et al., 2009; Moore et al., 2002; Ohayon, 1996; Ursin et al., 2005). Previous studies using longitudinal data show that poverty 'scars', that is the ability of past exposure to poverty in reducing current life satisfaction even when an individual is out of poverty (Clark et al., 2016). Given that sleep quality and quantity are significant predictors of life satisfaction, the scarring effects of poverty may take a toll on sleep health even when there is no current exposure to poverty (Ness and Saksvik-Lehouillier, 2018). Overall, the findings are in line with well documented evidence on social determinants of health including poverty and the role of sleep as a mediator between socioeconomic status (SES) and health (Grandner et al., 2015; Marmot and Wilkinson, 2005; Van Cauter and Spiegel, 1999).

Previous findings are mixed with studies reporting either direct or non-significant relationship between being a current smoker and sleep problems (Cohen et al., 2020; Ohayon and Paiva, 2005). However, we find a direct relationship between being a smoker and both sleep dissatisfaction and insufficient sleep among the younger population. When insufficient sleep is defined as ≤6 hours, our results on smoking as a predictor of insufficient sleep are in line with findings from a study conducted among Australian adults aged 45-65 years (Magee et al., 2009). In Finland, when insufficient sleep is defined as 1 hour difference between self-reports of sleep need and sleep length, women who are current smokers were also more likely to report insufficient sleep (Hublin et al., 2007). A lower prevalence of sleep problems among former smokers in the older population could be attributed to an improved perception of sleep quality after quitting nicotine in addition to reduced need for sleep with increasing age (Hirshkowitz et al., 2015). It should also be noted that the mean age of former smokers is sixty years in our sample indicating large lifestyle changes around that age, which could be driving these results.

Several theoretical models explain the positive link between being married and health in general via sharing of resources to achieve better health outcomes (Chen et al., 2015). A similar reasoning is applied to sleep outcomes with studies finding positive relationships between being married and better sleep outcomes including satisfaction with sleep (Abraham et al., 2017; Arber, 2012; Arber et al., 2009; Ohayon and Paiva, 2005). In particular, epidemiological studies have shown sleep problems that extend beyond 10 weeks after marital separation to translate into worsening health outcomes in general (Krietsch et al., 2014).

Prior literature shows the importance of not just presence or absence of children but also the age of the children in the household for sleep quality of the parents (Hagen et al., 2013). Contrary to previous findings showing new parents to face up to 6 years of disrupted sleep (Richter et al., 2019), our study finds no such effect. In fact, in our Italian population, we found an inverse relation between sleep dissatisfaction and living with children aged below 5 and 0-14 years. One explanation for these findings could be that our results are driven by older parents for a few possible reasons. In fact, older parents have been shown to have both more child rearing experience and positive attitude to parenting (Ragozin et al., 1982). They are also relatively highly educated and earn more, therefore placing them in a position with fewer worries in life. They have fewer emotional and behavioral problems and the same is reflected in their offspring's (Zondervan-Zwijnenburg et al., 2020). This gives an advantage to older parents in terms of both their own capabilities and that of the temperament of the offspring to enjoy parenthood better via either fewer sleep problems or better skills and resources to manage them.

Given the mean age of first time mothers in Italy is around 31 years (Eurostat, 2017), one could approximate that participants driving this negative relationship to be second or third time parents and therefore strengthen our findings in line with the above explanation (the age group driving our results is 45-64 years). However, our reasoning can only be validated if the findings in the subgroup for younger parents are in an expected direction (positive) which is true for insufficient sleep. The same does not hold for sleep dissatisfaction which could be due to the fact that although respondents may sleep fewer hours than an average person, they may still be satisfied with their sleep. Among younger parents, life satisfaction increases following childbirth (Baetschmann et al., 2016) which may correspond to higher sleep satisfaction but not necessarily sleep duration.

Our findings could also be explained as an adaptation effect (Clark et al., 2008) whereby living with a newborn baby will rescale your perception of a good quality sleep and after a time span the parent will perceive sleep quality much better than another person without a similar experience. However, the category 'living with children below 5 years' although not statistically significant also show a relationship in the same direction as other categories. Therefore, we need both studies with a larger sample size and information on children at even younger ages to see evidence of such an adaptation.

Furthermore, our findings on presence of children in the household are likely to be affected by potential confounding. In particular, there could be an omitted variable that is positively related to sleep problems and negatively related to presence of children in the household that could be confounding this relation. One could think of a variable like poor health of the respondent that may reduce fertility or fertility intentions (Sharma et al., 2013) and is also positively correlated to sleep problems (Epstein, 2010). Number of children is an endogenous variable (Baranowska-Rataj and Matysiak, 2016) and hence could also be subjected to reverse causality, which is sleep dissatisfaction and insufficient sleep duration predicting fewer children. Therefore, we require a good instrumental variable to tease out the true effect of presence of children in the household. Finally, our estimates could also be subjected to what is known as functional form approximation which arises due to fewer observations within a category of the variable (Kennedy, 2005). In our case, the number of observations reporting having children in the household is far fewer than the control category (no children) and therefore the linear term with the negative sign could be the best approximation of the data. Our finding on children needs to be investigated and validated with datasets with larger samples and more information on other individual characteristics to make sure that our results are not driven by confounding or lack of power.

The presence of pets in the house can modify the sleep environment including odor, noise and temperature (Krahn et al., 2015). Our study shows that pet ownership in general or owning a cat or dog specifically increases the likelihood of being dissatisfied with sleep and having insufficient sleep duration. Although it is widely accepted in the medical community that pets should not be allowed into the bedrooms at night; evidence on the role of pet ownership on owner's sleep quality is sparse. An Australian study based on a sample of 2036 adults found that participants who co-slept with their pets took longer to fall asleep, were more likely to wake up tired and disturbed during sleep by dog barking or other animal noises (Smith et al., 2014). Our results on pet ownership were also partly consistent with two other recent empirical studies. A study based on 40 healthy adults with no sleep disorders and objective measurement of sleep showed that having a dog in the bedroom may disrupt the sleep of the owners only if the dog shares the bed (Patel et al., 2017). Mein and Grant (2018) used elderly participants from the Whitehall II aged 59-79 years to show that dog owners were more likely to wake up more tired compared to non-dog owners (Mein and Grant, 2018). However, our findings could also be biased given the cross-sectional nature of our data. For instance, pet owners could be a selected population who suffer more frequently from sleep problems or other health characteristics leading to pet selection effects (Downes et al., 2009; Saunders et al., 2017). Hence reverse causality cannot be excluded and should be confirmed using prospective longitudinal data. Given the mixed evidence, more recent empirical works point towards the need to prioritize the sleep of owners over loyalty to the pet. Moreover, health care workers should consider inquiring about pet ownership to patients with sleep concerns and discuss the pros and cons of having a pet as their sleep companion.

Our study has some limitations. Our analysis includes self-reported measures of sleep dissatisfaction and sleep duration (Grandner, 2019). Previous studies show that participants tend to overestimate self-reported measures of sleep. Second, our survey measures sleep duration in whole number hours which is less precise than sleep duration in fractions or minutes. But this does not hinder us in defining insufficient sleep at the cutoff ≤6 hours. Also, the main sources of survey data in the US for sleep measures including BRFSS and NHANES uses sleep duration measured in whole number hours (Grandner, 2017). Third, previous studies show that prevalent health conditions such as cancer or lifestyle diseases such as obesity, heart disease or type-2 diabetes is either directly or indirectly related to sleep disruption through stress of disease burden(Epstein, 2010). Due to unavailability of information on any measure of health, our study could be exposed to omitted variability bias. However, controlling for smoking status could account for some measure of lifestyle which could be a proxy for current health status. Four, methodological drawbacks inherent to the cross-sectional study design, particularly problems of reverse causality, could not be ruled out. Finally, we performed several tests without taking into account multiple testing (Bender and Lange, 2001; Rubin, 2017). Therefore, the results from our exploratory study should be interpreted with caution and should be confirmed by future studies. The strengths of our study include the relatively large sample size that allowed us to obtain estimates adjusted for multiple potential confounding factors, the sample representativeness of the general Italian adult population by age, sex and socioeconomic characteristics, and the use of validated and standardized assessment measures for sleep dissatisfaction and insufficient sleep.

Our study shows that the prevalence of sleep dissatisfaction and insufficient sleep in the Italian adult population is likely increasing and relatively high as compared to other European countries. Interventions to improve sleep quality and quantity should be tailored to the population with a disadvantage socioeconomic background. Our findings on pets (and children) should be further investigated. If confirmed by longitudinal data, they might have important implications from a public health perspective.

BIBLIOGRAPHY

Abraham, O., *et al.* 2017. Factors contributing to poor satisfaction with sleep and healthcare seeking behavior in older adults. *Sleep Health,* 3(1), 43-48.

Akerstedt, T., *et al.* 2017. Sleep duration, mortality and the influence of age. *Eur J Epidemiol,* 32(10), 881-891.

Åkerstedt, T., *et al.* 2017. Sleep duration, mortality and the influence of age. *European journal of epidemiology*, 32(10), 881-891.

Arber, S. 2012. Gender, marital status and sleep problems in Britain. *Przegl Lek,* 69(2), 54-60.

Arber, S., Bote, M. and Meadows, R. 2009. Gender and socio-economic patterning of self-reported sleep problems in Britain. *Soc Sci Med,* 68(2), 281-289.

Baetschmann, G., Staub, K. E. and Studer, R. 2016. Does the stork deliver happiness? Parenthood and life satisfaction. *Journal of Economic Behavior & Organization,* 130, 242-260.

Baranowska-Rataj, A. and Matysiak, A. 2016. The causal effects of the number of children on female employment-do European institutional and gender conditions matter? *Journal of Labor Research,* 37(3), 343-367.

Basner, M., Spaeth, A. M. and Dinges, D. F. 2014. Sociodemographic characteristics and waking activities and their role in the timing and duration of sleep. *Sleep*, 37(12), 1889-1906.

Bender, R. and Lange, S. 2001. Adjusting for multiple testing—when and how? *Journal of clinical epidemiology,* 54(4), 343-349.

Boffi, M., Colleoni, M. and Del Greco, M. 2015. Night-time Hours and Activities of the Italians. *Articulo-Journal of Urban Research,* (11).

Buysse, D. J., *et al.* 1989. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry Res,* 28(2), 193-213.

Cappuccio, F. P., *et al.* 2010. Sleep duration and all-cause mortality: a systematic review and meta-analysis of prospective studies. *Sleep,* 33(5), 585-592.

Cespedes Feliciano, E. M., *et al.* 2018. Objective Sleep Characteristics and Cardiometabolic Health in Young Adolescents. *Pediatrics,* 142(1).

Chen, J. H., Waite, L. J. and Lauderdale, D. S. 2015. Marriage, Relationship Quality, and Sleep among U.S. Older Adults. *J Health Soc Behav*, 56(3), 356-377.

Clark, A. E., D'Ambrosio, C. and Ghislandi, S. 2016. ADAPTATION TO POVERTY IN LONG-RUN PANEL DATA. *Rev Econ Stat,* 98(3), 591-600.

Clark, A. E., *et al*. 2008. Lags and leads in life satisfaction: A test of the baseline hypothesis. *The Economic Journal*, 118(529), F222-F243.

Cohen, A., Ben Abu, N. and Haimov, I. 2020. The Interplay Between Tobacco Dependence and Sleep Quality Among Young Adults. *Behav Sleep Med,* 18(2), 163-176.

Curcio, G., *et al.* 2013. Validity of the Italian version of the Pittsburgh Sleep Quality Index (PSQI). *Neurol Sci,* 34(4), 511-519.

Dalmases, M., *et al.* 2018. Assessing sleep health in a European population: Results of the Catalan Health Survey 2015. *PLoS One,* 13(4), e0194495.

Downes, M., Canty, M. J. and More, S. J. 2009. Demography of the pet dog and cat population on the island of Ireland and human factors influencing pet ownership. *Prev Vet Med,* 92(1-2), 140-149.

Epstein, L. J., 2010. *Improving sleep: A guide to a good night's rest.* Harvard Health Publications.

Eurostat 2017. Fertility Indicators.

Gangwisch, J. E., *et al.* 2010. Earlier parental set bedtimes as a protective factor against depression and suicidal ideation. *Sleep,* 33(1), 97-106.

Grandner, M. A. 2017. Sleep, Health, and Society. *Sleep Med Clin,* 12(1), 1-22.

Grandner, M. A., 2019. Epidemiology of insufficient sleep and poor sleep quality. *Sleep and Health.* Elsevier, 11-20.

Grandner, M. A., *et al.* 2014. Habitual sleep duration associated with self-reported and objectively determined cardiometabolic risk factors. *Sleep Med,* 15(1), 42-50.

Grandner, M. A., *et al.* 2015. Social and Behavioral Determinants of Perceived Insufficient Sleep. *Front Neurol,* 6, 112.

Grandner, M. A., *et al.* 2006. Criterion validity of the Pittsburgh Sleep Quality Index: Investigation in a non-clinical sample. *Sleep Biol Rhythms,* 4(2), 129-139.

Groeger, J. A., Zijlstra, F. R. and Dijk, D. J. 2004. Sleep quantity, sleep difficulties and their perceived consequences in a representative sample of some 2000 British adults. *J Sleep Res,* 13(4), 359-371.

Hagen, E. W., *et al.* 2013. The sleep-time cost of parenting: sleep duration and sleepiness among employed parents in the Wisconsin Sleep Cohort Study. *Am J Epidemiol,* 177(5), 394-401.

Hirshkowitz, M., *et al.* 2015. National Sleep Foundation's sleep time duration recommendations: methodology and results summary. *Sleep Health,* 1(1), 40-43.

Hublin, C., *et al.* 2007. Sleep and mortality: a population-based 22-year follow-up study. *Sleep,* 30(10), 1245-1253.

Hume, K. I., Van, F. and Watson, A. 1998. A field study of age and gender differences in habitual adult sleep. *J Sleep Res,* 7(2), 85-94.

Kennedy, P. E. 2005. Oh no! I got the wrong sign! What should I do? *The Journal of Economic Education,* 36(1), 77-92.

Knutson, K. L., *et al.* 2017. The National Sleep Foundation's Sleep Health Index. *Sleep Health,* 3(4), 234-240.

Krahn, L. E., Tovar, M. D. and Miller, B. 2015. Are Pets in the Bedroom a Problem? *Mayo Clin Proc,* 90(12), 1663-1665.

Krietsch, K. N., Mason, A. E. and Sbarra, D. A. 2014. Sleep complaints predict increases in resting blood pressure following marital separation. *Health Psychol,* 33(10), 1204-1213.

Krueger, P. M. and Friedman, E. M. 2009. Sleep duration in the United States: a cross-sectional population-based study. *Am J Epidemiol,* 169(9), 1052-1063.

Leger, D., *et al.* 2008. An international survey of sleeping problems in the general population. *Curr Med Res Opin,* 24(1), 307-317.

Liu, X., *et al.* 2017. Poor sleep quality and nightmares are associated with non-suicidal self-injury in adolescents. *Eur Child Adolesc Psychiatry,* 26(3), 271-279.

Liu, X., *et al.* 2019. Electronic cigarettes in Italy: a tool for harm reduction or a gateway to smoking tobacco? *Tob Control*.

Liu, Y., *et al.* 2016. Prevalence of Healthy Sleep Duration among Adults--United States, 2014. *MMWR Morb Mortal Wkly Rep,* 65(6), 137-141.

Lugo, A., *et al.* 2017. Smoking in Italy in 2015-2016: prevalence, trends, roll-your-own cigarettes, and attitudes towards incoming regulations. *Tumori,* 103(4), 353-359.

Madrid-Valero, J. J., *et al.* 2017. Age and gender effects on the prevalence of poor sleep quality in the adult population. *Gac Sanit,* 31(1), 18-22.

Magee, C. A., Iverson, D. C. and Caputi, P. 2009. Factors associated with short and long sleep. *Prev Med,* 49(6), 461-467.

Manzar, M. D., *et al.* 2018. Dimensionality of the Pittsburgh Sleep Quality Index: a systematic review. *Health Qual Life Outcomes,* 16(1), 89.

Marino, M., *et al.* 2013. Measuring sleep: accuracy, sensitivity, and specificity of wrist actigraphy compared to polysomnography. *Sleep,* 36(11), 1747-1755.

Marmot, M. and Wilkinson, R., 2005. *Social determinants of health.* OUP Oxford.

Mein, G. and Grant, R. 2018. A cross-sectional exploratory analysis between pet ownership, sleep, exercise, health and neighbourhood perceptions: the Whitehall II cohort study. *BMC Geriatr,* 18(1), 176.

Mitchell, J. A., *et al.* 2013. Sleep duration and adolescent obesity. *Pediatrics,* 131(5), e1428-1434.

Moore, M. 2012. Looking for a good night's sleep. *Lancet,* 380(9839), 322-323.

Moore, P. J., *et al.* 2002. Socioeconomic status and health: the role of sleep. *Psychosom Med,* 64(2), 337-344.

Ness, T. E. B. and Saksvik-Lehouillier, I. 2018. The relationships between life satisfaction and sleep quality, sleep duration and variability of sleep in university students. *Journal of European Psychology Students,* 9(1).

Nowakowski, S., Meers, J. and Heimbach, E. 2013. Sleep and Women's Health. *Sleep Med Res,* 4(1), 1-22.

Ohayon, M. 1996. Epidemiological study on insomnia in the general population. *Sleep,* 19(3 Suppl), S7-15.

Ohayon, M., *et al.* 2017. National Sleep Foundation's sleep quality recommendations: first report. *Sleep Health,* 3(1), 6-19.

Ohayon, M. M., *et al.* 2018. A provisional tool for the measurement of sleep satisfaction. *Sleep Health,* 4(1), 6-12.

Ohayon, M. M. and Paiva, T. 2005. Global sleep dissatisfaction for the assessment of insomnia severity in the general population of Portugal. *Sleep Med,* 6(5), 435-441.

Ohayon, M. M. and Partinen, M. 2002. Insomnia and global sleep dissatisfaction in Finland. *J Sleep Res,* 11(4), 339-346.

Ohayon, M. M. and Smirne, S. 2002. Prevalence and consequences of insomnia disorders in the general population of Italy. *Sleep Med,* 3(2), 115-120.

Ohayon, M. M. and Zulley, J. 2001. Correlates of global sleep dissatisfaction in the German population. *Sleep,* 24(7), 780-787.

Patel, S. I., *et al.* 2017. The Effect of Dogs on Human Sleep in the Home Sleep Environment. *Mayo Clin Proc,* 92(9), 1368-1372.

Ragozin, A. S., *et al.* 1982. Effects of maternal age on parenting role. *Developmental Psychology,* 18(4), 627.

Richter, D., *et al.* 2019. Long-term effects of pregnancy and childbirth on sleep satisfaction and duration of first-time and experienced mothers and fathers. *Sleep,* 42(4).

Roberts, R. E. and Duong, H. T. 2014. The prospective association between sleep deprivation and depression among adolescents. *Sleep,* 37(2), 239-244.

Rubin, M. 2017. Do p values lose their meaning in exploratory analyses? It depends how you define the familywise error rate. *Review of General Psychology,* 21(3), 269-275.

Saunders, J., *et al.* 2017. Exploring the differences between pet and non-pet owners: Implications for human-animal interaction research and policy. *PLoS One,* 12(6), e0179494.

Sharma, R., *et al.* 2013. Lifestyle factors and reproductive health: taking control of your fertility. *Reproductive Biology and Endocrinology,* 11(1), 66.

Singh, M., *et al.* 2005. The association between obesity and short sleep duration: a population-based study. *J Clin Sleep Med,* 1(4), 357-363.

Smith, B., *et al.* 2014. The Prevalence and implications of human–animal co-sleeping in an Australian sample. *Anthrozoös,* 27(4), 543-551.

Soldatos, C. R., *et al.* 2005. How do individuals sleep around the world? Results from a single-day survey in ten countries. *Sleep Med,* 6(1), 5-13.

Ursin, R., Bjorvatn, B. and Holsten, F. 2005. Sleep duration, subjective sleep need, and sleep habits of 40- to 45-year-olds in the Hordaland Health Study. *Sleep,* 28(10), 1260-1269.

Van Cauter, E. and Spiegel, K. 1999. Sleep as a mediator of the relationship between socioeconomic status and health: a hypothesis. *Ann N Y Acad Sci,* 896, 254-261.

van den Berg, J. F., *et al.* 2009. Sex differences in subjective and actigraphic sleep measures: a population-based study of elderly persons. *Sleep,* 32(10), 1367-1375.

Watson, N. F., *et al.* 2015. Recommended Amount of Sleep for a Healthy Adult: A Joint Consensus Statement of the American Academy of Sleep Medicine and Sleep Research Society. *Sleep,* 38(6), 843-844.

Whinnery, J., *et al.* 2014. Short and long sleep duration associated with race/ethnicity, sociodemographics, and socioeconomic position. *Sleep,* 37(3), 601-611.

Zhang, B. and Wing, Y. K. 2006. Sex differences in insomnia: a meta-analysis. *Sleep,* 29(1), 85-93.

Zilli, I., Ficca, G. and Salzarulo, P. 2009. Factors involved in sleep satisfaction in the elderly. *Sleep Med,* 10(2), 233-239.

Zondervan-Zwijnenburg, M. A., *et al.* 2020. Parental Age and Offspring Childhood Mental Health: A Multi-Cohort, Population-Based Investigation. *Child development,* 91(3), 964-982.

TABLES

Table 1. Distribution of 3120 Italian participants aged ≥15 years, according to their sleep dissatisfaction and insufficient sleep, overall and by demographic and socio-economic characteristics, and smoking status. Corresponding odds ratios° (OR) and 95% confidence intervals (CI). Italy, 2019.

	N	Sleep dissatisfaction		Insufficient sleep duration	
		N (%)	OR (95% CI)	N (%)	OR (95% CI)
Total	3120	443 (14.2)		921 (29.5)	
Sex					
Male	1501	184 (12.3)	1.00^	449 (29.9)	1.00^
Female	1619	259 (16.0)	**1.30 (1.05-1.60)**	473 (29.2)	0.91 (0.78-1.07)
Age					
<45 years	1281	82 (6.4)	1.00^	223 (17.4)	1.00^
45-64 years	1039	172 (16.5)	**2.81 (2.12-3.71)**	354 (34.0)	**2.39 (1.96-2.91)**
≥65 years	800	189 (23.6)	**4.17 (3.08-5.64)**	345 (43.1)	**3.25 (2.60-4.07)**
P for trend			**<0.001**		**<0.001**
Education level					
Low	1028	192 (18.7)	1.00^	395 (38.5)	1.00^
Intermediate	1579	202 (12.8)	0.93 (0.72-1.19)	420 (26.6)	**0.81 (0.67-0.99)**
High	513	50 (9.8)	0.73 (0.51-1.05)	106 (20.7)	**0.61 (0.47-0.80)**
P for trend			0.108		**<0.001**
Perceived socioeconomic class					
Low	428	93 (21.8)	1.00^	193 (45.2)	1.00^
Intermediate	2294	310 (13.5)	**0.65 (0.49-0.87)**	628 (27.4)	**0.56 (0.44-0.70)**
High	398	40 (9.9)	**0.45 (0.29-0.70)**	100 (25.1)	**0.51 (0.37-0.71)**
P for trend			**<0.001**		**<0.001**
Self-reported income					
Low	1372	258 (18.8)	1.00^	515 (37.5)	1.00^
Intermediate	1004	118 (11.8)	**0.64 (0.50-0.83)**	271 (27.0)	**0.74 (0.61-0.89)**
High	744	67 (9.0)	**0.51 (0.37-0.70)**	136 (18.2)	**0.49 (0.39-0.62)**
P for trend			**<0.001**		**<0.001**
Smoking status					
Never smoker	2055	287 (14.0)	1.00^	570 (27.8)	1.00^
Ex-smoker	378	62 (16.5)	0.92 (0.68-1.26)	125 (33.1)	0.93 (0.72-1.18)
Current smoker	687	94 (13.6)	1.10 (0.84-1.44)	226 (32.9)	**1.32 (1.08-1.61)**

° ORs were estimated using unconditional multiple logistic regression models after adjustment for age groups (<25; 25-44; 45-64; ≥65 years), sex, level of education and geographic area. Estimates in bold are statistically significant at 0.05 level.

^ Reference category.

Table 2. Distribution of 3120 Italian participants aged ≥15 years, according to their sleep quality and quantity, by selected household characteristics. Corresponding odds ratios° (OR) and 95% confidence intervals (CI). Italy, 2019.

	N	Sleep dissatisfaction		Insufficient sleep duration	
		N (%)	OR (95% CI)	N (%)	OR (95% CI)
Marital Status					
Married/Cohabiting	1814	265 (14.6)	1.00^	575 (31.7)	1.00^
Single	879	68 (7.7)	1.15 (0.84-1.57)	171 (19.5)	1.09 (0.86-1.38)
Divorced/Separated	172	41 (24)	**1.75 (1.20-2.58)**	57 (33.2)	1.09 (0.77-1.53)
Widowed	255	69 (26.9)	1.36 (0.96-1.92)	118 (46.4)	1.32 (0.98-1.78)
Children 0-14 years					
No	2492	405 (16.3)	1.00^	791 (31.8)	1.00^
Yes	628	38 (6.1)	**0.48 (0.33-0.70)**	130 (20.7)	0.80 (0.64-1.01)
Children (Age categories)					
No children (0-14 years)	2492	405 (16.3)	1.00^	791 (31.8)	1.00^
Children below 5 but not between 6-14 years	150	5 (3.4)	**0.28 (0.11-0.70)**	25 (16.6)	0.66 (0.42-1.05)
Children between 6-14 but not below 5 years	414	26 (6.4)	**0.49 (0.32-0.75)**	89 (21.6)	0.81 (0.62-1.06)
Children below 5 and 6-14 years	64	7 (10.4)	0.97 (0.42-2.26)	16 (25.1)	1.11 (0.61-2.03)
Pets					
No	2118	279 (13.2)	1.00^	586 (27.7)	1.00^
Yes	1002	164 (16.4)	**1.35 (1.08-1.68)**	335 (33.4)	**1.46 (1.23-1.73)**
Pets (Cats/Dogs)					
No pets	2118	279 (13.2)	1.00^	586 (27.7)	1.00^
Cat but no dog	277	50 (17.9)	1.33 (0.94-1.88)	86 (31.1)	1.18 (0.89-1.57)
Dog but no cat	543	80 (14.8)	1.27 (0.96-1.68)	174 (32.0)	**1.43 (1.15-1.78)**
Cats and dogs	182	34 (18.8)	**1.61 (1.07-2.42)**	75 (41.1)	**2.12 (1.52-2.95)**

° ORs were estimated using unconditional multiple logistic regression models after adjustment for age groups (<25; 25-44; 45-64; ≥65 years), sex, level of education and geographic area. Estimates in bold are statistically significant at 0.05 level.
^ Reference category

Table S1. Distribution of 3120 Italian participants aged ≥15 years, according to their sleep dissatisfaction, stratified by age groups and by demographic and socio-economic characteristics, and smoking status. Corresponding odds ratios° (OR) and 95% confidence intervals (CI). Italy, 2019.

	Sleep dissatisfaction								
	15-44 years			45-64 years			≥ 65 years		
	N	N (%)	OR (95% CI)	N	N (%)	OR (95% CI)	N	N (%)	OR (95% CI)
Total	1281	83 (6.4)		1039	172 (16.5)		800	189 (23.6)	
Sex									
Male	647	38 (5.8)	1.00^	507	70 (13.9)	1.00^	346	76 (22.1)	1.00^
Female	634	45 (7.1)	1.26 (0.80-1.97)	532	101 (19.1)	**1.48 (1.06-2.06)**	454	113 (24.8)	1.16 (0.83-1.63)
Education level									
Low	198	13 (6.3)	1.00^	294	49 (16.7)	1.00^	536	130 (24.3)	1.00^
Intermediate	786	54 (6.8)	1.23 (0.64-2.34)	588	99 (16.8)	0.90 (0.61-1.34)	205	50 (24.1)	0.87 (0.59-1.28)
High	297	16 (5.5)	0.92 (0.43-1.97)	157	24 (15.4)	0.81 (0.47-1.40)	59	9 (15.9)	0.55 (0.26-1.14)
P for trend			0.830			0.457			0.107
Perceived socioeconomic class									
Low	137	10 (7.0)	1.00^	111	20 (18.1)	1.00^	179	64 (35.5)	1.00^
Intermediate	990	67 (6.8)	1.10 (0.52-2.31)	757	129 (17.0)	0.94 (0.55-1.61)	547	115 (21.0)	**0.49 (0.33-0.72)**
High	154	6 (3.8)	0.58 (0.19-1.75)	171	23 (13.4)	0.69 (0.34-1.40)	74	11 (14.8)	**0.28 (0.13-0.59)**
P for trend			0.335			0.305			**0.001**
Self-reported income									
Low	480	46 (9.6)	1.00^	430	83 (19.3)	1.00^	461	129 (27.9)	1.00^
Intermediate	445	24 (5.3)	**0.57 (0.34-0.98)**	329	50 (15.1)	0.71 (0.48-1.06)	230	45 (19.4)	**0.59 (0.39-0.90)**
High	356	13 (3.6)	**0.39 (0.20-0.77)**	280	39 (13.9)	**0.59 (0.37-0.93)**	109	16 (14.4)	**0.46 (0.25-0.83)**
P for trend			**0.007**			**0.023**			**0.011**
Smoking status									
Never smoker	870	47 (5.4)	1.00^	649	103 (15.9)	1.00^	536	137 (25.6)	1.00^
Current smoker	330	29 (8.8)	1.63 (0.99-2.67)	260	45 (17.1)	1.11 (0.75-1.63)	96	20 (20.8)	0.71 (0.40-1.25)
Ex-smoker	81	7 (8.1)	1.56 (0.66-3.70)	130	24 (18.6)	1.22 (0.75-2.00)	168	32 (18.9)	**0.60 (0.38-0.95)**

° ORs were estimated using unconditional multiple logistic regression models after adjustment for sex, level of education and geographic area. Estimates in bold are statistically significant at 0.05 level.

^ Reference category.

Table S2. Distribution of 3120 Italian participants aged ≥15 years, according to insufficient sleep duration, stratified by age groups and by demographic and socio-economic characteristics, and smoking status. Corresponding odds ratios° (OR) and 95% confidence intervals (CI). Italy, 2019.

	Insufficient sleep duration								
	15-44 years			45-64 years			≥ 65 years		
	N	N (%)	OR (95% CI)	N	N (%)	OR (95% CI)	N	N (%)	OR (95% CI)
Total	1281	223 (17.4)		1039	354 (34.0)		800	345 (43.1)	
Sex									
Male	647	108 (16.7)	1.00^	507	193 (38.1)	1.00^	346	148 (42.8)	1.00^
Female	634	115 (18.2)	1.13 (0.84-1.51)	532	161 (30.2)	**0.70 (0.54-0.91)**	454	197 (43.4)	1.02 (0.76-1.36)
Education level									
Low	198	39 (19.8)	1.00^	294	106 (36.1)	1.00^	536	250 (46.7)	1.00^
Intermediate	786	139 (17.6)	0.98 (0.66-1.47)	588	201 (34.2)	0.98 (0.72-1.33)	205	80 (38.9)	**0.62 (0.44-0.87)**
High	297	45 (15.3)	0.78 (0.48-1.25)	157	46 (29.5)	0.79 (0.51-1.21)	59	15 (24.8)	**0.33 (0.18-0.62)**
P for trend			0.296			0.270			**0.001**
Perceived socioeconomic class									
Low	137	31 (22.3)	1.00^	111	48 (43.0)	1.00^	179	115 (64.0)	1.00^
Intermediate	990	166 (16.8)	0.84 (0.53-1.34)	757	251 (33.2)	0.72 (0.47-1.09)	547	211 (38.5)	**0.41 (0.29-0.60)**
High	154	26 (17.1)	0.85 (0.46-1.58)	171	54 (31.9)	0.72 (0.42-1.23)	74	19 (25.9)	**0.21 (0.11-0.40)**
P for trend			0.615			0.233			**<0.001**
Self-reported income									
Low	480	108 (22.6)	1.00^	430	169 (39.2)	1.00^	461	238 (51.6)	1.00^
Intermediate	445	76 (17.0)	0.81 (0.58-1.14)	329	108 (32.9)	0.77 (0.57-1.06)	230	87 (37.7)	**0.63 (0.44-0.89)**
High	356	39 (11.0)	**0.52 (0.34-0.79)**	280	76 (27.3)	**0.63 (0.44-0.90)**	109	20 (18.5)	**0.26 (0.15-0.45)**
P for trend			**0.002**			**0.011**			**<0.001**
Smoking status									
Never smoker	870	124 (14.3)	1.00^	649	214 (33.0)	1.00^	536	232 (43.2)	1.00^
Current smoker	330	86 (26.1)	**2.02 (1.47-2.79)**	260	90 (34.4)	1.00 (0.73-1.35)	96	50 (51.9)	1.39 (0.87-2.23)
Ex-smoker	81	13 (15.8)	1.09 (0.58-2.06)	130	50 (38.1)	1.19 (0.80-1.77)	168	63 (37.5)	0.72 (0.49-1.06)

° ORs were estimated using unconditional multiple logistic regression models after adjustment for sex, level of education and geographic area. Estimates in bold are statistically significant at 0.05 level.

^ Reference category

Table S3. Distribution of 3120 Italian participants aged ≥15 years, stratified by age groups and according to their sleep dissatisfaction, by selected household characteristics. Corresponding odds ratios° (OR) and 95% confidence intervals (CI). Italy, 2019.

	Sleep dissatisfaction								
	15-44 years			45-64 years			≥ 65 years		
	N	N (%)	OR (95% CI)	N	N (%)	OR (95% CI)	N	N (%)	OR (95% CI)
Total	1281	223 (17.4)		1039	354 (34.0)		800	345 (43.1)	
Marital status									
Married/Cohabiting	545	42 (7.7)	1.00^	782	125 (16.0)	1.00^	487	98 (20.2)	1.00^
Single	700	36 (5.1)	**0.61 (0.38-0.98)**	136	21 (15.8)	1.03 (0.62-1.71)	43	11 (24.7)	1.38 (0.65-2.91)
Divorced/Separated	35	4 (12.8)	1.89 (0.65-5.46)	102	20 (19.9)	1.26 (0.74-2.13)	35	17 (47.1)	**4.06 (1.92-8.56)**
Widowed	1	0 (0)	-	19	5 (26.1)	1.83 (0.64-5.20)	235	64 (27.1)	**1.49 (1.01-2.20)**
Children 0-14 years									
No	835	58 (7.0)	1.00^	873	160 (18.3)	1.00^	784	187 (23.9)	1.00^
Yes	446	24 (5.4)	0.72 (0.43-1.18)	166	12 (7.2)	**0.37 (0.20-0.68)**	16	2 (11.5)	0.55 (0.12-2.57)
Children (Age categories)									
No children (0-14 years)	835	58 (7.0)	1.00^	873	160 (18.3)	1.00^	784	187 (23.9)	1.00^
Children below 5 but not between 6-14 years	133	5 (3.4)	0.46 (0.17-1.23)	11	1 (5.8)	0.29 (0.02-3.82)	6	0 (0)	-
Children between 6-14 but not below 5 years	257	13 (5.1)	0.66 (0.35-1.23)	147	11 (7.8)	**0.39 (0.21-0.74)**	10	2 (18.1)	0.78 (0.16-3.84)
Children below 5 and 6-14 years	56	7 (12.0)	1.62 (0.68-3.85)	8	0 (0)	-	0	-	-
Pets									
No	870	53 (6.1)	1.00^	668	98 (14.7)	1.00^	580	128 (22.1)	1.00^
Yes	411	29 (7.1)	1.23 (0.76-1.97)	371	74 (20.0)	1.39 (0.99-1.94)	220	61 (27.8)	1.27 (0.88-1.82)
Pets (Cats/Dogs)									
No pets	870	53 (6.1)	1.00^	668	98 (14.7)	1.00^	580	128 (22.1)	1.00^
Cat but no dog	95	7 (7.0)	1.25 (0.54-2.92)	90	17 (19.2)	1.34 (0.75-2.37)	92	26 (27.8)	1.24 (0.75-2.07)
Dog but no cat	249	18 (7.2)	1.20 (0.68-2.10)	207	37 (17.7)	1.19 (0.78-1.81)	87	26 (30.0)	1.38 (0.83-2.32)
Cats and dogs	67	5 (7.2)	1.33 (0.50-3.52)	74	20 (26.9)	**2.10 (1.20-3.69)**	41	10 (23.1)	1.07 (0.50-2.32)

° ORs were estimated using unconditional multiple logistic regression models after adjustment for sex, level of education and geographic area. Estimates in bold are statistically significant at 0.05 level.
^ Reference category.

Table S4. Distribution of 3120 Italian participants aged ≥15 years, stratified by age groups and according to insufficient sleep duration, by selected household characteristics. Corresponding odds ratios° (OR) and 95% confidence intervals (CI). Italy, 2019.

	Insufficient sleep duration								
	15-44 years			45-64 years			≥ 65 years		
	N	N (%)	OR (95% CI)	N	N (%)	OR (95% CI)	N	N (%)	OR (95% CI)
Total	1281	223 (17.4)		1039	354 (34.0)		800	345 (43.1)	
Marital status									
Married/Cohabiting	545	118 (21.7)	1.00^	782	259 (33.1)	1.00^	487	198 (40.5)	1.00^
Single	700	103 (14.7)	**0.58 (0.43-0.79)**	136	53 (39.4)	1.34 (0.91-1.96)	43	15 (35.0)	0.77 (0.39-1.51)
Divorced/Separated	35	2 (5.9)	0.24 (0.06-1.00)	102	34 (33.5)	1.07 (0.69-1.67)	35	21 (59.5)	**3.02 (1.43-6.36)**
Widowed	1	0 (0)	-	19	7 (36.9)	1.33 (0.52-3.43)	235	111 (47.4)	1.34 (0.95-1.89)
Children 0-14 years									
No	835	138 (16.5)	1.00^	873	314 (36.0)	1.00^	784	339 (43.3)	1.00^
Yes	446	85 (19.1)	1.15 (0.85-1.57)	166	39 (23.7)	**0.55 (0.37-0.81)**	16	6 (34.6)	1.05 (0.37-3.03)
Children (Age categories)									
No children (0-14 years)	835	138 (16.5)	1.00^	873	314 (36.0)	1.00^	784	339 (43.3)	1.00^
Children below 5 but not between 6-14 years	133	21 (16.0)	0.99 (0.60-1.64)	11	2 (22.2)	0.47 (0.11-2.00)	6	1 (19.7)	0.68 (0.09-5.26)
Children between 6-14 but not below 5 years	257	50 (19.5)	1.16 (0.81-1.67)	147	35 (23.6)	**0.55 (0.36-0.83)**	10	5 (43.1)	1.27 (0.36-4.51)
Children below 5 and 6-14 years	56	14 (24.6)	1.54 (0.80-2.94)	8	2 (28.3)	0.64 (0.14-2.94)	0	-	-
Pets									
No	870	148 (17.1)	1.00^	668	208 (31.1)	1.00^	580	231 (39.7)	1.00^
Yes	411	75 (18.2)	1.14 (0.84-1.57)	371	146 (39.3)	**1.49 (1.14-1.95)**	220	114 (51.9)	**1.66 (1.20-2.30)**
Pets (Cats/Dogs)									
No pets	870	148 (17.1)	1.00^	668	208 (31.1)	1.00^	580	231 (39.7)	1.00^
Cat but no dog	95	13 (13.9)	0.85 (0.46-1.57)	90	30 (33.5)	1.22 (0.76-1.96)	92	43 (46.5)	1.30 (0.82-2.06)
Dog but no cat	249	47 (18.9)	1.17 (0.81-1.69)	207	82 (39.5)	**1.46 (1.05-2.03)**	87	45 (51.6)	**1.61 (1.00-2.59)**
Cats and dogs	67	14 (21.5)	1.52 (0.82-2.82)	74	34 (46.1)	**1.98 (1.21-3.24)**	41	27 (64.4)	**3.14 (1.56-6.34)**

° ORs were estimated using unconditional multiple logistic regression models after adjustment for sex, level of education and geographic area. Estimates in bold are statistically significant at 0.05 level.

^ Reference category.

Table S5. Distribution of 800 Italian participants aged ≥65 years, according to insufficient sleep duration (≤5 hours per night) and individual and other household characteristics. Corresponding odds ratios° (OR) and 95% confidence intervals (CI). Italy, 2019.

	N	N (%)	OR (95% CI)
Total	800	144 (18.0)	
Sex			
Male	346	52 (15.0)	1.00^
Female	454	92 (20.2)	1.42 (0.97-2.08)
Education level			
Low	536	105 (19.5)	1.00^
Intermediate	205	32 (15.8)	0.66 (0.42-1.02)
High	59	7 (11.7)	0.50 (0.22-1.14)
P for trend			0.098
Perceived socioeconomic class			
Low	179	55 (30.5)	1.00^
Intermediate	547	76 (13.8)	**0.40 (0.26-0.62)**
High	74	14 (18.4)	0.51 (0.25-1.06)
P for trend			0.073
Self-reported income			
Low	461	104 (22.6)	1.00^
Intermediate	230	31 (13.6)	**0.54 (0.34-0.87)**
High	109	9 (7.9)	**0.32 (0.15-0.68)**
P for trend			**0.003**
Smoking status			
Never smoker	536	109 (20.3)	1.00^
Current smoker	96	14 (15.0)	0.69 (0.36-1.30)
Ex smoker	168	21 (12.3)	**0.50 (0.29-0.86)**
Marital Status			
Married/Cohabiting	487	74 (15.1)	1.00^
Single	43	6 (14.1)	0.92 (0.37-2.30)
Divorced/Separated	35	9 (27.0)	**2.33 (1.00-5.41)**
Widowed	235	55 (23.4)	**1.64 (1.07-2.50)**
Children 0-14 years			
No	784	139 (17.8)	1.00^
Yes	16	5 (27.5)	2.98 (0.92-9.66)
Children (Age categories)			
No children (0-14 years)	784	139 (17.8)	1.00^
Children below 5 but not between 6-14 years	6	0 (0)	-
Children between 6-14 but not below 5 years	10	5 (43.1)	**4.49 (1.23-16.48)**
Children below 5 and 6-14 years	0	-	-
Pets			
No	580	87 (15.1)	1.00^
Yes	220	57 (25.7)	**1.83 (1.24-2.71)**
Pets (Cats/Dogs)			
No pets	580	87 (15.1)	1.00^
Cat but no dog	92	14 (15.4)	0.92 (0.49-1.72)
Dog but no cat	87	30 (34.4)	**2.83 (1.68-4.77)**
Cats and dogs	41	12 (30.3)	**2.45 (1.17-5.11)**

° ORs were estimated using unconditional multiple logistic regression models after adjustment for sex, level of education and geographic area.

Estimates in bold are statistically significant at 0.05 level.

^ Reference category

Chapter 1.2

The role of technology and social media use on sleep-onset difficulties among Italian adolescents: a cross-sectional study2

ABSTRACT

The use of technology and social media among adolescents are an increasingly prevalent phenomenon. However, there is a paucity of evidence on the relationship between frequency of the use of electronic devices and social media and sleep-onset difficulties among the Italian population. The aim of this study is to investigate the association between the use of technology and social media, including Facebook and YouTube, and sleep-onset difficulties among adolescents from Lombardy, the most populous region in Italy. The relationship between use of technology and social media, and sleep-onset difficulties was investigated. Data came from the 2013-2014 wave of the Health Behavior in School-aged Children (HBSC) survey, a school-based cross-sectional study conducted on 3172 adolescents aged 11 to 15 years in Northern Italy. Information was collected on difficulties in falling asleep over the last 6 months. We estimated the odds ratios (OR) for sleep-onset difficulties and corresponding 95% confidence intervals (CI), using logistic regression models after adjustment for major potential confounders. The percentage of adolescents with sleep-onset difficulties were 34.3% overall, 29.7% in boys and 39.2% in girls. It was 30.3% in 11yo, 36.2% in 13yo and 37.3% in 15yo adolescents. Sleep onset difficulties were more frequent among adolescents with higher use of electronic device: for general use (OR for highest vs lowest tertile of use= 1.50; 95% CI: 1.21-1.85); use for playing games (OR= 1.35; 95% CI: 1.11-1.64); use of online social network (OR for always vs never or rarely= 1.40; 95% CI: 1.09-1.81); and YouTube (OR= 2.00; 95% CI: 1.50-2.66). This study adds novel information about the relationship between sleep-onset difficulties and technology and social media in a representative sample of school-aged children from a geographical location that has not been included in studies of this type previously. Exposure to screen-based devices and online social media is significantly associated with adolescent sleep onset-difficulties. Interventions to create a well-coordinated parent and school centered strategy, thereby

---

[2] This was a joint project with Eugenio Santoro[a], Alessandra Lugo[a], Juan J Madrid-Valero[b], Simone Ghislandi[c], Aleksandra Torbica[c] and Silvano Gallus[a] on behalf of the HBSC Lombardy 2014 Regional Group and published in *Journal of Medical Internet Research*.

a. Istituto di Ricerche Farmacologiche Mario Negri IRCCS, Italy
b. University of Alicante, Spain
c. Universita Bocconi, Italy

increasing awareness on the unfavorable effect of evolving technologies on sleep among adolescents are needed.

INTRODUCTION

Use of screen-based devices such as smartphones, computer, and tablets (Hysing et al., 2015), and online activities on web platforms, notably Facebook and YouTube (Gradisar et al., 2013; Uhls et al., 2017), have grown in the recent past, particularly among adolescents. With 11 new social media users every second and each user spending more time online than before, the proportion of active social media users increased worldwide by more than 14% between January and December 2017 (Global Digital Report, 2018). Specifically, among US teenagers, 95% have access to smartphones with around 45% being online continuously (Anderson and Jiang, 2018). The pervasive use of electronic devices has been often associated with poor sleep behaviors in adolescents. High frequency use of social media in the hours before bed was shown to be associated with inadequate sleep (Reynolds et al., 2019; Shimoga et al., 2019). A review of the literature found that in 90% of 67 studies conducted on the issue, the use of screen-based devices adversely affected sleep outcomes including both sleep quantity and quality (Hale and Guan, 2015). This association has been explained via three main underlying mechanisms (LeBourgeois et al., 2017), including direct time displacement from sleep duration to media use (Cain and Gradisar, 2010), increased mental, emotional or psychological stimulation based on media content (Cain and Gradisar, 2010; Garrison et al., 2011), and effects of artificial light emitting devices on alertness, sleep health and circadian timing (Chang et al., 2015).

More limited data are available regarding the association between social media use and sleep in the Italian population. The few studies on the issue in the US and Northern European countries showed that the time spent on social networks, such as Facebook, as well as the time spent on the internet for playing games were significantly associated with sleep difficulties (Arora et al., 2014; Hawi et al., 2018; Sampasa-Kanyinga et al., 2018). The share of teenagers using Facebook has decreased rapidly in the USA with YouTube and other online social media being more commonly used (Anderson and Jiang, 2018). The majority of studies investigating the effect of media or internet use employ measures that do not isolate the effect of online social media. Moreover, studies have rarely looked at the specific role of YouTube, the most relevant social media today (Anderson and Jiang, 2018). In a recent study conducted by the Pew Research Center, 85% of the adolescents aged 13-17 reported using YouTube with boys reporting YouTube as their go-to platform (Anderson and Jiang, 2018). Given the increasing prevalence of sleep problems, the continuously evolving technology, and the shift in usage preferences in social media among adolescents (Ferrie et al., 2011; Rowshan Ravan et al., 2010; Twenge et al., 2017), the role of social media on sleep is yet to be fully elucidated.

In Italy, to our knowledge, only two cross-sectional studies have investigated the relationship between the use of electronic devices and sleep quality (Bruni et al., 2015; Ghekiere et al., 2019). The first study is based on a non-representative and relatively limited sample of adolescents (n = 850) and finds a negative relationship between the use of electronic device use at night and sleep quality (Bruni et al., 2015). The second and most recent study uses four waves of HBSC data with a large sample of adolescents from 33 countries including that of Italy to show an overall negative relationship between screen time and sleep-onset difficulties (Ghekiere et al., 2019). The study presents an international overview of the relationship of interest, and we contribute to this literature both with additional data on use of social media and a regional focus of the issue.

Secondary data from 2008-2009 use of time survey shows that Italians sleep less compared to the European average (8 hours 25 minutes) and the region under study, Lombardy scores even lower on time spend sleeping (8 hours 18 minutes). To understand the relationship between screen time and sleep-onset difficulties among the Lombard's, it should be acknowledged that the temporal profile of daily activities of Italians is different from the rest of the world (Boffi et al., 2015). For example, using a large (n = 11000) 2017 representative sample of Italian adolescent population, Smorti et al. 2019, show very early use of smartphone (at 11 years), internet dependency, and long hours spent online among the Italian adolescents with females being at higher risk of dysfunctional use (Smorti et al., 2019). Findings from a 2013-2014 dated survey of 2000 children aged 9-16 years from Denmark, Italy, the UK, Romania and Portugal shows that Italian adolescents are least likely to use internet at own home or school and most likely to use it on the go, indicating less supervised settings (Mascheroni et al., 2013). This also shows that Italian adolescents are more likely to use internet on their phones. Only 7% of the Italian children own or have a mobile phone that does not connect to the mobile internet whereas these numbers ranges from 23-41% in other countries included in the sample (Mascheroni et al., 2013). The report shows that Denmark is the only country that has successfully integrated internet use into daily school activities which is also a strategic site for internet awareness and safety campaigns.

Moreover, the consequences of internet dependency per se and other lifestyle habits among Italian adolescents such as unsafe sexual habits, perception of being overweight, cyber bullying and a range of other physical and mental health problems seems to be very high (Smorti et al., 2019). Further comparing Italy to data from 41 countries for adolescents aged 11-15 years shows Italy to be among the top 2 countries to report the highest proportion of adolescents reporting multiple health problems at least once in a week, positioned among the top 11 to 20 countries that has the highest proportion of adolescents skipping breakfast and among the top 5 countries to have the highest number of 15 year adolescents who smoke or drink alcohol at least once a week (Currie, 2016).

A report compiled by the World Health Organization shows that the prevalence of recommended levels of physical activity (at least 60 minutes per day) among adolescents (aged 11-15 years) is also lower in Italy (11%) compared to several countries including Denmark, Estonia, Latvia, Poland, Finland (13-24%), Luxembourg (30% for 11-14 years) and Finland (32% for 10-15 years) (World Health Organization, 2018). Using a 2010-2014 representative sample of Italian individuals, it was found that smartphone use reduces quality of face-to-face social interactions, with Lombardy displaying both an increasing smartphone penetration and the lowest fraction of people reporting to see their friends at least once a week (Rotondi et al., 2017).

Based on the previous considerations, our study aims to understand the relationship between frequency of use of technology (including computer, tablet, and smartphone) and social media, and sleep-onset difficulties among adolescents in a large sample from a geographical location that has not been included in studies of this type previously.

METHODS

Sample

The Health Behavior in School-aged Children Study (HBSC) study population includes school going children aged 11, 13 and 15 years, marking the onset of adolescence. For the present analysis, we used data from the HBSC wave of 2013-2014 from Lombardy. This is the most populated Italian region, including, in 2014, 10.0 million inhabitants (i.e., 17% of the 60.5 million Italian people), and 0.3 million adolescents aged 11-15 years (i.e., 17% of the overall Italian adolescents aged 11-15 years) (Istat statistics, 2014). Within the Italian HBSC, in Lombardy schools were over-sampled to obtain sufficient statistical power to derive precise frequency estimates at a regional level. In this specific region, a set of questions on technology and social media use was added to the standard national questionnaire to gain in-depth knowledge in this area of interest. The HBSC primarily employs cluster sampling to choose the subjects. Sampling is done in three stages with the first being a random selection of schools from the database of all public schools in Lombardy. Within each of the 142 participating schools, a total of 235 classes were selected, with a response rate of 89.4% on average yielding a total of 210 participating classes. The classes were stratified on the basis of age with the purpose of ensuring a geographical coverage that represents the actual distribution of the population of 11, 13 and 15-year old in the Lombardy region. Design weights were applied to adjust for differences in sampling frequencies such as cases where students in certain schools were more likely to be included in the survey. Finally, all students from the selected sample of classes were invited to complete the questionnaire. Additionally, post stratification weighting of sample was used to ensure representativeness of pupils with respect to characteristics including school denominations, school urban-rural classification and equal representation of boys and girls. This ensures the representativeness and generalizability of the results. Details on the sampling methodology of HBSC surveys have been previously described elsewhere (Currie et al., 2014). In Lombardy, a total of 3172 adolescents aged 11, 13 and 15 years filled a self-completion questionnaire in their schools, administered in the classrooms by teachers.

Measures

HBSC is an international school-based survey on adolescent health, administered every 4 years in several North American and European countries, including Italy, as a World Health Organization (WHO) collaborative cross-national study (AARØ et al., 1986; Roberts et al., 2007). The survey is based on a standardized research protocol containing a theoretical framework for all the procedures including a selection of survey topics, data collection, and analysis with the objective of securing cross-national comparable data (Roberts et al., 2009). The survey instruments consist of three sets of questions: core questions similar for all participating countries to create the international dataset, optional questions on specific topics, and country-specific questions of national importance (Roberts et al., 2007; Roberts et al., 2009).

Besides general information on socio-economic characteristics of their family, adolescents were asked to provide information on own anthropometric variables (weight and height) and selected lifestyle habits, including tobacco smoking and alcohol drinking. The questionnaire also included specific sections on wellbeing, health behaviors, and social context.

Outcome

Difficulty to fall asleep was assessed from the question 'over the last 6 months, how many times did you have difficulties in falling asleep?' The responses provided were 1) every day; 2) more than once a week; 3) about once a week; 4) once a month and 5) rarely or never. For the analysis, we dichotomized the variable into two categories: participants with no sleep-onset difficulties (i.e., adolescents with difficulties in falling asleep either once a month, rarely or never) and participants with sleep-onset difficulties (i.e., adolescents with difficulties in falling asleep either every day, more than once a week or about once a week). This question and its cut-off have already been used in previous studies with the same sample

(Kosticova et al., 2020). Difficulty falling asleep is one of the four items of the HBSC symptom checklist (HBSC-SCL) used to measure psychological health and was shown to have high external and internal construct validity in a sample of Canadian adolescents (Gariepy et al., 2016). The question about sleep-onset difficulties was also used to disentangle adolescents with severe sleep-onset difficulties (difficulty falling asleep every day) from adolescents with less severe sleep-onset difficulties (more than once a week, once a week, once a month, rarely or never).

Exposures

We consider four independent variables in our main analysis. Students were asked to report: i) the usage frequency of electronic devices including computer, tablets (iPad types) or smartphone for purposes such as doing homework, sending emails, chatting, tweeting, staying on Facebook or surfing online, during their free time; and ii) the use of electronic devices such as a computer, consoles, tablets (iPad type), smartphones or other devices for playing during their spare time. For both questions, the response scale provided in the questionnaire had 9 categories reported in hours per day namely a) never; b) half an hour; c)1 hour; d) 2 hours; e) 3 hours; f) 4 hours; g) 5 hours; h) 6 hours; i) 7 hours or more per day, respectively. The responses were reported for week and weekend days separately and have been shown to have considerable test-retest reliability (Bobakova et al., 2015; Liu et al., 2010) and acceptable criterion validity (Busschaert et al., 2015). Following previous studies using the same data, we recode them to continuous measures and construct a weighted average mean (= $(5 \times \text{weekday} + 2 \times \text{weekend day})/7$ days) (Rey-Lopez et al., 2010). From this weighted average, tertile categories of screen time were constructed based on the distribution of the sample (Table 1). Several studies have used tertiles of screen time in similar analyses (Bucksch et al., 2019; Gopinath et al., 2012).

Students were asked to respond to how often they go on: iii) social networking sites and iv) frequency of using YouTube when they were connected to the internet. For both questions, the response categories provided were 1) always; 2) often, 3) sometimes and 4) never. Likert scales of similar groupings are commonly used to indicate frequency of use of online social networking sites in several studies (Buja et al.. 2018; Li et al., 2010). For the analysis, based on the distribution of responses, we transformed these variables into three categories namely always, often, and sometimes or never. Information on use of social networking sites and YouTube were not collected from 11 year olds as social media use is restricted for children less than 13 years in Italy as a result of Children’s Online Privacy Protection Act of 1998 (COPPA, 1998). Body mass index (BMI; calculated from the self-reported weight and height in kg/m^2) was categorized by considering the age and sex-specific cutoff points adapted from Cole et al., 2000 (Cole et al., 2000). Validity of self-reported weight and height using the same data tested in several adolescent populations show that there is a small underestimation of weight with very little consequence on overall results (Aasvee et al., 2015; Elgar et al., 2015). The students were also asked the number of times they smoked a cigarette and drank alcohol in a month. Studies using the HBSC dataset either use a dichotomous categorization of alcohol and smoking frequency (Lee et al., 2020) or retain the original variable as such (Horn-Hofmann et al., 2018). However, studies also highlight the highly skewed nature of these two variables, thereby suggesting a dichotomous categorization of alcohol and smoking frequency into never and ever (at least one time in a month) (Lee et al., 2020).

Statistical analysis

Odds ratios (OR) of sleep-onset difficulties (and severe sleep-onset difficulties) and corresponding 95% confidence intervals (CI) were estimated by multiple logistic regression models, after adjustment for a number of potential confounders based on previous literature. The model was run with all cofounders included simultaneously in the regression. Considered covariates included categories of age, sex, mother's level of education (primary/secondary; high school/university; do not know), father's level of education (primary/secondary; high school/university; do not know), BMI (underweight/normal weight; obese/overweight), tobacco smoking (never; ever) and alcohol drinking (never; ever). Further, interaction tests were also performed between use of electronic device (and social media), and sex and age respectively. Statistical analyses were performed in SAS (SAS Institute, Cary NC) and STATA 15 (STATA Corp, College Station, TX, US).

Sensitivity Analysis

We do a range of sensitivity analyses. Previous studies indicate that screen time may differ between schooldays and weekend (Sigmundova et al., 2018). Hence, we construct a total measure of screen time (total time spent on electronic devices for general purposes and playing games) for weekdays and weekends separately and construct tertiles of total screen time from this unweighted measure of total screen time. For the second sensitivity analysis, we include school effects in the statistical models. It could be that going to certain schools might affect sleep outcomes differently for reasons including higher academic pressure or other school-related characteristics (Alsaggaf et al., 2016; Ramamoorthy et al., 2019). Finally, we test for different definitions of exposure to screen time and sleep-onset difficulties based on international health guidelines. We categorize frequency of overall use of electronic devices (total of time spent on electronic devices for general purposes and playing games) into exceeding 2 hours of daily screen exposure (American Academy of Pediatrics, 2001) or not and sleep-onset difficulties is defined as having difficulties in falling asleep every day or more than once a week (Ghekiere et al., 2019; Irish et al., 2014). Finally, family income is a potential confounder affecting the relationship between use of technology (and social media) and sleep problems (Kelly et al., 2018). However, given the unavailability of information on family income, we repeated our main analysis controlling for the perceived economic status of the family. The responses are categorized as very good; quite good; average; bad and very bad with 43 missing observations. Given the very few observations for the 'very bad' category, we collapsed responses for bad and very bad as one category in our analysis.

RESULTS

Out of 3172 adolescents, 3151 (99.3%) had available information on sleep-onset difficulties. The percentage estimates of adolescents with sleep-onset difficulties were 34.3% of the total sample and 30.3%, 36.2%, 37.3% for 11yo, 13yo and 15yo participants respectively, (*p*=.001). Female participants presented high levels of sleep-onset difficulties (39.2%) than male participants (29.7%; *p* <.001).

Multivariate ORs for the relationship between adolescent sleep-onset difficulties and use of selected technologies and social media are presented in table 1.

[Table 1 about here]

Sleep-onset difficulties were more common among adolescents with a frequent use of electronic devices (OR for ≥ 2.2 vs. <0.9 hours/day = 1.50; 95% CI: 1.21–1.85; *p*<.001), use of electronic devices for playing games (OR for ≥1.8 vs <0.8 hours/day = 1.35; 95% CI: 1.11-1.64; *p*=.003), online social network use (OR for always vs never or rarely = 1.40; 95% CI: 1.09-1.81; *p*=.008) and YouTube use (OR for always vs never or rarely = 2.00; 95% CI: 1.50-2.66; *p*<.001). Table 1 also shows the sex-specific ORs for the relationship between adolescent sleep quality and the use of selected technologies and social media. OR for sleep-onset difficulties associated with the use of YouTube was consistently large and significant in both sexes (OR for always vs. never or rarely = 2.04; 95% CI: 1.35-3.10; *p*=.001 for males and OR = 1.91; 95% CI: 1.29-2.84; *p*=.001 for females). When considering electronic device use for general purpose and use of social networking sites, the OR was significant only for males (OR for high vs. low use= 1.74; 96% CI: 1.30-2.32; *p*<.001 and OR= 1.69; 95% CI: 1.17-2.45;*P*=.005, respectively), whereas when considering use of electronic device for playing games, the OR was significant only for females (OR= 1.46; 95% CI: 1.08-1.97; *p*=.007).

Table 2 shows the multivariate ORs for the relationship between sleep-onset difficulties and use of selected technologies and social media stratified by age. Sleep-onset difficulties were more common among 11- year old school aged children with frequent use of electronic devices for general purpose (OR for ≥ 2.2 vs. <0.9 hours/day = 1.48; 95% CI: 1.03–2.13; *p*=.034) and electronic devices for playing games (OR for ≥1.8 vs <0.8 hours/day = 1.55; 95% CI: 1.10-2.18; *p*=.013). Among 13-year old adolescents, Sleep-onset difficulties were more frequent among those with the use of electronic devices for general purposes (OR for ≥ 2.2 vs. <0.9 hours/day = 1.85; 95% CI: 1.29–2.66; p<.001) and YouTube (OR for always vs. never or rarely = 1.88; 95% CI: 1.26-2.81; p=.002). Finally, sleep-onset difficulties were more common among 15-year old adolescents with a frequent use of social networking sites (OR for always vs. never or rarely = 1.58; 95% CI: 1.08-1.78; p=.020) and use of YouTube (OR for always vs. never or rarely = 2.10; 95% CI: 1.39-3.18; p<.001).

[Table 2 about here]

In the analysis of severe sleep-onset difficulties among adolescents, l0.3% of adolescents had severe sleep onset difficulties; this percentage was 11.1% among 11yo; 10.7% among 13yo and 8.9% among 15 yo (p=0.24). Severe sleep onset difficulties were more common among adolescents with frequent use of electronic devices for general purposes (OR for ≥ 2.2 vs. <0.9 hours/day = 1.53; 95% CI: 1.10–2.13; p=.011), electronic devices for playing games (OR for ≥1.8 vs <0.8 hours/day = 1.73; 95% CI: 1.16–2.58; p=.001), online social network (OR for always vs. never or rarely = 1.73; 95% CI: 1.16-2.58; p=.007) and YouTube (OR for always vs. never or rarely = 3.18; 95% CI: 1.89-5.34; p<.001) (***Table* 3**).

[Table 3 about here]

Finally, we did not find any statistically significant interaction between use of electronic device, (and social media) and sex and age respectively (Results not shown). Table S1, S2 and S3 reports the results of three sensitivity analyses, respectively. Analysis of screen exposure by weekday and weekend (Table S1), including school effects in the statistical models (Table S2) and using a different definition of screen exposure and sleep-onset difficulties (Table S3). In all cases similar results were found in comparison with the previous analyses. Our results remain robust after controlling for perceived economic status (Table S4). Moreover, respondents perceiving worser economic status also reported higher sleep onset difficulties (Results not shown).

DISCUSSION

Our data confirms that exposure to screen-based devices, online social networking sites and video-sharing platform's is significantly associated with difficulties in falling asleep among adolescents. Our study adds to the literature by extending the findings to Lombardy, a geographical region that has not been included in studies of this type previously. Moreover, for the first time we show that use of YouTube, an online video sharing platform is related with sleep-onset difficulties among adolescents.

Our findings are in broad agreement with those from a meta-analysis of 20 cross-sectional studies (N=125,198), showing a 53% higher odds of poor sleep quality among those with consistent bedtime media use (Carter et al., 2016). Previous literature also shows that the use of computers and smartphones among adolescents were associated with day time sleep-related behaviors such as sleepiness and fatigue (Shochat et al., 2010), shorter sleep duration, later bed time and unfavorable changes in sleep habits over time (Lemola et al., 2015). Use of PC in particular has been shown to be associated with reduced sleep duration, poor sleep, and sleep-onset latency (Dorofaeff and Denny., 2016; Hysing et al., 2015; Punamaki et al., 2017). Smart phones are easy to carry around, easier to take to bed making it an imperative tool for adolescents. Both overall and late night cellphone use among adolescents in several countries has been associated with several sleep difficulties (Amra et al., 2017; Munezawa et al., 2011). Use of cellphones particularly for night-time texting in particular was associated with insufficient sleep (Troxel et al., 2015). In Italy, over 72% of adolescents aged 11–17 years access Internet through smartphones, with talks on a potential 'no mobile phone phobia' law being drafted in Italy, particularly aiming at the younger generation (Il Messaggero, 2019).

Given that adolescents use not just one device, but multiple devices at the same time, our results confirm previous findings on the relationship between multiple device use and sleep problems (Hysing et al., 2015). Our results are in line with the recent findings from the international study on use of technology and sleep-onset difficulties among adolescents using the same data and variables and extend the findings to a regional level (Ghekiere et al., 2015). Additionally, we also confirm an exposure-response relationship between the use of electronic devices and sleep difficulties (Hysing et al., 2015; Twenge et al., 2017).

Given the multifunctionality of electronic devices such as for doing homework or social networking, it is important to have specific information on the kinds of activities carried out on these platforms. Information on frequency of use of each individual social networking sites is relatively less available and more valuable (Shimoga et al., 2019). Information on the kind of social media used also signals the nature of the activity such as following vloggers on YouTube or connecting with peers on Facebook. We find that use of online social media and use of YouTube among adolescents is associated with higher odds of sleep-onset difficulties. A cross-sectional survey based on 467 Scottish adolescents (11-17 years old) found that overall social media use, particularly at night, was associated with poor sleep quality (Woods and Scott, 2016). In addition to sleep-onset difficulties as an outcome, various studies use standardized measurements of sleep including PSQI to find significant associations with social media use in general or specific sites such as Facebook (Levenson et al. 2016; Wolniczak et al., 2013). Using data from Millennium Cohort Study, Kelly et al., 2019 finds significant associations between higher frequency of weekday social media use (reported in hours) and sleep problems including inadequate sleep, sleep latency and sleep disruption (Kelly et al., 2019). Prior studies showed that the adverse effect of social media use on sleep is mediated by cognitive activation (Harbard et al., 2016) and through a behavioral effect of fear of missing out leading to delayed sleep-onset (Scott and Woods, 2018). Finally, our findings suggest a dose-response effect of online social media use on sleep-onset difficulties. These findings are similar to those from a recent cross-sectional study on Canadian adolescents linking cumulative effect of social media use on reduced sleep duration (Sampasa-Kanyinga et al., 2018).

The use of YouTube has been steadily growing every year since its launch in 2005 to become the most popular platform in 2018 among young adults in the US (Anderson and Jiang, 2018). To our knowledge, our study was the first to investigate the association between YouTube use and sleep-onset difficulties among adolescents. YouTube is a video intensive platform with some unique characteristics including asymmetric relationships, greater self-disclosure, higher personal revelation and developing parasocial relationships all contributing to supplanting real relationships (de Bérail et al., 2019). YouTube addiction follows the same model of internet addiction with the frequent use of YouTube being reinforced by the one-sided virtual relationships with YouTubers (de Bérail et al., 2019). We found a significant association between YouTube use and sleep-onset difficulties, overall and consistently in strata of sex and age. The same relationship was observed also when considering a stricter sleep definition (i.e., severe sleep difficulties).

Potential limitations of our analysis include methodological drawbacks inherent to the cross-sectional study design. First, cross-sectional designs have no dimension of time and hence it is impossible to conclude a causal relationship. Given that reverse causality (i.e., adolescents use more electronic devices and social media because of difficulties in falling asleep) could not be ruled out, our findings should be confirmed by longitudinal data.

Secondly, our study uses a self-reported measure of sleep latency and screen-time. Self-reported sleep measures in children need to be accurate for precise identification of sleep problems (Erwin and Bashore, 2017). Although some studies showed an overestimation of sleep latency among both adolescents and adults (Chambers, 1994; Wolfson et al., 2003), self-reported sleep latency and other sleep measures obtained some consensus with objective measurements (Wetter and Young, 1994; Zinkhan et al., 2014). Our measure of sleep latency may also suffer from recall bias given the 6-month period retrospective nature of the question. However, this question has already been used and validated in different studies (Gariepy et al., 2016; Ghekiere et al., 2019; Pallesen et al., 2008). These types of bias could be overcome in the future by the use of apps tracking sleep quality, being this one of the most used health apps among patients and citizens (Mosconi et al., 2019). Also, in the current study only sleep-onset difficulties are analysed, whereas various aspects of sleep including sleep duration, sleep quality, other sleep disturbances and their differences between week and weekend days should be explored in future studies. Although frequency of use of electronic devices and other social media has been self-reported, they have been shown to have considerable reliable and validity (Bobakova et al., 2015; Busschaert et al., 2015). Self-reported measures of screen time are an inexpensive and easier means of data collection in large samples and give us detailed information on the context of the activity.

Third, we do not have data at the national level, but only for Lombardy. Although the data is geographically limited, our study population was oversampled to obtain a representative sample of adolescents aged 11, 13, and 15 years at a regional level. Moreover, Lombardy represents the most populous and richest region in Italy, accounting for one-sixth of the country's population and one-fifth of the gross domestic product (GDP). Finally, the selection of a region from Northern Italy, being more digitally connected than Southern Italy, fits the goal of our study.

In addition, our sample includes only young adolescents aged 11-15 years old enrolled in schools. Although findings cannot be generalized to older adolescents, our results are crucial given that younger adolescents are at higher risk of compulsive internet use (Espinoza and Juvonen, 2011). The strengths of our study include the relatively large sample size, the availability of information to control for socioeconomic characteristics and risk behavior of adolescents, thereby strengthening the internal validity of our study and the representativeness of the sample at a regional level. Moreover, to our knowledge, this is the first study to examine the relationship between screen-time, social media use and sleep-onset difficulties in a geographical location that has not been studied before in the context of this research question. Exploiting the data on frequency of use of YouTube 'a parasocial networking site', we also study for the first time the relationship between YouTube use and sleep-onset difficulties among adolescents.

CONCLUSION

Our study shows that exposure to screen-based devices and online social media is associated with adolescent sleep-onset difficulties. These findings may have important public health implications. In fact, they claim for interventions and guidelines to minimize excess use of technology among adolescents. In line with the recommendations proposed by several pediatric societies, including the Italian Pediatric Society, it is advised to limit the use of tablets or other electronic devices after dinner to ensure better sleep (Società Italiana di Pediatria. 2017). The findings also call on social media industries to be an equal player in protecting the health and wellbeing of young users by considering them as potential users of their services (Childrens Commissioner, 2018). Moreover, interventions should include contributions from health professionals and educators to create a well-coordinated parent centered strategy to increase awareness on evolving technologies, social media platforms and the deleterious effects on sleep among adolescents.

BIBLIOGRAPHY

AARØ, L. E., *et al.* 1986. Health behaviour in schoolchildren A WHO cross-national survey*: A presentation of philosophy, methods and selected results of the first survey. *Health Promotion International*, 1(1), 17-33.

Aasvee, K., *et al.* 2015. Validity of self-reported height and weight for estimating prevalence of overweight among Estonian adolescents: the Health Behaviour in School-aged Children study. *BMC Res Notes,* 8, 606.

Alsaggaf, M. A., *et al.* 2016. Sleep quantity, quality, and insomnia symptoms of medical students during clinical years. Relationship with stress and academic performance. *Saudi Med J,* 37(2), 173-182.

American Academy of Pediatrics: Children, adolescents, and television. 2001. *Pediatrics,* 107(2), 423-426.

Amra, B., *et al.* 2017. The association of sleep and late-night cell phone use among adolescents. *J Pediatr (Rio J),* 93(6), 560-567.

Anderson, M. and Jiang, J. 2018. Teens, social media & technology 2018. *Pew Research Center,* 31, 2018.

Arora, T., *et al.* 2014. Associations between specific technologies and adolescent sleep quantity, sleep quality, and parasomnias. *Sleep Med,* 15(2), 240-247.

Bobakova, D., *et al.* 2015. Test-retest reliability of selected physical activity and sedentary behaviour HBSC items in the Czech Republic, Slovakia and Poland. *Int J Public Health,* 60(1), 59-67.

Boffi, M., Colleoni, M. and Del Greco, M. 2015. Night-time Hours and Activities of the Italians. *Articulo-Journal of Urban Research,* (11).

Bruni, O., *et al.* 2015. Technology Use and Sleep Quality in Preadolescence and Adolescence. *J Clin Sleep Med,* 11(12), 1433-1441.

Bucksch, J., *et al.* 2019. Associations between perceived social and physical environmental variables and physical activity and screen time among adolescents in four European countries. *Int J Public Health,* 64(1), 83-94.

Buja, A., *et al.* 2018. Problematic social networking site usage and substance use by young adolescents. *BMC Pediatr,* 18(1), 367.

Busschaert, C., *et al.* 2015. Reliability and validity of three questionnaires measuring context-specific sedentary behaviour and associated correlates in adolescents, adults and older adults. *Int J Behav Nutr Phys Act,* 12, 117.

Cain, N. and Gradisar, M. 2010. Electronic media use and sleep in school-aged children and adolescents: A review. *Sleep Med,* 11(8), 735-742.

Cao, W., *et al.* 2020. The psychological impact of the COVID-19 epidemic on college students in China. *Psychiatry research*, 112934.

Carter, B., *et al.* 2016. Association Between Portable Screen-Based Media Device Access or Use and Sleep Outcomes: A Systematic Review and Meta-analysis. *JAMA Pediatr*, 170(12), 1202-1208.

Chambers, M. J. 1994. Actigraphy and insomnia: a closer look. Part 1. *Sleep*, 17(5), 405-408; discussion 408-410.

Chang, A. M., *et al.* 2015. Evening use of light-emitting eReaders negatively affects sleep, circadian timing, and next-morning alertness. *Proc Natl Acad Sci U S A*, 112(4), 1232-1237.

Cole, T. J., *et al.* 2000. Establishing a standard definition for child overweight and obesity worldwide: international survey. *BMJ*, 320(7244), 1240-1243.

Childrens Commissioner, 2017. Life in 'Likes': Children's Commissioner Report into Social Media Use among 8–12 Year Olds. Children's Commissioner.

Currie, C., *et al.* 2014. Health Behaviour in School-aged Children (HBSC) study protocol: background, methodology and mandatory items for the 2013/14 survey.

Currie, D., 2016. *Growing up unequal: gender and socioeconomic differences in young people's health and well-being. Health Behaviour in School-aged Children (HBSC) study: international report from the 2013/2014 survey*. World Health Organization.

de Bérail, P., Guillon, M. and Bungener, C. 2019. The relations between YouTube addiction, social anxiety and parasocial relationships with YouTubers: A moderated-mediation model based on a cognitive-behavioral framework. *Computers in Human Behavior*, 99, 190-204.

Dipendenza da Internet, arriva la legge per curarla. [online], 2019. Available from: https://www.ilmessaggero.it/pay/edicola/dipendenza_internet_ultime_notizie_legge-4632455.html [Accessed 4 May 2020].

Dorofaeff, T. F. and Denny, S. 2006. Sleep and adolescence. Do New Zealand teenagers get enough? *J Paediatr Child Health*, 42(9), 515-520.

Elgar, F. J., *et al.* 2005. Validity of self-reported height and weight and predictors of bias in adolescents. *J Adolesc Health*, 37(5), 371-375.

Erwin, A. M. and Bashore, L. 2017. Subjective Sleep Measures in Children: Self-Report. *Front Pediatr*, 5, 22.

Espinoza, G. and Juvonen, J. 2011. The pervasiveness, connectedness, and intrusiveness of social network site use among young adolescents. *Cyberpsychol Behav Soc Netw*, 14(12), 705-709.

Ferrie, J. E., *et al.* 2011. Sleep epidemiology--a rapidly growing field. *Int J Epidemiol*, 40(6), 1431-1437.

Gariepy, G., *et al.* 2016. Validity and reliability of a brief symptom checklist to measure psychological health in school-aged children. *Child Indicators Research*, 9(2), 471-484.

Garrison, M. M., Liekweg, K. and Christakis, D. A. 2011. Media use and child sleep: the impact of content, timing, and environment. *Pediatrics*, 128(1), 29-35.

Ghekiere, A., *et al.* 2019. Trends in sleeping difficulties among European adolescents: Are these associated with physical inactivity and excessive screen time? *Int J Public Health,* 64(4), 487-498.

Global Digital Report 2018. Available from: https://wearesocial.com/blog/2018/01/global-digital-report-2018 [Accessed 4 May 2020].

Gopinath, B., *et al.* 2012. Physical activity and sedentary behaviors and health-related quality of life in adolescents. *Pediatrics,* 130(1), e167-174.

Gradisar, M., *et al.* 2013. The sleep and technology use of Americans: findings from the National Sleep Foundation's 2011 Sleep in America poll. *J Clin Sleep Med,* 9(12), 1291-1299.

Hale, L. and Guan, S. 2015. Screen time and sleep among school-aged children and adolescents: a systematic literature review. *Sleep Med Rev,* 21, 50-58.

Harbard, E., *et al.* 2016. What's Keeping Teenagers Up? Prebedtime Behaviors and Actigraphy-Assessed Sleep Over School and Vacation. *J Adolesc Health,* 58(4), 426-432.

Hawi, N. S., Samaha, M. and Griffiths, M. D. 2018. Internet gaming disorder in Lebanon: Relationships with age, sleep habits, and academic achievement. *J Behav Addict,* 7(1), 70-78.

Horn-Hofmann, C., *et al.* 2018. The relationship between pain severity and alcohol use among school-aged children and adolescents: the moderating role of drinking motives. *Pain medicine,* 19(2), 284-296.

Hysing, M., *et al.* 2015. Sleep and use of electronic devices in adolescence: results from a large population-based study. *BMJ Open,* 5(1), e006748.

Irish, L. A., *et al.* 2014. A 24-hour approach to the study of health behaviors: temporal relationships between waking health behaviors and sleep. *Ann Behav Med,* 47(2), 189-197.

Istat Statistics. 2014. *Resident municipal population by age, sex and marital status.*

Kelly, Y., *et al.* 2018. Social media use and adolescent mental health: Findings from the UK Millennium Cohort Study. *EClinicalMedicine,* 6, 59-68.

Kosticova, M., Husarova, D. and Dankulincova, Z. 2020. Difficulties in Getting to Sleep and their Association with Emotional and Behavioural Problems in Adolescents: Does the Sleeping Duration Influence this Association? *Int J Environ Res Public Health,* 17(5).

LeBourgeois, M. K., *et al.* 2017. Digital Media and Sleep in Childhood and Adolescence. *Pediatrics,* 140(Suppl 2), S92-S96.

Lee, J., *et al.* 2020. Exploring Sex Differences in the Association between Bullying Involvement and Alcohol and Marijuana Use among US Adolescents in 6th to 10th Grade. *Substance Use & Misuse,* 55(8), 1203-1213.

Lemola, S., *et al.* 2015. Adolescents' electronic media use at night, sleep disturbance, and depressive symptoms in the smartphone age. *J Youth Adolesc,* 44(2), 405-418.

Levenson, J. C., *et al.* 2016. The association between social media use and sleep disturbance among young adults. *Prev Med,* 85, 36-41.

Li, J., *et al.* 2016. Validation of the Social Networking Activity Intensity Scale among Junior Middle School Students in China. *PLoS One,* 11(10), e0165695.

Liu, Y., *et al.* 2010. Test-retest reliability of selected items of Health Behaviour in School-aged Children (HBSC) survey questionnaire in Beijing, China. *BMC Med Res Methodol,* 10, 73.

Mascheroni, G., *et al.* 2013. Mobile internet access and use among European children: initial findings of the Net Children Go Mobile project.

Mosconi, P., *et al.* 2019. Use of Health Apps and Wearable Devices: Survey Among Italian Associations for Patient Advocacy. *JMIR Mhealth Uhealth,* 7(1), e10242.

Munezawa, T., *et al.* 2011. The association between use of mobile phones after lights out and sleep disturbances among Japanese adolescents: a nationwide cross-sectional survey. *Sleep,* 34(8), 1013-1020.

Pallesen, S., *et al.* 2008. Time trends in sleep-onset difficulties among Norwegian adolescents: 1983--2005. *Scand J Public Health,* 36(8), 889-895.

Punamaki, R. L., *et al.* 2007. Use of information and communication technology (ICT) and perceived health in adolescence: the role of sleeping habits and waking-time tiredness. *J Adolesc,* 30(4), 569-585.

Ramamoorthy, S., *et al.* 2019. Effect of stress on sleep hygiene among school going adolescents in Chennai. *J Family Med Prim Care,* 8(9), 2917-2920.

Rey-Lopez, J. P., *et al.* 2010. Sedentary patterns and media availability in European adolescents: The HELENA study. *Prev Med,* 51(1), 50-55.

Reynolds, A. C., *et al.* 2019. Impact of high-frequency email and instant messaging (E/IM) interactions during the hour before bed on self-reported sleep duration and sufficiency in female Australian children and adolescents. *Sleep Health,* 5(1), 64-67.

Roberts, C., *et al.* 2007. Measuring the health and health behaviours of adolescents through cross-national survey research: recent developments in the Health Behaviour in School-aged Children (HBSC) study. *Journal of Public Health,* 15(3), 179-186.

Roberts, C., *et al.* 2009. The Health Behaviour in School-aged Children (HBSC) study: methodological developments and current tensions. *Int J Public Health,* 54 Suppl 2, 140-150.

Rotondi, V., Stanca, L. and Tomasuolo, M. 2017. Connecting alone: Smartphone use, quality of social interactions and well-being. *Journal of Economic Psychology,* 63, 17-26.

Rowshan Ravan, A., *et al.* 2010. Thirty-six-year secular trends in sleep duration and sleep satisfaction, and associations with mental stress and socioeconomic factors--results of the Population Study of Women in Gothenburg, Sweden. *J Sleep Res,* 19(3), 496-503.

Sampasa-Kanyinga, H., Hamilton, H. A. and Chaput, J. P. 2018. Use of social media is associated with short sleep duration in a dose-response manner in students aged 11 to 20 years. *Acta Paediatr,* 107(4), 694-700.

Scott, H. and Woods, H. C. 2018. Fear of missing out and sleep: Cognitive behavioural factors in adolescents' nighttime social media use. *J Adolesc,* 68, 61-65.

Shimoga, S. V., Erlyana, E. and Rebello, V. 2019. Associations of Social Media Use With Physical Activity and Sleep Adequacy Among Adolescents: Cross-Sectional Survey. *J Med Internet Res*, 21(6), e14290.

Shochat, T., Flint-Bretler, O. and Tzischinsky, O. 2010. Sleep patterns, electronic media exposure and daytime sleep-related behaviours among Israeli adolescents. *Acta Paediatr*, 99(9), 1396-1400.

Sigmundova, D., *et al.* 2016. Weekday-weekend patterns of physical activity and screen time in parents and their pre-schoolers. *BMC Public Health*, 16, 898.

Smorti, M., *et al.* 2019. Adolescent selfie: an Italian Society of Paediatrics survey of the lifestyle of teenagers. *Ital J Pediatr*, 45(1), 62.

Società Italiana di Pediatria. 2017. Un bambino su quattro soffre di disturbi del sonno. Le 10 regole d'oro per far dormire i piccoli, i più grandi (e anche mamma e papà). Available from: https://sip.it/2019/03/15/un-bambino-quattro-soffre-disturbi-del-sonno-le-10-regole-doro-far-dormire-piccoli-piu-grandi-anche-mamma-papa/ [Accessed 4 May 2020].

Troxel, W. M., Hunter, G. and Scharf, D. 2015. Say "GDNT": Frequency of Adolescent Texting at Night. *Sleep Health*, 1(4), 300-303.

Twenge, J. M., Krizan, Z. and Hisler, G. 2017. Decreases in self-reported sleep duration among U.S. adolescents 2009-2015 and association with new media screen time. *Sleep Med*, 39, 47-53.

Uhls, Y. T., Ellison, N. B. and Subrahmanyam, K. 2017. Benefits and Costs of Social Media in Adolescence. *Pediatrics*, 140(Suppl 2), S67-S70.

Wetter, D. W. and Young, T. B. 1994. The relation between cigarette smoking and sleep disturbance. *Prev Med*, 23(3), 328-334.

Wolfson, A. R., *et al.* 2003. Evidence for the validity of a sleep habits survey for adolescents. *Sleep*, 26(2), 213-216.

Wolniczak, I., *et al.* 2013. Association between Facebook dependence and poor sleep quality: a study in a sample of undergraduate students in Peru. *PLoS One*, 8(3), e59087.

Woods, H. C. and Scott, H. 2016. #Sleepyteens: Social media use in adolescence is associated with poor sleep quality, anxiety, depression and low self-esteem. *J Adolesc*, 51, 41-49.

World Health Organization 2018. Physical actvity country fact sheets.

Zinkhan, M., *et al.* 2014. Agreement of different methods for assessing sleep characteristics: a comparison of two actigraphs, wrist and hip placement, and self-report with polysomnography. *Sleep Med*, 15(9), 1107-1114.

TABLES

Table 1. Distribution of 3172 adolescents with sleep difficulties overall and by technology and social media use. Corresponding odds ratios° (OR) and 95% confidence intervals (CI) for the full sample and sex-specific samples. Lombardy, 2014.

	Total			**Males**			**Females**		
	N*	**Difficulty falling asleep %**	**Difficulty falling asleep OR (95% CI)**	**N**	**Difficulty falling asleep %**	**Difficulty falling asleep OR (95% CI)**	**N**	**Difficulty falling asleep %**	**Difficulty falling asleep OR (95% CI)**
Electronic device use for general purpose									
1st tertile (<0.9 hours/day)	1010	28.9	1.00^	557	26.6	1.00^	453	31.8	1.00^
2nd tertile (0.9-2.1 hours/day)	1018	32.6	1.15 (0.94-1.42)	544	24.6	0.89 (0.66-1.19)	474	41.8	1.46 (1.08-1.97)
3rd tertile (≥2.2 hours/day)	1059	41.3	1.50 (1.21-1.85)	483	39.1	1.74 (1.30-2.32)	576	43.1	1.30 (0.96-1.77)
p for trend			<.001			<.001			.094
Electronic device use for playing games									
1st tertile (<0.8 hours/day)	1149	32.7	1.00^	406	26.9	1.00^	743	35.9	1.00^
2nd tertile (0.8-1.7 hours/day)	853	33.5	1.21 (0.98-1.49)	519	28.7	1.13 (0.83-1.55)	334	41.0	1.24 (0.93-1.67)
3rd tertile (≥1.8 hours/day)	1093	36.7	1.35 (1.11-1.64)	665	32.2	1.25 (0.92-1.68)	428	43.7	1.46 (1.11-1.91)
p for trend			.003			.148			.007
Use of social networking sites									
Never/Rarely	701	32.0	1.00^	397	26.2	1.00^	304	39.5	1.00^
Often	650	36.9	1.16 (0.91-1.48)	345	31.6	1.30 (0.92-1.84)	305	43.0	1.04 (0.74-1.47)
Always	599	42.7	1.40 (1.09-1.81)	255	34.9	1.69 (1.17-2.45)	344	48.6	1.23 (0.87-1.73)
p for trend			.008			.005			.250
Use of YouTube									
Never/Rarely	412	27.4	1.00^	209	21.1	1.00^	203	34.0	1.00^
Often	854	36.5	1.55 (1.18-2.04)	434	29.3	1.41 (0.94-2.13)	420	44.1	1.66 (1.14-2.42)
Always	683	43.1	2.00 (1.50-2.66)	354	36.7	2.04 (1.35-3.10)	329	50.0	1.91 (1.29-2.84)
p for trend			<.001			.001			.001

* The sum does not add up to the total because of some missing values and exclusion of age 11 for social media measures.

° ORs were estimated using unconditional multiple logistic regression models after adjustment for age and sex of the child, mothers and fathers' highest level of education, tobacco and alcohol use among adolescents and BMI.

^ Reference category

Table 2. Distribution of 3172 adolescents with sleep difficulties overall and by technology and social media use. Corresponding odds ratios (OR) and 95% confidence intervals (CI) for age-specific samples. Lombardy, 2014.

	11 years old adolescents			13 years old adolescents			15 years old adolescents		
	N	Difficulty falling asleep %	Difficulty falling asleep OR (95% CI)	N	Difficulty falling asleep %	Difficulty falling asleep OR (95% CI)	N	Difficulty falling asleep %	Difficulty falling asleep OR (95% CI)
Electronic device use for general purpose									
1st tertile (<0.9 hours/day)	569	28.7	1.00^	264	26.9	1.00^	177	32.8	1.00^
2nd tertile (0.9-2.1 hours/day)	334	28.4	1.02 (0.73-1.42)	370	34.9	1.41 (0.98-2.05)	314	34.4	0.98 (0.65-1.49)
3rd tertile (≥2.2 hours/day)	244	36.5	1.48 (1.03-2.13)	417	44.1	1.85 (1.29-2.66)	398	41.2	1.15 (0.77-1.72)
p for trend			.034			<.001			.505
Electronic device use for playing games									
1st tertile (<0.8 hours/day)	468	27.6	1.00^	315	34.9	1.00^	366	37.4	1.00^
2nd tertile (0.8-1.7 hours/day)	328	29.9	1.09 (0.76-1.56)	292	35.3	1.34 (0.93-1.94)	233	36.5	1.21 (0.83-1.76)
3rd tertile (≥1.8 hours/day)	351	34.5	1.55 (1.10-2.18)	447	38.3	1.26 (0.90-1.77)	295	37.0	1.28 (0.90-1.83)
p for trend			.013			.185			.170
Use of social networking sites (Facebook)									
Never/Rarely				440	33.0	1.00^	261	30.3	1.00^
Often				323	37.5	1.13 (082-1.57)	327	36.4	1.22 (0.84-1.78)
Always				291	41.2	1.30 (0.93-1.83)	308	44.2	1.58 (1.08-2.32)
p for trend						.126			.020
Use of YouTube									
Never/Rarely				199	27.1	1.00^	213	27.7	1.00^
Often				451	36.4	1.55 (1.04-2.31)	403	36.7	1.50 (1.02-2.21)
Always				404	41.6	1.88 (1.26-2.81)	279	45.2	2.10 (1.39-3.18)
p for trend						.002			<.001

*The sum does not add up to the total because of some missing values and exclusion of age 11 for social media measures.

°ORs were estimated using unconditional multiple logistic regression models after adjustment for age and sex of the child, mothers and fathers' highest level of education, tobacco and alcohol use among adolescents and BMI. Estimates in bold are statistically significant at 0.05 level.

^ Reference category

Table 3. Distribution of 3172 adolescents with severe sleep difficulties overall and by technology and social media use. Corresponding odds ratios° (OR) and 95% confidence intervals (CI). Lombardy, 2014.

	N	Difficulty falling asleep	
		%	**OR (95% CI)**
Electronic device use for general purpose			
1st tertile (<0.9 hours/day)	1010	8.7	1.00^
2nd tertile (0.9-2.1 hours/day)	1018	9.2	1.08 (0.77-1.51)
3rd tertile (≥2.2 hours/day)	1059	13.0	1.53 (1.10-2.13)
p for trend			.011
Electronic device use for playing games			
1st tertile (<0.8 hours/day)	1149	9.2	1.00^
2nd tertile (0.8-1.7 hours/day)	853	9.0	1.11 (0.78-1.57)
3rd tertile (≥1.8 hours/day)	1093	12.6	1.73 (1.27-2.35)
p for trend			.001
Use of social networking sites (Facebook)			
Never/Rarely^	701	8.0	1.00^
Often	650	8.6	1.01 (0.66-1.54)
Always	599	13.9	1.73 (1.16-2.58)
p for trend			.007
Use of YouTube			
Never/Rarely^	412	5.3	1.00^
Often	854	8.0	1.42 (0.83-2.43)
Always	683	15.4	3.18 (1.89-5.34)
p for trend			<.001

*The sum does not add up to the total because of some missing values.

°ORs were estimated using unconditional multiple logistic regression models after adjustment for age and sex of the child, mothers and fathers' highest level of education, tobacco and alcohol use among adolescents and BMI.

^ Reference category

SUPPLEMENTARY TABLES

Table S1. Distribution of 3172 adolescents with sleep-onset difficulties overall and by technology use during school days and weekends. Corresponding odds ratios° (OR) and 95% confidence intervals (CI) for the full sample and sex-specific samples. Lombardy, 2014.

	Total			Males			Females		
	N*	Difficulty falling asleep		N	Difficulty falling asleep		N	Difficulty falling asleep	
		%	OR (95% CI)		%	OR (95% CI)		%	OR (95% CI)
Total screen time-weekday									
1st tertile (<2 hours/day)	1217	29.6	1.00^	592	26.0	1.00^	625	33.1	1.00^
2nd tertile (2-3.9 hours/day)	853	34.5	1.21 (0.98-1.48)	463	27.4	0.99 (0.74-1.34)	390	42.8	1.42 (1.06-1.90)
3rd tertile (≥4 hours/day)	1102	39.3	1.41 (1.16-1.72)	582	35.3	1.50 (1.14-1.98)	520	43.7	1.30 (0.98-1.72)
p for trend			.001			.004			.072
Total screen time-weekend									
1st tertile (<2.5 hours/day)	1160	29.9	1.00^	561	27.4	1.00^	599	32.2	1.00^
2nd tertile (2.5-4.9 hours/day)	872	32.7	1.19 (0.96-1.46)	475	23.9	0.84 (0.62-1.14)	397	43.2	1.61 (1.20-2.16)
3rd tertile (≥5 hours/day)	1140	40.0	1.51 (1.24-1.83)	601	36.5	1.48 (1.13-1.94)	539	44.0	1.47 (1.11-1.94)
p for trend			<.001			.004			.007

*The sum does not add up to the total because of some missing values and exclusion of age 11 for social media measures. °ORs were estimated using unconditional multiple logistic regression models after adjustment for age and sex of the child, mothers and fathers' highest level of education, tobacco and alcohol use among adolescents and BMI.

^ Reference category

Table S2. Odds ratios° (OR) and 95% confidence intervals (CI) of sleep onset difficulties when including indicators for school of the adolescent in addition to the baseline model. Lombardy, 2014

	OR (95% CI)
Electronic device use for general purpose	
1st tertile (<0.9 hours/day)	1.00^
2nd tertile (0.9-2.1 hours/day)	1.14 (0.91-1.42)
3rd tertile (≥2.2 hours/day)	1.44 (1.16-1.81)
p for trend	.001
Electronic device use for playing games	
1st tertile (<0.8 hours/day)	1.00^
2nd tertile (0.8-1.7 hours/day)	1.15 (0.92-1.44)
3rd tertile (≥1.8 hours/day)	1.28 (1.03-1.58)
p for trend	.024
Use of social networking sites	
Never/Rarely^	1.00^
Often	1.16 (0.89-1.52)
Always	1.47 (1.11-1.95)
p for trend	.007
Use of YouTube	
Never/Rarely^	1.00^
Often	1.63 (1.21-2.21)
Always	2.15 (1.57-2.95)
p for trend	<.001

°ORs were estimated using multiple logistic regression models after adjustment for age and sex of the child mothers and fathers highest level of education, tobacco and alcohol use among adolescents, and BMI and school indicators of the participant.

^Reference category

Table S3. Distribution of 3172 adolescents with sleep-onset difficulties overall and by technology use using alternative definitions of exposure and outcome variables. Corresponding odds ratios (OR) and 95% confidence intervals (CI) for the full sample and sex-specific samples. Lombardy, 2014.

	Total			**Males**			**Females**		
	N*	**Difficulty falling asleep**		**N**	**Difficulty falling asleep**		**N**	**Difficulty falling asleep**	
		%	**OR (95% CI)**		**%**	**OR (95% CI)**		**%**	**OR (95% CI)**
Total time spend on electronic devices									
Less than or equal to 2 hours/day	1251	17.9	1.00^	615	14.1	1.00^	636	21.5	1.00^
Exceeding 2 hours per day	1921	24.7	1.52 (1.24-1.87)	1022	19.6	1.46 (1.08-1.98)	899	30.5	1.55 (1.18-2.04)
Use of social networking sites									
Never/Rarely	701	19.8	1.00^	397	15.1	1.00^	304	26.0	1.00^
Often	650	22.3	1.10 (0.83-1.47)	345	18.6	1.29 (0.85-1.98)	305	26.6	1.00 (0.68-1.48)
Always	599	28.7	1.43 (1.07-1.90)	255	20.0	1.57 (1.00-2.44)	344	35.2	1.38 (0.94-2.02)
p for trend			.015			.048			.097
Use of YouTube									
Never/Rarely	412	17.7	1.00^	209	12.4	1.00^	203	23.2	1.00^
Often	854	22.1	1.32 (0.96-1.84)	434	16.6	1.25 (0.75-2.09)	420	27.9	1.39 (0.91-2.13)
Always	683	28.4	1.86 (1.34-2.59)	354	21.8	1.90 (1.14-3.17)	329	35.6	1.79 (1.15-2.78)
p for trend			<.001			.014			.010

*The sum does not add up to the total because of some missing values and exclusion of age 11 for social media measures. °ORs were estimated using unconditional multiple logistic regression models after adjustment for age and sex of the child, mothers and fathers' highest level of education, tobacco and alcohol use among adolescents and BMI.
^Reference category

Table S4. Odds ratios° (OR) and 95% confidence intervals (CI) of sleep onset difficulties controlling for perceived economic status of the family in addition to the baseline model. Lombardy, 2014

	OR (95% CI)
Electronic device use for general purpose	
1st tertile (<0.9 hours/day)	1.00^
2nd tertile (0.9-2.1 hours/day)	1.15 (0.94-1.42)
3rd tertile (≥2.2 hours/day)	1.49 (1.21-1.84)
p for trend	<.001
Electronic device use for playing games	
1st tertile (<0.8 hours/day)	1.00^
2nd tertile (0.8-1.7 hours/day)	1.21 (0.98-1.49)
3rd tertile (≥1.8 hours/day)	1.33 (1.09-1.62)
p for trend	.005
Use of social networking sites	
Never/Rarely^	1.00^
Often	1.17 (0.91-1.49)
Always	1.43 (1.11-1.84)
p for trend	.006
Use of YouTube	
Never/Rarely^	1.00^
Often	1.57 (1.19-2.08)
Always	2.02 (1.51-2.70)
p for trend	<.001

°ORs were estimated using multiple logistic regression models after adjustment for age and sex of the child, mothers and fathers' highest level of education, tobacco and alcohol use among adolescents, and BMI and perceived economic status of the family.

^Reference category

Chapter 1.3

Does infant sleep deprivation affect parity? Evidence from a longitudinal study[3]

ABSTRACT

In this article, we investigate if parents' experience of physical and socio-psychological aspects of childrearing, as measured by infant night wakings is important for further parity progression. Using Avon Longitudinal Study of Parents and Children (ALSPAC) cohort data, our analysis draws on information from parents and children during the prenatal period until 4 years of age of child. Around 5500 first time mothers reported information on the sleep patterns of their first child and subsequent pregnancies. Four cross sectional logistic regressions at four time points of age of child were used to analyze the relationship between infant night wakings and fertility of the mother. Odds ratio (OR) and corresponding 95% confidence intervals (CIs) for the models are reported. An increase in the night wakings of the first child predicts a decreased likelihood of intention to have a second child and this relationship is significant only until approximately 3 years of age of the first child (OR = 0.36; 95% CI = 0.14-0.90, OR = 0.62; 95% CI = 0.41-0.94 and OR = 0.45; 95% CI = 0.27-0.74 for more than two night wakings at 6, 18 and 30 months of age respectively). Given the role of infant night wakings in influencing fertility, policy discussion on parental wellbeing must consider this overlooked dimension of physical and emotional change experienced by new parents. Moreover, fertility can benefit from incorporating infant sleep patterns as a possible determinant of parity progression.

[3] This is a joint project with Simone Ghislandi[a] and we provide the latest version of the working paper.
[a.] Bocconi University, Italy.

INTRODUCTION

Night waking is a common sleep problem during infancy and childhood, affecting 14% of infants and 17% of preschool aged children (Martin et al., 2007; Sadeh et al., 2009). This, when accompanied by the inability of the infants to help themselves back to sleep without parental care, has adverse physical, mental (depression) and wellbeing effects on parents, mainly through parental sleep deprivation and exhaustion (Martin et al., 2007; Sadeh et al., 2009). Thus, infant night waking is a potential family stressor with the direct and swift ability to influence individual and familial decisions, including the decision to have another child.

Recently, there has been a renewed focus on literature that suggests the relevance of physical and socio-psychological experience of child rearing in further parity progression (Callan, 1985; Cartwright, 1976). A qualitative study on parenthood experiences, particularly on the 'baby stage' of parenting shows that phases of conception until early parenthood influenced their decisions on having the next child (Newman, 2008). To our knowledge, the first empirical study to investigate this issue used longitudinal data from Germany to show that a drop-in parental well-being surrounding first birth was associated with a decrease in the probability of having another child (Margolis and Myrskyla, 2015). One characteristic of parenting experience, that has a huge influence on parental well-being, is the sleep cycle of the infant (Karraker and Young, 2007). Disrupted sleep is a common complaint among new parents (Haig, 2014). In this article, we investigate how the physical and socio-psychological aspect of parenthood, measured by the variation in infant night wakings of the first child, can influence further parity progression.

Focusing on infants from birth until 4 years of age, we argue that night waking has adverse effects on parental wellbeing, further translating into a determining factor for fertility. This effect is considered to be more important for first time parents, as they do not have previous experience of parenthood (Dyck, 1990). Presser (2001) considers first time parenting as a shock for mothers who would not have anticipated the physical and emotional costs of child-rearing (Presser, 2001). If the firsthand experience of child rearing is more positive than anticipated, then parents are more likely to have a second child compared to parents who underestimate the task of childrearing (Margolis and Myrskyla, 2015; Newman, 2008).

Infant night wakings, mechanisms and fertility decisions

We draw on different framework and mechanisms from diverse strands of literature to understand the relationship between infant night wakings and parity progression.

Previous literature on parenting experiences and parity progression mentions learning frameworks as a means to theorize this relationship (Margolis and Myrskyla, 2015; Newman, 2008). People repeat behaviors that resulted in pleasant experiences and avoid activities that resulted in negative experiences in the past (Skinner, 1965). Past experiences also act as signal which lets an individual to be informed of her ability to handle similar situations in the future (Ajzen, 1991). Moreover, this concept was adapted to fertility intentions stating that child rearing experiences from the first child feeds as inputs into attitudes and norms about the potential consequences of having a second child, which further influences subsequent fertility decisions (Ajzen and Klobas, 2013).

Earlier studies on this topic are mostly based on anecdotal or qualitative evidence (Cartwright, 1976; Newman, 2008; Presser, 2001). Newman (2008) conducted in-depth interviews and incorporates various aspects of positive and negative parenthood experiences in shaping family size (Newman, 2008). The author focuses on difficulties in conception, pregnancy, birth and first year of parenting, all of which may contribute to what is known as 'parity progression hurdle'. In the early parenthood stage, exhaustion was a frequent complaint among new parents with broken sleep being a contributing factor (Giallo et al., 2011; Loutzenhiser et al., 2015; Martin et al., 2007). Fatigue arising from disrupted sleep can also result in a reduced frequency of sexual activities and lower levels of marital satisfaction leading to lower fertility (Elek et al., 2002; Favez et al., 2006; Germo et al., 2007; Newman, 2008). This could also lead to a conscious decision to both delay or forego more children (Newman, 2008).

In addition to these factors, several biological mechanisms can also help to understand this relationship. Breastfeeding has been long recognized to increase lactational amenorrhea, resulting in infertility for a minimum of 6 months since birth (Kennedy, 1988; Van der Wijden et al., 2003). This effect of delaying ovulatory cycles is most powerful during night-time suckling (Howie and McNeilly, 1982; Short, 1984) and can be attributed to the higher frequency of suckling bouts at night (Short, 1987) resulting in night-time suckling to delay the ovulation of mothers (Jones and da Costa, 1987).

The relation between child sleep and parity progression and its duration in time has received attention also from the evolutionary biology literature (Haig, 2014; Trivers, 1974). Parent offspring conflict theory states that the mother will start reproducing again only when the offspring is capable to survive without full-time maternal care. A newborn's goal is to maximize his/her survival, whereas the mothers' goal is to maximize her reproductive success, with the common aim of increasing the number of surviving descendants. As the child grows older, the child is less dependent on the mother and the benefit to the child from delayed birth of a sibling is less than the cost to the mother. At this point, the mother decides to have the next child. This period of parental investment is an opportunity cost wherein 'anything done by the parent for the offspring increases the offspring's chance of surviving while decreasing the parent's ability to invest in another offspring'(Trivers, 1974). Particularly, the parental investment here refers to maternal fatigue, which is an infant's strategy to extend the inter-birth interval (Haig, 2014). Maternal fatigue is also attributed to arise from night-time suckling, which prolongs lactational amenorrhea, which in turn delays fertility. In other words, infant night waking is considered to be an evolutionary adaptation (Stromberg, 2014). Following these approaches, one would expect the negative impact of child sleep on parity progression to fade out as time goes by. In other words, parents should keep no strong memory of past night awaking events. On the other hand, it is also possible that child sleeping problems might "scar" parents, affecting their decision making for longer periods.

The notion of 'scarring effects' is derived from the literature on economics and psychology. 'Scarring effects' refers to the process by which the evaluation of current situations depends on the retrospective evaluations of similar situations in the past (Kahneman, 1994). The idea is that past memories 'scar' because they 'scare' individuals, affecting their anticipated memories of the future, hence the expectations on which they base their decisions (De Witte, 2016). The literature on scarring effects has been applied mostly in the fields of unemployment and wellbeing (Clark et al., 2001; Heckman and Borjas, 1980). In the present analysis we will borrow from this stream of literature in order to investigate the duration of the sleeping effects on fertility.

More specifically, in the present study we investigate whether mothers' experience of infant night wakings of first child predicts the likelihood to have another child. The first step of the analysis investigates the contemporaneous relationship between infant night waking and parity progression. We examine this relationship for four consecutive time points (from birth to 4 years of age) at which information on infant night wakings and mothers' pregnancy is available. The chosen age group is also appropriate for our study given that it is during these ages that sleep patterns of the child can have the most effect on parents. Given all the mechanisms of how the relationship between infant night wakings and fertility may actualize, we hypothesize a negative effect of infant night wakings on the probability to go on to have one more child.

If lack of regular sleep affects fertility, it becomes important to understand the dynamic of this relation. In the second step of the analysis, we focus on the timing and the duration of this influence. In particular, we consider two directions in which the timing might be of interest. First the effect might fade off as time goes by. This is the case when sleeping problems persist over the months but their negative impact on fertility becomes weaker and weaker with the passing of time. Parents might indeed adapt to the situation and develop more or less effective coping strategies in order to "normalize" the situation (Damato and Zupancic, 2009; Newman, 2008; NHS UK). Second, parents might experience scarring effects of past infant night wakings on current decision of the mother to have another child. In particular, we test whether past infant night awakenings (say, at 6 months) have an effect on the decision to have a second child, given the current pattern of infant sleep (say at 1.5 years). Considering there is no literature on scarring effects of sleep or any related mechanism, it is ambiguous as to what we should expect to find in this direction.

METHODS

Data and Participants

We use data from Avon Longitudinal Study of Parents and Children (ALSPAC), consisting of a total of 15445 pregnant women, with expected dates of delivery between April 1991 and December 1992. This resulted in a total of 13978 children after excluding perinatal deaths, multiple births and children who died within 1 year of birth. Our analysis draws on information from parents and children during the prenatal period until 4 years of age of child. ALSPAC children are very similar to UK average in terms of anthropometric measures at birth, at 1 and 2 years of age (Barazzetta and Ghislandi, 2017). The average age of mothers is 28 years, with 43% of them having only one child. About half of the sample is male and only 5% of the sample accounts for non-white ethnicity. The study website contains other detailed information on variables, questionnaires and updated findings using ALSPAC data (Avon Longitudinal Study of Parents and Children). Figure 1 shows a flow chart depicting the sample selection used in our analysis.

[*Figure 1 Approximately Here*]

Key Variables

The key outcome variable is the subsequent pregnancy since the birth of the study child (Figure 2). We particularly focus on intention of parity progression from the first to the second child. Pregnancy is indicated by the question in the mother's questionnaire "if mother became pregnant after child was born". This question asked in different waves also captures the duration of the spell between the first child (born between April 1991 and December 1992) and a second one. More specifically, at each round of interview mothers would have to answer the question "if mother became pregnant after study child is of X months of age", where X in our data can be 8, 21, 33 and 47 months. The possible responses for each of the questions are "Yes, but not affected by the pregnancy; Yes, bit affected; Yes, moderately affected; Yes, affected a lot; and No, did not become pregnant". In our analyses, we create dichotomous variables for the mother becoming pregnant whenever the response was anything but "No" to the above questions.

[*Figure 2 Approximately Here*]

Our key explanatory variable is the number of night awakenings of the infant, measured at 6, 18, 30 and 42 months of age (Figure 3). Infant night awakenings are reported as a continuous measure. In our analyses, we construct a categorical measure of night waking with "0" corresponding to children with no night wakings, "1" corresponding to children with one- or two-night wakings and "2" corresponding to children with more than two-night wakings. We follow previous literature on the definition of infant night waking to construct the above categorical measure (Blair et al., 2012). In this way, we alleviate concerns of recall bias among mothers to correctly remember the times an infant woke at night and we can also differentiate between less frequent infant night wakings and frequent night wakings (Blair et al., 2012).

[Figure 3 *Approximately Here*]

Other Variables

We look at previous literature to choose covariates that are associated with the determinants of fertility behavior. Overall, our controls belong to a range of factors associated with parental, pre-natal and child characteristics. Related to parental characteristics, we control for mothers age at birth in three groups (<21, 21-34, >35 years), binary indicators for mother and father being employed across the four time periods, being divorced, or separated and if the pregnancy was intentional or not.

Socio economic status is measured using household income and highest parental education. Net household income per week is reported as a categorical variable of 5 bands (<100, 100-199, 200-299, 300-399, >400) at 3 and 4 years of child's age. We convert each income categories at 3 and 5 years to income figures in pounds by taking the midpoint of the bands. The resulting income figures are then averaged for the two time periods and divided into four quartiles to construct one permanent measure of income. Highest parental education is reported in 4 categories: cse/vocational (lowest), O-levels, A-levels, and degree (highest). Based on a previous study examining the socioeconomic determinants of sleep quality and quantity among children, we also include material deprivation as a control variable (Barazzetta and Ghislandi, 2017). We define an index of material deprivation in terms of access to household facilities including sole use of indoor toilet, sole use of bath or shower, working telephone, presence of damp or mould in the house and if the house is adequately warm during the winter. The index is constructed at 8, 21 and 33 months of child's age and can take values between 0 (not deprived in any item) to 5 (deprived in all times).

Prenatal characteristics included maternal depression, sleep and a proxy of parenting style. Prenatal depression is measured using the depression scale of Crown Crisp Experiential Index (CCEI) and is a continuous measure, ranging from 0-16. Prenatal sleep is measured using a categorical variable assessing the easiness of initiating sleep (very often, often, not very often and never). Both measures were assessed at 18 and 32 weeks of gestation and we take an average of the two weeks for our analysis. Prenatal personality of the mother was assessed at 18 months of gestation using the Interpersonal Sensitivity Measure (IPSM). IPSM consists of 36 items were scored on a 4-point Likert scale across 5 subscales namely, *Interpersonal Awareness, Need for Approval, Separation Anxiety, Timidity and Fragile Inner-Self* (Boyce and Parker, 1989). In our analysis, we carry out a principal component analysis of the five subscale scores and use the score based on the first component, following the Kaiser rule of retaining only components above the Eigen value of one. A higher score implies higher interpersonal sensitivity and is a predictor of dysfunctional and overprotecting parenting style (Otani et al., 2009a; Otani et al., 2009b). Maternal sleep quantity at night, measured at 2, 8, 21 and 30 months as a robustness check (<6 hours; 6-7 hours and >7 hours).

Birth outcomes include indicators for low birth weight (<2500g) and preterm birth (gestation week <37 weeks). Child characteristics include the sex of the child, an indicator for if the child was ever breastfed until 15 months of age and an indicator for parent-reported health status of the child (healthy or not healthy).

Statistical Analysis

To examine if infant night wakings influences fertility decisions, we use four cross sectional logistic regressions for each of the four time points (see Appendix). We estimate the odds ratio (OR) and corresponding 95% confidence intervals (CIs) for the cross-sectional logistic models. Mothers', who become pregnant at each point in time, are excluded from the subsequent analysis. Note that the estimation of the model for different periods in time provides evidence also about a possible fading off of the simultaneous effects.

In order to examine if there are any scarring effects of infant sleep, we estimate a logistic regression wherein we look at the effect of past infant night wakings on current parity progression while adjusting for current infant night wakings (see Appendix). Particularly, we use two cross sectional regressions for pregnancies at 8 to 21 months and 18 to 33 months. With respect to the simultaneous analysis, we thus exclude two periods: pregnancies at 0 to 8 months, due to absence of any prior sleep measures, and pregnancies at 30 to 42 months, due to insignificant effects for a contemporaneous relationship in the previous analysis (results shown later). The analysis of scarring effects also excludes mothers who became pregnant before 8 months in the first regression and those who became pregnant before 18 months in the second regression.

Following the literature on scarring effects in economics, we also include a "habituation" effect, which is simply the interaction between the simultaneous and the cumulative sleep variables used in the scarring equation. This additional variable is intended to measure the extent to which the longer cumulative exposures to sleep problems reduce the negative impact of simultaneous night wakening. It represents a further, more rigorous attempt to test whether, independently from the age of the child, families are able to adapt to persistent sleeping problems of their child, by gaining experience and developing effective coping strategies.

RESULTS

Table 1 presents the descriptive statistics for key variables separately for the full sample, for the sample that remain at one child and for the sample that go on to have a second pregnancy, respectively. In Column 5, we report the result (p-value) of the appropriate test (either a t-test or a Chi squared test) for the differences between those with one pregnancy and those with two or more. The proportion of 'more than 2 infant night wakings' is higher among the sample of mothers who stayed at parity one compared to the sample who went on to have a second pregnancy for three out of four time periods (7.8% vs 5.8% at 18 months, 7.3% vs 4.4% at 30 months and 4.1% vs 2.5% at 42 months) in our study. The proportion of infants that woke up 'one or two times' at night is higher among those who stayed at parity one than those who went on to a second pregnancy at 8 months (11.8% vs 10.8%). This preliminary evidence seems to suggest that sleep and parity progression might be positively linked.

Regarding the covariates, most of the major well-known demographic regularities are confirmed within our sample too. In particular, women that went on to have a second pregnancy are relatively younger, less likely to be employed, more likely to be separated or divorced than the women that stayed at parity one. Economically, mothers that are having a second pregnancy live in richer households.

[*Table 1 Approximately Here*]

Similarly, health and demographic covariates relate to parity progression as expected (Table 2). The proportion of mothers with low birth weight and pre-term infants were higher among those who stayed at parity one compared to those who had second pregnancies. There were no clear trends for children with ill-health among the two groups. Prenatal depression scores were higher among those who stayed at parity one, whereas these differences were not significant for prenatal maternal sleep and interpersonal sensitivity scores. The proportion of mothers with more than 7 hours of sleep per night was higher among those who went on to have a second pregnancy until 21 months since birth, with significant group differences. This descriptive evidence provides some preliminary information which will basically be confirmed by the results from logistic regression.

[*Table 2 Approximately Here*]

Table 3 examines whether infant night wakings at 6, 18, 30 and 42 months predicts contemporaneous intentions for parity progression to second birth, proxied by pregnancies, using multivariate logistic regressions. Model 1 in table 3 adjusts for simultaneous covariates including mothers and father's employment status, an indicator for divorced or separated, household income, material deprivation and child health status. Model 2 adjusts for predetermined covariates including mothers age at delivery, highest parental education, an indicator for intentional pregnancy, child gender, child ever breast fed, low birth weight, preterm birth, a score for prenatal depression, prenatal sleep in categories and prenatal personality score for mother in addition to simultaneous covariates included in model 1.

Overall, we find that infant night waking one or two times do not predict parity progression and effects mostly come from infant night waking more than two times at 6, 18 and 30 months from the birth of the child. Infant night waking more than two times at 6 months, 18 and 30 months predicts a decreased likelihood of having another child ($p < 0.05$), whereas infant night waking at 42 months do not predict parity progression (OR = 0.36; 95% CI = 0.14-0.90, OR = 0.62; 95% CI = 0.41-0.94 and OR = 0.45; 95% CI = 0.27-0.74 for more than two night wakings at 6, 18 and 30 months respectively). Overall, this evidence provides a picture of a U-shaped relation between child sleep and parity progression, with the probability of having a second child being heavily influenced by interrupted child sleep at two years and a half after the birth (i.e., the likelihood of having a second child reduces by 18% at 30 months).

[*Table 3 Approximately Here*]

Regarding the covariates, Table S1 in the supplementary file presents the full model with all the covariates. Most of the relations considered in the descriptive statistics are confirmed in this setting. Among the covariates, only few of them retain significance in the multivariate regression. In general, older, divorced or separated and employed mothers were less likely to go on to have a second pregnancy. At least one parent having a degree as the highest education was related to higher odds of mothers moving on to a second pregnancy. Although, differently from what noted in the descriptive statistics, there were no clear trends in the relationship between household income and fertility, having an employed partner was associated with higher odds of moving on to second pregnancy. None of the birth outcomes and child characteristics, including gender, and indicators of at-birth and simultaneous measures of health status retained statistical significance. Prenatal depression at 18 and 33 months of gestation were associated with lower odds of moving on to second pregnancy whereas having an intentional pregnancy and breast feeding practices for first birth were related to higher odds of moving on to second pregnancy.

Table 4 shows the results for scarring and habituation effects. The two panels in the table refer to two different spells from the first births. Two main results are confirmed across models and with different time horizon. First, current sleep problems, especially when severe, negatively affect the probability of having a second child. This reproduces the evidence provided in Table 3. A second regularity observed across our models is children's sleep problems do not seem to have a scarring effect. The stress associated to the lack of sleep seems to vanish with time, confirming the predictions of the evolutionary biology framework. When it comes to habituation (i.e., the extent to which previous experience of a negative situation can help adapting to the persistency of such situation), results are not so clear-cut. Having experienced night awakenings in the past seem to be associated to a habituation effect in the short run (8-21 months) but not in the longer term (18-33 months). This point towards a non-monotonic adaptation effect: while for low enough levels of experienced lack of sleep coping strategies can be effective, adaptive efforts lose efficacy for longer enough periods of sleep deprivation.

[*Table 4 Approximately Here*]

Additional evidence

In this section we provide some additional evidence to show that results are robust. First, we provide a robustness check of the presented results. Second, we investigate the role of mother sleep as the most likely mechanism through which child sleep can be related any parental decision and outcome.

Given the availability of an additional dataset of the same participants with not-strictly overlapping information, we carry out some robustness check using different alternatives of the main dependent and independent variable (Table 5). The two dependent variables used in this analysis is the (1) total number of pregnancies and (2) total number of children, both measured at 4 years since the birth of first child. The independent variable is an indicator for number of night wakings of the study child measured at 18 months and recoded to none, one or two and more than two night wakings respectively. The covariates used differ slightly from the previous analyses given that we have a different version of the dataset used for the main results (described in the table 5).

Model 1 of table 5 presents the unadjusted relationship between infant night wakings at 1.5 years of age and the two outcome variables. Model 2 of table 5 gives the results after adjusting for relevant covariates. These additional results confirm our main findings. Compared to experiencing no infant night wakings, being exposed to more than two infant night wakings per night reduces the probability of having both an additional pregnancy ($\beta = -0.11$, $p=0.019$) and an additional child ($\beta = -0.22$, $p=0.001$).

[*Table 5 Approximately Here*]

Table S2 presents the results from a fixed effects model, taking into account the panel structure of the data. The results from such a specification tells us if a change in infant sleep disruption is associated with a change in mothers' decision to have another child. Having another child may indicate not just a second pregnancy but a third or fourth since we are not able to drop mothers reporting a second pregnancy when moving on to the subsequent time points. Results show that a change (an increase) in infant sleep disruption (more than 2 night wakings) is related to a decrease in the mother's intention to have another child. Findings from the fixed effects specification confirm the robustness of our results to any time-fixed omitted variable bias.

If child sleep affects parents' fertility outcomes, it most likely happens through a reduction in parents' sleep (Haig, 2014; Newman, 2008). Indeed, one might notice that the quantity and the quality of the sleep of the mother has not been included as an explanatory variable for the analysis. This choice was motivated by two main reasons. First, the data did not allow us to simultaneously relate the sleep of the mother and the sleep of the child because the two variables are observed at different times. Second, it is more than plausible that child sleep could affect mother's decisions and wellbeing only if it affects her sleep in first place. In other words, mother sleep would be a mediator and, as such, it would get on the way of the estimation of interest. In order to test, although approximately, this hypothesis, we split our sample in two groups according to whether mothers sleep more or less than or equal to 7 hours per night and run our analysis separately for each subgroup.[4]

Results are reported in Table 6. Looking at the group of mothers sleeping well (>7 hours per night), it is evident that child sleep seems not to be related with parity progression. This is indeed consistent with a picture in which the intensity of the child sleeping problems should be strong enough to reduce mothers sleep in order to affect her decision and wellbeing. On the other hand, however, the impact of child sleep on the probability of having a second child conditioning on mothers sleeping less than 7 hours is still negative, suggesting that, for this subgroup, mother and child sleep are complementary to a certain extent.

[*Table 6 Approximately Here*]

DISCUSSION

Exploiting a sample of 5585 mother-child pairs, regression models are used to identify a possible relation between sleep patterns of the first child and mother's decision to have a second child. Our analysis builds on a specific aspect of parenting experience in understanding why parents who have one child do not go on to have another one. In our article, we argue that parents' first experience of childrearing, specifically infant night wakings, shapes their decision to have another child. We only include mothers with one living child as fertility differences across the developed world are largely driven by differences in parity progression from one to two (Bavel and Różańska-Putek, 2010). Our results show that having infants with more than two night wakings per night reduces the likelihood of parent's intentions to have a second child.

Infant waking at night causes fragmented sleep among parents leading to fatigue and exhaustion. This childrearing experience may induce negative feelings about having a next child. Therefore, disrupted sleep of first time parents, although rarely explored, is an important determinant of parity progression. Prior information on parenthood at an individual level can help to-be parents prepare better to welcome parenthood. This can help to reduce the adverse shocks associated with child rearing that many young couples face in their transition to parenthood (Presser, 2001). Prior information can also help put fatigue

[4] Higher maternal age at birth, higher parental education, higher prenatal depression, pre-term birth and having an ill child is negatively correlated with mothers sleep quantity. Higher income is negatively correlated with mothers sleep quantity in the first two years since child's birth (0-2 years) but positively correlated in the next two years (2-4 years) (results not shown)..

associated with pregnancy, birth, and child rearing at a perspective. Recent studies have already shown that infant night waking following birth of a first child causes sleep deprivation and sleep dissatisfaction in both father and mother (Richter et al., 2019). Although to our knowledge no studies have investigated the causal effect of poor sleep on separation or divorce, other related works show that sleep disruption during transition to parenthood severely affects marital satisfaction (Medina et al., 2009; Troxel et al., 2007). Overall, fatigue associated with disrupted sleep is an important factor that can affect all dimensions of everyday life including work and marital satisfaction thereby inhibiting parity progression. All of these factors come under the umbrella of parental wellbeing in general (Newman, 2008) and our study argues that among the many elements and complex interplay of factors that influences parity progression, disrupted sleep among parents is also crucial.

Another finding from our study is that the adverse effect of infant night wakings on parity progression fades off after 30 months or approximately two years and a little more. It could be that at this age, the child is more capable when compared to a new-born infant and is not as dependent on the mother as before and hence of less concern to the mother. These findings also resonate well with the parent offspring theory which states that the mother is more likely to have more children after she completes the period of parental investment or when the infant ceases the conflict via employing the psychological weapons (night waking) to maximize its survival. In light of our results, we could speculate that the parent-offspring conflict for parental investment via night wakings could end around approximately 2 years of the age of child as beyond this point infant night waking do not have significant effects on parity progression. We also find that there are no scarring effects of infant sleep problems in the past on mother's intention to have a second child. The effect comes from the disrupted sleep patterns of the child in the contemporaneous period. If the first child has sleep problems only in the past and is followed by a better sleeping pattern in the near future, then our results show that the probability of having a second child is not scarred by the earlier experiences of disrupted sleep patterns. Therefore, the current positive parenting experience triumphs over the previous negative experience on the decision to have a second child. In other words, our results point to the evidence that the role of infant night wakings in parity progression do not have memory effects.

These findings on scarring effects are important from a learning perspective on the possibilities of improving parenting experiences and wellbeing. Given the insights on memory formation and interpretation of past experiences, new parents can be made aware of the learning mechanisms to construct highlights and positive experiences in order to make parenting situations better. This is consistent with the literature on child development suggesting that parents of sleep-disturbed children, with time are more likely to limit parental nighttime involvement and to encourage infants to learn self-soothing (Sadeh et al., 2007; Tikotzky and Sadeh, 2009).

If child sleep can affect parents' fertility, it seems natural to assume that this happens mainly through a reduction of parents' sleep patterns. More specifically, we show that sleep-troubled mothers represent a necessary condition for child sleep affecting parity progression. In other words, when child sleep problems do not get to influence mothers sleep patterns, they do not manage to climb their way up to influence parents' decisions and wellbeing.

The availability of time-variant measures of infant sleep in our analysis helps us to link the effect of a specific dimension associated with child-rearing (fragmented sleep) to corresponding pregnancy outcomes. This is an advantage in our study when compared to existing studies that look at the effect of drop in overall wellbeing associated with first childbirth in general and does not specify what actually constitutes this drop in wellbeing (Margolis and Myrskyla, 2015). Although such specific factors associated with a drop in wellbeing has been studied qualitatively, never has it been studied quantitatively (Newman, 2008). Moreover, our study also opens up a new stream of potential research on the role of infant sleep in particular, which could affect a lot of family specific outcomes beyond fertility decision, making such as partnership dissolution, parental employment (already studied) and parental mental and physical health including many others.

The relevance of our findings at macro level depends on the size of the childless population and second birth rate (Margolis and Myrskyla, 2015). In the UK, the level of childlessness has seen an increase with childlessness among women born in 1944 and 1971 to be 11% and 18% respectively (Berrington et al., 2015; Office for National Statistics, 2016). Nevertheless, recent statistics from 1970 British cohort data shows evidence for declining childlessness among younger cohorts (Berrington, 2017). This suggests a rise in the proportion of population at risk of parity progression from first to second child in UK. Overall, levels of childlessness have also stabilized in much of western European countries, including the Nordic and three primarily German speaking countries (Germany, Switzerland and Austria) (Kreyenfeld and Konietzka, 2017). This evidence is related to the empirical regularity showing that the second birth rates are also declining in Europe (Frejka and Sardon, 2007). In England and Wales, the second birth rates reduced moderately from 73% for the 1950 cohort to 69% for the 1960 cohort (Frejka and Sardon, 2007). In general, our findings are relevant for countries with lower levels of childlessness and lower second birth rates.

Our study has limitations. First, our methods do not infer causality. Specifically, there might be concerns about confounding effects of income or other socio-economic factors affecting simultaneously the sleep of the infant and family fertility. We alleviate these concerns by including a good set of meaningful time-variant and predetermined controls, for both the mother and the child. Ultimately, our approach implies that, after controlling for these confounders infant sleep is a random event that escapes the capacity of control of the parents. Importantly, previous evidence on the same dataset supports this assumption (Blair et al., 2012). Second, our study focuses on a geographically limited sample. Normally, the richness and the consistency of the information obtainable from ad-hoc studies should be traded-off against their representativeness in more general context. The generalizability/internal validity trade-off associated even to well-designed studies, such as randomizations, is an indeed a common issue in the literature (Leamer, 2010). In our support, a few points should be raised. First, geographical restriction is the price to pay for getting access to a dataset that provides uniquely interesting and rich information at individual and household level. Second, the infant characteristics of the Avon study are similar to that of the infants in UK at birth, age 1 and 2. Moreover, the drop in total fertility rate (TFR) for Avon population over the years is lower than that of UK average, which makes the study more relevant for this population in particular. A third concern, in line with the previous one, is that our study is based on a 1990 survey (still ongoing), so that the validity for the current time period might be questionable. However, the same data on infant night wakings and other sleep measures have been recently used to address important research questions in social science, including parental employment and other associated demographic correlates, in addition to a vast number of epidemiological studies (Blair et al., 2012; Costa-Font and Flèche, 2017). Moreover, our research question is very broad and can be seen as a topic of interest outside of a particular time frame.

Despite these limitations, our study for the first time confirms that parents' experience of infant night wakings is an important factor in determining family size. Conditional on environmental or genetic factors, infant sleep is very random, and parents should learn to adapt to these changing patterns in the early days. Our findings also call for policy makers to consider the need to put in place, some form of support to improve the parenting experiences of first time parents so that such experiences do not contribute to inhibit further fertility intentions. Such support could include helping new parents "towards managing their postpartum sleep expectations" (Richter et al., 2019) and ways to reduce sleep deprivation after birth. For example, free weekly antenatal classes (also known as 'parentcraft classes') provided by the UK NHS includes discussions on managing sleep deprivation among parents (NHS UK). Access to such services could help new parents to be prepared for what to expect and focus on the good experiences of child rearing instead, which in turn would positively influence parity progress.

BIBLIOGRAPHY

Ajzen, I. 1991. The theory of planned behavior. *Organizational behavior and human decision processes,* 50(2), 179-211.

Ajzen, I. and Klobas, J. 2013. Fertility intentions: An approach based on the theory of planned behavior. *Demographic research,* 29, 203-232.

Avon Longitudinal Study of Parents and Children, [online]. Available from: http://www.bristol.ac.uk/alspac/ [Accessed 3 June 2020].

Barazzetta, M. and Ghislandi, S. 2017. Family Income and Material Deprivation: Do They Matter for Sleep Quality and Quantity in Early Life? Evidence From a Longitudinal Study. *Sleep,* 40(3).

Bavel, J. V. and Różańska-Putek, J. 2010. Second birth rates across Europe: interactions between women's level of education and child care enrolment. *Vienna Yearbook of Population Research*, 107-138.

Berrington, A., 2017. Childlessness in the UK. *In:* Kreyenfeld, M. and Konietzka, D. eds. *Childlessness in Europe: Contexts, Causes, and Consequences.* Cham: Springer International Publishing, 57-76.

Berrington, A., Stone, J. and Beaujouan, E. 2015. Educational differences in timing and quantum of childbearing in Britain: A study of cohorts born 1940–1969. *Demographic research,* 33, 733-764.

Blair, P. S., *et al.* 2012. Childhood sleep duration and associated demographic characteristics in an English cohort. *Sleep,* 35(3), 353-360.

Boyce, P. and Parker, G. 1989. Development of a scale to measure interpersonal sensitivity. *Aust N Z J Psychiatry,* 23(3), 341-351.

Callan, V. J., 1985. *Choices about children.* Melbourne, Australia: Longman Cheshire.

Cartwright, A. 1976. How many children?: Routledge direct editions.

Clark, A., Georgellis, Y. and Sanfey, P. 2001. Scarring: The psychological impact of past unemployment. *Economica,* 68(270), 221-241.

Costa-Font, J. and Flèche, S. 2017. Parental sleep and employment: evidence from a British cohort study.

Damato, E. G. and Zupancic, J. 2009. Strategies used by parents of twins to obtain sleep. *Appl Nurs Res,* 22(3), 216-220.

De Witte, H. 2016. On the scarring effects of job insecurity (and how they can be explained). *Scandinavian journal of work, environment & health,* 42(2), 99-102.

Dyck, I. 1990. Space, time, and renegotiating motherhood: an exploration of the domestic workplace. *Environment and Planning D: Society and Space,* 8(4), 459-483.

Elek, S. M., Hudson, D. B. and Fleck, M. O. 2002. Couples' experiences with fatigue during the transition to parenthood. *Journal of family nursing,* 8(3), 221-240.

Favez, N., Frascarolo, F. and Fivaz-Depeursinge, E. 2006. Family alliance stability and change from pregnancy to toddlerhood and marital correlates. *Swiss Journal of Psychology,* 65(4), 213-220.

Frejka, T. and Sardon, J.-P. 2007. Cohort birth order, parity progression ratio and parity distribution trends in developed countries. *Demographic research,* 16, 315-374.

Germo, G. R., *et al.* 2007. Child sleep arrangements and family life: Perspectives from mothers and fathers. *Infant and Child Development: An International Journal of Research and Practice,* 16(4), 433-456.

Giallo, R., Rose, N. and Vittorino, R. 2011. Fatigue, wellbeing and parenting in mothers of infants and toddlers with sleep problems. *Journal of Reproductive and Infant Psychology,* 29(3), 236-249.

Haig, D. 2014. Troubled sleep: Night waking, breastfeeding and parent-offspring conflict. *Evol Med Public Health,* 2014(1), 32-39.

Heckman, J. J. and Borjas, G. J. 1980. Does unemployment cause future unemployment? Definitions, questions and answers from a continuous time model of heterogeneity and state dependence. *Economica,* 47(187), 247-283.

Howie, P. W. and McNeilly, A. S. 1982. Effect of breast-feeding patterns on human birth intervals. *Reproduction,* 65(2), 545-557.

Jones, N. G. B. and da Costa, E. 1987. A suggested adaptive value of toddler night waking: delaying the birth of the next sibling. *Ethology and Sociobiology,* 8(2), 135-142.

Kahneman, D. 1994. New challenges to the rationality assumption. *Journal of Institutional and Theoretical Economics (JITE)/Zeitschrift für die gesamte Staatswissenschaft*, 18-36.

Karraker, K. H. and Young, M. 2007. Night Waking in 6-Month-Old Infants and Maternal Depressive Symptoms. *J Appl Dev Psychol,* 28(5-6), 493-498.

Kennedy, K. 1988. Consensus Statement: Breast-feeding as a Family Planning Method'. *The Lancet*, 1204-1205.

Kreyenfeld, M. and Konietzka, D., 2017. *Childlessness in Europe: Contexts, causes, and consequences.* Springer.

Leamer, E. E. 2010. Tantalus on the Road to Asymptopia. *Journal of Economic Perspectives,* 24(2), 31-46.

Loutzenhiser, L., McAuslan, P. and Sharpe, D. P. 2015. The trajectory of maternal and paternal fatigue and factors associated with fatigue across the transition to parenthood. *Clinical Psychologist,* 19(1), 15-27.

Margolis, R. and Myrskyla, M. 2015. Parental Well-being Surrounding First Birth as a Determinant of Further Parity Progression. *Demography,* 52(4), 1147-1166.

Martin, J., *et al.* 2007. Adverse associations of infant and child sleep problems and parent health: an Australian population study. *Pediatrics,* 119(5), 947-955.

Medina, A. M., Lederhos, C. L. and Lillis, T. A. 2009. Sleep disruption and decline in marital satisfaction across the transition to parenthood. *Fam Syst Health,* 27(2), 153-160.

Newman, L. 2008. How parenthood experiences influence desire for more children in Australia: A qualitative study. *Journal of Population Research,* 25(1), 1-27.

NHS UK, *Antenatal classes* [online]. Available from: https://www.nhs.uk/pregnancy/labour-and-birth/preparing-for-the-birth/antenatal-classes/.

Office for National Statistics, 2016. *Childbearing for women born in different years, England and Wales: 2016.* United Kingdom.

Otani, K., *et al.* 2009a. Parental overprotection increases interpersonal sensitivity in healthy subjects. *Compr Psychiatry*, 50(1), 54-57.

Otani, K., *et al.* 2009b. Dysfunctional parenting styles increase interpersonal sensitivity in healthy subjects. *J Nerv Ment Dis,* 197(12), 938-941.

Presser, H. B. 2001. Comment: A gender perspective for understanding low fertility in post-transitional societies. *Population and Development Review*, 27, 177-183.

Richter, D., *et al.* 2019. Long-term effects of pregnancy and childbirth on sleep satisfaction and duration of first-time and experienced mothers and fathers. *Sleep,* 42(4).

Sadeh, A., *et al.* 2007. Infant sleep and parental sleep-related cognitions. *J Fam Psychol*, 21(1), 74-87.

Sadeh, A., *et al.* 2009. Sleep and sleep ecology in the first 3 years: a web-based study. *J Sleep Res,* 18(1), 60-73.

Short, R. 1987. The biological basis for the contraceptive effects of breast feeding. *International Journal of Gynecology & Obstetrics*, 25, 207-217.

Short, R. V. 1984. Breast feeding. *Scientific American*, 250(4), 35-41.

Skinner, B. F., 1965. *Science and human behavior.* Simon and Schuster.

Stromberg, J., 16 April 2014. Do babies cry at night to prevent their parents from having more kids? *Vox*, p.

Tikotzky, L. and Sadeh, A. 2009. Maternal sleep-related cognitions and infant sleep: a longitudinal study from pregnancy through the 1st year. *Child Dev,* 80(3), 860-874.

Trivers, R. L. 1974. Parent-offspring conflict. *Integrative and Comparative Biology*, 14(1), 249-264.

Troxel, W. M., *et al.* 2007. Marital quality and the marital bed: examining the covariation between relationship quality and sleep. *Sleep Med Rev,* 11(5), 389-404.

Van der Wijden, C., Brown, J. and Kleijnen, J. 2003. Lactational amenorrhea for family planning. *Cochrane database of systematic reviews,* (4).

FIGURES

Figure 1: Flow chart for sample size of relevant variables at different points in time, 1991-95

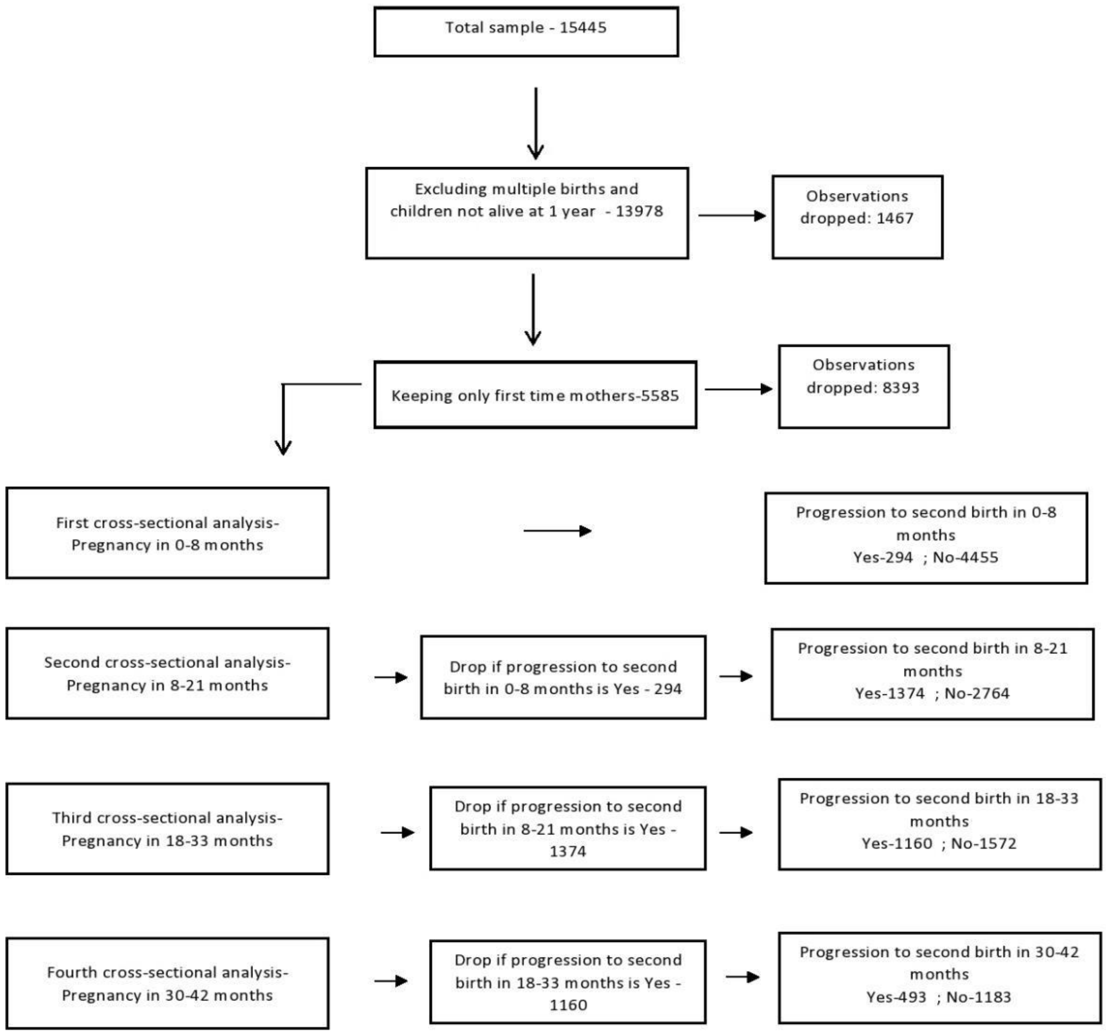


Figure 2 Progression to second pregnancy at four cross-sectional periods since birth of first child, 1991-95

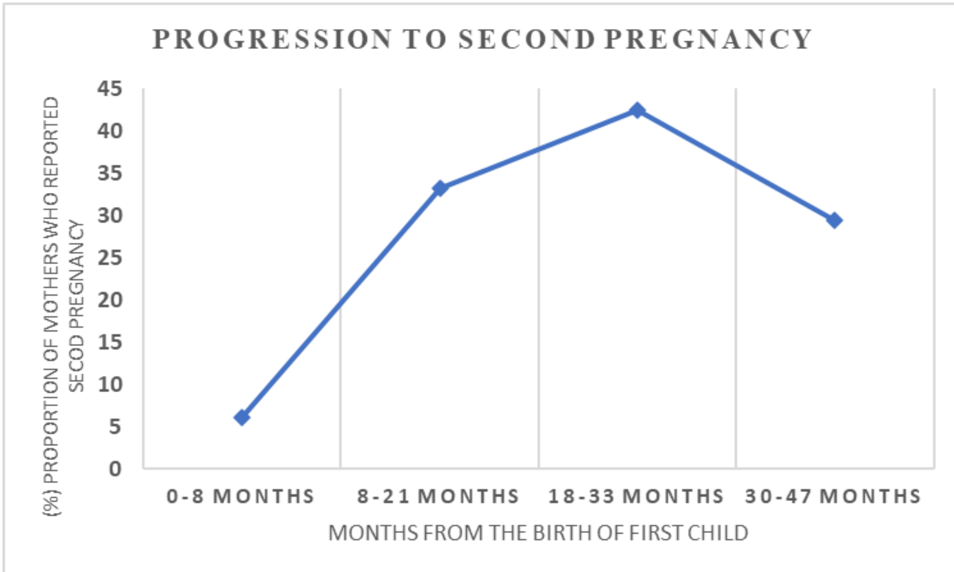
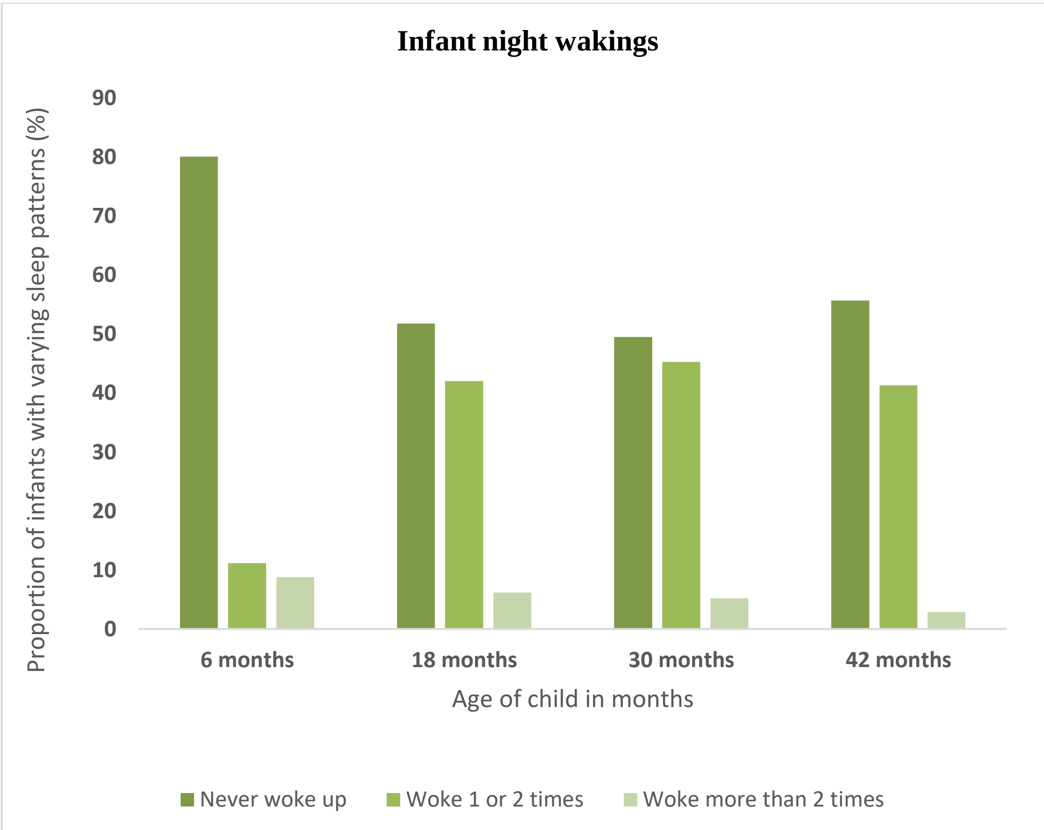


Figure 3 Distribution of infant night wakings from 0-4 years of age of child, 1991-95



TABLES

Table 1 Descriptive characteristics of the key variables in the analytic sample (N=5585) , 1991-95

	Total analytic sample	Observe only first birth	Observe second pregnancy	T Test or χ2 Test
Infant night wakings at 6 months (%)				
Never wake up	80.1	80.5	80.1	
Wakeup 1 or 2 times	11.2	11.8	10.8	
Wake up more than 2 times	8.8	7.7	9.1	
Infant night wakings at 18 months (%)				
Never wake up	51.8	48.1	53.4	*
Wakeup 1 or 2 times	42.0	44.1	40.8	
Wake up more than 2 times	6.2	7.8	5.8	
Infant night wakings at 30 months (%)				*
Never wake up	49.5	48.2	50.2	
Wakeup 1 or 2 times	45.3	44.5	45.4	
Wake up more than 2 times	5.2	7.3	4.4	
Infant night wakings at 42 months (%)				*
Never wake up	55.7	53.4	57.0	
Wakeup 1 or 2 times	41.4	42.5	40.5	
Wake up more than 2 times	2.9	4.1	2.5	
Mothers age at delivery (%)				*
< 21 years	12.3	8.6	8.3	
21-34 years	82.0	78.2	87.5	
≥ 35 years	5.8	13.2	4.2	
Highest level of parental education (%)				*
CSE / Vocational	7.2	5.6	6.0	
O-level	27.7	27.1	26.9	
A-level	37.2	44.2	35.6	
Degree	27.9	23.1	31.4	
Mother employed (%)				
0-8 months	42.0	44.4	42.2	
8-21 months	61.0	64.1	60.5	*
18-33 months	65.3	71.3	63.8	*
30-42 months	56.1	67.5	52.5	*
Partner employed (%)				
0-8 months	88.7	85.6	90.1	*
8-21 months	90.8	86.4	92.4	*
18-33 months	90.9	88.0	92.0	*
30-42 months	92.3	90.0	93.0	*
Divorced or separated (%)				
0-8 months	4.1	7.5	2.7	*
8-21 months	4.7	9.5	2.8	*
18-33 months	5.6	11.9	3.1	*
30-42 months	7.0	14.1	4.1	*
Income (%)				*
Q1	27.0	36.3	22.5	*
Q2	29.0	23.3	31.4	
Q3	22.7	21.7	23.5	
Q4	21.4	18.7	22.6	
Material deprivation				
8 months	0.67 (0.78)	0.66 (0.78)	0.64 (0.76)	
21 months	0.67 (0.70)	0.70 (0.75)	0.65 (0.67)	
33 months	0.61 (0.63)	0.63 (0.67)	0.60 (0.61)	

Note 1: Tests differences between those for whom only first births were observed and those who had a second birth. p*<0.5

Table 2 Descriptive characteristics of the other variables in the analytic sample (N=5585), 1991-95

	Total analytic sample	Observe only first birth	Observe second pregnancy	T Test or $\chi 2$ Test
Intentional pregnancy	67.5	62.4	74.4	*
Prenatal depression (mean)	4.6 (2.9)	4.8 (2.9)	4.4 (2.7)	*
Prenatal sleep-Got to sleep okay (%)				
Very often	24.1	24.7	24.1	
Often	48.7	51.0	50.6	
Not very often	24.8	22.6	23.5	
Never	2.3	1.7	1.8	
Maternal personality score (mean)	0 .0 (1.7)	-0.03 (1.7)	0.02 (1.6)	
Maternal sleep (>7 hours per night)				
2 months	20.7	17.8	21.7	*
8 months	41.8	37.8	43.1	*
21 months	40.5	37.5	41.4	*
33 months	32.7	38.0	30.9	*
Female child (%)	49.3	49.6	49.1	
Low birth weight (%)	6.6	7.4	5.4	*
Preterm birth (%)	6.4	6.3	5.6	
Ever breast fed (%)	78.3	77.1	81.3	*
Child health – ill (%)				
0-8 months	1.9	1.5	1.8	
8-21 months	3.0	3.2	2.8	
18-33 months	2.1	1.8	2.1	
30-42 months	3.3	3.0	3.1	

Note 1: Numbers shown in parentheses are standard deviations

Note 2: Tests differences between those for whom only first births were observed and those who had a second birth.

*p < .05

Table 3 Relationship between infant night waking and likelihood of second pregnancy, 1991-95

Infant night wakings	Probability of second pregnancy									
	Odds Ratio						Marginal Effects			
	Unadjusted		Model 1		Model 2		Model 1		Model 2	
	OR (95% CI)	P-value	OR (95% CI)	P-value	OR (95% CI)	P-value	dy/dx	P-value	dy/dx	P-value
At 6 months					***0-8 months***					
Never woke	1.00^		1.00^		1.00^		1.00^		1.00^	
Woke 1 or 2 times	0.59 (0.37-0.95)	0.029	0.43 (0.24-0.78)	0.005	0.56 (0.29-1.09)	0.090	-0.04	<0.001	-0.02	0.038
Woke more than 2 times	0.64 (0.39-1.06)	0.081	0.43 (0.22-0.84)	0.014	0.36 (0.14-0.90)	0.029	-0.04	0.001	-0.03	0.001
N	4505		3631		2740		3631		2740	
At 18 months					***8-21 months***					
Never woke	1.00^		1.00^		1.00^		1.00^		1.00^	
Woke 1 or 2 times	0.95 (0.83-1.09)	0.476	1.01 (0.87-1.18)	0.874	1.01(0.85-1.21)	0.885	0.00	0.874	0.00	0.885
Woke more than 2 times	0.72 (0.54-0.97)	0.030	0.64 (0.45-0.91)	0.013	0.62 (0.41-0.94)	0.024	-0.09	0.008	-0.10	0.015
N	3931		3160		2440		3160		2440	
At 30 months					***18-33 months***					
Never woke	1.00^		1.00^		1.00^		1.00^		1.00^	
Woke 1 or 2 times	1.13 (0.97-1.33)	0.126	1.20 (1.01-1.44)	0.040	1.19 (0.97-1.47)	0.098	0.05	0.039	0.04	0.098
Woke more than 2 times	0.61 (0.42-0.88)	0.009	0.61 (0.41-0.93)	0.020	0.45 (0.27-0.74)	0.002	-0.11	0.015	-0.18	0.001
N	2555		2145		1609		2145		1609	
At 42 months					***30-47 months***					
Never woke	1.00^		1.00^		1.00^		1.00^		1.00^	
Woke 1 or 2 times	0.91 (0.73-1.14)	0.428	0.95 (0.72-1.26)	0.723	1.00 (0.72-1.41)	0.979	-0.01	0.723	-0.00	0.979
Woke more than 2 times	0.53 (0.26-1.06)	0.074	0.96 (0.42-2.23)	0.932	0.78 (0.26-2.35)	0.662	-0.01	0.932	-0.05	0.652
N	1532		991		717		991		717	

^ Reference category

Note1 Model 1 adjusts for all simultaneous variables including a binary indicator for mother employed, father employed, divorced or separated, household income in categories, household material deprivation and child health

Note 2: Model 2 adjusts for all variables in model 1 in addition to pre-determined variables including mother's age at delivery, parents' highest level of education, , binary indicator for intentional pregnancy, prenatal depression score, prenatal sleep in categories and a score for prenatal maternal personality, child gender, ever breast fed, low birth weight, preterm birth

Table 4 Scarring and Habituation effects of past infant sleep on likelihood of second pregnancy, 1991-95

	Probability of second pregnancy			
	Model 1		Model 2	
	OR (95% CI)	P-value	OR (95% CI)	P-value
	8-21 months			
Scarring effect -Past Infant Sleep				
Infant night waking at 6 months				
Never woke	1.00^		1.00^	
Woke 1 or 2 times	1.14 (0.87-1.50)	0.345	1.30 (0.86-1.96)	0.218
Woke more than 2 times	1.15 (0.84-1.58)	0.387	1.01 (0.59-1.72)	0.975
Current Infant Sleep				
Infant night waking at 18 months				
Never woke	1.00^		1.00^	
Woke 1 or 2 times	1.00 (0.83-1.20)	0.987	1.04 (0.85-1.27)	0.732
Woke more than 2 times	0.56 (0.36-0.86)	0.008	0.37 (0.19-0.73)	0.004
Habituation effect- Current*Past Infant Sleep				
Never woke * Never woke			1.00^	
Woke more than 2 times*Woke more than 2 times			2.64 (0.93-7.53)	0.069
N	2388		2388	
	18-33 months			
Scarring effect- Past Cumulative Infant Sleep				
Infant night waking at 6 and 18 months	1.06 (0.94-1.19)	0.337	1.17 (0.96-1.41)	0.118
Current Infant Sleep				
Infant night waking at 30 months				
Never woke	1.00^		1.00^	
Woke 1 or 2 times	1.11 (0.88-1.39)	0.376	1.24 (0.93-1.66)	0.148
Woke more than 2 times	0.38 (0.22-0.66)	0.001	0.43 (0.16-1.16)	0.095
Habituation effect- Current*Past cumulative Infant night waking				
Never woke*Past cumulative sleep			1.00^	
Woke more than 2 times*Past cumulative night waking			0.88 (0.55-1.40)	0.577
N	1511		1511	

^ Reference category

Note1: Both models adjust for all time varying variables including mother employed, father employed, divorced or separated, household income, material deprivation and child health and pre-determined variables including mother’s age at delivery, parents’ highest level of education, child gender. low birth weight, preterm birth, intentional pregnancy, ever breast fed, prenatal depression, prenatal sleep and prenatal maternal personality.

Table 5 Effects of past infant sleep at 2 years of age on total number of pregnancies (Model 1) and total number of children 4 years since first birth (Model 2), 1991-95

	Probability of an additional pregnancy			
	Model 1		Model 2	
	Coefficient	P-value	Coefficient	P-value
Infant night waking at 2 years				
Never woke	1.00^		1.00^	
Woke 1 or 2 times	-0.05	0.005	-0.04	0.100
Woke more than 2 times	-0.10	0.005	-0.11	0.019
N	3922		2332	
	Probability of an additional child			
	Model 1		Model 2	
	Coefficient	P-value	Coefficient	P-value
Infant night waking at 2 years				
Never woke	1.00^		1.00^	
Woke 1 or 2 times	-0.06	0.024	-0.05	0.136
Woke more than 2 times	-0.19	0.000	-0.22	0.001
N	3859		2301	

Note 1: Model 1 does not adjust for any controls.

Note 2: Model 2 adjusts for covariates including mothers' age at birth, highest level of parental education, household income, indicator for child gender, low birth-weight, intentional pregnancy and prenatal personality score as used in the main specification. In addition model 2 includes controls for duration of mothers and fathers employment (until age 3 of child), indicator for divorced or separated and child's health status measured at 2 years of child. Maternal depression is computed as the average of depression subscale of CCEI at 18 weeks of gestation, 2 months and 2 years since child birth.

Table 6 Test for maternal sleep as mediator between infant night waking and second pregnancy, 1991-95

	Probability of second pregnancy	
	0-8 months	
Infant night wakings	Mothers who had ≤ 7 hours of sleep/night-2 months	Mothers who had > 7 hours of sleep/night-2 months
At 6 months	**OR (95%CI)**	**OR (95%CI)**
Never woke	1.00^	1.00^
Woke 1 or 2 times	0.47 (0.21-1.03)	0.73 (0.16-3.41)
Woke more than 2 times	0.33 (0.12-0.93)	0.75 (0.09-5.93)
N	2162	506
	8-21 months	
	Mothers who had ≤ 7 hours of sleep/night-8 months	Mothers who had > 7 hours of sleep/night-8 months
At 18 months	**OR (95%CI)**	**OR (95%CI)**
Never woke	1.00^	1.00^
Woke 1 or 2 times	1.04 (0.82-1.31)	0.90 (0.67-1.22)
Woke more than 2 times	0.69 (0.43-1.09)	0.44 (0.15-1.32)
N	1408	977
	18-33 months	
	Mothers who had ≤ 7 hours of sleep/night-21 months	Mothers who had > 7 hours of sleep/night-21 months
At 30 months	**OR (95%CI)**	**OR (95%CI)**
Never woke	1.00^	1.00^
Woke 1 or 2 times	1.34 (1.00-1.80)	1.01 (0.72-1.40)
Woke more than 2 times	0.43 (0.22-0.82)	0.74 (0.30-1.85)
N	847	663
	30-47 months	
	Mothers who had ≤ 7 hours of sleep/night-33 months	Mothers who had > 7 hours of sleep/night-33 months
At 42 months	**OR (95%CI)**	**OR (95%CI)**
Never woke	1.00^	1.00^
Woke 1 or 2 times	0.72 (0.46-1.13)	1.55 (0.85-2.81)
Woke more than 2 times	0.42 (0.10-1.74)	1.84 (0.21-16.3)
N	447	267

All models adjusts for covariates including mothers' age at birth, highest level of parental education, household income, indicator for child gender, low birth-weight, intentional pregnancy and prenatal personality score as used in the main specification. In addition model 2 includes controls for duration of mothers and fathers employment (until age 3 of child), indicator for divorced or separated and child's health status measured at 2 years of child. Maternal depression is computed as the average of depression subscale of CCEI at 18 weeks of gestation, 2 months and 2 years since child birth.

SUPPLEMENTARY FILE

Table S1 Relationship between infant night waking and likelihood of second pregnancy, 1991-95

	Probability of second pregnancy			
	0-8 months	8-21 months	18-33 months	30-42 months
Infant night wakings				
Never woke up	1.00^	1.00^	1.00^	1.00^
Wake up 1 or 2 times	0.56*	1.01	1.19*	1.00
Wake up more than 2 times	0.36**	0.62**	0.45***	0.78
Mothers age at delivery				
<21 years	1.00^	1.00^	1.00^	1.00^
21-34 years	1.13	0.90	1.20	0.97
≥ 35 years	0.87	0.49**	0.55*	0.12***
Highest level of parental education				
CSE/Vocational				
O-level	0.91	1.31	0.96	1.21
A-level	0.71	1.59**	0.87	0.99
Degree	1.13	2.29***	1.45	1.80
Mother employed	0.62**	0.50***	0.64***	0.51***
Partner employed	1.03	1.42*	1.51**	1.61
Divorced or separated	0.83	0.61	0.43***	0.45*
Income				
Q1	1.00^	1.00^	1.00^	1.00^
Q2	0.85	1.22	1.15	1.66*
Q3	0.79	1.09	0.86	1.47
Q4	0.84	1.10	0.70*	1.95**
Household material deprivation	1.07	0.99	1.00	0.90
Maternal personality score	0.89**	1.02	1.02	0.97
Female child	1.26	0.92	0.96	0.97
Low birth weight	0.75	0.90	0.79	0.73
Preterm birth	1.60	0.98	0.90	1.18
Child health				
Healthy	1.00^	1.00^	1.00^	1.00^
Unwell	1.47	0.64	2.40**	0.81
Prenatal depression	1.01	0.95**	0.96**	1.01
Prenatal sleep				
Very often	1.00^	1.00^	1.00^	1.00^
Often	1.01	1.29**	0.93	1.31
Not very often	1.74**	0.95	0.99	1.07
Never	3.70***	1.40	1.52	0.22
Intentional pregnancy	0.80	1.28**	1.44***	1.33
Ever breast fed	1.21	1.49***	1.02	1.48*
N	2740	2440	1609	717
R-squared	0.04	0.05	0.05	0.10

*p<0.1, **p<0.05, ***p<0.001

Table S2: Relationship between infant night waking and likelihood of second pregnancy, 1991-95

VARIABLES	(1) Likelihood of second/subsequent pregnancy	(2) Likelihood of second/subsequent pregnancy
Infant night wakings		
Never woke up		
Wake up 1 or 2 times	2.132***	2.176***
	(0.108)	(0.118)
Wake up more than 2 times	0.739***	0.777**
	(0.082)	(0.092)
Observations	11,008	9,659
Number of observations	2,965	2,717
Controls	No	Yes
Mother fixed effects	Yes	Yes
Model	Logistic	Logistic

Note 1: Model 1 does not adjust for any controls.

Note 2: Model 2 adjusts for time-varying covariates including duration of mothers and fathers employment, indicator for divorced or separated, child's health status and material deprivation..

APPENDIX

Statistical Analysis

1. Models for the relationship between infant night waking and likelihood of second pregnancy.

The logistic regressions for the four time periods are as follows:

$$\Pr(Pregnancy_t) = f(\beta_0 + \beta_1 Infant\ Sleep_{tx} + \gamma_1 X_{tx} + \gamma_2 X_0 + \varepsilon)$$

where $\Pr(Pregnancy_t)$ is the probability of a woman becoming pregnant of the second child *t* months after the first and *f( )* is the logit function. We have four combinations of *t* and *tx*: 1) *t*=0-8 months and *tx*=6 months; 2) t=8-21 months and *tx*=18 months, 3) *t*=18-33 months and *tx*=30 months, 4) *t*=30-47 months and *tx*=42 months. Given the different time points at which the outcome and independent variable were collected and the framing of questions; *t* and *tx* are not exactly the same. However, we have chosen *t* and *tx* in such a way that we can interpret a contemporaneous relationship. X_{tx} refers to time varying covariates and X_0 to predetermined covariates, as described in the Data section in the manuscript.

2. Models for Scarring effects of past infant sleep on likelihood of second pregnancy,

The models for the scarring effects can be expressed as follows:

$$\Pr(Pregnancy_t) = \beta_0 + \beta_1 Past\ Infant\ Sleep_{tx} + \beta_2 Current\ Infant\ Sleep_{ty} + \gamma_1 X_{tx} + \gamma_2 X_0 + \varepsilon$$

where: 1) *t*=8-21 months, *tx*=6 months and *ty*=18 months; 2) *t*=18-33 months, *tx*=6-18 months and *ty*=30 months. For the second period, $Infant\ Sleep_{tx}$ is a cumulative sum of past infant night wakings, ranging from 0 (never woke at night at both time points) – 6 (woke more than twice at both time points).

Chapter 1.4

Timing and Scarring Effects of Childhood Sleep Deprivation. Do They Matter for Later Human Capital Formation?[5]

ABSTRACT

The medical literature has linked sleep disruption to poor cognitive and non-cognitive outcomes, which are both essential for human capital development. However, evidence on the relationship between timing of this health shock (sleep disruption) in childhood and its future consequences is limited. Using a cohort of children born in 1990 in Avon in England, we find that parent/carer reported sleep disruption in early childhood is related to a range of cognitive and non-cognitive outcomes during adolescence. Applying a structured life course approach to our data, we compare Bayesian Information Criteria (BIC) for competing hypotheses on timing of sleep disruption in childhood (ages 1.5-10) in predicting cognitive and non-cognitive outcomes (ages 8-15). The knowledge of when sleep disruption may have severe consequences for human capital can aid in the efficient timing of appropriate interventions. Variables representing both critical periods (early or mid-childhood) and cumulative patterns of sleep disruption showed optimal fit when compared to the remaining models illustrating different life course pattern. To establish longer-term causal links and overcome potential sources of endogeneity, we exploit variations in individual specific parent/carer reported exposure to neighbor noise to instrument for sleep disruption. We find statistically significant causal effects of childhood sleep disruption on a range of outcomes including full-scale intelligence quotient (IQ) at age 8 and 15, verbal IQ at age 8, vocabulary IQ at age 15 and behavioral difficulties at age 10 and 12. Our results are robust to a range of sensitivity checks and elucidate the cognitive and non-cognitive related significance of both sleep and noise-related policies in developing human capital. Although, sleep disruption can have cumulative effects on human capital outcomes, a close tie with critical models indicates the need for early interventions. This may offer the best option given the evidence on cheaper and higher returns from early investments. We also observe scarring effects of past sleep disruption irrespective of the sleep problems at the present which strengthens our recommendation for early intervention.

INTRODUCTION

[5] This is a working paper and we provide the latest version of the manuscript.

The role of fetal and early childhood for shaping the outcomes over the life course has been long recognized (Almond and Currie, 2011; Barker, 1990; Cunha and Heckman, 2007). Evidence from landmark studies in this field suggests that disadvantaged childhood may have adverse effects on long term programming such as development of cognitive and non-cognitive skills, thereby affecting human capital formation in the adulthood (Adair et al., 2013; Cunha and Heckman, 2007; Heckman, 2007; Knudsen et al., 2006). The first thousand days of life is a window of opportunity for brain development (Cusick and Georgieff, 2016; Mputle, 2019). Moreover, the plasticity and vulnerability of brain is at the highest peak during early childhood (Kolb et al., 2017; Lidzba et al., 2009), suggesting the long term prospects of childhood interventions for human capital development. Given the existing evidence, the big question for pediatricians and policy makers is to understand the optimal point in time to execute the intervention to reap the largest benefit.

Gaps in the development of human capital potentials can be attributed to a variety of factors including in-utero and early life exposure to poor maternal health, malnutrition, poverty, diseases or exogenous shocks such as war, famine, and recession among many others (Almond and Currie, 2011). Given a large corpus of evidence on the role of early life conditions, we investigate the life course effects of sleep deprivation. Particularly, we quantify the cognitive and non-cognitive penalty of sleep disruption in childhood and examine if the timing and scarring effects of poor sleep matter for later human capital outcomes.

Early literature has linked sleep disruption to poor human capital development, particularly cognitive and non-cognitive outcomes (Killgore, 2010; Sadeh et al., 2014). Sleep plays a vital role in brain maturation which is required for development of memory consolidation, learning and behavioral functions (Blunden et al., 2000; Bub et al., 2011; Spruyt and Gozal, 2010). Animal models of sleep shows that sleep deprivation alters neural plasticity and affect learning capabilities during the time period equivalent to that of childhood (Row et al., 2007; Tartar et al., 2006). Although, we lack human data on such extensive duration of sleep deprivation, relatively identical studies on children also show similar findings (Beebe, 2006).

For instance, a meta-analysis of 50 studies conducted by authors Dewald et al. (2010) on children and adolescents shows that insufficient sleep, poor sleep quality and sleepiness significantly worsened school performance. The authors also find the negative effects to be larger in samples with younger participants, signifying the importance of pronounced developmental changes in prefrontal cortex during childhood and early adolescence for cognition. A meta-analysis of 86 studies on healthy children aged 5-12 years shows insufficient sleep to be positively correlated with lower cognitive performance and higher behavioral difficulties (Astill et al., 2012). Studies using longitudinal and experimental data also find similar results. Particularly, a study utilizing a modest life course approach found adverse effects of night waking's and insufficient sleep quantity at 2 years of age on cognitive outcomes at 6 years (Kocevska et al., 2017). These findings buttress the conclusion from an earlier study showing shorter sleep duration especially before 41 months to be associated with lower cognitive performance at 5 years (Touchette et al., 2007). Using a more robust causal approach, Kelly et al. (2013) employs a difference-in-difference methodology to find an increase in behavioral difficulties as children transitioned from regular to irregular bed time-schedules across childhood. Fallone et al. (2005) conducted a within-subject 3 week-long experiment under baseline, optimized and restricted sleep conditions among healthy normal children to show that restricted sleep conditions have a direct negative effect on teacher reported school performance compared to other two conditions. A similar experiment where children in the treatment group were subjected to an additional 1 hour sleep deprivation at night relative to other days performed worse on neurobehavioral functioning compared to the control group (Sadeh et al., 2003). Overall, experimental studies among healthy children are limited given the ethical issues in subjecting children to sleep deprivation. Studies have also differentiated between short and long term sleep deprivation effects on human capital outcomes as these may have differential effects on cognitive and non-cognitive outcomes (Jan et al., 2010).

However, there is a scarcity of literature investigating the characteristics of exposure to sleep problems, with a particular focus on developmental timing or duration and how this affects later cognitive and non-cognitive skills. In our study, we adapt several examples of life course research that investigates the relationship between early life shocks and later outcomes (Dunn et al., 2016; Green et al., 2018; Mishra et al., 2009; Pearce et al., 2014) to childhood sleep problems. This methodology allows for contrasting hypotheses where the effects may be dependent on timing, duration, and sequencing of exposure to adverse conditions. Such life course approaches sheds light on the mechanisms underlying the risk for poor skill development by indicating the development processes that are affected by exposure to sleep problems. Most importantly, childhood consists of several developmental windows where different interventions can be implemented to minimize the adverse effects of sleep problems. Life-course approach helps us to locate these optimal time points to intervene, both in terms of cost and effectiveness.

Life-course Approach

In life course epidemiology, there are several contrasting hypotheses regarding the timing, duration or sequence of a shock and the resulting exposure effects. This approach can best explain the pattern of relationship between exposure to sleep disruption across childhood and later outcomes. The first sets of theoretical models are the critical models where later outcomes depend on the timing of sleep disruption.

In particular, critical period refers to a period of exclusive risk where exposure to a shock at that particular time point may have permanent or irreversible effects (Ben-Shlomo and Kuh, 2002; Mishra et al., 2009). The second model is the accumulation model, which posits that it is the duration and not the timing of exposure that matters (Mishra et al., 2009). This implies that exposure to sleep problems in infancy has the same effect as exposures in early or mid-childhood and these effects compound in a dose-response manner to affect later human capital outcomes. These are contrasted against other models such as persistent or trajectory models where both sequence and direction of change over time matters (Green et al., 2018).

Previous literature provides evidence supporting the role of timing of adverse shocks (or investments) in predicting later outcomes. For instance, timing of in-utero health shocks such as early disruptions to prenatal visits compared to later prenatal visits have been shown to be causally related to worse child birth outcomes including birth weight and gestation (Evans and Lien, 2005). A similar study comparing in-utero investments at different gestational ages finds early exposure to higher-quality prenatal care (access to micronutrient supplements) to be causally related to higher educational attainment (Lavy et al., 2016). In addition to health shocks, timing of exposure to other types of shocks such as psychological, economic, or weather shocks has also been investigated in the past. For instance, being exposed to weather shocks in early years of childhood compared to later years were found to be associated with lower educational attainment (Duque et al., 2018). In another study, the authors contrasted competing hypotheses on the timing of exposure to certain family structures at various ages and identified age seven to be a sensitive period for later socio-emotional wellbeing (Pearce et al., 2014). Similarly, Duncan et al. (1998) found exposure to poverty at ages 0-5 to have heightened risk for poor academic achievement compared to exposure to poverty at other ages (Duncan et al., 1998).

Identifying which model best fits the data is important as they have different inferences for both the timing of the intervention and the underlying mechanism (developmental or proximal) (Green et al., 2018; Mishra et al., 2009). The variables encoding a priori hypothesis are separately added to the regression model and the sleep disruption variable that gives the best improvements to the model fit is chosen. This structured life course method is competent to differentiate between several contrasting models of association between exposures and later outcomes (Green et al., 2018; Mishra et al., 2009). In our study, we find evidence for mostly critical and cumulative models of sleep disruption exhibiting the best (or second best) fit. We also find sleep disruption to be negatively associated with our cognitive and non-cognitive (except for matrix IQ at age 15).

Establishing causality

Although reverse causality between past sleep problems and later human capital outcomes can be ruled out, these observational correlations cannot be interpreted as causal effects. It could be that sleep disruption may arise from some unobserved conditions which could be simultaneously correlated with cognition. The few studies addressing the problem of causal inference when studying the effects of sleep deprivation on human capital outcomes have exploited within country variation in time zones or within country sun set times (Gibson and Shrader, 2014; Jagnani, 2018; Nissenbaum et al., 2012). When considering populations that are geographically restricted such as Avon (as in our case), the above instruments are not applicable. Here we use an instrument which is non-experimental but causally identifies the effect of past episodes of sleep disruption on cognitive outcomes. Following the identification strategy used by Fan and Weinhold (2018), we estimate the causal effects of sleep disruption on human capital outcomes using exposure to neighbor noise as an instrument. We isolate the causal impact of sleep disruption based on the identifying assumption that exposure to neighbor noise is exogenous to our set of human capital outcomes and affects the outcomes only through sleep disruption (after controlling for a range of socio-economic, dwelling and neighborhood characteristics). We test for the validity assumptions of our IV and also employ techniques to construct bounds when assuming that the IV is only plausibly exogenous (Conley et al., 2012). The instrumental variable estimation finds statistically significant causal effects for outcomes including verbal IQ at age 8, full-scale and vocabulary IQ at age 15, and behavioral difficulties at age 10 and 12.

Scarring effects

The notion of 'scarring effects' is derived from the literature on economics and psychology. 'Scarring effects' refers to the process by which the outcomes in the current period depend on the past experiences independent of the current experiences (Kahneman, 1994). These past experiences or exposures may offset the current exposure to influence simultaneous outcomes. This hypothesis about how past experiences 'scar' has been mostly applied in the literature of unemployment and wellbeing (Heckman and Borjas, 1980). Clark et al. (2001) shows duration of past unemployment to be negatively associated with current wellbeing independent of the current employment status. Applying the theory of scarring effects to the current context, we hypothesize that the past cumulative measure of exposure to sleep problems (across childhood) may have significant adverse effects on cognition, independent of the more proximal measures of sleep exposure.

To summarize, this study first uses a structured life course approach to show that a critical period and cumulative model best fits the relationships between sleep disruption and human capital outcomes. Then we establish the causal relationship between measures of sleep disruption and human capital outcomes using neighbour noise as an instrumental variable. We show that sleep disruption during early childhood has significant adverse effects on cognitive and non-cognitive outcomes during adolescence. Finally, we show that past sleep disruptions have scarring effects on later human capital outcomes independent of the sleep quality at the present.

DATA

We use data from Avon Longitudinal Study of Parents and Children (ALSPAC), a population based longitudinal study, based in three health districts in South West of England (Golding, 2010). A total of 15445 pregnant women, with expected dates of delivery between April 1991 and December 1992 were enrolled in the study. We have a final sample of 14504 children after excluding perinatal deaths, children who died within 1 year of birth and duplicate observations. Boyd et al. (2013) provides a full description of the cohort profile and methods used for data collection. The study website contains other detailed information including survey methodology, questionnaires, variables, and updated findings using ALSPAC data (www.alspac.bris.ac.uk). Ethical approval for the study was obtained from the ALSPAC Ethics and Law Committee and the Local Research Ethics Committee.

Outcome variables

The key outcome variables are cognitive and non-cognitive outcomes measured during 8 to 15 years of age of the child. Cognitive outcomes include full-scale, verbal and performance IQ measured at age 8 and full-scale, vocabulary and matrix IQ measured at age 15. All children in the sample were invited to participate in the clinical sessions held at the University of Bristol. However, the response rate was lower compared to the full sample and varies for each outcome.

IQ at 8 years (mean age = 8.7 years) was measured using Wechsler Intelligence Scale for Children (WISC-III) during the 'Focus at 8' session by trained psychologists using an adapted form of WISC-III (Wechsler, 1991). The test consisted of five verbal subtests (information, similarities, arithmetic, vocabulary, and comprehension) and five performance subtests (picture completion, coding, picture arrangement, block design, object assembly). The raw scores for each subtest were age-scaled based on WISC- III manual instructions. The age-scaled verbal subtests were combined to form verbal IQ and the age-scaled performance subtests were combined to form the performance IQ. These scores were summed and converted to obtain the full-scale IQ. The full-scale IQ at age 15 (mean age =15.5 years) was measured using the Wechsler Abbreviated Scale of Intelligence (WASI) (Wechsler, 1999). It consists of two verbal subtests (vocabulary and similarities) and two performance subtests (block design and matrix reasoning). Raw scores were age-scaled and converted to T-scores based on the WASI manual. For the ALSPAC study, only the vocabulary and matrix reasoning subtests were conducted, and these were used to obtain the full-scale IQ. Full-scale IQ, verbal, and performance IQ at age 8 and full-scale IQ at age 15 were standardized to achieve a mean of 100 and standard deviation of 15 for ease of interpretation. Similarly, vocabulary and matrix reasoning subtests were standardized to achieve a mean of 50 and a standard deviation of 10 for an easier interpretation.

Non-cognitive skills for the ALSPAC children are measured using the Strengths and Difficulties Questionnaire (SDQ) (Clark et al., 2019; Kelly et al., 2013; Lekfuangfu et al., 2015). The SDQ is a validated tool consisting of 25 questions for identifying behavioral difficulties among children (Goodman, 1997, 2001). For example, some of the questions answered by parent/carer included items relating to child's concentration span, behavior in new situations, temper tantrums, if they were liked by other children and if they were considerate about other people's feelings (see Table A1 for the complete list of SDQ items). These responses were used to construct five sub-scales as follows: hyperactivity scale, emotional symptoms scale, conduct problems scale, peer problems scale and prosocial scale. For our analysis, we use the total difficulties score which is the sum of all subscales excluding prosocial scale. We use the parent-reported version of the SDQ for our main analysis and repeat all analyses with teacher-reported SDQ at age 12 as a robustness check (results available from the author upon request). This helps to overcome the bias from using responses collected from only one source (Fleche, 2017). For our analysis, we reverse the total SDQ score so that higher values indicate lower behavioral difficulties.

Exposure/ Independent Variables

The mothers or caretakers of the children in the ALSPAC cohort were asked "if in the past one year, the child continued to get up after only few hours of sleep". The responses were originally categorized into yes, did not worry; yes, worried a bit; worried greatly and no and was reported at 1.5, 2.5, 3.5, 5, 6, 7 and 10 years of age (mean age at the time of completing the questionnaire) of child.[6] For our analysis, we

[6] The time period for which the information on sleep disruption holds is considered to begin from the year prior to the mean age of completing the questionnaire. For example, if the mean age at which the questionnaire was completed is 10 years; then the corresponding information refers to sleep disruption during 9-10 years of age.

constructed a binary variable for the child experiencing '*sleep disruption*' if the mother answered everything except no.[7]

Information on *sleep disruption* at 1.5 and 2.5 years is combined to construct a measure of '*Infancy sleep disruption*' which takes a value of 1 if *sleep disruption* is reported at either of these age points and 0 otherwise. In similar fashion, information on *sleep disruption* at 3.5 and 5 years is combined to construct a measure of '*early childhood sleep disruption'* and information at 6 and 7 years is combined to construct a measure of '*mid-childhood sleep disruption'* respectively[8]. This is done to reduce 6 time points for *sleep disruption* into 3 for ease of adapting the data to a life course model. Figure 1 illustrates a graphical representation depicting the aggregation of the exposure variable.

Other covariates

We include a range of covariates including child and family characteristics, socio-economic status, prenatal and birth outcomes. Socio-economic status is measured as the highest level of parental education. The lowest level is CSE/vocational followed O-level, A-level, and university degree. Net household income per week is reported as a categorical variable of 5 bands measured at 3, 4, 7 and 8 years of the child's age. For each age, we convert the bands into a figure by taking the midpoint of the interval and these new income variables are then averaged across the four ages to construct a permanent measure of household income and divided into four quartiles. The average income is £151, £285, £454, £663 per week respectively for the first (lowest), second, third and fourth quartile. We also include a measure of material deprivation indicating the quality of dwelling and neighbourhood characteristics in our models. The material deprivation index is constructed at 2, 3, 5 and 7 years by assessing the number of items that the household is deprived of: indoor toilet, shower or bath for sole use, working telephone, presence of damp or mould in the household and cleanliness of the neighbourhood. We constructed a permanent measure of material deprivation by taking an average of the number of items deprived in each period.

Maternal age at delivery was categorized into three: <21, 21-34 and ≥ 35 years. Mother's employment status was measured at 8 months, 2, 3, 4, 8, 10 and 11 years respectively since birth of study child. For outcomes measured up to 8 years of age of child, duration of maternal employment was computed as sum of first five periods recorded to be employed (until age 8). For outcomes measured at age 10, we include information until age 10 and for outcomes measured during 12 to15 years of age, we include maternal

[7] Imputation to exposure variables is carried out at this point; specifically, before constructing sleep disruption at infancy, early-childhood and mid-childhood. More information on imputation is given in the next section.

[8] When the outcomes correspond to those measured at 12-15 years, mid childhood sleep problems take a value of 1 if a sleep problem was reported at 6, 7 or 10 years of age.

employment status for all periods.[9] Marital status of the mother was recorded at 2, 3, 4, 7, 8, 9, 10 and 11 years since the birth of the study child. For outcomes measured at 8 years, we consider a mother to be divorced or separated if they responded yes to this question at any point in first 5 periods (until age 8), for outcomes measured at 10 years, we include the information until age 10 and for outcomes measured at age 12-15, we include information at all periods.

Other controls include maternal smoking (yes/no) measured at 32 weeks of gestation, age of the child in months, child gender, child ethnicity (white vs. others), an indicator for first born, low birth weight (< 2500g), ever breastfed breast fed within 1 year of birth (yes/no). Parent reported health status of the child was measured every year from birth until 7 years of age (very healthy, healthy with minor problems, sometimes ill, always ill). We consider a child to be unwell if the parent/carer reported the child to be sometimes or always ill at any point in time.

For additional robustness check, we also control for household having ever moved residences and the child's sensitivity to noise. Mother/carer was asked to report 'if the child moved home' at 2, 3, 4, 5, 6 and 7 years. We consider the study child to have ever moved if the responses to the question at any point was reported to be 'yes'. At 2 years, mother/carer responded to questions on child's ability to continue any activity in general or in particular such as playing or looking at books despite a noisy environment. The responses ranged from 1 (almost always) to 5 (almost never). For our analysis, we take an average of the responses over the 3 domains to construct a measure of child sensitivity to noise. A higher score indicates higher sensitivity to noise (or easily disturbed).

ESTIMATION STRATEGY

Model Selection

Table 1 shows 10 (9 binary and 1 continuous) variables constituting contrasting hypotheses regarding exposure to sleep disruption across three periods of childhood (infancy, early and mid-childhood). Critical period hypothesis is represented by binary indicators of sleep disruption in infancy, early and mid-childhood. The accumulation hypothesis is indicated by a cumulative sum of the periods in which the child has been exposed to disruption. Other variables indicate alternative hypotheses including sleep trajectories, persistent exposure to sleep disruption and a threshold hypothesis whereby any exposure to sleep disruption is associated with an equal change in the cognitive outcome.

[9] Duration of maternal employment was considered to be missing only if the information was missing for more than 2 time points.

For the model selection, analysis proceeds as follows: First, we predict all outcomes with all covariates but no *sleep disruption* exposure variables using linear regression models (or non-linear models when using binary outcomes). Second, we add each *sleep disruption* exposure variable defined in Table 1 separately to the model. The improvement in each model was estimated using a Wald test ($p < 0.05$ for significance), followed by comparison using the Bayesian Information Criterion (BIC) (Green et al., 2018). We choose the exposure variables that gives the lowest BIC (best fit).

Our measures on sleep deprivation are subjected to attrition over time (around 22% from infancy to mid childhood). This implies that the contrasting models may have different sample sizes, with the variable encoding cumulative model of sleep disruption having the least number of observations. Given that the sample size is a parameter for the BIC[10] and could affect the final estimates, we impute a proportion of the missing observations in the exposure variable.

For the imputation, we employ the multiple imputation technique specifically developed for categorical time series data to create 25 imputed datasets (Halpin, 2016). The method carries out a monotonic series of imputations, thereby preserving the longitudinal consistency of the exposure variable. Therefore, this method has an advantage over standard imputations methods such as the multiple imputation techniques developed by Rubin (1987) where the longitudinal nature of an information is usually ignored. The method starts with rearranging the exposure variables in a wide format, with a focus on gaps in the sequence rather than missing observations in each time indexed variable. Two types of gaps are observed: internal and initial or terminal. For internal gaps, the leads and lags in the sequence is not missing and each element in the gap is imputed using information observed on the prior and subsequent states which was originally present or imputed in the process along with the duration of time spend in that state. However, initial and terminals gaps are imputed using information on the prior and subsequent states respectively. For the *sleep disruption* variable observed at 7 time points, maximum internal gaps of 3 and maximum initial and terminal gaps of 2 were imputed. Additionally, to improve the quality of the model, we add the gender of the child (fully observed time fixed variable) to the model.[11]

About 43% of the observations with information on sleep disruption at any one age point were imputed. Given that not all gaps in an imputed sequence are necessarily imputed[12], all contrasting models do not have same sample size. Therefore, we restrict the sample size in our model estimation (after imputation) to that of the cumulative sleep disruption model, which has the least number of observations.

---

[10] BIC = -2 * LL + log(N) * k where *log()* has the base-e called the natural logarithm, *LL* is the log-likelihood of the model, *N* is the sample size, and *k* is the number of parameters in the model.

[11] Child gender was the only time fixed variable also appropriate as a predictor of the exposure variable available in the dataset. Adding predictor variables with missing observations is not recommended.

[12] Each exposure sequence may have several gaps.

Instrument variable estimation

Our reduced form associations between *sleep disruption* and cognitive outcomes as carried out in model selection analysis are specified as below:

$$Cognitive\ outcomes_i = \beta_1 Sleep\ disruption_i + X_i'\beta_2 + \varepsilon_i \ (1)$$

Given that cumulative model of sleep disruption provides either the best or second best fit to all outcomes, we consider the cumulative model of exposure as the independent variable in our main analyses. This is for the ease of presenting and reading the results.

For all outcomes measured during 8 to 10 years of age, the cumulative measure for sleep is a sum of all binary indicators of sleep disruption from 1.5 up to 7 years of age. For outcomes measured during ages 12 to15, the cumulative measure for sleep is a sum of all binary indicators of sleep disruption from 1.5 up to 10 years of age. This is done so that exposure variables always constitute the past and not the current exposure. The vector X constitutes all controls that might confound the relationship between *Sleep disruption* and cognitive outcomes and ε_i corresponds to the random error term.

The estimate of $\hat{\beta}_1$ may be biased due to possible underlying unobservable conditions that may lead to both *sleep disruption* and poor cognitive outcomes (simultaneity).[13] To overcome possible endogeneity issues and identify the causal effect of *sleep disruption* on cognitive outcomes, we adopt an instrumental variable estimation. Specifically, we exploit mother/carer reported exposure to neighbor noise as an instrument for child sleep disruption (Fan and Weinhold, 2018). We use plausibly exogenous exposure to neighbor noise to predict sleep disruption and then examine the effect of this neighbor noise induced sleep disruption to estimate the causal effect on cognition outcomes.

To estimate the causal effect of sleep disruption on cognitive outcomes, I estimate a two stage least-squares (2SLS) model as below:

First stage:

$$Sleep\ disruption_i = \alpha_1 Neighbor\ noise_i + X_i'\alpha + \mu_i \ (2)$$

[13] Reverse causality is less likely to be an issue here as our exposure is measured in the past.

where $Neighbor\ noise_i$ is an index of severity of noise problem due to neighbors. Respondents were asked to respond to the question 'if noise from other homes is a problem for the family' at 2, 3, 5 and 7 years of age of child. Possible responses were serious problem; minor problem; no problem and no opinion. We consider neighbor noise to be a problem if respondents consider neighbor noise to be a serious or minor problem. 'No opinion' responses were considered to be missing. These responses were summed over the four time periods to construct an index of noise problem. A higher score on this index indicates higher severity of problems due to neighbor noise. The vector X corresponds to the covariates described in the section above.

Second stage:

$$Cognitive\ outcomes_i = \gamma_1 Sleep\ \widehat{disruption_i} + X_i'\gamma_2 + \pi_i \ (3)$$

Linear regression models are used in all analyses. The causal effect of sleep disruption on cognitive outcomes is identified by the parameter γ_1.

Scarring effects

Scarring effects are modeled using a linear regression model where a subset of outcomes measured during 8 to 10 years of age, namely full-scale, verbal and performance IQ measured at 8 years and SDQ measures at age 10 are considered.[14] The two key independent variables of interest include a cumulative measure of past sleep disruption (similar to the one used for the main estimation) and a measure of current sleep disruption measured at 10 years of age. Particularly, the model can be expressed as follows;

$$Y_{i\,(8.7-9.7\ years)} = \beta_0 + \beta_1\ Sleep\ disruption_{i\,(1.5-7\ years)} + \beta_2\ Sleep\ disruption_{i\,(9-10\ years)} + X_i'\beta_3 + \varepsilon_i \ (4)$$

The time period for the outcome variable includes the mean age at which 'Focus at 8' and 'Focus at 9' sessions took place. We expect both $\beta_1 < 0$ and $\beta_2 < 0$.

RESULTS

Table 2 displays the summary statistics for the exposure and outcome variables. Overall, four-fifth of the sample has experienced sleep disruption at any one time point in their childhood whereas persistent sleep disruption is observed in only one-third of the sample. The prevalence rates of sleep disruption during

[14] We do not include total SDQ at age 12 and IQ scores at age 15 years as we do not have data on sleep disruption measured at approximately the same age. This is important when testing for scarring effects.

infancy, early and mid-childhood ranges from 53-60%. Table 3 displays the summary statistics for other variables used in the analyses. A higher proportion of mothers belong to the age group 21-34 years (83%), and only less than 10% of the mothers report a university degree as the highest level of education. The proportion of mothers who were continuously employed across all time periods range from 10-14%. Around half of the children are male with only 5% being non-white. Around 38% of the children are first-borns with less than 10% of them reported to be ill at any time point in the sample.

In Table 4, the BIC estimates for the sleep disruption variables of the best fitting (and second best) model are compared to the alternate models described in Table 1. The first and second column displays the outcomes of interest and the age at which the outcome is measured. The second (third, fourth or fifth) column displays the BIC for the model with the best or second best fit. BIC values highlighted in bold represents the best fit. We provide both as it is common to have models with very similar estimates for the same outcome (Green et al., 2018). For IQ (full-scale, verbal and performance) measured at age 8 and full-scale IQ measured at age 15, 'critical period' model for sleep disruption during early childhood and cumulative sleep disruption models predicted the best fit. Cumulative and persistent models of sleep disruption provided the best fit for vocabulary IQ at age 15. No sleep disruption variables significantly predicted matrix IQ. For non-cognitive outcomes measured by total SDQ score, critical model for sleep disruption in mid-childhood and cumulative models provide the best fit. The fifth column shows that the BIC estimates for the remaining models is higher than for the best (and second-best) fitting model. Overall, critical, cumulative, and persistent models provide the best fit to the cognitive and non-cognitive outcomes under consideration.

Table 5a and 5b presents the reduced form relationship between sleep disruption and cognitive outcomes after controlling for our full set of covariates. We consider two alternative specifications, (1) and (2). The baseline results are presented in specification (1) whereas in specification (2) we investigate how adding controls for child sensitivity to noise (*Noise sensitivity*) and moving houses *(Ever moved*) to the baseline model influence the results. This additional check ensures that the baseline correlation is not driven by increased sensitivity to noise for a child or residential moves (Fan and Weinhold, 2018).

Results for specification (1) in Table 5a shows each additional episode of sleep disruption predicting a 0.82 point decrease in full-scale IQ ($p<0.001$), 0.89 point decrease in verbal IQ ($p<0.001$), and 0.55 point decrease in performance IQ ($p<0.05$) measured at 8 years of age. For every additional episode of sleep disruption, full-scale IQ and vocabulary IQ measured at 15 years of age reduces by 0.62 points ($p<0.05$) and 0.37 points ($p<0.05$) respectively. Results for specification (1) in Table 5b shows each additional episode of sleep disruption predicting a decrease in total SDQ measured at age 10 and 12 years by 0.76 points ($p<0.001$) and by 0.67 points ($p<0.001$) respectively.

Results for the robustness specification (2) in Table 5a and 5b shows that a higher sensitivity to noise is negatively related to all cognitive and non-cognitive outcomes whereas the relationship between having ever moved one's residence and all outcomes is statistically not significant. Our coefficients for *sleep disruption* are robust to the inclusion of these control variables suggesting that unobserved heterogeneity in sensitivity to noise and self-selection via moving houses do not drive the observed correlations between sleep disruption and cognitive outcomes. However, given the endogenous nature of *Noise sensitivity* and *ever moved*, the results should be interpreted with caution.

Tables 6a and 6b presents the IV estimation results for cognitive and non-cognitive outcomes respectively. Again, we consider two alternative specifications, (1) and (2). Specification (2) investigates how adding controls for child sensitivity to noise (*Noise sensitivity*) and moving houses *(Ever moved*) to the set of baseline covariates in specification (1) influence the results. Results from Table 6a and 6b shows statistically significant causal effects of sleep disruption on verbal IQ at age 8, full-scale and vocabulary IQ at age 15.5, total SDQ at age 10 and 12, respectively. Each additional episode of sleep disruption is associated with a decrease in verbal IQ at age 8 by approximately 10.79 points ($p<0.05$), full-scale IQ at age 15.5 by 9.58 points ($p<0.05$), vocabulary IQ at age 15.5 by 5.83 points ($p<0.05$), total SDQ at age 10 and 12 by 8.14 points ($p<0.001$) and 7.77 points ($p<0.001$) respectively. Our results are robust to the additional controls included in specification (2). Table 6a and 6b also shows the corresponding first stage relationship between $Neighbor\ noise$ and $Sleep\ disruption$. A higher severity of neighbor noise is positively correlated with sleep disruption ($p<0.001$). The tables also report the Kleibergen-Paap Wald F-Statistics are larger than 10, indicating the strength of the instrument.

Our 2SLS estimates are larger in magnitude when compared to the OLS estimates. This could be due to several reasons. One reason could be omitted variable bias as discussed in section 3.2. For example, children might be growing up in a stressful family environment (for which the data is unavailable) that might be positively correlated with sleep disruption but at the same time hindering the development of cognitive and non-cognitive skills. This combination of positive and negative correlations might bias the OLS estimates in a downward direction. Second, we rely on parent/carer-reported measures of sleep disruption which might be noisily measured when compared to more objective measures of sleep quality. Our instrument variable $Neighbor\ noise$ could also be subjected to measurement error but to a lesser extent when compared to $Sleep\ disruption$. Finally, the 2SLS estimator appliers to compliers (local treatment effect), which is the effect of sleep disruption driven by neighbor noise which could be more intensive. This effect might be diluted when averaged over all observations as done with the OLS estimator.

Validity of the instrument

The validity of 'exposure to neighbor noise' as an instrument variable requires the following three conditions to be met: relevance, exogeneity and the exclusion restriction. Relevance assumption requires that neighbor noise should significantly predict sleep disruption. Exogeneity requires that those households having children with poor cognitive or non-cognitive outcomes should not report more noisy neighbors or self-select into dwellings with noisier neighbours. Exclusion criteria require that the effect of noisy neighbours on child outcomes should actualize only through sleep problems. We investigate all the three assumptions.

First, we showed that neighbour noise significantly predicts child sleep disruption. Next, we provide qualitative evidence to show that our IV satisfies the exogeneity condition. A fall in house prices (and a downgrading of housing equity) contributed to a downward trend in residential mobility in the early 1990s in the UK (Henley, 1998). Using data from BHPS from 1992-2008; matched to area-level data on housing market, Ermisch and Washbrook (2012) studies the residential mobility of home-owners under 45 years of age in UK. Their findings show mobility in UK to be higher at younger ages, and most likely before individuals form partnerships and decide to begin child-rearing, suggesting fewer chances of residential mobility after having children as an overall trend. Consistent with Henly (1998), this study also confirmed the negative effect of fall in house prices and low housing equity on the likelihood of a move within UK (Ermisch and Washbrook, 2012). A much earlier study on patterns of residential mobility in the private housing market of Bristol in 1975 also showed similar patterns (Short, 1978). A higher proportion of movers were the young with newly formed households given that residential mobility was a response to changing space requirements. All of the studies mentioned above provide evidence to show that residential mobility was limited in our study period. In case households did move across neighborhoods, we find that moving to a new house is uncorrelated with cognitive outcomes of children.

For the exclusion restriction, we run the risk of over-estimating the magnitude of $Sleep\ disruption$ if we are unable to control for any channel (other than sleep disruption) through which neighbour noise may be causally related to cognitive and non-cognitive outcomes. Note that, our IV estimates are 5-10 times larger than the OLS estimates. To ensure that our IV estimates do not suffer from this bias, we provide evidence to show that moving houses or individual level variation in sensitivity to noise are not the primary channels through which neighbor noise affects cognitive and non-cognitive outcomes.

Additionally, we also control for a large set of socio-demographic, child and maternal characteristics to rule out any alternate first-order causal channels that may violate our exclusion restriction.

As a robustness check, we also control for potential channels through which neighbor noise might affect child human capital outcomes independent of sleep problems in children. Neighbor noise might also affect parental sleep, in turn affecting parental mood, family functioning and a potentially a less stimulating environment for the child. Kohen et al. (2008) draws a similar hypothesis starting with poorer neighbourhood cohesion affecting maternal behaviors such as depression or parenting practices which then may have adverse effects on the verbal skills of the child. To check if our IV estimates are robust to this additional pathway, we control for a measure of frequency of mother child activities. Mothers reported the frequency of time they spend doing the following activities with the child at ages 3, 5, 7 and 10 years: sings to the child; reads; plays (toys or imitation games); active or physical play; walks with the child; and other activities. Frequency was measured as almost daily, 3-5 times per week; 1-2 times per week; less the once per week and never. For our analysis, we take an average of the frequency of the activities at each age and then average them across the childhood.[15] Table A2 shows that our IV estimates are robust when controlling for this additional measure.

Nevertheless, we consider the possibility that neighbor noise might be correlated with human capital outcomes via channels other than sleep disruption. One such pathway could be manifested through the effect of neighbor noise on the unobserved measures of sleep disruption occurring between the last episode of sleep disruption measured and the age at which the corresponding outcomes are measured. To estimate the sensitivity of the 2SLS estimates to violations of exclusion restriction, I follow a methodology by Conley et al. (2012) and consider our IV to be only plausibly exogenous. Based on Nybom (2017), a plausibly exogenous IV is allowed to have a direct effect λ on the outcomes, thereby relaxing the exclusion restriction. We vary λ in an interval [0,1] where 0 indicates an IV which is perfectly exogenous whereas any value greater than 0 in the interval indicates a plausibly exogenous IV. This approach helps us to identify the threshold at which the 2SLS estimates from Table 6a and 6b is statistically insignificant at 10% level.

Figure A1 shows the 90% confidence intervals for the 2SLS estimates of the equation () for each value of λ plotted on the x-axis and varying in the interval [0,1]. In the figures, we see that the value 0 represented by the red line is not included in the confidence intervals, indicating that our IV estimates are robust to possible violations of exclusion restrictions.

[15] For outcomes measured at age 8, we include information from the first 3 time points when constructing the measure of frequency of mother child activities. For the remaining outcomes, we include information from all time points.

Table 7 presents the scarring effects of all sleep measures for a subset of human capital outcomes. Controlling for the current sleep problem, past cumulative sleep problem predicts lower full-scale, verbal and performance IQ at 8 years ($p<0.001$) and total SDQ at 10 years ($p<0.001$).

DISCUSSION

Sleep problems among children and adolescents is an increasingly prevalent problem today (Varghese et al., 2020). Existing medical and observational studies have tried to identify the potential causes and consequences of sleep problems. However, there is limited empirical (and causal) evidence on the dynamics of sleep problems in early childhood and how they influence later outcomes. Using a longitudinal cohort study from UK, this study, for the first time provides evidence on timing of exposure to sleep disruption and human capital outcomes. Our results show that the length of time exposed to sleep disruption across childhood increases the risk for adverse cognitive and non-cognitive outcomes and this may be driven by exposure to sleep disruption in early/mid childhood. Findings from scarring effects strengthen our conclusion that sleep in the past matters for later outcomes irrespective of the sleep in the present. Furthermore, using a novel methodology, we show that our findings are causal and highlights the cognitive penalty of noise induced sleep disruption

It is important to understand the meaning and implications of our findings regarding the model fit in order to locate the most appropriate time for intervention. For all IQ measures at age 8 and total IQ at age 15, both critical period model indicating early childhood and accumulation model provided the best fit. A good fit for critical period model means that early childhood is a period of exclusive risk during which exposure to sleep disruption may have lasting effects on IQ attainment. Outside this window of early childhood, there is no excess risk associated with sleep disruption for IQ attainment. On the other hand, a good fit for accumulation model indicates that exposure to sleep disruption gradually accumulates (irrespective of the timing) to increase the risk for a lower IQ score. Attaining very close BIC estimates for different models is common in the literature and could be attributed to a potential confounding between accumulation and critical periods of exposure (Hallqvist et al., 2004). For example, an accumulation model with no exposure to risk in infancy and mid-childhood but exposure in early childhood supports an accumulated dose of one, with early-childhood as a critical period. In this way, the accumulation and critical period hypothesis may be interrelated and empirically disentangling the mutual confounding effects using simple life course methods (such as the ones we use) might be difficult. The same reasoning applies to behavioral difficulties score for which critical period model indicating mid-childhood and accumulation model provided the best fit. Additionally, the slightly later critical period for non-cognitive outcomes compared to cognitive outcomes could be attributed to the differences in the developmental timeline of these skills. Non-cognitive skills develop and peak much later compared to cognitive skills and therefore have a later critical period for development and intervention (Borghans et al., 2008; Popli et al., 2013).

On balance, our findings indicate the need to intervene during early childhood for several reasons. One, early life has been shown to have the strongest impact on later outcomes, especially human capital development. Two, we provide evidence suggesting a causal effect of cumulative exposure to sleep disruption and therefore an intervention in early period will contribute to dampening the cumulative effect of this shock. Three, our findings corresponds with the timing of intervention indicated by some earlier literature on educational outcomes (Duncan et al., 1998) and medical literature suggesting evidence for sleep debt which accumulates over time (Dimitriou et al., 2015). Lastly, our findings on scarring effects of past sleep support the importance of intervening in early childhood as exposure to past sleep matters irrespective of the sleep quality at the time when the outcomes were measured.

Given the optimal time point of intervention, we next study the causal effect of cumulative measure of sleep disruption on cognitive and non-cognitive outcomes using an instrumental variable approach. Controlling for a range of socioeconomic, maternal and child characteristics, we show that sleep disruption instrumented by neighbor noise has a statistically significant causal effect on IQ and total SDQ measures. We show that the negative relationship between sleep disruption and human capital outcomes are not driven by any selection bias via moving houses or heterogeneity in sensitivity to noises among children and also robust to possible violations of exclusion restriction. We also find that the correlation and the causal relationship between sleep disruption and verbal/vocabulary IQ is larger compared to that of performance IQ or matrix subtest. Findings from longitudinal twin studies show that heritability in childhood and early adolescence was highest for performance IQ (64-72%) compared to full-scale IQ (34-65%) and verbal IQ (37-51%) and similarly for IQ measured during adolescence (Van Soelen et al., 2011). This implies that genetics significantly predicts a large variation in performance IQ and environment factors like noise induced sleep disruption may only have smaller effects on the same. Also, SDQ measures have much smaller heritability estimates (10-23%) which is also stable over time (Morris et al., 2021), and is therefore consistent with our OLS and 2SLS estimates which are larger than that of cognitive measures.

Finally, our paper also highlights the cognitive and non-cognitive penalty of being exposed to neighbor noise which has not received much attention as a public health issue. As shown by our findings, neighbour noise significantly predicts sleep disruption, which in turn affects human capital development. Potential solutions for intervention may therefore be supplemented with identifying and minimizing noise induced sleep disruption in childhood. In future work, we need to explore whether the same pattern of relationship between sleep disruption and human capital outcomes can be replicated in other samples. Future studies may also investigate the underlying mechanisms explaining the relationship as the traits developed across life course can contribute towards both the later exposure and final outcome. Understanding the features of these traits such as if they are remediable or malleable are important for effective interventions by policy makers.

BIBLIOGRAPHY

Adair LS, Fall CH, Osmond C, Stein AD, Martorell R, Ramirez-Zea M, Sachdev HS, Dahly DL, Bas I, Norris SA. Associations of linear growth and relative weight gain during early life with adult health and human capital in countries of low and middle income: findings from five birth cohort studies. The Lancet 2013;382; 525-534.

Almond D, Currie J. Killing me softly: The fetal origins hypothesis. Journal of economic perspectives 2011;25; 153-172.

Astill RG, Van der Heijden KB, Van IJzendoorn MH, Van Someren EJ. Sleep, cognition, and behavioral problems in school-age children: A century of research meta-analyzed. Psychological bulletin 2012;138; 1109.

Barker DJ. The fetal and infant origins of adult disease. BMJ: British Medical Journal 1990;301; 1111.

Beebe D. Neurobehavioral effects of childhood sleep-disordered breathing (SDB): A comprehensive review. Sleep 2006;29; 1115-1134.

Ben-Shlomo Y, Kuh D. 2002. A life course approach to chronic disease epidemiology: conceptual models, empirical challenges and interdisciplinary perspectives. (Ed)^(Eds). Oxford University Press; 2002.

Blunden S, Lushington K, Kennedy D, Martin J, Dawson D. Behavior and neurocognitive performance in children aged 5-10 years who snore compared to controls. Journal of Clinical and Experimental neuropsychology 2000;22; 554-568.

Borghans L, Duckworth AL, Heckman JJ, Ter Weel B. The economics and psychology of personality traits. Journal of human Resources 2008;43; 972-1059.

Boyd A, Golding J, Macleod J, Lawlor DA, Fraser A, Henderson J, Molloy L, Ness A, Ring S, Davey Smith G. Cohort profile: the ‘children of the 90s’—the index offspring of the Avon Longitudinal Study of Parents and Children. International journal of epidemiology 2013;42; 111-127.

Bub KL, Buckhalt JA, El-Sheikh M. Children's sleep and cognitive performance: a cross-domain analysis of change over time. Developmental psychology 2011;47; 1504.

Clark A, Georgellis Y, Sanfey P. Scarring: The psychological impact of past unemployment. Economica 2001;68; 221-241.

Clark AE, D’Ambrosio C, Barrazzetta M. Childhood circumstances and young adult outcomes: the role of mothers' financial problems. 2019.

Conley TG, Hansen CB, Rossi PE. Plausibly exogenous. Review of Economics and Statistics 2012;94; 260-272.

Cunha F, Heckman J. The technology of skill formation. American Economic Review 2007;97; 31-47.

Cusick S, Georgieff MK. The first 1, 000 days of life: the brain's window of opportunity. Retrieved from www. unicef-irc. org/article/958-the-first-1000-days-of-life-the-brainswindow-of-opportunity. html 2016.
Dewald JF, Meijer AM, Oort FJ, Kerkhof GA, Bögels SM. The influence of sleep quality, sleep duration and sleepiness on school performance in children and adolescents: A meta-analytic review. Sleep medicine reviews 2010;14; 179-189.

Dimitriou D, Le Cornu Knight F, Milton P. The role of environmental factors on sleep patterns and school performance in adolescents. Frontiers in Psychology 2015;6; 1717.

Duncan GJ, Yeung WJ, Brooks-Gunn J, Smith JR. How much does childhood poverty affect the life chances of children? American sociological review 1998; 406-423.

Dunn EC, Busso DS, Raffeld MR, Smoller JW, Nelson CA, Doyle AE, Luk G. Does developmental timing of exposure to child maltreatment predict memory performance in adulthood? Results from a large, population-based sample. Child abuse & neglect 2016;51; 181-191.

Duque V, Rosales-Rueda M, Sanchez F. 2018. How do early-life shocks interact with subsequent human-capital investments? Evidence from administrative data. (Ed)^(Eds), IZA World of Labor Conference. 2018.

Ermisch J, Washbrook E. Residential mobility: wealth, demographic and housing market effects. Scottish Journal of Political Economy 2012;59; 483-499.

Evans WN, Lien DS. The benefits of prenatal care: evidence from the PAT bus strike. Journal of Econometrics 2005;125; 207-239.

Fallone G, Acebo C, Seifer R, Carskadon MA. Experimental restriction of sleep opportunity in children: effects on teacher ratings. Sleep 2005;28; 1561-1567.

Fan Y, Weinhold D. Sleep Disruption and Health. 2018. Available a: https://fanyinus.weebly.com/uploads/1/1/7/3/117361814/sleep-disruption-and-health_20201016_v3_appendix.pdf

Fleche S. Teacher Quality, Test Scores and Non-Cognitive Skills: Evidence from Primary School Teachers in the UK. CEP Discussion Paper No. 1472. Centre for Economic Performance 2017.

Gibson M, Shrader J. Time use and productivity: The wage returns to sleep. 2014.

Golding J. Determinants of child health and development: the contribution of ALSPAC—a personal view of the birth cohort study. Archives of disease in childhood 2010;95; 319-322.

Goodman R. The Strengths and Difficulties Questionnaire: a research note. Journal of child psychology and psychiatry 1997;38; 581-586.

Goodman R. Psychometric properties of the strengths and difficulties questionnaire. Journal of the American Academy of Child & Adolescent Psychiatry 2001;40; 1337-1345.

Green MJ, Stritzel H, Smith C, Popham F, Crosnoe R. Timing of poverty in childhood and adolescent health: evidence from the US and UK. Social Science & Medicine 2018;197; 136-143.

Hallqvist J, Lynch J, Bartley M, Lang T, Blane D. Can we disentangle life course processes of accumulation, critical period and social mobility? An analysis of disadvantaged socio-economic positions and myocardial infarction in the Stockholm Heart Epidemiology Program. Social science & medicine 2004;58; 1555-1562.

Halpin B. Multiple imputation for categorical time series. The Stata Journal 2016;16; 590-612.
Heckman JJ. The economics, technology, and neuroscience of human capability formation. Proceedings of the national Academy of Sciences 2007;104; 13250-13255.

Heckman JJ, Borjas GJ. Does unemployment cause future unemployment? Definitions, questions and answers from a continuous time model of heterogeneity and state dependence. Economica 1980;47; 247-283.

Henley A. Residential mobility, housing equity and the labour market. The Economic Journal 1998;108; 414-427.

Jagnani M. 2018. Poor sleep: Sunset time and human capital production. (Ed)^(Eds). mimeo; 2018.

Jan JE, Reiter RJ, Bax MC, Ribary U, Freeman RD, Wasdell MB. Long-term sleep disturbances in children: a cause of neuronal loss. european journal of paediatric neurology 2010;14; 380-390.

Kahneman D. New challenges to the rationality assumption. Journal of Institutional and Theoretical Economics (JITE)/Zeitschrift für die gesamte Staatswissenschaft 1994; 18-36.

Kelly Y, Kelly J, Sacker A. Changes in bedtime schedules and behavioral difficulties in 7 year old children. Pediatrics 2013;132; e1184-e1193.

Killgore WD. 2010. Effects of sleep deprivation on cognition. (Ed)^(Eds), Progress in brain research, vol. 185. Elsevier; 2010.

Knudsen EI, Heckman JJ, Cameron JL, Shonkoff JP. Economic, neurobiological, and behavioral perspectives on building America's future workforce. Proceedings of the national Academy of Sciences 2006;103; 10155-10162.

Kocevska D, Rijlaarsdam J, Ghassabian A, Jaddoe VW, Franco OH, Verhulst FC, Tiemeier H. Early childhood sleep patterns and cognitive development at age 6 years: the Generation R Study. Journal of pediatric psychology 2017;42; 260-268.

Kohen DE, Leventhal T, Dahinten VS, McIntosh CN. Neighborhood disadvantage: Pathways of effects for young children. Child development 2008;79; 156-169.

Kolb B, Harker A, Gibb R. Principles of plasticity in the developing brain. Developmental Medicine & Child Neurology 2017;59; 1218-1223.

Lavy V, Schlosser A, Shany A. 2016. Out of africa: Human capital consequences of in utero conditions. (Ed)^(Eds). National Bureau of Economic Research; 2016.

Lekfuangfu WN, Powdthavee N, Clark AE, Ward G. Early maternal employment and non-cognitive outcomes in early childhood and adolescence: evidence from British birth cohort data. 2015.

Lidzba K, Wilke M, Staudt M, Krägeloh-Mann I. Early plasticity versus early vulnerability: the problem of heterogeneous lesion types. Brain 2009;132; e128-e128.

Mishra G, Nitsch D, Black S, De Stavola B, Kuh D, Hardy R. A structured approach to modelling the effects of binary exposure variables over the life course. International journal of epidemiology 2009;38; 528-537.

Morris TT, Smith GD, van Den Berg G, Davies NM. Consistency of non-cognitive skills and their relation to educational outcomes in a UK cohort. bioRxiv 2021; 470682.

Mputle LDP. 2019. Nurturing care during the first 1000 days of life: a systematic review. (Ed)^(Eds). North-West University (South-Africa); 2019.

Nissenbaum MA, Aramini JJ, Hanning CD. Effects of industrial wind turbine noise on sleep and health. Noise and Health 2012;14; 237.

Nybom M. The distribution of lifetime earnings returns to college. Journal of Labor Economics 2017;35; 903-952.

Pearce A, Hope S, Lewis H, Law C. Family structure and socio-emotional wellbeing in the early years: a life course approach. Longitudinal and life course studies 2014;5; 263-282.

Popli G, Gladwell D, Tsuchiya A. Estimating the critical and sensitive periods of investment in early childhood: A methodological note. Social Science & Medicine 2013;97; 316-324.

Row BW, Kheirandish L, Cheng Y, Rowell PP, Gozal D. Impaired spatial working memory and altered choline acetyltransferase (CHAT) immunoreactivity and nicotinic receptor binding in rats exposed to intermittent hypoxia during sleep. Behavioural brain research 2007;177; 308-314.

Rubin DB. 1987. Multiple imputation for survey nonresponse. (Ed)^(Eds). New York: Wiley; 1987.

Sadeh A, Gruber R, Raviv A. The effects of sleep restriction and extension on school-age children: What a difference an hour makes. Child development 2003;74; 444-455.

Sadeh A, Tikotzky L, Kahn M. Sleep in infancy and childhood: implications for emotional and behavioral difficulties in adolescence and beyond. Current opinion in psychiatry 2014;27; 453-459.

Short JR. Residential mobility in the private housing market of Bristol. Transactions of the Institute of British Geographers 1978; 533-547.

Spruyt K, Gozal D. Sleep in children: the evolving challenge of catching enough and quality Zzz's. Sleep, Health and Society: From Aetiology to Public Health. 1st ed. Oxford: OUP 2010; 215-238.

Tartar JL, Ward CP, McKenna JT, Thakkar M, Arrigoni E, McCarley RW, Brown RE, Strecker RE. Hippocampal synaptic plasticity and spatial learning are impaired in a rat model of sleep fragmentation. European Journal of Neuroscience 2006;23; 2739-2748.

Touchette É, Petit D, Séguin JR, Boivin M, Tremblay RE, Montplaisir JY. Associations between sleep duration patterns and behavioral/cognitive functioning at school entry. Sleep 2007;30; 1213-1219.

Van Soelen IL, Brouwer RM, Van Leeuwen M, Kahn RS, Pol HEH, Boomsma DI. Heritability of verbal and performance intelligence in a pediatric longitudinal sample. Twin Research and Human Genetics 2011;14; 119-128.

Varghese NE, Lugo A, Ghislandi S, Colombo P, Pacifici R, Gallus S. Sleep dissatisfaction and insufficient sleep duration in the Italian population. Scientific reports 2020;10; 1-8.

Wechsler D. Manual for the WISC-III. San Antonio, TX: Psychological Corporation 1991.

Wechsler D. Manual for the Wechsler abbreviated intelligence scale (WASI). San Antonio, TX: The Psychological Corporation 1999.

FIGURES

Figure 1a: Aggregation of *sleep disruption* information at different time points for cognitive and non-cognitive outcomes measured during age 8-10 years

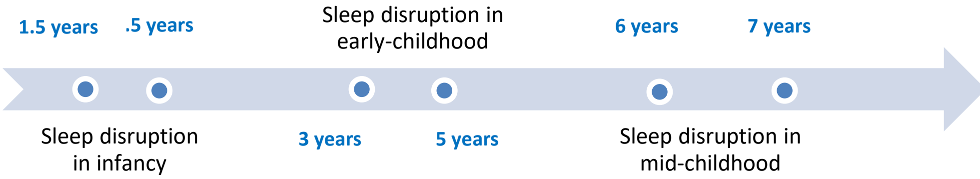
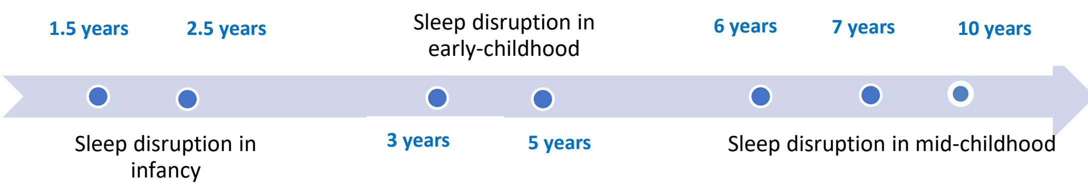


Figure 1b: Aggregation of *sleep disruption* information at different time points for cognitive and non-cognitive outcomes measured during age 12-15 years



TABLES

Table 1: Sleep disruption exposure variable definitions

Variable	Definition
Infancy sleep disruption	1 = sleep disruption at age 1.5 or 2.5 years 0 = no sleep disruption at age 1.5 or 2.5 months
Early childhood sleep disruption	1 = sleep disruption at age 3 or 5 years 0 = no sleep disruption at age 3 or 5 years
Mid-childhood sleep problem[+]	1 = sleep disruption at age 6 or 7 years 0 = no sleep disruption at age 6 or 7 years
Cumulative sleep disruption	Sum of infancy, early childhood and mid-childhood variables (range: 0-3)
Persistent sleep disruption	1 = sleep disruption in infancy, early childhood and mid-childhood 0 = all else
Any sleep disruption	1 = sleep disruption in infancy, early childhood or mid-childhood
Sleep trajectories	0 = all else
Early upward sleep disruption trajectory	1 = no sleep disruption in infancy but sleep disruption in early childhood 0 = all else
Early downward sleep disruption trajectory	1 = sleep disruption in infancy but no sleep disruption in early childhood. 0 = all else
Late upward sleep disruption trajectory	1 = no sleep disruption in early childhood but sleep disruption in mid-childhood 0 = all else
Late downward sleep disruption trajectory	1 = sleep disruption in early childhood but no sleep disruption in mid-childhood 0 = all else

[+]Mid childhood sleep disruption takes a value of 1 if sleep disruption is reported at 6, 7 or 10 years of age when the cognitive and non-cognitive outcomes are measured during 12-15 years of age.

Table 2: Descriptive statistics for outcome and exposure variables

Variables	Categories	% or $\bar{X}$	Min	Max	N
Sleep disruption variables					
Infancy sleep disruption		33.8	0	1	11225
Early childhood sleep disruption		24.4	0	1	10152
Mid-childhood sleep disruption		18.8	0	1	9519
Mid-childhood sleep disruption^{+}		22.5	0	1	9567
Cumulative sleep disruption	0	55.9	0	3	9433
	1	22.8			
	2	12.9			
	3	8.5			
Cumulative sleep disruption^{+}	0	54.8	0	3	9433
	1	22.8			
	2	12.9			
	3	9.6			
Persistent sleep disruption		7.5	0	1	10623
Persistent sleep disruption+		8.5	0	1	10620
Any sleep disruption		48.7	0	1	10276
Any sleep disruption+		49.9	0	1	10320
Sleep disruption trajectories					
Early upward trajectory		7.1	0	1	10557
Early downward trajectory		15.4	0	1	10820
Late upward trajectory		7.3	0	1	9796
Late upward trajectory^{+}		9.0	0	1	9793
Late downward trajectory		11.5	0	1	9875
Late downward trajectory^{+}		9.9	0	1	9926
Cognitive outcomes					
Verbal IQ at 8 years		99.5 (15.0)	45	142	7284
Performance IQ at 8 years		99.5 (15.0)	53	145	7276
Reading score at 9 years		7.5 (2.5)	0	10	7541
Spelling score at 9 years		10.2 (3.5)	0	15	7525
Total GCSE/GNVQ score at 16 years		341.9 (154.3)	0	898	11998
Five or more A*−C grades at GCSE at 16 years		50.4	0	1	11998

Note 1: Standard deviations for the means are presented in the brackets.

Note 2: The descriptive statistics for the exposure variables are derived from the imputed sample.

$^{+}$These are variables representing contrasting hypotheses when the outcomes are measured during ages 12-15 years.

Table 3: Descriptive statistics for covariates and other variables used in the analyses

Variables	Categories	% or $\bar{X}$	Min	Max	Sample size
Socio-economic characteristics					
Parental highest level of education			1	4	12744
	CSE/Vocational	20.7			
	O-level	29.1			
	A-level	30.4			
	Degree	19.8			
Household income			1	4	10688
	Q1	25.2			
	Q2	25.3			
	Q3	24.8			
	Q4	24.7			
Dwelling - material deprivation		0.53 (0.61)	0	3	14504
Maternal characteristics			1	3	13978
Mothers age at birth			1	3	13799
	<21 years	7.3			
	21-34 years	82.8			
	>35 years	9.9			
Maternal duration of employment		2.26 (1.74)	0	5	10940
Maternal duration of employment+		3.62 (1.81)	0	7	14504
Divorced or separated		15.0	0	1	11644
Divorced or separated+		17.0	0	1	14504
Prenatal smoking		21.3	0	1	11199
Child characteristics					
Age at 'Focus at 8 session' (in months)		103.82 (3.91)	89	127	7385
Age at 'Focus at 9 session' (in months)		118.49 (3.89)	105	140	7621
Age at start of KS4 academic year (in years)		15.01 (0.09)	15	16	11998
Female		48.7	0	1	14504
Ethnicity- non-white		5.0	0	1	11927
First-born		38.2	0	1	14504
Low-birth weight (<2500g)		4.9	0	1	13691
Ever breast fed since 1 year of birth		73.9	0	1	10857
Child health-unwell		9.3	0	1	14504
Other variables					
Ever moved		49.5	0	1	11815
Sensitivity to noise		2.81 (0.71)	1	5	10234
Instrumental variable					
Index of severity of neighbor noise		0.52 (0.94)	0	4	14504

Note: Standard deviations for the means are presented in the brackets.

+These are variables representing contrasting hypotheses when the outcomes are measured during ages 12-15 years.

Table 4: Model fit statistics

		Bayesian Information Criteria					N
Model		Early childhood sleep disruption	Mid childhood sleep disruption	Cumulative sleep disruption	Persistent sleep disruption	Other models^	
Outcomes	Age (years)						
Total IQ	8.5	**42528.1**		42529.4		>42529.4	5276
Verbal IQ	8.5	42741.4		**42740.2**		>42741.4	5299
Performance IQ	8.5	**43226.4**		43228.2		>43228.2	5292
Total IQ	15.5	29201.0		**29200.2**		>29201.0	3615
Vocabulary IQ	15.5			**28071.6**	28072.5	>28072.5	3866
Matrix IQ	15.5	-	-	**-**	-	-	-
Total SDQ (Parent reported)	10		36145.1	**36141.3**		>36145.1	6105
Total SDQ (Parent reported)	12		**33017.7**	33027.9		>33027.9	5547

Note: We provide best and second best BIC estimates for each outcome. Best fit BIC estimates are displayed in bold.

^Other models encode for different hypotheses described in Table 1 and not provided in Table 4.

*None of the models for matrix IQ were statistically significant at 5% level of significance and hence omitted.

All models are adjusted for maternal age at delivery, gender of child, household income in categories, highest parental education, indicator for first born, material deprivation, mother's duration of employment, marital status, age and ethnicity of child, indicator for low birth weight, ever breast fed and health status of the child.

Table 5a: Correlation between sleep disruption and cognitive outcomes-Reduced form regressions

VARIABLES	Full-scale IQ		Verbal IQ		Performance IQ		Full-scale IQ		Vocabulary IQ	
Age	8 years		8 years		8 years		15 years		15 years	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Cumulative index of sleep disruption	-0.817***	-0.803***	-0.894***	-0.906***	-0.547**	-0.514**	-0.619**	-0.628**	-0.373**	-0.376**
	(0.209)	(0.213)	(0.209)	(0.212)	(0.221)	(0.225)	(0.251)	(0.257)	(0.166)	(0.170)
Noise sensitivity		-0.834***		-0.869***		-0.543*		-0.838**		-0.677***
		(0.274)		(0.275)		(0.287)		(0.326)		(0.210)
Ever moved		-0.532		-0.533		-0.315		0.486		0.314
		(0.381)		(0.381)		(0.402)		(0.466)		(0.300)
Observations	5,276	5,132	5,299	5,154	5,292	5,148	3,615	3,526	3,866	3,770

Robust standard errors in parentheses.
All models are adjusted for maternal age at delivery, gender of child, household income in categories, highest parental education, indicator for first born, material deprivation, mother's duration of employment, marital status, age and ethnicity of child, indicator for low birth weight, ever breast fed and health status of the child.
*** p<0.01, ** p<0.05, * p<0.1

Table 5b: Correlation between sleep disruption and non-cognitive outcomes-Reduced form regressions

VARIABLES	Total SDQ		Total SDQ	
Age	10 years		12 years	
	(1)	(2)	(3)	(4)
Cumulative index of sleep disruption	-0.764***	-0.740***	-0.667***	-0.668***
	(0.072)	(0.073)	(0.073)	(0.075)
Noise sensitivity		-0.421***		-0.403***
		(0.087)		(0.093)
Ever moved		-0.077		0.005
		(0.122)		(0.131)
Observations	6,105	5,924	5,547	5,406

Robust standard errors in parentheses.
All models are adjusted for maternal age at delivery, gender of child, household income in categories, highest parental education, indicator for first born, material deprivation, mother's duration of employment, marital status, age and ethnicity of child, indicator for low birth weight, ever breast fed and health status of the child.
*** p<0.01, ** p<0.05, * p<0.1

Table 6a: Instrumental variables estimation for cognitive outcomes: Two stage least squares

VARIABLES	Full-scale IQ		Verbal IQ		Performance IQ		Full-scale IQ		Vocabulary IQ	
Age	8 years		8 years		8 years		15 years		15 years	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Cumulative index of sleep disruption	-5.990	-6.008	-10.785**	-10.639**	1.077	0.884	-9.584**	-9.478**	-5.829**	-6.033**
	(4.404)	(4.394)	(4.950)	(4.908)	(4.215)	(4.188)	(4.127)	(3.966)	(2.608)	(2.558)
Noise sensitivity		-0.877***		-0.979***		-0.533*		-0.790**		-0.671***
		(0.296)		(0.335)		(0.288)		(0.377)		(0.242)
Ever moved		-0.243		0.012		-0.391		0.997*		0.592
		(0.471)		(0.532)		(0.458)		(0.588)		(0.368)
First stage:										
Neighbour noise	0.046***	0.047***	0.047***	0.048***	0.047***	0.048***	0.065***	0.068***	0.066***	0.068***
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.016)	(0.016)	(0.015)	(0.015)
K-P F statistic	12.68	12.74	13.32	13.38	13.39	13.48	17.28	18.59	18.52	19.64
Observations	5,276	5,132	5,299	5,154	5,292	5,148	3,615	3,526	3,866	3,770

Robust standard errors in parentheses
All models are adjusted for maternal age at delivery, gender of child, household income in categories, highest parental education, indicator for first born, material deprivation, mother's duration of employment, marital status, age and ethnicity of child, indicator for low birth weight, ever breast fed and health status of the child.
*** p<0.01, ** p<0.05, * p<0.1

Table 6b: Instrumental variables estimation for non-cognitive outcomes: Two stage least squares

VARIABLES	Total SDQ		Total SDQ	
Age	10 years		12 years	
	(1)	(2)	(3)	(4)
Cumulative index of sleep disruption	-8.141***	-7.908***	-7.765***	-8.163***
	(2.369)	(2.238)	(2.199)	(2.279)
Noise sensitivity		-0.469***		-0.390**
		(0.148)		(0.157)
Ever moved		0.234		0.418
		(0.231)		(0.263)
First stage:				
Neighbour noise	0.046***	0.048***	0.051***	0.052***
	(0.012)	(0.012)	(0.013)	(0.013)
K-P F statistic	14.53	15.69	16..12	(16.21)
Observations	6,105	5,924	5,547	5,406

Robust standard errors in parentheses

All models are adjusted for maternal age at delivery, gender of child, household income in categories, highest parental education, indicator for first born, material deprivation, mother's duration of employment, marital status, age and ethnicity of child, indicator for low birth weight, ever breast fed and health status of the child.

***p<0.01, **p<0.05, *p<0.1

Table 7: Scarring effects of exposure to sleep deprivation

	Cumulative sleep measure from 7 years		Current sleep measure at 9-10 years		
Outcomes at 8-10 years^	Coefficient	P-value	Coefficient	P-value	N
Full-scale IQ	-0.797	P < 0.001	-0.279	NS	5276
Verbal IQ	-0.832	P < 0.001	-0.833	NS	5299
Performance IQ	-0.573	P < 0.05	0.345	NS	5292
Total SDQ	-0.573	P < 0.001	-2.652	P < 0.001	6090

Note: All models are adjusted for maternal age at delivery, mother's duration of employment, marital status, sex of the child, age, ethnicity, binary variables indicating first born, low birth weight, ever breast fed and health status of the child respectively. Socio-demographics include household income in quartiles, highest parental education and material deprivation.
Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

APPENDIX

Fig A1: Relaxing instrument exogeneity

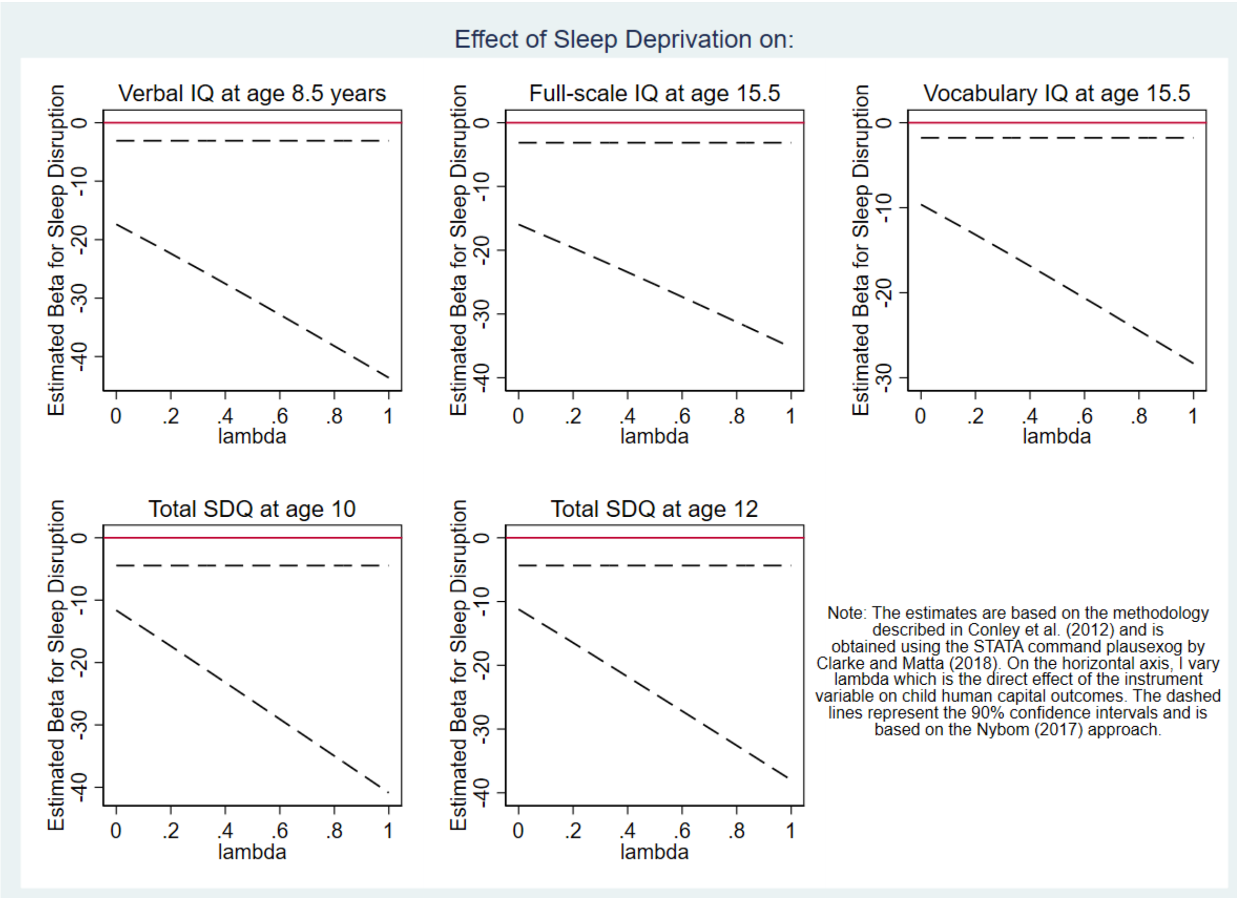


Table A1: Strengths and Difficulties Questionnaire (SD)

Please think about this child's behaviour over the last 6 months if you can

This child:	Not True	Somewhat True	Certainly True
1. Is considerate of other people's feelings [PS]	2	1	0
2. Is restless, overactive, cannot stay still for long [H]	0	1	2
3. Often complains of headaches, stomach-aches or sickness [E]	0	1	2
4. Shares readily with other children, for example toys, treats, pencils [PS]	2	1	0
5. Often loses temper [B]	0	1	2
6. Is rather solitary, prefers to play alone [P]	0	1	2
7. Is generally well behaved, usually does what adults request [B]	2	1	0
8. Has many worries or often seems worried [E]	0	1	2
9. Is helpful if someone is hurt, upset or feeling ill [PS]	2	1	0
10. Is constantly fidgeting or squirming [H]	0	1	2
11. Has at least one good friend [P]	2	1	0
12. Often fights with other children or bullies them [B]	0	1	2
13. Is often unhappy, depressed or tearful [E]	0	1	2
14. Is generally liked by other children [P]	2	1	0
15. Is easily distracted, concentration wanders [H]	0	1	2
16. Is nervous or clingy in new situations, easily loses confidence [E]	0	1	2
17. Is kind to younger children [PS]	2	1	0
18. Often lies or cheats [B]	0	1	2
19. Is picked on or bullied by other children [P]	0	1	2
20. Often volunteers to help others (parents, teachers, other children) [PS]	2	1	0
21. Thinks things out before acting [H]	2	1	0
22. Steals from home, school or elsewhere [B]	0	1	2
23. Gets along better with adults than with other children [P]	0	1	2
24. Has many fears, is easily scared [E]	0	1	2
25. Has a good attention span, sees chores or homework through to the end [H]	2	1	0

Notes: [E], [B], [H], [P], and [PS] respectively refer to the following dimensions of the SDQ: "Emotional health", "Behaviour problems", "Hyperactivity/inattention", "Peer relationship problems", and "Pro-social behaviour". The internalising SDQ is the sum of the "Emotional health" and "Peer relation problems" scores and the externalising SDQ is the sum of the "Hyperactivity/inattention" and "Behaviour problems" scores. The total SDQ is the sum of the internalising and externalising SDQ.

Table A2: Robustness check for 2SLS estimates (controlling for frequency of mother-child activities)

VARIABLES	Verbal IQ	Full-scale IQ	Vocabulary IQ	Total SDQ	Total SDQ
Age	8 years	15 years	15 years	10 years	12 years
Cumulative index of sleep disruption	-15.858**	-13.335**	-8.457**	-15.326***	-14.127***
	(8.026)	(5.902)	(3.872)	(5.592)	(5.081)
Observations	5,254	3,585	3,835	6,042	5,494

Note: All models are adjusted for maternal age at delivery, mother's duration of employment, marital status, sex of the child, age, ethnicity, binary variables indicating first born, low birth weight, ever breast fed and health status of the child respectively. Socio-demographics include household income in quartiles, highest parental education and material deprivation. Additionally, the models control for the frequency of mother-child activities across childhood.

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

CHAPTER 2

Hospital Closure and Patient Outcomes: Evidence from AMI emergency admissions in Italy[16]

ABSTRACT

In the European Union (EU), cost containment strategies have been particularly targeted towards reorganization of hospital networks given that hospitals account for large shares of healthcare budgets. Despite the increasing trend of hospital closures and mergers, few empirical studies have provided credible evidence to resolve the ambiguity regarding the consequences on patient outcomes. This paper uses a difference-in-differences approach to study the causal effect of hospital closures in Italy from 2008-2015 on AMI patient outcomes including in-hospital mortality, 30, 90 and 365-day cardiac and cardiology related readmission and length of stay. Patient outcomes for AMI admissions from municipalities exposed to versus not exposed to a home hospital closure are compared before and after the closure year. The analysis is conducted at both individual and municipality level. Results show that hospital closures increase in-hospital mortality by 1.2% and length of stay by about 0.4 days. Home hospital closures increase cardiac/circulatory related and AMI compatible 30-day, 90-day and 365-day readmissions for AMI by about 2.7% to 5.1% , however these results are statistically significant for only some specifications. The effects for in-hospital mortality are consistent with that of the individual level estimates. The effect of hospital closure on in-hospital mortality and readmission is persistent across the post-closure years indicating that adaptation to this negative shock (closure) in the availability of public health care may take a long time. We further check two potential mechanisms, increased travel time and congestion, and show that both mechanisms are relevant in explaining the causal effect of hospital closures on AMI outcomes in the case of Italy.

---

[16] This is a joint project with Simone Ghislandi[1], Anna Renner[2], Benedetta Scotti[1]. This is a working paper and we provide the latest version of the manuscript.

[1]Bocconi University, Milan, Italy
[2]Technische Universität Wien, Vienna, Austria

INTRODUCTION

Addressing increasing expenditure in the healthcare sector is a key policy challenge among most modern welfare states. In the European Union (EU), cost containment strategies have been particularly targeted towards reorganization of hospital networks given that hospitals account for large shares of healthcare budgets (Aiken et al., 2001; Schwierz, 2016). Such restructuring policies mainly involve hospital closures or merging of several hospitals into a larger units potentially leading to the closure of one or more hospital sites[17] (Arcà et al., 2020; Burkey et al., 2017; Dranove and Lindrooth, 2003; Wenzl et al., 2017). Historically, there has been an increasing trend of hospital closures or downsizing of hospital capacity in the EU since the 1960s (Aiken et al., 2001), which has been revived following the financial crises of 2008. However, it has been pointed out that apart from cutbacks in public spending, there could be other external (market) or internal (hospital) reasons motivating hospital or emergency department (ED) closures (Kaufman et al., 2016). Targeted hospital (wards) are therefore often characterized by low volume and reduced efficiency, leading to improved patient outcomes after a site-closure due to economies of scale and scope (Gujral and Basu, 2019). On the other hand, the closure of hospitals, especially remote ones, may hinder patients' access to care (Burkey et al., 2017). This may operate via longer travel times or increased case load in close-by hospitals and is especially crucial for those patients that require timely provision of emergency medical services (Avdic, 2016; Avdic et al., 2018; Bentham, 1986). Making use of individual-level panel data, we are able to address this conceptual inconclusiveness to study the effects of hospital closures in Italy from 2008 to 2015 on patient outcomes after an acute myocardial infarction (AMI).

Despite the increasing trend of hospital closures and mergers, very few empirical studies have provided credible evidence to resolve the ambiguity regarding the consequences on patient outcomes. This might be in part due to the econometric challenges accompanying the setup of this problem. First, a bias might emerge from unobservable characteristics that differ between patient groups exposed and not exposed to closures, respectively. This is because closures are not necessarily exogenous events but might occur due to low quality and efficiency before the closure. Second, the true effect of closures on patient outcomes might be contaminated by the changes in patient composition mix and service delivery patterns in the remaining open hospitals. The scarcity of studies may also be attributed to a lack of reliable administrative data on hospital closures or of centralized, yearly updated lists of operating hospitals.

[17] Hospital mergers may or may not lead to hospital closures depending on if they share a common physical facility (Dranove and Lindrooth, 2003).

Existing studies have tried to overcome the empirical challenges by relying on a variety of quasi-experimental methods. Likewise, data limitations were circumvented by researchers collating, sometimes for the first time, a list of hospital or ward-specific closures and linking several data sets to minimize validity concerns. For example, Avdic (2016), using Swedish patient level data recorded between 1990 and 2010, exploited the variation in closure induced changes in geographical distance to find increased mortality among patients who experienced an AMI. Similarly, Avdic et al. (2018) reported adverse child-birth outcomes following a series of hospital mergers in Sweden. On the contrary, the same authors found that closure of cancer surgical wards in Sweden was associated with improvements in patient outcomes (Avdic et al., 2014). The study exploits a regional variation in closures of cancer clinics, and identified learning-by-doing effects emerging from an increase in surgery volumes to be a potential mechanism underlying the relationship. Literature also shows evidence for no effects of hospital closures on patient outcomes. Using a matching approach to construct a control group, a study investigating hospital mergers in England between 1997 and 2006 reports no changes in patient outcomes (Gaynor et al., 2012). Similarly, Grytten et al. (2014) analyze local hospital closures in Norway using a propensity score matching method to find no effects on neonatal and infant mortality. These trends of contradictory findings were found in the studies conducted in the US.

Buchmueller et al. (2006) study hospital closures in the Los Angeles county between 1997 and 2003 and find an increase in mortality from heart attacks and unintentional injuries among patients who experienced an increase in distance to the closest hospital. More recent studies including Gujral and Basu (2019), Song and Saghafian (2019), Carroll (2019) also report hospital closures in several US states to negatively impact patient welfare. On the other hand, a number of studies show closure of inefficient hospitals to increase efficiency via an increase in inpatient volume which in turn reduces the cost per admission (Capps, 2005; Lindrooth et al., 2003). Further studies in the US context find no significant differences in patient outcomes, including mortality and hospitalization rates resulting from hospital closures (Hsia et al., 2012; Joynt et al., 2015; Rosenbach and Dayhoff, 1995).

Conceptually, a detrimental effect of hospital closures on patient outcomes can be expected to be most pronounced for time-sensitive conditions such as AMI, stroke, accidents or child-birth owing to increased travel times to the hospital (Gujral and Basu, 2019). For instance, in case of an AMI event, blood clots are formed in the coronary arteries leading to oxygen deficiency in the heart, disrupting the blood supply to all parts of the body. After five minutes of the event, the body begins to experience damage and death is practically unavoidable if treatment is not ensured within 15 minutes (Antman, 2008; Avdic, 2016; Ryan et al., 1996). Similarly, unanticipated complications are common during the period towards birth and require timely and efficient care to ensure the health of mother and the newborn (Avdic et al., 2018).

Given the existing findings, we use emergency AMI admissions recorded between 2008 and 2015 in Italy to provide new evidence on the causal effect of hospital closures on patient outcomes and the potential mechanisms underlying this relationship. First, we employ a staggered difference-in-difference approach, and show that hospital closures significantly increase the probability of in-hospital mortality, increase cardiac/circulatory and AMI compatible 30-day, 90-day, 365-day hospital readmissions and increase length-of-stay for AMI admissions at the individual and municipality level. Next, we use a two-step approach similar to Heckman et al. (2013) to provide evidence on two potential mechanisms through which hospital closures may affect patient outcomes: travel time and congestion (i.e. lack of resources proxied by bed utilization rate). In line with this recent literature (Avdic, 2016; Buchmueller et al., 2006), we find a strong positive association between hospital closures and increased travel time and congestion, which translates to worse AMI outcomes, particularly in-hospital mortality.

Besides the high policy relevance of the relationship between austerity measures in the public sector and population wellbeing (Depalo, 2019; Karanikolos et al., 2013; Quaglio et al., 2013; Reinhard et al., 2018; Stuckler et al., 2017), our paper contributes to the existing literature in three ways. First, we are able to identify a causal effect of hospital closures due to public austerity considerations on health outcomes based on a common and often deadly medical condition. Second, we are able to disentangle two potential mechanisms explaining this effect, travel time and congestion. This is crucial for the efficient allocation of limited resources without jeopardizing patient health. Finally, this is the first study investigating the effect of hospital closures on patient outcomes in an Italian context. Italy represents an interesting case study since in the aftermath of the Great Recession of 2008 it adopted tough fiscal consolidation policies, also affecting the public healthcare system (Arcà et al., 2020). As such, it gives the opportunity to evaluate the health consequences of austerity-induced cost containment strategies.

BACKGROUND AND INSTITUTIONAL CONTEXT

In Italy, healthcare is a fundamental right of the citizen and guarantees free national medical coverage (with some co-payment) to all. In 1978, the Italian government established the National Health Service (NHS) based on the Beveridge model where healthcare is financed by general taxation (Brenna, 2011). Since then, the system has been through some substantial reforms which led to the transfer of fiscal, financial and managerial responsibilities to the regional level (Mauro et al., 2017).

According to the OECD, in 2017 the per capita health expenditure in Italy was about 15% below the EU average (OECD, 2019). General health spending in 2017 was 8.8%, one percentage point below the EU average of 9.8%. Although the system went through many reforms in the nineties, the tightest containment policies were experienced during the first decade of the 2000s. According to Eurostat data, for example, from 2012 to 2017 Italy was among the bottom three countries in Europe in terms of health expenditure growth.

Cutback management strategies were implemented at all levels. Following the fiscal federalism reforms, the Ministry of Health fixed the spending targets which were then adjusted according to the regional healthcare needs and goals. Soon after, ten regions (Abruzzo, Molise, Apulia, Campania, Calabria, Sicily, Lazio, Piedmont, Sardinia and Liguria) ran into budget deficits due to shortfalls in management skills and service levels (Arcà et al., 2020). As a consequence, a set of budget recovery plans (*Piani di Rientro*, PdR) were introduced in 2006 (and ongoing) in order to restore the financial and economic balance in the budget deficit regions. This region-specific recovery schemes included active involvement by the Ministry of Health in setting spending targets, designing and evaluation of healthcare services delivery via a system of ex-ante and ex-post monitoring and tightening of overall autonomy in healthcare decision making at the regional level (Aimone Gigio et al., 2018; Depalo, 2019). Although the reforms were implemented in a staggered manner with varying intensities across the PdR regions, these financial recovery plans resulted in an overall reduction in number of hospital beds, workforce, number of hospitals and hospitalization rates (Mauro et al., 2017). Regarding their consequences, the scientific literature has not yet reached a consensus. While Arcà et al. (2020), Depalo (2019) and Bobini et al (2019) showed that the PdR scheme was successful in reducing the healthcare expenditure, but simultaneously resulted in a worsening of health outcomes, Bordignon et al (2020) conclude that budget containment policies increased efficiency by reducing expenditure with no relevant health effects on the affected population.

As these cost containment measures were rolled out, Italy was also struck by the 2008 financial crisis which escalated the public debt and drastically reduced the health care expenditure (6% in 2000-2007 to 2.3% in 2008-2010) (Falco, 2019). In addition to an increase in out of pocket payments and increased waiting times, there was a significant reduction in both hospitals and hospital beds during the crisis (Falco, 2019). According to the OECD statistics on health care resources, Italy witnessed around 144 hospital closures between 2008 and 2015 whereas total hospital beds per 1000 inhabitants went down from 3.8 in 2008 to 3.2 in 2015 (3.2 to 2.6 for acute care beds) (OECD, 2018). So far, no studies have analyzed the impact of hospital closures on health outcomes in Italy.

DATA AND EMPIRICAL STRATEGY

Data sources

We use data from various sources to build our final admission-level analytic file, as is detailed in Figure 1. We start with the hospital discharge data from the Ministry of Health for the years 2008-2015 with information on all patients admitted for AMI during this time period in Italy. This data set consisting of patient-level characteristics is then merged to several data sets that provide the final dataset with information on hospital, municipality level information and treatments. This includes the hospital level data that includes the characteristics of hospital such as the name of the hospital, geographical and administrative identifiers of the hospital and type of hospital (public/private). This is then linked to the dataset on beds available from the Italian ministry of health and the dataset on municipality-to-municipality minimum travel time which was constructed for this study. Finally, we merge this on to the data on population statistics ad income at municipality level available on the ISTAT and Ministry of Economy websites respectively.

[Figure 1 about here]

The outcome variables of our econometric analyses include inpatient mortality, cardiac/circulatory related and AMI compatible 30–day, 90-day, 365-day readmission for AMI and length of stay. Inpatient mortality is a binary variable indicating whether an AMI hospitalization resulted in a death or not. The second outcome is a dummy measuring whether an AMI hospitalization resulted in a readmission for cardiac or circulatory or AMI compatible reasons (see Table S1) that happened within 30, 90 and 365 days of the index AMI admission. We exclude planned admissions when constructing the readmission variables. The third outcome is the length of stay for each episode of AMI hospitalization.

Definition of home hospital and hospital closure

In order to identify patients affected by a hospital closure it is necessary to define a 'home hospital' for each patient. A home hospital is attributed to each patient using AMI hospitalizations from the Hospital discharge file and the data on travel time for each municipality-of-residence/municipality-of-hospital combination. First, all AMI admissions are aggregated to municipality level for each year. Following Avdic (2016), home hospital is the modal hospital to which most of the residents of the municipality were admitted to in a given year (Avdic, 2016). When there are ties between two hospitals that receive the same number of admissions in a year, we choose the hospital that is the closest (i.e. minimum travel time by car between the municipalities' centroids) from the municipality of residence. If two hospitals still compete to be a home hospital for a municipality, we choose one of them at random.

In a next step we identify which of the home hospitals closed during our study period. Since hospital closures have never been studied in the context of Italy, a centralized list of hospital closures does not exist. We rely on previous literature on hospital closures in the context of other countries for our definition of closures (Avdic, 2016; Gujral and Basu, 2019; Song and Saghafian, 2019). Using patient discharge data from Italian Ministry of Health, we consider hospitals to be closed if the volume of emergency AMI admissions in a hospital reduces by 90% or less between two years and stays the same in the subsequent years in the study sample[18]. Any remaining observations in the closed hospitals are dropped from the sample. We define the closure year as the year prior to the actual year of closure (Gujral and Basu, 2019).

The list on hospital closures was then reconfirmed using several external sources including Italian newspapers, online articles on the list of hospitals under risk of closure and other related news or documents on the official website for Italian Ministry of Health. We also went through the websites of closed hospitals (if available) to check for any news on renovation, mergers or closures. A total of 47 hospitals during 2008-2015 were identified to have had admissions drop to 10% or less and therefore identified as closures. Out of these, a total of 43 hospital closures were considered as home hospital based on the definition outlined above (see Table 1 for details). Our list of closures which is mainly based on a year to year comparison of AMI admissions should overcome any discrepancies arising from a simple collation of hospitals that were reported to be closed from various sources. Similarly, the definition of closure year as the year prior to closure captures the earliest interruptions in any services provided. In the year prior to closure, not all services are necessarily cut off and this coupled with the definition of closures (90% reduction in AMI) indicates that our main estimates reported in the paper are conservative (Gujral and Basu, 2019; Troske and Davis, 2019).

Analysis

Given the variation in timing of hospital closures between 2008-2015, we use a "staggered" difference-in-differences (DID) or DID with multiple time periods (Stevenson and Wolfers, 2003) with municipality and year fixed effects to examine changes in patient outcomes before and after a hospital closure event. This approach allows us to control for observed and unobserved heterogeneity between the treatment and control group that is constant over time. If the parallel trend assumption is met, DID analysis provide a

[18] Given the system of coding hospitals at the national level, the same hospitals are sometimes assigned different codes in different years, or two different hospitals are assigned same codes in different years. Mergers between hospitals at the same location are also assigned different codes, indicating a false closure. To deal with this, we manually identify hospitals over the years by matching them on hospital names, address, and postcodes (and telephone numbers). Cases that had inputting errors or inconsistencies in entries after the first stage were double checked using external sources. Finally, we used the trend of the admission volume to double check the final list of hospitals over time.

causal interpretation of the treatment effect. The treatment is defined at the municipality level. We define a municipality to be treated if it experiences at least one home hospital closure. Figure 2 shows the geographical distribution of treated municipalities based on the definition of 90% decrease in admissions.

The model is estimated at the individual level, and it reads as follows:

$$Y_{ihjt} = \alpha_0 + \alpha_1(Treat\ Post_{jt}) + \boldsymbol{X}\beta'_{ihjt} + \boldsymbol{H}\gamma'_{hjt} + \boldsymbol{M}\mu'_{jt}\,\delta_j + \lambda_t + \epsilon_{ihjt} \#(1)$$

where Y_{ihjt} is a binary variable indicating a) whether an admission i to hospital h from municipality j resulted in death at time t, or b) whether an individual admission i to hospital h from municipality j occurred within 30, 90 and 365 days from the previous index admission (30-day, 90-day and 365-day readmission), or a continuous variable indicating the length of stay. $Treat\ Post_{jt}$ indicates whether an admission is from an affected municipality j and occurs in the post closure period, α_1 being the parameter of interest. $\boldsymbol{X}_{ihjt}$, $\boldsymbol{H}_{ht}$ and $\boldsymbol{M}_{jt}$ are admission, hospital and municipality level covariates that can impact health outcomes. Admission-level controls include age, gender, marital status, education level and Elixhauser comorbidities index (Elixhauser et al., 1998). The comorbidities index is constructed as the weighted sum of the presence of 16 secondary diagnoses[19]. The index is then transformed into a categorical variable (a sum of 0 or 1 is indicated by 0 and 1 respectively and a sum of 2 or above is indicated by 2). Hospital level controls include type of hospital and volume of AMI hospitalizations per hospital-year. Municipality level controls include resident population and the share of resident population above 65. Municipality and year fixed effects are used in all regressions, whereas macro-region year fixed effects are only included in some individual level specifications. Information on local economic deprivation is considered for sensitivity analysis. Local economic deprivation is measured by income per capita and by the share of resident population living below the poverty threshold. Measures of economic deprivation are constructed from taxpayers' data provided by the Ministry of Interior and are available for years 2010-2015.

Municipality and year fixed effects are captured by terms δ_j and λ_t, respectively. The estimates for in-hospital death and 30-day, 90-day and 365-day readmission are obtained through linear probability models whereas the estimate for average length of stay is obtained through linear regression models. Robust standard errors are clustered at the municipality level.

[19] 16 secondary diagnoses include congestive heart failure, peripheral vascular disease, cerebrovascular disease, dementia, chronic pulmonary disease, rheumatologic disease, peptic ulcer disease, mild liver disease, diabetes without and with chronic complications, hemiplegia or paraplegia, renal disease, cancer, moderate or severe liver disease, metastatic carcinoma and AIDS/HIV.

In the second step of the analysis, we extend our model by including leads and lags of the treatment effect to check if the effects of hospital closures were consistent over the long term or faded off after few years. That is, we interact the treatment dummy with dummies capturing years from hospital closure. As we do not know the precise timing of hospital closure over the year, we use the year prior to closure as baseline. The specification looks as follows:

$$y_{ihjt} = \alpha + \sum_{t=s}^{-2} \pi_t T_{jt} + \sum_{t=0}^{k} \phi_t K_{jt} + \boldsymbol{X}\beta'_{ihjt} + \boldsymbol{H}\gamma'_{hjt} + \boldsymbol{M}\mu'_{jt}\, \delta_j + \lambda_t + \epsilon_{ihjt} \#(3)$$

Where $t = (s,\ldots,-2,-1,0,1,2,\ldots,k)$ represents time from home hospital closure, T_{jt} are interactions of the treatment indicator (which equals 1 if municipality m experienced a home hospital closure) and time dummies for all periods before time -1 (i.e. year prior to closure). Likewise, K_{jt} are interactions of the treatment indicator and time dummies for all periods from time 0 (i.e. year of closure) onwards. Lack of statistical significance on coefficients π_t provides indirect evidence in support of the parallel-trends assumption. We would expect either the long-term benefits of hospital closure in terms of improved efficiency to offset the short run adverse effects or no such benefits in the long run.

Exploration of potential mechanisms

Hospital closures may impact health outcomes in several ways. First, closure increases the time it takes to travel to another hospital. Given that AMI is a time sensitive condition, even a small increase in travel time can result in adverse health outcomes. Second, closures may induce congestion issues in non-closing hospitals if staffing and facilities are not fully adjusted to meet the increased case load from closing hospitals. We test the first two mechanisms. For this purpose, we combine the above staggered difference-in-differences approach with an instrumental variable (IV) estimation. Specifically, the difference-in difference estimator is used in the first stage as an instrument for the mechanisms of interest, i.e. travel time and beds utilization rate in CCUs as a proxy for congestion. In the second stage, health outcomes are regressed on predicted values of travel time and beds utilization rate obtained from the first stage. The two-stage system is specified as follows:

$$1st\ stage: W_{i(hj)t} = \psi_0 + \psi_1\left(Treat\ Post_{jt}\right) + \delta_j + \lambda_t + \epsilon_{ihjt} \#(5)$$

$$2nd\ stage: Y_{ijt} = \alpha + \theta^{IV}\widehat{W}_{i(hj)t} + \boldsymbol{X}\beta'_{ihjt} + \boldsymbol{H}\gamma'_{hjt} + \boldsymbol{M}\mu'_{jt} + \delta_j + \lambda_t + \epsilon_{ihjt}\ (6)$$

Where $W_{i(hj)t}$ is (i) the travel time from municipality of residence *j* to municipality of hospital *h*, (ii) the beds utilization rate in CCU of hospital *h* in year *t*, measured as the ratio between total AMI admissions in hospital *h* in year *t* and the number of CCU beds on January 1 of year *t.* The centroid-to-centroid travel time by car is computed for every municipality of residence and municipality of hospital combination for 2008-2015 using the R's *osrm* package. A detailed explanation on the computation of the travel time can be found in the a. Information on the number of beds in each CCU is provided by the Ministry of Health. Specifically, we have data on the number of beds in CCUs of all hospitals active in Italy between 2010 and 2015, on January 1 of each year. We combine data on beds in CCUs with hospital discharge data to construct yearly measures of beds utilization rate in CCUs of hospitals serving each municipality for years 2010-2015.

RESULTS

Table 2 reports the descriptive sample statistics for a set of variables, both overall and by treatment group (affected and not affected by home hospital closures), calculated over the entire period of analysis. In-hospital mortality and length of stay is lower for patients affected by home hospital closures compared to the control group, while 30-day, 90-day and 365-day readmission is higher for the treated group. It is worth stressing that these statistics are not informative about closure effects. Table S2 reports detailed descriptive statistics for control and treated groups in year prior to closure, in the year of closure and in the year after closure. It shows that in-hospital mortality and length of stay has increased in the year following closure. Regarding readmission rates, we see a decrease in 30-day, 90-day and 365-day readmission in the year following closure. However, it should be noted that readmission rates increase over time in the treated group (8.3% for 30-day readmission, 8.9% for 90-day readmission and 9.9% for 365-day readmission at t+3) . Regarding length of stay, there is no consistent increase or decrease in the years following the treatment. Overall, AMI patients in the treatment group are slightly younger, more likely to be male, married, educated and healthier (Table 2, Panel A). Treated municipalities are less populated in terms of overall population and also have relatively larger proportion of people exposed to economic deprivation (Table 2, Panel B).

[Table 1 about here]

Table 3 reports the results on the impact of home hospital closure on in-hospital mortality after an AMI. When accounting for municipality and year fixed effects only, we find home hospital closure to increase the probability of in-hospital death following an AMI by about 1.2% (Table 3, Column 1). This result is robust to the inclusion of individual, hospital and municipality characteristics (Table 3, Column 2). When including region-year fixed effects to account for yearly region-specific shocks, the effect decreases in magnitude to 0.7%, but remains statistically significant at conventional levels (p-value <0.05). Table 4 displays results on the impact of home hospital closures on the probability of cardiac/circulatory related and AMI compatible 30-day, 90-day and 365-day hospital readmission for an AMI index admission. We do not find any statistically significant effect of home hospital closure on the probability of any of the three hospital readmission variables. Table 5 displays results for the impact of home hospital closure on the length (days) of in-hospital stay following an AMI. We find that home hospital closure increases the length of in-hospital stay by about 0.4 and 0.3 days for the first two specifications respectively (Columns 1 and 2). For the last specification including region-year fixed effects, the effect size returns to 0.4 days (Table 4, Column 3)[20].

[Table 3, 4 and 5 about here]

As the validity of the DID estimates rests on the assumption of parallel trends, we formally test this by running a series of regressions including leads and lags of the treatment effect as specified in the methods section. Incidentally, this approach allows also to inspect the effect of home hospital closure on outcomes of interest over time. Results are displayed in Table 6 for all outcomes of interest. We find evidence in support of the parallel trends' assumption. Indeed, the coefficients for periods prior to the year before closure are non-significant (with the exception of 90-day readmission at t-2 and length of stay at t-4) (Figure 3). In addition, coefficients relative to post-closure periods reveal that the effect of hospital closure on in-hospital mortality and 30-day, 90-day and 365-day readmission is rather persistent. Regarding length of stay, the coefficients for each post-closure year are not statistically significant. In contrast with findings from prior work (Avdic, 2016), these results suggest that adaptation to an exogenous negative shock in the availability of public healthcare may take relatively long.

To examine the channels through which hospital closures may affect AMI patients' outcomes we adopt a 2-stage approach, as described in the Section 3. We present results from the first stage for both mechanisms of interest, i.e. travel time and utilization rate of beds in CCUs. Table 7 reports results for the effect of hospital closures on travel time. Hospital closure increases average travel time to hospitals by around 4 minutes on average, which is about 19% of pre-closure average travel time for treated municipalities (Column 1). To put this result into context, prior studies find that in the event of cardiac arrest, a common consequence of AMI, the brain may suffer irreversible damages after only five minutes (Pell et al., 2001). Column 2 displays the results when adding leads and lags to the baseline specification. We find the effect of hospital closure on average travel time to the hospital to be persistent over time, suggesting limited relocation of patients' following a hospital closure. This provides indirect evidence

[20] We obtain qualitatively analogous results when applying Poisson panel data models.

supporting the validity of the parallel trends' assumption. Table 8 displays estimates of the effect of hospital closures on utilization rate of beds in CCUs. On average, a hospital closure increases the yearly utilization rate by about 7 patients per CCU bed (Column 1). This effect is consistent with the parallel trends assumption and is persistent over time (Column 2).

Table 9 reports second-stage results for travel time. *Ceteris paribus,* an exogenous increase of one minute increase in travel time induced by a hospital closure increases the probability of in-hospital mortality by 0.6%, increases the probability of 30-day, 90-day and 365-day hospital readmission by 3% and increases the length of hospital stay by about 0.15 days ($p<0.1$ for in-hospital mortality and length of stay). However, the effect on hospital readmission is not statistically significant at conventional levels

In interpreting these results, some caveats apply. First, one should keep in mind that our data do not allow to compute the precise travel time to hospital from individuals' own residence. As explained in the Data section, we can only measure travel time for each municipality of residence – municipality of hospital combination. This limitation has two implications. First, for AMI patients who are admitted to a hospital located in their municipality of residence, travel time is set to be zero. Second, AMI patients residing in municipality *m* admitted to hospital *h,* located in a municipality other than *m,* face the same travel time, regardless of their specific home address. Both considerations raise concerns mostly for individuals residing in large municipalities. As a robustness check, we will repeat our analyses focusing on AMI patients' residing in small-medium sized municipalities only, presented in the robustness check section below.

Table 10 displays second-stage results for the CCU bed utilization rate. *Ceteris paribus*, an exogenous increase of 10 patients per CCU bed per year induced by hospital closure increases the probability of in-hospital mortality by 20%, increases the probability of 30-day, 90-day and 365-day hospital readmission by 10% and increases the length of hospital stay by about 0.2 days. However, the effect on readmission and length of stay is not statistically significant at conventional levels.

It is worth noting that higher congestion, i.e. lower availability of CCU beds, in non-closing hospitals may negatively affect AMI patients who are not directly affected by a home hospital closure based on their municipality of residence. This means that our estimated effect of hospital closures on patients' outcomes provides a lower bound estimate of the actual impact.

Robustness checks

We perform a series of robustness checks. First, anticipating hospital closure, individuals may relocate to municipalities nearby non-closing hospitals. If this were the case, the composition of population exposed to hospital closures would be endogenously affected by closures themselves. We check for the presence of anticipation effects by checking trends in AMI hospitalizations in treated and non-treated municipalities. As one can see from Table S3, we find essentially no evidence of diverging trends prior to hospital closure in terms of AMI hospitalizations and AMI hospitalization rate. In fact, we find no evidence of diverging trends after closure, too. This suggests that AMI patients from municipalities served by closing hospitals were absorbed into non-closing ones. If this were not the case, we would most likely observe a decline in AMI hospitalizations from treated municipalities, possibly associated to higher out-of-hospital AMI mortality compared to municipalities in the control group.

Second, we address the limitations implied by our measurement of travel time by either excluding large municipalities from the empirical analysis or replicating the results with subsamples of municipalities with varying population sizes. Table S4 presents the results for the effect of hospital closure on AMI patient outcomes when including only small-medium municipalities (population <50k). With the exception of 365-day readmission and length of stay which does not remain statistically significant, other outcomes are statistically significant. In particular, home hospital closures increase the probability of 30-day and 90-day readmission by 1.3% and 1.2% respectively and are statistically significant at conventional levels. Table S5 displays the parallel trends for the effect of closure on patient outcomes at individual level for different subsets of municipalities, with resident population below pre-defined thresholds (<100k, <50k, <30k)[21]. The estimates are consistent with those presented in Table 6.

Table S6 and S7 shows the estimated impact of hospital closure on travel time for different subsets of municipalities, with resident population below pre-defined thresholds at individual and municipality level respectively. Estimates are qualitatively and quantitatively analogous to the main results displayed in Table 7, and provide evidence supporting the parallel trends assumption. Table S8 shows the estimated impact of hospital closure on utilization rate of CCU beds in hospitals receiving AMI patients, restricting the sample to AMI patients residing in municipalities with population below the pre-defined thresholds. The estimates are consistent with the main results reported in Table 8, and support the parallel trends assumption.

Tables S9 and S10 display second-stage results for testing the mechanisms of interest, travel time and utilization rate of CCU beds respectively, restricting the sample to resident populations below pre-defined thresholds at the individual level. As one can see, the effect of an increase in travel time induced by hospital closure on AMI patients' outcome is imprecisely estimated and does not retain statistical

[21] Table S5 does not report results for 90-day and 365-day hospital readmission. The results for 90-day and 365-day are similar to that of 30-day readmission and are available upon request from authors.

significance across any specifications.[22] On the contrary, the effects of an increase in utilization rate of CCU beds are comparatively more consistent with the main results presented in Table 10.The effects remain statistically significant only for in-hospital mortality among patients residing in municipalities below 100k (Table S10, Column 1).

Third, we control for measures of local economic deprivation to check the robustness of our main results. Table S11 and S12 presents the results for the effect of hospital closure on AMI patient outcomes at the individual and municipality level respectively when controlling for the share of resident population living below the poverty threshold. At the individual level, except for length of stay which is statistically significant only at 10% (Column 3 in Table S11), other estimates are consistent with the main results. At the municipality level, only in-hospital mortality retains statistical significance which is similar to the main results presented in Table S13. Effects are similar when controlling for average tax-payer income instead of share of resident population living below the poverty threshold to the baseline regression for the effect of hospital closure on patient outcomes (results available upon request to the authors).

Finally, we repeat our main empirical exercises by using municipality as unit of analysis. The purpose of this check is twofold. First, we want to make sure that results hold when outcomes and treatment are defined at the same level (i.e., the municipality). Second, we want to test whether results for in-hospital mortality and hospital readmission are robust to alternative non-linear specifications. While the estimation of high dimensional fixed effect non-linear models at the patient level poses computational challenges, this is not the case at the municipality level. Table S13 reports the main results for patients' outcomes measured at the municipality level, namely annual number of in-hospital deaths, annual number of 30-day, 90-day and 365-day hospital readmissions, and yearly average length of stay. All models include municipality and year fixed effects, and control for resident population and share of resident population aged 65+. The estimates for in-hospital death and hospital readmissions were obtained through Poisson models using the log of yearly AMI hospitalizations at the municipality level as the exposure variable. The estimate for average length of stay is obtained through linear regression models. Robust standard errors are clustered at the municipality level. In line with individual-level specifications, we find that the municipalities affected by home hospital closure recorded, on average, 26% ($ore^{0.235}$) more in-hospital deaths and 0.3 days longer hospital stay after closure compared to municipalities in the control group. Similar to individual level results, the effect of home hospital closure on readmission rates are not consistent and do not remain statistically significant at conventional levels. Table S14 displays results of

[22] Table S9 and S10 does not report results for 90-day and 365-day hospital readmission. The results for 90-day and 365-day are similar to that of 30-day readmission and are available upon request from authors.

regression analyses including treatment leads and lags. In this case too, we find evidence in support of the parallel trends' assumption and results are qualitatively similar to the individual level results. Table S15 displays first-stage for results average travel time. At the municipality level, the coefficient reduces to 2.6 minutes but remains statistically significant (Table S15, Column 1). Municipality-level analysis confirms that the effect of hospital closure on average travel time to hospital to be persistent over time (Column 2 of Table S15). Finally, Table S16 reports second-stage results for travel time with AMI patients' outcomes measured at the municipality level. *Ceteris paribus,* an exogenous increase of 1 minute in average travel time induced by a hospital closure increases the number of in-hospital deaths by 9% (=$or\ e^{0.090}$) and increases the mean length of hospital stay by about 0.7 days. Similar to individual level results, hospital closure induced travel time does not have a statistically significant effect on hospital readmission.

Recently, there is new evidence noting that traditional two-way fixed effects may lead to biased estimates (Goodman-Bacon, 2021; Callaway and Sant'Anna, 2020). ATT estimates for a difference-in-difference model with multiple time-periods with heterogeneity in treatment timing are biased because such estimates are a weighted average of comparisons of treated units with units treated in the past, units that are yet to be treated and never treated units. While, the never treated units provide a perfect control group, and similarly comparison with units that are yet to be treated are not bad controls either especially as they help to test the parallel trend assumption. However, comparison of treated units with units treated in the past violates the parallel trend assumption contributing to a biased ATT. To check if our estimates are robust to this source of bias, we use the methodology developed by Callaway and Saint'Anna (2020) to compare treated municipalities with municipalities that are not yet treated and never treated (using the csdid package available in Stata). Using this methodology, we can generate group-time average treatment effect estimates for each set of municipalities experiencing a home hospital closure in the same year. These estimates are then aggregated to produce an unbiased estimate of ATT. Results are presented in Table S17. Our results remain the same for in-hospital mortality and readmission rates but not for length of stay. Results show that compared to either not-yet treated municipalities and never treated municipalities, hospitalizations from treated municipality's leads to a decrease in length of stay.

The divergence from the main results could be due to several potential reasons (1) the existing findings for the effect of hospital closures on length of stay is mixed. Using a methodology similar to ours, Gujral and Basu (2019) reports an increase in length of stay, consistent with our main findings. On the other hands, studies analyzing the same research question but employing different methodologies (matching and Instrument variable analysis) find that hospital closures may lead to a 'speed up' behavior in remaining hospitals via reducing the length of stay. This points to a potential explanation that results for length of stay attained through a two-way fixed effects difference-in-difference model not taking into account the sources of bias explained above may lead to unreliable estimates (in our case, for the 'length of stay' variable) and (2) We could not carry out this analysis at the municipality level of analysis due to missing observations in the matrix for calculating the aggregated estimates. This left us with the only option of running the analysis at hospitalization level. At the hospitalization level, we could not specify the cross-sectional identifier given multiple observations of municipalities in a given year. Therefore, our specification is not ideal and could be another reason that could have potentially led to a flip in the sign of the main coefficient.

DISCUSSION

In this article, we study the causal impact of hospital closures on AMI patient outcomes in Italy during a period characterized by high levels of decentralization and cost containment measures. Using a staggered difference-in-difference approach, we compare health outcomes among AMI patients from municipalities that were exposed to home hospital closures with those that were not. With a sample of over 819,412 AMI cases and 43 home hospital closures identified between 2008 and 2015, our results show a significantly increased probability of experiencing in-hospital mortality following a home hospital closure at the individual level. Specifically, AMI patients affected by home hospital closures display a 0.7 to 1.2% increase in the likelihood of dying in the hospital. Results also show a increase in cardiac or circulatory related and AMI compatible 30-day, 90-day and 360-day readmission and an increase in length of stay following a home hospital closure. Overall, our empirical results reveal adverse effects of hospital closure on patient outcomes.

Our findings on the adverse effects of hospital closures on in-hospital mortality are broadly consistent with those reported by previous studies. Using emergency department (ED) closures in California between 1999-2010, Liu et al. (2014) detected 5% higher odds of inpatient mortality for admissions occurring near a closure compared to farther away among the general patient population. However, this estimate increased to 15% when only patients with an AMI were considered, indicating a higher vulnerability of patients with time sensitive conditions. Similar studies investigating the effect of hospital or ED closures on mortality outcomes found effects ranging between 3.0 and 6.2% for AMI admissions (Carroll, 2019; Gujral and Basu, 2019; Shen and Hsia, 2011). Our results can also be discussed in light of studies analyzing the effect of a closure induced changes in geographical distance on access to care. Buchmueller et al. (2006) finds an increase in one mile driving distance to the nearest hospital following closures to result in a 6.5% increase in deaths from heart attacks. Similarly, Avdic (2016) exploits an exogenous change in minimum hospital distance resulting from closures of emergency hospitals in

Sweden to detect a decrease of 2% points per 10 km increase in distance) in surviving an AMI one year after closure. Compared to the existing studies, our estimates for in-hospital mortality is smaller and this might be due to reasons including unavailability of out-of-hospital mortality or differences in the selection of primary diagnosis included in the study. Our findings are also consistent with existing literature on increase in travel time as a potential mechanism through which closures might affect patient outcomes (Avdic, 2016).

However, our findings differ from a large collection of studies that finds no significant effect of hospital closures or closure induced changes in distance or travel time on patient outcomes (Hsia et al., 2012; Joynt et al., 2015; Rosenbach and Dayhoff, 1995). There are several potential explanations for these non-findings. One, most studies limit the dependent variable to mortality measures. However, the negative effect of closure might not be as severe as death and could be limited to smaller effects which could be captured through readmission or morbidity measures (Hsia et al., 2012). Two, changes in travel time post closure could be very small and may not be enough to detect a significant change in patient outcomes (Hsia et al., 2012). Three, results may vary depending on the empirical strategy used. For instance, Gujral and Basu (2019) employ a methodology that accounts for all potential mechanisms including improved efficiency and quality in the nearby hospitals which may offset some of the adverse effects of closures, thereby leading to a null result. Five, results may also vary depending on the health system across countries. It could be that in certain geographical regions, lower quality hospitals were closed, resulting in an increase in increased average quality off-setting the adverse effects of the closures.

Our study does not find statistically significant effects of home hospital closure on hospital readmission for the main (reduced form) specification. However, when including leads and lags to the main specification and when considering only small municipalities, we find that home hospital increases the probability of 30-day, 90-day and 365-day readmission for AMI admissions. These findings on readmission is consistent with our hypothesis that hospital closures may induce a lack in quality of care delivered in the nearby hospitals (due to factors such as congestion), thereby increasing the readmission rates. Prior evidence on the role of hospital closure or mergers on hospital readmission is mixed. For instance, Ho and Hamilton (2000) uses a difference-in-difference methodology to estimate the effects of hospital consolidation in California between 1992 and 1995 to find a positive effect on 90-day readmission following a heart attack. In contrast, several studies have also found statistically non-significant effects of hospital closures on readmission (Beaulieu et al., 2020; Joynt et al., 2015; Song and Saghafian, 2019). Literature on hospital closures and patient outcomes highlight heterogeneity in the characteristics of closed hospitals (and remaining hospitals) as the main reason for studies finding mixed results. This may include the quality, size, type (teaching, non-profit etc.), competitive capabilities of the hospital in the market and location (rural versus urban) (Joynt et al., 2015; Song and Saghafian, 2019). With regards to readmission rates, another reason for null findings in some studies could be also be due to simultaneous external policies targeted specifically to reduce readmission rates, thereby nullifying the effects of hospital closures on readmission rates (Wasfy et al., 2017).

Our findings are also in line with several studies that found hospital closures to have a positive effect on length of stay (LOS) (Gujral and Basu, 2019). Higher LOS indicates overcrowding which could also be a reason for a decrease in 30 day readmission following a closure. Also, a higher LOS may indicate that the AMI admissions were of higher severity (and therefore increased complications) following a delay in time to treatment due to a closure. Studies have also found longer LOS to be positively related to hospital readmissions, particularly among heart failure patients (Rachoin et al., 2020; Reynolds et al., 2015). This implies that the patients with longer LOS are more complicated patients and given this, they are also more likely to be at the risk of readmission. However, studies also find an inverse relationship between longer length of stay and readmission following an AMI (Eapen et al., 2013). For instance, a higher LOS may indicate a speed-up behavior to manage the new inflow of patients which might have consequences for patient such as an increase in hospital readmission. Further research is needed to examine how the trade-off between length of stay and readmission affects patient welfare when exposed to shocks in public healthcare supply.

Our study has some limitations. First, our measure of mortality fails to account for out-of-hospital mortality, for example deaths occurring on the way to hospital or right after discharge. This may reduce the observed mortality rate in the treated group. Therefore, our results should be considered as a conservative measure of the effect of hospital closure on mortality outcomes. Second, municipalities as a geographical unit may not be a “perfect surrogate” of an area around a hospital that is affected by its closure (Liu et al., 2014). Nearby municipalities relying on the same hospital, and therefore likely to be affected, might be excluded from the treatment group. Similarly, unaffected admissions relying on hospitals in another nearby municipality might be falsely included in the treatment group. In this sense, our results reflect a lower bound for the increased odds of in-hospital mortality for AMI patients. Three, closures are defined on the basis of a drop in only AMI admissions, therefore our results reflect the consequences of AMI ward closures and not necessarily complete hospital closures. Also, another reason for smaller effects on patient outcomes when compared to existing findings could be due to the fact that other wards in the hospital may mitigate some of the adverse effects of AMI ward closure. Fourth, as mentioned, our measure of travel time does not capture the precise distance between AMI patients’ residence and hospital of admission. As such, it does not allow to fully exploit heterogeneity in travel time among individuals residing in the same municipality. Likewise, bed utilization rate does not capture the occupancy rate faced by AMI patients at the time of admission, but rather hospital bed turnover rate on a yearly basis. For this reason, it is an imperfect proxy of congestion.

Given the existing findings on the disproportionate impact of ED closures on health outcomes for vulnerable populations in the US (Liu et al., 2014), future research could examine if the hospitals or AMI wards that closed in Italy were those with "greater experience and expertise" in caring for socially disadvantaged populations, and if the adverse effects of such closures further aggravated existing disparities in health outcomes. Further, another underlying mechanism explaining increased mortality after hospital closures, namely the inability to shift patients admitted to an ED to inpatient beds in a timely manner could be explored. This phenomenon known as 'boarding' in an ED has been shown to result in worse clinical outcomes (Rabin et al., 2012) and may arise due to an overall reduction in beds arising from closures of entire hospitals, as opposed to not just AMI wards. Such an analysis could be a valuable contribution in the context of Italy but requires a clear distinction between entire hospital closures versus AMI ward closure. Finally, given the policy-induced wave of maternity ward closures in Italy since the early 2010s[23] and the time sensitive nature of delivery complications, it would be an interesting research future ally to repeat our analysis for birth outcomes.

CONCLUSION

Using data from Italy, our study shows that hospital closure may severely affect patient outcomes, and that such effects can be persistent over time. While hospital closures could be motivated by financial and safety reasons, effective policies to minimize human and social costs by ensuring continued quality of care are needed. One solution that is widely discussed in the literature is selective closures of hospitals and bailouts (Song and Saghafian, 2019). Policy makers need to think carefully about which hospitals to bail out in case of risk of closures. Song and Saghafian (2019) shows that bailing out hospitals in markets with fewer patient choices or those with relatively less desirable characteristics compared to neighboring hospitals may minimize the adverse effect of hospital closures. Regulations and monitoring efforts to reduce emergency department overcrowding is also a potential solution to reduce the adverse impact of closures. This could include simple strategies such as 'moving boarders to inpatient halls' or 'distributing procedures evenly over the week' to optimize bed management (Rabin et al., 2012). Adequate availability of acute and emergency care capacity and minimizing the risk of overcrowding is particularly important in the light of the COVID-19 pandemic. Italy, being on the higher end of the spectrum among the European countries that implemented sizeable reductions of acute care beds and hospitals since the 90s,

[23] See the "Guidelines for the promotion and improvement of quality, safety and appropriateness of care interventions in the birth path" approved by Agreement at the Joint Conference (State-Regions) of December 16, 2010.

faced significant challenges in tackling the consequences of COVID-19. This has also affected the provision of non-COVID related emergency care. For example, De Rosa et al. (2020) showed a significant reduction in AMI admissions in conjunction with an increase in fatality and complication rates (given a delay in time to treatment) during COVID-19 pandemic in Italy. Therefore, policy discussion should focus on the need for increased availability of hospital capacity and to move away from the implementation of fiscal constraints that do not provide adequate protection to population health.

BIBLIOGRAPHY

Aiken L, Buchanan J, Edwards N, Freemantle N, Harding A, Harison A, Healy J, Hensher M, Jakab M, Langenbrunner J. Hospitals in a Changing Europe. 2001.

Aimone Gigio L, Alampi D, Camussi SA, Ciaccio G, Guaitini P, Lozzi M, Mancini AL, Panicara E, Paolicelli M. La Sanità in Italia: Il Difficile Equilibrio Tra Vincoli Di Bilancio E Qualità Dei Servizi Nelle Regioni in Piano Di Rientro (The Italian Healthcare System: The Difficult Balance between Budget Sustainability and the Quality of Services in Regions Subject to a Financial Recovery Plan). Bank of Italy Occasional Paper 2018.

Antman EM. Time is muscle: translation into practice. Journal of the American College of Cardiology 2008;52; 1216-1221.

Arcà E, Principe F, Van Doorslaer E. Death by austerity? The impact of cost containment on avoidable mortality in Italy. Health Economics 2020;29; 1500-1516.

Avdic D. Improving efficiency or impairing access? Health care consolidation and quality of care: Evidence from emergency hospital closures in Sweden. Journal of Health Economics 2016;48; 44-60.

Avdic D, Lundborg P, Vikström J. 2014. Learning-by-doing in a highly skilled profession when stakes are high: Evidence from advanced cancer surgery. (Ed)^(Eds). Working Paper; 2014.

Avdic D, Lundborg P, Vikström J. Mergers and birth outcomes: evidence from maternity ward closures. 2018.

Beaulieu ND, Dafny LS, Landon BE, Dalton JB, Kuye I, McWilliams JM. Changes in quality of care after hospital mergers and acquisitions. New England Journal of Medicine 2020;382; 51-59.

Bentham G. Proximity to hospital and mortality from motor vehicle traffic accidents. Social science & medicine 1986;23; 1021-1026.

Brenna E. Quasi-market and cost-containment in Beveridge systems: the Lombardy model of Italy. Health Policy 2011;103; 209-218.

Buchmueller TC, Jacobson M, Wold C. How far to the hospital?: The effect of hospital closures on access to care. Journal of health economics 2006;25; 740-761.

Burkey ML, Bhadury J, Eiselt HA, Toyoglu H. The impact of hospital closures on geographical access: Evidence from four southeastern states of the United States. Operations Research Perspectives 2017;4; 56-66.

Callaway, Brantly, and Pedro HC Sant'Anna. "Difference-in-differences with multiple time periods." Journal of Econometrics (2020).

Capps CS. The quality effects of hospital mergerss. US Department of Justice, Antitrust Division; 2005.

Carroll C. Impeding access or promoting efficiency? effects of rural hospital closure on the cost and quality of care. NBER Working Paper 2019.

De Rosa S, Spaccarotella C, Basso C, Calabrò MP, Curcio A, Filardi PP, Mancone M, Mercuro G, Muscoli S, Nodari S. Reduction of hospitalizations for myocardial infarction in Italy in the COVID-19 era. European heart journal 2020;41; 2083-2088.

Depalo D. The side effects on health of a recovery plan in Italy: A nonparametric bounding approach. Regional Science and Urban Economics 2019;78; 103466.

Dranove D, Lindrooth R. Hospital consolidation and costs: another look at the evidence. Journal of health economics 2003;22; 983-997.

Eapen ZJ, Reed SD, Li Y, Kociol RD, Armstrong PW, Starling RC, McMurray JJ, Massie BM, Swedberg K, Ezekowitz JA. Do countries or hospitals with longer hospital stays for acute heart failure have lower readmission rates? Findings from ASCEND-HF. Circulation: Heart Failure 2013;6; 727-732.

Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity measures for use with administrative data. Medical care 1998; 8-27.

Falco RD. Access to Healthcare and the Global Financial Crisis in Italy: A Human Rights Perspective. e-cadernos CES 2019.

Gaynor M, Laudicella M, Propper C. Can governments do it better? Merger mania and hospital outcomes in the English NHS. Journal of health economics 2012;31; 528-543.

Goodman-Bacon, Andrew. "Difference-in-differences with variation in treatment timing." Journal of Econometrics (2021).

Grytten J, Monkerud L, Skau I, Sørensen R. Regionalization and local hospital closure in Norwegian maternity care—the effect on neonatal and infant mortality. Health services research 2014;49; 1184-1204.

Gujral K, Basu A. 2019. Impact of rural and urban hospital closures on inpatient mortality. (Ed)^(Eds). National Bureau of Economic Research; 2019.

Heckman J, Pinto R, Savelyev P. Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. American Economic Review 2013;103; 2052-2086.

Ho V, Hamilton BH. Hospital mergers and acquisitions: does market consolidation harm patients? Journal of Health Economics 2000;19; 767-791.

Hsia RY, Kanzaria HK, Srebotnjak T, Maselli J, McCulloch C, Auerbach AD. Is emergency department closure resulting in increased distance to the nearest emergency department associated with increased inpatient mortality? Annals of emergency medicine 2012;60; 707-715. e704.

Joynt KE, Chatterjee P, Orav EJ, Jha AK. Hospital closures had no measurable impact on local hospitalization rates or mortality rates, 2003–11. Health affairs 2015;34; 765-772.
Karanikolos M, Mladovsky P, Cylus J, Thomson S, Basu S, Stuckler D, Mackenbach JP, McKee M. Financial crisis, austerity, and health in Europe. The Lancet 2013;381; 1323-1331.

Kaufman BG, Thomas SR, Randolph RK, Perry JR, Thompson KW, Holmes GM, Pink GH. The rising rate of rural hospital closures. The Journal of Rural Health 2016;32; 35-43.

Lindrooth RC, Sasso ATL, Bazzoli GJ. The effect of urban hospital closure on markets. Journal of health economics 2003;22; 691-712.

Liu C, Srebotnjak T, Hsia RY. California emergency department closures are associated with increased inpatient mortality at nearby hospitals. Health Affairs 2014;33; 1323-1329.

Mauro M, Maresso A, Guglielmo A. Health decentralization at a dead-end: towards new recovery plans for Italian hospitals. Health Policy 2017;121; 582-587.

OECD. 2018. Health Care Resources: Hospitals. (Ed)^(Eds). 2018.

Pell JP, Sirel JM, Marsden AK, Ford I, Cobbe SM. Effect of reducing ambulance response times on deaths from out of hospital cardiac arrest: cohort study. Bmj 2001;322; 1385-1388.

Quaglio G, Karapiperis T, Van Woensel L, Arnold E, McDaid D. Austerity and health in Europe. Health policy 2013;113; 13-19.

Rabin E, Kocher K, McClelland M, Pines J, Hwang U, Rathlev N, Asplin B, Trueger NS, Weber E. Solutions to emergency department 'boarding'and crowding are underused and may need to be legislated. Health Affairs 2012;31; 1757-1766.

Rachoin J-S, Aplin KS, Gandhi S, Kupersmith E, Cerceo E. Impact of Length of Stay on Readmission in Hospitalized Patients. Cureus 2020;12.

Reinhard E, Layte R, McCrory C, Panico L, Avendano M. The great recession and the health of young children: a fixed-effects analysis in Ireland. American journal of epidemiology 2018;187; 1438-1448.

Reynolds K, Butler MG, Kimes TM, Rosales AG, Chan W, Nichols GA. Relation of acute heart failure hospital length of stay to subsequent readmission and all-cause mortality. The American journal of cardiology 2015;116; 400-405.

Rosenbach ML, Dayhoff DA. Access to care in rural America: impact of hospital closures. Health care financing review 1995;17; 15.

Ryan TJ, Anderson JL, Antman EM, Braniff BA, Brooks NH, Califf RM, Hillis LD, Hiratzka LF, Rapaport E. ACC/AHA guidelines for the management of patients with acute myocardial infarction: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Committee on Management of Acute Myocardial Infarction). Journal of the American College of Cardiology 1996;28; 1328-1419.

Schwierz C. 2016. Cost-containment policies in hospital expenditure in the European Union. (Ed)^(Eds). Directorate General Economic and Financial Affairs (DG ECFIN), European …; 2016.

Shen Y-C, Hsia RY. Association between ambulance diversion and survival among patients with acute myocardial infarction. Jama 2011;305; 2440-2447.

Song L, Saghafian S. Do Hospital Closures Improve the Efficiency and Quality of Other Hospitals?s. Harvard Kennedy School, John F. Kennedy School of Government; 2019.

Stevenson B, Wolfers J. 2003. Bargaining in the shadow of the law: Divorce laws and family distress. (Ed)^(Eds). National Bureau of Economic Research; 2003.

Stuckler D, Reeves A, Loopstra R, Karanikolos M, McKee M. Austerity and health: the impact in the UK and Europe. European journal of public health 2017;27; 18-21.

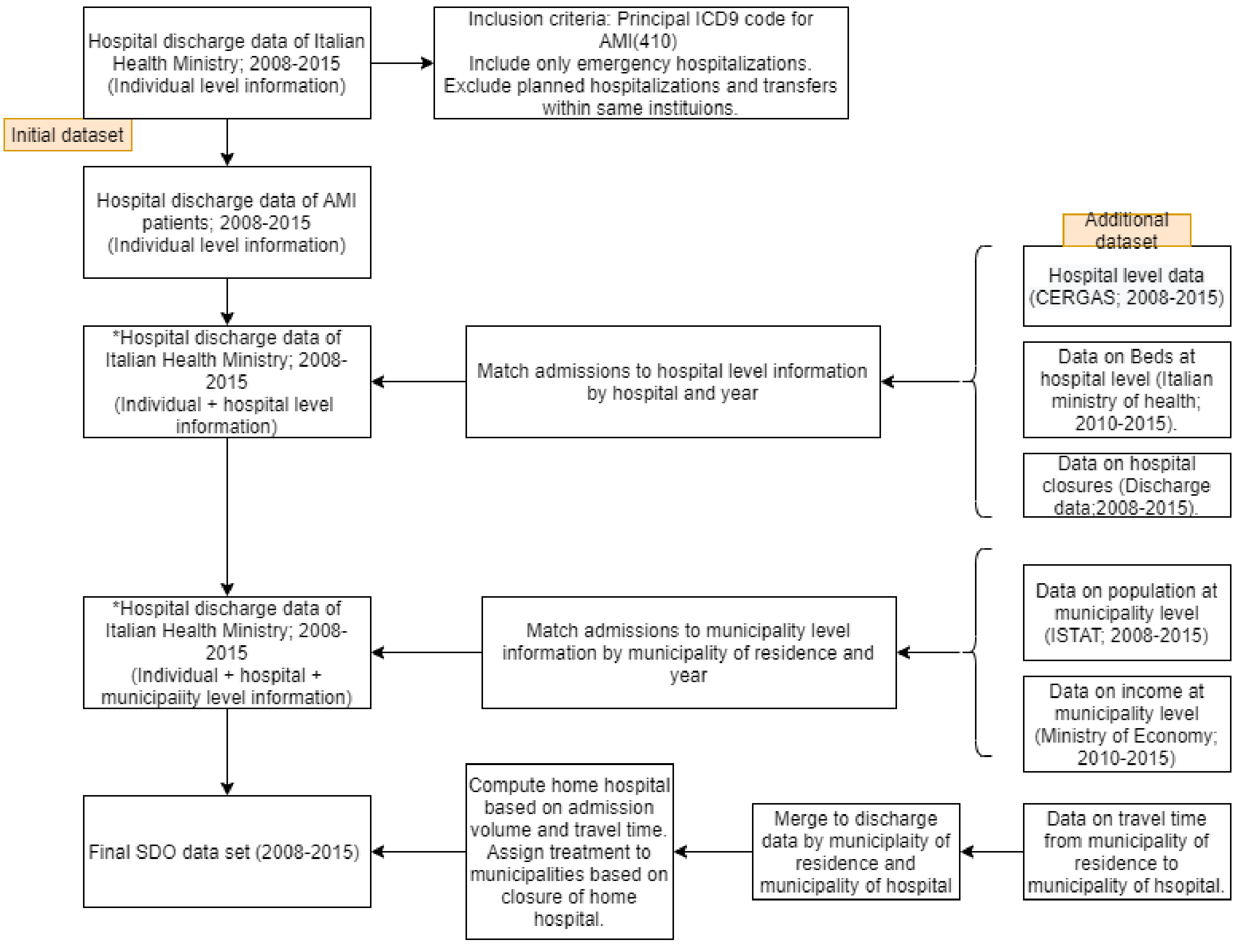
Troske S, Davis A. Do Hospital Closures Affect Patient Time in an Ambulance? 2019.

Wasfy JH, Zigler CM, Choirat C, Wang Y, Dominici F, Yeh RW. Readmission rates after passage of the hospital readmissions reduction program: a pre–post analysis. Annals of internal medicine 2017;166; 324-331.

Wenzl M, Naci H, Mossialos E. Health policy in times of austerity—A conceptual framework for evaluating effects of policy on efficiency and equity illustrated with examples from Europe since 2008. Health Policy 2017;121; 947-954.

FIGURES

Figure 1: Data flow used to create primary analytic files



Note: When dataset includes information on beds at hospital level and income at municipality level; final dataset includes a time period of 2010-2015. These variables are used in robustness checks or as potential mechanisms. Hence, the main results cover a time period of 2008-2015.

Figure 2: Spatial distribution of municipalities affected by home hospital closure: 2008-2015

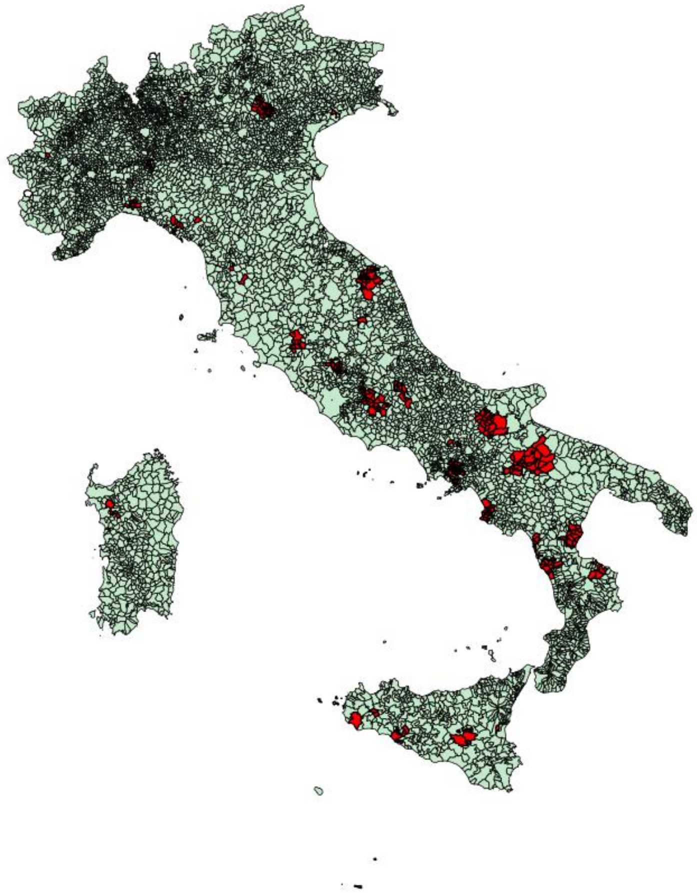


Figure 3: Parallel trends for AMI patient outcomes

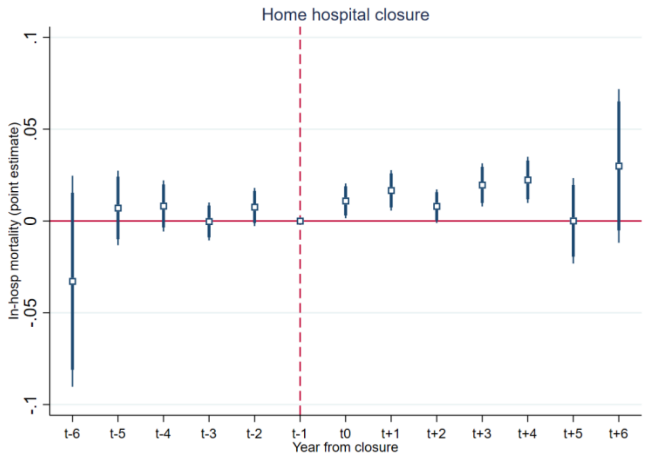


Fig 3a: Effect of hospital closure on in-hospital mortality (Linear model).

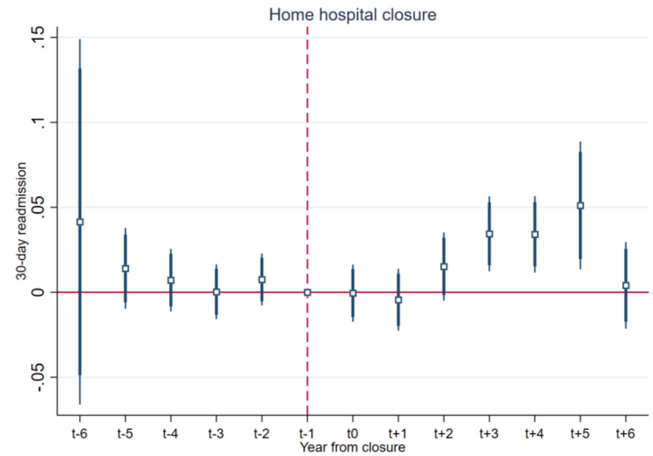


Fig 3b: Effect of hospital closure on 30-day readmission (Linear model).

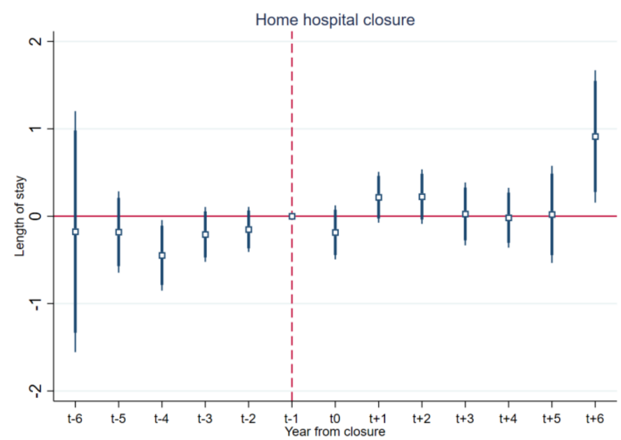


Fig 3c: Effect of hospital closure on length of stay (Linear model).

Note: Point estimates from staggered difference-in-difference models (LPM) with 90% (thicker lines) and 95% (thinner lines) confidence intervals. Controls include education, gender, age, Charlson comorbidity index, marital status, type of hospital, AMI admission volume by hospital, size of municipality population and municipality population aged 65+ (%).

TABLES

Table 1: List of closed hospitals

Hospital ID	Municipality of hospital	Year of closure	Home hospital
1	CATANIA	2008	YES
2	BRONI	2008	YES
3	ROMA	2008	YES
4	GUBBIO	2008	NO
5	CLUSONE	2008	YES
6	EMPOLI	2008	YES
7	BISCEGLIE	2009	NO
8	PAGANI	2009	NO
9	TODI	2010	NO
10	PESCINA	2010	YES
11	TORINO	2010	YES
12	NAPOLI	2010	NO
13	POLLENA TROCCHIA	2010	YES
14	ROMA	2010	NO
15	PIAZZA ARMERINA	2010	YES
16	ITTIRI	2010	YES
17	CERRETO SANNITA	2010	YES
18	NOLA	2011	YES
19	ACQUAPENDENTE	2011	YES
20	MAGLIANO SABINA	2011	YES
21	VENOSA	2011	YES
22	TREBISACCE	2011	YES
23	SCAFATI	2011	YES
24	SAN FELICE A CANCELLO	2011	YES
25	NAPOLI	2011	NO
26	MAZARA DEL VALLO	2012	YES
27	RIBERA	2012	YES
28	CETRARO	2012	YES
29	CARIATI	2012	YES
30	SCHIO	2012	YES
31	THIENE	2012	YES
32	GENOVA	2012	YES
33	ANAGNI	2012	YES
34	RECANATI	2012	YES
35	AVIGLIANA	2013	YES
36	EMPOLI	2013	YES
37	JESI	2013	YES
38	SUBIACO	2013	YES
39	LUCERA	2013	YES
40	BARI	2013	NO
41	CANOSA DI PUGLIA	2013	YES
42	AGROPOLI	2013	YES
43	SAN DONÀ DI PIAVE	2014	YES
44	LA SPEZIA	2014	YES
45	SAN SEVERINO MARCHE	2014	YES
46	CASTEL SAN PIETRO TERME	2014	NO
47	COPPARO	2014	NO

Table 2: Descriptive Statistics

VARIABLES	Total	Group averages		
		Untreated	Treated	T test or $\chi 2$ test
Panel A: individual-level				
Main outcomes				
In-hospital death	6.5	6.5	4.3	<0.001
30-day readmission	4.6	4.6	7.5	<0.001
90-day readmission	5.2	5.2	8.0	<0.001
365-day readmission	6.2	6.2	9.2	<0.001
Length of stay	7.26 (5.29)	7.27 (5.30)	6.72 (4.94)	<0.001
Health characteristics				
Maximum comorbidities				<0.001
0	0.1	0.1	0.3	
1	0.1	0.1	0.4	
2 (2 or more)	99.8	99.8	99.3	
Socio-demographic characteristics				
Age	70.16 (13.46)	70.19 (13.46)	68.12 (13.36)	<0.001
Female	35.0	35.0	32.2	<0.001
Married	45.8	45.7	50.6	<0.001
Level of education				<0.001
Elementary/No qualification	76.5	76.6	64.1	
Junior high school diploma	14.0	13.9	23.3	
High school diploma	7.6	7.5	10.3	
University degree/Short degree	0.4	0.4	0.7	
Graduate degree	1.5	1.5	1.7	
Channels				
Travel time	14.70 (17.97)	14.55 (17.93)	24.04 (17.92)	<0.001
Utilization rate beds CCU*	44.97 (22.51)	44.98 (22.43)	44.45 (26.09)	0.011
Observations (2008-2015)	818,835	806,219	12,616	
Observations (2010-2015)	618,431	606,003	12,428	
Panel B: municipality-level				
Municipality characteristics				
Resident population (1,000)	8.57 (44.50)	8.63 (44.96)	6.10 (8.18)	0.054
% of resident population aged 65+	22.10 (5.33)	22.09 (5.32)	22.20 (5.62)	0.497
Income per capita (1,000 euro)*	17.33 (3.87)	17.38 (3.88)	15.39 (3.11)	<0.001
% of population under poverty*	34.3 (11.0)	34.1 (10.9)	40.5 (10.7)	<0.001
Observations (2008-2015)	41,514	40,364	1,150	
Observations (2010-2015)	55,350	54,173	1,177	

Note: Panel A reports proportions or means (standard deviations) for variables at the individual (AMI patient) level. Panel B reports means (standard deviations) for variables at the municipality of residence level. In both panels, descriptive statistics are reported for the treatment and control group separately. *Data available for years 2010-2015 only.

Table 3: Effect of home hospital closure on in-hospital death

VARIABLES	(1) Dead	(2) Dead	(3) Dead
Post x Treated	0.012***	0.011***	0.007**
	(0.003)	(0.003)	(0.003)
Constant	0.072***	-0.171***	-0.133***
	(0.001)	(0.021)	(0.020)
Observations	818,835	818,459	818,459
Number of municipalities	7,992	7,992	7,992
Patient/hospital/municipality characteristics	No	Yes	Yes
Municipality FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Region x Year FE	No	No	Yes
Model	LPM	LPM	LPM

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Dead is a dummy measuring whether hospitalization ends with death. Controls include education, gender, age, Charlson comorbidity index, marital status, type of hospital, AMI admission volume by hospital, size of municipality population and municipality population aged 65+ (%).

Table 4: Effect of home hospital closure on readmission

VARIABLES	(1) 30-day readmission	(2) 30-day readmission	(3) 30-day readmission	(4) 90-day readmission	(5) 90-day readmission	(6) 90-day readmission	(7) 365-day readmission	(8) 365-day readmission	(9) 365-day readmission
Post x Treated	0.001	0.009	-0.003	-0.001	0.007	-0.005	-0.001	0.007	-0.005
	(0.007)	(0.006)	(0.006)	(0.007)	(0.007)	(0.006)	(0.007)	(0.007)	(0.007)
Constant	0.044***	0.156***	0.145***	0.051***	0.142***	0.129***	0.062***	0.118***	0.108***
	(0.001)	(0.028)	(0.028)	(0.001)	(0.027)	(0.028)	(0.001)	(0.028)	(0.029)
Observations	752,045	751,699	751,699	742,728	742,383	742,383	725,142	724,806	724,806
Number of municipalities	7,991	7,991	7,991	7,991	7,991	7,991	7,991	7,991	7,991
Patient/hospital/municipality characteristics	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Municipality FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region x Year FE	No	No	Yes	No	No	Yes	No	No	Yes

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. 30-day, 90-day and 365-day readmission is a dummy measuring whether an AMI hospitalization ends a cardiac/circulatory related or AMI compatible admission within 30, 90 and 365 days respectively. Controls include education, gender, age, Charlson comorbidity index, marital status, type of hospital, AMI admission volume by hospital, size of municipality population and municipality population aged 65+ (%).

Table 5: Effect of home hospital closure on length of stay

VARIABLES	(1) Length of stay	(2) Length of stay	(3) Length of stay
Post x Treated	0.349***	0.271**	0.418***
	(0.119)	(0.123)	(0.104)
Constant	7.383***	0.664	1.378***
	(0.024)	(0.599)	(0.530)
Observations	818,835	818,459	818,459
Number of municipalities	7,992	7,992	7,992
Patient/hospital/municipality characteristics	No	Yes	Yes
Municipality FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Region x Year FE	No	No	Yes

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Controls include education, gender, age, Charlson comorbidity index, marital status, type of hospital, AMI admission volume by hospital, size of municipality population and municipality population aged 65+ (%).

Table 6: Effect of home hospital closure on AMI patient outcomes – Individual level parallel trends

VARIABLES	(1) Dead	(2) 30-day readmission	(3) 90-day readmission	(4) 365-day readmission	(5) Length of Stay
Treated x t-6	-0.033	0.041	0.054	0.055	-0.177
	(0.029)	(0.055)	(0.045)	(0.046)	(0.704)
Treated x t-5	0.007	0.014	0.010	0.000	-0.181
	(0.010)	(0.012)	(0.013)	(0.014)	(0.238)
Treated x t-4	0.008	0.007	0.012	0.008	-0.448**
	(0.007)	(0.009)	(0.010)	(0.010)	(0.206)
Treated x t-3	-0.000	0.000	0.006	0.000	-0.209
	(0.005)	(0.008)	(0.009)	(0.009)	(0.160)
Treated x t-2	0.008	0.008	0.014*	0.013	-0.152
	(0.005)	(0.008)	(0.008)	(0.009)	(0.132)
Treated x t=0	0.011**	-0.000	0.002	-0.005	-0.185
	(0.005)	(0.009)	(0.009)	(0.010)	(0.158)
Treated x t+1	0.017***	-0.004	-0.001	-0.004	0.217
	(0.006)	(0.009)	(0.010)	(0.011)	(0.148)
Treated x t+2	0.008*	0.015	0.017	0.018	0.224
	(0.005)	(0.010)	(0.011)	(0.011)	(0.159)
Treated x t+3	0.020***	0.034***	0.037***	0.035***	0.026
	(0.006)	(0.011)	(0.012)	(0.012)	(0.184)
Treated x t+4	0.022***	0.034***	0.035***	0.027**	-0.017
	(0.006)	(0.011)	(0.012)	(0.013)	(0.175)
Treated x t+5	0.000	0.051***	0.051***	0.039**	0.021
	(0.012)	(0.019)	(0.020)	(0.020)	(0.283)
Treated x t+6	0.030	0.004	0.007	0.024	0.913**
	(0.021)	(0.013)	(0.013)	(0.030)	(0.386)
Constant	-0.158***	0.100***	0.093***	0.086***	3.806***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.139)
Observations	818,687	751,911	742,595	725,012	818,687
Number of municipalities	7,992	7,991	7,991	7,991	7,992
Year FE	Yes	Yes	Yes	Yes	Yes
Municipality FE	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes
Hospital controls	Yes	Yes	Yes	Yes	Yes
Model	LPM	Linear	Linear	Linear	Linear

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Dead is a dummy measuring whether hospitalization ends with death in the hospital. 30-day, 90-day and 365-day readmission is a dummy measuring whether an AMI hospitalization ends a cardiac/circulatory related or AMI compatible admission within 30, 90 and 365 days respectively. Controls include education, gender, age, Charlson comorbidity index, marital status, type of hospital and AMI admission volume by hospital.

Table 7: Mechanism: effect of closure on travel time

VARIABLES	(1) Travel time	(2) Travel time
Post x Treated	3.808***	
	(0.848)	
Treated x t-6		-2.051
		(1.650)
Treated x t-5		0.067
		(0.910)
Treated x t-4		-0.319
		(1.018)
Treated x t-3		0.120
		(0.544)
Treated x t-2		0.452
		(0.457)
Treated x t=0		1.426**
		(0.575)
Treated x t+1		5.701***
		(0.999)
Treated x t+2		4.711***
		(0.942)
Treated x t+3		3.942***
		(1.120)
Treated x t+4		4.595***
		(1.037)
Treated x t+5		3.387***
		(0.803)
Treated x t+6		4.254***
		(1.513)
Constant	11.016***	10.950***
	(1.648)	(1.631)
Observations	818,477	818,347
Number of municipalities	7,992	7,992
Municipality FE	Yes	Yes
Year FE	Yes	Yes
Model	Linear	Linear

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Travel time measures time taken to travel from the center of municipality of residence to center of municipality of hospital. Controls include size of municipality population and municipality population aged 65+ (%).

Table 8: Mechanism: Effect of closure on bed utilization rate

VARIABLES	(1) Utilization rate beds CCU	(2) Utilization rate beds CCU
Post x treated	7.162***	
	(1.957)	
Treated x t-4		3.252
		(3.422)
Treated x t-2		0.726
		(2.126)
Treated x t-1		2.297
		(2.448)
Treated x t=0		0.135
		(3.016)
Treated x t+1		13.397***
		(3.975)
Treated x t+2		9.812***
		(3.359)
Treated x t+3		11.876***
		(3.091)
Treated x t+4		12.935***
		(3.294)
Treated x t+5		14.130***
		(3.557)
Treated x t+6		11.344**
		(4.431)
Constant	43.473***	41.226***
	(0.388)	(2.413)
Observations	548,420	548,302
Number of municipalities	7,887	7,887
Municipality FE	Yes	Yes
Year FE	Yes	Yes
Model	Linear	Linear

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Utilization rate beds CCU measures the number of yearly AMI admissions at the hospital level over the number of beds in the coronary care unit.

Table 9: Mechanism-effect of closure induced change in travel time on AMI patient outcomes

VARIABLES	(1) Dead	(2) 30-day readmission	(3) 90-day readmission	(4) 365-day readmission	(5) Length of Stay
Travel time	0.006*	0.003	0.003	0.003	0.148*
	(0.003)	(0.003)	(0.003)	(0.003)	(0.077)
Observations	818,348	751,578	742,261	724,683	818,348
Number of municipalities	7,881	7,870	7,869	7,868	7,881
Kleibergen-Paap F-stat.	4.1	7.9	7.9	7.9	4.1
Baseline controls	Yes	Yes	Yes	Yes	Yes
Municipality FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Years	2008-2015	2008-2015	2008-2015	2008-2015	2008-2015

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

Dead is a dummy measuring whether hospitalization ends with death in the hospital. 30-day, 90-day and 365-day readmission is a dummy measuring whether an AMI hospitalization ends in a cardiac/circulatory related or AMI compatible admission within 30, 90 and 365 days respectively. Travel time measures time taken to travel from the center of municipality of residence to center of municipality of hospital. Controls include education, gender, age, Charlson comorbidity index, marital status, type of hospital and AMI admission volume by hospital, population size of municipality and municipality population aged 65+ (%).

Table 10: Mechanism: effect of closure-induced change in utilization rates of bed on AMI patient outcomes

VARIABLES	(1) Dead	(2) 30-day readmission	(3) 90-day readmission	(4) 365-day readmission	(5) Length of Stay
Utilization rate beds CCU	0.002**	0.000	0.001	0.001	0.024
	(0.001)	(0.001)	(0.001)	(0.001)	(0.020)
Observations	547,901	502,740	496,620	484,191	547,901
Number of municipalities	7,666	7,605	7,601	7,593	7,666
Kleibergen-Paap F-stat.	11.1	12.8	12.8	12.6	11.1
Baseline controls	Yes	Yes	Yes	Yes	Yes
Municipality FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes

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

Dead is a dummy measuring whether hospitalization ends with death in the hospital. 30-day, 90-day and 365-day readmission is a dummy measuring whether an AMI hospitalization ends in a cardiac/circulatory related or AMI compatible admission within 30, 90 and 365 days respectively. Utilization rate beds CCU measures the number of yearly AMI admissions at the hospital level over the number of beds in the coronary care unit. Controls include education, gender, age, Charlson comorbidity index, marital status, type of hospital and AMI admission volume by hospital, population size of municipality and municipality population aged 65+ (%).

Supplementary Tables

Table S1: List of primary diagnosis categories included in the estimation of readmission rates

Diagnosis category	ICD-9 code
Rheumatic heart disease	390-398
Hypertensive heart disease	401, 402, 402.01, 402.11, 402.91, 403, 404, 404.01, 404.03, 404.11, 404.13, 404.91, 404.93, 405
Ischemic heart disease	410-414
Disease of pulmonary circulation	415–417
Other forms of heart disease	420, 421, 421.9, 422, 422.90, 422.99, 423, 424, 425, 426, 427, 428, 429
Cerebrovascular Disease	437, 438
Diseases Of Arteries, Arterioles, And Capillaries	440, 441 441.01 441.1 441.2, 442, 443, 444, 444.1, 445, 446, 447
Diseases Of Veins And Lymphatics, And Other Diseases Of Circulatory System	451, 453, 459
Other AMI compatible disease categories	
Fluid overload disorder	276.6
Pulmonary collapse	518
Vascular insufficiency of intestine	557
General symptoms	780
Symptoms involving cardiovascular system	785
Other ill-defined and unknown causes of morbidity and mortality	799
Complications of surgical and medical care	997, 998

Table S2: Descriptive Statistics

Variable	Group averages						
	Untreated	Treated			T test or χ2 test		
		t-1	t	t+1	t-1	t	t+1
Main outcomes							
In-hospital death	6.5%	3.4%	4.5%	4.8%	<0.001	<0.001	<0.001
30-day readmission	4.6%	7.4%	7.7%	6.1%	<0.001	<0.001	<0.001
90-day readmission	5.2%	7.8%	8.3%	6.8%	<0.001	<0.001	<0.001
365-day readmission	6.2%	9.2%	9.0%	8.0%	<0.001	<0.001	<0.001
Length of stay	7.27 (5.30)	6.64 (4.76)	6.55 (4.78)	6.83(4.84)	<0.001	<0.001	<0.001
Channels							
Travel time	14.55 (17.93)	20.91 (19.87)	22.43 (18.79)	26.40 (18.57)	<0.001	<0.001	<0.001
Utilization rate beds CCU*	44.98 (22.43)	35.78 (19.90)	33.23 (21.73)	48.26 (34.44)	<0.001	<0.001	<0.001

Note: The table reports means (standard deviations) for variables at the individual (AMI patient) level. Descriptive statistics are reported for the treatment and control group separately. Variables in the treatment group are measured in year prior to hospital closure (*t-1*) and year of hospital closure (*t*). * Data available for years 2010-2015 only.

Table S3: Trends in AMI hospitalizations – Staggered difference-in-difference

VARIABLES	(1) Total AMI hospitalizations	(2) AMI hospitalization rate
Treated x t-6	0.164*	-0.044
	(0.085)	(0.159)
Treated x t-5	0.049	0.078
	(0.066)	(0.070)
Treated x t-4	0.006	-0.017
	(0.038)	(0.057)
Treated x t-3	0.030	0.011
	(0.034)	(0.051)
Treated x t-2	0.082***	0.080*
	(0.028)	(0.044)
Treated x t=0	-0.010	-0.008
	(0.030)	(0.046)
Treated x t+1	-0.089**	-0.102*
	(0.038)	(0.055)
Treated x t+2	0.065	0.007
	(0.041)	(0.045)
Treated x t+3	0.161***	0.161***
	(0.051)	(0.050)
Treated x t+4	0.010	-0.044
	(0.056)	(0.050)
Treated x t+5	-0.016	-0.107
	(0.065)	(0.100)
Treated x t+6	0.145*	-0.010
	(0.084)	(0.135)
Observations	55,168	55,168
Number of municipalities	7,852	7,852
Municipality FE	Yes	Yes
Year FE	Yes	Yes
Model	Poisson	Linear

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1
AMI hospitalization rate is calculated as (total AMI hospitalizations in a municipality/ total population in a municipality). Controls include municipality population aged 65+ (%).

Table S4: Effect of home hospital closure on patient outcomes- Individual level (excluding large municipalities).

VARIABLES	(1) Dead	(2) 30-day readmission	(3) 90-day readmission	(4) 365-day readmission	(5) Length of Stay
Post x Treated	0.009***	0.013**	0.012*	0.011	0.181
	(0.003)	(0.006)	(0.006)	(0.007)	(0.111)
Constant	-0.124***	0.266***	0.258***	0.240***	0.008
	(0.026)	(0.048)	(0.048)	(0.052)	(0.812)
Observations	532,260	482,006	476,015	464,803	532,260
Number of municipalities	7,854	7,853	7,853	7,853	7,854
Pop	<50k	<50k	<50k	<50k	<50k
Patient/hospital characteristics	Yes	Yes	Yes	Yes	Yes
Municipality FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Region x Year FE	No	No	No	No	No

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.
Dead is a dummy measuring whether a hospitalization ends in death. 30-day, 90-day and 365-day readmission is a dummy measuring whether an AMI hospitalization ends a cardiac/circulatory related or AMI compatible admission within 30, 90 and 365 days respectively. Controls include education, gender, age, Charlson comorbidity index, marital status, type of hospital and AMI admission volume by hospital.

Table S5: Effect of home hospital closure on patient outcomes- Individual level parallel trends (by varying municipality size thresholds).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
VARIABLES		Dead			30-day readmission			Length of stay	
Treated x t-6	-0.031	-0.031	-0.028	0.045	0.043	0.040	-0.177	-0.176	-0.117
	(0.029)	(0.029)	(0.029)	(0.055)	(0.055)	(0.055)	(0.703)	(0.705)	(0.703)
Treated x t-5	0.008	0.008	0.017	0.016	0.015	0.021	-0.186	-0.191	-0.056
	(0.010)	(0.010)	(0.012)	(0.012)	(0.012)	(0.013)	(0.237)	(0.242)	(0.238)
Treated x t-4	0.009	0.009	0.014*	0.009	0.004	-0.001	-0.440**	-0.404*	-0.437*
	(0.007)	(0.008)	(0.008)	(0.009)	(0.010)	(0.009)	(0.207)	(0.220)	(0.228)
Treated x t-3	0.000	-0.000	0.003	0.002	-0.002	-0.008	-0.205	-0.141	-0.100
	(0.005)	(0.006)	(0.006)	(0.008)	(0.008)	(0.009)	(0.160)	(0.173)	(0.183)
Treated x t-2	0.008	0.007	0.010	0.008	0.006	0.004	-0.149	-0.097	-0.120
	(0.005)	(0.006)	(0.006)	(0.008)	(0.008)	(0.009)	(0.132)	(0.141)	(0.153)
Treated x t=0	0.011**	0.008*	0.007	-0.001	0.000	0.003	-0.179	-0.282*	-0.227
	(0.005)	(0.005)	(0.006)	(0.009)	(0.009)	(0.010)	(0.157)	(0.157)	(0.165)
Treated x t+1	0.016***	0.017***	0.019***	-0.004	-0.002	-0.010	0.231	0.161	0.275*
	(0.006)	(0.006)	(0.007)	(0.009)	(0.010)	(0.010)	(0.148)	(0.157)	(0.163)
Treated x t+2	0.007	0.007	0.004	0.016	0.020*	0.013	0.244	0.189	0.257
	(0.005)	(0.005)	(0.006)	(0.010)	(0.011)	(0.010)	(0.159)	(0.168)	(0.182)
Treated x t+3	0.018***	0.015**	0.018***	0.035***	0.042***	0.037***	0.042	-0.096	-0.037
	(0.006)	(0.006)	(0.007)	(0.011)	(0.012)	(0.012)	(0.183)	(0.178)	(0.183)
Treated x t+4	0.021***	0.021***	0.023***	0.035***	0.035***	0.029**	-0.001	-0.027	-0.092
	(0.006)	(0.007)	(0.007)	(0.011)	(0.012)	(0.013)	(0.175)	(0.187)	(0.207)
Treated x t+5	-0.001	-0.002	-0.002	0.052***	0.055***	0.040**	0.045	-0.012	0.081
	(0.012)	(0.012)	(0.015)	(0.019)	(0.019)	(0.020)	(0.280)	(0.281)	(0.339)
Treated x t+6	0.027	0.026	0.027	0.003	0.007	0.000	0.921**	0.827**	0.894**
	(0.021)	(0.021)	(0.022)	(0.013)	(0.013)	(0.014)	(0.378)	(0.378)	(0.386)
Constant	-0.152***	-0.150***	- 0.152***	0.118***	0.120***	0.118***	3.778***	3.767***	3.809***
	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)	(0.094)	(0.083)	(0.082)
Size of municipality	Pop<100k	Pop<50k	Pop<30k	Pop<100k	Pop<50k	Pop<30k	Pop<100 k	Pop<50k	Pop<30k
Observations	628,182	532,130	443,010	569,722	481,884	402,080	628,182	532,130	443,010
Number of municipalities	7,947	7,854	7,689	7,946	7,853	7,688	7,947	7,854	7,689
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Municipality FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Hospital controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Model	LPM	LPM	LPM	LPM	LPM	LPM	LPM	LPM	LPM

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Dead is a dummy measuring whether hospitalization ends in death. 30-day readmission is a dummy measuring whether hospitalization ends with cardiac/circulatory related and AMI compatible admission within 30 days of an index AMI admission. Controls include education, gender, age, Charlson comorbidity index, marital status, type of hospital and AMI admission volume by hospital.

Table S6: Mechanism: effect of closure on travel time- Individual level (by varying municipality size threshold)

VARIABLES	(1) Travel time	(2) Travel time	(3) Travel time
Treated x t-6	-1.508	-1.396	-1.284
	(1.640)	(1.643)	(1.645)
Treated x t-5	0.515	0.518	0.701
	(0.915)	(0.926)	(1.110)
Treated x t-4	-0.045	0.563	0.721
	(1.021)	(0.874)	(0.887)
Treated x t-3	0.308	0.715	0.877
	(0.546)	(0.535)	(0.592)
Treated x t-2	0.541	0.625	0.615
	(0.457)	(0.500)	(0.567)
Treated x t=0	1.366**	0.902	0.858*
	(0.575)	(0.550)	(0.509)
Treated x t+1	5.561***	4.593***	4.903***
	(0.994)	(0.831)	(0.947)
Treated x t+2	4.517***	3.742***	3.990***
	(0.946)	(0.892)	(0.967)
Treated x t+3	3.701***	2.743***	3.515***
	(1.126)	(1.054)	(1.011)
Treated x t+4	4.288***	3.595***	4.540***
	(1.038)	(1.001)	(1.061)
Treated x t+5	3.028***	2.392***	2.974***
	(0.804)	(0.752)	(0.799)
Treated x t+6	3.724**	2.844**	3.440**
	(1.493)	(1.416)	(1.434)
Constant	20.387***	26.996***	27.344***
	(2.849)	(2.575)	(2.215)
Size of municipality	Pop <100k	Pop <50k	Pop <30k
Observations	628,199	532,146	443,023
Number of municipalities	7,947	7,854	7,689
Municipality FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Model	Linear	Linear	Linear

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses.
Travel time measures time taken to travel from the center of municipality of residence to center of municipality of hospital. Controls include size of municipality population and municipality population aged 65+ (%).

Table S7: Mechanism: effect of hospital closure on average travel time- municipality level (by varying municipality size threshold)

VARIABLES	(1) Average travel time	(2) Average travel time	(3) Average travel time
Treated x t-6	1.756	1.800	1.841
	(1.797)	(1.797)	(1.800)
Treated x t-5	0.602	0.617	0.672
	(1.074)	(1.075)	(1.110)
Treated x t-4	0.017	0.064	0.097
	(0.938)	(0.943)	(0.964)
Treated x t-3	1.253*	1.298*	1.348*
	(0.753)	(0.759)	(0.779)
Treated x t-2	1.666**	1.686**	1.720**
	(0.762)	(0.768)	(0.789)
Treated x t=0	0.806	0.759	0.737
	(0.746)	(0.751)	(0.767)
Treated x t+1	5.503***	5.414***	5.399***
	(0.821)	(0.825)	(0.846)
Treated x t+2	4.037***	3.950***	3.997***
	(0.850)	(0.855)	(0.875)
Treated x t+3	2.914***	2.805***	2.906***
	(0.995)	(1.002)	(1.020)
Treated x t+4	4.367***	4.284***	4.395***
	(1.092)	(1.099)	(1.126)
Treated x t+5	2.382**	2.291**	2.336*
	(1.157)	(1.158)	(1.195)
Treated x t+6	2.468	2.387	2.458
	(2.508)	(2.508)	(2.513)
Constant	26.456***	27.926***	28.555***
	(1.577)	(1.658)	(1.792)
Size of municipality	Pop <100k	Pop <50k	Pop <30k
Observations	54,941	54,138	52,845
Number of municipalities	7,947	7,854	7,689
Municipality FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Model	Linear	Linear	Linear

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1
Average travel time is computed as the mean of travel time for all residents of a municipality in a year. Controls include size of municipality population and municipality population aged 65+ (%).

Table S8: Mechanism: effect of closure on bed utilization rate- Individual level (by varying municipality size threshold)

VARIABLES	(1) Utilization rate beds CCU	(2) Utilization rate beds CCU	(3) Utilization rate beds CCU
Treated x t-4	3.432	3.266	2.811
	(3.435)	(3.452)	(3.270)
Treated x t-2	0.518	1.271	0.853
	(2.117)	(2.033)	(2.329)
Treated x t-1	1.864	2.094	1.061
	(2.434)	(2.438)	(2.122)
Treated x t=0	-0.595	-0.813	-1.545
	(2.986)	(3.009)	(2.462)
Treated x t+1	12.472***	12.437***	11.779***
	(3.947)	(4.075)	(3.433)
Treated x t+2	8.739***	9.432***	7.794***
	(3.332)	(3.375)	(2.534)
Treated x t+3	10.783***	12.159***	10.392***
	(3.060)	(3.043)	(2.382)
Treated x t+4	11.806***	13.228***	10.666***
	(3.266)	(3.200)	(2.463)
Treated x t+5	12.738***	13.762***	11.181***
	(3.522)	(3.536)	(2.871)
Treated x t+6	9.789**	11.298***	9.147**
	(4.396)	(4.383)	(3.883)
Constant	40.187***	40.185***	41.052***
	(2.413)	(2.415)	(2.098)
Observations	413,858	348,597	287,324
Number of municipalities	7,842	7,748	7,582
Municipality FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Model	Linear	Linear	Linear
Resident Population	<100k	<50k	<30k

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.
Utilization rate beds CCU measures the number of yearly AMI admissions at the hospital level over the number of beds in the coronary care unit.

Table S9: Mechanism: effect of closure-induced change in travel time on AMI patient outcomes- Individual level (by varying population size)

VARIABLES	(1) Dead	(2) Dead	(3) Dead	(4) 30-day readmission	(5) 30-day readmission	(6) 30-day readmission	(7) Length of stay	(8) Length of stay	(9) Length of stay
Travel time	0.013	-0.018	-0.140	0.005	0.025	0.014	0.388	-0.488	-4.532
	(0.015)	(0.025)	(2.130)	(0.006)	(0.041)	(0.019)	(0.414)	(0.989)	(69.386)
Observations	628,201	402,198	443,029	569,723	481,885	402,081	628,201	532,149	443,029
Number of municipalities	7,836	7,665	7,578	7,825	7,732	7,567	7,836	7,743	7,578
Kleibergen-Paap F-stat.	0.7	0.6	0.0	3.1	0.5	1.0	0.7	0.3	0.0
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Municipality FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Years	2008-2015	2008-2015	2008-2015	2008-2015	2008-2015	2008-2015	2008-2015	2008-2015	2008-2015
Size of municipality	Pop<100k	Pop<50k	Pop.<30k	Pop.<100k	Pop<50k	Pop<30k	Pop<100k	Pop<50k	Pop<30k

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Travel time measures time taken to travel from the center of municipality of residence to center of municipality of hospital. Controls include education, gender, age, Charlson comorbidity index, marital status, type of hospital and AMI admission volume by hospital, size of municipality population and municipality population aged 65+ (%).

Table S10: Mechanism: effect of closure-induced change in utilization rates of bed on AMI patient outcomes- Individual level (by varying population size)

VARIABLES	(1) Dead	(2) Dead	(3) Dead	(4) 30-day readmission	(5) 30-day readmission	(6) 30-day readmission	(7) Length of stay	(8) Length of stay	(9) Length of stay
Utilization rate beds CCU	0.002*	0.002	0.001	0.001	0.001	0.001	0.033	0.020	0.030
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.025)	(0.025)	(0.031)
Travel time	-0.000	-0.000	-0.000	-0.001***	-0.001***	-0.001***	-0.006	-0.003	-0.002
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.006)	(0.006)	(0.007)
Observations	413,755	348,494	287,221	374,092	314,654	260,065	413,755	348,494	287,221
Number of municipalities	7,621	7,527	7,361	7,560	7,466	7,300	7,621	7,527	7,361
Kleibergen-Paap F-stat.	7.5	6.5	5.7	9.0	7.5	6.4	7.5	6.5	5.7
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Municipality FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Municipality size	<100k	<50k	<30k	<100k	<50k	<30k	<100k	<50k	<30k

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

Dead is a dummy measuring whether hospitalization ends with death in the hospital. 30-day readmission is a dummy measuring whether hospitalization ends with cardiac/circulatory related and AMI compatible readmission within 30 days. Utilization rate beds CCU measures the number of yearly AMI admissions at the hospital level over the number of beds in the coronary care unit. Controls include education, gender, age, Charlson comorbidity index, marital status, type of hospital and AMI admission volume by hospital. size of municipality population and municipality population aged 65+ (%).

Table S11: Effect of home hospital closure on patient outcomes- Individual level (controlling for measure of economic deprivation)

VARIABLES	(1) Dead	(2) 30-day readmission	(3) 90-day readmission	(4) 365-day readmission	(5) Length of Stay
Post x Treated	0.012***	0.006	0.007	0.005	0.254*
	(0.004)	(0.007)	(0.007)	(0.008)	(0.130)
Resident population living below poverty threshold (%)	0.087*	0.007	0.022	0.035	3.591**
	(0.052)	(0.080)	(0.079)	(0.075)	(1.640)
Constant	-0.209***	0.171***	0.148***	0.115***	-0.917
	(0.024)	(0.039)	(0.038)	(0.038)	(0.873)
Observations	617,908	566,662	559,609	545,327	617,908
Number of municipalities	7,962	7,961	7,960	7,960	7,962
Patient/hospital/municipality characteristics	Yes	Yes	Yes	Yes	Yes
Municipality FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Region x Year FE	No	No	No	No	No

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

Dead is a dummy measuring whether a hospitalization ends in death. 30-day, 90-day and 365-day readmission is a dummy measuring whether hospitalization ends with cardiac/circulatory related readmission within 30, 90 and 365 days respectively. Controls include education, gender, age, Charlson comorbidity index, marital status, type of hospital and AMI admission volume by hospital.

Table S12: Effect of home hospital closure on patient outcomes- Municipality level (controlling for measure of economic deprivation)

VARIABLES	(1) Total in hospital deaths	(2) Total 30 day readmission	(3) Total 90 day readmission	(4) Total 365 day readmission	(5) Mean Length of Stay
Post x treated	0.259**	-0.022	-0.004	0.001	0.248
	(0.101)	(0.078)	(0.079)	(0.074)	(0.204)
Resident population living below poverty threshold (%)	1.074	-0.611	-0.334	-0.261	1.625
	(0.841)	(1.144)	(1.052)	(0.894)	(1.783)
Constant					4.751***
					(0.939)
Observations	32,099	25,103	26,857	28,734	41,439
Number of municipalities	5,687	4,403	4,717	5,057	7,962
Patient/hospital/municipality characteristics	Yes	Yes	Yes	Yes	Yes
Municipality FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Region x Year FE	No	No	No	No	No

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.
Total in-hospital deaths measures the number of AMI in-hospital deaths for each municipality. Total 30-day, 90-day and 365-day readmission measures the number of hospital readmissions within 30, 90 and 365 days for each municipality. Mean length of stay measures the average length of stay for each municipality. Controls include size of municipality population and municipality population aged 65+ (%).

Table S13: Effect of home hospital closure on patient outcomes - Municipality level

VARIABLES	(1) Total in-hospital deaths	(2) Total 30-day readmissions	(3) Total 90-day readmissions	(4) Total 365-day readmissions	(5) Mean length of stay
Post x treated	0.235***	0.006	-0.003	0.011	0.317**
	(0.075)	(0.076)	(0.071)	(0.064)	(0.149)
Constant					6.545***
					(0.512)
Observations	45,997	36,984	39,264	41,728	55,322
Number of municipalities	6,180	4,905	5,215	5,556	7,992
Municipality FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Model	Poisson	Poisson	Poisson	Poisson	Linear

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Total in-hospital deaths measures the number of AMI in-hospital deaths for each municipality. Total 30-day, 90-day and 365-day readmission measures the number of hospital readmissions within 30, 90 and 365 days respectively for each municipality. Mean length of stay measures the average length of stay for each municipality. Controls include size of municipality population and municipality population aged 65+ (%).

Table S14: Effect of home hospital closure on AMI patient outcomes– Municipality level parallel trends

VARIABLES	(1) Total in-hospital deaths	(2) Total 30-day readmission	(3) Total 90-day readmission	(4) Total 365-day readmission	(5) Mean length of stay
Treated x t-6	-0.366	0.432	0.494	0.424	-3.219***
	(0.373)	(0.573)	(0.400)	(0.362)	(1.177)
Treated x t-5	0.158	0.169	0.116	0.010	-0.032
	(0.174)	(0.134)	(0.133)	(0.143)	(0.472)
Treated x t-4	0.240	0.094	0.147	0.084	-0.744**
	(0.153)	(0.123)	(0.117)	(0.106)	(0.347)
Treated x t-3	-0.013	-0.016	0.056	-0.005	-0.523*
	(0.143)	(0.111)	(0.108)	(0.101)	(0.299)
Treated x t-2	0.146	0.074	0.137	0.104	-0.430
	(0.141)	(0.098)	(0.093)	(0.090)	(0.273)
Treated x t=0	0.311**	-0.014	0.012	-0.054	-0.459*
	(0.124)	(0.111)	(0.113)	(0.109)	(0.272)
Treated x t+1	0.382***	-0.170	-0.110	-0.090	-0.099
	(0.142)	(0.135)	(0.133)	(0.123)	(0.260)
Treated x t+2	0.177	0.061	0.089	0.113	0.291
	(0.129)	(0.124)	(0.120)	(0.109)	(0.280)
Treated x t+3	0.417***	0.219*	0.246*	0.211*	0.255
	(0.162)	(0.131)	(0.130)	(0.120)	(0.300)
Treated x t+4	0.477***	0.244*	0.256*	0.192	-0.294
	(0.179)	(0.131)	(0.134)	(0.128)	(0.300)
Treated x t+5	-0.044	0.426**	0.430**	0.311*	0.965*
	(0.402)	(0.181)	(0.179)	(0.170)	(0.538)
Treated x t+6	0.534*	-0.469	-0.452	0.275	0.442
	(0.318)	(0.566)	(0.565)	(0.641)	(0.670)
Constant					6.553***
					(0.513)
Observations	45,984	36,975	39,254	41,719	55,308
Number of municipalities	6,180	4,905	5,215	5,556	7,992
Year FE	Yes	Yes	Yes	Yes	Yes
Municipality FE	Yes	Yes	Yes	Yes	Yes
Municipality controls	Yes	Yes	Yes	Yes	Yes
Hospital controls	Yes	Yes	Yes	Yes	Yes
Model	Poisson	Poisson	Poisson	Poisson	Linear

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Total in-hospital deaths measures the number of AMI in-hospital deaths for each municipality. Total 30-day, 90-day and 365-day readmission measures the number of hospital readmissions within 30, 90 and 365 days respectively for each municipality. Mean length of stay measures the average length of stay for each municipality. Controls include size of municipality population and municipality population aged 65+ (%). For Poisson models, we use number of AMI hospitalizations at the municipality as offset variable.

Table S15: Mechanism: effect of closure on travel time-Municipality level

VARIABLES	(1) Average travel time	(2) Average travel time
Post x treated	2.595***	
	(0.556)	
Treated x t-6		1.661
		(1.799)
Treated x t-5		0.546
		(1.075)
Treated x t-4		-0.018
		(0.938)
Treated x t-3		1.232
		(0.753)
Treated x t-2		1.656**
		(0.762)
Treated x t=0		0.815
		(0.746)
Treated x t+1		5.521***
		(0.821)
Treated x t+2		4.060***
		(0.850)
Treated x t+3		2.941***
		(0.995)
Treated x t+4		4.400***
		(1.093)
Treated x t+5		2.438**
		(1.157)
Treated x t+6		2.503
		(2.511)
Constant	23.151***	23.095***
	(1.393)	(1.393)
Observations	55,322	55,308
Number of municipalities	7,992	7,992
Municipality FE	Yes	Yes
Year FE	Yes	Yes
Model	Linear	Linear

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. Average travel time is computed as the mean of travel time for all residents of a municipality in a year. Controls include size of municipality population and municipality population aged 65+ (%).

Table S16: Mechanism: effect of closure-induced change in travel time on AMI patient outcomes- Municipality level

VARIABLES	(1) Total in-hospital deaths	(2) Total 30 day readmission	(3) Total 90 day readmission	(4) Total 365 day readmission	(5) Mean length of stay
Average time	0.090***	0.002	-0.001	0.004	0.077*
	(0.029)	(0.029)	(0.028)	(0.025)	(0.045)
Observations	45,997	36,984	39,264	41,728	55,182
Number of municipalities	6,180	4,905	5,215	5,556	7,852
Municipality FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Model	Poisson	Poisson	Poisson	Poisson	Linear
Years	2008-2015	2008-2015	2008-2015	2008-2015	2008-2015

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Total in-hospital deaths measures the number of AMI in-hospital deaths for each municipality. Total 30-day, 90-day and 365-day readmission measures the number of hospital readmissions within 30, 90 and 365 days respectively for each municipality. Mean length of stay measures the average length of stay for residents from each municipality. Controls include population size o0f municipality and municipality population aged 65+ (%). For Poisson models, we use number of AMI hospitalizations at the municipality as offset variable.

Table S17: Effect of home hospital closure on patient outcomes (Callaway and Sant'Anna estimates)

Outcomes	Control group	
	Never treated	Not-yet treated
In-hospital mortality	0.019 (p=0.026)	0.020 (p=0.027)
30-day readmission	0.016 (p=0.276)	0.016 (p=0.278)
Length of stay	-0.823 (p<0.001)	-0.825 (p<0.001)

Dead is a dummy measuring whether hospitalization ends with death in the hospital. 30-day readmission is a dummy measuring whether an AMI hospitalization ends in a cardiac/circulatory related or AMI compatible admission within 30 days respectively. Controls include education, gender, age, Charlson comorbidity index, marital status, type of hospital and AMI admission volume by hospital, population size of municipality and municipality population aged 65+ (%).

CHAPTER 3

European Covid Survey (ECOS)[24]

This chapter comprises of four sets of research findings (3 published articles and 1 working paper) that studies the commitment and acceptability to COVID-19 preventive measures, and in general a snapshot of the public perception of the European population during the pandemic. All findings in this chapter are based on data from the European Covid Survey (ECOS), that was developed and implemented during the first wave of the COVID-19 (April 2-15, 2020) in Europe. The English version of the questionnaire developed for this study can be found in the appendix.

Ethical approval for this study was provided by the University of Hamburg, Germany under the umbrella project "Countering COVID-19: A European survey on the acceptability of and commitment to preventive measures". Subject recruitment and payment were done through the Agency Dynata. Proprietary panels used double opt-in recruitment and a written informed consent was obtained from individual participants by Dynata. Confidentiality and anonymity of the participants were ensured by issuing a unique identifier to each respondent. Study participants were informed about their freedom to opt-out of the study at any point of time. No patients are involved in this research, and we strictly followed the RESPECT Code of practice for the conduct of socio-economic research in Europe.

[24] The development of the survey, data collection and findings from the subsequent analyses in this chapter is the result of the collaboration between the principal investigators of this project: Iryna Sabat[a], Sebastian Neuman-Böhme[b], Nirosha ElsemVarghese[c], Pedro Pita Barros[a], Aleksandra Torbica[c], Werner Brouwer[b, d], Job van Exel[b,d], Jonas Schreyögg[e], Tom Stargardt[e]

- a Nova School of Business and Economics, Portugal
- b Erasmus School of Health Policy and Management, Erasmus University Rotterdam, the Netherlands
- c Centre for Research on Health and Social Care Management, CERGAS, Bocconi University, Italy
- d Erasmus School of Economics, Erasmus University Rotterdam, the Netherlands
- e Hamburg Center for Health Economics, University of Hamburg, Germany

Chapter 3.1

United but divided: Policy responses and people's perceptions in the EU during the COVID-19 outbreak[25]

ABSTRACT

To understand the public sentiment toward the measures used by policymakers for COVID-19 containment, a survey among representative samples of the population in seven European countries was carried out in the first two weeks of April 2020. The study addressed people's support for containment policies, worries about COVID-19 consequences, and trust in sources of information. Citizens were overall satisfied with their government's response to the pandemic; however, the extent of approval differed across countries and policy measures. A north-south divide in public opinion was noticeable across the European states. It was particularly pronounced for intrusive policy measures, such as mobile data use for movement tracking, economic concerns, and trust in the information from the national government. Considerable differences in people's attitudes were noticed within countries, especially across individual regions and age groups. The findings suggest that the epidemic acts as a stressor, causing health and economic anxieties even in households that were not directly affected by the virus. At the same time, the burden of stress was unequally distributed across regions and age groups. Based on the data collected, we draw lessons from the containment stage and identify several insights that can facilitate the design of lockdown exit strategies and future containment policies so that a high level of compliance can be expected.

[25] The paper constituting this chapter was published in *Health Policy* journal. Author list as follows: Iryna Sabat, Sebastian Neuman-Böhme, Nirosha Elsem Varghese, Pedro Pita Barros, Werner Brouwer, Job van Exel, Aleksandra Torbica, Jonas Schreyögg and Tom Stargardt

INTRODUCTION

The outbreak of COVID-19 triggered a wide range of responses from governments in the European Union. Given that the disease was new and effective medical countermeasures did not exist in early 2020, governments had to adopt non-medical measures aiming at the containment and mitigation of COVID-19. With the aim of "flattening the curve," these policies included bans on public gatherings, closures of academic institutions and public places, national and international mobility restrictions, confinement, and several others (IMF Policy Tracker, 2020).

Italy was the first country in Europe to apply intervention measures from the beginning of March 2020 in response to the severity of the COVID-19 outbreak. Other EU countries followed soon afterward, using similar countermeasures around mid-March 2020 (Flaxman et al., 2020). The adoption of these policies varied in their scale, stringency, and pace across countries. While most European states implemented confinement measures, the extent of limitations of people's freedoms differed across individual countries. Lockdowns were usually strictest where the pandemic was deadliest (Italy, Spain, and France), imposing severe limitations on population movements. Some governments chose less stringent versions of confinement or no lockdown at all, for instance, "an intelligent lockdown" in the Netherlands or "freedom under responsibility" in Sweden (The Economist, 2020).

Forced to react swiftly to the unfolding epidemic situation, policymakers in every country tried to balance the implementation of containment policies against numerous important factors with the priority mostly given to the protection of the population's health. Consequently, there has been a lot of debate in every society about whether measures taken by the government were appropriate or not. Some parts of the population have been voicing support for more severe containment policies to minimize the spread of the virus. Such attitudes were likely fueled by people's worries about their health and the potential of their national healthcare system to withstand the epidemic. Meanwhile, others expressed their concerns about the social and economic consequences of such policies, thereby advocating for less severe containment measures (Financial Times, 2020).

As the pandemic began to abate, governments started designing the lockdown exit strategies and restarting their economies. However, the risk that the new wave of the epidemic may happen did not disappear, especially given that the vaccine development takes a long time, and herd immunity was not achieved (Graeden et al., 2020). In this light, the issue of lifting lockdowns has become a new subject of public debate across and within European countries raising discussions about the appropriateness of timing, risks, and potential consequences of ending the confinement (The NewYork Times, 2020). Lifting lockdown restrictions creates acute dilemmas to the policymakers since the economic and human costs of any exit strategy seem to be closely linked together. Taking a utilitarian approach in this situation could backfire if the society's understanding is not preliminarily secured or expectations are not fulfilled.

Policymakers and public health experts have to persuade their citizens to make behavior changes and respect future containment interventions while facing the difficulty of enforcing such regulations. Therefore, it becomes crucial to understand people's worries about the pandemic and their perceptions of the effects of containment policies, so that the design of further policies and contingency measures is well-informed, and a high level of compliance can be expected from the population. Moreover, trust in the government and social institutions may become central to achieving a successful implementation of future measures, whereas lack of it may turn detrimental to the fight against the pandemic. Hence it is of paramount importance to understand who people trust most so that public health messages can be amplified using correct means of communication.

We provide a timely description of the current situation and draw lessons from the containment stage to inform the design and implementation of the lockdown exit policies.

MATERIALS AND METHODS

In order to understand the public sentiment towards the COVID-19 containment measures and to inform future policy development, we collected information on people's support for these policies, their worries in relation to the unfolding epidemic, and their trust in different sources of information. We surveyed over 7,000 people representative of the adult population in seven European countries: Denmark, France, Germany, Italy, Portugal, the Netherlands, and the UK. The fieldwork was conducted online during April 2-15, 2020, using multi-sourced online panels provided by the market research company Dynata. To ensure that the sampling frame was representative given the online nature of the study, the company applied diverse recruiting procedures to reach the general population (through open recruitment, loyalty programs, affiliate networks, mobile apps). It then used quotas to match the national census shares in each country.

The questionnaire was designed by the authors of the study except for the worry items that were adopted from the World Health Organization (WHO) COVID-19 Snapshot Monitoring project (Betsch et al., 2020). The questionnaire was carefully translated into six other languages by native speakers and then implemented using the Qualtrics platform first as a pilot (10% of the sample in every country) and next as a large-scale survey. The data from the pilot study were included in the total sample.

In each country, we collected data from a sample of 1,000 respondents' representative of the national population in terms of region, age, gender, and education. Given that the Italian region Lombardy was the most severely hit by the COVID-19 outbreak, we collected 500 additional responses in this region representative in terms of age and gender. Learning about perceptions and attitudes of people who reside there could provide essential insights to researchers and policymakers. The extra data collected from Lombardy were not included in the representative sample of Italy. Thus, no weighting was used as the additional Lombardy sample was analyzed separately and denoted as Lombardy in the results section.

RESULTS

Policy support

We assessed people's approval of policy measures that were taken (or were likely to be taken) by their national government in response to the COVID-19 outbreak. In particular, we covered such issues as school closures, bans on public gatherings, border closures, bans imposed on the export of medical equipment, fines for quarantine violations, random temperature checks, curfews, public transport suspensions and utilization of mobile phone data for tracking COVID-19 cases and their contacts.

On average, 68% of people in the seven European countries approved of the policies taken in their country in response to the pandemic, implying considerable public support. Nevertheless, the extent of approval differed by country and by policy measure.

The most approved measures were fining 14-day quarantine violations, ban of public gatherings, and border closures (each supported by 83% of respondents). By the time of the survey's fieldwork, restrictions on public gatherings had been adopted in all countries covered by the study, whereas international travel controls had been imposed to a certain extent everywhere, except the UK (Hale et al., 2020).

Prior to complete border closures in mid-March 2020, some countries (for example, Italy, France, Germany, Denmark) had been requiring screening and 14-day quarantine for arrivals from high-risk regions already since February. In contrast, other countries, such as Portugal and the Netherlands, started later and turned directly to strict measures, such as banning arrivals from high-risk areas and imposing partial border closures. The latter typically implied either limitation on entries of nonresidents or closure of only certain types of borders (land, sea, air), while ensuring "green lanes" for freight vehicles transporting goods. However, complete border closures occurred haphazardly and led to disrupted commerce and stranding citizens. Among countries covered in our study, Denmark was the first to close all borders in mid-March, whereas the UK did so only in the second half of May 2020. Moreover, at the time of fieldwork, the UK did not have routine screenings at its airports or quarantine requirement for travelers (European Commission, 2016; Hale et al., 2020). Thus, the results for the UK showed the extent of public support that these measures would have received, had they been implemented earlier.

Meanwhile, the most opposed containment policies were public transport suspension (37% of respondents against it), ban of medical export, use of mobile phone data for tracking, and the imposition of a curfew (each disapproved by approximately 23% of respondents).

These trends might reflect within-country regional and age structure of the population. For example, older individuals and those living in remote areas tended to be the most strongly opposed to public transport suspension. In fact, among countries covered by the survey, public transport suspension was implemented only in Italy, whereas its volume was reduced in all other states except for Germany (Hale et al., 2020). The stay-at-home orders were most significantly opposed by the youngest respondents aged below 25. This measure was enforced in all countries covered by the survey except for Denmark, where it was introduced as a recommendation (Hale et al., 2020).

Overall, a north-south gradient could often be noticed in the EU regarding policy support: people living in the southern states (Portugal, Italy, and France) tended to approve of the containment policies more than residents in the northern countries (Denmark, Germany and the Netherlands). Noteworthy, the largest share of supporters for every containment measure was noticed among the residents of Italy and particularly in Lombardy. Here, on average, 79% of the population approved of the government's response to the pandemic.

Fig.1 illustrates the average degree of approval of several selected countermeasures across seven European countries (measured on a Likert scale from 1-strongly disapprove to 5-strongly approve), which highlights how diverse Europe is in the perceptions of COVID-19 policy responses. Higher intensity of the color reflects a higher level of approval of a specific policy by the population in each country.

[Figure 1 about here]

Interestingly, the most significant share of the population who explicitly opposed each of the containment policies taken by their government was identified in Denmark. Here, for example, 22% of respondents disapproved of school closures and 48% disapproved of the imposition of a curfew. In comparison, the average disapproval of these measures in other countries was around 8% for schools and 20% for curfews.

The most polarizing opinions were observed concerning the use of mobile data for tracking COVID-19 cases and their contacts. The most significant share of people explicitly opposing such policy was identified in Denmark (34%), the Netherlands (31%), and Germany (25%). It was particularly disfavored by the youngest age group (33% of respondents aged below 25 against it).

This policy received significant media attention as some countries and the European Commission started the collaboration with telecom providers to access individual geolocation data for prediction and surveillance of COVID-19 spread (Jarman et al., 2020; OECD, 2020). As of March 2020, Deutsche Telekom provided German authorities with the anonymized data on the movement of its users. In Italy, Vodafone, WindTre and Telecom Italia offered aggregated user data provision to the government for the same purpose. Authorities in the Lombardy region used mobile phone data to check compliance with the lockdown restrictions (Jarman et al., 2020; OECD, 2020; Pollina and Busvine, 2020). Other countries either initiated the development of their own mobile phone tracking apps or cooperated on the creation of common software, such as the Pan-European Privacy-Preserving Proximity Tracing (PEPP-PT) project led by Germany. However, the launch of the PEPP-PT was delayed at the end of April due to the data protection concerns voiced by experts and even some of the project participants (Deutsche Welle, 2020).

While proponents of the contact-tracing measures claim that using mobile data is of paramount importance in response to the COVID-19 pandemic, many people worry about the government's use of technology due to possible privacy violations, thereby raising debates about the appropriateness of such social control measures (Ienca and Vayena, 2020; Deutsche Welle, 2020; OECD, 2020). According to our data, people in some European countries expressed considerable reluctance about supporting such policy, which therefore makes future compliance questionable. Moreover, such privacy disputes, as in the case of the PEPP-PT project launch, might trigger higher reluctance among the potential users to use any contact-tracing app in the future, which could be detrimental for the implementation of a viable tracing technology (Deutsche Welle, 2020).

To better understand public opinion on certain policies, it is essential to look at the big picture and place obtained results into the national contexts. People's attitudes were likely based on their perceptions of the general state of affairs in their country, particularly in terms of the epidemic situation and restrictions they were subject to at that moment.

In view of that, Table 1 summarizes the scale of the pandemic and the stringency of government's response in seven European countries at four points of time spaced around April 12 (when the survey's fieldwork was 99% complete in every country). The public health situation in each state is described using total confirmed cases of COVID-19 and total deaths attributed to COVID-19, both measured per 1 million people and reported by the European Centre for Disease Prevention and Control (Roser et al., 2020). The stringency of government's response is measured with the COVID-19 Government Response Stringency Index, a composite measure of containment policies ranging from 1 to 100, where a higher value denotes a stricter response (Hale et al., 2020).

At the time of the survey's fieldwork, the epidemic situation was worst, and the stringency index was highest in Italy and France (Hale et al., 2020; Roser et al., 2020). Clearly, there was a north-south gradient in the stringency of government response: Italy, France and Portugal imposed more demanding policies than Denmark, Germany, the Netherlands, and the UK. Nevertheless, although people in southern countries were exposed to more severe containment measures, they approved of them more than people residing in northern states, who experienced less stringent restrictions.

[Table 1 about here]

Turning now to within-country variations, we observed considerable heterogeneity of attitudes towards many policy responses within individual countries with particularly marked differences between regions and age groups in Italy, France, and the Netherlands.

Hereinafter, we grouped regions based on the severity of the COVID-19 outbreak distinguishing between the most and the least affected areas. Noteworthy, Lombardy denotes the extra sample collected in Italy and was analyzed separately from the representative Italian sample. Overall, we did not find significant differences in policy support between Lombardy and the rest of Italy.

To illustrate within-country differences, Fig.2. reflects regional and age-related heterogeneity of public opinions in France and Italy toward banning the export of medical equipment, such as masks. In fact, this measure was briefly undertaken by Germany and France at the onset of the pandemic in early March 2020, leading to political tensions between the EU member states. Germany declared that the reason was to avoid shortages of masks, gloves and safety glasses within the country, whereas France argued that the ban was needed for the assessment of inventory and storage capacity (Tsang, 2020). Following the call for solidarity, both countries lifted the within-EU export ban on equipment in mid-March (Breton, 2020).

[Figure 2 about here]

While support for this policy tended to be similar in the most and the least severely affected parts of Italy and France, the approval of the export ban conspicuously differed across age groups. Older individuals approved more of this policy than younger people, which, besides other factors, may be related to the levels of worry people in these age categories have about the risks that COVID-19 poses to their health. We found that 51% of French and 46% of Italian respondents aged above 65 perceived risks to their health from COVID-19 as high or very high, while the corresponding share among people aged below 25 equaled 30% in France and 17% in Italy.

Worries about health and the economy

To address the mental health implications of the COVID-19 outbreak and subsequent containment measures, we assessed levels of worry prevailing in European societies over several domains (health, economic, emotional, work, and future). More specifically, we addressed concerns about losing a close person, becoming unemployed, health system getting overloaded, school closures, small companies running out of business, recession, restricted access to food supplies, blackouts, and society getting more egoistic. These items were adopted from the WHO COVID-19 Snapshot Monitoring project, which will allow future comparisons with similar data collected for other countries and at different points in time (Betsch et al., 2020).

We found that the mean trend was similar in all countries: people worried most of all about the health system getting overloaded so that the capacities could become insufficient to cope with the surge in COVID-19 cases. We observed that even in case of households that had not been directly hit by the novel coronavirus (above 75% of respondents in the total sample), the pandemic might have acted as a stressor causing health and economic anxieties.

Fig.3 presents people's worry about selected issues across seven EU countries (measured on a Likert scale from 1-not worry at all to 5-worry a lot), where the higher intensity of color reflects a larger share of the population who worry "quite a bit" or "a lot". Cross-country differences look substantial, and a north-south divide in the worry caused by the COVID-19 outbreak is conspicuous.

[Figure 3 about here]

For instance, 84% of respondents in Portugal and 81% in Italy mentioned that they worried "quite a bit" or "a lot" about the national health system becoming overloaded, while the corresponding shares in Denmark and Germany were 54% and 62%, respectively. These health concerns might have reflected the development of the pandemic. As showed in Table 1, the progress of the epidemic had a north-south pattern with more COVID-19 cases and deaths per million of the population in southern states than in northern. The exception was the UK, where the epidemic was third deadliest after Italy and France, but government response was less strict than in countries with a better epidemiological situation (Roser et al., 2020; Hale et al., 2020).

Similarly, more people in Portugal and Italy were concerned with the economic consequences of the pandemic than in other European countries. For example, 68% of Portuguese and 56% of Italians were worried about losing their jobs, while respective shares in the Netherlands and Denmark were 27% and 16%, correspondingly.

These cross-country differences in economic anxieties may be related to people's perceptions of the economic and financial countermeasures taken by their national government and the EU. During the pandemic, European countries implemented several fiscal and monetary measures to mitigate the economic impact of the COVID-19 outbreak. These policies typically included support of wages under the reduced-hour scheme, postponement of tax payments for companies, direct financial supports and grants to small enterprises and self-employed, the extension of unemployment benefits, provision of capital buffers to banks, etc. (IMF Policy Tracker, 2020). Nevertheless, there were substantial variations in the timing and specific content of these countermeasures across the states.

To briefly overview the scale of economic support provided by the government in each of the seven countries, Table 2 summarizes values of the economic support index, a composite measure reflecting income support and debt/contract relief provided by the national government to households (Hale et al., 2020). It is measured on a 0 to 100 scale, where a higher value refers to a more substantial economic assistance.

[Table 2 about here]

At the time of the survey's fieldwork, all countries provided some type of economic relief to their residents. Nevertheless, the extent of such support was conspicuously different: France and the UK ranked highest, while Denmark, Germany, and Italy ranked lowest (Hale et al., 2020). Hence, it may be possible that higher levels of economic concerns in some countries indicated people's beliefs in the insufficiency of the government's response, which will be subject to the analysis in the next waves of the survey.

Moreover, the composition of employment varies across the EU, especially in terms of informal and temporary employment. Temporary contracts provide lower levels of social protection and job security to employees, but their prevalence has increased over the last years, particularly in the Netherlands, Italy, and France. As of 2019, the share of temporary employees in the total number of employed was highest in southern European countries: Portugal (17.4%), France (13.3%), and Italy (13.1%). In contrast, it was significantly lower in northern states: the UK (3.8%), Denmark (8.3%), and Germany (9.3%). The only exception was the Netherlands, where temporary workers constituted 13.6% of all employees (Eurostat, 2020). Thus, such differences in the employment composition may be in part responsible for the cross-country dissimilarities in economic concerns.

We also observed differences in the levels of concern within individual countries. Fig.4 shows the extent of worry about the health system and a recession in Italy. We grouped regions based on the severity of the COVID-19 outbreak and distinguished the levels of anxiety across age categories. Higher intensity of the color reflects a greater extent of worry.

[Figure 4 about here]

Overall, the level of worry in the highly affected regions of the country was not higher than elsewhere in Italy, except for the youngest age group. However, economic concerns tended to be unequally distributed across the age groups. For instance, worries about the recession and small companies running out of business were higher among older individuals than younger age cohort. This pattern was similar in all countries covered by the survey.

Trust in sources of information

We asked people about the main sources of information from where they received news about COVID-19. The data show that overall, 94% of respondents closely followed the news on the situation with

COVID-19, implying a high level of public awareness. Regarding the sources of information, 86% of respondents mentioned receiving updates from the TV and 50% additionally searched for information on the Internet. Presumably, reliable information presented through the television emerged as the best channel to reach the population at large.

Next, we assessed the extent of people's trust in the information received from various sources in the context of the COVID-19 situation. The trust in the following information sources was addressed: national government, the EU, the WHO, hospitals and GPs, national news channels and newspapers, social media, relatives and friends.

Fig.5 shows mean values of trust in information from six selected sources across seven European states (measured on a Likert scale from 1-no trust at all to 5-trust very much). Higher intensity of the color reflects a higher level of trust in the information from a specific source.

[Figure 5 about here]

The data show that overall people had the highest levels of trust in information from hospitals, family doctors, and the WHO, followed by information from the national government and main national news channels. This ranking of sources by trust was similar in all countries covered by the survey, except for France, where citizens had a high level of confidence only in healthcare providers and placed relatively little trust in all other sources.

Moreover, a north-south divide could be noticed in the level of trust in information from the national government. Trust was highest in Denmark and the Netherlands (more than 70% of respondents trusted "much" or "very much"), whereas it was lowest in France (27% of respondents had a high level of trust).

Furthermore, a similar north-south gradient was observed concerning the trust in the EU: trust was highest in Denmark (45%), Germany (40%), the Netherlands (39%) and the UK (35%), whereas it was lowest in Italy (24%) and France (21%). Portugal was an exception to this case since the corresponding value here constituted 46%.

Finally, we also observed considerable regional heterogeneities in levels of trust within countries with particularly noticeable differences across individual regions in Italy, France, and Germany. Fig.6 shows people's trust in information from the national government in the context of COVID-19 in Germany and France as an example, where the higher intensity of the color indicates a greater extent of trust. While trust did not differ significantly between regions grouped with respect to the COVID-19 severity, it was heterogenous across the age groups.

[Figure 6 about here]

Although the survey asked about the level of trust in information from different sources in the context of the COVID-19 situation and not about the overall trust in institutions, these two are likely to be related. Generally, trust reflects people's perceptions of whether institutions are doing what is right. Thus, trust in the information they provide can be considered an indicator of the confidence that citizens have in these institutions (OECD, 2013).

POLICY IMPLICATIONS

The COVID-19 pandemic raised new challenges for policymakers across the EU. The imminent threat to public health at the onset of the pandemic led most governments to impose a lockdown on society. However, as the peak of the pandemic abated, the focus of attention turned to the social and economic consequences of the containment measures. Given that without acquired herd immunity the risk of a new wave of the epidemic remains high, and the production and distribution of vaccines may take 12 to 18 months (Nature Nanotechnology, 2020), governments must try to strike the right balance between effects on public health, social life and the economy when considering possible exit-strategies from the current lockdown situation.

In the absence of medical intervention, policymakers and public health officials must resort to non-medical behavioral interventions. Lifting the lockdown requires that citizens support and adhere to the policy measures that aim to contain the spread of the virus as social and economic activity gradually restarts. Given the difficulty of enforcing such regulations, future measures need to be both well-designed and well-communicated to the public. The more people are willing to comply voluntarily with the new measures, the less enforcement and supervision will be needed to achieve high compliance. For this, people's perceptions and attitudes need to be factored in at the policy-design and implementation stages.

Our survey sought to capture the public sentiment toward measures previously taken by policymakers to contain COVID-19 and addressed people's support for policies, worries about the consequences of COVID-19, and trust in different sources of information. The first insights obtained from the data showed that containment and mitigating policies undertaken by national governments in response to the initial stages of the COVID-19 pandemic were generally well-received by the population in all countries covered by the survey. Nevertheless, the extent of approval varied across states and specific policy measures.

Several lessons can be drawn for the design and implementation of policies for the prolongation or gradual removal of lockdown restrictions.

First, we observed a north-south divide in people's perceptions, worries and trust across the European countries. This finding suggests that further containment measures and lockdown exit strategies need to be balanced against the factors that worry people in each specific country. One noteworthy example is the level of importance that people in European countries attribute to the concepts of individual freedom and privacy. Using mobile data for tracking COVID-19 cases and their contacts may be a controversial decision to take even though it is believed by many experts to be a useful tool to manage the COVID-19 outbreak. The effectiveness of this policy critically depends on a sufficient level of adoption of the technology by the population (Hale et al., 2020). Our data suggest that this may not be achieved easily in some European countries.

A clear takeaway is that an open dialogue with society on this matter is needed. Explaining the need for and the advantages of such intrusive policies through trusted means of communication, while addressing people's concerns explicitly and being open about the risks of using such policy measures may help raise the support and compliance in society to a sufficient degree.

Another critical issue is the balance between saving lives and saving livelihoods. According to the survey, people in southern European countries are substantially more concerned about the economic aspects of the COVID-19 outbreak than people in northern European countries. Economic anxieties, if left unaddressed, may have adverse effects on the mental health and wellbeing of the population, as well as cause downward adjustments in consumption behavior, thereby exacerbating the economic situation in a country if the recession indeed happens.

Second, we found considerable heterogeneities in people's approval of policies within individual countries. This tendency was particularly noticeable in France and Italy. One possible determinant of regional differences in public support could be the extent of the devolution of decision-making in the country. On the one hand, devolution could enable regional or local authorities to make better decisions due to their better awareness of region-specific circumstances. On the other hand, it could harm the coordination of policy responses between the central and regional authorities within individual countries. Thus, it is crucial to understand the determinants of such differences and address them to secure public support of future policies and ensure high compliance with government measures.

Furthermore, our results showed that the burden of stress tended to be unequally distributed across and within countries. Even in case of households that were not directly hit by COVID-19, the pandemic may have acted as a stressor causing health and economic anxieties. Such worries may be detrimental to individual mental health and wellbeing, and they may become further exacerbated by the imposition of self-isolation policies. Thus, it may be reasonable to consider an asymmetric approach to the design of exit strategies taking region-specific levels of support and worry into account. This includes the identification of vulnerable categories of the population not only in terms of health risks but also with respect to social and economic activities and addressing their concerns satisfactorily.

Third, during a pandemic, public trust in the government and the information it provides is of paramount importance. To expect high compliance over extended periods of time, policymakers need to adopt effective strategies and means of communication whereby securing a sufficient level of trust and confidence from the society. As our results suggest, some countries were more successful in this respect than others.

Society needs to be well-informed about the dilemmas faced by policymakers, and for this, the communication between the government and the citizens must be clear and transparent. The data showed that 94% of respondents closely followed the news on the situation with COVID-19 mainly using television to keep themselves updated. Thus, television emerged as the best channel to reach the population at large, suggesting that presenting reliable information through this means is an effective strategy to follow.

Nevertheless, given that the data show regional and age-related heterogeneities in trust and policy support, it may be worth tailoring messages and means of communication to specific groups of the society. For example, cooperation with public figures and well-known experts can be used to deliver government and public health messages in a simple language, or local voices could be used to amplify such messages in individual regions of the country.

Overall, information provision, public education and effective communication strategies should be among the key guidelines for policymakers when implementing exit strategies and designing future containment measures so that these policies have public support and high compliance.

Additional waves of the survey are scheduled in June and August 2020. This will allow us to investigate in more detail how the population copes with the health, social and economic consequences of the COVID-19 pandemic as the situation evolves.

BIBLIOGRAPHY

Betsch, C., *et al.* 2020. COVID-19 Snapshot MOnitoring (COSMO): monitoring knowledge, risk perceptions, preventive behaviours, and public trust in the current coronavirus outbreak. *Psych Archives*, 1-6.

Breton, T., 2020. After intense discussions, I welcome that DE and FR now allow for export of #COVID19 protective equipment. I will continue to follow supply to IT closely. No single EU country can win this battle alone. #cooperation #solidarity. *Twitter* [online]. Available from: https://twitter.com/thierrybreton/status/1239159582821408770?lang=en [Accessed June 11, 2020].

Deutsche Welle, 2020. EU push for coronavirus contact tracing suffers setback. [online]. Available from: https://www.dw.com/en/eu-push-for-coronavirus-contact-tracing-suffers-setback/a-53219372 [Accessed June 12, 2020].

European Commission, 2016. Temporary Reintroduction of Border Control. [online]. Available from: https://ec.europa.eu/home-affairs/what-we-do/policies/borders-and-visas/schengen/reintroduction-border-control_en [Accessed June 11, 2020].

Eurostat, 2020. Temporary employees as percentage of the total number of employees. [online]. Available from: https://ec.europa.eu/eurostat/databrowser/view/tesem110/default/table?lang=en [Accessed June 11, 2020].

Financial Times, 2020. Lockdown fatigue hits as Europe enforces coronavirus restrictions. [online]. Available from: https://www.ft.com/content/eb81dc96-c2bd-4c1c-ac8a-bf88653b0b8d [Accessed April 1, 2020].

Flaxman, S., *et al.* 2020. Estimating the number of infections and the impact of non-pharmaceutical interventions on COVID-19 in European countries: technical description update. *arXiv preprint arXiv:2004.11342*.

Graeden, E., Carlson, C. and Katz, R. 2020. Answering the right questions for policymakers on COVID-19. *The Lancet Global Health,* 8(6), e768-e769.

Hale, T., *et al.*, 2020. Oxford COVID-19 government response tracker (OxCGRT). *Blavatnik School of Government*, [online]. Available from: www.bsg.ox.ac.uk/covidtracker.

Ienca, M. and Vayena, E. 2020. On the responsible use of digital data to tackle the COVID-19 pandemic. *Nature medicine,* 26(4), 463-464.

IMF Policy Tracker, 2020. [online]. Available from: https://www.imf.org/en/Topics/imf-and-covid19/Policy-Responses-to-COVID-19 [Accessed April 24, 2020].

Jarman, H., *et al*. 2020. How will governments know when to lift and impose restrictions. *Health System Response Monitor–Cross-Country Analysis. WHO, European Commission, European Observatory on Health Systems and Policies*, 7.

Nature Nanotechnology, 2020. The race against COVID-19. [online], 15. [Accessed 2020].

OECD, 2013. *Trust in government, policy effectiveness and the governance agenda*.

OECD, 2020. Tracking and tracing COVID: Protecting privacy and data while using apps and biometrics. [online]. Available from: https://www.oecd.org/coronavirus/policy-responses/tracking-and-tracing-covid-protecting-privacy-and-data-while-using-apps-and-biometrics/ [Accessed May 6, 2020].

Pollina, E. and Busvine, D. 2020. European mobile operators share data for coronavirus fight. *Reuters https://www. reuters. com/article/us-health-coronavirus-europe-telecoms/european-mobile-operators-share-data-for-coronavirus-fight-idUSKBN2152C2*.

Roser, M., *et al*. 2020. Coronavirus pandemic (COVID-19). *Our world in data*.

The Economist, 2020. Do low-trust societies do better in a pandemic? . [online]. Available from: https://www.economist.com/europe/2020/05/02/do-low-trust-societies-do-better-in-a-pandemic [Accessed May 7, 2020].

The NewYork Times, 2020. Reopenings Mark a New Phase: Global' Trial-and-Error' Played Out in Lives. [online], (May 2020). [Accessed May 8, 2020].

Tsang, A. 2020. EU seeks solidarity as nations restrict medical exports. *The New York*.

FIGURES

Fig.1. Mean support of government policies

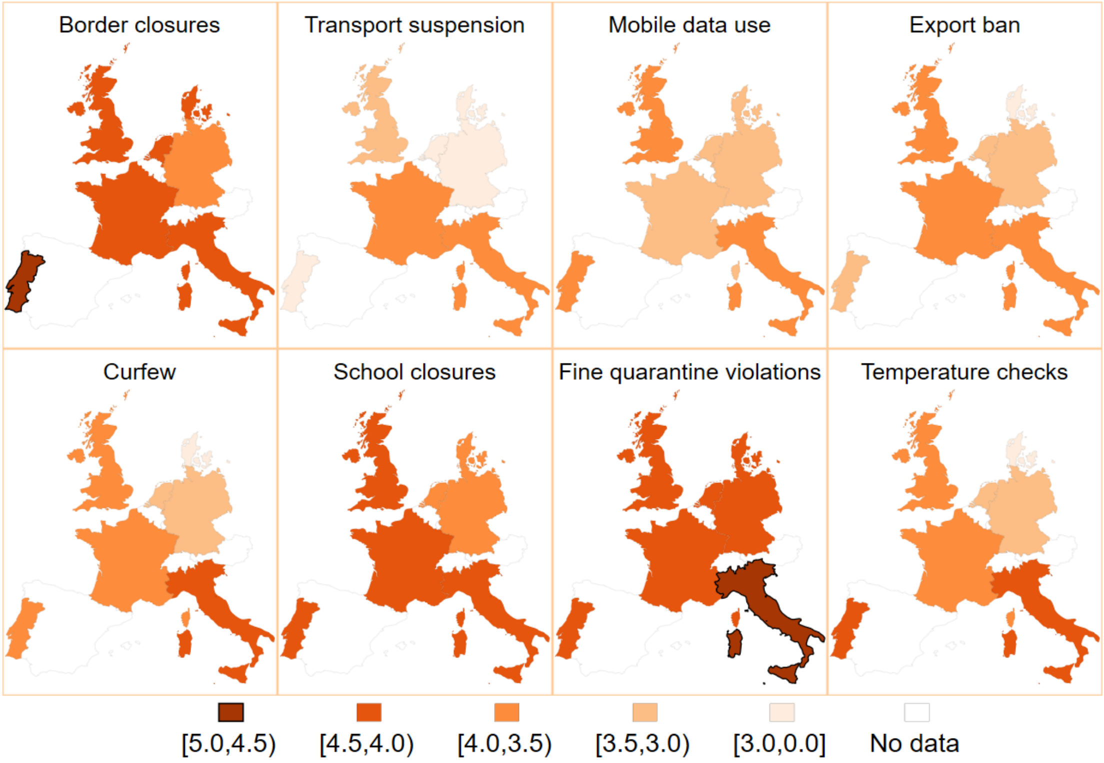


Fig.2. Heterogeneity of public attitudes within countries by region and age category in France and Italy

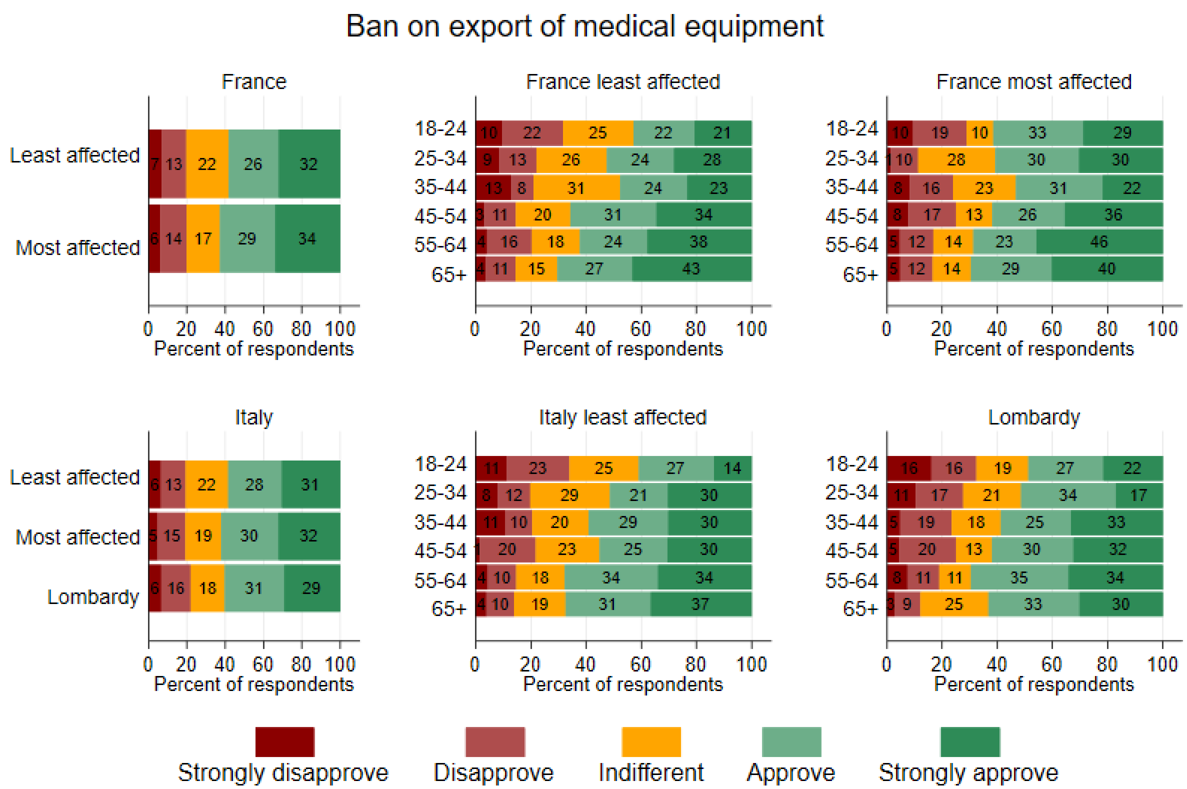


Fig.3. The proportion of respondents who worry "quite a bit" or "a lot"

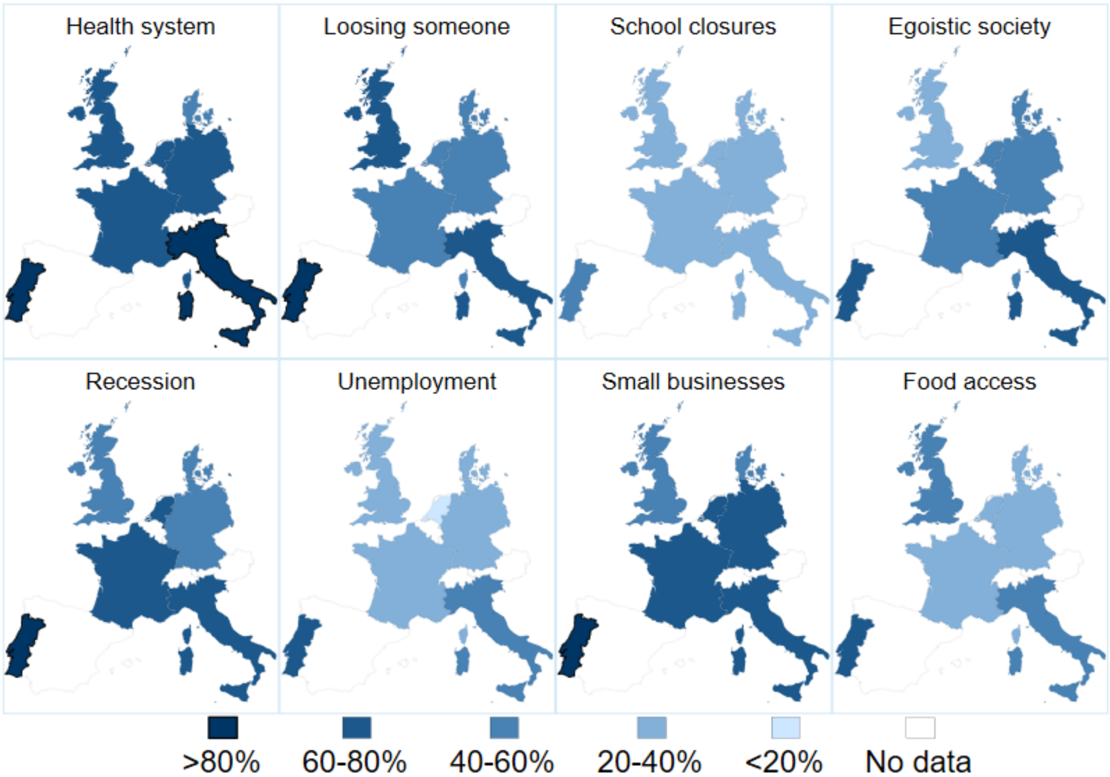


Fig.4. Heterogeneity of people's worries in Italy by region and age category

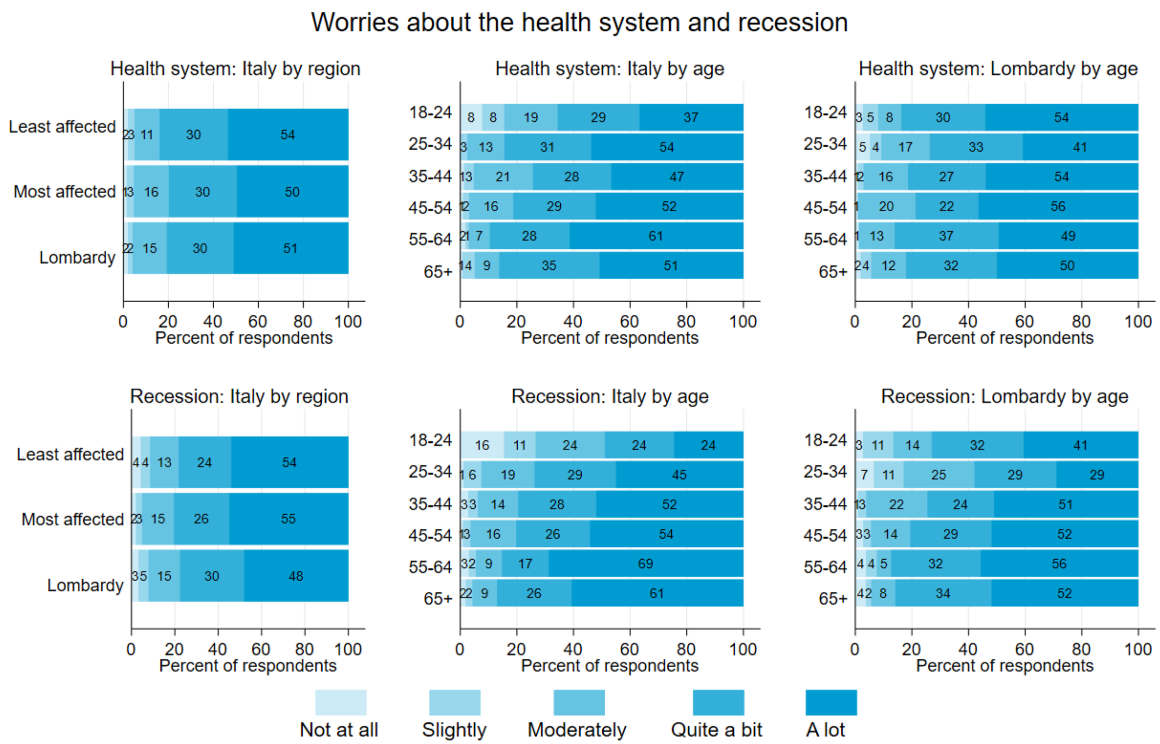


Fig.5. Mean trust in information sources in the context of COVID-19 situation

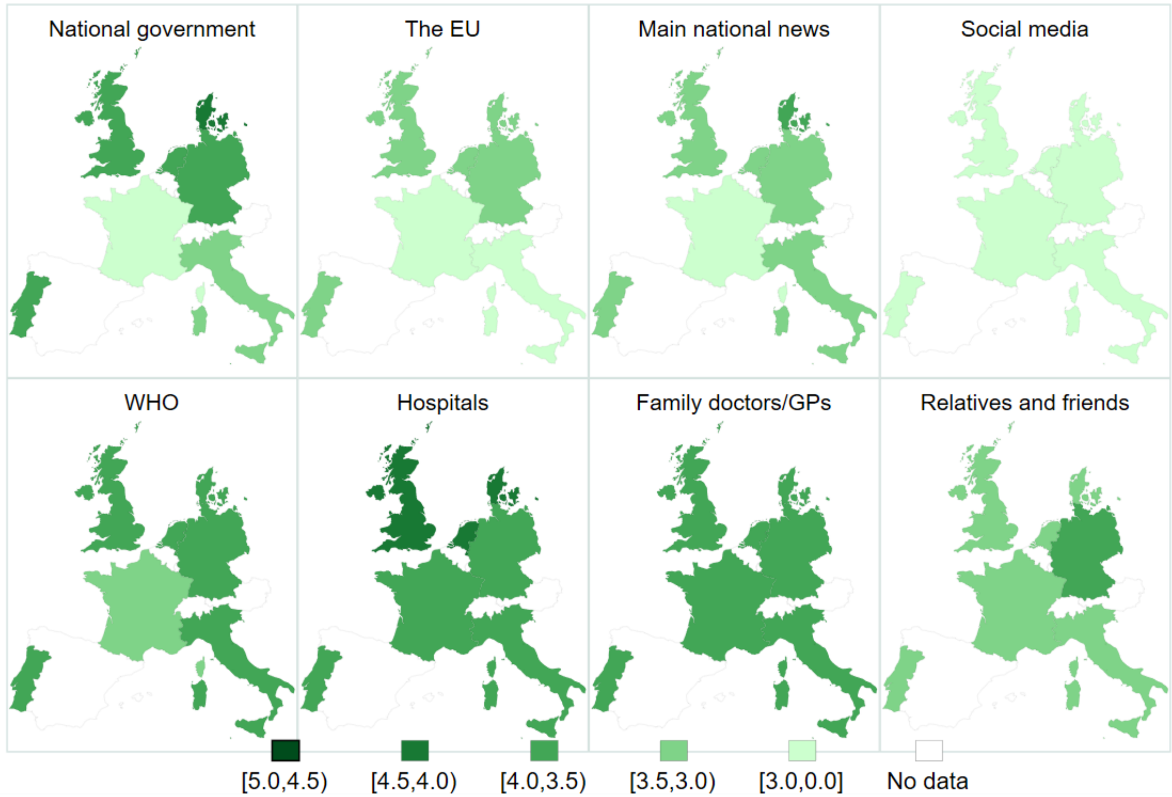
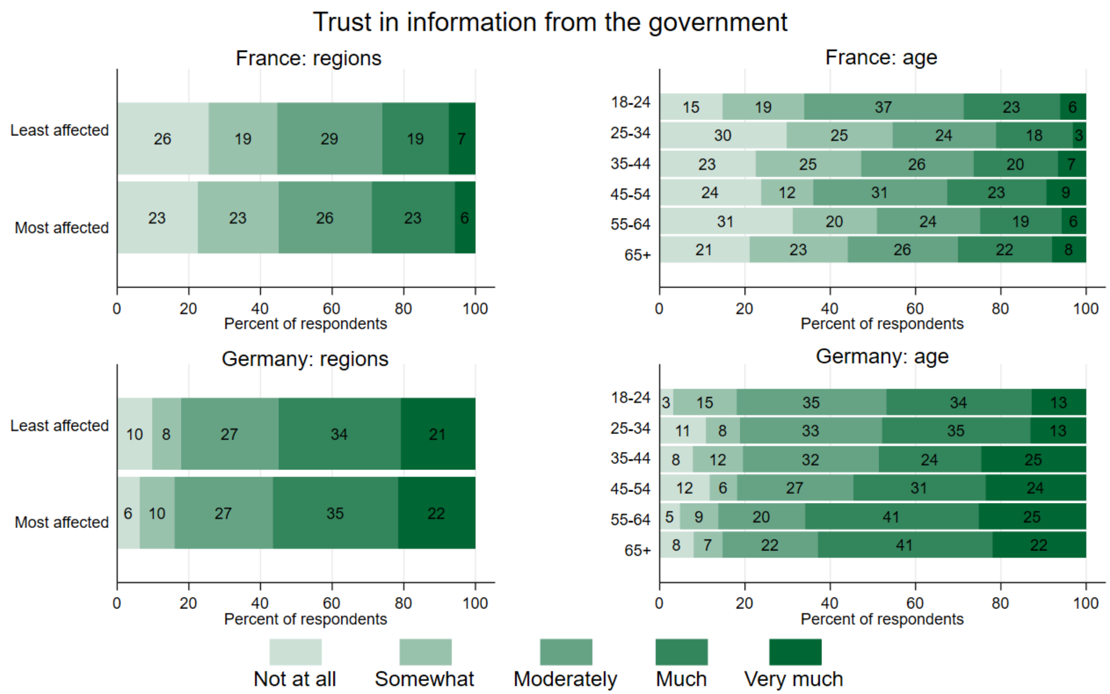


Fig.6. Heterogeneity of levels of trust in information from the national government



TABLES

Table 1. Total confirmed COVID-19 cases and deaths (per 1 million people) and government response stringency index

Country		Date			
		March 12, 2020	April 12, 2020	May 12, 2020	June 12, 2020*
Denmark	Cases	89	1 035	1 815	2 078
	Deaths	0	45	91	102
	Stringency index	37.96	72.22	65.74	62.96
France	Cases	35	1 437	2 138	2 383
	Deaths	0.74	212	408	450
	Stringency index	28.7	90.74	76.85	60.19
Germany	Cases	19	1 438	2 035	2 216
	Deaths	0.04	32	90	105
	Stringency index	32.87	73.15	64.35	50
Italy	Cases	206	2 519	3 636	3 906
	Deaths	14	322	508	565
	Stringency index	85.19	93.52	62.96	48.15
Netherlands	Cases	29	1 425	2 497	2 816
	Deaths	0.29	154	318	353
	Stringency index	41.67	79.63	68.52	62.96
Portugal	Cases	6	1 568	2 715	3 522
	Deaths	0	46	112	148
	Stringency index	32.41	87.96	75	71.3
United Kingdom	Cases	7	1 164	3 286	4 293
	Deaths	0.1	171	472	608
	Stringency index	11.11	75.93	75.93	70.37

*or closest available date

Table 2. Economic support index

Country	Date			
	March 12, 2020	April 12, 2020	May 12, 2020	June 12, 2020*
Denmark	37.5	37.5	87.5	87.5
France	0	100	100	75
Germany	0	37.5	87.5	62.5
Italy	0	50	50	75
Netherlands	0	62.5	62.5	62.5
Portugal	25	75	75	50
United Kingdom	0	100	100	100

*or closest available date

Chapter 3.2

Risk communication during COVID-19: A descriptive study on familiarity with, adherence to and trust in the WHO preventive measures[26]

ABSTRACT

Risk communication is a key component of public health interventions during an outbreak. As the coronavirus pandemic unfolded in late 2019, the World Health Organization (WHO) was at the forefront in the development of risk communication strategies. The WHO introduced a range of activities with the purpose of enabling the public to avail verified and timely information on COVID-19 prevention behaviors. Given the various WHO activities to protect the public health during COVID-19, it is important to investigate the extent of familiarity and uptake of the WHO recommendations among the public during the first wave of the pandemic. To do this, we conducted a large-scale Pan-European survey covering around 7500 individuals that are representative of populations from seven European countries, collected online during April 2-April 15, 2020. We use descriptive statistics including proportions and correlations and graphical representations such as bar charts to analyze and display the data. Our findings suggest that information from the WHO in the context of COVID-19 is well trusted and acted upon by the public. Overall familiarity and adherence were quite high in most countries. Adherence was higher for social distancing recommendations compared to hygiene measures. Familiarity and adherence were higher among older, female, and highly educated respondents. However, country level heterogeneities were observed in the level of trust in information from the WHO, with countries severely affected by the pandemic reporting lower levels of trust. Our findings call for efforts from health authorities to get regular feedback from the public on their familiarity and compliance with recommendations for preventive measures at all stages of the pandemic, to further develop and adapt risk communication as the pandemic evolves.

[26] The paper constituting this chapter was published in *Plos One* journal. Author list as follows: Nirosha Elsem Varghese, Iryna Sabat, Sebastian Neuman-Böhme, Jonas Schreyögg, Tom Stargardt, Aleksandra Torbica, Job van Exel, Pedro Pita Barros, Werner Brouwer

INTRODUCTION

Risk communication is key to improving familiarity with and adherence to preventive measures, in normal times but also particularly during health emergencies. Failure to communicate the right message effectively can result in loss of trust, damage to the economy and loss of lives (World Health Organization, 2020e). For risk communication to be effective, risk messages have to be shared with the public in an openly and timely manner, so as to reduce the knowledge gap and to convince the public to adjust their behavior during a crisis (Zhang et al., 2020). In addition to disseminating recommendations that are easy for the public to understand and comply with, trust in the source of the message is important for an effective risk communication (Slovic, 1993; World Health Organization, 2020e).

The World Health Organization (WHO) has been in the frontline in its operations to contain and mitigate the spread of the COVID-19 pandemic. The WHO is a key player in disseminating up to date information and recommendations on COVID-19 preventive behaviors to the public (Vaezi and Javanmard, 2020). With a physical presence in 149 countries, these recommendations are also adapted to national and local considerations, thereby setting the WHO protocol as a foundation for further containment strategies at various levels of government (World Health Organization, 2007; World Health Organization, 2019b).

As the coronavirus pandemic unfolded in late 2019, the WHO was quick to realize the need for a tailored risk communication strategy. The WHO Information Network for Epidemics (or EPI-WIN) was introduced when COVID-19 was declared a public emergency of international concern on 30th January 2020 (World Health Organization, 2019a). EPI-WIN provides customized information and guidance to specific target groups in addition to fighting the 'infodemic'. For example, this involved increasing the public awareness on preventive measures against COVID-19 through easy to understand behavioral messages using infographics and videos on the WHO website. EPI-WIN also guides national governments in risk communication and community engagement according to the transmission scenario of each country with the purpose of developing, implementing, and monitoring a communication plan that can help protect the public health during the health crisis (World Health Organization, 2020a). Another such WHO and national government collaboration in risk communication is the Global Outbreak Alert and Response Network (GOARN), a network of 250 technical institutes across the globe that has been actively involved in co-creating and co-implementing risk communication messages so as to adapt to the local context (World Health Organization, 2020b).

Additionally, the WHO undertook a range of other innovative steps to improve risk communication during this pandemic. They teamed up with social media companies and Google to ensure that any search queries related to COVID-19 directs the user to the WHO pages (Zarocostas, 2020). The WHO introduced an online training course on COVID-19 and collaborated with celebrities on the safe hands challenge to demonstrate hand hygiene on social media (World Health Organization, 2020f). Given all the actions undertaken by the WHO to promote public awareness on COVID-19, it seems important to investigate the familiarity of the public with the WHO recommended preventive measures, whether familiarity translates into adherence to these measures, and the role of trust in the information in this relationship.

MATERIALS AND METHODS

We use individual level data covering 7000 respondent's representative of the adult population (aged 18 and above) in seven European countries: Denmark, France, Germany, Italy, Portugal, the Netherlands, and the UK. The online survey was conducted during April 2-15, 2020 by the market research company Dynata (https://www.dynata.com). To maximize reach and capacity, the respondents were recruited using variety of contact methods (such as websites, emails, social media influencers, TV ads, loyalty partnerships and so on) which builds into a combined panel which is more representative of the offline population. Project details are not included in the invite to reduce self-selection bias. The questionnaire was initially developed in English by the authors of the study and was then translated and adapted to country specific context by native speakers. The questionnaire was first administered as a pilot to collect 10% of the sample which was included to the final sample. Potential participants go through an initial eligibility check using personal information (if available) and other screening questions. Upon informed consent, data was collected from 1000 respondents' in each country representative of the national population in terms of region, age, gender, and education. Representativeness of the sample was achieved by using quotas for the demographic characteristics based on the national census statistics. Quality checks were carried out on the final sample to eliminate and replace any speeders (below one-third of median time duration taken to complete the survey), slow respondents (above 95^{th} percentile of time duration) and bad responses (answers with no logical consistency or straight liners). Upon completion of the survey, panelists received incentives which they could redeem for a range of gifts, charitable contribution or other services.

As part of a larger survey, respondents were asked about their familiarity with the preventive measures recommended by the WHO, their adherence to these measures, and their trust in the information from the WHO. Socio-demographic variables used in our study include age, gender and level of education (low, medium and high). Level of education was defined based on country specific education system and was provided to us directly by Dynata (See S1 Table). We also collect information on household composition and consider a household to be vulnerable if respondents report having children, disabled, family members with diagnosed chronic medical conditions or elderly at home. Summary statistics including percentages and Spearman's rank correlations were used to analyze the data. The statistical significance for difference in proportions between groups was tested using Chi-squared test. Additionally, we use graphical representations and simple ranking for summarizing results. To investigate the independent relationship between two variables after partialling out potential confounders, we use multivariate non-linear regressions. Statistical analyses were performed on STATA 15 (STATA Corp, College Station, TX, US).

We also use external country-specific data on the severity of the pandemic and stringency of containment measures to present these country specific differences as a potential factor influencing the relationship between familiarity, trust and adherence. Data on COVID-19 prevalence and COVID-19 attributed deaths per 1 million people are reported by the European Centre for Disease Prevention and Control (Worldometer, 2020). To assess the degree of strictness of containment policies in each country, we use the COVID-19 Government Response Stringency Index introduced by the Oxford University and measured on a range from 0-100, with higher values indicating stricter measures (Hale and Webster, 2020). Familiarity and adherence to preventive measures may depend on the pandemic situation of the country and the governmental response at the time of survey. Country-specific data on COVID-19 death (and prevalence) around the time of the survey show a north-south divide in the development of the pandemic with a higher number of COVID-19 deaths being recorded in Italy and France followed by the UK (Worldometer, 2020). The timing and stringency of the containment measures also varied across the countries. Italy, France and Portugal imposed strict containment policies compared to other countries included in the survey (Hale and Webster, 2020). Strict measures were implemented in all countries displaying a higher severity of COVID-19 pandemic except for the UK where the government response was less strict in response to the epidemiological crisis faced by the country (Hale and Webster, 2020; Worldometer, 2020). To highlight these variations as a means to further understand the country level differences in public response to WHO recommendations, we provide a timely description of the epidemiological situation and stringency of containment policies in each country (Table 1).

[Table 1 about here]

RESULTS

Familiarity with the WHO recommendations

In the wake of the COVID-19 crisis, the WHO put forward six basic preventive measures to help contain and mitigate the spread of coronavirus. The recommendations were first released on January 10, 2020 on the WHO website, around 11 weeks prior to the release of our survey. The recommendations included timely and easy-to-understand measures such as regularly washing hands with soap for at least 20 seconds, covering nose and mouth while coughing or sneezing, keeping a social distance of at least 1 meter, avoid shaking hands, hugging or kissing when greeting others, using alcohol-based hand rub and avoid touching nose, eyes and mouth. In our survey, respondents were shown the graphic presentation of the six measures that was used for communication in their country and asked to rate their familiarity with the measures on a scale from "not at all familiar" (1) to "very familiar" (5). Respondents reporting a score of 4 (moderately familiar) or 5 (very familiar) are classified as being familiar with the WHO recommendations.

[Figure 1 here]

On average 86.3% of the respondents reported being familiar with the WHO recommended measures. Looking at country level variations (Figure 1), we see that the proportion of respondents who reported being familiar with the recommendations was the highest in Portugal (95.2%) and the lowest in the United Kingdom (81.4%; p<0.001). The other countries in the sample reveal similar levels of familiarity (84-87%) with the WHO recommendations. It is noteworthy that in the Lombardy region familiarity was also very high (91.0%), but especially also that the proportion of the population 'not at all familiar' (0.4%) or only 'slightly familiar' (0.6%) was the lowest.

Across countries, we find a higher proportion of female (88.4% for females vs. 84.0% for males; p<0.001), older (89.5% for 65+ vs. 78.1 for 18-24 yo; p<0.001) and highly educated (87.5% for high/medium vs. 83.6% for low; p<0.001) respondents reporting familiarity with the WHO recommendations. The same patterns are observed within each country as well. Finally, it should be noted that respondents may overstate their familiarity resulting in self-reporting bias. Also, we assume that the participant's response to the familiarity question reflects the same level of familiarity for each WHO recommendation (given that we show the poster before asking the question). Nevertheless, it is possible that this may not be the case and respondent's answers may be biased in a way that a higher level of familiarity with one (or more) recommendation may dominate low levels of familiarity with other recommendation or vice-versa.

Adherence to WHO recommendations on preventive behavior

In our study, we asked respondents to rate their adherence to the six preventive measures over the past four weeks using four levels: no; yes, a bit; yes, quite strongly; yes, fully. We consider respondents to adhere to the recommendations if they reported ‘yes, quite strongly’ or ‘yes, fully’ to each of the six recommendations. Overall, we see that 92.1% of the respondents reported to have adopted the WHO recommendations. Avoiding physical contact by not shaking hands, kissing or hugging when meeting others (93.6%) and keeping a social distance of 1 meter (91.5%) had an overall higher adherence rate, whereas using an alcohol-based hand rub (67.5%) and avoiding touching nose, eyes and mouth (62.4%) had the lowest rates of adherence (Figure 2).

[Figure 2 approximately here]

Comparing countries, Portugal and Italy perform best in adhering to all the WHO recommendations whereas France and Denmark perform the worst (Table 2). The difference between the top and worst adhering countries for each WHO recommendation is statistically significant ($p<0.001$).

[Table 2 approximately here]

Finally, the proportion reporting adherence is higher among female (94.0% for females vs. 90.0% for males; $p<0.001$) and older (95.1% for 65+ vs. 87.5% for 18-24 yo; $p<0.001$). If we consider 25-64 years as our reference, we find that respondents aged 18-24 are less likely to comply (87.5% vs. 91.9%; $p<0.001$) whereas respondents aged 65 and above are more likely to comply (95.1% vs. 91.9%; $p<0.001$). Respondents reporting to ‘adhere fully’ are higher among those with high/medium level of education (51.0%) compared to low (46.5%; $p<0.001$) and also among those who have family members that are vulnerable, such as elderly and those with additional comorbidities (49.7%), compared to those who indicate they do not have vulnerable family members (45.6%, $p<0.001$). Similar results are observed within countries except for levels of education which does not follow a consistent pattern in all countries.

Perception of adherence to the WHO recommendations by others

We also asked respondents if, according to them, others in the community adhered to the six WHO recommendations over the past four weeks. The possible responses are: no; yes, a bit; yes, quite strongly; yes, fully. We take a mean of the reported responses across the six recommendations to compute an overall score of respondent’s perception of adherence by others. We then consider respondents to adhere to the recommendations if they reported ‘yes, quite strongly’ or ‘yes, fully’ to each of the six recommendations.

Overall, the proportion of respondents who report that others adhere to the WHO recommendations is 81.3%, which is considerably lower than their own adherence (92.2%). This difference is highest in the UK (19% points difference; $p<0.001$) and the lowest in the Netherlands (3%; $p= 0.022$) and in France

(1%; p<0.001). Also, it should be noted that respondents could be overstating their own adherence to avoid judgement whereas adherence estimates of others could be a truer estimate of their own actual adherence.

Trust in information from the WHO

Furthermore, we asked respondents to rate their level of trust in information from the WHO in the context of COVID-19 on a scale "no trust at all" (1) to "trust very much" (5). Respondents reporting a score of 4 or 5 are classified as having trust in the information from the WHO, and those with a score of 1 or 2 as having no trust in this information. We find that on average 59.8% of the respondents from the countries included in this study trust the information on COVID-19 from the WHO, while 15.5% do not trust this information.

[Table 3 about here]

Table 3 shows the proportion of respondents in each country reporting trust or no trust in the information from the WHO and the relative ranking among the countries in terms of trust. Marked differences in trust is observed between the countries. In particular, we find that trust is highest in Denmark and Netherlands and the lowest in France with the differences between countries being statistically significant (p<0.001). Similarly, Denmark scores the lowest on distrust whereas France scores the highest followed by Italy, two countries that were the most impacted by the COVID-19 (p<0.001).

Finally, we emphasize that the level of trust in information from the WHO corresponds to COVID-19 information in general and trust in WHO recommended preventive measures may constitute only a part of this information. Therefore, we assume that the respondents have the same level of trust in all of the COVID-19 related information from the WHO and also the same level of trust in each of the WHO recommended preventive measure.

Do Familiarity and Trust Breed Adherence?

We present evidence suggesting that familiarity and trust could be driving factors for adherence. First, looking at the piecewise relationship between familiarity and adherence, we find that overall familiarity with the six WHO recommendations is significantly correlated with adherence to these recommendations (Table 4), especially for hygiene measures (R1, R2) and avoiding physical contact (R3, R4).

[Table 4 about here]

Trust could also be a facilitator for adherence (OECD, 2013). In our study, we see that distrust was lower among those who adhered (14.5%) compared to those who did not adhere to the WHO recommendations (29.3%). Table 5 presents the correlation between level of trust and level of adherence for each WHO recommended preventive measure. Trust in COVID-19 information from the WHO is positively correlated with adherence. The correlation is stronger for hygiene measures (R1, R2) and avoiding physical contact (R3, R4).

To further examine if familiarity drives adherence independent of trust, we run ordered logistic regression models of the following econometric specification.

$$Y_i = \alpha_0 + F_i \alpha_1 + T_i \alpha_2 + \varepsilon_i$$

where Y_i denotes adherence to each preventive measure, F_i denotes overall familiarity, T_i denotes trust in WHO information on COVID-19 and ε_i corresponds to the random error term. Robust standard errors were used in all regressions.

Figure 3 presents a graphical description of the Odds Ratio (OR) and 95% Confidence Interval (CI) for the relationship between overall familiarity and adherence while simultaneously controlling for trust in information from the WHO. Our findings strengthen the evidence from simple correlations. Controlling for trust, being very familiar compared to not at all familiar increases the likelihood of adherence with statistical significance retained only for hygiene measures ($p<0.001$ for R1 and R2) and avoiding physical contact ($p<0.001$ for R3 and R4). Trust in information from the WHO is also positively related to adherence ($p<0.001$ for R1-R6).

[Figure 3 approximately here]

As a robustness check, we also recode our familiarity variable to a binary measure (familiar or not). Respondents choosing any response other than “very familiar” are considered to be not familiar with WHO recommendations or otherwise. We repeat our analysis with this categorization to ensure enough observations in each group to carry out a robust analysis. Controlling for trust, being familiar compared to not familiar increases the likelihood of adherence for all preventive measures ($p<0.001$ for R1-R6). Trust in information from the WHO is also positively related to adherence ($p<0.001$ for R1-R6). Results are robust when controlling for respondent's perception of others adherence of WHO recommendations, which has been shown to have a sizeable effect on one's own adherence (Tunçgenç et al., 2021). The relationship also holds for all countries separately (results are available from the authors on request). The results indicate that both overall familiarity and trust is positively related to adherence independent of one another. We also notice that the effect size for familiarity is larger than that for trust (almost 2 times larger for R1-R4), highlighting a potential larger influence of the familiarity variable in driving adherence.

Overall, at first sight this would mean familiarity implies adherence and trust is a catalyst for this relationship. However, the relationship between familiarity, adherence and trust is not so direct. Factors such as the severity of the COVID-19 crisis and other perceived worries could be influencing each of these factors independently and together. For instance, respondents from Italy and Portugal reported the highest levels of familiarity and adherence, but at the same time showed diverging profiles on COVID-19 attributed death counts (Table 1) and trust in WHO information during the period of our study. In Italy, adherence to physical distancing recommendations is as high as in Portugal, although Italy reports lower trust compared to Portugal. That this high adherence in Italy is narrowed to only physical distancing measures could be attributed to the necessity of adherence given the severity of the pandemic (Table 1). However, Portugal still tops adherence in all measures (including hygiene), which could be facilitated by the high levels of trust in information. This is suggestive of the ability of the WHO to act without any coercion when there are high levels of trust, especially when adherence corresponds to recommendations that are difficult to enforce socially or legally such as hand hygiene.

DISCUSSION

The ongoing threat to global health from COVID-19 poses critical challenges to governments, medical communities, health organizations, businesses and the public in responding to the evolving pandemic. With an abundance of misinformation on the disease, governments and health organizations need to be meticulous in disseminating up to date and evidence-based information to the public. The guidelines and recommended preventive behaviors as put forward by WHO and other national level public health agencies is of immense importance given the increasing prevalence of cases and emergence of new variants. The WHO mainly recommends hygiene and physical contact precautions to the public given that coronavirus is mainly transmitted through droplets and aerosols. As the COVID-19 vaccination process has just begun in many countries and we still have a long way to go before achieving herd immunity on a global scale, the importance of non-pharmaceutical interventions such as social distancing, use of protective equipment such as face masks and other hygiene behaviors in containing the coronavirus needs to be highlighted (World Health Organization, 2020d). Given that the pandemic is still ongoing, we used data from a pan European survey collected in April 2020 to evaluate the efficiency and effectiveness of the risk communication strategies put in place by the WHO so far during the pandemic. Following are some insights and attention points on risk communication as learned from our findings.

First, our survey results suggest that overall familiarity and adherence with the recommendations is quite high in most countries in Europe. This indicates both the effectiveness of the WHO risk communication strategy and the interest among the public to seek and follow better practices. Countries reporting high levels of familiarity (Portugal and Italy) were also the top adherers. Similarly, countries reporting lower levels of familiarity (UK, Netherlands and Germany) performed worse on adherence. Although there could be other factors influencing this relationship, our results suggest that increasing familiarity with preventive measures could lead to higher levels of adherence among the public, and hence is an effective way to help contain and mitigate the spread of infectious diseases.

Second, we observe considerable heterogeneity in adherence to the different recommendations. Overall, people complied better with avoiding physical contact, but less with hand hygiene and avoiding touching eyes, nose or mouth. Both sets of recommendations involve behavioral modifications with the exception that during the first stage of the pandemic, social distancing was legally and socially enforced, which could be one explanation for the higher adherence rates. Literature also shows that non-adherence to be high especially when recommendations involve behavioral modifications (Walters-Salas, 2012). Given that countries are moving back and forth between lockdown and exit strategies and the simultaneous warnings from the WHO on further waves of coronavirus transmission (or emergence of new variants of coronavirus) (World Health Organization, 2021), it is crucial that the public keep up with social distancing measures even when not legally enforced.

Although social distancing measures has been mostly recommended given the nature of coronavirus transmission, hand washing is also important given that there could be indirect transmission via infected surfaces (World Health Organization, 2020d). However, hand washing has a lower adherence rate globally given the complex interaction of many behavioral aspects that drives compliance to hand hygiene (Jumaa, 2005). Hence, there is an increased need to put higher emphasis on improving adherence to hand hygiene and, most importantly, designing policies to ensure that adherence to social distancing does not fade off without legal enforcement over time.

Third, our analysis suggests evidence for heterogeneities in adherence based on socio-demographic characteristics of the respondents. Particularly, we find older, female, and higher educated respondents to report higher levels of familiarity and adherence. Additionally, we also find those respondents with vulnerable household members to have higher levels of adherence. Therefore, steps should be taken to increase awareness among the groups that are less likely to be familiar with or adhere to the preventive measures, in particular the young, males, less educated and households with non-vulnerable family members, since they also play a role in transmitting the virus.

Older people, who are more vulnerable to COVID-19, report higher levels of familiarity and adherence. Possibly they seek more information, or risk communication has been tailored to them better. But, it is equally important to increase awareness among younger people about the risks of not adhering to recommendations, because even if they themselves are less vulnerable, as potential carriers of the virus they may infect others who are. Similarly, households that do not have a vulnerable family member might be less worried about getting infected and hence show poor adherence. Literature shows higher levels of adherence among women in general, attributing this to several factors including early cognitive maturation, capacity for self-care and the stronger perceived need to comply to social expectations (Boucquemont et al., 2019). Higher educated respondents might have higher levels of health literacy, which is required to critically assess the information provided in relation to their behaviors (Abel and McQueen, 2020). Thus, we might conclude that risk messages may not fit to all groups alike and, therefore, need to be customized to the specific risks and concerns in that group.

Fourth, trust in information from the WHO could influence adherence to its recommendations. Trust is an overlooked aspect in crisis management (World Health Organization, 2009). Public health organizations need to be more transparent and receptive in their communication to gain the trust of citizens. Most importantly, if the severity of the pandemic in a country is high, this could imply that trust levels are already low, and people are more worried. Hence strategies to improve both adherence and trust should take into consideration the severity of the pandemic in the country and the level of worries among the population. Finally, low trust in authoritative bodies could also be associated with low interpersonal trust (perception of the adherence of others to WHO recommendations) in the society as a whole (Borum, 2010), resulting in covid induced worries, social fear or acts of self-interest such as panic buying and stock piling, which makes crisis management more difficult. Having a perception that others in the community do not adhere to the WHO recommendations could also reduce one's own level of adherence (Fetzer et al., 2020). Thus, during these hard times, risk communication should not miss out on messages that could improve the public's trust in their community members and organizations that provide credible information.

Our findings are based on data collected during the first wave of the pandemic. This was the initial and escalating phase of COVID-19, characterized by panic and worries attributed to limited knowledge of COVID-19. Lessons from the first wave of COVID-19 may help all stake holders to be better prepared for further waves of the pandemic, especially with regards to risk communication which is a key element of outbreak control strategy. Identification of subgroup populations that are less likely to trust, to be familiar or adhere signals the need to tailor risk communication strategies that is different from what was implemented in the first wave. However, measures that were effective in the first wave may not necessarily prove to be always successful in future. Pandemic fatigue, misalignment of individual and collective benefits and/or shifts in government response may bring about changes to public adherence over time. Our results from the initial wave of COVID-19 also provides the ground for building a longitudinal dataset to understand changes in public adherence over time and a robust evidence for factors driving these changes, which will in turn lead to "precision policy planning and evaluation" (Wu et al., 2021).

Finally, we address some limitations of our study. The online nature of our survey may generate a sampling bias in our results. Particularly, all respondents may not have an equal chance of being invited to participate in the survey depending on their level of online engagement. Access to internet may be further correlated with age and socio-economic status. To reduce this bias, Dynata employs a multi-source panel recruitment model which uses a wide range of sources to reach respondents, thereby ensuring a diverse population to participate in the survey. For example, Dynata collaborates with loyalty programs which allow to survey people who would normally never participate in online research. They then bring all these sources together on one integrated platform, thereby ensuring total control over the exact blend composition of the respondents. Another limitation of the survey could be the nature of the dissemination of the WHO recommendations. If the WHO risk communication activities are primarily web-based, this might bias our results as the offline population may not have access to this information. However, the WHO has been active in implementing strategies to ensure that the message reaches everyone. For instance, the WHO has collaborated with the International Telecommunication Union (ITU) to send texts to public on preventive measures against COVID-19 (World Health Organization, 2020c). Moreover, the WHO has also set in place the Risk Communication and Community Engagement (RCCE) system to provide guidance for countries in implementing effective strategies during COVID-19 such as public communication, community engagement and capacity building, all of which may further help to engage with masses that are otherwise difficult to reach (World Health Organization, 2020e).

CONCLUSION

Overall, we find that information from WHO in the context of COVID-19 is well trusted and acted upon by the public. However, our results suggest the need to strengthen efforts to reach the less vulnerable parts of the population in information campaigns, and to take the worries of the public into account in the design and dissemination of risk communication strategies. Furthermore, our findings call for efforts to get regular feedback from the public on their familiarity with the most recent recommendations and their support for policy measures that increase compliance with these recommendations. As both the pandemic and the recommendations evolve, risk communication needs to be tailored to the different groups in society in order to be more effective.

BIBLIOGRAPHY

Abel, T. and McQueen, D. 2020. Critical health literacy and the COVID-19 crisis. *Health promotion international*.

Borum, R. 2010. The science of interpersonal trust.

Boucquemont, J., *et al.* 2019. Gender differences in medication adherence among adolescent and young adult kidney transplant recipients. *Transplantation*, 103(4), 798-806.

Fetzer, T. R., *et al.*, 2020. *Global Behaviors and Perceptions at the Onset of the COVID-19 Pandemic.* National Bureau of Economic Research, 0898-2937.

Hale, T. and Webster, S. 2020. Oxford COVID-19 government response tracker.

Jumaa, P. 2005. Hand hygiene: simple and complex. *International Journal of Infectious Diseases,* 9(1), 3-14.

OECD 2013. Trust in government, policy effectiveness and the governance agenda. *Government at a Glance*, 19-37.

Roser, M., *et al.* 2020. Coronavirus pandemic (COVID-19). *Our world in data*.

Slovic, P. 1993. Perceived risk, trust, and democracy. *Risk analysis*, 13(6), 675-682.

Tunçgenç, B., *et al.* 2021. Social influence matters: We follow pandemic guidelines most when our close circle does. *British Journal of Psychology*.

Vaezi, A. and Javanmard, S. H. 2020. Infodemic and risk communication in the era of CoV-19. *Advanced Biomedical Research*, 9.

Walters-Salas, E. b. T. 2012. The Challenge of Patient Adherence. *Bariatric Nursing and Surgical Patient Care*, 7(4), 186-186.

World Health Organization, 2007. *WHO interim protocol: rapid operations to contain the initial emergence of pandemic influenza* [online]. Available from: https://www.who.int/influenza/resources/documents/RapidContProtOct15.pdf.

World Health Organization, 2009. *Pandemic influenza preparedness and response: a WHO guidance document* [online]. Geneva: World Health Organization. Available from: https://www.who.int/influenza/resources/documents/pandemic_guidance_04_2009/en/.

World Health Organization, 2019a. *EPI-WIN: WHO information network for epidemics* [online]. Available from: https://www.who.int/teams/risk-communication [Accessed 01-05-2020 2020].

World Health Organization, 2019b. *WHO presence in countries, territories and areas: 2019 report* [online]. World Health Organization. Available from: https://www.who.int/country-cooperation/publications/who-presence-report-2019/en/.

World Health Organization, 2020a. *Critical Preparedness, readiness and response actions for COVID-19* [online]. Available from: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/critical-preparedness-readiness-and-response-actions-for-covid-19 [Accessed May 5 2020].

World Health Organization, 2020b. Global Outbreak Alert and Response Network. [online], 2020 (May 1). Available from: https://extranet.who.int/goarn/.

World Health Organization, 2020c. *ITU-WHO Joint Statement: Unleashing information technology to defeat COVID-19* [online]. Available from: https://www.who.int/news/item/20-04-2020-itu-who-joint-statement-unleashing-information-technology-to-defeat-covid-19.

World Health Organization, 2020d. *Modes of transmission of virus causing COVID-19: implications for IPC precaution recommendations* [online]. Available from: https://www.who.int/news-room/commentaries/detail/modes-of-transmission-of-virus-causing-covid-19-implications-for-ipc-precaution-recommendations [Accessed May 5 2020].

World Health Organization, 2020e. *Risk communication and community engagement (RCCE) readiness and response to the 2019 novel coronaviruses (2019- nCoV): interim guidance, 26 January 2020* [online]. Available from: https://www.who.int/publications/i/item/risk-communication-and-community-engagement-readiness-and-initial-response-for-novel-coronaviruses.

World Health Organization, 2020f. The #Safe hands challenge,. [online]. Available from: https://www.who.int/news-room/campaigns/connecting-the-world-to-combat-coronavirus/safehands-challenge.

World Health Organization, 2021. *Weekly epidemiological update-12 January 2021* [online]. Available from: https://www.who.int/publications/m/item/weekly-epidemiological-update---12-january-2021.

Worldometer, 2020. COVID-19 CORONAVIRUS PANDEMIC. [online], 2020 (May 5). Available from: https://www.worldometers.info/coronavirus/.

Wu, J. T., *et al.* 2021. Nowcasting epidemics of novel pathogens: lessons from COVID-19. *Nature Medicine*, 27(3), 388-395.

Zarocostas, J. 2020. How to fight an infodemic. *The Lancet*, 395(10225), 676.

Zhang, L., Li, H. and Chen, K., Effective risk communication for public health emergency: Reflection on the COVID-19 (2019-nCoV) outbreak in Wuhan, China. ed. *Healthcare*, 2020, 64.

FIGURES

Figure 1: Familiarity with the WHO recommendations, by country

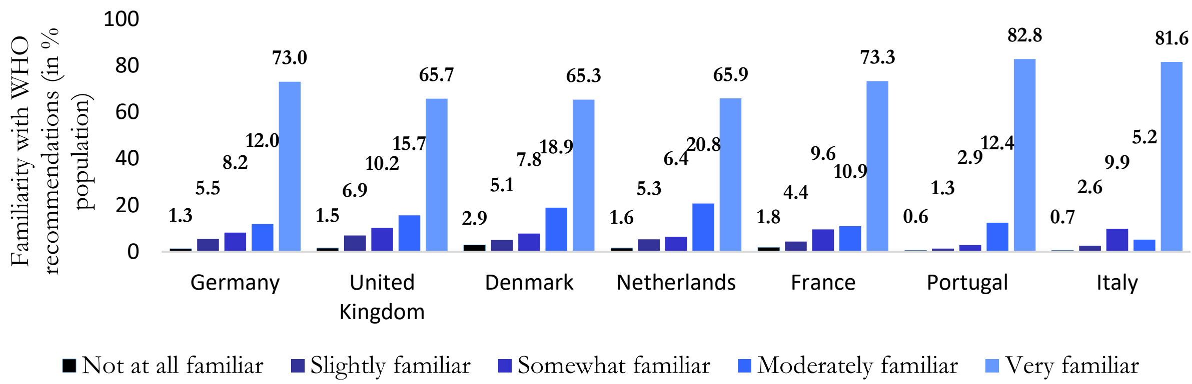
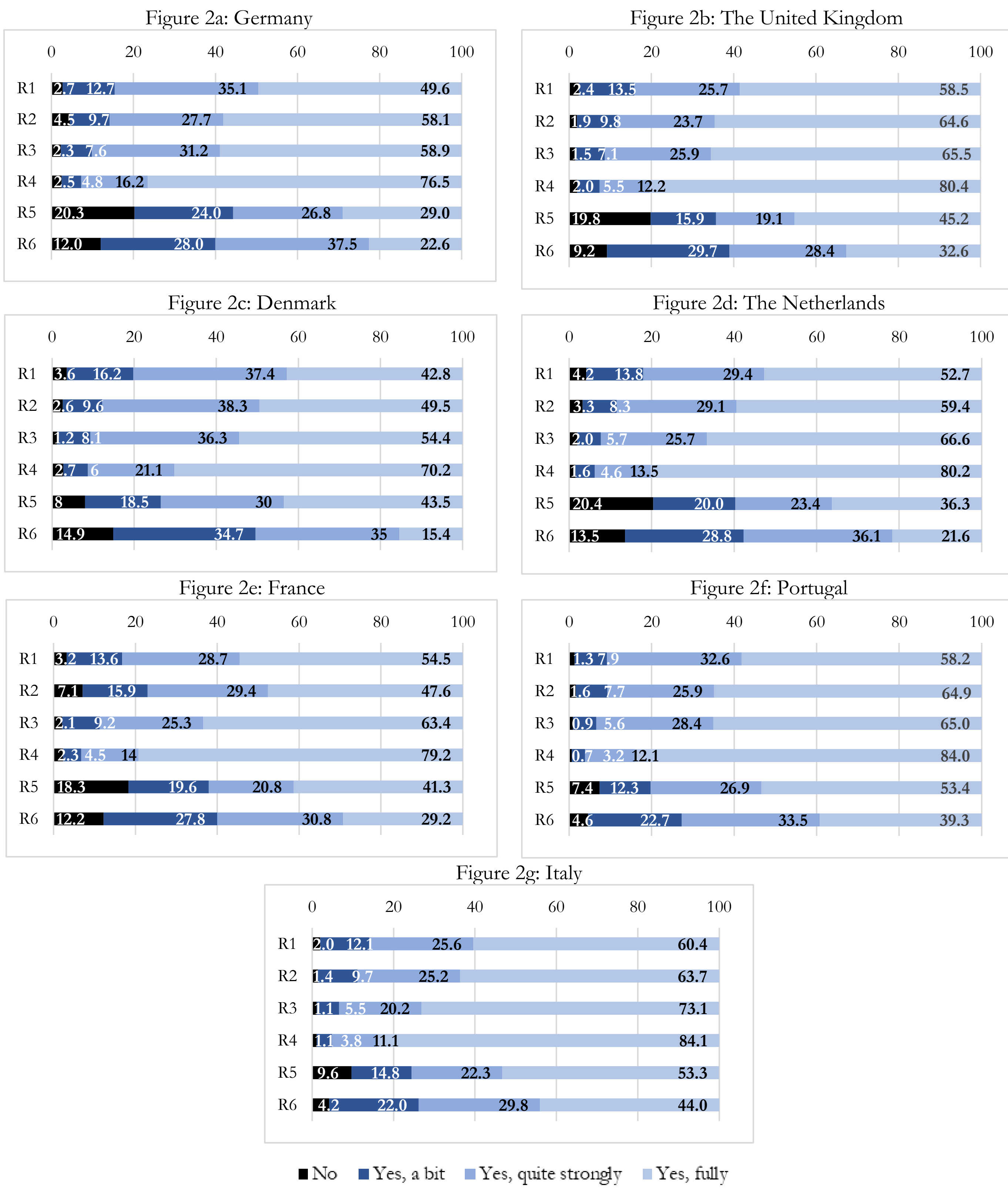
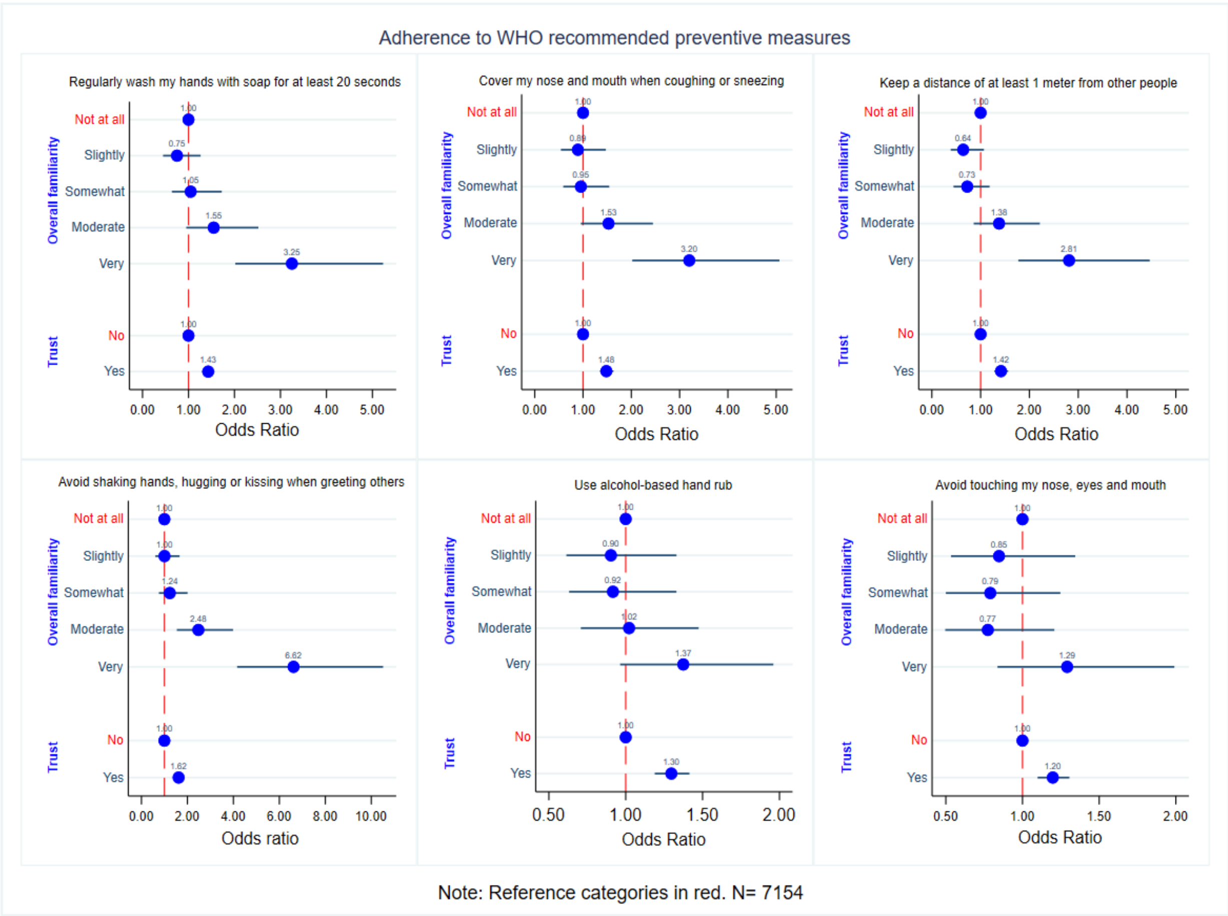


Figure 2a-2h: Adherence to WHO recommendations, by country



*Note: **R1**: Regularly wash my hands with soap for at least 20 seconds, **R2**: Cover my nose and mouth when coughing or sneezing, **R3**: Keep a distance of at least 1 meter from other people, **R4**: Avoid shaking hands, hugging or kissing when greeting others, **R5**: Use alcohol-based hand rub' and **R6**: Avoid touching my nose, eyes and mouth.*

Figure 3: Odds ratio for adherence to WHO recommended preventive measures by overall familiarity and trust in information from the WHO



TABLES

Table 1: Relative ranking of countries according to the severity of the pandemic and stringency of containment measures

	Rank		
	Cases	Deaths	Stringency index
Italy	1	1	1
France	4	2	2
UK	6	3	5
Netherlands	5	4	4
Portugal	2	5	3
Denmark	7	6	6
Germany	3	7	7

Note 1: Cases and deaths refer to total confirmed COVID-19 cases and COVID-19 attributed deaths (per 1 million people). Stringency index refers to degree of strictness of government response to COVID-19 (measured between 0-100).
Note 2: The rankings are based on data collected from the source (Hale and Webster, 2020; Roser et al., 2020) as of April 12, 2020

Table 2: Rank of WHO recommendations in the order of their relative adherence.

Rank	WHO recommendations	Top adherers	Worst adherers
1	Avoid shaking hands, hugging or kissing when greeting others.	Portugal, Italy	Denmark
2	Keep a distance of at least 1 meter from other people.	Portugal, Italy	France
3	Cover my nose and mouth when coughing or sneezing.	Portugal	France
4	Regularly wash my hands with soap for at least 20 seconds.	Portugal	Denmark
5	Use alcohol-based hand rub.	Portugal	Germany
6	Avoid touching my nose, eyes and mouth.	Portugal, Italy	Denmark

Table 3: Relative ranking of countries according to the proportion of respondents who trust and distrust information from the WHO

Country	**Trust (%)**	**Rank**
Denmark	64.7	1
Portugal	64.5	2
UK	61.6	3
Italy	60.9	4
Netherlands	59.7	5
Germany	56.8	6
France	49.9	7
Country	**Distrust (%)**	**Rank**
France	22.6	1
Italy	16.4	2
Germany	15.7	3
UK	14.1	4
Netherlands	14.0	5
Portugal	13.0	6
Denmark	12.7	7

Table 4: Correlations between level of familiarity and level of adherence

Adherence	**Familiarity**							
	Germany	United Kingdom	Denmark	Netherlands	France	Portugal	Italy	Overall
R1	0.303*	0.254*	0.255*	0.245*	0.239*	0.142*	0.218*	0.247*
R2	0.314*	0.251*	0.230*	0.258*	0.219*	0.195*	0.223*	0.246*
R3	0.275*	0.221*	0.260*	0.236*	0.308*	0.179*	0.243*	0.250*
R4	0.375*	0.276*	0.258*	0.313*	0.318*	0.261*	0.306*	0.305*
R5	0.009	0.051	0.141*	0.048	0.091*	0.105*	0.139*	0.096*
R6	0.087*	0.097*	0.110*	0.091*	0.085*	0.102*	0.093*	0.120*

*Note**:** R1-R6 corresponds to the six recommendations released by the world health organization. They are as follows. **R1**: Regularly wash my hands with soap for at least 20 seconds, **R2**: Cover my nose and mouth when coughing or sneezing, **R3**: Keep a distance of at least 1 meter from other people, **R4**: Avoid shaking hands, hugging or kissing when greeting others, **R5**: Use alcohol-based hand rub' and **R6**: Avoid touching my nose, eyes and mouth. Spearman rank correlation test is used for this analysis. *, and ** denote significance at 1 and 5 percent levels respectively.*

Table 5: Correlations between level of trust and level of adherence

Adherence	**Trust**							
	Germany	United Kingdom	Denmark	Netherlands	France	Portugal	Italy	Overall
R1	0.132*	0.113*	0.122*	0.110*	0.154*	0.041	0.140*	0.116*
R2	0.140*	0.101*	0.133*	0.076**	0.144*	0.075**	0.143*	0.123*
R3	0.126*	0.167*	0.093*	0.110*	0.126*	0.073**	0.080**	0.110*
R4	0.130*	0.198*	0.107*	0.161*	0.111*	0.099**	0.102*	0.128*
R5	0.028	0.058	0.124*	0.000	0.119*	0.046	0.130*	0.081*
R6	0.050	0.049	0.040	-0.001	0.122*	0.056	0.102*	0.061*

Note: *R1-R6 corresponds to the six recommendations released by the world health organization. They are as follows.* ***R1***: *Regularly wash my hands with soap for at least 20 seconds,* ***R2***: *Cover my nose and mouth when coughing or sneezing,* ***R3***: *Keep a distance of at least 1 meter from other people,* ***R4***: *Avoid shaking hands, hugging or kissing when greeting others,* ***R5***: *Use alcohol-based hand rub' and* ***R6***: *Avoid touching my nose, eyes and mouth. Spearman rank correlation test is used for this analysis. * and ** denote significance at 1 and 5 percent levels respectively.*

Table S1: Classification of level of education into low/medium/high categories

Country	Level of education	Category
United Kingdom	Combined Junior and Infant School/ Infant School	Low
	Junior School	
	Comprehensive School	
	Comprehensive School (GCSE)/ Secondary Modern (GCSE)/ Grammar School (GSCE)/ City Technology College	Medium
	College and Institution of Higher education	High
	Open College -College of Technology - Institute/ Teacher Training College	
	University/ Open University	
Portugal	Sem Estudos	Low
	Primário Incompleto	
	Primário Completo	
	Nível Médio Incompleto	Medium
	Nível Médio Completo	
	Superior Incompleto	High
	Superior Completo	
Netherlands	LO (lagere school, LAVO, VGLO)	Low
	LBO (LBO, LTS, ITO, LEAO, Huishoudschool, LLO)	
	MAO (MAVO, IVO, MULO, ULO, 3jr HBS, 3jr VWO, 3jr VHMO)	
	MBO (MTS, UTS, MEAO)	Medium
	HAO (HAVO, VWO, Atheneum, Gymnasium, NMS, HBS, Lyceum)	
	HBO (HTS, HEAO, Wetensch. kand., Univers. onderwijs kand.)	High
	WO (Universitair onderwijs, Doctoraalopleiding, TH)	
Italy	Scuola elementare	Low
	Scuola media inferiore	
	Istituto professionale	Medium
	Scuola superiore	
	Università	High
	Master	
	Dottorato	
Germany	Grundschule	Low
	Hauptschule	
	Realschule	
	Gymnasium/ Berufliches Gymnasium/ Fachgymnasium, Gesamtschule	Medium
	Fachoberschule, Fachschule, Berufsschule, Berufsfachschule	
	Technische Hochschule, Pädagogische Hochschule, Kunsthochschule/ Musikhochschule	High
	Fachhochschule	
	Universität, Technische Universität	
France	École Primaire	Low
	Collège	
	Lycée d`Enseignement général et technologique	Medium
	Lycée professionnel	
	Grande École de Commerce et de Gestion/ scientifique, Établissement d`Enseignement supérieur catholique/ artistique/ agricole	High
	Grand Établissement/ École normale supérieure/ d`Ingénieur/ d`Architecture/ nationale vétérinaire	
	Université, Institut universitaire de Technologie/ national polytechnique/ d`Études politiques/ universitaire de Formation des Maîtres	
Denmark	Folkeskolen - f.eks. 9. eller 10. klasse	Low
	Gymnasial uddannelse - f.eks. Almen Gymnasium, HHX, HTX osv.	Medium
	En videregående erhvervsuddannelse - f.eks. landbrugs-, social- og sundheds uddannelser, produktionsskole	High
	En mellemlang videregående uddannelse	
	Universitets uddannelse	

Chapter 3.3

Social and Health disparities in Public Sentiments towards COVID-19: A Cross-Sectional study[27]

ABSTRACT

We study the determinants of public attitude to COVID-19 containment policies using the European COVID survey (ECOS), a large-scale Pan-European survey covering around 7000 individuals that are representative of populations from seven European countries, collected online during April 2-April 15, 2020. We find individual level socio-demographic and health factors to be significant predictors of public sentiments during the pandemic. The elderly, female and respondents from poor socio-economic position (SEP) were more likely to support the governments containment policies. They were also more likely to be worried regarding issues including loss of loved ones, overburdened hospital capacity, economic crises and food security. The social gradient in the impact of the virus is evident. Although higher support to the government policies may increase adherence to COVID-19 preventive measures, simultaneous worries might increase anxieties related to future and may question the tenacity of their support in the longer term. Overall, the study highlights the socially patterned public response to the outbreak and calls attention to the issue of wide polarities in societal response in exacerbating the social and health inequalities in the population.

[27] This is a working paper and we provide the latest version of the manuscript constituting this chapter. Author list as follow: Nirosha Elsem Varghese, Sebastian Neuman-Böhme, Iryna Sabat, Aleksandra Torbica, Pedro Pita Barros, Werner Brouwer, Job van Exel, Jonas Schreyögg and Tom Stargardt

INTRODUCTION

Initially, governmental responses to COVID-19 were limited to non-pharmaceutical interventions given the lack of definitive treatments. In response, mitigation measures like nation-wide lockdowns, border closures and curfews were adopted in Europe (IMF Policy Tracker, 2020). Nevertheless, countries varied in the timing, time span and stringency of the mitigation measures as well as the public approval of these policies (Sabat et al., 2020). Public willingness to cooperate during a crisis is crucial for the smooth implementation of government policies (Anderson et al., 2020). This willingness to comply depends on public belief in the effectiveness and support for the containment measures during the pandemic (Galasso et al., 2020).

According to KAP (Knowledge, attitudes and practices) theory, changes in human behavior depend on the attainment of knowledge, development of attitudes and formation of new practices (Ajilore et al., 2017). Lower levels of KAP have been shown to result in ill health and lower adherence of preventive behavior (Ajilore et al., 2017). A study analyzing online comments in response to news articles on a new vaccine captured unfavorable attitudes explaining public resistance to vaccination during the 2009 H1N1 pandemic (Henrich and Holmes, 2011). Similarly, in a representative sample of Italian residents, a mismatch between public expectations of COVID-19 preventive measures and government policies resulted in a lower willingness to comply (Briscese et al., 2020). A coherent understanding of attitudes and beliefs among the general population may help to identify and rectify perceptions and practices obstructing the effective compliance of safety measures during a public health emergency.

Similarly, varied worries during a pandemic have been shown to increase individual perception of risk to oneself (Khosravi, 2020). According to Protection Motivation Theory (PMT), higher levels of perceived risk indicate higher adherence to protective behaviors during a pandemic (Khosravi, 2020; Rogers and Prentice-Dunn, 1997). Studies have reported public worries during previous health crises to be related to higher likelihood of engaging in preventive actions, especially in the early stages of the pandemic (Jones and Salathe, 2009; Leung et al., 2005). On the other hand, worries during a crisis can spew misinformation and rumors (regarding etiology, prevention and cure of the disease) hindering the effective adherence of preventive measures. Therefore, individual worries may reflect either a healthy or an unfavorable psychological response to a public health crisis.

Recent studies on societal inequalities during COVID-19 have primarily contributed to evidence on disparities in COVID-19 attributed morbidity and mortality. A report prepared by Public Health England (PHE) using surveillance data shows COVID-19 has not only replicated existing inequalities but has worsened it with the older, male, respondents with low socioeconomic status (SES) and multiple comorbidities having a higher likelihood of being diagnosed and dying from COVID-19 (Public Health England, 2020). Similar patterns have been found in several European and non-European populations (Ahrenfeldt et al., 2020; Noor and Islam, 2020; Sannigrahi et al., 2020). Although these fatality estimates improve our knowledge on the severity of the disease and risk of transmission among population subgroups, the story remains incomplete in many ways. Studies on public attitudes and worries may help elucidate the higher risk of COVID-19 attributed morbidity and mortality among specific population groups.

A study conducted in 6 EU countries identified respondents who are older, female and economically better off reporting higher beliefs in the effectiveness of the lockdown policies (Margraf et al., 2020). Consistent findings were reported by studies conducted in other European countries (Meier et al., 2020; Naumann et al., 2020; Peretti-Watel et al., 2020). Gender differences in public perception towards EU policies during the pandemic has received special attention (Galasso et al., 2020). However, these findings contrasts with a recent study conducted in the Netherlands, Germany and Italy reporting no heterogeneous effects of public perceptions by gender (Meier et al., 2020). Findings on age were also contrasted in an online survey of 1500 Americans showing perceived personal health risks associated with COVID-19 to fall sharply with age (Bordalo et al., 2020). This could explain the lower adherence among the elderly in a study conducted in the Italian population (Carlucci et al., 2020). Using data from 7 EU countries, it was found that the elderly reported the least support for ban of public transportation (Sabat et al., 2020) which could be attributed to fear of loss of freedom of movement or the need to maintain a social life (Carlucci et al., 2020). Along the same vein, economic indicators did not retain statistical significance in a multivariate analysis on public attitudes conducted in the French population (Peretti-Watel et al., 2020). Moreover, existing studies also suffer from using convenience samples (Carlucci et al., 2020; Meier et al., 2020).These inconsistent findings for social and health correlates of public support could be attributed to differences in sampling methods, differences in methodologies, non-comparable definitions of socio-demographic correlates and outcomes measuring public attitudes. Similar challenges exist for the analysis of public worries as well.

Existing studies on COVID-19 attributed worries have consistently shown the elderly, male and respondents belonging to higher SES reporting fewer worries (Barber and Kim, 2021; Fitzpatrick et al., 2020; Maaravi and Heller, 2020). Yet, some studies fail to find a statistically valid relationship (Horesh et al., 2020) or such analysis is not the goal even when the required data is available (Varga et al., 2021). Moreover, a multivariate analysis is often omitted which reduces the robustness of the findings (Fitzpatrick et al., 2020; Maaravi and Heller, 2020). The literature also lacks in regard to the comparability of findings along different domains of worry. For instance, older individuals might be more worried about risk to life or health but less about unemployment. Finally, there is a scarcity of literature on COVID-19 worries that is geographically comparable, particularly in a European context.

In light of this literature, we conducted a large scale survey focusing on public perceptions during early stages of the COVID-19 pandemic to study individual level variations in socio-demographic and health characteristics in public support and worries across several domains. Thereby, we contribute to the literature in several ways. One, we provide evidence on the determinants of public sentiments during the early stages of COVID-19. Two, our results can be generalized to the European population under study. Three, our results inform policy makers to target and customize information to the less responsive groups for maximum benefit of the non-pharmaceutical interventions for themselves and the public at large and four, our findings may add to the discussion on the role of societal inequalities in public attitudes during a pandemic.

Based on previous literature, age, gender, education, financial and health status were selected as factors that could influence public support and worries during COVID-19 (Azlan et al., 2020; Fitzpatrick et al., 2020; Galasso et al., 2020; Margraf et al., 2020; Meier et al., 2020; Peretti-Watel et al., 2020). Additionally, we include factors such as household composition, intensity of following news on COVID-19, source of information on COVID-19 and knowledge of COVID-19 positive case among social contacts as correlates of public perception during a pandemic (Meier et al., 2020; Peretti-Watel et al., 2020). We also include trust in information from national government as a predictor of public support as trust ensures the smooth implementation of all policies that requires citizen compliance (Marien and Hooghe, 2011). Participants from different countries are hypothesized to react differently during the pandemic and therefore we account for this by including country fixed effects.

MATERIALS AND METHODS

Data

The data comes from the first wave of the European Covid survey (ECOS) conducted during April 2-15, 2020, by the market research company Dynata. Around 7000 respondents from seven European countries: Denmark, France, Germany, Italy, Portugal, the Netherlands, and the UK participated in the online cross-sectional survey. The questionnaire was initially developed in English by the authors of the study, with the core parts of the survey addressing topics related to public sentiments towards COVID-19 containment policies and COVID-19 related worries among the public. The questionnaire was then translated and adapted to country specific context by native speakers with a background in health economics. "In each country, data was collected from 1000 respondents' representative of the national population in terms of region, age, gender, and education" (Varghese et al., 2020). A detailed description of the survey methodology and variables used in this study can be found in an earlier ECOS publication (Sabat et al., 2020).

Outcome variables

Our two outcome variables include public support towards COVID-19 containment policies and public worries during the COVID-19 pandemic. Public support in seven domains including school closures, mobility restrictions (mean of support for curfews, cancelling public gatherings and ban on public transportation), border closure, ban on export of medical device, fine for violating quarantine, random temperature checks and using mobile phones for tracking movements were measured on a 5 point Likert scale: 1=strongly disagree, 2=disagree, 3=indifferent; 4=agree and 5=strongly agree. For our analysis, we consider an individual to disagree if they 'strongly disagree' or 'disagree' and agree if they 'strongly agree' or 'agree' (1= disagree, 2=indifferent, 3=agree). Additionally, we also construct a measure of overall public attitude as a mean of policy support in all domains.

Participants responded to concerns regarding losing loved ones, health system, school closures, food security and economic situation (mean of worries reported on small companies running out of business, unemployment and recession) in their country. The possible responses were: 1=not at all worried, 2=slightly, 3=moderately, 4=quite a lot; 5=a lot. For our analysis, we consider the respondents to be not worried if they respond as 'not at all worried' or 'slightly worried' and as worried if they choose 'quite a lot' or 'a lot' (1=not worried, 2=moderately worried, 3=worried). We also construct a measure of overall worries as a mean of reported level of worries in all domains. The above categorizations were done to ensure sufficient observations in each group to conduct a robust analysis.

Independent variables

Individual level predictors include information on age, gender, level of education, respondent's ability to meet household financial needs and self-reported health. EQ-5D-5L, a five-item questionnaire encompassing the domains of mobility, self-care, usual activities, pain/discomfort and anxiety/depression was used to proxy self-reported health status (Herdman et al., 2011). Respondents indicated their level of health problems in each domain on a scale of 1 (no problem) to 5 (unable to do the activity). For our analysis, we take the sum of responses across the five domains of EQ5D and rescaled to 0-100 with a higher score indicating worse health (Himmler et al., 2020). Similarly, we use ICECAP-A as an alternate measure of health as a predictor of worries among adults. ICECAP-A is assessed through five attributes of capability wellbeing namely stability, attachment, autonomy, achievement, and enjoyment. Each attribute is measured on a scale of 1 (attribute applicable in all areas of life) to 4 (attribute not applicable in any areas of life). For our analysis, we take the sum of responses across the five domains of ICECAP-A and rescaled to 0-100 with a higher score indicating poor quality of life (Keeley et al., 2015). Studies comparing ICECAP-A and EQ-5D-5L shows that ICECAP-A provides complementary information in terms of wellbeing and quality of life to EQ-5D-5L (Goranitis et al., 2016).

Control variables

Other confounders include an indicator for household consisting of vulnerable family members (elderly or those with comorbidities), categorical variable indicating the intensity of following news on COVID-19, using social media as one of the primary source of information on COVID-19 and knowledge of a COVID-19 positive case among social contacts (confirmed or not confirmed vs. no). Trust in COVID-19 information from the government (1= do not trust, 2= moderately, 3=trust) is also included as a covariate in the models predicting public support.

Statistical methods

To investigate the socio-demographic and health determinants of public support and worries, Odds Ratio (OR) and 95% Confidence Interval (CI) were estimated using ordered logistic regression models, simultaneously controlling for potential confounders. Our statistical specification is as follows:

$$Y_i = \alpha_0 + X_{1i}\,\alpha_1 + X_{2i}\,\alpha_2 + X_{3i}\,\alpha_3 + \varepsilon_i$$

where Y_i corresponds to outcomes for each individual; X_1 is the set of socio-demographic and health determinants of interest; X_2 corresponds to other covariates of interest and X_3 corresponds to either country fixed effects (in the main analyses) or country level variables (for robustness checks). Robust standard errors were used in all regressions.

RESULTS

Main results

Table 1 shows the descriptive statistics for the socio-demographic, health and other variables used in the study. The mean age of the respondents is 47 years with approximately 50% of the sample composed of females, 30% reporting low level of education and 50% reporting difficulty in their ability to meet household financial needs. Approximately half of the respondents reported ‘slight’ to ‘extreme’ pain/discomfort or anxiety/depression. Overall wellbeing also suffered with only one-fifth of the respondents reporting feelings of achievement and progress in all areas of their lives. Table S1 reports the descriptive statistics for the outcome variables by socio-demographic and health characteristics. Overall, respondents who approved of the COVID-19 policies were relatively older, female, less healthy and more likely to be those with low level of education and facing difficulties in meeting household financial needs. They were also more likely to report overall worries. A description on the distribution of the outcomes used in our analysis can be found elsewhere (Sabat et al., 2020).

[Place Table 1 about here]

Figure S1 graphs the OR and 95% CI for overall policy support. The older and those reporting poor health reported higher support whereas ability to meet ends meet easily was negatively related to overall policy support. Gender and level of education were not statistically related to overall policy support.

Table 2 reports the OR for socio-demographic and health determinants for public support in each policy domain. Compared to the 18-24 age categories, respondents aged 55 and above reported higher support in all domains except for suspension of public transportation which received the least support from the elderly. Although the relationship between age and support for overall mobility restrictions is positive, respondents aged 65 and above were less likely to support ban on public transportation (results not shown). Being female was positively related to support for border closures and monetary penalties but negatively related to support for ban on export of medical device, temperature checks and mobile phone tracking. Results on gender were not statistically significant for school closures and restriction of overall mobility. Being female was directly related to policy support for restricting public gatherings (results not shown). Higher educated were less likely to support all policies except for ban on school closure and monetary policies whereas those with better financial ability were less likely to support all policies except for monetary penalties. The relationship between level of education and public support is statistically significant in all domains except for school closure, ban on export of medical devices and monetary penalties whereas the corresponding relationship for financial ability did not attain statistical significance for support for mobile phone tracking. Respondents with poor health reported higher support for COVID-19 related government policies. However, this relationship was not statistically significant when the policies concerned were school closure, border closure and monetary penalties. The results examining each component of mobility restriction can be obtained from authors on request.

[Place Table 2 about here]

Table S2 reports the OR for public support for COVID-19 policies by other covariates included in our analysis. Respondents with vulnerable family members reported lower support for school closures but higher support for ban on export of medical device and using mobile phones for tracking movement. Higher trust in information from government and using social media as one of the primary source of information was positively related to policy support whereas respondents not following COVID-19 related news closely reported lower support. Awareness of a COVID-19 infected among social contacts was related to higher support for temperature checks and mobile phone tracking but lower support for border closure. Compared to Italy, residents of other countries were less likely to support policies in all domains with the exception of Portugal reporting higher support for border closures.

Figure S2 graphs the OR and 95% CI for overall worries by socio-demographic and health variables. The figure plotted on the left uses EQ-5D as a proxy for health status and the figure on the right side uses ICECAP-A to represent health status. Age, female, poor health and wellbeing is directly related to higher overall worries whereas higher education and better financial situation predicts lower overall worries.

Table 3 presents the OR for socio-demographic and health determinants for COVID-19 related worries. Compared to the 18-24 age group, respondents belonging to older age categories were more likely to be worried in all domains. The elderly reported the highest worries for health system being overburdened whereas worries regarding the overall economic situation and school closures were higher among the younger age groups. Looking at each component of economic worries, respondents aged 55 and above report higher worries regarding recession and small companies running out of business but relatively lower worries regarding unemployment (results not shown). Being female is directly related to worries in all domains. Respondents with better financial ability reported lower worries in all domains whereas the higher educated were less likely to be worried regarding losing someone, school closures and restricted access to food supply but more worried about health system being overburdened and overall economic situation. When looking at each component of economic worries, we find that higher educated were more worried about recession but less worried about being unemployed (results not shown). Poor health (EQ-5D) and wellbeing (ICECAP-A) was directly related to worries in all domains except school closures. The positive relationship between poor health and economic worries is not statistically significant. However, redoing the analysis for each component of economic worries, we find poor health to be positively related to worries about recession but negatively related to unemployment (results not shown). The results examining each component of economic worries can be obtained from authors on request.

[Place Table 3 about here]

Table S3 reports the OR for COVID-19 related worries by other covariates included in our analysis. Respondents with vulnerable family members reported higher worries although this relationship was not statistically significant for economic worries and health system being overburdened. A lower intensity of following COVID-19 news was related to lower worries whereas awareness of a COVID-19 infected among social contacts and using social media as one of the primary sources of information positively predicted higher worries. Compared to Italy, residents of other countries reported lower worries in all domains with the exception of Portugal reporting relatively higher worries.

Robustness checks

We replace country indicators with external country-specific data on COVID-19 prevalence (and death), stringency of containment policies and income support on the date of survey participation as predictors in our analysis. The public health situation in the country and the severity of the containment policies are likely to influence public attitude to the COVID-19 policies (Sabat et al., 2020). Also income support during the pandemic could reduce some of the worries among the public (Sabat et al., 2020). Country specific data on COVID-19 prevalence per 1 million people were collected from the data repository of European Centre for Disease Prevention and Control (Roser et al., 2020). The COVID-19 Government Response Stringency Index introduced by the Oxford University measures the degree of strictness in the government's containment policies on a range from 0-100, with higher values indicating stricter measures (Hale et al., 2020). Income support during the COVID-19 pandemic was measured on a scale of 0-2 (0= no income support; 1= covers <50% of lost salary and 2= covers >50% of lost salary (Hale et al., 2020). We do not include country level information simultaneously with indicators for each country as these indicators absorbs all of these country level variations. Prevalence of COVID-19 cases, level of stringency and income support are also highly correlated ($\rho = 0.7 - 0.8$) with each other and are therefore not simultaneously included in the same model.

Results are qualitatively similar when replacing country indicators with the external country specific information. Higher number of COVID-19 cases and higher stringency of government response were both positively related to higher public support. A higher prevalence of cases was positively related to worries whereas a higher income support was negatively related to worries. Higher number of deaths is positively related to both higher public support and worries.

We also repeat our main analyses after removing the middle category from the two dependent variables. Previous research has shown that responses style behaviors arising from specific question formats may cause non-random response errors (Moors, 2008)**.** For example, respondents are more likely to opt for a middle category (neutral response) when this option is available. Our results remain robust to this check (results available from author upon request).

DISCUSSION

The multi-country cross-sectional study showed individual level variation in socio-demographic (age, gender, education, and financial ability) and health variables to be significant predictors of public sentiments during the early and escalating phase of the COVID-19 pandemic. Additionally, we also show severity of the outbreak, stringency of the containment policies and economic support in each country also explained some of the variation in public perception and physiological response to COVID-19.

In particular, we find age to be directly related to public support for all COVID-19 containment policies. Our findings for age and public support is consistent with studies conducted in both European (Margraf et al., 2020; Meier et al., 2020; Peretti-Watel et al., 2020) and non-European countries (Azlan et al., 2020; Margraf et al., 2020). This could be due to the higher institutional trust among the elderly (Marien and Hooghe, 2011) and also the general understanding that age is a major risk factor for COVID-19 patients (Public Health England, 2020). Furthermore, some of the lockdown restrictions affected the lifestyle and preferred leisure activities of some age-groups more than others. Younger people are likely to be more negatively affected in their daily life by school closures and mobility restrictions than older individuals.

The higher policy support among elderly is accompanied by higher worries regarding recession and access to health care. Similar worries among the elderly have been reported in recent COVID-19 studies (Bergman et al., 2020; Neuman and Koma, 2020). According to Eurostat, 1 in 7 pensioners are at risk of poverty in the EU as of 2018 and this trend has been increasing since 2013 (12.6% in 2013 to 14.2% in 2017) (EUROSTAT, 2017). Moreover, the pandemic can worsen the economic status of elderly via fluctuations in stock market on retirement savings (Neuman and Koma, 2020) and difficulties in collecting benefits and compensations due to mobility restrictions. They are also affected by obstructions to access to health care in mainly two ways during the pandemic. First, the pandemic brought about a discrimination against the elderly via 'priority setting in health care' whereby a priori age-cuts were set to decide who can access healthcare (Carrieri et al., 2020). For example, SIAARTI guidelines on admissions to ICUs in Italy propose potential age-cuts based on the argument that older individuals need longer recovery periods and have fewer chances of survival (SIAARTI, 2003). Previous research has suggested such prioritization of COVID-19 patients by health services to be one of the reasons for excess mortality in several regions (Vandoros, 2020). Second, overcrowding in hospitals during pandemic implies breaks in other health care services which in turn increase morbidity and mortality during the pandemic. However younger age groups reported higher worries regarding unemployment. Also the effect size for overall economic worries among 35-44 years is almost 1.5 times that of respondents aged 65 and above and reinforces evidence from previous studies (Horesh et al., 2020; Mann et al., 2020). Overall, our findings suggest that elderly worry more about health related aspects whereas the young worry more about the economy.

The risk of COVID-19 is four times higher in patients with COPD compared to those without (Zhao et al., 2020). This evidence supports our findings on participants with poor health reporting higher overall worries. A study on the health related quality of life (measured using EQ-5D) in the Chinese population during the pandemic found pain/discomfort and anxiety/depression to be the most frequently reported problems which is also true in our dataset of European adults (Ping et al., 2020). Supporting our findings, the study also shows EQ5D to be correlated with higher worries during COVID-19 (Ping et al., 2020).

Our results indicate lower policy support among the highly educated and financially well off. This is consistent with evidence from the Ipsos survey showing higher income households to be more in favor of reopening the economy even if the virus is not contained (Ipsos Survey, 2020). They were also more likely to have concerns regarding the possibility of the government or a third party to access their personal information via contact tracing apps or temperature checks. Although higher SES groups are more likely to trust in information from the government (Alesina and La Ferrara, 2002), our results show that this does not necessarily mean higher trust in government's use of personal information (Altmann et al., 2020).The higher educated were less likely to support mobility restrictions probably because of the short and potential long term negative consequences on the economy as evident from the reported worries concerning unemployment or recession. Worries among college students have been especially linked to anxiety of future employment (Cao et al., 2020).

We also provide evidence on the disproportionate risk of COVID-19 worries among the socially disadvantaged. Poor households have more to lose from reopening of the economy. They are constantly exposed to co-workers as many low-skilled jobs are manual or not suited for working remotely. These findings are in line with studies reporting an income gradient in the COVID-19 outbreak in several countries (Baena-Díez et al., 2020; Sannigrahi et al., 2020). Given the higher support among the low SES subgroups, their ability to adhere to the COVID-19 preventive measures might be limited if not protected by some level of economic support. Households reliant on hourly earnings may be forced to go out and look for jobs, increasing their likelihood of being infected. Literature shows communities of similar economic status to cluster together, thereby multiplying the risk (Fisher and Bubola, 2020). Overall, these social inequities in COVID-19 perceptions could be attributed to the ability of public to access and apprehend the right health information and most importantly the privilege to act on this knowledge.

Evidence from our study highlights a gendered vulnerability to the COVID-19 outbreak. Women were more likely to be worried in all domains compared to men. Our results are consistent with evidence from eight OECD countries investigating gender differences in COVID-19 attitudes and behaviors (Galasso et al., 2020). The authors call for gender-based public health communication given the behavioral differences when facing new crises (Galasso et al., 2020). Given the roles majority of women hold in the society as that of a primary caregiver in both at home and in a health care setting, they are at a higher risk of being exposed to the virus. School and childcare closures during the COVID-19 pandemic have intensified this workload. While this may not be the case for every woman, there is for example already evidence that female researchers, especially those with young children report a significant decline in the time they can devote to their research. All else being equal, female scientists in the US and Europe reported a 5% larger decline in research time (Myers et al., 2020). Women also reported higher worries regarding the health system being overburdened. Previous epidemics have shown an increase in maternal and neonatal mortality and a cutoff from sexual and reproductive services in general (Frontieres, 2020). Although mortality attributed to COVID-19 has been shown to be lower for females (Public Health England, 2020), this could be due to higher willingness to cooperate with the COVID-19 recommendation as a result of higher institutional trust (Marien and Hooghe, 2011) and worries among women.

Descriptive evidence using the same data shows Italy and Portugal reporting higher health related worries and attributes this pattern to the north-south divide in the progression of the epidemic with Italy, France and UK being the most affected states in the initial phase of the epidemic (Sabat et al., 2020). These findings are supported in our multivariate analysis where all countries with the exception of Portugal reports lower health related worries compared to Italy. Likewise, Italy was the most affected country whose national government provided the least economic support during the COVID-19 crisis, thereby explaining the lower economic worries in all countries compared to Italy. The only exception was Portugal reporting higher economic worries compared to Italy. One reason for this could be the relatively high share of temporary workers in Portugal resulting in lower social protection and job security. Compared to Italy, all EU countries in our sample also reported lower support for containment measures. The exceptions were France and Portugal reporting higher support for containment measures including support for school closure, mobility restrictions, border closure and ban on export of medical device. This could again be attributed to the severity of the pandemic despite the stringent measures implemented in these countries.

Overall, subgroups reporting higher support for containment policies were also more likely to be worried suggestive of the protection motivation theory (Rogers and Prentice-Dunn, 1997). Our findings are also consistent with evidence on higher consumption of COVID-19 news to be related to higher worries (Gao et al., 2020). However, not following any COVID-19 related news might also result in the public missing out on important information leading to lower support for containment measures. Thus, the aim should be to achieve a balance with adequate level of knowledge, choosing the right source of information and protecting oneself from the misinformation and infodemic.

Our study has some limitations. First, our results are not causal and should be interpreted accordingly. Second, the health variables used in our study, EQ5D-5L and ICECAP-A require rescaling by country specific weights since the distance between the categories is not uniform across countries. Given weights are not available for all countries and hence not comparable, we omit this step. Third, our results could be contaminated by selection and social desirability bias given the online nature of the survey and self-reported behaviors. Four, the dependent variables used in this study measures public attitude during COVID-19. Although positive attitudes to containment policies may predict adherence (Ajilore et al., 2017), this may not always be true (Mohamed et al., 2021). Five, we find qualitatively similar results when repeating the analyses for overall policy support and worries for each country separately (results available on request). However, the coefficients for socio-demographic variables varied in their statistical validity and effect sizes. Some of the relationships that were not statistically valid in the pooled sample retain their statistical significance in the country specific data or vice-versa. One example is females reporting significantly higher overall policy support in the UK and Italy. Moreover, country specific results may also vary by the type of containment measures or worries. Although these findings are important from a country specific viewpoint, this is not our aim and should be considered for future research.

POLICY RECOMMENDATIONS

Several lessons can be drawn from our study on socio-demographic correlates of public response to the COVID-19 pandemic:

First, monetary penalties for violating quarantines are an additional burden on low-income earners. Given that they disproportionately bear the economic costs of quarantine in forms of job loss, wage reduction and increased risk to health, a universal penalty system should be replaced by policies encouraging them to comply with quarantine rules. One example could be the implementation of wage replacement programme by local governments for eligible workers following the quarantine (Stieg, 2020).

Second, our findings signal concerns regarding use of mobile phones for surveillance purposes (Naumann et al., 2020) especially among women and higher educated. Overall, there have been varying degrees of success with contact tracing apps across the EU. For example, Portugal's official contract tracing app 'StayAway Covid' was reported to have fewer than required downloads necessary to break the spread of the pandemic whereas in the UK, the government has been criticized for not introducing legislative basis for the app (The Law Library of Congress, 2020). There is a need for a transparent and homogenized system of surveillances such as the interconnectivity of apps across national borders. Policies should be set in place to protect women against cyber security threats or harassment. Moreover, local governments need to ensure that low income neighbourhood without proper accesses to broadband signals are not excluded from these activities compounding socio-economic inequalities.

Third, the relatively lower support on the temporary ban of public transportation among the elderly might suggest a regular use of such services. New solutions to support the mobility related needs of the elderly should be designed and implemented. Volunteering programs such as 'Caido' ensuring continued and timely access to food, medicines and monetary benefits for the elderly (CAIDO) or measures to ensure proper social distancing among passengers is required.

Fourth, Policy makers could try to increase the acceptability and adherence to containment policies among younger generations by creating spaces to meet outside of (their parents) home simultaneously maintaining measures to reduce the risk of infections. Furthermore, technical solutions such as mobile air purifiers have been discussed as effective solutions to allow schools to re-open and mitigate the social, mental health and educational effects of the pandemic on children (Curtius et al., 2021). Our results show that support for school closures is higher among older respondents, while the negative effects of school closure are largely borne by school children and their parents.

Finally, females and respondents with economic and health vulnerabilities report higher worries. Any government response addressing such worries should prioritize at-risk population groups that are unprepared to battle the consequences of the pandemic. Possible solutions may include prioritized access to reproductive and mental health services, recognizing additional risk to women health care workers by providing uninterrupted supply of Personal Protective Equipment (PPE) and paid parental leave or children allowances during COVID-19. Targeted emergency financial and psychological support provisions for at-risk groups may also help to tackle the inequalities of the COVID-19 crisis.

Overall, our study focuses on the social patterning in the public response to COVID-19. Our findings call for further efforts to collect individual level data on socio-demographics to identify at-risk groups as the pandemic evolves. Development of pandemic recommendations and public health interventions will benefit from further research on the role of social determinants of public response to COVID-19 across countries and over time.

BIBLIOGRAPHY

Ahrenfeldt, L. J., *et al.* 2020. Sex and age differences in COVID-19 mortality in Europe.

Ajilore, K., Atakiti, I. and Onyenankeya, K. 2017. College students' knowledge, attitudes and adherence to public service announcements on Ebola in Nigeria: Suggestions for improving future Ebola prevention education programmes. *Health Education Journal*, 76(6), 648-660.

Alesina, A. and La Ferrara, E. 2002. Who trusts others? *Journal of public economics,* 85(2), 207-234.

Altmann, S., *et al.* 2020. Acceptability of App-Based Contact Tracing for COVID-19: Cross-Country Survey Study. *JMIR mHealth and uHealth,* 8(8), e19857.

Anderson, R. M., *et al.* 2020. How will country-based mitigation measures influence the course of the COVID-19 epidemic? *The Lancet*, 395(10228), 931-934.

Azlan, A. A., *et al.* 2020. Public knowledge, attitudes and practices towards COVID-19: A cross-sectional study in Malaysia. *Plos one,* 15(5), e0233668.

Baena-Díez, J. M., *et al.* 2020. Impact of COVID-19 outbreak by income: hitting hardest the most deprived. *Journal of Public Health*.

Barber, S. J. and Kim, H. 2021. COVID-19 worries and behavior changes in older and younger men and women. *The Journals of Gerontology: Series B,* 76(2), e17-e23.

Bergman, Y. S., *et al.* 2020. COVID-19 health worries and anxiety symptoms among older adults: the moderating role of ageism. *International psychogeriatrics*, 32(11), 1371-1375.

Bordalo, P., *et al.*, 2020. *Older people are less pessimistic about the health risks of COVID-19.* National Bureau of Economic Research, 0898-2937.

Briscese, G., *et al*., 2020. *Compliance with covid-19 social-distancing measures in italy: the role of expectations and duration.* National Bureau of Economic Research, 0898-2937.

CAIDO, *Assistenza per anziani a Udine: associazione con personale specializzato* [online]. Available from: caido.it.

Cao, W., *et al.* 2020. The psychological impact of the COVID-19 epidemic on college students in China. *Psychiatry research*, 112934.

Carlucci, L., D'Ambrosio, I. and Balsamo, M. 2020. Demographic and attitudinal factors of adherence to quarantine guidelines during COVID-19: The Italian Model. *Frontiers in Psychology,* 11, 2702.

Carrieri, D., Peccatori, F. A. and Boniolo, G. 2020. COVID-19: a plea to protect the older population. *International Journal for Equity in Health*, 19(1), 1-4.

Curtius, J., Granzin, M. and Schrod, J. 2021. Testing mobile air purifiers in a school classroom: Reducing the airborne transmission risk for SARS-CoV-2. *Aerosol Science and Technology*, 1-18.

EUROSTAT 2017. At-risk-of-poverty rate by poverty threshold and most frequent activity in the previous year - EU-SILC and ECHP surveys.

Fisher, M. and Bubola, E. 2020. As Coronavirus deepens inequality, inequality worsens its spread. *New York Times*.

Fitzpatrick, K. M., Drawve, G. and Harris, C. 2020. Facing new fears during the COVID-19 pandemic: The State of America's mental health. *Journal of anxiety disorders,* 75, 102291.

Frontieres, M. S., 2020. Women and girls face greater dangers during COVID-19 pandemic. [online]. Available from: https://www.msf.org/women-and-girls-face-greater-dangers-during-covid-19-pandemic.

Galasso, V., *et al.* 2020. Gender differences in COVID-19 attitudes and behavior: Panel evidence from eight countries. *Proceedings of the National Academy of Sciences,* 117(44), 27285-27291.

Gao, J., *et al.* 2020. Mental health problems and social media exposure during COVID-19 outbreak. *Plos one,* 15(4), e0231924.

Goranitis, I., *et al.* 2016. Measuring health and broader well-being benefits in the context of opiate dependence: the psychometric performance of the ICECAP-A and the EQ-5D-5L. *Value in health,* 19(6), 820-828.

Hale, T., *et al.* 2020. Variation in government responses to COVID-19. *Blavatnik school of government working paper*, 31.

Henrich, N. and Holmes, B. 2011. What the public was saying about the H1N1 vaccine: perceptions and issues discussed in on-line comments during the 2009 H1N1 pandemic. *PloS one,* 6(4), e18479.

Herdman, M., *et al.* 2011. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Quality of life research,* 20(10), 1727-1736.

Himmler, S., *et al.* 2020. Willingness to pay for an early warning system for infectious diseases. *The European Journal of Health Economics*, 1-11.

Horesh, D., Kapel Lev-Ari, R. and Hasson-Ohayon, I. 2020. Risk factors for psychological distress during the COVID-19 pandemic in Israel: Loneliness, age, gender, and health status play an important role. *British journal of health psychology,* 25(4), 925-933.

IMF Policy Tracker, 2020. [online].

Ipsos Survey 2020. People are divided over whether the economy should open if COVID-19 is not fully contained.

Jones, J. H. and Salathe, M. 2009. Early assessment of anxiety and behavioral response to novel swine-origin influenza A (H1N1). *PLoS one,* 4(12), e8032.

Keeley, T., *et al.* 2015. A longitudinal assessment of the responsiveness of the ICECAP-A in a randomised controlled trial of a knee pain intervention. *Quality of life research,* 24(10), 2319-2331.

Khosravi, M., 2020. Perceived Risk of COVID-19 Pandemic: The Role of Public Worry and Trust. Electron J Gen Med. 2020; 17 (4): em203.

Leung, G. M., *et al.* 2005. Longitudinal assessment of community psychobehavioral responses during and after the 2003 outbreak of severe acute respiratory syndrome in Hong Kong. *Clinical Infectious Diseases,* 40(12), 1713-1720.

Maaravi, Y. and Heller, B. 2020. Not all worries were created equal: the case of COVID-19 anxiety. *Public Health,* 185, 243-245.

Mann, F. D., Krueger, R. F. and Vohs, K. D. 2020. Personal economic anxiety in response to COVID-19. *Personality and Individual Differences,* 167, 110233.

Margraf, J., Brailovskaia, J. and Schneider, S. 2020. Behavioral measures to fight COVID-19: An 8-country study of perceived usefulness, adherence and their predictors. *Plos one,* 15(12), e0243523.

Marien, S. and Hooghe, M. 2011. Does political trust matter? An empirical investigation into the relation between political trust and support for law compliance. *European Journal of Political Research,* 50(2), 267-291.

Meier, K., *et al.* 2020. Public perspectives on protective measures during the COVID-19 pandemic in the Netherlands, Germany and Italy: A survey study. *PloS one,* 15(8), e0236917.

Mohamed, A. A. O., *et al.* 2021. Knowledge, attitude and practice of the Sudanese people towards COVID-19: an online survey. *BMC Public Health,* 21(1), 1-7.

Moors, G. 2008. Exploring the effect of a middle response category on response style in attitude measurement. *Quality & quantity,* 42(6), 779-794.

Myers, K. R., *et al.* 2020. Unequal effects of the COVID-19 pandemic on scientists. *Nature human behaviour,* 4(9), 880-883.

Naumann, E., *et al*. 2020. COVID-19 policies in Germany and their social, political, and psychological consequences. *European Policy Analysis*.

Neuman;, T. and Koma, W., 2020. Retirement Insecurity in the Time of COVID-19: The Next Shoe to Drop? [online]. Available from: https://www.kff.org/policy-watch/retirement-insecurity-in-time-of-covid-19-next-shoe-to-drop/ [Accessed February 16, 2021].

Noor, F. M. and Islam, M. M. 2020. Prevalence and Associated Risk Factors of Mortality Among COVID-19 Patients: A Meta-Analysis. *Journal of community health*, 1-13.

Peretti-Watel, P., Verger, P. and Launay, O. 2020. The French general population's attitudes toward lockdown against COVID-19: a fragile consensus. *BMC public health*, 20(1), 1-8.

Ping, W., *et al.* 2020. Evaluation of health-related quality of life using EQ-5D in China during the COVID-19 pandemic. *PloS one*, 15(6), e0234850.

Public Health England 2020. Disparities in the risk and outcomes of COVID-19.

Rogers, R. W. and Prentice-Dunn, S. 1997. Protection motivation theory.

Roser, M., *et al.* 2020. Coronavirus pandemic (COVID-19). *Our world in data*.

Sabat, I., *et al.* 2020. United but divided: policy responses and people's perceptions in the EU during the COVID-19 outbreak. *Health Policy*.

Sannigrahi, S., *et al.* 2020. Examining the association between socio-demographic composition and COVID-19 fatalities in the European region using spatial regression approach. *Sustainable cities and society*, 62, 102418.

SIAARTI, G. d. S. B. 2003. SIAARTI Guidelines for admission to and discharge from Intensive Care Units and for limitation of treatments in Intensive Care. *Minerva Anestesiol*, 69, 101-111.

Stieg, C., 2020. Could you get paid to quarantine during the Covid-19 pandemic? Some local governments are already doing it. [online]. [Accessed 5 October 2020].

The Law Library of Congress, 2020 Regulating Electronic Means to Fight the Spread of COVID-19. [online]. Available from: https://www.loc.gov/law/help/coronavirus-apps/coronavirus-apps.pdf [Accessed January 2021].

Vandoros, S. 2020. Excess Mortality during the Covid-19 pandemic: Early evidence from England and Wales. *Social Science & Medicine*, 113101.

Varga, T. V., *et al.* 2021. Loneliness, worries, anxiety, and precautionary behaviours in response to the COVID-19 pandemic: a longitudinal analysis of 200,000 Western and Northern Europeans. *The Lancet Regional Health-Europe*, 100020.

Varghese, N. E., *et al.*, 2020. Europeans know and act on WHO recommendations during COVID-19. [online]. Available from: https://voxeu.org/article/europeans-know-and-act-who-recommendations-during-covid-19.

Zhao, Q., *et al.* 2020. The impact of COPD and smoking history on the severity of COVID-19: a systemic review and meta-analysis. *Journal of medical virology*.

TABLES

Table 1: Descriptive statistics for demographic and other variables

Age (%)	
18-24 years	10.4
25-34 years	16.5
35-44 years	18.9
45-54 years	18.1
55-64 years	15.5
65+ years	20.7
Age	47.1
Female (%)	51.8
Level of Education (%)	
Low	29.6
Medium	44.1
High	26.3
Ability to meet financial needs (%)	
With great difficulty	11.6
With some difficulty	38.1
Fairly easily	38.5
Easily	11.8
***EQ5D score (0-100)**	29.88 (11.57)
Mobility (no problems)	75.5
Self-care (no problems)	90.2
Usual activities (no problems)	74.5
Pain/Discomfort (no problems)	47.3
Anxiety/Depression (no problems)	50.7
***ICECAP-A (0-100)**	49.75 (14.51)
Feeling settled and secure (all areas of my life)	23.4
Love, friendship and support (all areas of my life)	36.3
Being independent (all areas of my life)	39.8
Achievement and progress (all areas of my life)	18.7
Enjoyment and pleasure (all areas of my life)	28.3
Trust in COVID-19 information from government (%)	
Do not trust	19.1
Indifferent	24.5
Trust	56.4
Vulnerable household members (%)	55.9
Follow news on COVID-19 (%)	
Very closely	57.0
Somewhat closely	37.1
Not closely at all	5.0
Do not know anything about it	0.9
Use social media for information on COVID-19 (%)	32.9
Contact with infected (%)	22.3

Note: Numbers shown in parenthesis are standard deviations.
*Higher values of EQ5D-5L and ICECAP-A indicate worse health and wellbeing.

Table 2: Odds ratio (OR) for public support for specific COVID-19 policies by socio-demographic and health determinants

VARIABLES	School closure	Restrictions to mobility*	Border closure	Ban on export of medical device	Fine for violating quarantine	Temperature check	Mobile phone for tracking movement
Age (Reference category: 18-24 years)							
25-34 years	1.174	1.294***	1.239*	1.291***	1.198	1.324***	1.455***
	(0.138)	(0.127)	(0.148)	(0.113)	(0.137)	(0.137)	(0.135)
35-44 years	1.197	1.214**	1.414***	1.482***	1.354***	1.258**	1.487***
	(0.139)	(0.118)	(0.170)	(0.129)	(0.155)	(0.126)	(0.137)
45-54 years	1.390***	1.260**	1.590***	1.503***	1.453***	1.331***	1.605***
	(0.164)	(0.123)	(0.197)	(0.133)	(0.171)	(0.134)	(0.148)
55-64 years	1.572***	1.358***	1.616***	1.905***	2.019***	1.280**	1.743***
	(0.202)	(0.137)	(0.211)	(0.177)	(0.267)	(0.133)	(0.170)
65+ years	1.502***	1.306***	1.632***	2.013***	2.395***	1.292***	2.204***
	(0.184)	(0.129)	(0.201)	(0.180)	(0.313)	(0.128)	(0.206)
Gender (Reference category: Male)							
Female	0.979	1.077	1.530***	0.911**	1.322***	0.868***	0.900**
	(0.061)	(0.054)	(0.102)	(0.043)	(0.089)	(0.044)	(0.044)
Level of education (Reference category: Low)							
Middle	1.047	0.890*	0.907	0.957	1.030	0.828***	0.926
	(0.079)	(0.055)	(0.075)	(0.054)	(0.083)	(0.050)	(0.054)
High	1.112	0.875*	0.737***	0.975	1.010	0.792***	0.862**
	(0.098)	(0.062)	(0.068)	(0.062)	(0.092)	(0.056)	(0.057)
Ability to meet financial needs (Reference category: with very difficulty)							
With some difficulty	0.993	0.902	1.213*	0.954	1.162	0.896	0.998
	(0.103)	(0.086)	(0.134)	(0.077)	(0.118)	(0.083)	(0.083)
Fairly easily	0.889	0.798**	0.927	0.954	1.223*	0.820**	1.021
	(0.095)	(0.077)	(0.104)	(0.080)	(0.130)	(0.077)	(0.087)
Easily	0.729**	0.738***	0.707***	0.839*	0.957	0.813*	1.132
	(0.094)	(0.085)	(0.095)	(0.086)	(0.127)	(0.091)	(0.120)
Health: EQ5D-5L score	1.002	1.010***	1.003	1.007***	0.999	1.006**	1.004*
	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.002)
Observations	7,040	7,040	7,040	7,040	7,040	7,040	7,040
Country indicators	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Other controls include indicator for vulnerable household members, follow COVID-19 related news, social media is a primary source of COVID-19 information, any COVID-19 positive person in social contacts and trust in information from government. Robust standard errors in parenthesis; ***p<0.001, **p<0.05, *p<0.1

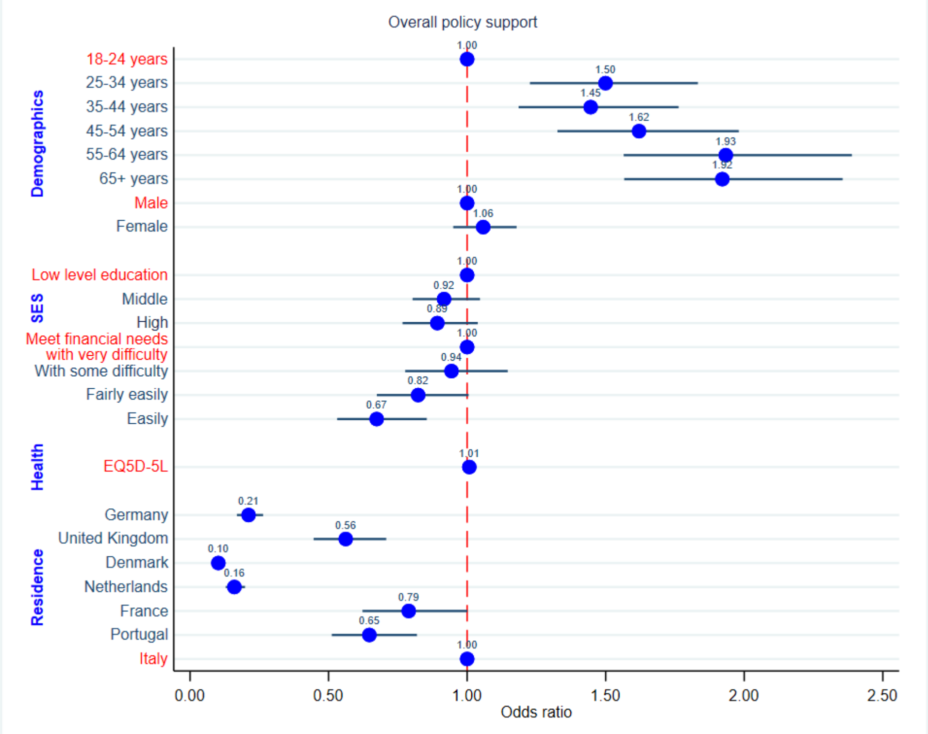
Table 3: Odds ratio (OR) for public worries regarding COVID-19 by socio-demographic and health determinants

VARIABLES	Losing someone		School closure		Health system overburdened		Economic worries*		Restricted access to food supply	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Age (Reference category: 18-24 years)										
25-34 years	1.103	1.104	1.123	1.124	1.267**	1.275**	1.679***	1.681***	1.346***	1.339***
	(0.113)	(0.113)	(0.105)	(0.105)	(0.132)	(0.133)	(0.169)	(0.169)	(0.128)	(0.127)
35-44 years	1.045	1.024	1.421***	1.436***	1.509***	1.477***	2.046***	2.036***	1.234**	1.219**
	(0.103)	(0.101)	(0.131)	(0.132)	(0.156)	(0.153)	(0.201)	(0.201)	(0.114)	(0.113)
45-54 years	1.101	1.093	1.239**	1.254**	1.720***	1.711***	2.213***	2.204***	1.209**	1.210**
	(0.111)	(0.110)	(0.115)	(0.117)	(0.183)	(0.181)	(0.220)	(0.219)	(0.112)	(0.112)
55-64 years	1.018	1.040	1.180*	1.187*	2.092***	2.138***	1.860***	1.862***	1.112	1.134
	(0.105)	(0.107)	(0.114)	(0.114)	(0.233)	(0.238)	(0.191)	(0.191)	(0.107)	(0.109)
65+ years	1.057	1.093	1.382***	1.383***	2.021***	2.099***	1.358***	1.363***	1.197*	1.235**
	(0.105)	(0.108)	(0.128)	(0.128)	(0.217)	(0.225)	(0.130)	(0.130)	(0.112)	(0.115)
Gender (Reference category: Male)										
Female	1.494***	1.498***	1.181***	1.193***	1.645***	1.639***	1.159***	1.154***	1.217***	1.226***
	(0.075)	(0.075)	(0.054)	(0.055)	(0.091)	(0.090)	(0.056)	(0.056)	(0.057)	(0.058)
Level of education (Reference category: Low)										
Middle	0.913	0.903*	0.811***	0.811***	1.099	1.090	1.066	1.065	0.843***	0.837***
	(0.056)	(0.055)	(0.045)	(0.045)	(0.075)	(0.074)	(0.063)	(0.063)	(0.048)	(0.048)
High	0.868**	0.853**	0.893*	0.892*	1.196**	1.183**	1.191***	1.190***	0.734***	0.725***
	(0.061)	(0.059)	(0.056)	(0.056)	(0.093)	(0.091)	(0.080)	(0.080)	(0.047)	(0.046)
Ability to meet household financial needs (Reference category: With very difficulty)										
With some difficulty	0.993	0.965	0.836**	0.825**	0.966	0.945	0.758***	0.759***	0.776***	0.754***
	(0.086)	(0.083)	(0.068)	(0.067)	(0.091)	(0.089)	(0.066)	(0.066)	(0.063)	(0.061)
Fairly easily	0.860**	0.832**	0.759***	0.735***	0.904	0.889	0.605***	0.612***	0.639***	0.610***
	(0.076)	(0.073)	(0.063)	(0.060)	(0.088)	(0.086)	(0.054)	(0.055)	(0.054)	(0.051)
Easily	0.666***	0.656***	0.593***	0.560***	0.652***	0.665***	0.464***	0.476***	0.468***	0.449***
	(0.072)	(0.071)	(0.061)	(0.058)	(0.075)	(0.078)	(0.050)	(0.051)	(0.049)	(0.047)
Health: EQ5D-5L score	1.016***		0.998		1.017***		1.002		1.013***	
	(0.002)		(0.002)		(0.003)		(0.002)		(0.002)	
Capability wellbeing-ICECAP score		1.010***		0.994***		1.014***		1.003*		1.005***
		(0.002)		(0.002)		(0.002)		(0.002)		(0.002)
Observations	7,040	7,040	7,040	7,040	7,040	7,040	7,040	7,040	7,040	7,040
Country indicators	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

*Economic worries are computed as the mean of worries reported on small companies running out of business, unemployment, and recession.
Other controls include indicator for vulnerable household members, follow COVID-19 related news, social media is a primary source of COVID-19 information and any COVID-19 positive person in social contacts. Robust standard errors in parentheses; ***p<0.01, **p<0.05, *p<0.1

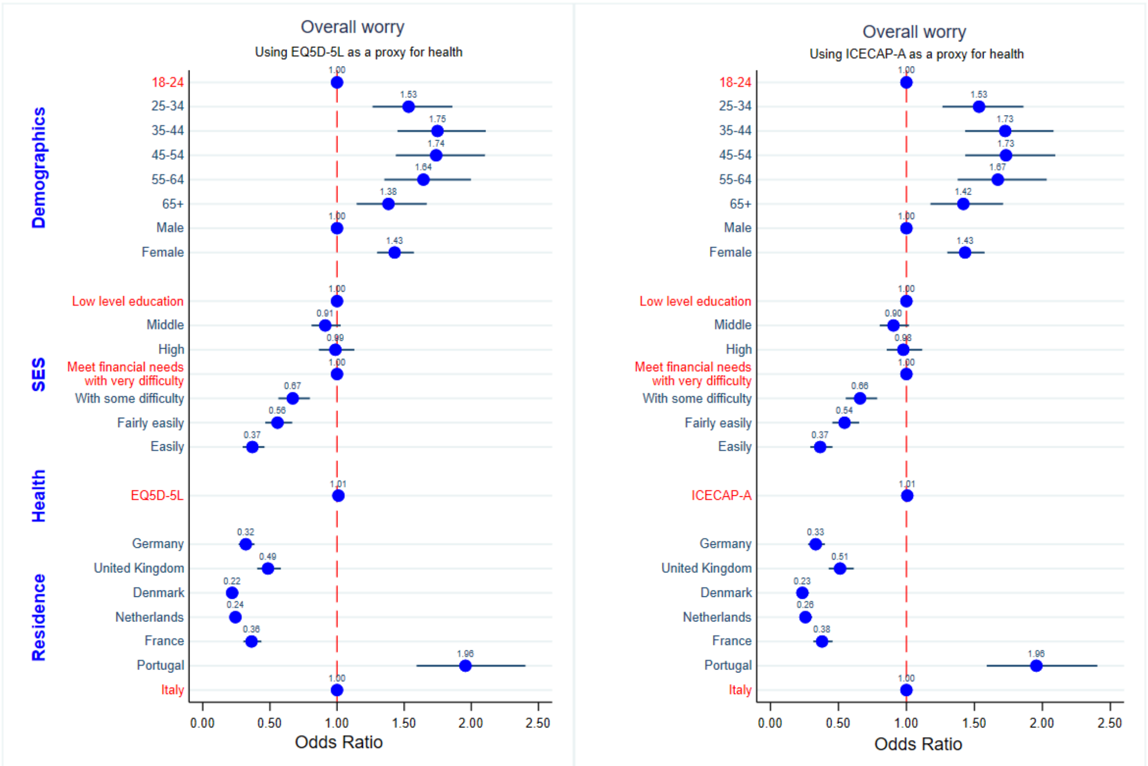
Supplementary tables

Figure S1: Odds ratio for overall policy support by socio-demographic and health determinants.



Note: Reference categories are in red. Other controls include indicator for vulnerable household members, intensity of following COVID-19 related news, social media is a primary source of COVID-19 information, any COVID-19 positive person in social contacts and trust in information from national government; N=7040

Figure S2: Odds ratio for overall worries by socio-demographic and health determinants.



Note: Reference categories in red. Other controls include an indicator for vulnerable household members, intensity of following COVID-19 related news, social media is a primary source of COVID-19 information and knowledge of COVID-19 positive case among social contacts. N=7040.

Table S1: Public support and worries by socio-demographic and health variables

		Age (mean)	Female (%)	Middle/high education (%)	Easily/fairly easily meet financial needs (%)	EQ5D-5L (mean)	ICECAP-A (mean)
Policy support	**Approve (%)**						
School closure	79.1	47.70 (16.54)	51.9	68.6	50.1	29.70 (11.37)	
Impose curfews	61.5	47.76 (16.45)	52.8	66.5	47.7	30.21 (11.83)	
Suspend public transportation	39.7	45.62 (16.26)	52.4	66.5	47.1	30.16 (11.87)	
Cancel public gatherings	83.0	47.70 (16.55)	53.3	68.5	50.3	29.82 (11.36)	
Restricted mobility*	59.2	47.52 (16.51)	53.0	66.5	47.6	30.19 (11.70)	
Close borders	82.6	47.59 (16.44)	53.7	68.1	49.3	29.86 (11.44)	
Ban export of medical device	49.9	49.23 (16.29)	50.4	67.4	49.2	30.36 (12.06)	
Regular temperature checks	60.4	47.27 (16.43)	51.5	66.0	48.6	29.84 (11.63)	
Use mobile phone to track mobility	55.8	48.55 (16.41)	51.0	67.5	50.9	29.93 (11.62)	
Overall policy support**							
Approve	66.6	48.12 (16.36)	52.5	67.0	48.6	30.00 (11.56)	
Indifferent	28.2	45.08 (16.89)	51.5	73.3	53.3	29.6 0 (11.37)	
Disapprove	5.2	45.15 (16.57)	44.9	75.8	56.2	29.90 (12.63)	
Worries	**Worried (%)**						
Losing someone I love	61.2	46.89 (16.42)	55.9	67.3	48.1	30.48 (11.69)	50.94 (14.62)
Health system overloaded	70.3	48.06 (16.31)	55.2	68.5	48.8	30.39 (11.54)	50.84 (14.38)
School closure	30.4	46.49 (16.09)	54.7	64.4	45.3	30.12 (11.81)	49.81 (14.87)
Small companies losing business	66.6	48.74 (16.27)	52.5	68.2	50.6	29.76 (11.25)	49.95 (14.30)
Recession	63.6	49.32 (15..82)	51.8	69.1	50.0	30.03 (11.29)	50.61 (14.31)
Becoming unemployed	36.7	41.93 (14.24)	56.5	65.1	41.2	29.34 (10.82)	51.41 (14.98)
Economic worries+	49.9	46.30 (15.42)	54.6	67.0	45.7	29.77 (11.02)	51.09 (14.56)
Restricted access to food supplies	44.3	46.61 (16.35)	55.1	64.3	46.0	30.88 (12.18)	51.26 (15.07)
Overall worries++							
Worried	47.7	46.19 (15.67)	56.4	64.8	44.8	30.36 (11.48)	51.50 (14.83)
Not very worried	37.5	47.76 (17.17)	50.4	73.6	53.5	29.96 (11.78)	49.08 (13.62)
Not worried	14.8	48.38 (17.74)	40.6	72.4	59.9	28.17 (11.14)	45.78 (14.72)

Note: Numbers shown in parenthesis are standard deviations

*Restricted mobility is computed from taking the mean of responses to three individual mobility related policies namely impose curfews, suspend public transportation and cancel public gatherings.

**Overall policy support is computed from taking the mean of responses to all nine individual policy support questions.

+Economic worries is computed from taking the mean of responses to three questions related to economic worries including small companies losing business, recession and being unemployed.

++ Overall worries are computed from taking the mean of responses to seven individual worry related questions.

Table S2: Odds ratio (OR) for public support for COVID-19 policies by household composition, trust in government, level and sources of knowledge regarding COVID-19, social contacts and country of residence.

VARIABLES	**School closures**	**Restrictions to mobility**	**Border closure**	**Ban on export of medical device**	**Fine for violating quarantine**	**Temperature checks**	**Mobile phone for tracking movement**
Vulnerable family members	0.843***	1.080	0.954	1.091*	1.063	1.042	1.107**
Level of trust in information from government							
Do not trust^	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Indifferent	1.273***	1.325***	1.153	1.099	1.563***	1.334***	1.666***
Trust well	2.250***	1.759***	1.686***	1.234***	2.944***	1.792***	2.309***
Follow COVID-19 news							
Very closely^	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Somewhat closely	0.615***	0.676***	0.640***	0.820***	0.600***	0.684***	0.644***
Not closely at all	0.302***	0.496***	0.312***	0.824**	0.253***	0.477***	0.607***
I don't know anything about it	0.132***	0.372***	0.163***	0.686	0.180***	0.455***	0.450***
COVID-19 infected in social contacts	0.997	1.016	0.859*	0.961	0.940	1.094	1.127**
Social media is one of the primary source of COVID-19 information	1.121*	1.262***	1.028	1.094*	0.979	1.223***	1.182***
Country of residence							
Italy^	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Germany	0.216***	0.271***	0.333***	0.648***	0.357***	0.174***	0.426***
United Kingdom	0.475***	0.716***	0.972	1.020	0.441***	0.409***	0.578***
Denmark	0.161***	0.150***	0.487***	0.336***	0.369***	0.101***	0.303***
Netherlands	0.274***	0.199***	0.393***	0.606***	0.297***	0.172***	0.281***
France	0.624***	0.984	1.143	1.127	0.706**	0.686***	0.664***
Portugal	1.047	0.648***	2.436***	0.545***	0.570***	0.716***	0.834*
Observations	7,040	7,040	7,040	7,040	7,040	7,040	7,040

^Reference category.

All regressions are controlled for socio-demographic and health indicators.

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table S3: Odds ratio (OR) for worries regarding COVID-19 by household composition, level and sources of knowledge regarding COVID-19, social contacts and country of residence.

VARIABLES	**Losing someone**	**School closure**	**Health system overburdened**	**Economic worries**	**Restricted access to food supply**
Vulnerable family members	1.291***	1.597***	1.067	1.010	1.208***
Follow COVID-19 news					
Very closely^	1.00	1.00	1.00	1.00	1.00
Somewhat closely	0.650***	0.818***	0.597***	0.678***	0.807***
Not closely at all	0.308***	0.826*	0.249***	0.369***	0.746***
I don't know anything about it	0.265***	1.247	0.219***	0.604*	0.605*
COVID-19 infected in social contacts	1.389***	1.300***	1.059	1.320***	0.982
Social media is one of the primary source of COVID-19 information	1.201***	1.366***	1.169**	1.354***	1.312***
Country of residence					
Italy^	1.00	1.00	1.00	1.00	1.00
Germany	0.412***	0.963	0.412***	0.314***	0.620***
United Kingdom	0.715***	0.677***	0.618***	0.324***	1.653***
Denmark	0.467***	0.688***	0.251***	0.217***	0.416***
Netherlands	0.288***	0.945	0.442***	0.247***	0.396***
France	0.509***	0.829**	0.677***	0.319***	0.719***
Portugal	2.000***	1.540***	1.177	1.502***	2.227***
Observations	7,040	7,040	7,040	7,040	7,040

^Reference category.

All regressions are controlled for socio-demographic and health indicators.

*** p<0.01, ** p<0.05, * p<0.1

Chapter 3.4

Once we have it, will we use it? A European Survey on willingness to be vaccinated against COVID-19.[28]

INTRODUCTION

While the focus of attention currently is on developing a vaccine against the Coronavirus SARS-CoV-2 to protect against the disease COVID-19, policymakers should prepare for the next challenge: uptake of the vaccine among the public. Having a vaccine does not automatically imply it will be used. Compliance with the anti-H1N1 vaccine during the 2009 influenza pandemic for instance was low (Blasi et al., 2012), and in the decade since, vaccination rates have remained an issue of concern (Kata, 2012) while vaccination hesitancy has become more prevalent, leading to increases in disease outbreaks in multiple countries (Larson et al., 2018). It is therefore important to understand whether or not people are willing to be vaccinated against COVID-19, as this can have large consequences for the success of a vaccination program – with potentially large health and economic consequences. In this chapter, we provide some first insights into this willingness to be vaccinated, based on a multi-country European study (Sabat et al., 2020), which hopefully result in more attention for this important issue.

A vaccine against COVID-19

On April 26, the WHO counted seven COVID-19 candidate vaccines in the clinical evaluation phase and 82 more in the preclinical evaluation phase (World Health Organization, 2020). This underlines the unprecedented current efforts worldwide to find an effective vaccine against the Coronavirus SARS-CoV-2. Some expect that first vaccines may become available under emergency use protocols as soon as early 2021, given the speed and scale of research and development efforts globally, while others argue it will take longer (Callaway, 2020; Tundzhay, 2020; Welcome trust, 2020). In both cases, the development phase should be followed by large scale vaccination programs to attain herd immunity (Fine et al., 2011).

[28] The paper constituting this chapter was published as an editorial in the '*European Journal of Health Economics*'. Author list as follows: Sebastian Neuman-Böhme, Nirosha Elsem Varghese, Iryna Sabat, Pedro Pita Barros, Werner Brouwer, Job van Exel, Jonas Schreyögg and Tom Stargardt

That way, we can protect the lives of the most vulnerable people and reduce the social and economic burden of the current crisis.

Vaccination programs can lead to herd immunity without requiring a large proportion of the population to be infected. The latter is mostly seen as an undesirable option, given the potentially high numbers of deaths as a result of infection. Especially so if the health systems are overwhelmed by a large number of patients with severe COVID-19 symptoms (D'Souza and Dowdy, 2020). Herd immunity through vaccination, however, requires a sufficient proportion of the population to be vaccinated. While vaccination is widely recognized as an effective way to reduce or eliminate the burden of infectious diseases by health authorities and the medical community (Andre et al., 2008), its effectiveness also depends on the individual willingness to be vaccinated. This willingness could be negatively affected by doubts and worries that exist in the population about the safety and appropriateness of vaccines. This is sometimes labelled vaccine hesitancy (Siciliani et al., 2020). If too many individuals hesitate about being vaccinated, herd immunity may not be reached. Besides objective trade-offs of costs and benefits of a vaccine, risk-attitude, pro-social considerations, and misinformation or misperceptions about a vaccine may play a role in this (Betsch et al., 2013; Kata, 2012; Korn et al., 2018).

At present, it is unclear whether a sufficient proportion of the population would decide to get vaccinated when a vaccine becomes available. In the EU, vaccine delays and refusals are contributing to declining immunization rates in several countries and lead to increases in disease outbreaks (Larson et al., 2018). Hence, and the question is whether enough Europeans trust the effectiveness and safety of vaccines and the healthcare system that delivers them (The Vaccine Confidence Project, 2015).

MATERIAL

In order to shed more light on the issue of willingness to be vaccinated, we investigated people attitudes about vaccination against COVID-19 in an online survey among representative samples of the population (in terms of region, gender, age-group and education) in seven European countries (N=7.662). The sample consisted of about 1.000 respondents per country, and an additional 500 from the highly affected region Lombardy, since we expected that results might differ from the rest of Italy. In this first wave of the data collection, respondents were inquired about worries and beliefs about COVID-19, as well as attitudes about vaccination and their willingness to be vaccinated between 2 and 15 April 2020 (Sabat et al., 2020). In this editorial, we provide some first insights into the findings, in order to stimulate further research and policy in this area.

RESULTS

In total, 73.9 % of the 7,664 participants from Denmark, France, Germany, Italy, Portugal, the Netherlands, and the UK stated that they would be willing to get vaccinated against COVID-19 if a vaccine would be available. A further 18.9% of respondents stated that they were not sure, and 7.2% stated that they don't want to get vaccinated. As shown in Figures 1 and 2, the willingness ranged from 62% in France to approx. 80% in Denmark and the UK. The largest proportions of the population opposed to a COVID-19 vaccination were observed in Germany (10%) and France (10%), while France also has the largest group of people who were unsure about getting vaccinated (28%).

[figures 1, 2 and 3 about here]

Looking closer, we found considerable differences in willingness to get vaccinated across genders and age groups (figure 3). A significantly higher proportion of men were willing to get vaccinated (77.94%, Chi-squared, $p < 0.001$) than women (70.15%). The willingness to be vaccinated is largest among men above the age of 55, while uncertainty ranged between 14-17% across all age groups. Males who were unwilling to get vaccinated tended to be younger with the largest share of 12% among the 18-24-year-olds. Similarly, the trend for women who were unwilling to vaccinate seems also to follow the age categories. The uncertainty among women was higher in all age groups and largest for women between the ages of 45 to 54 (26%).

One might argue that the group who is currently unsure about getting a vaccine may be the most relevant. These are the people who potentially can be persuaded more easily to get vaccinated to achieve herd immunity. Based on our results, these efforts could best be aimed at persons below the age of 55 and at females in general, where the willingness is lower.

We asked respondents who were unsure about being vaccinated about their main reasons (Figure 4). More than half (55%) said they were concerned about potential side effects of a vaccine, although this concern was more frequent among women (36%) than men (19%). Around 15% of respondents stated that a vaccine might not be safe, with no notable differences between genders. These findings are in the literature on frequent reasons for vaccine hesitancy (The Vaccine Confidence Project, 2015). Looking at the open text explanations given to the category "other", we saw that a common concern seems to be that a COVID-19 vaccine might be experimental, without any studies on side-effects, and that the vaccine may not be safe for specific groups, such as for pregnant woman, people with pre-existing conditions like MS, allergic persons etc.

[Figure 4 about here]

This finding highlights that while the current focus seems to be on developing a vaccine about ten times faster than usual (Welcome trust, 2020), the public should also be reassured that any vaccine that becomes available that quickly is safe and effective. Otherwise, there is a risk to lose the public trust in the particular vaccine, and coronavirus vaccination altogether (Jiang, 2020), potentially compromising herd immunity.

[Figure 5 about here]

We find a similar trend regarding the most frequently mentioned reasons and the gender differences for the concerns about side effects among those who were not willing to get vaccinated. Notable gender differences could also be observed among those respondents who stated that they think COVID-19 is not dangerous to their health (11%), comprised of almost twice as many men (7%) than women (4%). Furthermore, we see that an overall rejection of vaccination was more than twice as common among women (7%) than among men (3%). When looking at the open text answers of respondents who choose other reasons (11%), we found concerns about safety but also comments about conspiracy theories and a general rejection of vaccines.

DISCUSSION

The literature suggests multiple steps that could be taken by policymakers to decrease vaccine hesitancy and convince doubters to get vaccinated after all. One approach for vaccine advocacy suggests "vaccine adoption = access + acceptance" (Thomson and Watson, 2012). Looking at access, it is essential to translate the willingness to be vaccinated into actual vaccination decisions. Our study measured the intention to vaccinate; this rate might differ from actual vaccination uptake (vaccination decision) depending on potential constraints, such as the price of the vaccine and the ease of access of vaccination sites. Vaccines should thus be available in a timely manner and an easily accessible way to have as little attrition as possible (Siciliani et al., 2020). In the case of the coronavirus vaccine, access will prove quite challenging since, at the early stages of availability, the demand for this vaccine worldwide will be much greater than the (short term) production capacities. Currently, about 5 billion doses of vaccine are produced yearly worldwide, of which 30% are seasonal flu vaccines (The Economist, 2020). So even when a vaccine becomes available, access to it will probably be limited in the short-run. Therefore, policymakers need to prepare how access can be organized equitably and effectively.

Our results on acceptability suggest that substantial gains could be made among the sizeable proportion of the population (i.e. 18.9%) that is unsure whether they want to get vaccinated. If this group needs to be convinced to be vaccinated to get to herd immunity, clear communication about safety, and potential side-effects of the vaccine is especially important. This could help stimulate the hesitant part of European citizens to get vaccinated after all.

This is especially important since it is unclear whether the group of people who are willing to be vaccinated in itself is large enough to achieve herd immunity. The basic reproduction number R_0 shows the transmission potential of diseases (Rothman et al., 2008), i.e., to how many people the infection is expected to be passed on by one infected individual in a fully susceptible population, on average. The herd immunity threshold describes the proportion of the population that needs to be immune, so that the infectious diseases is stable (R=1) and is calculated as (Nishiura and Chowell, 2009):

$$Herd\ immunity\ threshold = 1 - \frac{1}{R_0}$$

This means that the higher the basic reproductive number R_0 is, the higher the herd immunity threshold becomes. A recent study estimated a COVID-19 R_0 of around 3.87 for Europe (Flaxman et al., 2020), implying a herd immunity threshold for Europe of 74%. For the US, it was estimated at around 3.45, implying a herd immunity threshold of 71% (Pitzer et al., 2020), while a recent study argues these values may be lower if there is heterogeneity in the individual susceptibility to the virus (Gomes et al., 2020). Of course, these estimates are uncertain, but comparing this 71-74% threshold range with our results indicates that the current willingness levels in France, Germany and the Netherlands, in particular, may prove insufficient to reach this threshold.

Our survey highlighted important differences between citizens from European countries in terms of willingness to be vaccinated against COVID-19. The levels don't follow trends that we see in other vaccination rates, e.g. against measles, which are generally higher, but in most countries below the recommended 95% threshold (World Health Organization, 2019).

Understanding which groups in the population are not willing to be vaccinated and why remains vital for the design of policy responses to vaccination hesitancy. One of the avenues to explore could be to emphasize the social benefits of vaccination more strongly so that they weigh the public health dimension more heavily in their decision whether to vaccinate (Betsch et al., 2013). A recent study, for example, found that people are more willing to get vaccinated when they were informed that this would protect others who have are willing but unable to get vaccinated themselves (Böhm et al., 2019). Consequently, one of the communication strategies could be to emphasize how vaccination against COVID-19 helps to protect vulnerable members of society. Furthermore, the distribution of vaccinated individuals in the population matters, pockets of non-vaccinated groups could be highly problematic even when overall vaccination rates are high. Unvaccinated individuals may be in contact with other unvaccinated individuals relatively often (Barclay et al., 2014). Outbreak in particular communities may then occur, even if overall vaccination rates are high. Examples of measles outbreaks in the Netherlands (van den Hof et al., 2001) and the US (Centers for Disease Control and Prevention, 2020), for instance highlight the role of religious communities and travelers in this context.

Alternative strategies range from restrictive measures against those who chose not to be vaccinated to mandatory vaccination schemes for certain target groups or the whole population. Experimental evidence suggests that individuals under specific conditions may be willing to support mandatory vaccination policies, but this support seems very sensitive to adverse events (Meier et al., 2020). Such a policy may be less appropriate in the context of COVID-19.

CONCLUSION

Our findings highlight that considerable policy effort may be required to come from having a vaccine to adequate vaccination rates, especially in some countries. Targeting those in the population who are currently hesitant seems most promising and cost-effective, but this requires convincing evidence and clear communication on the safety and effectiveness of the vaccine. This may be at odds with the current push for having a vaccine available as soon as possible. A campaign emphasizing the social benefits of vaccination could increase the willingness to be vaccinated among those amenable to such pro-social motives. Finally, a sizeable proportion of the population indicates not to be open to vaccination. This group may remain at risk of spreading the virus and contracting the disease, even after herd immunity has been achieved. Concluding, improving our understanding of vaccination hesitancy in the context of COVID-19, as well as finding and using policies to overcome it, may be as important as discovering a safe and effective vaccine.

BIBLIOGRAPHY

Andre, F. E., *et al.* 2008. Vaccination greatly reduces disease, disability, death and inequity worldwide. *Bulletin of the World health organization,* 86, 140-146.

Barclay, V. C., *et al.* 2014. Positive network assortativity of influenza vaccination at a high school: implications for outbreak risk and herd immunity. *PloS one,* 9(2), e87042.

Betsch, C., Böhm, R. and Korn, L. 2013. Inviting free-riders or appealing to prosocial behavior? game-theoretical reflections on communicating herd immunity in vaccine advocacy. *Health Psychology,* 32(9), 978.

Blasi, F., *et al.* 2012. Compliance with anti-H1N1 vaccine among healthcare workers and general population. *Clinical Microbiology and Infection,* 18, 37-41.

Böhm, R., *et al.* 2019. The willingness to vaccinate increases when vaccination protects others who have low responsibility for not being vaccinated. *Journal of behavioral medicine,* 42(3), 381-391.

Callaway, E. 2020. CORONAVIRUS VACCINES. *Nature,* 580, 577.

Centers for Disease Control and Prevention, 2020. Measles Cases and Outbreaks. [online]. Available from: https://www.cdc.gov/measles/cases-outbreaks.html.

D'Souza, G. and Dowdy, D. 2020. What's herd immunity and how can we achieve it with COVID-19. *Johns Hopkins School of Public Health Expert Insights*.

Fine, P., Eames, K. and Heymann, D. L. 2011. "Herd immunity": a rough guide. *Clinical infectious diseases,* 52(7), 911-916.

Flaxman, S., *et al.* 2020. Estimating the number of infections and the impact of non-pharmaceutical interventions on COVID-19 in European countries: technical description update. *arXiv preprint arXiv:2004.11342*.

Gomes, M. G. M., *et al.* 2020. Individual variation in susceptibility or exposure to SARS-CoV-2 lowers the herd immunity threshold. *MedRxiv*.

Jiang, S. 2020. Don't rush to deploy COVID-19 vaccines and drugs without sufficient safety guarantees. *Nature,* 579(7798), 321-322.

Kata, A. 2012. Anti-vaccine activists, Web 2.0, and the postmodern paradigm–An overview of tactics and tropes used online by the anti-vaccination movement. *Vaccine,* 30(25), 3778-3789.

Korn, L., *et al.* 2018. Social nudging: The effect of social feedback interventions on vaccine uptake. *Health Psychology,* 37(11), 1045.

Larson, H., *et al.* 2018. State of vaccine confidence in the EU 2018. *Luxembourg: Publications Office of the European Union,* 10, 241099.

Meier, N. W., *et al.* 2020. Individual preferences for voluntary vs. mandatory vaccination policies: an experimental analysis. *European Journal of Public Health,* 30(1), 50-55.

Nishiura, H. and Chowell, G., 2009. The effective reproduction number as a prelude to statistical estimation of time-dependent epidemic trends. *Mathematical and statistical estimation approaches in epidemiology.* Springer, 103-121.

Pitzer, V. E., *et al.* 2020. The impact of changes in diagnostic testing practices on estimates of COVID-19 transmission in the United States. *MedRxiv*.

Rothman, K. J., Greenland, S. and Lash, T. L., 2008. *Modern epidemiology.* Lippincott Williams & Wilkins.

Sabat, I., *et al.* 2020. Countering COVID-19: A European survey on acceptability of and commitment to preventive measures. *European Journal of Public Health*, 30(Supplement_5), ckaa166. 619.

Siciliani, L., *et al.* 2020. Strengthening vaccination programmes and health systems in the European Union: A framework for action. *Health Policy*, 124(5), 511-518.

The Economist, 2020. Can the world find a good covid-19 vaccine quickly enough? [online]. Available from: https://www.economist.com/briefing/2020/04/16/can-the-world-find-a-good-covid-19-vaccine-quickly-enough.

The Vaccine Confidence Project, 2015. The State of Vaccine Confidence. [online]. Available from: https://www.vaccineconfidence.org/s/VCP_The-State-of-Vaccine-Confidence_2015.pdf [Accessed 26 March 2020].

Thomson, A. and Watson, M., 2012. Listen, understand, engage. American Association for the Advancement of Science.

Tundzhay, B., 2020. Covid-19 vaccine availability unrealistic in 12-18 month period. *Pharma Technology Focus* [online]. Available from: https://www.pharmaceutical-technology.com/comment/covid-19-vaccine-availability/ [Accessed 26 April 2020].

van den Hof, S., *et al.* 2001. Measles outbreak in a community with very low vaccine coverage, the Netherlands. *Emerging infectious diseases*, 7(3 Suppl), 593.

Welcome trust, 2020. How can we develop a COVID-19 vaccine quickly? . [online]. Available from: https://wellcome.ac.uk/news/how-can-we-develop-covid-19-vaccine-quickly.

World Health Organization, 2019. Measles: fighting a global resurgence. [online]. Available from: https://www.who.int/news-room/feature-stories/detail/measles-fighting-a-global-resurgence [Accessed 20 April, 2020].

World Health Organization, 2020. Draft of the landscape of COVID-19 candidate vaccines. [online]. Available from: https://www.who.int/blueprint/priority-diseases/key-action/novel-coronavirus-landscape-ncov.pdf?ua=1.

FIGURES

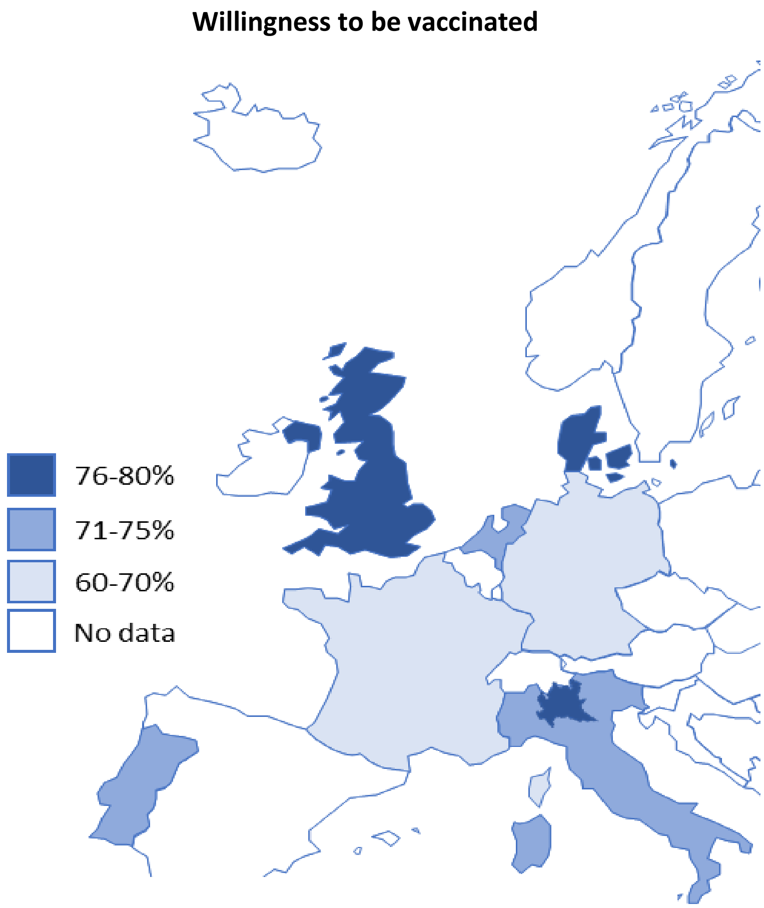


Figure 1: Proportion of respondents who stated they would be willing to be vaccinated against the novel coronavirus per country

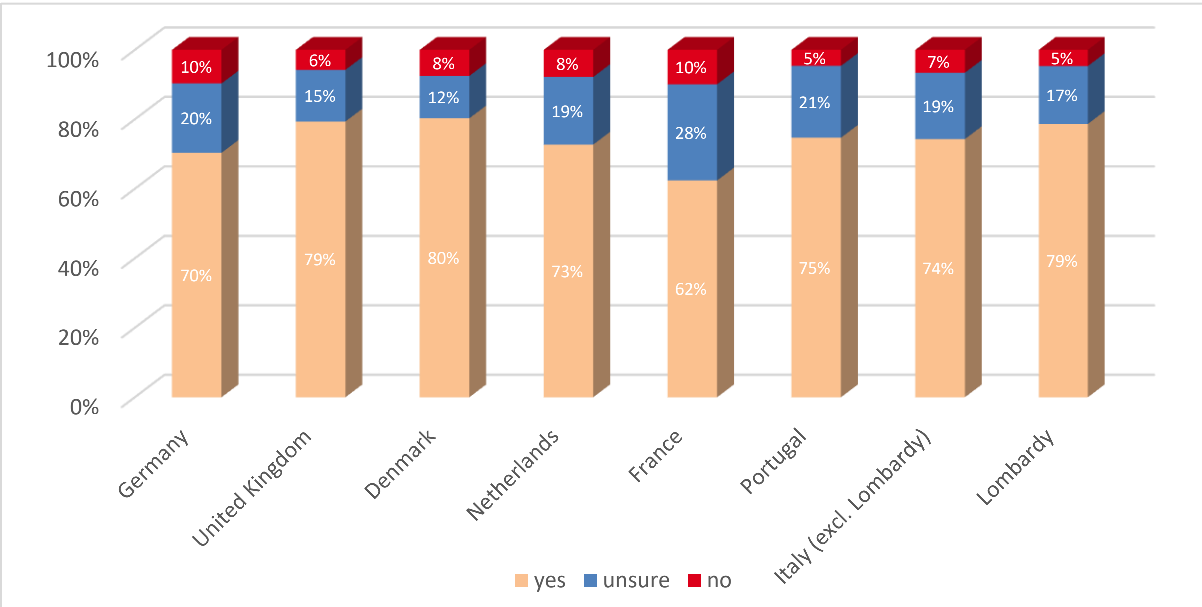


Figure 2: Willingness to be vaccinated against the coronavirus by country

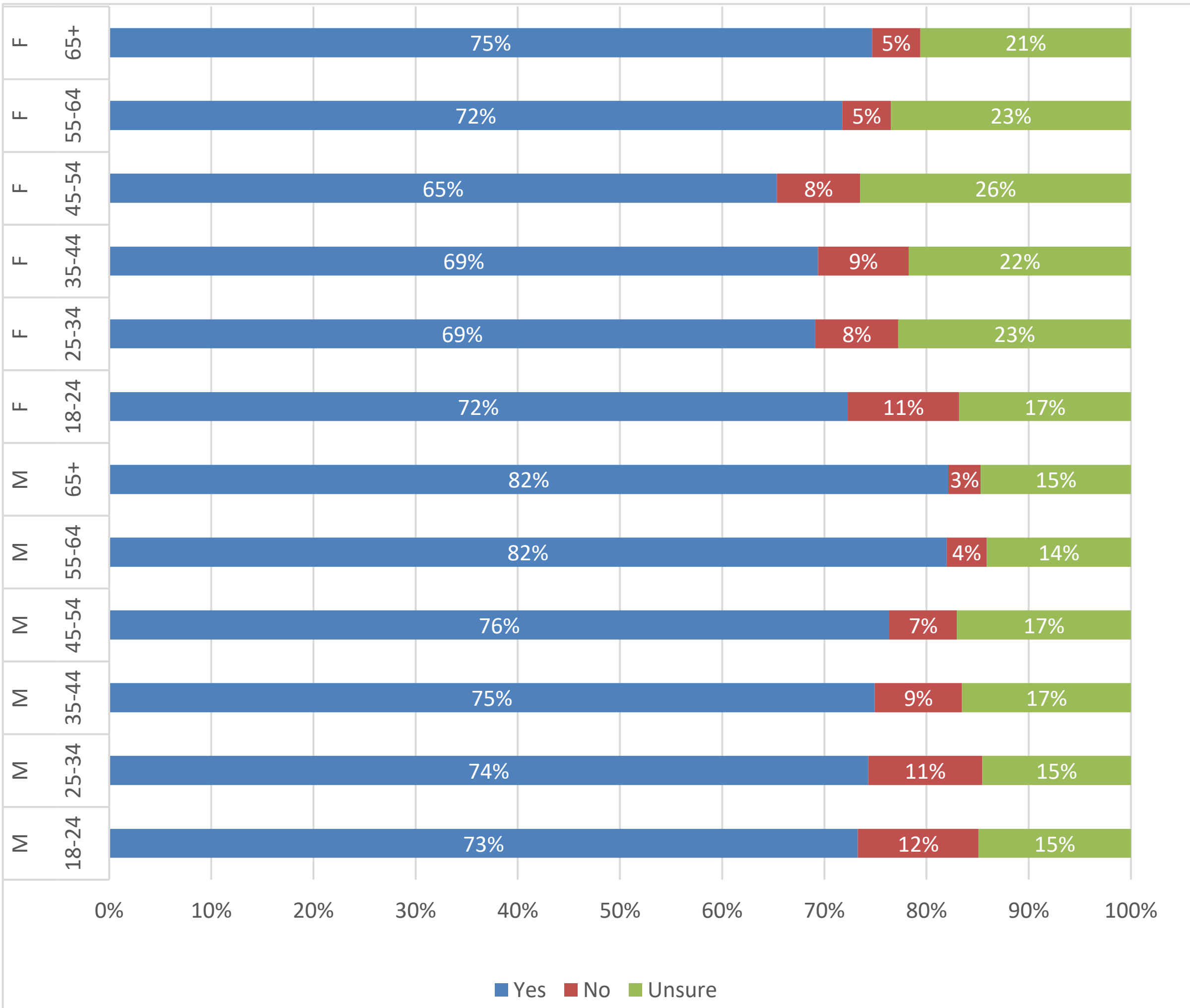


Figure 3:Willingness to be vaccinated against COVID-19 by age group and gender

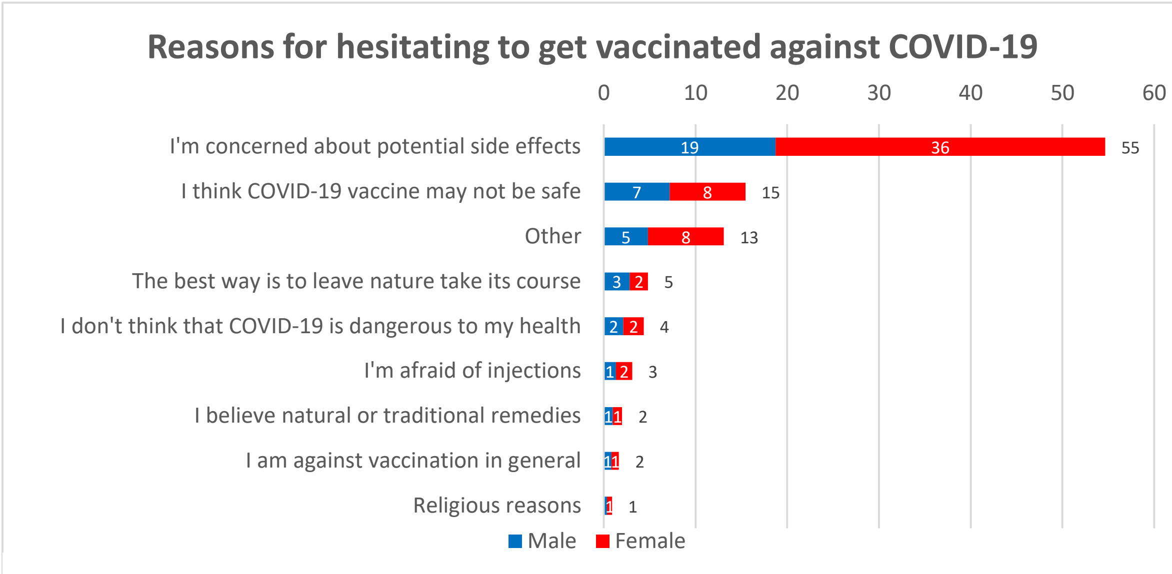


Figure 4: Reasons given by people who were unsure if they would like to be vaccinated against COVID-19 in percent, N=1451

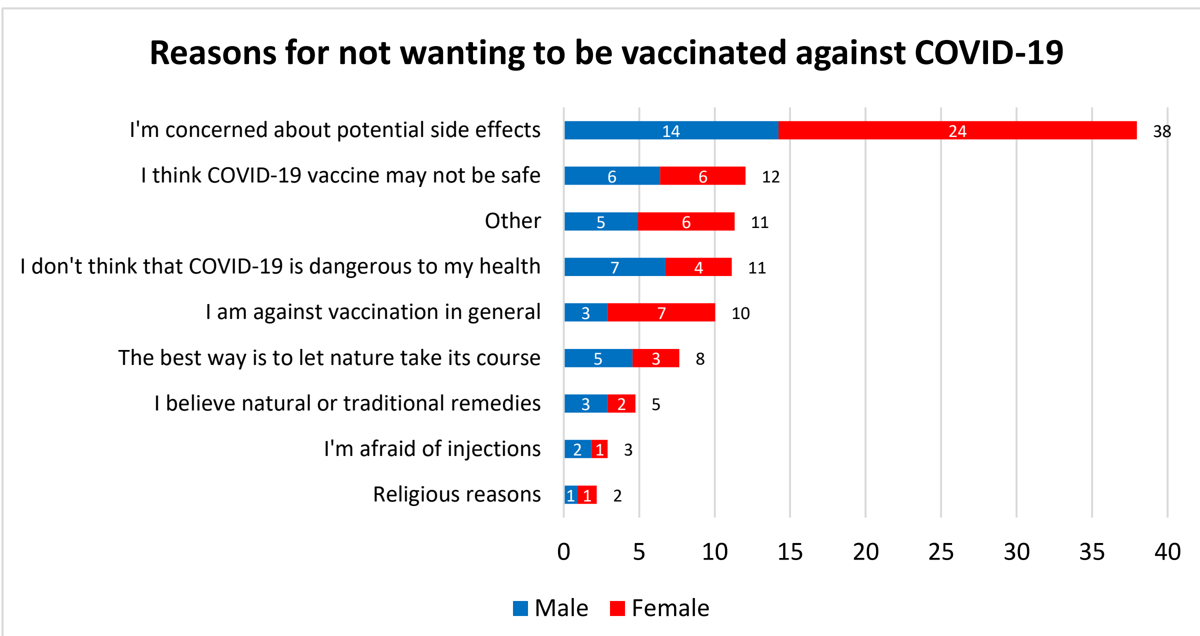


Figure 5: Reasons for not getting vaccinated against COVID-19 in percent, N=548

Appendix: ECOS Wave-1 Survey Questionnaire

Consent form

Dear Participant,
Thank you for participating in this survey on the novel coronavirus (COVID-19). The goal of this study is to understand people's attitudes towards the disease and the associated risks. Please answer the survey to the best of your knowledge and abilities.

Before you start, please:

- Make sure you have about 20 minutes of uninterrupted time;
- Maximize your browser window;
- Switch your phone to a silent mode;
- Switch off your e-mail, phone notifications and anything else that may distract you.

Please, do not use external sources of information like the Internet to search for information. Many people may not know the answers to some questions, but please answer every question according to your belief if you are not sure what the right answer is.

Your data will be treated in accordance with the provisions of the European Data Protection Regulation (GDPR EU).

CONSENT FORM
I consent to participate in this survey. I understand that all data will be kept confidential by the researcher. My personal information will not be stored with the data. I am free to withdraw at any time, without giving a reason.

○ I voluntarily consent to participate in this study.

○ I do not wish to participate.

Block: Demographics

How old are you currently?

What is your gender?

- ○ Male
- ○ Female

Does your household include any of the following members (other than you)?

- ☐ Very young children and babies
- ☐ Children
- ☐ Disabled person(s)
- ☐ Someone with diagnosed chronic medical conditions (such as heart or lung conditions or diabetes)
- ☐ Elderly person(s)
- ☐ ⊗None of the ones mentioned above

Block: EQ5D-5L
Now, we would like you to rate your own health.
Under each heading, please click the ONE box that best describes your health TODAY.

MOBILITY

- ○ I have no problems in walking about (1)
- ○ I have slight problems in walking about (2)
- ○ I have moderate problems in walking about (3)
- ○ I have severe problems in walking about (4)
- ○ I am unable to walk about (5)

SELF-CARE

- ○ I have no problems washing or dressing myself (1)
- ○ I have slight problems washing or dressing myself (2)
- ○ I have moderate problems washing or dressing myself (3)
- ○ I have severe problems washing or dressing myself (4)
- ○ I am unable to wash or dress myself (5)

USUAL ACTIVITIES *(e.g. work, study, housework, family or leisure activities)*

- ○ I have no problems doing my usual activities (1)
- ○ I have slight problems doing my usual activities (2)
- ○ I have moderate problems doing my usual activities (3)
- ○ I have severe problems doing my usual activities (4)
- ○ I am unable to do my usual activities (5)

PAIN / DISCOMFORT

- ○ I have no pain or discomfort (1)
- ○ I have slight pain or discomfort (2)
- ○ I have moderate pain or discomfort (3)
- ○ I have severe pain or discomfort (4)
- ○ I have extreme pain or discomfort (5)

ANXIETY / DEPRESSION

- ○ I am not anxious or depressed (1)
- ○ I am slightly anxious or depressed (2)
- ○ I am moderately anxious or depressed (3)
- ○ I am severely anxious or depressed (4)
- ○ I am extremely anxious or depressed (5)

Block: ICECAP-A

We would like to know about your overall quality of life.
Please indicate which statements best describe your overall quality of life at the moment

Feeling settled and secure

- ○ I am able to feel settled and secure in **all** areas of my life (1)
- ○ I am able to feel settled and secure in **many** areas of my life (2)
- ○ I am able to feel settled and secure in **a few** areas of my life (3)
- ○ I am **unable** to feel settled and secure in **any** areas of my life (4)

Love, friendship and support

- ○ I can have **a lot** of love, friendship and support (1)
- ○ I can have **quite a lot** of love, friendship and support (2)
- ○ I can have **a little** love, friendship and support (3)
- ○ I **cannot** have **any** love, friendship and support (4)

Being independent

- ○ I am able to be **completely** independent (1)
- ○ I am able to be independent in **many** things (2)
- ○ I am able to be independent in **a few** things (3)
- ○ I am **unable** to be at all independent (4)

Achievement and progress

- I can achieve and progress in **all** aspects of my life (1)
- I can achieve and progress in **many** aspects of my life (2)
- I can achieve and progress in **a few** aspects of my life (3)
- I **cannot** achieve and progress in **any** aspects of my life (4)

Enjoyment and pleasure

- I can have **a lot** of enjoyment and pleasure (1)
- I can have **quite a lot** of enjoyment and pleasure (2)
- I can have **a little** enjoyment and pleasure (3)
- I **cannot** have **any** enjoyment and pleasure (4)

Do you know people in your immediate social environment who are or have been infected with the novel coronavirus?

- Yes, confirmed (1)
- Yes, but not yet confirmed (2)
- No (3)
- Don't know (4)

How closely have you been following the news about the COVID-19 outbreak?

- I don't know anything about it (1)
- Not closely at all (2)
- Somewhat closely (3)
- Very closely (4)

What is the main source of information from where you get information on COVID-19? (several answers possible)

- ☐ TV (1)
- ☐ Internet search (2)
- ☐ Social media (3)
- ☐ Newspapers (4)
- ☐ Relatives and friends (5)
- ☐ ⊗I don't follow any (6)
- ☐ Other sources (7)

Block: Policy support

Next, we would like to ask you about your opinion on different governmental policies that could be/were taken to contain the spread of the novel coronavirus.

Please indicate on the scale below to which extent you approve or disapprove the following government measures related to COVID-19 outbreak.

	Strongly disapprove (1)	Disapprove (2)	Indifferent (3)	Approve (4)	Strongly approve (5)
Closing schools and universities for three months (1)	○	○	○	○	○
Fine people who violate the 14 days home quarantine knowingly (2)	○	○	○	○	○
Closing all borders to deny entry to foreign travellers for three months (3)	○	○	○	○	○
Impose a curfew (allowed to go out only to buy groceries/medicine) for three months (4)	○	○	○	○	○
Random and regular temperature checks on the streets (5)	○	○	○	○	○
Ban on export of medical equipment (e.g. masks) from your country (6)	○	○	○	○	○

Suspend public transport for three months (7)	○	○	○	○	○
Suspend all public gatherings (e.g. concerts, religious services, cinemas) for three months (8)	○	○	○	○	○
Use mobile phone data for tracking people infected with coronavirus and others they had contact with to identify and quarantine them (9)	○	○	○	○	○

Block: Worry

Crises often involve fears and worries. Please let us know:
At the moment, how much do you worry about:

	Do not worry at all				Worry a lot
	1	2	3	4	5
losing someone I love					
health system being overloaded					
school closures					
small companies running out of business					
recession					
restricted access to food supplies					
blackouts					
society getting more egoistic ()					
becoming unemployed ()					

Block: Vaccination

Researchers are currently working hard to find a safe and effective vaccine against the novel coronavirus. For the following questions please imagine that the vaccine became available

Would you be willing to get vaccinated against the novel coronavirus?

- ○ Yes (1)
- ○ No (2)
- ○ Not sure (3)

Please, explain why you wouldn't consider taking a vaccination or why you are hesitant? (more than one answer can be selected)

- ▢ I don't think that COVID-19 is dangerous to my health (1)
- ▢ I think COVID-19 vaccine may not be safe enough (2)
- ▢ I am against vaccination in general (3)
- ▢ I believe natural or traditional remedies can treat COVID-19 (4)
- ▢ The best way is to leave nature take its course (5)
- ▢ I'm afraid of injections (6)
- ▢ I'm concerned about potential side effects (7)
- ▢ Religious reasons (8)
- ▢ Other (10)

Block: Trust

On the scale below, please indicate to which extent you trust the information from the following sources in the context of COVID-19 situation.

	Not at all			Very much	
	1	2	3	4	5
Your national government					
The European Union					
Main national news channels / newspapers					
Social media (Facebook, Twitter, Instagram)					
Hospitals					
General practitioner/ Family doctor					
World Health Organization (WHO)					
Your relatives and friends					

Block: Familiarity and compliance with WHO recommendations

Now we would like you to carefully read the 5 basic protective measures against the novel coronavirus recommended by the World Health Organization (WHO).

How familiar were you with these basic protective measures against the novel coronavirus before you saw the poster?

- ○ Not at all familiar
- ○ Slightly familiar
- ○ Somewhat familiar
- ○ Moderately familiar
- ○ Very familiar

Next, we would like to know about your own practices related to the novel coronavirus.
Thinking about the last four weeks, did you adhere to the following activities due to concerns about the novel coronavirus?

	No	Yes, a bit	Yes, quite strongly	Yes, fully
Regularly wash my hands with soap for at least 20 seconds	○	○	○	○
Cover my nose and mouth when coughing or sneezing	○	○	○	○
Keep distance of at least 1 meter from other people	○	○	○	○
Avoid shaking hands, hugging or kissing when greeting others	○	○	○	○
Use alcohol-based hand rub	○	○	○	○
Avoid touching my nose, eyes and mouth	○	○	○	○

In your opinion, do <u>others in your community</u> adhere to the WHO basic protective measures against the novel coronavirus these days?

	No	Yes, a bit	Yes, quite strongly	Yes, fully
Regularly wash their hands with soap for at least 20 seconds	○	○	○	○
Cover their nose and mouth when coughing or sneezing	○	○	○	○
Keep distance of at least 1 meter from other people	○	○	○	○
Avoid handshakes, kisses and hugs when greeting others	○	○	○	○
Use alcohol-based hand rub	○	○	○	○
Avoid touching their eyes, nose and mouth	○	○	○	○

****