The Life with Corona survey

Wolfgang Stojetz\textsuperscript{a,b,*}, Neil T.N. Ferguson\textsuperscript{a}, Ghassan Baliki\textsuperscript{c}, Oscar Díaz\textsuperscript{a}, Jan Elses\textsuperscript{a}, Damir Esenaliev\textsuperscript{a,c}, Hanna Freudenreich\textsuperscript{c}, Anke Koebach\textsuperscript{d,e}, Liliana Abreu\textsuperscript{d}, Laura Peitz\textsuperscript{a}, Ani Todua\textsuperscript{a,c}, Monika Schreiner\textsuperscript{c}, Anke Hoeffler\textsuperscript{d}, Patrícia Justino\textsuperscript{f}, Tilman Brück\textsuperscript{a,c,g}

\textsuperscript{a} ISDC – International Security and Development Center, Augustastr. 89, 10117 Berlin, Germany
\textsuperscript{b} Department of Economics, Humboldt University of Berlin, Spandaue Str. 1, 10178 Berlin, Germany
\textsuperscript{c} Leibniz Institute of Vegetable and Ornamental Crops, Thedor-Eckermeyer-Weg 1, 14979 Großbeeren, Germany
\textsuperscript{d} Department of Politics and Public Administration, University of Konstanz, Universitätsstr. 10, 78464 Konstanz, Germany
\textsuperscript{e} Department of Psychology, University of Konstanz, Universitätsstr.10, 78464 Konstanz, Germany
\textsuperscript{f} World Institute for Development Economic Research, United Nations University, Katajanokanlanturi 6 B, FI-00160 Helsinki, Finland
\textsuperscript{g} Natural Resources Institute, University of Greenwich, Medway Campus, Central Avenue Chatham Maritime, Kent ME4 4TB, UK

\textbf{A B S T R A C T}

The COVID-19 pandemic is a global crisis affecting everyone. Yet, its challenges and countermeasures vary significantly over time and space. Individual experiences of the pandemic are highly heterogeneous and its impacts span and interlink multiple dimensions, such as health, economic, social and political impacts. Therefore, there is a need to disaggregate “the pandemic” analyzing experiences, behaviours and impacts at the micro level and from multiple disciplinary perspectives. Such analyses require multi-topic pan-national survey data that are collected continuously and can be matched with other datasets, such as disease statistics or information on countermeasures. To this end, we introduce a new dataset that matches these desirable properties - the Life with Corona (LwC) survey - and perform illustrative analyses to show the importance of such micro data to understand how the pandemic and its countermeasures shape lives and societies over time.

1. Introduction

The COVID-19 pandemic has affected everyone, but challenges and countermeasures vary significantly over time and space. Individual experiences are far from uniform (Patel et al., 2020; Benzeval et al., 2020) with impacts present in multiple dimensions, such as (mental) health (Pfefferbaum and North, 2020; Nathiya et al., 2020; Kim and Jung, 2021; Abreu et al., 2021), and economic (Bloom et al., 2020; Bottan et al., 2020; Brodeur et al., 2020), social (Jin et al., 2021; Brück et al., 2020) and political outcomes (Barrios and Hochberg, 2020; Kerr et al., 2021). Significant high-quality data have been collected to provide understanding of the pandemic but data that can connect the dots between places, domains and points in time are so far lacking. Most surveys focus on single issues (e.g. Aristovnik et al., 2020; on education; Kpan et al., 2021, on behaviour; Lazarus et al., 2020 on attitudes to government). Some others collect data in multiple domains but only in narrow time windows, such as Fetzer et al. (2020). A third important set of surveys focuses on specific places (e.g. Blom et al., 2020 in Germany; Heuvel et al., 2020, USA; Varshney et al., 2020, India). The addition of COVID-specific questions into long-standing panels, such as Understanding Society and the German Socio-Economic Panel will provide important long-term insights but do not capture the earliest phases of the pandemic. In this article, we introduce Life with Corona (LwC), a new, public survey dataset that collects data in multiple domains, across space, in continuous time since the earliest phases of the pandemic. We present illustrative analyses that show, both, the need for such data and the power of our data to answer questions of this sort.

2. Desirable survey properties

Given the comprehensive nature of the pandemic, we note an urgent demand for social science surveys that facilitate micro-level research across domains, time, and exposure rates. Specifically, this implies the need for a survey with the following desirable properties: (C1) collected throughout the pandemic, including before, during and after lockdowns and peaks in infections; (C2) composed of large numbers of observations, between and within countries; (C3) captures COVID-19 exposure and experiences across multiple domains; (C4) collected from different socio-demographic groups; (C5) allows matching with secondary data, such as data on the spread of the disease or social instability data. To our knowledge, the LwC Survey is currently the only survey to satisfy these criteria.

* Corresponding author. ISDC – International Security and Development Center, Augustastr. 89, 10117, Berlin, Germany.
E-mail address: stojetz@isd.org (W. Stojetz).

https://doi.org/10.1016/j.socscimed.2022.115109
Received 6 December 2021; Received in revised form 30 May 2022; Accepted 31 May 2022
Available online 4 June 2022
0277-9536/© 2022 Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND IGO license (http://creativecommons.org/licenses/by-nc-nd/3.0/igo/).
3. The Life with Corona survey

Life with Corona (LwC) is a multi-year research project (see: https://lifewithcorona.org). The project provides online survey data that were continuously collected in real time during the COVID-19 pandemic. It is led by an international consortium of researchers, supported by a worldwide network of collaborating institutions (see: https://lifewithcorona.org/network). The online survey is based on a publicly-accessible online questionnaire that can be answered in 27 languages and was launched in March 2020 (fulfilling C1). On October 1, 2020, some modules were slightly adapted, starting “round 2” of data collection. Similarly, the questionnaire was updated one more time on April 29, 2021, starting “round 3” of data collection. At the time of submission, the dataset contains 39,449 observations (including 32,958 pure cross-sectional and \( n \times T = 6491 \) panel observations), collected from 167 countries (C2). The majority of our responses (60%) come from Germany. However, there are a significant number of observations from Argentina (1701), Portugal (1309), USA (1089) and the UK (870). In total, there are 21 countries with more than 150 observations in the data. For more details on the number of responses by country and survey round see: Table 1.

The survey covers three themes: Health; Economy; and State & Society (C3). Given the diversity of experiences of the pandemic across individuals, a particular emphasis is placed on measuring “SARS-CoV-2 exposure”, including COVID-19 experiences, perceptions of the pandemic, and behaviours to counter it. The three main themes comprise 12 specific modules: 1) Personal information; 2) Household; 3) Location; 4) Living conditions, 5) SARS-CoV-2 exposure; 5) Work; 6)

<table>
<thead>
<tr>
<th>Country</th>
<th>Number of respondents</th>
<th>Cross-sectional data</th>
<th>Panel</th>
<th>Round 1</th>
<th>Round 2</th>
<th>Round 3</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Germany</td>
<td>5020</td>
<td>15835</td>
<td>2684</td>
<td>23539</td>
<td>3145</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Argentina</td>
<td>1013</td>
<td>611</td>
<td>77</td>
<td>1701</td>
<td>356</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Portugal</td>
<td>454</td>
<td>822</td>
<td>33</td>
<td>1309</td>
<td>289</td>
<td></td>
<td></td>
</tr>
<tr>
<td>USA</td>
<td>785</td>
<td>289</td>
<td>15</td>
<td>1089</td>
<td>390</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GB</td>
<td>540</td>
<td>311</td>
<td>19</td>
<td>870</td>
<td>290</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brazil</td>
<td>327</td>
<td>148</td>
<td>6</td>
<td>481</td>
<td>173</td>
<td></td>
<td></td>
</tr>
<tr>
<td>India</td>
<td>281</td>
<td>74</td>
<td>1</td>
<td>356</td>
<td>83</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Finland</td>
<td>194</td>
<td>115</td>
<td>9</td>
<td>318</td>
<td>143</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spain</td>
<td>195</td>
<td>79</td>
<td>6</td>
<td>280</td>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Austria</td>
<td>131</td>
<td>105</td>
<td>13</td>
<td>249</td>
<td>71</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mexico</td>
<td>59</td>
<td>176</td>
<td>4</td>
<td>239</td>
<td>33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Australia</td>
<td>165</td>
<td>68</td>
<td>5</td>
<td>238</td>
<td>76</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kyrgyzstan</td>
<td>175</td>
<td>50</td>
<td>1</td>
<td>226</td>
<td>22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Belgium</td>
<td>64</td>
<td>142</td>
<td>7</td>
<td>213</td>
<td>49</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Switzerland</td>
<td>96</td>
<td>104</td>
<td>10</td>
<td>210</td>
<td>60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Italy</td>
<td>142</td>
<td>53</td>
<td>5</td>
<td>200</td>
<td>67</td>
<td></td>
<td></td>
</tr>
<tr>
<td>France</td>
<td>122</td>
<td>69</td>
<td>4</td>
<td>195</td>
<td>76</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Colombia</td>
<td>110</td>
<td>57</td>
<td>27</td>
<td>194</td>
<td>49</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indonesia</td>
<td>28</td>
<td>161</td>
<td>3</td>
<td>192</td>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Canada</td>
<td>102</td>
<td>65</td>
<td>8</td>
<td>175</td>
<td>60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Netherlands</td>
<td>86</td>
<td>68</td>
<td>9</td>
<td>163</td>
<td>54</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>1705</td>
<td>2678</td>
<td>2629</td>
<td>7012</td>
<td>883</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: The table lists all countries with participants >150. All responses from other countries were pooled into the “Other” category.

Fig. 1. Democratic preferences and mental health

Note: Kernel density estimates of participants’ depression, anxiety, aggression and symptomatic scores by political views. For each mental health outcome, the mean score is significantly higher among those opposing a democratic system than those supporting it (p < .01). The dashed lines indicate thresholds for severe levels of symptoms. Period: October 1, 2020–March 25, 2021. Sample: N = 18,438, weighted based on age, gender and education (observations from countries with less than 150 responses were excluded).
Income; 7) Food Security; 8) Well-being; 9) Trust; 10) Social relations; 11) Public life; and 12) Personal views. Wherever possible, the survey draws on standard survey tools and specific items that have been validated extensively in various cultures and are considered the ‘gold standard’ across research disciplines. All final survey questions were pre-tested and reviewed by an IRB board.

As face-to-face interviews are difficult to implement during a pandemic, telephone and online interviews have become the predominant questionnaire-based survey modes for collecting micro data pertaining to COVID-19. Both modes have strengths and weaknesses. For phone surveys, random and representative samples may be available, but implementation is costly, sample sizes are modest, and the scope of the questionnaire is limited. Online surveys can be made instantly and freely available to anyone with internet access at any time, which means more individuals can be reached than via phone-based alternatives. A key drawback is that the samples are typically not random and not representative. Potential biases due to self-selection into the sample can be mitigated through the use of statistical weights based on population data (as described below). That way, statistics can be made

Fig. 2. Worries about getting COVID-19 a) Oneself and friends and family, by gender
Note: Unconditional means of “being worried” by sex, with 95% confidence intervals. Period: October 1, 2020–March 25, 2021. Sample: N = 18,438, weighted based on age, gender and education (observations from countries with less than 150 responses were excluded).

b) Oneself, by gender, income provider status and co-habitant children.
Note: Predicted conditional means of “being worried about getting COVID-19”, with 95% confidence intervals. Estimates are based on linear regression of participants’ infection worries on gender, income provider status, living with children, and their full interactions, controlling for country fixed effects. Period: October 1, 2020–March 25, 2021. Sample: N = 18,438, weighted based on age, gender and education (observations from countries with less than 150 responses were excluded).

Income; 7) Food Security; 8) Well-being; 9) Trust; 10) Social relations; 11) Public life; and 12) Personal views. Wherever possible, the survey draws on standard survey tools and specific items that have been validated extensively in various cultures and are considered the ‘gold standard’ across research disciplines. All final survey questions were pre-tested and reviewed by an IRB board.

As face-to-face interviews are difficult to implement during a pandemic, telephone and online interviews have become the predominant questionnaire-based survey modes for collecting micro data pertaining to COVID-19. Both modes have strengths and weaknesses. For phone surveys, random and representative samples may be available, but implementation is costly, sample sizes are modest, and the scope of the questionnaire is limited. Online surveys can be made instantly and freely available to anyone with internet access at any time, which means more individuals can be reached than via phone-based alternatives. A key drawback is that the samples are typically not random and not representative. Potential biases due to self-selection into the sample can be mitigated through the use of statistical weights based on population data (as described below). That way, statistics can be made

Fig. 3. Support for government measures a) Country level
Note: Country-level shares of respondents who support their government’s measures to counter the pandemic. Period: October 1, 2020–March 25, 2021. Sample: N = 18,438, weighted based on age, gender and education (observations from countries with less than 150 responses were excluded).

b) Government measures too weak (among non-supporters), country level
Note: Country-level shares of non-supportive respondents who say that their governments’ measures to counter the pandemic are too weak. Period: October 1, 2020–March 25, 2021. Sample: N = 2,379, weighted based on age, gender and education (observations from countries with less than 30 ‘unsupportive’ respondents were excluded).

c) Worried about getting ill and support for government measures in Berlin, Germany, postcode level
Note: Relationship at the postcode level between the share of respondents worried about getting ill and the mean level of support for countermeasures in Berlin, Germany (scatterplot and linear regression). Period: October 1, 2020–March 25, 2021. Sample: N = 329, responses from 172 5-digit postcodes (observations from postcodes with less than 5 responses were excluded).
representative in terms of various demographic dimensions at any level for which population data is available. For this reason, we opted to collect our data via an online survey.

We use snowball and panel sampling to survey individuals across social strata (C4). Participants are invited to retake the survey each quarter. The survey is advertised via Google, social media, newspapers, and networks. This strategy maximises the number of respondents, meeting basic sample size requirements for intra- and international comparisons (see: C2). To mitigate potential biases stemming from this sampling approach, we statistically weight the data. The dataset is weighted by gender and age, and in rounds 2 and 3, also by education. For Germany we additionally include weights by income groups (see: Section S3 in the Online Supplementary Materials for more detailed information). Observations include self-reported sub-national location information (e.g. postcodes in Germany) and are time stamped, enabling spatiotemporal matching with other data (C5). The resulting dataset, including weights, is available in real-time to the LwC network. Members of the research teams at the collaborating institutions have access to weekly updated draws of the data. The data are currently available to anyone upon submission of a request to the LwC team. The most recent draw of the data before the point in time when the request is submitted is provided.

We collect comparable data via phone surveys in four African countries (Brück and Regassa, 2022; see: https://lifewithcorona.org/lwc-africa). The merged phone and online survey dataset will be made publicly available upon publication of this article on the project website and via GESIS (see: https://www.gesis.org/en) Findings are regularly posted on the project website, circulated via our newsletter and social media, and used for teaching.

4. Results from the Life with Corona survey

Below, we discuss various examples of analyses that LwC data enable, due to the defining criteria: across domains (C1), socio-demographic groups (C2), space (C3), time (C4), and data sources (C5). Detailed information on analyzed measures is provided in Section S4 of the Online Supplementary Materials.

4.1. Across domains

LwC facilitates analyses of interrelated outcomes of key life aspects. For example, we can show that poorer mental health, including depression, anxiety, aggression and somatization, is significantly related to opposing democracy (on average). Comparing kernel density estimates of mental health outcomes between supporters and non-supporters of democracy, we find that medium to high levels of depression and somatic symptom burdens are particularly associated with lower support for democracy (Fig. 1). These results not only add evidence for deteriorating mental health during the pandemic but also demonstrate associated risks to political stability and social cohesion, thus informing evidence-based mitigation policies.

4.2. Across socio-demographic groups

LwC enables us to contrast pandemic experiences across socio-demographic groups. For example, comparisons of weighted mean responses reveal that women are significantly more worried about the disease than men, being both more concerned for themselves and for others (Fig. 2a). The magnitude of the gender gap is particularly pronounced among those who neither have children nor are the main breadwinner of their household (Fig. 2b). These results have three key
Second, they reveal that the pandemic creates new gender disparities in outcomes. For example, simple inspection of the data shows implications. First, they highlight the importance of directly surveying worries during a pandemic (and perceptions and beliefs more broadly). Second, they reveal that the pandemic creates new gender disparities (such as pandemic worries) beyond those in commonly discussed domains (such as income). Third, gender critically intersects with other socio-demographic factors in shaping pandemic inequalities.

4.3. Across space

LwC also makes it possible to analyse intra- and international variations in outcomes. For example, simple inspection of the data shows that non-support for government countermeasures ranges from less than 10% to almost 40% across countries (Fig. 3a). We see a broad range of people who think that measures are too weak, ranging from 10% to 90% (Fig. 3b). In Berlin, Germany, linear regression analysis at the postcode level shows that support is positively and significantly associated with worries about infection at the neighbourhood level (Fig. 3c). These results demonstrate that people’s social and political attitudes can vary strongly across localities and that they can also spatially cluster with perception and beliefs pertaining to public health.

4.4. Across time and panel structure

LwC facilitates the study of intertemporal variation in exposure to and consequences of the pandemic, including longitudinal analyses. For example, local polynomial smoothing shows that life satisfaction in Germany almost monotonically decreased from October 2020 until January 2021 (Fig. 4a). Using our panel observations, linear fixed effects models suggest that intra-household tension is strongly associated with lower life satisfaction (Fig. 4b). This demonstrates that during a pandemic the evolution of individual well-being over time is accompanied by household-level stressors.

4.5. Across data sources

LwC responses can be spatially and temporally matched with external data, based on survey timestamps and respondents’ self-reported location. Secondary data may include a variety of data sources at the national and sub-national levels, ranging from public health outcomes (such as daily level of new COVID-19 infections) to policy measures (such as the stringency of countermeasures) to macro-economic indicators (such as the unemployment rate) to social stability indicators. Merging LwC data with countermeasure stringency (Hale et al., 2021), local polynomial smoothing reveals a complex relationship with prosocial behaviour: at low stringency levels, slight increases are associated with more prosocial behaviour, but the opposite holds at higher stringency levels (Fig. 5a). At the same time, individuals perceive social relations to worsen as stringency rises (Fig. 5b). These results emphasise that the impacts of the pandemic and its countermeasures on social cohesion are complex and potentially non-linear. While ‘low-intensity’ lockdowns can bring communities together, high-intensity lockdowns can drive them apart.

5. Discussion

Even though the ‘acute phase’ of the pandemic might now have ended in countries with high rates of immunisation, the crises created by the pandemic will have lasting heterogeneous impacts. New, multi-topic, micro-level data are required to understand these impacts, how they are interrelated and how they shape the post-pandemic world. In this article, we show the power of such data in the form of the LwC data. Analyses using LwC data allow us to advance our understanding of the multidimensional challenges, inequalities and other impacts the pandemic and its countermeasures have created, and continue to create, for societies around the world.

Funding source

German Federal Ministry of Education and Research (BMBF).
UK Economic and Social Research Council (ESRC).
Alexander von Humboldt Foundation.

Role of funding source

The funding sources had no involvement in the data collection, analysis, interpretation of data, in the writing of the articles, and in the decision to submit it for publication.
Credit author statement

Wolfgang Stojetz: Conceptualization, Writing - original draft, Writing - review & editing
Neil T. N. Ferguson: Conceptualization, Writing - original draft, Writing - review & editing

Ghassan Balliki: Methodology, Writing - review & editing
Oscar Diaz: Data analysis, Writing - review & editing
Jan Elles: Data curation, Writing - review & editing
Damir Esmaliiev: Methodology, Writing - review & editing
Hanna Freudenreich: Data curation, Writing - review & editing
Ani Koeabach: Conceptualization, Writing - review & editing
Lilliana Abreu: Conceptualization, Writing - review & editing
Laura Peitz: Conceptualization, Writing - review & editing
Monika Schreiner: Conceptualization, Writing - review & editing
Ani Koeabach: Conceptualization, Writing - review & editing
Patricia Justino: Conceptualization, Funding acquisition, Writing - review & editing
Tilman Brück: Supervision, Conceptualization, Funding acquisition, Methodology, Writing - review & editing

Submission declaration and verification

The work described has not been published previously. It is not under consideration for publication elsewhere and its publication is approved by all authors. If accepted, it will not be published elsewhere in the same form, in English or in any other language, including electronically without the written consent of the copyright holder.

Declaration of competing interest

None.

Acknowledgment

We gratefully acknowledge support from the German Federal Ministry of Education and Research (BMBF), the UK Economic and Social Research Council (ESRC), and the United Nations University World Institute for Development Economics Research (UNU-WIDER). Anke Hoeffler acknowledges funding from the Alexander von Humboldt Foundation. For excellent support with information technology, software, and website design, we thank Klaus Hartl and Nils Henrik Johansson. For excellent support with project and partnership management, we thank Philip Albers, Ilna Kuttita, Evaa Nyssonen, Myrslava Purska, Babette Regisser, Petra Rietzler, Gabija Verbaite, and Julia Vogt. For excellent support with data collection and analysis, we thank Samuel Carleial, Lea Ellmanns, Helena Meier, Mekdim D. Regassa, Andrej Smirnov, and Dorothee Weifen.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.socscimed.2022.115109.

References