IZA Institute of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14903

A Year of Pandemic: Levels, Changes and Validity of Well-Being Data from Twitter. Evidence from Ten Countries

Francesco Sarracino
Talita Greyling
Kelsey O'Connor
Chiara Peroni
Stephanie Rossouw

NOVEMBER 2021

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14903

A Year of Pandemic: Levels, Changes and Validity of Well-Being Data from Twitter. Evidence from Ten Countries

Francesco Sarracino
Institut national de la statistique et des études économiques du Grand-Duché du Luxembourg

Talita Greyling
University of Johannesburg and Auckland University of Technology

Kelsey O'Connor
Institut national de la statistique et des études économiques du Grand-Duché du Luxembourg, Research division, IZA and University of Johannesburg

Chiara Peroni
STATEC

Stephanie Rossouw
Auckland University of Technology and University of Johannesburg

NOVEMBER 2021

Any opinions expressed in this paper are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but IZA takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The IZA Institute of Labor Economics is an independent economic research institute that conducts research in labor economics and offers evidence-based policy advice on labor market issues. Supported by the Deutsche Post Foundation, IZA runs the world's largest network of economists, whose research aims to provide answers to the global labor market challenges of our time. Our key objective is to build bridges between academic research, policymakers and society.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

A Year of Pandemic: Levels, Changes and Validity of Well-Being Data from Twitter. Evidence from Ten Countries*

In this article, we describe how well-being changed during 2020 in ten countries, namely Australia, Belgium, France, Germany, Great Britain, Italy, Luxembourg, New Zealand, South Africa, and Spain. Our measure of well-being is the Gross National Happiness (GNH), a country-level index built applying sentiment analysis to data from Twitter. We aim to describe how GNH changed during the pandemic within countries, assess its validity as a measure of well-being, and analyse its correlations. We take advantage of a unique dataset of daily observations about GNH, generalised trust and trust in national institutions, fear concerning the economy, loneliness, infection rate, policy stringency and distancing. To assess the validity of the data sourced from Twitter, we exploit various survey data sources, such as the Eurobarometer and consumer satisfaction, and Big Data sources, such as Google Trends. Results indicate that sentiment analysis of tweets can provide reliable and timely information on well-being. This can be particularly useful to timely inform decision-making.

JEL Classification: C55, I10, I31, H12

Keywords: happiness, COVID-19, Big Data, Twitter, Sentiment Analysis, well-being, public policy, trust, fear, loneliness

Corresponding author:
Francesco Sarracino
Institut national de la statistique et des études économiques
du Grand-Duché du Luxembourg
Research division
Luxembourg

E-mail: Francesco.Sarracino@statec.etat.lu

* The authors contributed equally to this work.

1. Introduction

Improving individuals' well-being is not only a desirable outcome "per se": it also carries wider individual and societal benefits. Happier people tend to live longer and healthier lives (Danner et al. [1], Guven and Saloumidis [2], Frijters et al. [3], Graham and Pinto [4]), have better employment outcomes (O'Connor [5]) and share creative, altruistic and problem-solving traits (Lyubomirsky et al. [6]). Happier people are more satisfied with their jobs, are more productive, cooperative and less absent (Bryson et al. [7], DiMaria et al. [8], Oswald et al. [9]). What is more, higher levels of past and present happiness predict higher levels of compliance during COVID-19 (Krekel et al. [10]). COVID-19, however, is having a deep impact on individuals' well-being. In general, the literature shows that traumatic events alter well-being in rapid and persistent ways (Bonanno et al. [11], Kessler et al. [12], Norris et al. [13]); this, in turn, has long-lasting collateral social and economic effects (Arampatzi et al. [14]). The novel coronavirus pandemic is undoubtedly one such event.

This article analyses the well-being changes during 2020 using the *Gross National Happiness.today* (GNH hereafter) index. The GNH is an aggregate country-level index of well-being, comparable across countries (Greyling et al. [15]), compiled by applying sentiment analysis to Twitter posts (tweets). The motivation of this analysis is twofold: on the one hand, the spread of COVID-19 has deeply affected well-being in societies, both directly and through the socioeconomic consequences of containment measures; on the other hand, well-being has important societal consequences and can also affect the effectiveness of responses to COVID-19. This latter point is especially relevant to policymaking, as individuals' responses are key to health policies' effectiveness and affect the successfulness of "exit" strategies to ease lockdowns and recovery plans. Timely well-being data is particularly relevant during the pandemic, as they can facilitate rapid policy responses to changing conditions.

One of the main features of the GNH is timeliness. In contrast, well-being data are typically collected via large scale population surveys by statistical sources or institutional and academic bodies. Surveys take time to administer, and data are available to researchers and analysts with delay. Also, the pandemic disrupted the ability to collect survey data, and therefore data on well-being, its changes and correlates over 2020 are scarce. For instance, the Eurobarometer, a European Commission's survey, usually provides the most recent comparable well-being series for EU countries at a biannual frequency. In 2019, the survey was administered twice; in 2020, however, the survey was only administered once. As table 1 shows, Eurobarometer data for Luxembourg indicates that life satisfaction, a valid and reliable measure of well-being, decreased by eight percentage points between Autumn 2019 and Summer 2020. It is plausible that this large observed fall in well-being is due to the pandemic, but questions remain on the size of the decrease and the mechanisms underlying the fall. The actual decrease might have been larger or smaller than eight percentage points; the recorded fall was due to the Eurobarometer survey's timing. With one reference point distant in time, this sole observation does not tell us anything about the evolution of well-being during the period between the waves. Up-to-date and more frequent observations are needed to address these issues, particularly relevant during a pandemic.

Table 1. Life satisfaction in Luxembourg from Spring 2019 to Summer 2020.

	Spring 2019	Autumn 2019	Summer 2020
people not satisfied with their lives (%)	4	6	14
people satisfied with their lives (%)	96	94	86

Source: Eurobarometer data (European Commission [16]). Well-being is measured by life satisfaction. The original variable is organised into four categories. For ease of interpretation, the bottom and top two categories have been collapsed.

The main contributions of our work are as follows. Firstly, we provide a timely account of changes in well-being during the pandemic in Australia, Belgium, France, Germany, Great Britain, Italy, Luxembourg, New Zealand, South Africa, and Spain. Secondly, we add to the literature on the use of social media data to study social and economic variables (Brodeur et al. [17], Greyling et al. [18]). In particular, we show that sentiment analysis of tweets can be used to measure well-being and other important variables, such as generalised trust, trust in institutions, economic fear, and loneliness. Importantly, we assess the validity of the data issued from sentiment analysis applied to Big Data. This also contributes to the methodological research on developing new and timely measures of socioeconomic variables using Big Data and machine learning. Finally, we provide an assessment of what affected well-being during the pandemic by regressing changes in well-being on pandemic indicators, socioeconomic variables, containment policies and behavioural responses.

The paper is organised as follows. The next section reviews the literature on the impact of COVID-19 on well-being and the use of Big Data to measure well-being. Section 3 presents the data and the methodology to construct the GNH and the additional variables used in this study. Section 4 describes the evolution of GNH and the main variables in the analysis. It reports the results of the regression analysis of changes in well-being on the variables of interest. The techniques used to analyse the data are presented alongside the results, whereas the validity tests of our variables are provided in the Appendices. Finally, Section 5 summarises the main results and provides some concluding remarks.

2. Literature review

An extensive and interdisciplinary literature discusses the negative impact of the COVID-19 pandemic on populations' well-being. Much of this literature focuses on the direct and indirect consequences of the pandemic - through the social, emotional and economic consequences of lockdowns and health policies - on mental health (Brooks et al. [19], Holmes et al. [20], Blasco-Belled et al. [21], Cooke et al. [22], Rajkumar [23], Xiong et al. [24], Saladino et al. [25], Li and Wang [26], Cao et al. [27]). Salari et al. [28]'s meta-analysis of published studies on mental health in the general population reveal that five studies indicate a prevalence of stress (30%), 17 studies indicate a prevalence of anxiety (31.9%), and 14 studies indicate a prevalence of depression (33.7%) (Please note these rates refer to the first wave of COVID-19 pandemic). Kawohl and Nordt [29] modelled the effect of COVID-19 on suicide rates through rising unemployment. Krendl and Perry [30] found that older adults reported higher depression and greater loneliness following the onset of the pandemic in a sample of American respondents. Sibley et al. [31] documented an increase in anxiety/depression following lockdown and warned about long-term challenges to mental health in New Zealand. Patrick et al. [32] reported marked changes in the mental health of parents and children in the United States. O'Connor and Peroni [33] documented a decline in mental health for nearly a third of residents in Luxembourg. The most important factors associated with the decline in mental health were worsening physical health, income, and a decline in job security. There are, however, exceptions. Sønderskov et al. [34], for instance, documented an increase in the psychological well-being of the Danish population from the first wave

(March 31 – April 6, 2020) to the second one (April 22 – April 30, 2020), probably because symptoms of anxiety and depression decreased. Recchi et al. [35] reach a similar conclusion. Using panel data from France (administered at three points in time between the 1st of April and the 6th of May 2020), the team found that, in general, self-reported health and well-being improved during the lockdown with respect to the previous year. Although the result hides some heterogeneity within the population (blue-collar workers seem to have suffered more from the crisis), the authors explained their finding, arguing that individuals not affected by the virus judged their situation better than they normally would have. As the authors warn, their findings are based on data from the first six weeks of lockdown in France, and they consider the possibility that the pandemic will affect the population in the long run.

Most studies indicate that well-being decreased in correspondence with the pandemic. However, the literature is not unanimous about which channels matter the most. Some studies, for instance, tried to track down the impact of specific policy responses to the pandemic on well-being. This is the case of recent work done by (Rossouw et al. [36]), in which the authors show that lockdown regulations hampered happiness - measured by Gross National Happiness - in South Africa. Additionally, they argued that the determinants of happiness under lockdown were factors directly linked to the regulations that were implemented: lack of access to alcohol (and tobacco), increased social media usage, concerns over future employment and more time to spend at home negatively impacted happiness. Similarly, Greyling et al. [37] found a negative effect of lockdown on Gross National Happiness in South Africa, New Zealand and Australia. Unobserved factors can, however, confound the association between lockdown and well-being. To address this issue, Foa et al. [38] distinguish the effect of lockdown from that of the pandemic using weekly data issued from YouGov's Great Britain Mood Tracker Poll and Google Trends. They found that lockdowns positively impact subjective well-being and that the main threat to mental health is the severity of the pandemic. Additionally, they suggest that lockdowns help relieve the negative impact of the pandemic on well-being by relieving anxiety and stress. The result that the negative impact on well-being is due to the severity of the pandemic is consistent with the evidence provided by Kivi et al. [39]. The authors showed that the well-being of Swedish older adults (as measured by life satisfaction and loneliness) increased in the early stage of the pandemic compared to previous years (2015-2020). On the contrary, well-being decreased for those who were more worried about the pandemic.

With only a few exceptions, previous studies mainly observed well-being at a given point in time. This is because well-being is prevalently measured with surveys that take time to administer and elaborate before results can be published. Only a few studies used Big Data to track the evolution of well-being during the pandemic. As the impact of the pandemic on societies changes over time, it is important to monitor the dynamics of well-being to fully understand their causes and economic, social, and political consequences.

The few available studies on well-being changes during the pandemic reached varied conclusions. Wang et al. [40] found no differences in the Chinese population's stress, anxiety, and depression levels when comparing two waves (January 31 - February 2, to February 28 - March 1, 2020). Brülhart and Lalive [41] compared helpline calls to Switzerland's most popular free helpline during the pandemic to the previous year's records to infer the impact of COVID-19 on people's suffering. They concluded that the impact of COVID-19 was negligible as the number of calls has grown in line with the long-run trend. Using Gross National Happiness data from New Zealand, Rossouw et al. [42] documented a decline in well-being that persisted over time. Sibley et al. [31] reported limited changes in well-being during the first stages of the pandemic, possibly due to increased community connectedness. Brodeur et al. [17] analysed data from Google Trends collected between the 1st of January 2019 and the 10th of April 2020 in nine Western European countries and the American States. Although their main aim was to assess the causal link between lockdown and the search intensity for various proxies (of lack) of well-being (terms included boredom, contentment, divorce, impairment, irritability, loneliness, panic, sadness, sleep, stress, suicide, well-being, and worry), they found evidence of mean-reversion in several measures of well-being, but not in all. They concluded that the level of well-being at the beginning of lockdown could be a poor guide to its level later on. Cheng and colleagues [43] reached a similar conclusion. The team applied the difference-in-difference technique to monthly data from a longitudinal survey administered on middle-aged and older adults in Singapore. They found large declines in overall life satisfaction and domain satisfaction during the outbreak: the magnitude of the effects are comparable to those of a major health shock or the loss of a beloved person. Figures also indicate that the impact of COVID-19 on life satisfaction is persistent over time, as it remained below its pre-pandemic levels even after the lockdown was lifted.

There are various possible explanations for the contrasting results reported above. For instance, previous studies are generally based on data for just one country, often comparing the same country over time or comparing the pandemic to a "normal" year, i.e., 2019. Such analyses hold broad institutional characteristics constant but prevent researchers from comparing these characteristics. Different country contexts, infection rates, and governmental responses are another potential explanation for the heterogeneity of results. This motivates our cross-country investigation over time.

Big Data, such as those sourced from social media platforms like Twitter, are a possible source of timely and internationally comparable data on individuals' well-being, sentiments and behaviours. However, only a few studies used Big Data to track the evolution of well-being during the pandemic. Big Data can be useful to decision-makers, especially in situations requiring rapid decisions or in the presence of incomplete information, as data are typically available with a delay. Big Data allow authorities to observe people's behaviour, and not only their opinions. In particular, applying sentiment analysis to data issued from online social media allows researchers to "listen" to what people deem important in their lives.

Moreover, Big Data does not suffer from non-response bias (Callegaro and Yang [44]). However, these advantages come at the expense of possibly large measurement errors, limited representativeness, and difficulties in establishing the validity and robustness of the measures. The use of Big Data implies a trade-off between timeliness of information and solidity and reliability of knowledge. These data offer a solution to the first issue, but their validity must be constantly assessed.

Previous studies that use Big Data to calculate a happiness index are scarce. The Hedonometer, created by Dodds and Danforth [45] and their team, is the first measure of happiness that started at the end of 2008. They use the Twitter Decahose Application Programming Interface (API) feed, a streaming API feed that continuously sends a sample of roughly 10 per cent of all tweets. This allows Dodds and the team to continuously measure happiness levels per day, thus resulting in a time series from 2008 to the present (Dodds et al. [46]). However, the Hedonometer cannot deal with the context in which words are used. Words in themselves are evaluated and not the sentiment of the construct. For example, a phrase such as "I did not enjoy the holiday" will attract a score of 7.66 for 'enjoy' and 7.96 for 'holiday', thus reflecting an overwhelmingly positive sentiment when actually the sentiment is negative. Furthermore, the Hedonometer calculates a happiness index on a scale of 1 (sad) to 9 (happy), but it cannot detect the emotions underpinning the words or the tweets. Thus, it cannot determine if the changes in the levels of happiness are due to negative emotions such as fear or anger or positive emotions such as joy.

The second known measure was developed in 2012 by Ceron et al. [47]. They used an Integrated Sentiment Analysis (a human supervised machine learning method) on Big Data extracted from Twitter for both Italy and Japan. They created a composite index of subjective and perceived well-being that captures various aspects and dimensions of individual and collective life (Iacus et al. [48]). Up until 2017, the researchers extracted and classified 240 million tweets over 24 quarters. They applied a new human supervised sentiment analysis to analyse the sentiment and did not rely on lexicons or special semantic rules.

2.1 Our contribution

The present study contributes to the literature on the changes and correlates of well-being during the pandemic. To this purpose, we use the Gross National Happiness Index (Greyling et al. [15]), a measure of well-being extracted from tweets and available daily throughout 2020. We focus on a set of ten countries, including a set of Western European countries severely affected by the pandemic during the first wave (Belgium, France, Great Britain, Germany, Italy, Luxembourg and Spain), and Australia, New Zealand, and South Africa for which data are readily available (Greyling et al. [37], Rossouw et al. [36] [41]). Secondly, we exploit the timeliness and frequency of our data to study how well-being changed during the pandemic. Thirdly, we analyse the correlates of well-being changes using daily observations on emotions, trust in national institutions, loneliness, economic fear, as well as indicators of the pandemic and economic conditions. In doing so, we also contribute to the methodological literature on applying sentiment analysis to Big Data. In particular, for the first time, we use this technique to produce data on economic fear, trust in national institutions, and sadness about loneliness. Finally, we establish the validity of the measures sourced from Twitter by analysing their correlation across countries and time with other measures issued from surveys and Big Data.

3. Data

3.1 The Gross National Happiness Index

To measure well-being (the dependent variable), we make use of the Gross National Happiness Index (GNH), which was launched in April 2019 (Greyling et al. [15]). This project measures the evaluative mood of a country's citizens over time by analysing the sentiment content of tweets. As a measure of mood, the GNH captures the more volatile part of well-being, commonly referred to as happiness (Diener et al. [49]). However, the evaluative qualification indicates tweets reflect individuals' conscious decisions - they evaluate what they want to say.

The GNH index is compiled by extracting posts (tweets) from Twitter, a voluntary information-sharing social media platform. Sentiment analysis is applied to a live Twitter feed. Each tweet is labelled as having either a positive, neutral or negative sentiment. Then, this classification is applied to a sentiment-balance algorithm to derive an aggregate happiness score - the GNH. The resulting GNH ranges from 0 to 10, with higher values indicating higher happiness, and is available at the aggregate (country) level and hourly frequency. In addition to the GNH, an algorithm extracts each tweet's emotions, which is novel in the literature. The method used to compile the GNH figures is described in detail in Greyling et al. [15].

Sentiment analysis is an automated process to determine the feelings and attitudes of the author of a written text (Hailong et al. [50]). Authors from many social sciences have applied sentiment analysis to address various issues (Eichstaedt et al. [51], Caldarelli et al. [52], Gayo-Avello [53], Bollen et al. [54], Asur and Huberman [55], O'Connor et al. [56]). For instance, Twitter messages have been used to track the influence rate in the United Kingdom and the United States (Lampos and Cristianini [57], Culotta [58]). Paul and Dredze [59] found a positive association between public health data and the data issued from sentiment analysis of tweets.

The GNH has been available on an hourly basis since April 2019 for Australia, New Zealand and South Africa (see Greyling et al. [37], Rossouw et al. [41], Greyling et al. [15]). This paper extends the coverage of the GNH to Belgium, France, Great Britain, Germany, Italy, Luxembourg and Spain, for which data are available from January 1st to December 31st, 2020. The GNH is available live on the project's website (https://gnh.today/).

The number of tweets is large, and user numbers represent significant proportions of the countries' populations investigated. Every day, people exchange more than 4,600 tweets in New Zealand and approximately 93,500 tweets in the United Kingdom. In Luxembourg, the smallest of the considered countries, tweets are approximately 257 per day. This indicates that data traffic should be sufficient to compile a daily GNH index for all the considered countries.

One of the advantages of Twitter data is that they are abundant, and users are heterogeneous. Twitter accounts include individuals, groups of individuals, organisations and media outlets, thus providing the moods of a vast blend of users, which is not found in survey data. Another advantage is that Twitter can provide timely information about the mood of a country. However, compared to data from statistical surveys, Twitter data also have limitations. A major known drawback of social media data, including Twitter, consists in their lack of representativeness of the population object of the study (in this case, the general population of the countries). It is, therefore, vital to assess the ability of the GNH to correctly capture the (changes in) well-being, i.e., whether the GNH is a valid indicator of well-being.

The validity of the GNH can be assessed in several ways. We assume that content validity, i.e., the ability of the GNH to represent well-being correctly, is satisfied. The reason is that the algorithm for sentiment analysis is logically built to measure the affection content of a tweet. We test the criterion validity of the GNH by checking whether GNH significantly correlates with external data that are known to represent the same or similar concepts. This validation exercise is presented in Appendix A. We proceed as follows: firstly, we compare cross-sectional country rankings produced from the GNH to those available from alternative measures of well-being, such as the Eurobarometer's life satisfaction and the World Happiness Report (Helliwell et al. [60]). Secondly, we analyse the time-series correlation between GNH changes and changes in consumer confidence and other well-being indicators available for the same period from Google search. The validation exercise provides encouraging results: GNH-based country rankings are not significantly different from those based on alternative indicators. Results from time-series correlations are mixed: the correlation of GNH to well-being from survey data for four European countries is uncertain. At the same time, it performs relatively well in relation to the index of negative emotions and consumer confidence.

In addition to our tests, previous literature suggests that the GNH index correctly reflects the evaluative mood of a nation. Greyling et al. [37] showed a negative and statistically significant association between the GNH index and 'depression' and 'anxiety' for Australia, New Zealand and South Africa. Moreover, data indicate that the GNH index variations reflect various events, including the COVID-19 pandemic. Data from South Africa show that the GNH dropped well below previous daily averages following the outbreak of COVID-19. Later, when distancing regulations were implemented, the GNH recovered slightly but remained lower than normal (Greyling et al. [37]). In a different field, Lampos and Cristianini [57] and Culotta [58] showed that information extracted from tweets could track health variables.

In summary, evidence from previous studies, and validation results from this study, indicate that the GNH can be regarded as a valid measure of well-being. Descriptive statistics on the GNH are available, along with information on additional variables, in Table 3. Fig 1 depicts the time series of GNH (solid line) for seven European countries (panel 1a) and Australia, New Zealand, and South Africa (panel 1b) for 2020. Noticeably, in Europe, the GNH dipped at the outbreak of the first and second pandemic waves. At the pandemic onset, the fall in GNH is also apparent for Australia, New Zealand, and South Africa.

3.2 Additional variables from Twitter

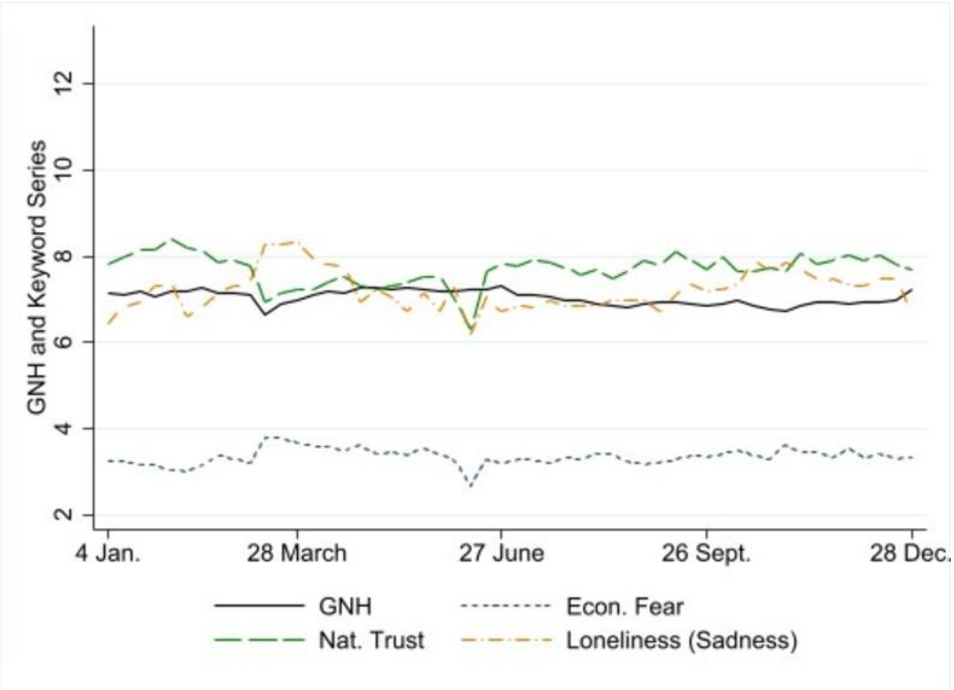
The dataset also includes the following variables derived from the emotions expressed in tweets:

- anger, fear, anticipation, trust (also referred to as generalised trust hereafter), surprise, sadness, joy, and disgust (as already done, for instance, see Greyling et al. [37]),
- fear in relation to the economic situation (economic fear),

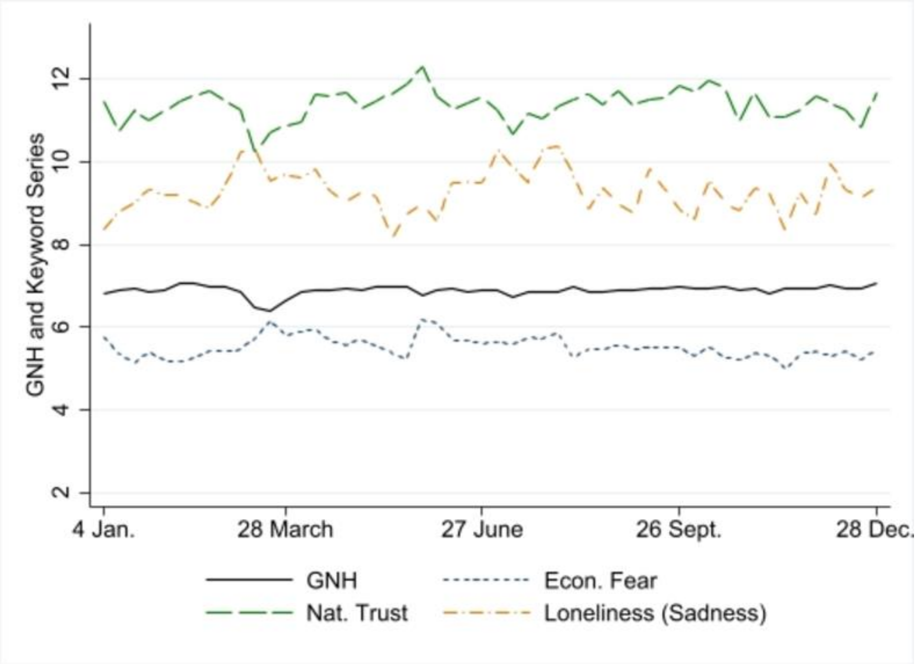
- sadness in relation to loneliness (referred to as loneliness hereafter), and
- trust in national institutions.

The abovementioned variables were extracted from tweets using sentiment analysis and constructed as daily averages of the associated emotions. Table 2 provides three examples of how emotions are extracted. Each tweet is attributed a score according to the presence and intensity of one or more emotions. For example, the tweet "I love dogs: they are such good companions"; resulted in a score of 1 being assigned to the emotion called 'trust' and a score of 2 to the emotion called 'joy'. The daily score of a given emotion A corresponds to the average of the scores that emotion A received on a given day. As an example, table 2's tweets generate a score of 2.3 for 'joy', i.e. (2 + 5 + 0)/3.

The remaining variables - economic fear, loneliness, and trust in national institutions - were obtained by applying the method illustrated above to tweets that included selected keywords. The detailed list of keywords used to produce each variable is available in Appendix D.



a) Average daily data across six European countries.



b) Average daily data across Australia, New Zealand and South Africa.

Fig 1. Gross National Happiness, economic fear, trust in national institutions, and loneliness in 2020.

Note: Data are presented using seven-day (centered) moving averages. Figures for Luxembourg are missing due to the scarcity of tweets using the selected keywords.

Source: Data are all sourced from the project "Preferences Through Twitter" with the support of FNR, UJ and AUT.

Fig 1 shows the time series of economic fear, loneliness, trust in national institutions, and the GNH for six European countries (panel 1a) and Australia, New Zealand, and South Africa (panel 1b) for the year 2020. Appendix E gives details on the validity of these measures. Please note that the number of European countries decreases to six because figures are not available for Luxembourg.

One can see from Fig 1 that Australia (A), New Zealand (NZ) and South Africa (SA) experienced higher levels of economic fear, sadness linked to loneliness, and trust in national institutions than the European countries (EU). Trust and sadness exhibit higher volatility in A-NZ-SA than in the EU. We notice a marked decrease of trust in national institutions in correspondence with the first severe coronavirus outbreak (March), followed by a slow recovery in both groups of countries. At the same time, loneliness increased, with a further increase in EU countries during the second wave. Economic fear increased during the first wave. Afterwards, it remained at levels higher than the initial one in the EU countries. At the same time, it decreased slightly but steadily in A-NZ-SA (Fig 12 in Appendix B provides detailed trends for each variable for each country). The limited number of tweets produced in Luxembourg did not allow us to compute trust in national institutions, sadness in relation to loneliness, and economic fear.

Table 2. Examples of coding tweets for emotions.

"I love dogs; they are such good companions"

Anger	Fear	Anticipation	Trust	Surprise	Sadness	Joy	Disgust
0	0	0	1	0	0	2	0

"Judith's doing a great job boosting the party vote in her new role as leader of the Nat Party, hope they get rid of that Bridges guy now"

Anger	Fear	Anticipation	Trust	Surprise	Sadness	Joy	Disgust
4	0	1	2	0	0	5	0

"Mask-wearing is really reducing in inner Auckland – I've been virtually the only one I've seen today. (Lack of) distancing pretty much the same….#COVID19NZ"

Anger	Fear	Anticipation	Trust	Surprise	Sadness	Joy	Disgust
6	0	2	0	2	0	0	4

Source: (Greyling et al. [15]).

3.3 Additional explanatory variables

We integrate the variables derived from Twitter data with additional information to account for the evolution of the pandemic (daily new cases), policy responses, behavioural responses (distancing), and economic conditions (unemployment rate). Data on COVID-19 are sourced from Our World in Data (Roser et al. [61]). Among available series, we retain the number of new confirmed cases per day per million in population. We adjust by population to account for countries' sizes. In much of the analyses, we further transform new cases using an inverse hyperbolic sine transformation, which is roughly equivalent to a log transformation but is identified for zeros. Please note that the European countries' data are from the ECDC (European Centre for Disease Prevention and Control). Available series also include the number of tests performed, deaths and hospitalisations.

The indicator of policy stringency is the Containment and Health Index from the University of Oxford's COVID-19 Government Response Tracker (Hale et al. [62]). The tracker includes multiple indices summarising 18 indicators of policy response to the COVID-19 pandemic in different dimensions. The Containment Index, also known as the stringency index, is based on the following nine indicators:

school closing, workplace closing, cancel events, restrictions of gathering, close public transport, stay at home requirements, restrictions on internal movement, international travel controls, and public information campaigns. Details on the construction of the index and the underlying indicators are available online (www.bsg.ox.ac.uk/covidtracker). The data and methodology are frequently updated as the pandemic evolves.

We use Google Mobility Reports (Google [63]) to measure distancing and account for behavioural responses to the pandemic and government policies. Google Mobility Reports provide daily aggregate mobility/visitation data by geographic locations. The data, collected from users' devices that have opted-in to location history on their Google account, permit mobility trends across several types of visited places: retail and recreation, groceries and pharmacies, parks, transport hubs, workplaces, and residential. The measure of distancing we consider is an index reflecting the time people spend at home. Our choice is motivated by requiring fewer assumptions about people's movements during the pandemic. The figures are compiled as relative movements (visits' numbers) compared to the number of visits during the baseline period, 3 January to 6 February. Mobility is also normalised for each day of the week. Seven baseline days are used, corresponding to the median values observed during the five-week baseline period. For this reason, we cannot compare daily movements. We instead use weekly average values or daily data smoothed using a seven-day centered moving average.

The monthly unemployment rate is made available for the European countries by Eurostat (Eurostat, [64]). We use the raw, not seasonally adjusted series. Table 3 provides summary statistics for the variables described in this section.

Table 3. Descriptive statistics by country. Average values over the year 2020.

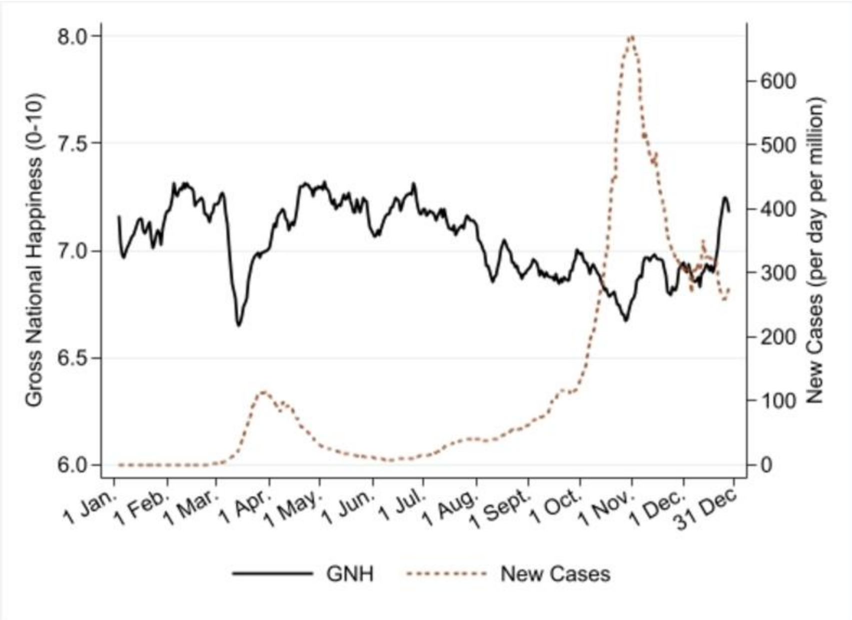
Country	GNH (0-10)	Confirmed cases (Cum.) per million	Containment policy (0-100)	Residential mobility (%)	Unemployment rate (%)	Consumer confidence Pos. Bal. (%)	Tweets per day
Australia	7.27	1,115	54.62	8.58			24,354
Belgium	7.05	55,782	51.24	11.57	5.63	-12.16	6,335
France	6.29	41,022	54.90	10.52	8.18	-12.89	37,250
Germany	7.43	21,013	51.94	6.75	4.19	-9.55	22,318
Italy	7.22	34,851	58.56	10.37	9.12	-16.65	27,677
Luxembourg	7.14	74,148	43.16	12.29	6.75	-11.69	257
New Zealand	7.06	448	35.48	8.00			4,624
South Africa	6.34	17,825	53.18	15.18			57,256
Spain	6.81	41,242	56.27	10.22	15.54	-22.86	55,289
United Kingdom	7.42	36,771	57.06	13.16	4.20	-16.56	93,500

Note: The unemployment rate is unadjusted, consumer confidence is adjusted.
Source: all sources are described in the text. They are omitted for brevity.

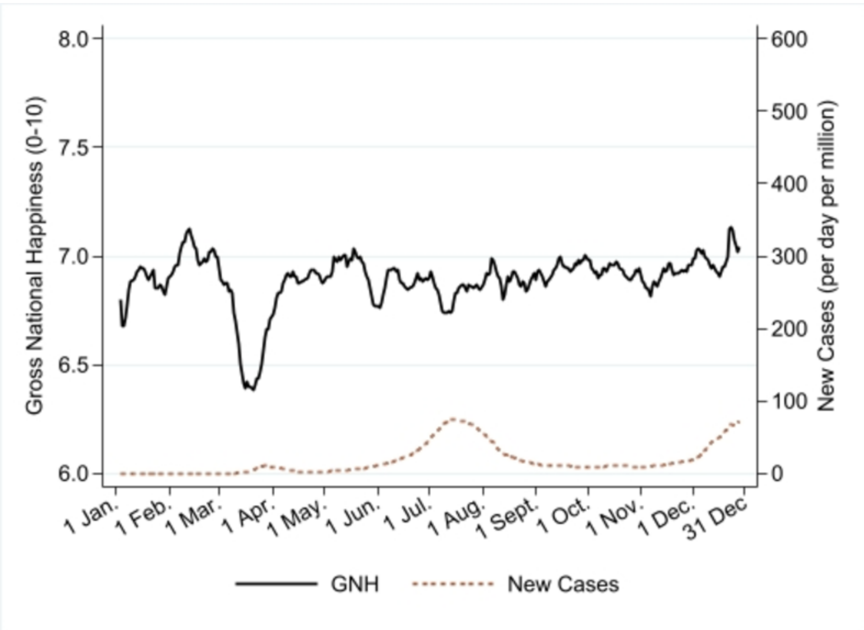
4. Results

Before turning to the regression analysis, we describe the changes of our main variables in correspondence with two marking events of 2020, namely the evolution of the pandemic and the policy responses to the outbreak. For brevity, we group countries in two groups: the seven European countries (EU hereafter), and Australia, New Zealand and South Africa (A-NZ-SA hereafter). Please note detailed variables' evolution for each country is presented in Appendix B.

Fig 2 contrasts the respective evolutions of average GNH (solid line) and COVID-19 infections (dashed line) for the EU (panel 2a) and A-NZ-SA (panel 2b). Please note that infections are the average number of daily new confirmed positive cases per million. Note also that mass testing was not performed during the first pandemic wave.



a) Average daily data across seven European countries.



b) Average daily data across Australia, New Zealand and South Africa.

Fig 2. Gross National Happiness and new COVID-19 cases per day in 2020.
Note: GNH and new cases are presented using seven-day (centered) moving averages.
Source: GNH data (Greyling et al. [15]) are sourced from the project "Preferences Through Twitter" with the support of FNR, UJ and AUT. The number of new positive cases is sourced from OurWorldinData.org.

Panel 2(a) shows that GNH dipped in correspondence to the two pandemic peaks of March and November 2020. In Europe, GNH dropped suddenly (-8.6%) during the first wave, recovering afterwards quickly (+9.84%). In correspondence to the slow but steady increase in the number of cases during the late European summer-autumn, GNH showed a steady decline culminating with a sharp fall at the beginning of November, when infections reached a second peak.

Panel 2(b) shows a similar pattern for A-NZ-SA's GNH during the first peak (GNH suddenly dropped by 9.33%). The evolution of GNH declined slightly during the emergence of the second pandemic wave (May- July) and recovered afterwards. We also observe that the number of new positive cases has been substantially lower than the one recorded in Europe in this group of countries. GNH changes in 2020 are more volatile in Europe than in A-NZ-SA (table 13 in Appendix C provides average scores of the GNH by sub-periods, while the changes of GNH for each country are shown in Fig 10 in Appendix B). Hence, while the average GNH is about 7, our daily data reveal a varied picture in terms of intensity and duration of the shock and across waves of infection.

However, other variables may influence GNH, besides the number of infections. Containment policies, economic conditions, trust in others and institutions may have affected the overall well-being, both directly and through their impact on the pandemic. We account for the joint effect of these variables using regression analysis in section 4.1. In the remainder of this section, we briefly describe GNH changes in relation to changes in containment policies and trust in others. These variables are relevant for their direct effects on well-being, and because of their effect on the pandemic: containment policies limited the spread of COVID-19 (Fong et al. [65], Chinazzi et al. [66]), thus benefiting well-being; trust in others promoted cooperation and solidarity with positive spillovers on compliance and well-being (Bargain and Aminjonov [67]).

Fig 3 in section 4.1 reports the changes of GNH (solid line) along with those of containment policies (dashed line). We notice the jump in policy stringency and GNH drop, which occurred at the pandemic's onset. After this, the months of May – July were characterised by a gradual relaxation of containment policies. Then, policies in the two groups of countries took different directions. In A-NZ-SA, increases in stringency during July-September were followed by a marked relaxation of policies. In European countries, the summer coincided with a relaxation of policies, down to a degree of stringency maintained throughout October. Stringency jumped again from about 50 to nearly 80 points during October, accompanied by the sharp fall in GNH. However, the decline in GNH had begun earlier. In A-NZ-SA, the increase in stringency saw a dip in GNH, but also, in this case, the latter's decline had started before.

It is worth noticing that countries' policy responses to the pandemic have been widely heterogeneous, both within and between the two groups of countries. Fig 11 in Appendix B depicts changes in GNH, the number of new positive cases and the level of containment policies separately for each country. The graphs show that policy response in the EU shares a similar pattern across countries – after the initial shock, stringency in containment policies increased with a rise in new positive cases. In contrast, countries in the Southern hemisphere saw nearly zero new infections, contrasted with heterogeneous governments' responses: Australia maintained strict containment policies throughout 2020, whereas South Africa gradually relaxed the measures; in New Zealand, the initial suppression strategy was followed by a sharp drop in stringency, interrupted only by a sharp increase in stringency during August-September.

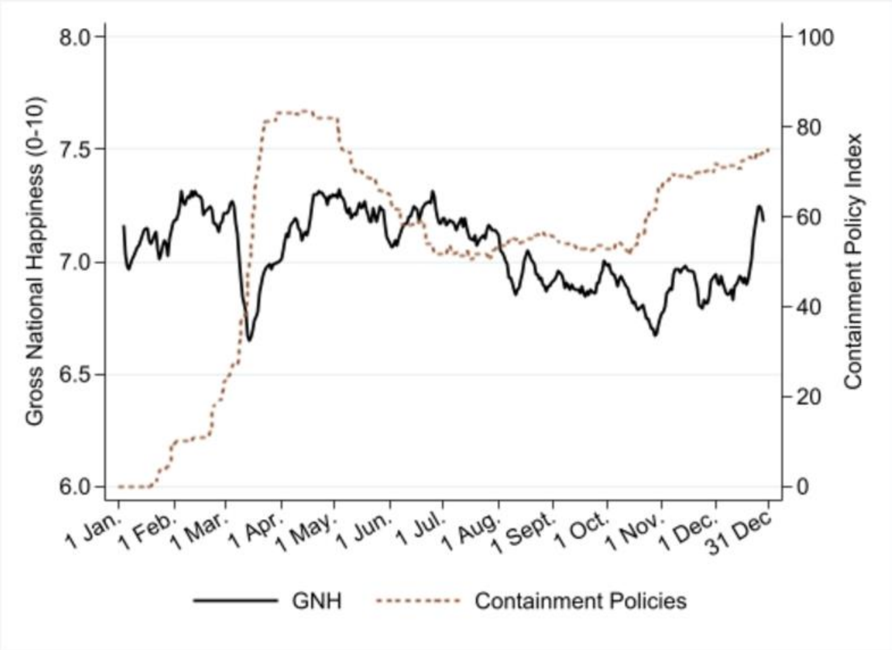
A-NZ-SA report, on average, higher trust in others than European countries (respectively, 7.6 and 6.2 throughout 2020). Our measure of trust correlates at 81% (statistically significant at 10%) with survey-based measures of trust. For more details, see Fig 16 in Appendix E and the notes therein. Also, the changes in trust over time differ in the two groups of countries (see Fig 4). Trust increased in both groups of countries in the first half of the year, markedly in A-NZ-SA (on average, 8.33% in European countries and 13.7% in A-NZ-SA). The upward trend was shortly interrupted in correspondence with the first outbreak of COVID-19. Subsequently, trust started declining (with an initial sudden drop by about 6%), corresponding to a renewed growth in infections (from June onward), with different patterns. In A-NZ-SA, it levelled off before decreasing again in September-October and started increasing again in November-December. In Europe, trust declined (nearly 7.7%) in the period from June to September and exhibited an increasing trend afterwards. Fig 12 in Appendix B shows changes of trust for each country in the study.

In summary, the daily evolution of GNH reveals considerable variations in well-being responses during 2020. We found a clear indication that people suffer when the infection worsens and that the recovery takes longer than the decline. A descriptive examination of the changes in the number of new cases, containment policies, and generalised trust indicates a mixed relationship with GNH changes. This suggests that multiple factors should be considered jointly to explain the changes in well-being during the pandemic.

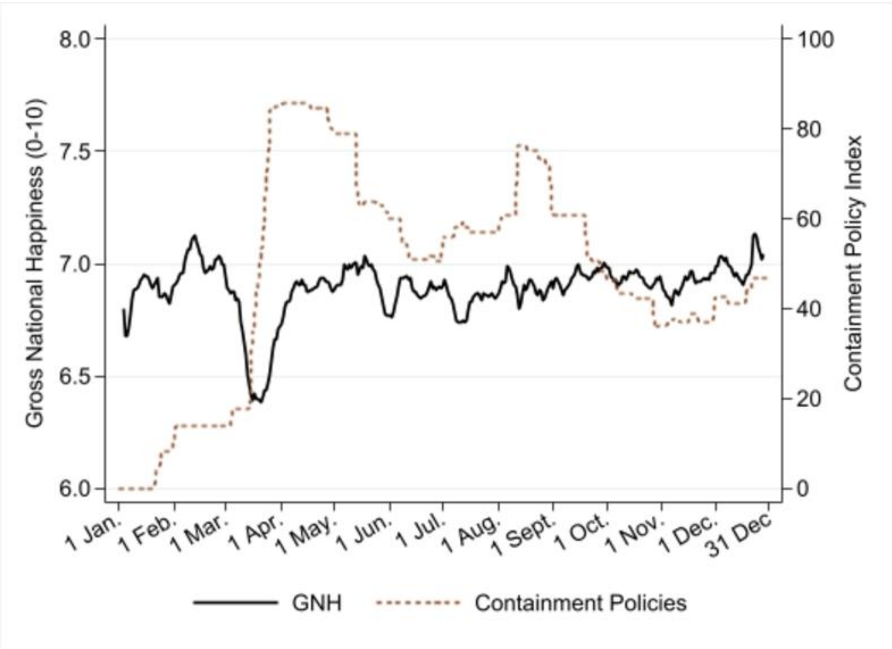
In the next section, we explore the joint effects of multiple variables on GNH using regression analysis.

4.1 Regression results

Previous literature indicates that the pandemic negatively affects well-being through multiple channels. To simultaneously account for the role of the various possible explanatory factors, we adopt regression analysis. For policies, we distinguish between the effect of an expected increase in containment policies from the one of an expected decrease. The expectation that policies will become more stringent should hamper well-being, as it indicates less freedom and signals bad times ahead. On the contrary, we expect that the relaxation of policies should correlate positively with people's well-being. We also assess the role of physical distancing, economic fear, trust in national institutions, loneliness, and generalised trust. We expect that distancing negatively affects well-being due to reduced freedom, autonomy and the possibility to gather socially. We expect that GNH decreases if people are concerned about the economy and feel lonely. On the contrary, generalised trust and trust in national institutions should positively correlate with well-being. We also include controls for the remaining emotions (anger, anticipation, disgust, fear, joy, sadness, and surprise, besides generalised trust) and seasons and months to account for unobserved factors such as weather.



a) Average daily data across seven European countries.

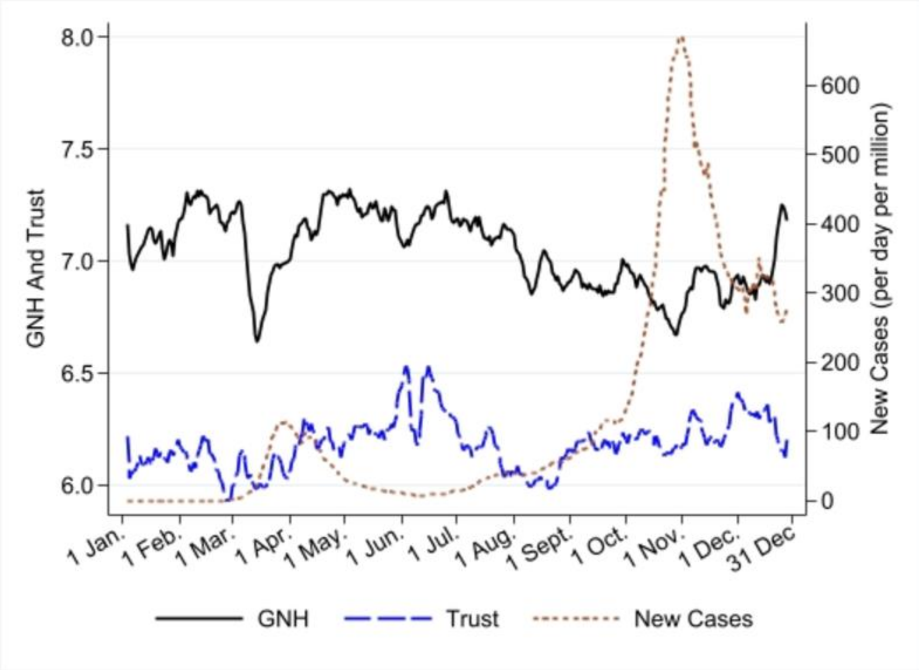


b) Average daily data across Australia, New Zealand and South Africa.

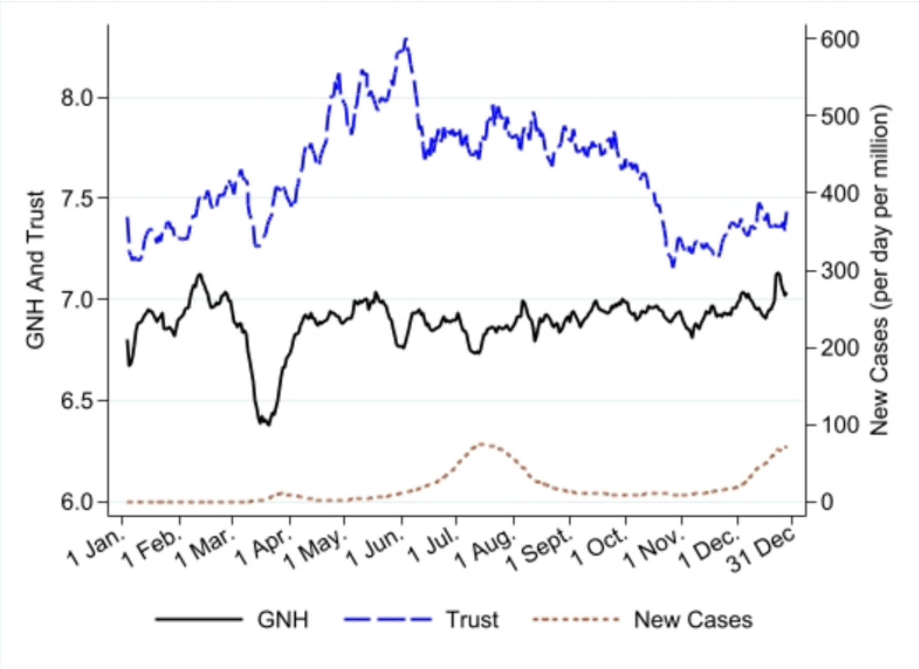
Fig 3. Gross National Happiness and Containment Policies. Average daily data across ten countries.

Note: GNH is presented using seven-day (centered) moving averages.

Source: GNH data (Greyling et al. [15]) are sourced from the project "Preferences Through Twitter" with the support of FNR, UJ and AUT. The policy index is sourced from Oxford Policy Tracker.



a) Average daily data across seven European countries.



b) Average daily data across Australia, New Zealand and South Africa.

Fig 4. Gross National Happiness, broad trust and new positive cases of COVID-19 in 2020.
Note: Data are presented using seven-day (centered) moving averages.
Source: GNH data (Greyling et al. [15]) and generalised trust are sourced from the project "Preferences Through Twitter" with the support of FNR, UJ and AUT. The number of new positive cases of COVID-19 is sourced from OurWorldinData.org.

The estimated regression model, which accounts for the time-series properties of the data, is given by the following equation:

$$GNH_{it} = \alpha + \rho GNH_{it-1} + \beta_1 IHS(Cases)_{it} + \beta_2 Distancing_{it} + \beta_3 Decr.Cont.Policies_{it+1} + \beta_4 Incr.Cont.Policies_{it+1} + \beta_5 Emotions_{it} + \beta_6 EconFear_{it} + \beta_7 GenTrust_{it} + \beta_8 InstTrust_{it} + \beta_9 Loneliness_{it} + \beta_1 0X_{it} + \epsilon_{it} \quad (1)$$

Where GNH_{it}represents the average Gross National Happiness for country i in week t. $IHS(Cases)$ represents the inverse hyperbolic sine of the average number of new cases per million in a week. $Distancing$ is the index of residential mobility. $Decr.Cont.Policies_{it+1}$ is the expected decrease of the containment policy index in the following week. It is a dummy variable set to one if the index decreases at $t+1$, zero otherwise. In detail, $Decr.Cont.Policies_{it+1} = 1\ if\ Cont.Policies_{it+1} - Cont.Policies_{it} < 0; or\ Decr.Cont.Policies_{it+1} = 0\ if\ Cont.Policies_{it+1} - Cont.Policies_{it} \geq 0$

$Incr.Cont.Policies_{it+1}$ is the reciprocal: the expected increase of the policy index in the following week. It is set to one if containment policies increase at $t+1$, zero otherwise. We also control for economic fear, generalised trust, trust in national institutions and loneliness. $Emotions$ is a vector including the emotions previously mentioned. X is a vector of control variables, including dummies for each month and season.

Statistical significance is assessed using Wild Cluster Bootstrap methods. Clustering standard errors at the country level are necessary because of the strong persistence in both the dependent and independent variables within a country. Bootstrap methods are needed because the number of countries is small. A small number of clusters leads to rejecting the null hypothesis relatively more frequently, in some cases at more than double the critical value (Bertrand et al. [68]). Wild Cluster Bootstrap methods resample over clusters, and using Webb weights, are particularly intended to accommodate scenarios with less than ten clusters. The limitation of the Wild Cluster Bootstrap method is that only the p-values from the bootstrap distribution can be obtained to assess the significance of coefficients (Cameron and Miller [69]).

Table 4 presents the first set of results. Regressors are included step-wise to check their association to GNH, before and after controlling for additional variables. Results are presented sequentially in columns 1 to 6.

Table 4. Association between the number of positive cases, physical distancing, expected increase and decrease of policy stringency, and GNH.

Variable	(1)	(2)	(3)	(4)	(5)	(6)
Lag GNH	0.907	0.906	0.93	0.908	0.922	0.906
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
∆ IHS New Cases		-0.064			-0.039	-0.048
		[0.010]			[0.007]	[0.021]
Residential - mobility			0.006		0.004	
			[0.000]		[0.003]	
F. Decr. Stringency				-0.026	-0.032	-0.029
				[0.079]	[0.022]	[0.043]
F. Incr. Stringency				-0.051	-0.049	-0.044
				[0.005]	[0.007]	[0.035]
Spring	0.197	0.173	0.186	0.199	0.175	0.18
	[0.078]	[0.110]	[0.058]	[0.082]	[0.090]	[0.107]
Summer	-0.013	-0.012	0.005	-0.004	0.01	-0.004
	[0.339]	[0.270]	[0.613]	[0.679]	[0.696]	[0.689]
Fall	0.179	0.157	0.174	0.189	0.17	0.17
	[0.087]	[0.115]	[0.045]	[0.091]	[0.093]	[0.117]
Constant	0.63	0.643	0.486	0.652	0.536	0.664
Month Controls	yes	yes	yes	yes	yes	yes
N	510	510	460	500	450	500
Adj. R Sq.	0.837	0.839	0.842	0.844	0.849	0.845
# of Countries	10	10	10	10	10	10

Bootstrapped p-values in brackets.

Coefficients on dummies for the months of the year are omitted for brevity. The auto- regressive term, the lagged value of GNH, reported in the first row, has a high and significant coefficient, which indicates that the variation in current GNH depends largely on its previous realisations: lag GNH, along with the month and season controls, explains nearly 84% of the overall variance. This indicates that GNH is a relatively stable variable that is not easily affected by volatile events. Auto-correlation is common in daily time series. In this case, it also indicates a desirable feature in a measure of well-being: "in a world of bread and circuses" (Deaton [70]), GNH seems to capture the bread more than the circuses.

Dummies for Spring and Fall have positive and significant (10 per cent) coefficients, indicating that GNH tends to be on average higher during those seasons. The coefficient for Summer is not statistically different from zero. Models 2 to 4 add, respectively, new cases, the index of residential mobility, and the expected changes in containment policy to the baseline (model 1). We find a negative and significant coefficient for new cases and both expected increase and decrease of containment policies. The index of residential mobility attracts a positive and statistically significant coefficient, indicating that, ceteris paribus, staying at home favoured GNH. Notice that the number of observations reduces by 50 when we control for mobility. The reason is that Google Mobility data are available from the beginning of the pandemic (mid-February); thus, the initial weeks of 2020 are missing. Model 6 shows that the results of the complete model (column 5) do not depend on the smaller sample size due to the inclusion of the control for residential mobility.

Results do not change when we include all the controls at the same time (Model 5). The coefficient of the adjusted R-squared indicates that the full model explains 85% of the total variance, which slightly improves the baseline. All coefficients maintain their sign, magnitude and statistical significance.

Table 5 presents results for an extended model, which includes the emotions, controls for economic fear, trust in national institutions and loneliness. For ease of comparison, the first column reports the same specification of column 5 in Table 4. Also, in this case, we included variables step-wise. To minimise the number of controls and preserve degrees of freedom, we applied a step-wise selection process for the emotions. We kept the emotions with a p-value (after Wild Cluster Bootstrap) smaller than 0.4. After this selection process, only disgust, fear, surprise and trust are retained (column 3). Months and season controls are included in the estimates but omitted from the table. The full model's adjusted R squared indicates that 92.7% of the total variance is explained, an improvement over the initial model of nearly seven percentage points, despite the decrease in the number of countries. Note that Luxembourg is excluded from the analysis because tweets about economic conditions, national institutions and loneliness are scarce.

Table 5. Regressions of GNH on the complete set of control variables. Average weekly values by country.

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Lag GNH	0.922	0.756	0.76		0.788	0.788	0.778	0.788	0.776
	[0.000]	[0.002]	[0.000]		[0.001]	[0.001]	[0.001]	[0.001]	[0.001]
Δ IHS New Cases	-0.039	-0.03	-0.027	-0.067	-0.031	-0.03	-0.031	-0.031	-0.032
	[0.007]	[0.005]	[0.025]	[0.092]	[0.011]	[0.021]	[0.014]	[0.010]	[0.020]
Residential - mobility	0.004	0.007	0.007	0.001	0.006	0.006	0.006	0.006	0.006
	[0.003]	[0.001]	[0.000]	[0.975]	[0.000]	[0.000]	[0.001]	[0.000]	[0.000]
F. Decr. Stringency	-0.032	-0.027	-0.026	-0.028	-0.017	-0.017	-0.017	-0.017	-0.016
	[0.022]	[0.166]	[0.124]	[0.975]	[0.184]	[0.218]	[0.207]	[0.118]	[0.145]
F. Incr. Stringency	-0.049	-0.036	-0.034	-0.032	-0.03	-0.03	-0.031	-0.03	-0.031
	[0.007]	[0.018]	[0.012]	[0.348]	[0.013]	[0.001]	[0.016]	[0.014]	[0.004]
Anger		-0.079							
		[0.657]							
Anticipation		-0.075							
		[0.319]							
Disgust		-0.13	-0.28	-1.376	-0.357	-0.344	-0.332	-0.359	-0.337
		[0.519]	[0.000]	[0.106]	[0.003]	[0.014]	[0.016]	[0.003]	[0.019]
Fear		-0.051	-0.124	-0.025	-0.041	-0.037	-0.056	-0.042	-0.064
		[0.220]	[0.219]	[0.918]	[0.407]	[0.533]	[0.170]	[0.433]	[0.178]
Joy		0.048							
		[0.366]							
Sadness		-0.12							
		[0.642]							
Surprise		0.327	0.329	0.942	0.301	0.3	0.289	0.302	0.286
		[0.040]	[0.037]	[0.027]	[0.038]	[0.533]	[0.170]	[0.433]	[0.026]
Trust		0.128	0.11	0.343	0.102	0.107	0.127	0.102	0.13
		[0.124]	[0.103]	[0.423]	[0.036]	[0.028]	[0.029]	[0.035]	[0.010]
Economic Fear						-0.009			0.009
						[0.854]			[0.813]
Nat. Trust							-0.011		-0.014
							[0.786]		[0.595]
Loneliness (Sad)								0.001	0
								[0.933]	[0.993]
Constant	0.536	1.074	0.966	5.127	0.801	0.77	0.84	0.803	0.882
Month Controls	yes	yes	yes	-	yes	yes	yes	yes	yes
Season Controls	yes	yes	yes	-	yes	yes	yes	yes	yes
N	450	450	450	450	405	405	405	405	405
Adj. R Sq.	0.849	0.881	0.88	0.527	0.928	0.928	0.928	0.927	0.927
# of Countries	10	10	10	10	9	9	9	9	9

Bootstrapped p-values in brackets.

The auto-correlation term still explains a large part of this variability, but its coefficient decreased from 0.922 to 0.776. The changes in the number of new infections and the expected increase of containment policies maintain their negative sign, magnitude and significance. The coefficient of the expected decrease in policy stringency is no longer statistically significant. Ceteris paribus, an increase in the number of people staying at home, remains positively and significantly associated with GNH changes.

GNH also grows when trust and surprise increase and decreases when disgust increases. These results are not surprising: it is well established that trust correlates with well-being both cross-sectionally and over time. Surprise and disgust are emotions that change with mood and correlate with more volatile aspects of well-being, sometimes referred to as effective or momentary subjective well-being. Economic fear, trust in national institutions, and sadness about loneliness are not significant. The reasons for this result are not clear.

In a further specification, we split changes in new positive cases into two variables: increases and decreases of new cases. This allows us to study the symmetry of the effect of contagion on GNH (results are omitted for reasons of space and can be made available upon request to the authors). We found that an increase in new cases correlates negatively and significantly with GNH, whereas a decrease does not attract a statistically significant coefficient. This suggests that a worsening pandemic situation affects GNH much more than an improvement.

If we compare model 3, which includes Luxembourg, with model 5, we observe that the exclusion of Luxembourg increases our ability to explain the overall variance by 4.8 percentage points (from 88% to 92.8%). This is probably because the time series for Luxembourg are more volatile than those of the other countries. In terms of results, the exclusion of Luxembourg does not change our general findings. Model 4 is identical to model 3, except it excludes the autoregressive term and month and seasonal controls. It is intended to check the robustness of results when increasing the degrees of freedom and removing the strong influence of lagged GNH. As expected, the adjusted R squared decreases considerably, from 88% to 52.7%. The coefficients of the variables of interest maintain their signs, but most of them lose significance: only the change in the number of new positive cases and surprise remain statistically significant.

In sum, our final model seems to explain GNH changes during 2020 in the studied countries rather well. A large part of the overall variation is explained by the autoregressive term - which is to be expected and signals that GNH is rather stable throughout the weeks of the year - as well as the month and seasonal controls. GNH decreases when the virus spreads, particularly when new positive cases increase and containment measures become more stringent. Under these circumstances, an increase in people staying at home predicts an increase in GNH. This is probably because people feel safe if they stay at home in a difficult or dangerous situation. Disgust and GNH are negatively associated, whereas we find a positive association with surprise and trust.

5. Conclusions

This work aimed to describe and explain changes in well-being that occurred in 2020 - the year marked by the outbreak of the novel coronavirus pandemic - using novel, timely data on happiness. The considered countries include Australia, New Zealand, South Africa, and seven European countries: Belgium, France, Germany, Great Britain, Italy, Luxembourg and Spain. To this purpose, we created a dataset that includes daily observations on well-being, emotions, economic conditions (unemployment), new infections, distancing behaviour, and containment policy. Well-being is measured by Gross National Happiness (GNH), an aggregate country-level variable derived by applying sentiment analysis to Twitter data. We also used sentiment analysis to derive eight emotions and three additional variables: economic fear, trust in national institutions, and sadness about loneliness. As far as we know, this is the first time that this kind of information has been derived from sentiment analysis and used in an empirical analysis. A final contribution of this work is testing the validity of the measures produced using sentiment analysis. Thus, we indirectly contribute to exploring using Big Data and machine learning for compiling and analysing social and economic statistics.

Results indicate that GNH is a valid measure of well-being, as it correlates meaningfully with alternative measures of well-being, and ill-being, from surveys and other Big Data sources, such as Google Trends. The same holds for economic fear, trust in national institutions, and generalised trust.

The availability of data from Twitter reveals a much more varied picture than the snapshots provided by surveys: while countries had on average a GNH score of seven, our data indicate that GNH changed substantially in correspondence with the evolution of the pandemic. This means we could get a distorted view of well-being and other variables, depending on when a survey (snapshot) is administered. For instance, our data indicate that well-being exhibited considerable variation over the studied year: the first pandemic wave featured a sudden decline of GNH followed by a rapid recovery in all countries. Following this, the evolution of GNH exhibited varied patterns across countries. In particular, the second wave of contagion was accompanied by a prolonged decline in GNH in Europe. In Australia, New Zealand, and South Africa, a second period of decline of GNH started in mid-May. It reached a peak at the beginning of July, before recovering to its pre-pandemic levels.

What explains the changes in GNH during 2020? GNH decreases in correspondence with rising infection waves and when strict containment policies are in place. To account for the simultaneous effect of various factors on the changes of GNH over time, we used regression analysis. Once accounted for the time series structure of the data and seasonal factors, we found that GNH changes correlate negatively with changes in new positive cases (and, in particular, the increases) and with the expected increase in containment policy stringency. An increase in people staying at home predicts an increase in GNH. In other words, ceteris paribus, the more time spent at home, the higher was GNH. This can be explained by an increased sense of protection and "altruism" - intended as own contribution to the fight against the spread of the virus - associated with increased distancing. Results also indicate that economic fear, trust in national institutions and sadness about loneliness are not significantly associated with changes in GNH. This is puzzling but could indicate that health and lockdown concerns dominated individuals' mood during the pandemic. Finally, we found that GNH correlates positively when surprise and generalised trust increase and disgust decreases. Among these variables, trust is a relevant one, as previous studies showed that higher trust correlates with higher compliance to containment policies and contributes to social cohesion and economic activity.

In sum, our study provided several insights. Firstly, this study showed that sentiment analysis applied to Twitter data could generate timely and frequent measures of well-being and other variables relevant for economic and political decisions, such as generalised trust, trust in institutions, loneliness and economic fear. Secondly, we find that such data are valid as they correlate meaningfully with data from surveys and other sources of Big Data. Thus, figures issued from sentiment analysis of Twitter data can valuably complement survey data to provide insights for the general public, the research community, and policymakers. Finally, changes in GNH during the pandemic correlate significantly with new infections, policy stringency, residential mobility and trust. These correlations suggest that GNH covers both cognitive and affective aspects of life as a measure of well-being.

Acknowledgements

We would like to thank the Luxembourg National Research Fund (FNR) for the generous funding of the project. Additionally, we thank the University of Johannesburg and Auckland University of Technology for the funds contributed. We also thank AFSTEREO for the I.T. service provided.

References

1. Danner DD, Snowdon DA, Friesen WV. Positive emotions in early life and longevity: findings from the nun study. Journal of Personality and Social Psychology. 2001; 80(5):804.
2. Guven C, Saloumidis R. Life satisfaction and longevity: longitudinal evidence from the German socioeconomic panel. German Economic Review. 2014; 15(4):453-472.
3. Frijters P, Haisken-DeNew JP, Shields MA. The increasingly mixed proportional hazard model: An application to socioeconomic status, health shocks, and mortality. Journal of Business & Economic Statistics. 2011; 29(2):271-281.
4. Graham C. Pinto S. Unequal hopes and lives in the USA: Optimism, race, place, and premature mortality. Journal of Population Economics. 2019; 32(2):665-733.
5. O'Connor KJ. Life satisfaction and noncognitive skills: Effects on the likelihood of unemployment. Kyklos. 2020; 73(4):568-604.
6. Lyubomirsky S, Sheldon KM, Schkade D. Pursuing happiness: The architecture of sustainable change. Review of General Psychology. 2005; 9:111-131.
7. Bryson A, Clark AE, Freeman RB, Green CP. Share capitalism and worker wellbeing. Labour Economics. 2016; 42:151-158.

8. DiMaria CH, Peroni C, Sarracino F. Happiness matters: productivity gains from subjective well-being. Journal of Happiness Studies. 2020; 21(1):139-160.
9. Oswald AJ, Proto E, Sgroi D. Happiness and productivity. Journal of Labor Economics. 2015; 33(4):789-822.
10. Krekel C, Swanke S, De Neve JE, Fancourt D. Are happier people more compliant? Global evidence from three large-scale surveys during Covid-19 lockdowns. IZA Discussion Paper. 2020; 13690, IZA.
11. Bonanno GA, Brewin CR, Kaniasty K, Greca AML. Weighing the costs of disaster: Consequences, risks, and resilience in individuals, families, and communities. Psychological Science in the Public Interest. 2010; 11(1):1-49.
12. Kessler RC, Galea S, Jones RT, Parker HA. Mental illness and suicidality after Hurricane Katrina. Bulletin of the World Health Organization. 2006; 84(12):930-939.
13. Norris FH, Friedman MJ, Watson PJ. 60,000 disaster victims speak: Part ii. Summary and implications of the disaster mental health research. Psychiatry: Interpersonal and Biological Processes. 2002; 65(3):240-260.
14. Arampatzi E, Burger M, Ianchovichina E, Rohricht T, and Veenhoven R. Unhappy development: Dissatisfaction with life on the eve of the Arab Spring. Review of Income and Wealth. 2018; 64: S80-S113.
15. Greyling T, Rossouw S, Afstereo. Gross National Happiness Index. 2019. Available at http://gnh.today Accessed 5 January 2021.
16. European Commission. Standard Eurobarometer 93 - public opinion in the European Union, first results, annex. 2020.
17. Brodeur A, Clark AE, Fleche S, Powdthavee N. Assessing the impact of the coronavirus lockdown on unhappiness, loneliness, and boredom using google trends. Technical Report preprint arXiv. 2020; 2004.12129.
18. Greyling T, Rossouw S, Adhikari T. The good, the bad and the ugly of lockdowns during COVID-19. PLoS ONE. 2021; 16(1):e0245546.

19. Brooks SK, Webster RK, Smith LE, Woodland L, Wessely S, Greenberg N, Rubin GJ. The psychological impact of quarantine and how to reduce it: rapid review of the evidence. The Lancet. 2020; 395(10227):912-920.
20. Holmes EA, O'Connor RC, Perry VH, Tracey I, Wessely S, Arseneault L, Ballard C, Christensen H, Cohen Silver R, Everall I, Ford T, John A, Kabir T, King K, Madan I, Michie S. Przybylski AK, Shafran R, Sweeney A, Worthman CM, Yardley L, Cowan K, Cope C, Hotopf M, Bullmore E. Multidisciplinary research priorities for the COVID-19 pandemic: a call for action for mental health science. The Lancet Psychiatry. 2020; 7(6):547-560.
21. Blasco-Belled A, Tejada-Gallardo C, Torrelles-Nadal C, Alsinet C. The costs of the COVID-19 on subjective well-being: an analysis of the outbreak in Spain. Sustainability. 2020; 12(15):6243.
22. Cooke JE, Eirich R, Racine N, Madigan S. Prevalence of posttraumatic and general psychological stress during COVID-19: a rapid review and meta-analysis. Psychiatry Research. 2020; 292:113347.
23. Rajkumar RP. Covid-19 and mental health: A review of the existing literature. Asian Journal of Psychiatry. 2020; 52:102066.
24. Xiong J, Lipsitz O, Nasri F, Lui LM, Gill H, Phan L, Chen-Li D. Iacobucci M, Ho R, Majeed A, et al. Impact of COVID-19 pandemic on mental health in the general population: A systematic review. Journal of Affective Disorders. 2020; 277:55-64.
25. Saladino V, Algeri D, Auriemma V. The psychological and social impact of Covid-19: new perspectives of well-being. Frontiers in Psychology. 2020; 11:2550.
26. Li LZ, Wang S. Prevalence and predictors of general psychiatric disorders and loneliness during COVID-19 in the United Kingdom. Psychiatry Research. 2020; 291:165-1781.
27. Cao W, Fang Z, Hou G, Han M, Xu X, Dong J, Zheng J. The psychological impact of the COVID-19 epidemic on college students in China. Psychiatry Research. 2020; 287:112934.
28. Salari N, Hosseinian-Far A, Jalali R, Vaisi-Raygani A, Rasoulpoor S, Mohammadi M, Rasoulpoor S, Khaledi-Paveh B. Prevalence of stress, anxiety, depression among the general

population during the COVID-19 pandemic: a systematic review and meta-analysis. Globalization and Health. 2020; 16(1):1-11.

29. Kawohl W, Nordt C. COVID-19, unemployment, and suicide. The Lancet Psychiatry. 2020; 7(5):389-390.
30. Krendl AC, Perry BL. The impact of sheltering in place during the COVID-19 pandemic on older adults' social and mental well-being. The Journals of Gerontology: Series B. 2021; 76(2): e53-e58.
31. Sibley CG, Greaves L, Satherley N, Wilson M, Lee C, Milojev P, et al. Short-term Effects of the COVID-19 Pandemic and a Nationwide Lockdown on Institutional Trust, Attitudes to Government, Health and Wellbeing. APA PsycNet 2020; 75(5): 618–63.
32. Patrick SW, Henkhaus LE, Zickafoose JS, Lovell K, Halvorson A, Loch S, Letterie M, Davis MM. Well-being of parents and children during the COVID-19 pandemic: a national survey. Pediatrics. 2020; 146(4) https://doi.org/10.1542/peds.2020-0836
33. O'Connor KJ, Peroni C. One in three Luxembourg residents report their mental health declined during the covid-19 crisis. International Journal of Community Well-Being. 2020; 1-7.
34. Sønderskov KM, Dinesen PT, Santini ZI, Ostergaard SD. Increased psychological well-being after the apex of the COVID-19 pandemic. Acta Neuropsychiatrica. 2020; 32(5):277-279.
35. Recchi E, Ferragina E, Helmeid E, Pauly S, Safi M, Sauger N, Schradie J. The "eye of the hurricane" paradox: an unexpected and unequal rise of well-being during the Covid-19 lockdown in France. Research in Social Stratification and Mobility. 2020; 68:100508.
36. Rossouw S, Greyling T, Adhikari T. Happiness-lost: Did Governments make the right decisions to combat Covid-19? South African Journal of Economic and Management Sciences. 2021; 24(1): a3795. https://doi.org/10.4102/sajems.v24i1.3795
37. Greyling T, Rossouw S, Adhikari T. A tale of three countries: How did Covid-19 lockdown impact happiness? South African Journal of Economics. 2021; 89(1): 25-43.
38. Foa R, Gilbert S, Fabian MO. Covid-19 and subjective well-being: Separating the effects of lockdowns from the pandemic. 2020. Available at SSRN 367000.

39. Kivi M, Hansson I, Bjälkebring P. Up and about: Older adults' well-being during the COVID-19 pandemic in a Swedish Longitudinal Study. The Journals of Gerontology: Series B. 2021; 76(2): e4-e9.
40. Wang C, Pan R, Wan X, Tan Y, Xu L, McIntyre RS, Choo FN, Tran B, Ho R, Sharma VK, et al. A longitudinal study on the mental health of general population during the COVID-19 epidemic in China. Brain, Behavior, and Immunity. 2020; 87:40-48.
41. Brülhart M, Lalive R. Daily suffering: Helpline calls during the COVID-19 crisis. Covid Economics. 2020; 19:143-58.
42. Rossouw S, Greyling T, Adhikari T. The evolution of happiness pre and peri-COVID-19: a Markov Switching Dynamic Regression Model. Forthcoming in PLoS ONE. 2021.
43. Cheng TC, Kim S, Koh K. The Impact of COVID-19 on Subjective Well-Being: Evidence from Singapore.Callegaro M, Yang Y. The role of surveys in the era of "big data". In Vannette, D. and Krosnick, J., editors, The Palgrave handbook of survey research. Palgrave Macmillan, Cham. 2020. pp 175-192.
44. Dodds PS, Danforth CM. Measuring the happiness of large-scale written expression: Songs, blogs, and presidents. Journal of Happiness Studies. 2010; 11:441-456.
45. Dodds PS, Harris KD, Kloumann IM, Bliss CA, Danforth CM. Temporal patterns of happiness and information in a global social network: Hedonometrics and Twitter. PLoS ONE. 2011; 6(12):e26752.
46. Ceron A, Curini L, Iacus SM. Isa: A fast, scalable and accurate algorithm for sentiment analysis of social media content. Information Sciences. 2016; 367:105-124.
47. Iacus S, Porro G, Salini S, Siletti E. An Italian composite subjective well-being index: The voice of Twitter users from 2012 to 2017. Social Indicators Research. 2020;1-19.
48. Diener E, Lucas R. Schimmack U, Helliwell J. Well-being for public policy. 2009. Oxford University Press, New York.
49. Hailong Z, Wenyan G, Bo J. Machine learning and lexicon-based methods for sentiment classification: A survey. In 2017 11th web information system and application conference. IEEE. 2014. pp 262-265.

50. Eichstaedt JC, Schwartz HA, Kern ML, Park G, Labarthe DR, Merchant RM, Jha S, Agrawal M, Dziurzynski LA, Sap M. Psychological language on Twitter predicts county-level heart disease mortality. Psychological Science. 2015; 26(2):159-169.
51. Caldarelli G, Chessa A, Pammolli F, Pompa G, Puliga M, Riccaboni M, Riotta G. A multi-level geographical study of Italian political elections from Twitter data. PLoS ONE. 2014; 9(5):e95809.
52. Gayo-Avello D. A meta-analysis of state-of-the-art electoral prediction from Twitter data. Social Science Computer Review. 2013; 31(6):649-679.
53. Bollen J, Mao H, Zeng X. Twitter mood predicts the stock market. Journal of Computational Science. 2011; 2(1):1-8.
54. Asur S, Huberman BA. Predicting the future with social media. In 2010 IEEE/WIC/ACM international conference on web intelligence and intelligent agent technology, volume 1. IEEE. 2010; 492-499.
55. O'Connor B, Balasubramanyan R, Routledge B, Smith N. From tweets to polls: Linking text sentiment to public opinion time series, volume 4. 2010.
56. Lampos V, Cristianini N. Tracking the flu pandemic by monitoring the social web. In 2010 2nd international workshop on cognitive information processing. IEEE. 2010; 411-416.
57. Culotta A. Towards detecting influence epidemics by analyzing Twitter messages. In Hong, L. and Davison, B., editors, Proceedings of the first workshop on social media analytics. 2010. pp 115-122.
58. Paul M, Dredze M. You are what you tweet: Analyzing Twitter for public health. In Proceedings of the International AAAI Conference on Web and Social Media, volume 5. 2011.
59. Helliwell JF, Layard R, Sachs J, De Neve JE, editors. World Happiness Report. 2021. Sustainable Development Solutions Network, New York.
60. Roser M, Ritchie H, Ortiz-Ospina E, Hasell J. Coronavirus pandemic (COVID-19). 2021. Accessed: 11 January 2021. Available from: https://ourworldindata.org/coronavirus
61. Hale T, Angrist N, Cameron-Blake E, Hallas L, Kira B, Majumdar S, Petherick A, Phillips T, Tatlow H, Webster S. Oxford COVID-19 government response tracker. 2020. Accessed: 11

January 2021. Available from https://www.bsg.ox.ac.uk/research/research-projects/covid-19-government-response-tracker .

62. Google. Google COVID-19 Community Mobility Reports. 2020. Accessed: 11 January 2020Available from https://www.google.com/covid19/mobility/
63. Eurostat. Harmonised unemployment rates (%) - monthly data [ei_lmhr] total unemployment, percentage of active population (unadjusted). 2020b.
64. Fong MW, Gao H, Wong JY, Xiao J, Shiu EY, Ryu S, Cowling BJ. Nonpharmaceutical measures for pandemic influence in nonhealthcare settings-social distancing measures. Emerging Infectious Diseases. 2020; 26(5):976.
65. Chinazzi M, Davis JT, Ajelli M, Gioannini C, Litvinova M, Merler S, Piontti AP, Mu K, Rossi L, Sun K, et al. The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. Science. 2020; 368(6489):395-400.
66. Bargain O, Aminjonov U. Trust and compliance to public health policies in times of covid-19. Journal of Public Economics. 2020; 192:104316.
67. Bertrand M, Duflo E, Mullainathan S. How much should we trust differences-in-differences estimates? Quarterly Journal of Economics. 2004; 119(1):249-275.
68. Cameron AC, Miller DL. A practitioner's guide to cluster- robust inference. Journal of Human Resources. 2015; 50(2):317-72.
69. Deaton A. The financial crisis and the well-being of Americans 2011 OEP Hicks lecture. Oxford Economic Papers. 2012; 64(1):1-26.
70. OECD. Guidelines on Measuring Subjective Well-being. 2013. OECD.
71. University of Luxembourg. Come-here (covid-19, mental health, resilience and self-regulation) longitudinal survey. 2020.
72. Google LLC. Google trends. 2021. Accessed: January 2021.
73. Eurostat. Business and consumer surveys (dg ecfin) (ei_bcs) meta- data in euro sdmx metadata structure (esms). 2020a.

Appendix A. Validity of the Gross National Happiness Index.

This Appendix presents the results of the Gross National Happiness (GNH) validity tests as a measure of well-being. Assessing the validity of metrics based on unstructured data, such as Twitter data, is difficult because their features - timeliness, large (non-representative) samples, and high frequency - make them unique, thus limiting the availability of comparable measures. In other words, (objective or subjective) measures of well-being that are available with the same frequency and timeliness of GNH are scarce. One possibility is to correlate GNH with measures of well-being issued from other sources of Big Data, such as Google. However, the downside of this approach is that it relies on the assumption that data issued from Google are themselves valid.

We assess the validity of GNH using the following approaches: firstly, we check whether GNH correlates with survey measures of subjective well-being using cross-sectional country-level data; secondly, we test whether changes in GNH over time correlate significantly with some of the few sources of repeated observations on well-being available in 2020, and with Google data.

Correlation across countries

We first assess the validity of GNH by calculating its correlation with life satisfaction, a widely used measure of subjective well-being whose validity and reliability has been largely confirmed (OECD [71]). Observing a high correlation between GNH and life satisfaction indicates that GNH reflects similar factors affecting life satisfaction and suggests it is a valid measure of subjective well-being.

Our measure of life satisfaction is available from Eurobarometer surveys conducted by the European Commission. Eurobarometer surveys have been conducted biennially since the 1980s to measure public opinion in the European Union. Each survey wave is comprised of approximately 1000 face-to-face interviews in each country. Due to the COVID-19 pandemic, many countries could not conduct face-to-face interviews (including Luxembourg and the United Kingdom). Respondents answered online and were recruited using a probabilistic method by telephone. See European Commission [16] for additional details. It is measured as the response to the question, "On the whole, are you very satisfied, fairly satisfied, not very satisfied or not at all satisfied with the life you lead?" Response options are coded from one to four, with greater values representing greater satisfaction. We used the Standard Eurobarometer 93.1 fielded from 9 July to 26 August 2020 in the European Union, United Kingdom, and five EU candidate countries (European Commission [16]).

Fig 5 depicts the correlation of GNH with life satisfaction. Average GNH is computed by country over the period mid-July to the end of August, i.e., the months when Eurobarometer surveys were administered. The scatterplot indicates that there is a positive association between the two measures. The Spearman correlation coefficient is 0.37 (not statistically significant), which appears as an outlier if we exclude Italy. If we include Italy, the correlation coefficient is 0.32 but not statistically significant.

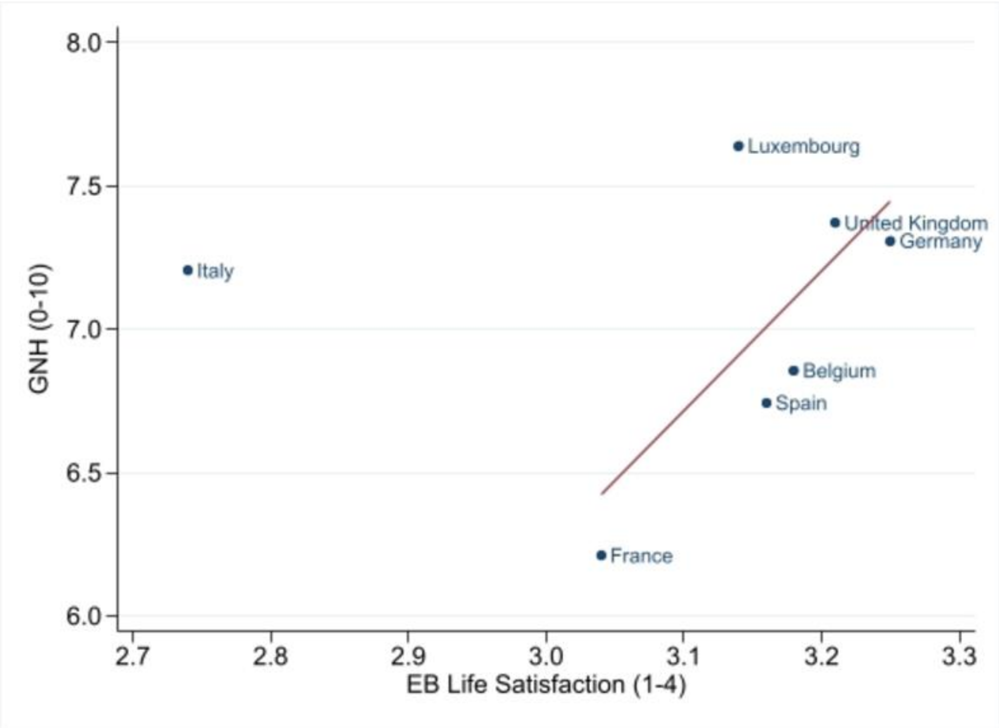


Fig 5. Average Gross National Happiness correlates positively with average life satisfaction.
Note: The GNH score is the average by country over the same period that the Eurobarometer was collected, from 9 July to 26 August 2020.
Source: GNH data (Greyling et al. [15]) are sourced from the project "Preferences Through Twitter" with the support of FNR, UJ and AUT. Life satisfaction data are from the Eurobarometer (European Commission [16]), Summer 2020.

Average GNH correlates meaningfully with the measure of well-being reported by the World Happiness Report 2021 (Helliwell et al. [60]), i.e., the average life evaluation from 2018 to 2020 (see Fig 6). Note the data on life evaluation in 2020 for Luxembourg are missing. In this case, the authors report the average over the years 2018-19. The report uses the Gallup World Poll data to rank countries from the happiest to the least happy. The Spearman correlation between GNH and average life evaluation is 48% ($Prob > |t| = 0.16, N = 10$). After excluding South Africa and France, which appear as outliers, the Spearman correlation is 20% ($Prob > |t| = 0.65, N = 8$).

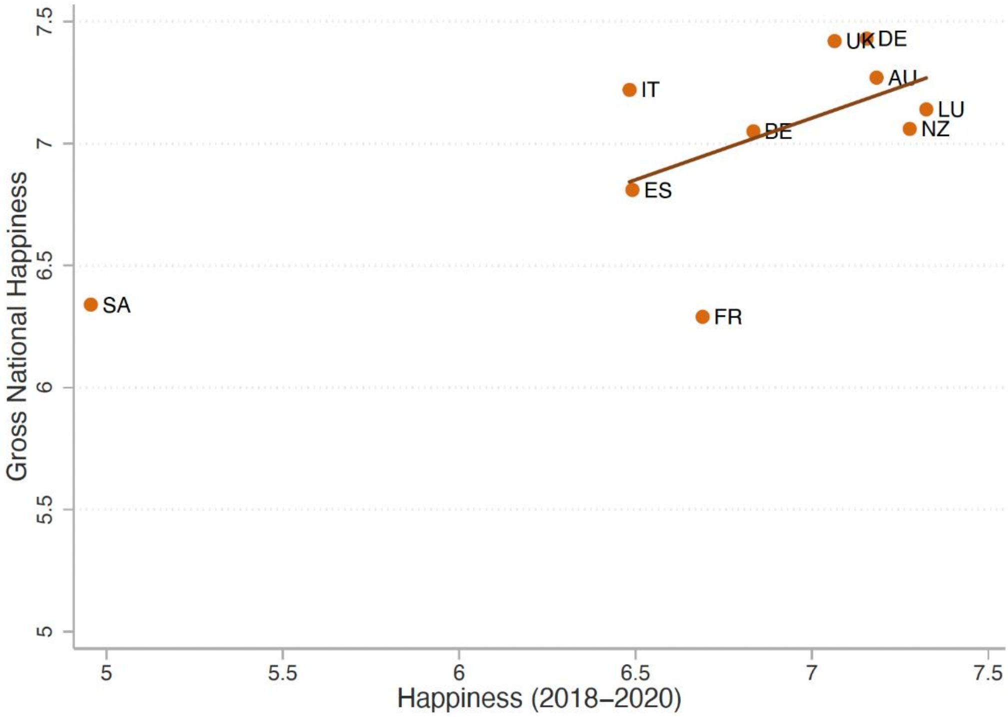


Fig 6. Correlation between average GNH in 2020 and average life evaluation (2018-2010) from the World Happiness Report 2021.
Note: average life evaluation is computed over the years 2018, 2019 and 2020. Data on Luxembourg exclude the year 2020. Average GNH is computed over the year 2020.
Source: GNH data (Greyling et al. [15]) are sourced from the project "Preferences Through Twitter" with the support of FNR, UJ and AUT. Average life evaluation is sourced from the World Happiness Report, 2021 (Helliwell et al. [60]).

Correlation over time

Correlation over time is an important test of validity for a measure that has the benefit of timeliness and frequency. For this purpose, we use three sources of repeated observations on well-being and ill-being for 2020: a survey conducted by the University of Luxembourg, Google Trends, and consumer confidence data provided by Eurostat.

University of Luxembourg' data on life satisfaction (University of Luxembourg [72]) have been collected via three surveys administered online to a convenience sample of residents in a selected number of European countries (for our purposes, data are available for France, Germany, Italy and Spain). Fig 7 shows that the two measures are poorly associated (the correlation coefficient is -0.26

$Prob > |t| = 0.622, N = 6$). GNH and life satisfaction seem to be trending together between August and November, but not between May and August. Another possibility is that GNH anticipates the changes in life satisfaction (the changes taking place between March and August would match well the figures from the University). Still, we did not find any support for this hypothesis.

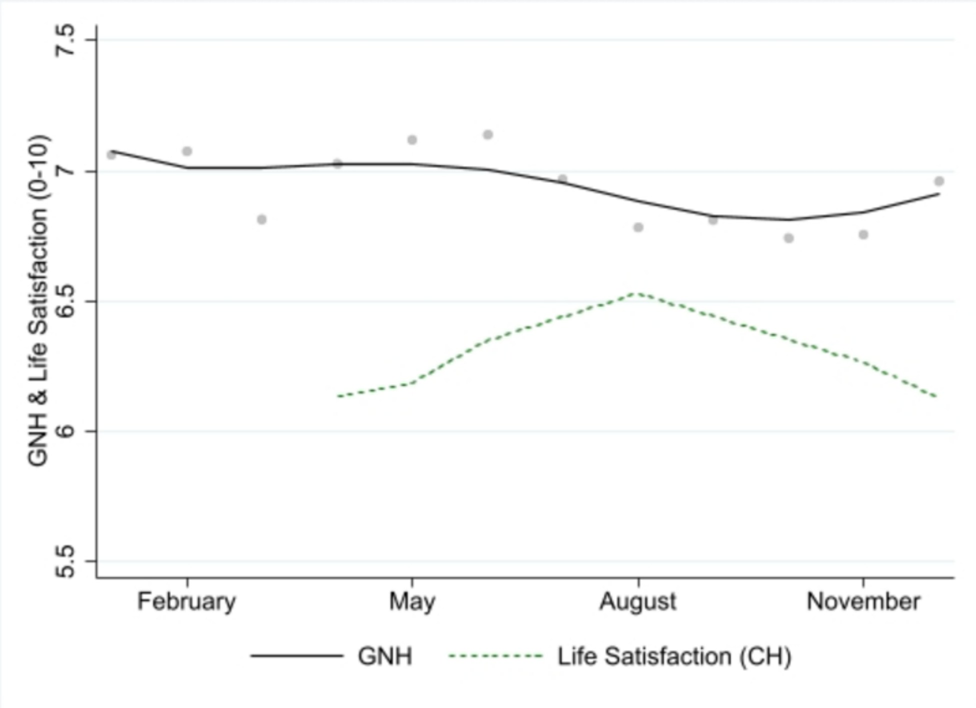


Fig 7. Gross National Happiness and average life satisfaction over time in four European countries (France, Germany, Italy and Spain).
Source: GNH data (Greyling et al. [15]) are sourced from the project "Preferences Through Twitter" with the support of FNR, UJ and AUT. Life satisfaction data are from COME-HERE (COVID-19, MEntal HEalth, REsilience and Self-regulation) longitudinal survey conducted by the University of Luxembourg (University of Luxembourg [72]).

Google Trends is a source of frequent data. Search results are available daily by country and used in numerous research projects ranging from assessing economic conditions to individuals' feelings (see, for instance, Brodeur et al. [17]). Rather than focusing on the trends of topics such as "happiness, well-being, or life satisfaction", which may not accurately reflect the well-being of Google users, we created an index of negative emotions (dashed line in Fig 8) by averaging daily Google search scores for three topics: fear, sadness, and anger. The correlation coefficient between GNH and the index of negative emotions is -0.27 ($Prob > |t| = 0.39, N = 12\ months$). The negative sign is to be expected, as the index of negative emotions should correlate negatively with a measure of well-being. A visual inspection of Fig 8 reveals that indeed GNH and the index of negative emotions move in the expected direction, as they document worsening well-being over 2020. The main discrepancy is observed for the first half of the year when GNH decreases less and more slowly than the index of negative emotions.

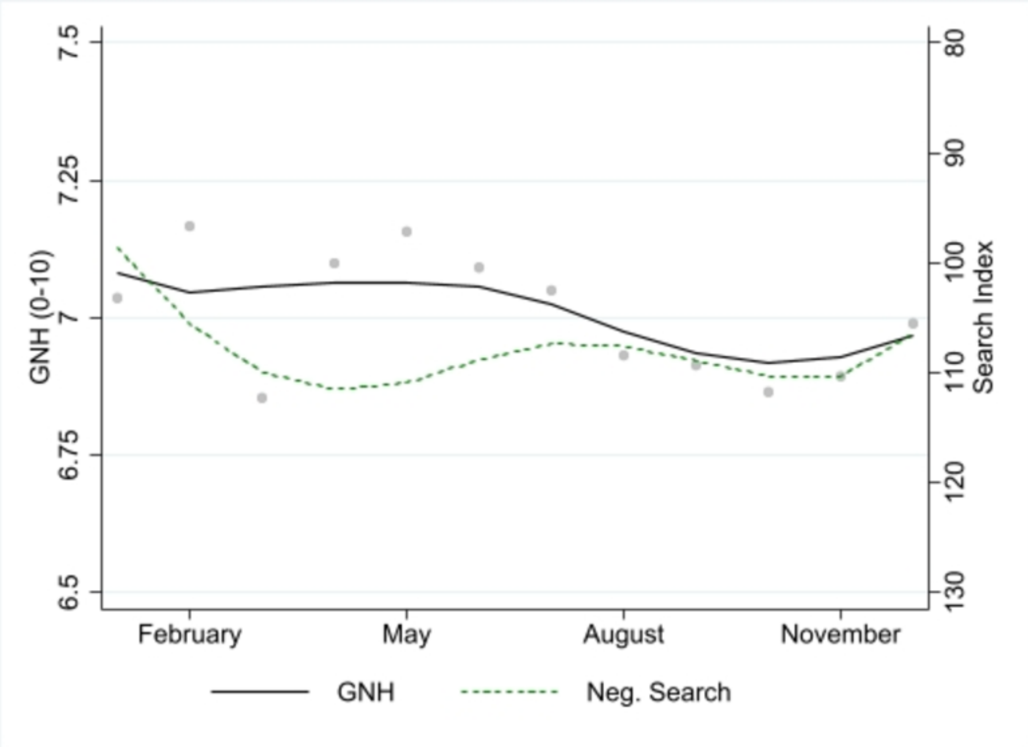


Fig 8. Gross National Happiness correlates meaningfully with the index of negative emotions over time.
Note: The index of negative emotions is the average of weakly averages of negative emotions (fear, sadness and anger) within countries.
Source: GNH data (Greyling et al. [15]) are sourced from the project "Preferences Through Twitter" with the support of FNR, UJ and AUT. The index of negative emotions is sourced from Google LLC [73].

Finally, we use consumer confidence data (Eurostat [74]) as a source of repeated observations to validate GNH. Although consumer confidence relates more to the economic and material domain of people's lives, it should positively correlate with GNH and are available at a relatively high frequency. Consumer confidence is monitored via monthly surveys administered by Eurostat to residents of European Union Member States. The final score is an index that averages positive and negative feelings of consumers in relation to their economic conditions and perspectives. Fig 9 shows that the monthly changes of GNH correlate positively with the changes in consumer confidence: the Spearman correlation coefficient is 0.5, ($Prob > |t| = 0.17, N = 9\ months$). Please note that three observations are missing because of the initial month, for which we cannot compute the change, and because data for Italy in the month of April are missing. Therefore, it was not possible to compute the changes relative to March and April and April and May.

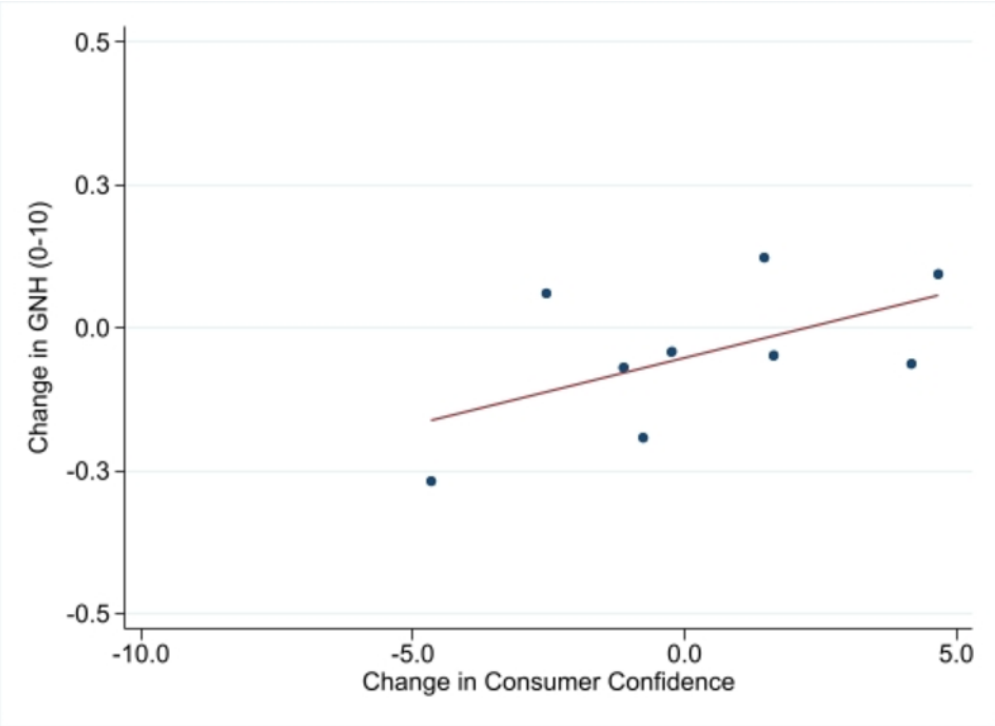
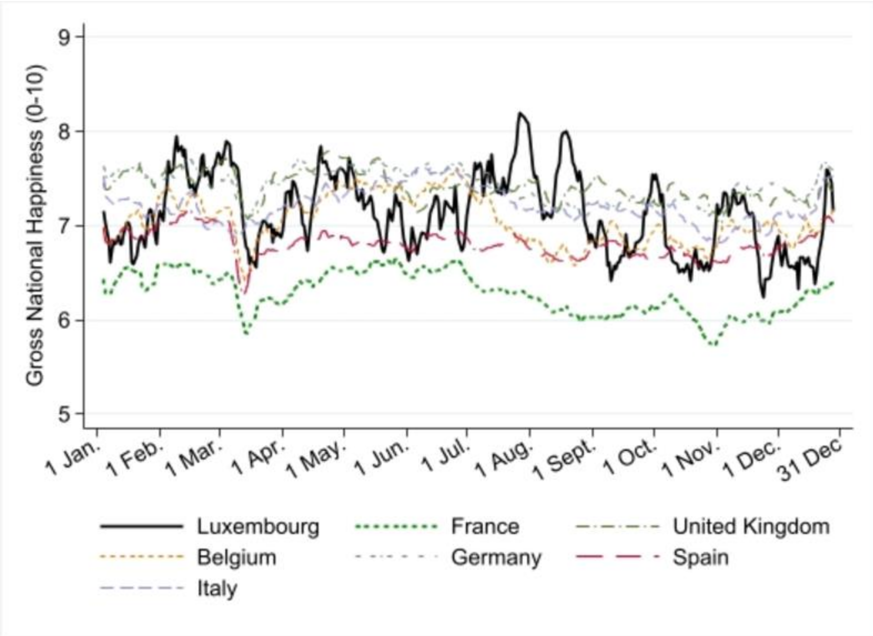
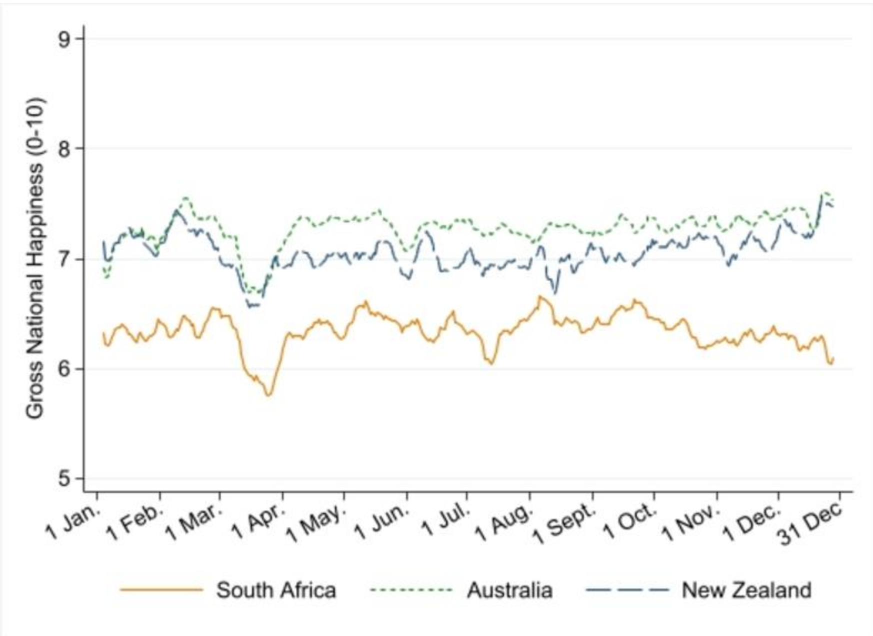


Fig 9. Gross National Happiness correlates meaningfully with consumer confidence data.
Note: Consumer confidence is a monthly index averaging positive and negative feelings about economic conditions and perspectives.
Source: GNH data (Greyling et al. [15]) are sourced from the project "Preferences Through Twitter" with the support of FNR, UJ and AUT. Consumer confidence data are from the European Commission (Eurostat [74]).

Appendix B. Evolution of GNH and other variables by country.



a) Average daily data across seven European countries.



b) Average daily data across Australia, New Zealand and South Africa.

Fig 10. Gross National Happiness by country in 2020.
Note: GNH is presented using seven-day (centered) moving averages.
Source: GNH data (Greyling et al. [15]) are sourced from the project "Preferences Through Twitter" with the support of FNR, UJ and AUT.

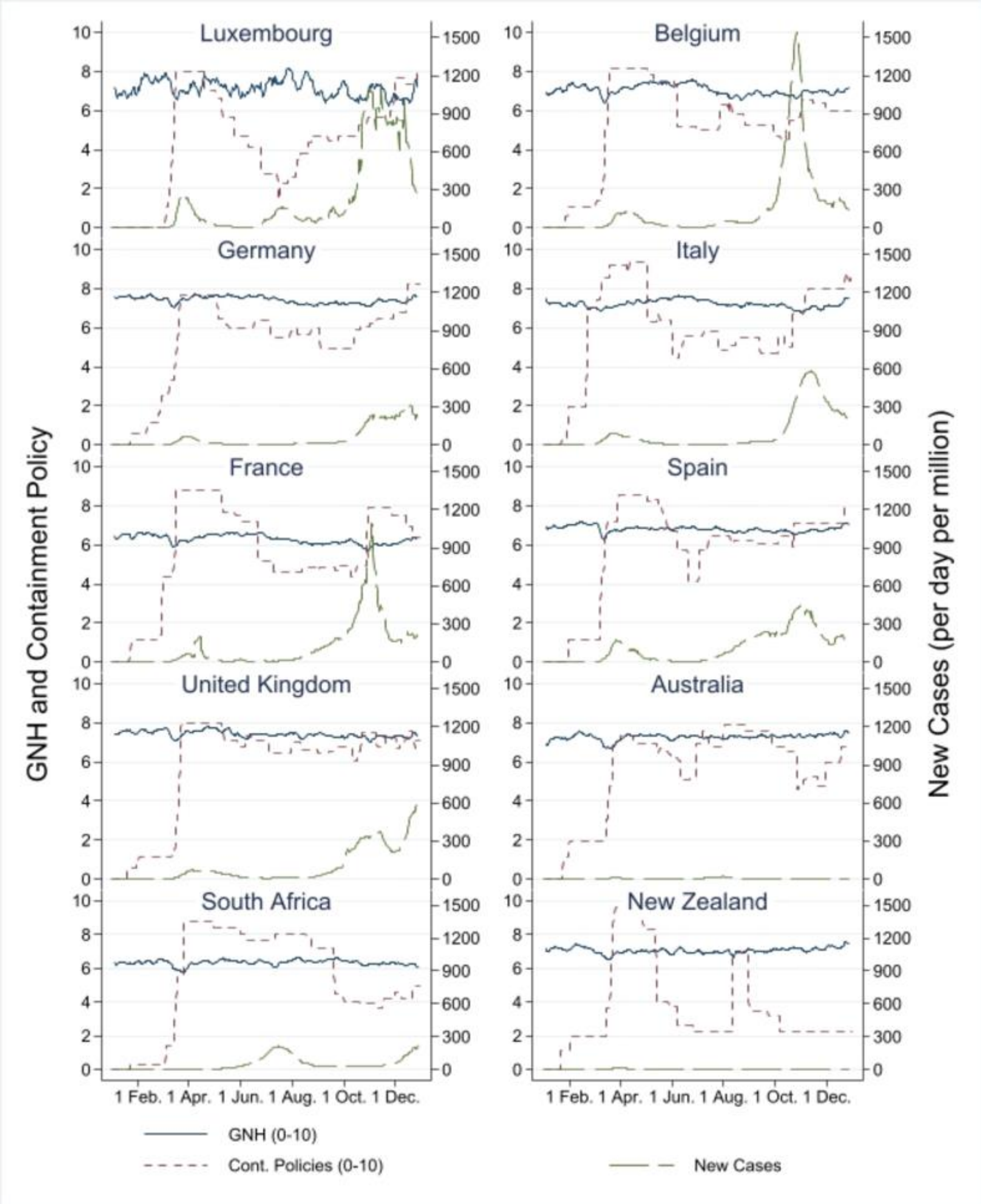


Fig 11. GNH, new positive cases, and containment policies by country.
Note: GNH and new positive cases are smoothed using seven-day (centered) moving averages. The Containment Policy Index values were divided by 10 to put them on the same scale as GNH.
Source: GNH data (Greyling et al. [15]) are sourced from the project "Preferences Through Twitter" with the support of FNR, UJ and AUT. The policy index is sourced from Oxford Policy Tracker.

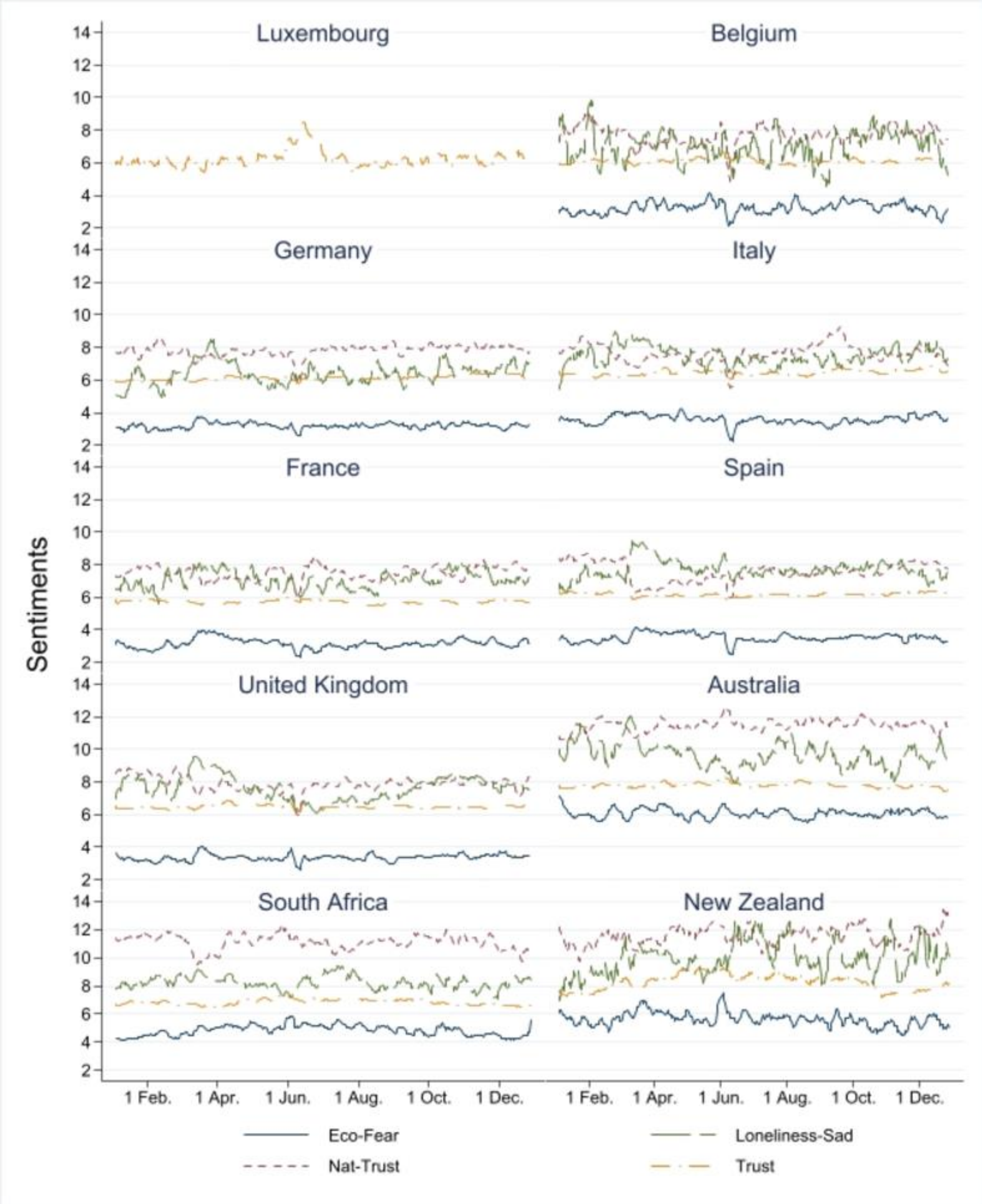


Fig 12. Economic fear, loneliness, trust in national institutions, and generalised trust by country in 2020.

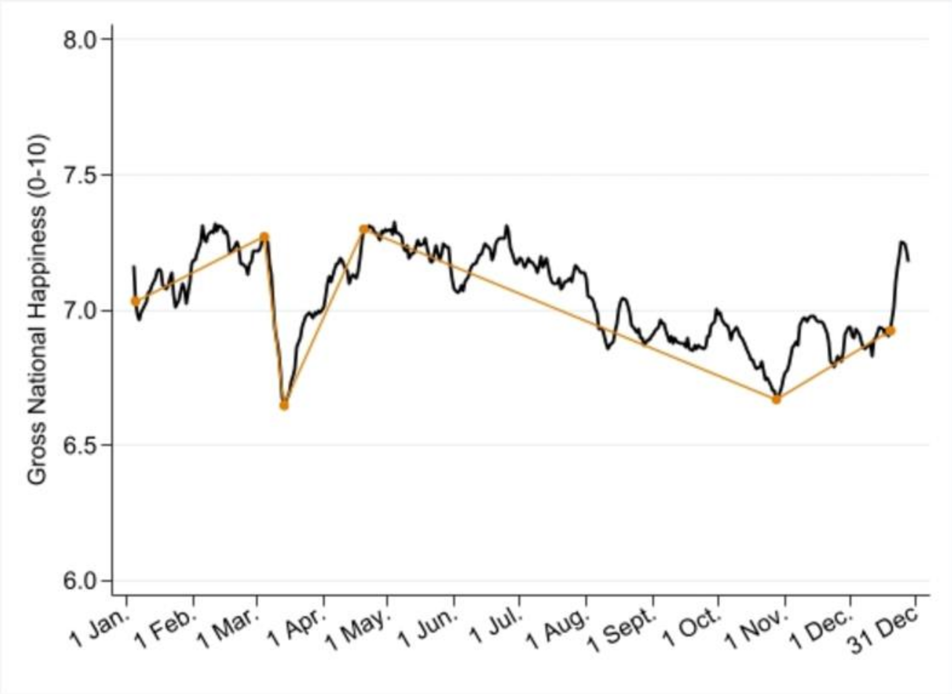
Note: data are smoothed using seven-day (centered) moving averages.

Source: All data are sourced from the project "Preferences Through Twitter" with the support of FNR, UJ and AUT.

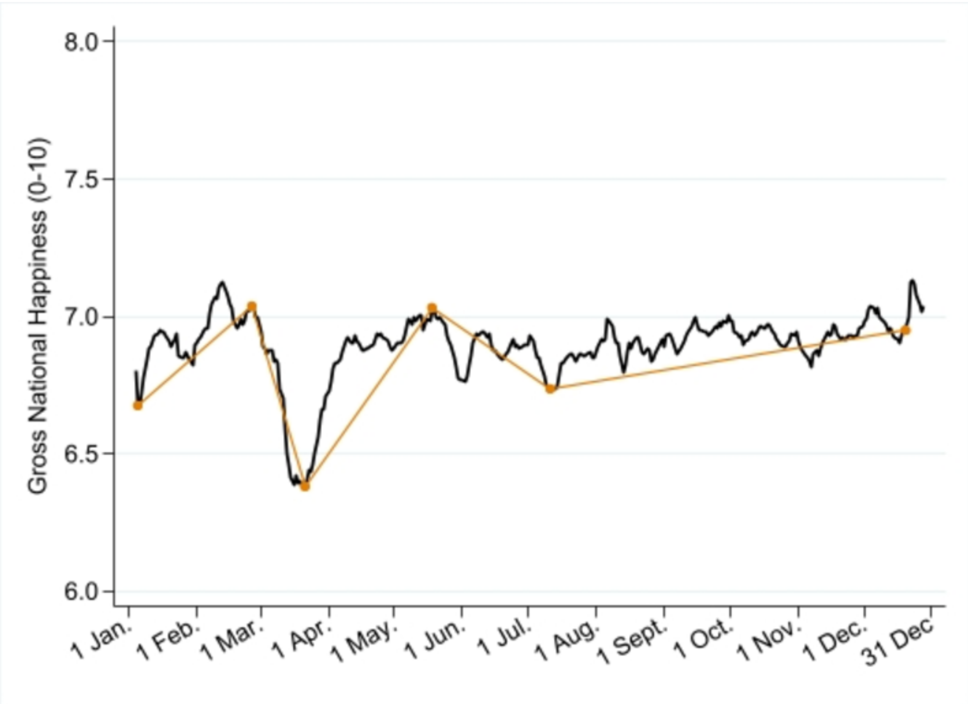
Appendix C. GNH evolution in different periods.

Table 6. Descriptive statistics by subperiods.

Date	GNH	Elapsed Time (Days)	Change	% Change	Change / days * 100
European countries					
5-Jan	7.03				
5-Mar	7.27	60	0.24	3.35	0.39
14-Mar	6.65	9	-0.62	-8.58	-6.93
20-Apr	7.3	37	0.65	9.84	1.77
28-Oct	6.67	191	-0.63	-8.62	-0.33
20-Dec	6.92	53	0.25	3.81	0.48
Australia, New Zealand, and South Africa					
5-Jan	6.68				
26-Feb	7.04	52	0.36	5.43	0.7
21-Mar	6.38	24	-0.66	-9.33	-2.74
18-May	7.03	58	0.65	10.2	1.12
11-Jul	6.74	54	-0.3	-4.2	-0.55
20-Dec	6.95	162	0.21	3.15	0.13



a) Average daily data across seven European countries.



b) Average daily data across Australia, New Zealand and South Africa.

Fig 13. Sub-periods of Gross National Happiness.

Note: GNH series use seven-day (centered) moving averages.

Source: GNH data (Greyling et al. [15]) are sourced from the project "Preferences Through Twitter" with the support of FNR, UJ and AUT.

Appendix D: List of keywords used for additional variables

Economic situation: fear in relation to jobs, economy, saving, work, wages, income, inflation, stock market, investment, unemployment, unemployed, employment rate, tech start-up, venture capital.

National institutions: trust in relation to government, parliament, ministry, minister, senator, MPs, legislator, political, politics, prime minister.

Loneliness: lonely, loneliness, alone, isolation, abandoned, social distancing, lonesome, by oneself, solitary, outcast, companionless, solitary, homesick.

Appendix E: Validity of additional variables sought from Twitter

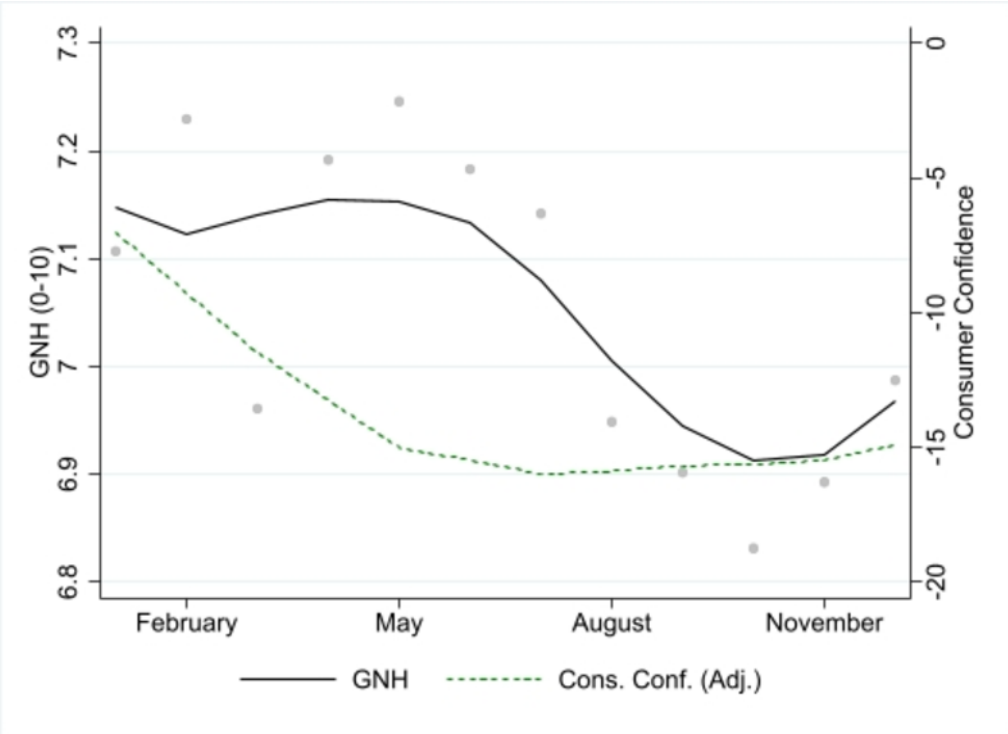


Fig 14. Economic fear correlates meaningfully with consumer confidence data.
Note: Consumer confidence is a monthly index averaging positive and negative feelings about economic conditions and perspectives. The Spearman correlation coefficient is -0.87 ($Prob > |t| = 0.003, N = 9\ months$). The reduced number of months is because of the difference for the initial month, which is missing, and because consumer confidence in April is missing for Italy.
Source: Economic fear data are sourced from the project "Preferences Through Twitter" with the support of FNR, UJ and AUT. Consumer confidence data are from Eurostat (Eurostat [74]).

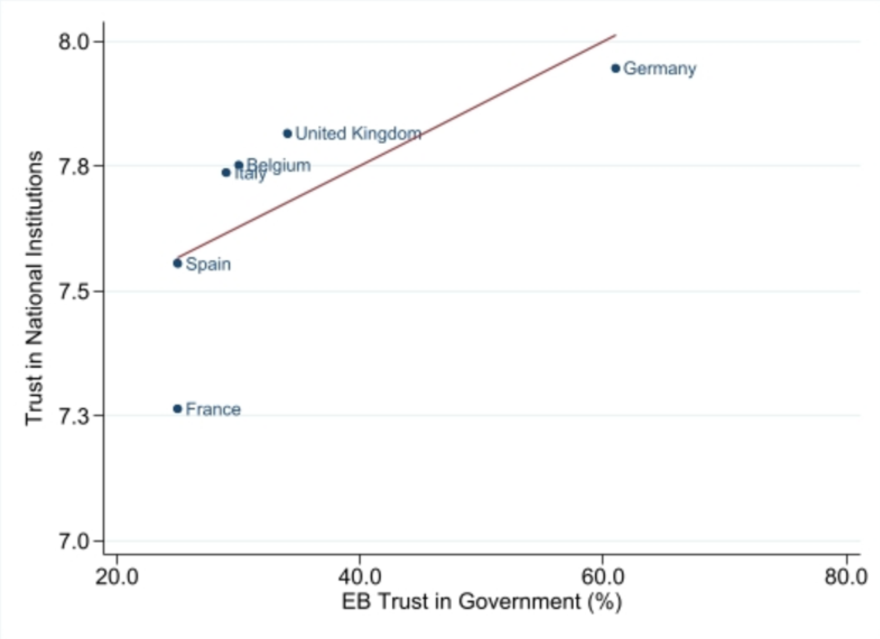


Fig 15. Trust in national institutions correlates with the share of people trusting the government.
Note: Trust in National Institutions and the share of people trusting the government are the averages by country over the period mid-July to the end of August, i.e. the same time the Eurobarometer was collected. The Spearman correlation coefficient is 0.98 ($Prob > |t| = 0.000, N = 6$).
Source: Trust in National Institutions is sourced from the project "Preferences Through Twitter" with the support of FNR, UJ and AUT. The share of people trusting the government is from the Eurobarometer (European Commission [16]), Summer 2020.

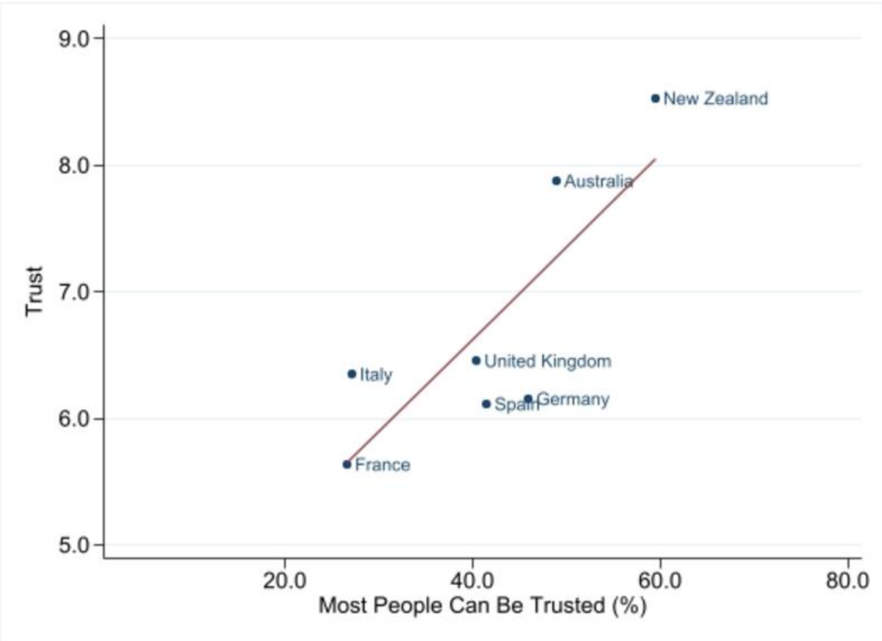


Fig 16. Association between trust and trust sourced from survey data.
Note: Trust is the average score for each country in 2020. The Spearman correlation coefficient is 0.714 ($Prob > |t| = 0.0713, N = 7$).
Source: Trust is sourced from the project "Preferences Through Twitter" with the support of FNR, UJ and AUT. Survey measures of trust are sourced from the World Values Survey (2018) – European Value Study (2018-2020) integrated data. Data for Belgium, Luxembourg and South Africa are missing.