

# Centre for Global Finance

## Working Paper Series

**No.14 / 2020**

### **Online data collection for monitoring the impact of COVID-19**

By Victor Murinde, Athina Petropoulou and Meng Xie

The Centre for Global Finance (CGF) Working Paper Series features recent studies by resident members of CGF as well as visiting researchers, altogether demonstrating the depth and breadth of research being undertaken at CGF. The papers are published to facilitate preliminary dissemination of ongoing research, enhance quality of work and contribute to the advancement of knowledge. We acknowledge, without implication, financial support from the DEGRP Research Grant (ES/N013344/2) on “Delivering Inclusive Financial Development and Growth”, funded by the ESRC and the former UK Department for International Development, which merged with the Foreign & Commonwealth Office on 2 September 2020 to become the Foreign, Commonwealth & Development Office (FCDO), the ESRC-NSFC (ES/P005241/1) Research Grant on “Developing financial systems to support sustainable growth in China – The role of innovation, diversity and financial regulation”, and the AXA Research Fund.

List of previous Working Papers of CGF:

- No.1/2018      *Capital, risk and profitability of WAEMU banks: Does cross-border banking matter?* By Désiré Kanga, Victor Murinde, and Issouf Soumaré
- No.2/2018      *Capital flows and productivity in Africa: The angel is in the details.* By François A. B. Bationo, Stephany Griffith-Jones, Victor Murinde, Issouf Soumaré and Judith Tyson
- No.3/2018      *The persistence of bank fragility in Africa: GMM dynamic panel data evidence.* By Abbi M. Kedir, Syed Faizan Iftikhar, Victor Murinde and Bernadette Kamgnia
- No.4/2018      *Reflections on central banking.* By Victor Murinde and Patrick Njoroge
- No.5/2018      *Let beholders behold: Can banks see beyond oil booms and mitigate the Dutch disease?* By Morakinyo O. Adetutu, John E. Ebireri, Victor Murinde and Kayode A. Odusanya
- No.6/2018      *National culture, CEO power and risk-taking by global banks.* By Eilnaz Kashefi Pour and Victor Murinde
- No.7/2018      *Corporate investment, financing and payout decisions under financial constraints and uncertainty: Evidence from UK firms.* By Qingwei Meng, Victor Murinde and Ping Wang
- No.8/2018      *Bank opacity and risk-taking: Evidence from analysts' forecasts* By Samuel Fosu, Collins G. Ntim, William Coffie, and Victor Murinde
- No.9/2018      *Does microcredit increase hope, aspirations and well-being? Evidence from Sierra Leone.* By Adriana Garcia, Robert Lensink, and Maarten Voors
- No.10/2018     *Lessons from emerging economies for African low income countries on managing capital flows.* By Stephany Griffith-Jones and José Antonio Ocampo

- No.11/2018 *Financial inclusion and economic growth: What do we know?* By Joshua Y. Abor, Haruna Issahaku, Mohammed Amidu, and Victor Murinde
- No.12/2018 *Climate vulnerability and the cost of debt.* By Gerhard Kling, Yuen C Lo, Victor Murinde, and Ulrich Volz
- No.13/2018 *Pan-African banks on the rise: Does cross-border banking increase firms' Access to finance in WAEMU?* By Désiré Kanga, Victor Murinde, Lemma Senbet, and Issouf Soumaré
- No.14/2018 *The peer monitoring role of the interbank market and implications for bank regulation: Evidence from Kenya.* By Victor Murinde, Ye Bai, Christopher J. Green, Isaya Maana, Samuel Tiriongo, and Kethi Ngoka-Kisinguh
- No.1/2019 *Central bank independence: What are the key issues?* By Désiré Kanga and Victor Murinde
- No.2/2019 *Banking services and inclusive development in sub-Saharan Africa.* By Haruna Issahaku, Mohammed Amidu and Aisha Mohammed Sissy
- No.3/2019 *A survey of literature on financial literacy and financial behaviour: Is there a gender gap?* By Maryam Sholevar and Laurence Harris
- No.4/2019 *Capital adjustment over the cycle: Evidence from microfinance institutions.* By Issouf Soumaré, Hubert Tchakoute Tchuigoua, and Hélyoth T.S. Hessou
- No.5/2019 *Provisioning and business cycle: Evidence from microfinance institutions.* By Hélyoth T.S. Hessou, Robert Lensink, Issouf Soumaré, and Hubert Tchakoute Tchuigoua
- No.6/2019 *Lending and business cycle: evidence from microfinance institutions.* By Hubert Tchakoute Tchuigoua, Issouf Soumaré, and Hélyoth T.S. Hessou
- No.7/2019 *Term structure of CDS spreads & risk-based capital of the protection seller: an extension of the dynamic Nelson-Siegel model with regime switching.* By Standley R. Baron and Issouf Soumaré
- No.8/2019 *Confidence, financial inclusion and sustainable economic development.* By Ayse Demir, Reinhard Bachmann, Victor Murinde, Laurence Harris, Christine Oughton and Meng Xie
- No.9/2019 *The network structure of the Malawi interbank market: implications for liquidity distribution and contagion around the banking system.* By Esmie Koriheya Kanyumbu
- No.10/2019 *Aid and Exchange Rates in sub-Saharan Africa: No More Dutch Disease?* By Oliver Morrissey, Lionel Roger and Lars Spreng

- No.11/2019 *Does credit deepening increase financial fragility?* By Peng Yiqing, Niels Hermes, and Robert Lensink
- No.12/2019 *Does microcredit increase aspirational hope? Evidence from a group lending scheme in Sierra Leone.* By Adriana Garcia, Robert Lensink, and Maarten Voors
- No.13/2019 *Do better formal institutions promote financial inclusion?* By Peng Yiqing, Niels Hermes, and Robert Lensink
- No.14/2019 *Do interbank interest rates reflect the financial soundness of borrowing banks?* By Thomas Bwire, Martin Brownbridge, Doreen K. Rubatsimbira and Grace A. Tinyinondi
- No.15/2019 *Institutional environment and the microstructure of the interbank market.* By Thomas Bwire, Martin Brownbridge, Doreen K. Rubatsimbira, and Grace A. Tinyinondi
- No.16/2019 *Segmentation of the interbank money market in Zambia.* By Jonathan M Chipili, Francis Z Mbao, Alick B Lungu, Shula M Sikaona, Anthony Bwalya, and Cosam S Chanda
- No.1/2020 *How has the rise of Pan-African banks impacted bank stability in WAEMU?* By Désiré Kanga, Victor Murinde, and Issouf Soumaré
- No.2/2020 *Threshold effects of financial inclusion on income inequality.* By Ayse Demir, Vanesa Pesqué-Cela, and Victor Murinde
- No.3/2020 *FinTech, financial inclusion and income inequality: A quantile regression approach.* By Ayse Demir, Vanesa Pesqué-Cela, Yener Altunbas, Victor Murinde
- No.4/2020 *Director reputation and earnings management: evidence from the British honours system.* By Tolulola Lawal
- No.5/2020 *Financial inclusion and the growth-inequality-poverty triangle: New evidence from Africa.* By Ayse Demir and Victor Murinde
- No.6/2020 *Fellowship amongst peers: A systematic and selective survey of literature on the analysis of interbank lending networks.* By Anosi F. Ikimalo and Victor Murinde
- No.7/2020 *Exploring the impact of COVID-19 pandemic on Africa's FinTech space.* By Joshua Yindenaba Abor
- No.8/2020 *Financial market integration in sub-saharan Africa: How important is contagion?* By Robert Akunga, Ahmad Hassan Ahmad and Simeon Coleman
- No.9/2020 *Finance and well-being in developing countries: Does access to mobile money improve household well-being?* By Fei Jiang, Christopher J. Green, Ahmad Hassan Ahmad and Carlos Sakyi-Nyarko

- No.10/2020    Mobile money, ICT, financial inclusion and inclusive growth: How different is Africa? By By Fei Jiang, Christopher J. Green, Ahmad Hassan Ahmad and Victor Murinde
- No.11/2020    Financial inclusion and welfare improvement: Empirical evidence from A households survey data. By Carlos Sakyi-Nyarko, Ahmad Hassan Ahmad and Christopher J. Green
- No.12/2020    The FinTech revolution: What are the opportunities and risks? By Victor Murinde and Efthymios Rizopoulos
- No.13/2020    The COVID-19 pandemic and its impact on African economies and financial markets: A review. By Elikplimi Komla Agbloyor and Joshua Yindenaba Abor

Any reproduction, publication and reprint in the form of a different publication, whether printed or produced electronically, in whole or in part, is permitted only with the explicit written authorisation of the authors of this paper. The views expressed in the paper are those of the authors and do not necessarily reflect those of the CGF.

All CGF Working Papers can be downloaded from CGF Website.

Centre for Global Finance  
SOAS University of London  
10 Thornhaugh Street, Russell Square  
London  
WC1H 0XG

Email: [cgf@soas.ac.uk](mailto:cgf@soas.ac.uk)

Website: <https://www.soas.ac.uk/centreforglobalfinance/publications/>

# Online Data Collection for Monitoring the Impact of COVID-19

*Victor Murinde<sup>a</sup>, Athina Petropoulou<sup>a</sup> and Meng Xie<sup>a,\*</sup>*

*<sup>a</sup>Centre for Global Finance, School of Finance and Management, SOAS University of London, UK*

## Abstract

The dreaded coronavirus pandemic, which struck the world economy in February this year and now has woken up to a winter wave in some countries, makes the need for relevant socio-economic data more imperative than ever – be it micro, macro, primary or secondary data. This paper attempts to provide a comprehensive overview of the ongoing online data collection, towards monitoring and evaluating the impact of COVID-19 across a broad spectrum of economic activities. We review and classify several data survey designs according to their focus, the type of information collected and possible methods of data analysis. Given that the pandemic struck the world economy with a synchronised shock, our focus is global but revolves around households, companies, banks and governments, in a flow-of-funds context. We also review and reflect on conventional methods as well as digital approaches to data collection and identify their strengths and weaknesses. Finally, we propose an innovative dynamic approach that combines multiple methods for collecting data, including Application Programming Interface (API) and artificial intelligence, to generate ‘big data’ to support research.

**Keywords:** online data collection; COVID-19; flow-of-funds; API; artificial intelligence

---

\* We acknowledge research funding for the SOAS Centre for Global Finance, under: the AXA Research Fund; the ESRC and the former UK Department for International Development, which merged with the Foreign & Commonwealth Office on 2 September 2020 to become the Foreign, Commonwealth & Development Office (FCDO), Research Grant No. ES/N013344/2 on “Delivering Inclusive Financial Development and Growth”; and the ESRC-NSFC (ES/P005241/1) Research Grant on ‘Developing financial systems to support sustainable growth in China – The role of innovation, diversity and financial regulation’. Useful comments and suggestions were received from Efthymios Rizopoulos, Richard Alexander, Ayse Demir, Mustabsar Awais, Orly Levy, Esmie Kanyumbu and other participants at the SOAS Centre for Global Finance Research Seminar on 14<sup>th</sup> October 2020, where an earlier version of this paper was presented. We are responsible for all surviving errors.

*E-mail addresses:* [vm10@soas.ac.uk](mailto:vm10@soas.ac.uk) (Victor Murinde), [ap102@soas.ac.uk](mailto:ap102@soas.ac.uk) (Athina Petropoulou), [xm1@soas.ac.uk](mailto:xm1@soas.ac.uk) (Meng Xie).

# 1. Introduction

The COVID-19 pandemic has created a systemic shock to the global economy affecting households, companies, governments, banks and financial markets. At the time of writing over 27 million cases of coronavirus have been recorded globally since the start of the pandemic in more than 200 countries. Governments worldwide have allocated more than \$13trillion to tackle the economic recession. Even though the long-term impact of the coronavirus is still unknown, researchers and organizations around the world are undertaking research activities to explore the early implications of COVID-19. This task is challenging given the extreme speed at which the crisis unfolds and the uncertainty around the epidemiology of the virus and the effectiveness of containment measures; AXA (2020) indicates COVID-19 as the number one global risk.<sup>1</sup> Circumstances like these increase the uncertainty about what the economic landscape will look like when we emerge from this crisis and intensify the need for better understanding of the ways through which the coronavirus is re-shaping the norms in our society. For this reason, there is a plethora of online surveys on the impacts of COVID-19 across different areas. For example, the National Institute for Digital Learning (NIDL) at Dublin City University (DCU) launched an online survey to investigate the affective experience of educators as they changed to online teaching in response to COVID-19. The World Health Organization surveys health professionals globally about the medical effectiveness of the measures taken to contain the spread of the COVID-19 pandemic. The Southern Health NHS Foundation Trust is undertaking a global study that examines the psychological impact of the pandemic and the lockdown upon individuals. Imperial College London, King's College London and the University of Cambridge are looking into the wellbeing and cognitive facilities of individuals before, during and after the lockdown in the UK, while trying to understand the factors that may mitigate the COVID-19 impact. The University of Granada is running an online survey on the impacts of the quarantine on the dietary habits of people in Greece, Italy, Portugal, France, Germany, and Spain. UNESCO conducted a world-wide survey on the impact of COVID-19 on museum visiting patterns. Hence, the dreaded coronavirus pandemic, which struck the world economy in February this year and now has woken up to a winter wave in some countries, makes the need for relevant socio-economic data more imperative than ever – be it micro, macro, primary or secondary data.

The objective of this paper is twofold. First, we offer a comprehensive overview of online data collection that has been undertaken or is being undertaken now with regards to COVID-19 impact on households, firms, banks and governments – the major components of economic activity in a flow of funds framework. We start with an overview of the online surveys that have been done, review the motivation behind each of them, the data collection methods used and the type of collected data. For

---

<sup>1</sup> The AXA Future Risks Report 2020 finds that the main risks in the next five to ten years will relate to pandemics and infectious diseases, climate change and cyber risks; see AXA (2020).

the surveys that are reporting their findings, we evaluate their key findings as a way of preview of the survey's results. Second, we propose ways through which data collection on COVID-19 can go forward. Under these highly uncertain circumstances, the efficacy of data has become increasingly important. However, traditional data sources fall short to provide stakeholders with timely or adequate insights. For example, field work is limited with countries going in lockdown. Generating data online is the only way researchers and practitioners can obtain timely and granular perceptions. After reflecting on traditional and digital approaches to data collection, we propose a novel and dynamic approach which is a viable strategy for data collection on COVID-19. Due to the unprecedented economic impact of the COVID-19, a substantial part of the online data collection on COVID-19 that we are proposing is focusing on the economic impact of the pandemic and the lockdown measures across three main sectors, i.e. households, firms and banks.

In this paper we discuss both primary and secondary data, as both of them have their advantages and disadvantages. It depends on the research question(s) and research design to choose which type of data serves its objectives better. Many online surveys that we discuss in this paper collect primary data on the impact of COVID-19 by directly asking those that are affected. For example, the "COVID-19 Global Survey: Inside business-impacts and responses" conducted by the ACCA collects data by directly asking organisations about the challenges that the pandemic has generated for them. When collecting primary data, researchers have better control over the information collected as they can tailor the data collection tools (i.e. questionnaires, interviews etc) to meet the exact purpose of the research in hand. On the other hand, secondary data may be more general and not tailored specifically to meet the exact needs of the research. However, secondary data have many advantages. They are easily accessible because they are often shared publicly. Also, the abundance of available secondary data suggests that researchers can have access to numerous datasets and do plethora of readings from pre-made datasets with minimum or at some cases zero cost. Pre-existing data can be used in many different ways and generate new insights into existing concepts. For example, Demirguc-Kunt et al. (2020) use pre-made, available datasets on electricity consumption and NO<sub>2</sub> emissions in order to measure the economic activity during the pandemic.

The rest of the paper is organized as follows. Section 2 provides a brief review of the methodologies used for data collection. Section 3 explicitly discusses the role of big data and artificial intelligence in research. Section 4 describes the online data collection on companies. Section 5 describes the online COVID-19 data collection initiatives for households. Section 6 is dedicated to surveys on the government sector. Section 7 compares conventional, digital, and dynamic approaches to data collection. Finally, section 8 concludes and proposes ways of moving forward with online data collection for monitoring and analysing the economic impact of the coronavirus.

## 2. Data Collection Survey Methods

### 2.1 A survey as a research design

In scientific research, a survey is a research design that entails the collection of quantitative and qualitative data from a representative sample(s) of the population. A survey comprises a list of questions, which are specifically crafted to obtain information needed for the purpose of a research. Therefore, the objective of a research project is a prime factor that determines how to devise survey questions. In this paper, we discuss surveys by categorising them into three main types, namely factual, inferential and exploratory (Easterby-Smith, Thorpe and Jackson, 2015).

Factual surveys are usually used to collect descriptive data, mainly seeking factual information from people, such as demographic characteristics, household consumption or attitude towards certain subjects. This type of survey is often employed in opinion polls, market research and studies of household or consumer behaviour. For example, national statistical institutes and international organisations have a long tradition of using factual surveys to collect information on key variables that reflect diverse aspects of human societies. For example, The United Nation Statistics Division has used surveys since 1948 to collect factual data on a wide range of topics for [The Demographic Yearbook](#). The data ranges from population size and composition, births, deaths, marriage and divorce, as well as respective rates, on an annual basis. The Demographic Yearbook census datasets also cover additional topics including economic activity, educational attainment, household characteristics, housing characteristics, ethnicity, language, foreign-born and foreign population.

Polling institutes regularly employ this type of survey to collect and monitor people's voting preferences via Voting Intention polls in the run-up to regional or national elections (e.g. YouGov, IFOP, ICM, Gallup Poll etc.). International consultancy firms are also in favour of factual surveys to reveal trends and attitudes with connection to business activities. For example, Mckinsey & Company has conducted various [Mckinsey Global Surveys](#) with global executives to reveal the pressing business, economic, and management issues. It has also developed global surveys of consumer sentiment during the coronavirus crisis, which is dedicated to track consumer sentiment to assess how people's expectations, incomes, spending, and behaviours change throughout the crisis across multiple countries over time.

Therefore, the fact-collecting surveys are commonly used for developing an overview or descriptive interpretation of research subjects. In this respect, the design of questions and subsequent analysis have less interest in examining internal relationships between variables.

Inferential surveys are popular amongst explanatory research that aims to examine associations or causal relationships between variables. They are often used in psychology, marketing, and organisational behaviour studies for data collection. For examples, Spitzmuller et al (2006) conducted surveys to investigate whether organisational members' evaluations of their organisational experiences

affect survey response to organisationally sponsored surveys. Delmas and Pekovic (2013) examined the influence of environmental standards on the productivity of employees by using survey data from 5220 firms. Koschate-Fischer et al. (2014) conducted a survey with 17,324 consumers to collect their perceived degree of commoditisation and the product category involvement of given product categories. This data was used together with relevant factual data, such as purchasing amount, retailer information, product category and household information, to identify factors that moderate the relationship between private label share and store loyalty.

One distinctive character shared across these studies is that researchers usually hold prior assumptions or established hypotheses about the relationship between selected variables, and they are interested in obtaining quantitative data. The common starting point of their survey is defining independent variables (i.e. factors can be controlled or manipulated), and dependent variables (i.e. factors are assumed to be affected by the change/difference of independent variables). It is then followed by a justification of how to measure each variable. Therefore, the use of inferential surveys for data collection requires researchers to decide the nature of variables (i.e. dependent or independent) and develop robust measurements. Furthermore, the design of questions must serve the purpose and nature of measurements. For instance, researcher's choice on whether to use binary questions (e.g. yes or no), categorical questions (e.g. A, B, or C) or scaling questions (e.g. how much do you agree with this statement? 1=disagree, 10=agree) derives from their judgement on what type of data is needed for answering research questions and how that data can be effectively captured (i.e. measurements).

In contrast to explanatory research, many studies are carried out without any intention of testing relationship between factors. Instead, they use surveys to explore new characters and patterns in relation to research subjects. Hofstede's (1980/1984) survey of national cultural differences is among one of them. The objective of his research was to create a universal set of principles against which any culture can be measured. The survey constitutes a list of structured questions with regards to respondents' views and values on personal and professional life. The responses were collected in a quantitative manner where respondent's opinion and attitude were recorded by their rating on given statements. However, in other exploratory studies, researchers rely on surveys to collect qualitative data. For example, Eckhardt, Belk and Devinney's (2010) survey on why consumers don't consume ethnically employed face-to-face depth interviews. Participants spoke about their beliefs, views and behaviours in three different consumer ethics situations through open-ended questions and statements.

The time scale is also an essential element to be considered in conducting surveys. Some studies adopt cross-sectional design that collects data at a single point or a short period of time, whereas others require longitudinal surveys to gather continuous information and allow the comparison of changes in variables over time.

As discussed above, the variation on research objectives contribute to diverse usefulness of surveys, and subsequently to different ways of constructing questions and interacting with participants. The next section discusses major instruments used in surveys for data collection together with communication channels employed to reach participants and collect responses.

## **2.2 How to collect primary data using survey instruments?**

### *2.2.1 Questionnaire*

Questionnaire comprises a series of questions that expect participants to answer in a survey. In most cases, the given questions come with a choice of answers for respondents to select from, but some questionnaires tend to be open-ended for participants to freely express thoughts and attitudes. Questionnaire is an efficient tool to collect large-scale data in a standardised and uniform manner.

### *2.2.2 Self-completion questionnaire*

Self-completion questionnaire, also referred to as self-administration or self-reporting questionnaire, is one of the main instruments for collecting data in surveys. Participants need to answer questions themselves without assistance from the survey designers or administrator. This type of questionnaire can be executed in diverse forms.

The very traditional but still predominant form is paper-and-pencil, usually distributed through post or individuals who are recruited to hand out questionnaires. With the mail questionnaire, questions and instructions of how to answer are printed out on papers and sent to the participant through post. Following the completion of the questionnaire, the respondent is usually asked to return it by post, or to deposit it in a certain location. For example, Statistics Canada used self-reporting questionnaire via post to collect data for the Business Conditions Survey in 1985. Self-completion surveys could also be administrated with help of individuals who hand out printed questionnaires to respondents and then collect them back. However, since the innovation of computers and internet, an increasing number of surveys have started distributing questionnaires and collecting answers by using digital platforms. Researchers can design questionnaires by using online survey tools (e.g. Qualtrics, SurveyGizmo, SogoSurvey, SurveyMonkey and Google Forms) and invite participants by sending them the link to the questionnaire through emails. For example, the Online Services Questionnaire (OSQ) and Member States Questionnaire (MSQ) of the [2020 UN E-Government Survey](#) were conducted over internet.

This digital platform has brought many benefits to the design and administration of questionnaires. First, it reduces the cost and time on printing, posting and hiring researchers for handing out paper-based questionnaires. Second, it allows a more controlled navigation of questions. For example, when a study needs the response of each question independent of the others, the online questionnaire can be designed to only present one question at a time and thus avoid the situation where respondents might see all questions in paper questionnaires. It can also allow an accurate diversion of respondents during

the completion of questionnaire if researchers need to offer different questions to various groups based upon defined characteristics (e.g. age, gender, occupation). Third, the online survey tools enable a high degree of flexibility for respondents to complete questionnaires at convenient time and location, and without the concern of losing the paper or how to return the form. Last, online survey tools can collect and store data simultaneously while the respondent enters answers. This function has significantly improved the efficiency of organising data after receiving completed questionnaire.

Whilst the digital channels have contributed to many positive changes to the practice of questionnaires, the underpinning core mechanism remains the same – requesting participants to provide answers to given questions. This form of obtaining information faces many challenges, such as nonresponse, low rate of response, ineffective questions and incomplete answers. For example, most people find answering questionnaires boring, time consuming and unrewarding, thus decide to ignore the invitation of participation or return incomplete questionnaires. Because of this responding fatigue, the design of questionnaires often tries to avoid long questions, which in return affects the quantity and depth of required data. In most cases, it is difficult to have control over the involvement of non-respondents (e.g. friends, families, colleagues) in answering questions. When a questionnaire is sent to senior management in a firm, it is very likely to be answered by someone else in their office. It is also common that respondents do not fully understand questions but have no one to ask for clarifications. If a questionnaire contains open-ended questions without appropriate prompts, respondents would either skip the question or provide short answers. The use of self-completion questionnaires is also not suitable in surveys of respondents whose literacy is limited.

Since self-reporting questionnaires offer limited or, sometimes, no flexibility of explaining, elaborating or providing additional information, the way of asking questions is vital for determining the quality of collected data. Open questions are very useful to collect qualitative data and explore new areas, such as people's thoughts with relation to a pandemic such as COVID-19, however, they often face low response in questionnaires as mentioned above. Closed questions are questions/statements with fixed answers/scores to choose from. They demand less effort for respondents to provide answers and allow the measuring of variables. However, because the given choices are determined by researchers based on prior knowledge, there is great risk of excluding responses that are not covered by the fixed answers. Diverse forms of closed questions have been developed to help collect data more effectively. For obtaining factual data, the binary and categorical questions are often used to ask for gender, age group, occupation, and respondents' simple approval/disapproval on given statements. The Likert scale is a predominate tool for capturing the degree of opinions and attitude. Examples of various closed questions forms are presented in Appendix 1.

### *2.2.3 Questionnaire by interviews*

Interviews are directed conversations evolving around questions and answers about a topic (Lofland and Lofland, 1984). Interviewing can be categorised into different formats based on the extent to which researchers have control over asking and recording questions. This include structured interviews, semi-structured interviews and unstructured interviews. It can also be differentiated by how researchers interact with participants, such as face-to-face interviewing, remote interviewing and one-to-one/group interviewing.

Since the main disadvantages of self-reporting questionnaire are associated with the lack of interactions between researchers and respondents, the form of structured interviews is often adopted by survey researchers to overcome this limitation. However, in order to maintain the standardisation of collecting data, interviewers need to follow instructions to read out the given questions, statements and prompts as they are presented in a questionnaire to every participant. In addition, they are supposed to read out questions in the same order to every respondent. However, both sides can elaborate and clarify ambiguous questions or statements when necessary.

Most studies started structured interviews in person using a paper-and-pencil form. Since the broad use of telephones, more studies choose to conduct structured interviews remotely by using Telephone Interview. With the development of mobile devices and computers, Computer-assisted Interview (CAI) has become a popular tool in research, where the respondent or interviewer uses an electronic device, such as a tablet, a mobile or a computer to answer the questions in face-to-face interviews. In recent years, internet-based virtual communication tools have become a widely adopted alternative for conducting in-person interviews when interviewers and respondents are not able to meet in the same physical space.

Structure interviews are particularly favoured by large-scale research that usually has a national or global coverage. For example, The World Bank's Enterprise Surveys (ES) collect questionnaire data from enterprises in manufacturing and key service sectors in every region of the world via face-to-face interviews. Statistics Denmark has used structured interviews for conducting the Consumer Expectation Surveys since 1974. The questionnaire was first conducted in face-to-face structured interviews, and then moved to computer-assisted telephone interviewing (CATI). The questionnaire data for the Labour Force Survey in the Integrated Household Survey by the Office for National Statistics (ONS) are collected through a combination of face-to-face interviews using computer-assisted questionnaires and telephone interviews.

### *2.2.4 Semi- and unstructured interviews*

For studies that require in-depth understanding of people's behaviour or thoughts, semi-structured and unstructured interviews are useful to generate information that is often hidden or new to our current

knowledge. For instance, unstructured interviewing is one of the techniques frequently used in consumer behaviour research (Levy, 1981) for drawing out consumers' deep thoughts. Semi-structured interviews refer to a context where interviewers have a series of questions scheduled to ask but are able to change the order of questions, amend questions or add additional questions. For example, Eckhardt, Belk and Devinney (2010) used semi-structured interviews to find out why consumers do not consume ethically. In unstructured interviews, interviewers may only prepare a list of topics and expect to develop questions and comments through highly interactive conversations. Researchers establish a "grand tour question" (Spradley, 1979) and participants are asked to tell their story as they see it, feel it, experience it (Corbin and Morse, 2003). This type of interview is popular amongst sociology and anthropology studies. For example, Professor Anne Oakley is the pioneer in using unstructured interviews; she first used them in the 1970s to achieve an in-depth understanding of women's experiences, thoughts and feelings regarding their lives and roles in family and society.

An example of face to face interview is the Family Resources Survey. The Family Resources Survey is one of the largest UK household surveys and is carried out by the Office for National Statistics (ONS) with the National Centre for Social Research (NatCen) and the Northern Ireland Statistics and Research Agency (NISRA) and is sponsored by the Department for Work and Pensions (DWP). It is a continuous annual survey that collects information on the financial circumstances of individuals living in private households in the UK. Topics covered by the survey include income and state support, tenure, disability, carers, pensions, savings, investments and self-employment. Research is conducted through face-to-face interviews where all adult residents in the household over 16 years old are interviewed. The mean time of the interview is approximately one hour and households who take part receive a £10 Post Office voucher. Data from the survey is released annually.

#### *2.2.5 Group interview*

Interviews can be conducted on a one-to-one basis where a researcher speaks with only one respondent at a time. They can also be held between one researcher and a group of participants. A focus group is generally involves engaging a small number of people in an informal group discussion, focused around a particular topic or set of issues (Wilkinson, 2004). In comparison to an individual interview, focus groups are considered to be more effective in terms of generating abundant and diverse information. This is because the interaction between the group members encourages each person to clarify their ideas against others, to compare their views with others, to contest against divergent opinions, and to help each other recall and elaborate shared thoughts or experiences (Fontana and Frey, 2005; Kitzinger, 1994; Morgan, 1996). In addition, the coverage of multiple participants within the same time and space is useful to reduce the cost and improve the efficiency of data collection (Krueger and Casey, 2000).

One of the most challenging choices researchers have to make is on deciding which data collection method is optimal for the research in hand. The optimal data collection method is the best method,

given the research question and certain restrictions (Biemer and Lyberg 2003). Different populations, research questions and circumstances require different modes of survey to reduce sources of survey error and to minimize the costs. Stern et al. (2014) discuss the challenges, dilemmas and opportunities for survey methodologies in social sciences. De Leeuw (2008) offers a detailed overview of computer assisted data collection methods namely the computer assisted telephone interviewing (CATI), the computer assisted personal interviewing (CAPI) and the computer assisted self-interviewing (CASI, CSAQ and the Web). The use of telesurveys for household surveys is also examined in detail by Nathan (2001), whereas Buskirk and Andrus (2012) provide a detailed overview of three approaches for conducting online surveys via mobile devices. Researchers often opt for mixed-mode approaches as this compensates for the weaknesses of each individual method. De Leeuw (2005) offers an outline of the advantages and disadvantages of mixed-mode survey designs.

### **3. Big Data and Artificial Intelligence**

Recent advances in data revolution have been shaping economic research along several dimensions. Traditional statistics and household surveys have been successful in tracking medium to long-term trends but ineffective in delivering real time information which is essential in order to take timely action. Thanks to the digital revolution, real-time digital data are generated in increasing volumes. The availability of big data is changing the way researchers approach empirical research and the statistical tools that they apply. The Billion Prices Project uses prices collected from many online retailers worldwide on a daily basis to conduct research in macro and international economics. Researchers construct daily price indexes that complement official statistics such as the Consumer Price Index (CPI). This is particularly interesting for countries such as Argentina where government statistics are not regularly published or are criticized for being misreported (Cavallo 2013). Social media have been used to collect high frequency and in real time information on the aggregate economic activity. The University of Michigan Social Media Job Loss Index for example is derived from signals counting job-related phrases in Twitter and is updated weekly. The “Hedonometer”, developed by the Computational Story Lab of the University of Vermont, is another example of a project where researchers use real-time, big data generated by the online lives of individuals in order to calculate the state of their mental health. The Hedonometer measures word choices across Twitter posts to calculate a moving measure of well-being<sup>2</sup>. Yu et al. (2019) develop an online big data- driven oil consumption forecasting model which uses Google trends in predicting oil consumption trends and values. Global Pulse is an initiative by the United Nations that uses big data and artificial intelligence to address development challenges. Through the Global Pulse, the UN are discovering ways to harness the power of big data to promote sustainable development and human well-being. Companies such as I know First use artificial intelligence- based algorithm to generate daily market predictions for over 10,000 financial assets.

---

<sup>2</sup> [http://hedonometer.org/timeseries/en\\_all/?from=2020-01-01&to=2020-10-14](http://hedonometer.org/timeseries/en_all/?from=2020-01-01&to=2020-10-14)

Eagle Alpha is connecting alternative data that are used by firms to make data-driven investment decisions. RavenPack is a data analytics provider in the financial services that helps clients extract value and insight from unstructured data. The efficacy of data has become increasingly important since the start of the COVID-19 pandemic. Taiwan is an example of a country that was able to stop some of the COVID-19 transmission early by using big data analytics. The country integrated its national health insurance database with its immigration and customs database and used the travel ticket scan to obtain data on the origin of the flight and the route and then use these data to classify travellers' infectious risks. It also generated real-time alerts during a clinical visit based on travel history data to help aid case identification.

Big data abundance accelerates advances in computing. Artificial intelligence techniques such as Machine Learning Algorithms are general purpose technologies that provide innovative solutions to traditional challenges in theory and application. Modern advances in big data have made it possible for machines to replicate the way human brain works. Neural Networks for example take in numerous stimuli, pass it through layers and learn to associate the input with the output. Big data is essential in order to train the NN to learn the best relationship between inputs and outputs because the learning improves as more and more examples are considered and the machine is able to perform the task repeatedly. ML models have been extremely popular because of their ability to handle heterogenous and large number of inputs without the worrying about issues such as co-linearity. The evolution of big data makes it possible to construct long-run indices and create aggregate statistics and new tracking measures of the economy. Chetty et al. (2014) for example study the long-run impacts of teachers' quality and students' long-term outcomes by linking information from student records to administrative tax data. This allows the researchers to track approximately one million individuals from elementary school to early adulthood. Artificial Intelligence tools have been widely commoditized in the field of finance where market data is abundant. Hutchinson et al (1994) first attempted to approximate the market's option pricing function using a NN. Since then several developments have been made towards the same direction. Recent studies that apply ML to option pricing include Culkin and Das (2017), Fang and George (2017) and Gan et al (2020) among others. Also, researchers have used ML architectures for financial forecasting with some promising results, see for example Rundo et al. (2019), Siami-Namini et al. (2018), Sun et al. (2018) and Selvin et al. (2017). The main argument for applying ML techniques to financial time series analysis revolves around the fact that ML captures non-linearities in the data, whereas traditional approaches assume a linear relationship or a non-linear model that is simplified to a linear one (De Prado, 2016). ML models do not need any prior specification of the structure of the model and thus can reveal non-linear and high dimensional complex relationships between many input variables and the outputs. Furthermore, the availability of big data helps to develop collaborations between academics and data-rich companies and provides researchers with the opportunity to look inside the "black box" and generate new opportunities for experiments.

The big data revolution does not come without challenges. First, it raises privacy and confidentiality concerns which enhance the need for research transparency. Researchers that work directly with private companies often agree to keep the data confidential. This however limits opportunities for other researchers to replicate or extend their studies. Second, large datasets with complex structure can be difficult to organize and study which makes the need for a simplifying conceptual framework even more profound. Advanced data mining approaches from the fields of statistics and computer science are available to the researchers however the trouble of seeking a meaningful conclusion from this vast amount of data remains a challenge. Third, figuring out ways to organize the data is a crucial challenge in empirical research. Programming capabilities and statistical methods are adapting to make the most out of these granular datasets. Forth, complex ML methods require a large number of observations to be properly tested. This can be a barrier for applications which only have a limited number of observations e.g. macroeconomics (Ghoddusi et al. 2019).

## **4. Online Data Collection on Companies**

### **4.1 Online data collection by ONS**

The Office for National Statistics (ONS), in response to the coronavirus pandemic, undertook the “Business Impact of Coronavirus Survey (BICS)”. The aim of this survey is to capture business’ views on the impact of the coronavirus on turnover, workforce, prices, trade and business resilience. Participants of the survey are businesses of all sizes from various industrial sectors and regions in the UK registered for Value Added Tax (VAT) and/or Pay As You Earn (PAYE). Businesses that participated in the survey also received either a letter or an email explaining how to complete the survey online.

The study has been done in seven fortnight waves, starting on the 9<sup>th</sup> of March 2020 and covering up until the 14<sup>th</sup> of June 2020. Table 1 presents the reference and collection period for each wave of the survey. The questionnaire of the survey was regularly updated so that the survey complies with changing policy and user needs. However, these changes limit the comparability of the survey results between waves. Up to Wave 5, the number of businesses who responded to every wave is 3,521 which allows the researchers to track this panel over the whole survey period and examine how the impact of COVID-19 has evolved for them. The total amount of unique businesses that responded is 9,036.

[Table 1 around here]

Table 2 presents the total sample and response rates for each wave of the survey. There has been a sample boost of 2,000 businesses on Wave 5, primarily located in Scotland and Wales. Responses to

the survey are qualitative. Data are available to download in the form of descriptive statistics.<sup>3</sup> Information from this survey is being used by decision-makers to identify the strategies needed to respond to the impact of coronavirus on the UK economy and society.

[Table 2 around here]

A snapshot of the data is presented in table 3. This tables presents the changes in turnover for Wave 5 broken down by industry and answers the question “In what way was your enterprise’s turnover different in the last two weeks?”

[Table 3 around here]

The voluntary nature of the survey, the difference in response rates and the dependency on participants that only responded in particular waves suggest that the survey results should be treated with caution and comparisons of the proportions of businesses trading status is not recommended. For example, 1,366 businesses that continue to trade in Wave 5 did not respond in Wave 4. This suggests that researchers’ ability for results comparisons between Waves is only limited, thus undermining the panel nature of the survey.

#### **4.2 Online data collection by ACCA**

The “COVID-19 Global Survey: Inside business-impacts and responses” is conducted by the Association of Chartered Certified Accountants (ACCA). This survey aims to identify a wide range of business and financial challenges that organizations of all sizes and sectors face. The survey was issued on the 13<sup>th</sup> of March 2020 and ran until the 26<sup>th</sup> of March 2020. The number of participants is over 10,000, including members of the ACCA and other stakeholders from over 100 countries. The country with the highest number of responses is the UK with 2,774 responses and Malaysia ranks second with 919 responses. Table 4 shows the top 20 countries in number of responses.

[Table 4 around here]

Figure 1 presents the survey demographics in terms of sector and size of the participating organizations. The majority of the respondents is in the corporate or financial services sector (56%) and has small size- fewer than 50 people (31%).

[Figure 1 around here]

The multi- country element of this survey is valuable for researchers as it allows for regional comparisons of the impact of COVID-19 (see for example figure 2). However, responses from the UK and Malaysia account for more than 36% of the sample so generalizing the findings if this study as

---

<sup>3</sup> Data available

<https://www.ons.gov.uk/economy/economicoutputandproductivity/output/datasets/businessimpactofcovid19surveybicsresults>

worldwide is questionable. Moreover, the survey ran only until the 26<sup>th</sup> of March, whereas the COVID-19 pandemic continues to unfold until much later. A second wave for this survey could be valuable in capturing the post-COVID-19 impact on the same cohort of participants.

[Figure 2 around here]

### **4.3 Online data collection by IGC and IPA**

The “COVID-19 Economic Impact Survey” is a joint initiative designed by the International Growth Centre (IGC) and Innovations for Poverty Action (IPA). The aim of this survey is to collect data on the economic impacts of COVID-19 in developing countries. The survey is designed in such a way that it can be used to measure the economic consequences of the lockdowns on a variety of economic agents, such as large companies, informal and small businesses, self-employed, workers and farmers. This study will inform governments across developing countries on the economic implications of particular responses. Participants are asked about their economic activities prior to being affected by COVID-19, how the outbreak affected them and how the government could support them.

The survey is conducted via Computer-assisted telephone interviewing (CATI), a telephone interviewing method of data collection in which a computer displays the question on a screen, the interviewer reads it to them over the phone and then enters the respondent’s answer to the computer.<sup>4</sup> The survey can also be taken online via the Survey CTO platform and is available in English, Portuguese, Spanish, French, Bengali, Urdu and Burmese.<sup>5</sup> The IPA/ICG survey can be used as a stand-alone survey in nationally- representative samples and as an online survey to be distributed across countries. It can also be integrated as an add-on to existing research projects. There are 37 groups of researchers that have incorporated or are planning to incorporate this questionnaire into their research projects covering data from 17 developing countries. Data from this survey will be publicly available in the upcoming months.

Additionally, IPA and IGC are teaming up with the World Bank. Through their website they will be sharing data and results from the World Bank’s Business Pulse Survey. This survey has been harmonized with the IPA/ICG survey questionnaire so that results can be more comparable across projects and countries. The World Bank’s Business Pulse Survey is a phone/online survey created to capture the effects of the pandemic on companies through each of the four channels of impact i.e.

---

<sup>4</sup> Examples of surveys that have used CATI; Lafuma et al. (2009), Baltas & Argouslidis (2007) and Derr et al. (1992).

<sup>5</sup> The questionnaire of this survey is available <https://docs.google.com/spreadsheets/d/14D154dpD37pgEDdir9uyHqHNHwtrMg96ZQTZQGcWC5Q/edit#gid=1166996848>

supply, demand, financial and uncertainty as well as the ways through which the firms mitigated these effects and their views on public support programs.

Generally, CATI is a popular mobile phone survey mode for low- and middle- income country research, where computer/internet penetration may be lower compared to developed economies. Traditionally, CATI uses a centralized call centre, which however was not manned during the pandemic. As of April, call centres around the world transitioned to virtual call centres where interviewers call from home. This brings new challenges such as ensuring a quiet space to make phone calls or a reliable internet connectivity, among others. For this reason, in March 2020 J-PAL published guidance about how to transition to work from home (WFH) based CATI as an attempt to present suggestions and best practices to help address these new challenges (Bhajibhakare et al., 2020). In general, using CATI can pose certain challenges to this particular survey. In low- and middle- income countries, there is still a gender gap in mobile phone use. In sub-Saharan Africa for example 85% of adult men own a mobile phone compared with 74% of adult women (GSMA 2020). Phone ownership is also lower in rural areas where the phone coverage is even non-existent in certain parts of the area. Research in Nigeria and Ghana indicate that general population mobile phone surveys tend to be over-representative of the male, young, urban and educated people (Lau et al., 2016; L’Engle et al., 2018). Especially during COVID-19 it can be the case that some of the people that are mostly affected are no longer reachable by phone if for example they are hospitalized, or they cannot top up their phone credit.

#### **4.4 Online surveys on CEOs**

Some surveys have attempted to examine the effects of COVID-19 by directly asking the corporate decision makers. The “YPO Chief Executive Global Survey on the Business Impact of COVID-19” was conducted by YPO<sup>6</sup> via online questionnaire from 10 to 13 March 2020. Sample of this survey includes 2,754 chief executives and YPO members from 110 countries, ranging from 25 to 93 years old. Half of the respondents are from the US. The aim of the survey is to get a snapshot of the business impact of COVID-19 and gain insight into the actions that chief executives are taking as a result of the pandemic. In mid-April, YPO conducted a second global COVID-19 survey of chief executives five weeks after the first survey. The sample size was 3,534 chief executives from 109 different countries at the ages of 23-91. Again, half of the survey respondents were from the US.

PwC surveyed CFOs globally about the impact of coronavirus on their business on the “PwC’s COVID-19 CFO Pulse”. The survey was conducted during the week starting at the 23<sup>rd</sup> of March and included 150 finance executives from 8 countries. To expand the global scope, PwC released a second survey on the week of 6<sup>th</sup> of April that included 824 CFOs from 21 countries. In the subsequent third and fourth surveys, PwC is expanding its country and company coverage, resulting in a total of 24 countries in the

---

<sup>6</sup> YPO (formerly Young Presidents’ Organization) is a global leadership community of chief executives with more than 29,000 members in 130 countries.

fourth pulse, namely Central and Southern Africa, Brazil, the Caribbean, China/Hong Kong, Cyprus, the Czech Republic, Denmark, France, Germany, Greece, Kazakhstan, Malaysia, Malta, Mexico, the Middle East, Netherlands, Portugal, Singapore, Slovakia, Sweden, Switzerland, Turkey, the US and Vietnam.

Mercer is surveying companies around the world on a variety of topics relating to the impact of the coronavirus. Questions cover current and future impact on business, preparing to return to work and changes to benefits, compensation, and incentives. The survey is updated with new questions on a monthly basis. By the time of writing, five waves of the survey have been completed. The surveys have been conducted via Keysurvey software and live results are reported in the same platform.<sup>7</sup>

The International Trade Centre launched a worldwide survey in order to assess the economic impact of the coronavirus pandemic on global businesses on the 20<sup>th</sup> of April 2020. The focus of the survey is on how companies have been affected by the pandemic and how government and business support institutions have supported them. The survey is conducted online via the International Trade Centre's website and is available in many languages.<sup>8</sup>

All the surveys conducted by the organizations discussed above offer valuable information regarding the impact of COVID-19 on the corporate sector by reaching to their wide, multi-national network of corporate leaders. However, emphasis is mostly given to the results of the surveys and only little information is publicly available with regards to the survey design. It would be useful to know more on the demographics of the survey, the type of questions asked, the response rate or the data collected from them. Furthermore, as these corporations draw their sample from their existing clients, this may lead to substantial sample selection bias issues, particularly in low- and middle-income economies where typically the most prosperous firms may be clients of big audit firms.

## **5. Online Data Collection on Households**

Other surveys are focusing on the effects of the coronavirus on households. The “Consumers and COVID-19: A real-time survey”, conducted by the Federal Reserve Bank of Cleveland evolves around consumers' expectations for how the economy is likely to be affected by the outbreak of the pandemic and how their own behaviour has changed in response to it (Dietrich et al. 2020). The survey began in March 10, 2020 and is still ongoing. The sampling methodology uses repeated cross sections; there is not a panel aspect that tracks individuals across time. The majority of questions asked are of the “Yes” or “No” type and participants are asked about the expected impact of the COVID-19 pandemic on output and inflation over a one-year horizon. The online survey is administered on the Qualtrics platform and

---

<sup>7</sup> Results can be accessed through <https://taap.mercer.com/covid19results0> .

<sup>8</sup> Survey is available <https://www.intracen.org/survey/Tell-us-how-your-business-is-affected-by-COVID-19/> .

respondents are US consumers. The number of responses per day range between 50 and 208 and by April researchers had solicited around 4,000 households.

Table 5 shows demographic characteristics of the survey respondents. Survey's results are updated frequently on Cleveland Fed's website so as to inform policymakers and researchers about consumers' belief during this time of high uncertainty. Data from this survey are available through Federal Bank of Cleveland's website.<sup>9</sup>

[Table 5 around here]

In response to COVID-19, longitudinal studies have launched special surveys for their participants to understand the immediate and long-term impacts on individuals and households. The 1946 National Survey of Health and Development, managed by the MRC Unit for Lifelong Health and Ageing at UCL and the UCL Centre for Longitudinal Studies' 1958 National Child Development Study, 1970 British Cohort, the Millennium Cohort Study, and Next Steps have launched together an online, nationwide survey to over 50,00 participants to examine the impact of COVID-19. This study has also been developed in collaboration with King's College London, Twins UK team, and the organisations who are running the COVID-19 symptoms tracker app. Members of the tri-ethnic Southall & Brent Revisited cohort, SABRE have also been invited to take part. The focus of this survey is to gather information on various aspects of respondents lives, such as physical health, time-use, household composition, relationships and caring, housing and financial situation, finances and benefits, employment (pre- and post- outbreak), education (pre- and post- outbreak) among others. Data from this survey will help researchers understand the health, social and economic impacts of the COVID-19 outbreak and by following participants' lives into the future, the studies will be able to track the longer-term consequences of the pandemic too. The first data collection round took place in May and the first set of data will be released in summer 2020 through UK Data Service. Further data collections will be made available later in the year.

*Understanding Society*, which follows 40,000 UK households annually, has launched a study among its participants asking them to complete a web-based survey once a month. This survey is part of the UK Household Longitudinal Study and allows researchers to explore how the pandemic is impacting individuals, families and communities across the UK. Participants complete a regular survey, which includes core content designed to track changes, alongside variable content adapted as the coronavirus situation develops. Researchers will be able to link the data from this web survey to answers respondents have given in previous (and future) waves of the annual *Understanding Society* survey. The monthly survey is implemented as a web survey and there is an additional telephone survey for households without internet access. Fieldwork is carried out by Ipsos MORI for the online survey and by Kantar

---

<sup>9</sup> Data available <https://www.clevelandfed.org/en/our-research/indicators-and-data/consumers-and-covid-19.aspx>

for the telephone survey. Data from this study will be made available to researchers each month via the UK Data Service.<sup>10</sup>

The European Foundation for the Improvement of Living and Working Conditions (Eurofound) launched a large-scale online survey entitled “Living, working and COVID-19” on 9 April, 2020 across the European Union to capture the immediate economic and social effects of COVID-19. The aim of this survey is to investigate the impact on well-being, work, remote working (telework) and on the financial situation of EU residents. Most of the questions asked in this survey are based on Eurofound’s European Quality of Life Survey (EQLS) and European Working Conditions Survey (EWCS), while other questions are new or were adapted from other sources, such as the EU Statistics on Income and Living Conditions (EU-SILC). The survey is field in two waves and an additional follow-up for those respondents that agreed to be part of the panel. The first wave took place in April and the second one in May while the follow-up in July. Final results will be published in September 2020. The questionnaire has five main sections- well-being, work and telework, living conditions and financial situation of European Union, socio-demographic and the household composition of the respondent, altogether containing a total of 34 questions. A final question asks the respondents to leave their email address so that it is possible to include a panel element to the survey. The survey is available in 22 languages: Bulgarian, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, German, Greek, Hungarian, Italian, Latvian, Lithuanian, Polish, Portuguese, Romanian, Slovak, Slovene, Spanish and Swedish. The online questionnaire was generated using SoSci Survey. Respondents were recruited using online snowball sampling methods and social media advertisements.

Given the restrictions on face-to-face data collection due to COVID-19, surveys on households have changed to telephone- or web-based modality exclusively. Even though this type of survey can capture valuable information, it fails to capture non-verbal cues, emotions and behaviours that can shed light to various aspects of the households’ *status quo*. Furthermore, in cases where we have repeated cross sections, the loss of participants should be taken into account since there will be households that will stop responding to the survey because they are contacted repeatedly. In addition, it is likely that the respondent households do not have similar characteristics with the non-respondent and the reason for non-response to the survey is associated with the phenomenon being surveyed. For example, non-respondent households could be sick with COVID-19 and hospitalized. This suggest that any analysis on the information obtained could lead to biased results.

---

<sup>10</sup> Other longitudinal studies that have launched special surveys for COVID-19 include Avon Longitudinal Study of Parents and Children, Born in Bradford, English Longitudinal Study of Ageing, Generation Scotland and Aberdeen Children of the 1950s and TwinsUK.

## 6. Online Data Collection on the Government Sector

COVID-19 has prompted a wide range of responses from governments around the globe. The University of Oxford launched the Oxford COVID-19 Government Response Tracker (OxCGRT) during March in order to track governments' policy responses to the coronavirus outbreak. The OxCGRT systemically collects information on several different common policy responses that governments have taken to respond to the pandemic such as income support to citizens, contact tracing or emergency investments into healthcare and can be used as a tool to compare policy responses among countries.

The tracker includes data from over 160 countries and collects publicly available information on 17 indicators of government response. Of these indicators, eight relate to containment and closure policies such as school closures and restrictions in movement, four record information on economic policies such as income support to citizens or foreign aid provision and five relate to health system policies such as the COVID-19 testing regime or emergency investments into healthcare. Data from the indicators are aggregated into four common indices, that range from 1 to 100 based on the level of government action on these areas. The overall government response index reports the government response over all indicators in the database, the containment and health index records lockdown restrictions as well as testing policy, contact tracing or short term investment in healthcare, the economic support index records measures such as income support and debt relief and the original stringency index relates to the strictness of the lockdown measures (Hale et al. 2020). Figure 3 shows mean values for the four indices for a sample of 165 countries starting from the 1<sup>st</sup> of January. The tracker's data and its indices are regularly updated and are available online.<sup>11</sup>

[Figure 3 around here]

The OxCGRT tracker can be a valuable tool for researchers to incorporate into their research in order to explain the differences in the impact of COVID-19 across countries. However, there has been some criticism over the fact that for a few countries there has been a lag in updating the data, for example in the cases of Israel and France. The value of this tool is promising but is yet to be assessed.

## 7. Innovative Online Data collection

### 7.1. What is innovative?

In this section, we discuss conventional methods of collecting data together with their advantages and limitations. We also introduce a digital approach to data collection whereby researchers can access

---

<sup>11</sup> Data is accessible through <https://www.bsg.ox.ac.uk/research/research-projects/coronavirus-government-response-tracker#data>

concealed raw data in open virtual platforms by using data mining tools and data analytical software. This capability of reaching raw data is effective in unlocking timely and rich information that are often packed away or unstructured in the digital world. In seeking a solution for improving the quality and efficiency of data collection for research, we propose a dynamic approach that encourages a hybrid use of conventional and digital methods.

## **7.2. Conventional approach to data collection: to buy or download ready-to-use databases**

Traditionally, researchers rely on using ready-to-use data published by certain organisations (e.g. IMF, World Bank) or buying packaged datasets from database vendors (e.g. Gallup). This source of data is useful for a number of reasons. First, the data are already being collected, processed and organised in standard formats for analysis (e.g. spreadsheets, tables etc.). Second, those datasets usually cover a good range of information that is commonly needed for general research purposes. Third, these dataset providers are generally well acknowledged by the research world (e.g. academic journals, peer researchers), therefore, research works based on their data are often regarded as valid and reliable.

However, many limitations have been experienced while using pre-packaged databases in research. For example, researchers do not have freedom in deciding which variables to be included in the datasets because they had been determined by the database providers or vendors. This lack of choices imposes restrictions on conducting certain research works. Researchers often struggle to find appropriate variables that are needed for answering specific questions. They have to either compromise on research questions based on the nature of the data, or they continue using the data but produce results that cannot reflect the truth. In addition, ready-to-use datasets rarely offer daily data as it takes time for database vendors to process and present raw data. Therefore, when using pre-packaged datasets, researchers are completely dependent on what the data vendors collected and how often they collect and publish data.

Alongside the use of secondary data as mentioned above, there is a long tradition of collecting primary data via survey methods, such as questionnaires and interviews. Despite of the advantage that researchers have the control of what data to collect, surveys are often challenged by issues such as accessibility, coverage, time limitation, sampling, costs and the quality of survey design. See Figure 4. for conventional approach to data collection.

[Figure 4 around here]

## **7.3. Digital approach to data collection**

**Method 1:** To obtain structured raw data via APIs and bring that data back with Tableau®

API refers to Application Programming Interface. It is a computing interface which defines interactions between multiple software intermediaries. API can be understood as a digital key to open a space where companies and organisations store the raw data that record their activities and that they are ready to

share with outsiders; the data are constantly captured by computers and structured in computer languages. The rapid digitalisation in business and daily life encourages more organisations and companies to store, manage and share their data via APIs. Some APIs are offered for free (e.g. ONS, NHS, World Bank, OECD), some APIs charge according to usage (e.g. World Trading Data, Alphavantage).

The increasing popularity of APIs provide researchers with a great opportunity of unlocking abundant near real-time raw data in the open virtual world. Therefore, instead of relying on pre-packaged datasets, we can use APIs to explore and harvest raw data directly from target companies/organisations with the freedom of deciding which information/variables to be extracted and at which frequency (e.g. daily, weekly). This means that we can use this system to meet our exact data requirements/needs. For example, if we wanted to see World Bank's raw data on the weekly difference of the Expenditure of Country A and Country B in 2019, we can programme the system to deliver this, but with a pre-condition that the raw data came to World Bank at such frequency.

Twitter is an example of a social media platform that provides users and companies with access to Twitter data via their APIs<sup>12</sup>. Twitter APIs provide data on the information that users have chosen to share publicly as well as private information (e.g. personal messages) given that additional permission is warranted. Data from Twitter APIs can be used in numerous ways and have applications across many industries, from public health to marketing. For example, NGOs can access Tweets in order to identify misinformation associated with public health issues or spot early indicators for a disease outbreak. The Northeastern University used Twitter data in 2017 to estimate the spread of the seasonal flu in basically real time. When tested against official statistics from surveillance systems, the model has been shown to precisely predict the evolution of the disease up to six weeks in advance. Marketing firms can use Twitter data to recognize topics of interest and tailor their advertising campaigns to reach a wider audience.

Any programming language can be used for developing APIs. It is up to the developer to decide which language works better for him. This choice is guided by a number of parameters, for example prior knowledge of this language, the frameworks available for building APIs or how big and active the community of developers that use this language is. Common choices of languages for developing APIs include Python, R, Ruby and others.

After discovering a resourceful API, we need developer tools to connect it, bring the data back and organise the data in a standard format for analysis. Tableau® is an intelligence software that can help complete this connecting, transporting and organising process. Tableau® enables a powerful tool called

---

<sup>12</sup> Other examples include LinkedIn and Glassdoor.

Web Data Connectors (WDCs), which function as a glue to connect APIs and transport data to our hands. When the data is transported via WDCs to Tableau®, the software automatically converts and organises the data into formats that we are familiar with, i.e. tables, which we can export it into a spreadsheet or start analysis within Tableau®. Developer skills are required when building/programming WDCs in order to tell the software what we need (e.g. parameters) in computer languages.

**Method 2:** To extract unstructured raw data and bring that data back with Tableau®

Apart from collecting structure raw data via APIs, we are able to extract unstructured raw data from almost any open-source web using Tableau® . WDCs allow us to connect to data that is accessible over HTTP and that doesn't already have a connector. This means that we can build our own WDCs to connect to almost any data that is accessible over the web and to create extracts for our own purposes.

Therefore, this digital approach to data collection will enable us to obtain live and accessible raw data directly from original data sources with tailored parameters, thereby enhancing the volume, diversity and quality of incoming data. This will improve the overall quality of research and policy-making capabilities. Nonetheless, the digital methods also face some limitations due to its focus on online data sources. For example, it is difficult to capture non-digital information, such as cash transactions, off-line behaviour and attitude, which are usually captured via survey methods.

#### **7.4. A dynamic approach to data collection**

As discussed above, both conventional and digital methods of data collection have their strengths and weaknesses. This led us to propose a dynamic approach that recommends the use of both methods to complement each other in future research in order to acquire comprehensive and diverse data. Figure 5 illustrates the structure of a dynamic approach to data collection.

[Figure 5 around here]

## **8. Conclusion and going forward**

The COVID-19 pandemic has had large economic consequences worldwide by unleashing sudden changes in household economic behaviour, challenging the long-term survival of many companies, creating a pressing need for appropriate government responses, and testing the resilience of national financial structures and the international financial system. As the crisis caused by the pandemic continues to unfold, it is imperative to understand how stakeholders have fared during this time. The aim of this paper is to evaluate the conventional and digital methods used in an attempt to collect data on COVID-19 for researchers, practitioners and policy makers and to propose ways through which data collection process on COVID-19 can go forward. We are identifying the main mechanisms used to

collect the data, the type of data collected as well as gaps and limitation in the current approaches. To the best of our knowledge, this is the first paper that embarks on such a topic since the coronavirus outbreak. As more and more countries are moving out of the lockdown measures, there is an imperative need for knowledge creation on the impact of COVID-19 and how the post-COVID-19 era will look like for households, companies, banks and governments. The sooner we have data on how these four pillars have reacted to the pandemic, the more informed interventions and planning can be made to go back to normal functioning systems.

More particularly, we reflect upon the type of data that households, companies, banks and international organizations need, the ways through which this data can be collected and how it should be stored and shared. This will result in a compilation of new data using existing primary sources, survey methods and digital approaches. Open data can further enhance the potential of data revolution and by making this data openly available we encourage to guide further research on the ways through which the coronavirus pandemic has impacted the main pillars of the global economy. A comprehensive dataset that will allow for dynamics between households, companies and banks can provide a useful tool for generating projections and statistical forecasts and doing a plethora of readings through big data mining.

It is noted in this paper that questionnaires serve well as a large-scale data collection tool with a choice of open-ended or closed answers. Open questions are useful to collect qualitative data and explore new answers as participants can answer freely, whereas closed questions avoid the fatigue associated with open questions but may exclude answers participants want to use. Self-completion questionnaires are traditionally sent via post or distributed via person, both methods naturally incur costs. However, it is also pointed out that the use of computers has changed all this and digital distribution has brought many benefits including cost and time saved on printing, less need to post or hire infield researchers, a controlled presentation of questions or one at a time questions to reduce bias, flexibility for respondents and efficient collection and storage of data. Nevertheless, data obtained in this way is not without its limitations such as nonresponse, low response rate, ineffective questions and incomplete answers. Since the main disadvantage of self-completion questionnaires are associated with a lack of support between researcher and respondent, structured interviews may be used to overcome this, which would be particularly useful for large scale global research suggested in the paper, to cover the COVID-19 outbreak. The case for unstructured interviews remains, as these are useful for in-depth answers through highly interactive discussions.

Also, it is noted that big data has accelerated advances in computing, marketing and medicine, with machine learning algorithms being a consequence of this, making it possible to study certain topics and assess them empirically – something which was difficult before. Despite these breakthroughs, disadvantages include privacy and confidentiality breaches and the challenge of organising data from a granular dataset to more digestible. However, online data collection methods change due to the

objectives and motivations of the research piece. For example, ONS found using closed, categorical questions the most useful but found it difficult to compare responses with questions changing with dynamic policies, resulting in uneven responses. Another example was that ACCA conducted a global survey which proved valuable for researchers as it allowed comparisons. Nevertheless, there was an uneven response rate among the 100 countries which participated, reducing its validity and its ability to generalise results. The IGC and IPA conducted telephone interviewing via CATI, which is a popular survey among low- and middle- income countries where internet usage is lower, hence leading to a more inclusive and reliable research. The noted disadvantages of telephone interviews included gender gaps in phone usage, phone ownership being lower in rural areas and the general demographic of those who own a mobile being male, young, urban and educated.

Data collection for households were generally closed, yes or no type questions. The data collected was considered very useful as participant's lives were followed into the future, allowing to see the consequences of the pandemic. However, due to face-to-face interviews being cancelled in light of COVID-19, there was a lack of depth in the data that would have been possible outside of a pandemic.

The OxCGRT tracker proved to be a valuable data collection method, comparing common policy responses from governments around the world. Which, unlike data collection from the ACCA, can be compared and seen as reliable. Although not fully assessed, the OxCGRT tracker has collected data globally and is deemed, at the moment, to be a useful tool for researchers.

Up until this point, researchers have relied on premade datasets provided by credible organisations, such as IMF or World Bank. Benefitting from already collected, processed and formatted data – ready to analyse. Though useful, in this paper we aim to evolve to overcome the limitations of using such data, namely, researchers not having freedom to decide what variables to include in datasets, which can be restricting. Ultimately, researchers are dependent on those who define the variables in the datasets and thus their research may be guided by the available data.

The use of API's, across industries, has been increasingly popular (e.g Media, Sports and Travel). Researchers are now able to freely select the data that is the most useful while it being near to real time. Data can then be manipulated into a more digestible form using intelligent software called Tableau<sup>□</sup>. Raw and unstructured data can also be put into Tableau and turned into a customised data set, from original data sources. Increasing the overall quality of research and informing policy responses more accurately. Limitations of a purely digital method include being unable to capture non-digital information i.e. cash transactions and offline behaviour.

In light of the non-digital and digital data collection methods, a combination of the two is what will drive meaningful research on COVID-19 and allow researchers to forecast the probable impact on households, firms, banks and the government.

Often in research, cross-country comparability can be a challenge. There is a high degree of difficulty when cross-country comparisons are made because of differences in the ways in which certain information is reported for each country<sup>13</sup>. Also, institutional differences across countries make cross-national research inherently more difficult than research focused in a single country. The goal of data harmonization is to generate variables that capture the same conceptual information and in the same units. When evaluating a number of harmonization projects, Burkhauser and Lillard (2005) find that these efforts bring significant value to the research community and to policy makers.

At a time when COVID-19 is not fully under control in the world, research on COVID-19 remains a hot topic. Research in the post-COVID-19 world is likely to face important challenges with regards to data availability. The fast pace of the developments (i.e., the outbreak of the disease, the policy interventions) is difficult to be reflected on the current datasets that are available. For example, macroeconomic data are typically available at annual or quarterly frequencies either of which is too coarse to reflect the economic impact of the lockdowns. Consumption has taken a direct hit; yet it is difficult to accurately quantify it due to lack of appropriate consumption datasets at a meaningful frequency and coverage (typically consumption datasets rely on representative surveys and run on an annual basis – see for example the General Household Survey (GHS) for South Africa, the Annual Population Survey (APS) for the UK and the Scottish Household Survey. Against this background new data collection tools need to be designed and offered.

A number of working papers are already seeking to employ alternative tools of data collection for quantifying the impact of COVID-19 on households. Baker et al (2020) use transaction-level daily financial data to examine the impact of COVID-19 on the spending behaviour of Americans. They obtain this data from a non-profit platform that links checking, savings and credit card accounts and facilitates individuals in sustaining saving habits. This kind of data allows researchers for a timely and dynamic approach to identify how household spending has changed during the pandemic. Similarly, for China, Chen et al. (2020) collect daily city-level data on offline consumption by acquiring data from China UnionPay Merchant Services Corporation (UMS) for all spending that is done through UMS POS machines and QR scanners.

Accounting and financial data on firms and banks are typically on an annual (and seldomly on a quarterly) basis at the very best. Shen et al. (2020) and Fu and Shen (2020) use financial data up until the first quarter of 2020 to study the impact of the COVID-19 on the performance of listed Chinese firms. For start-ups and SMEs, the availability of data is an even bigger issue with data on their economic performance often being only annual - see for example Kober et al. (2012) and Lappalainen and Niskanen (2012). Daily stock market data on listed firms and banks have also been used by

---

<sup>13</sup> For example, the abbreviation Plc is used for describing public limited companies in the UK, whereas the equivalent for the US is Inc and for France is SA.

researchers to examine the impact of COVID-19 -see for example Rizwan et al. (2020) and Iyke (2020). However, this kind of data limits the sample to only firms and banks that are listed. The lack of higher frequency and more inclusive financial data has encouraged researchers to employ alternative data collection strategies in order to form timely insights on the impact of the COVID-19 on firms and banks. Kraus et al. (2020) employ a qualitative research design based on 27 semi-structured interviews to evaluate how family firms in Europe cope with the crisis caused by the pandemic. Brown et al. (2020) and Brown and Rocha (2020) use real-time data from Crunchbase for equity investments in the UK and China respectively. Flögel and Gärtner (2020) in their discussion acknowledge the need for collecting additional information from companies and banks on how they have fared during the pandemic. Seven qualitative telephone interviews were conducted with five enterprises and a savings bank in order to investigate the role that regional banks in Germany played during the pandemic.

Official economic indicators are typically available with a lag and for this reason they fail to capture in near-real time the implications of the disease on the economy. In an attempt to overcome this issue, researchers are tracking the evolution of high-frequency variables which can proxy the economic activity. Proxy measures allows researchers to investigate the early economic impact of the pandemic without depending on traditional economic activity indicators which are available on a much slower pace and at lower frequencies. Demirguc-Kunt et al. (2020) use electricity consumption and Nitrogen dioxide ( $NO_2$ ) emissions as proxies for measuring economic activity during the pandemic both of which are available at daily frequency and can provide a near real-time picture of economic changes. Almost every economic activity requires electricity and  $NO_2$  emissions are related to traffic and many construction and industry activities. Similarly, Gu et al. (2020) collect daily electricity usage data from over 34,000 enterprises in Suzhou, China which allows them to construct a measure of firm-level exposure to a particular shock, including the COVID-19 shock.

Research initiatives like the aforementioned highlight the importance of novel data sources. This accelerates the need to design data collection tools and implement them in such a way so that we gain deeper understanding in how COVID-19 has changed the behaviour of households, firms and banks. The data collection on households that we propose includes data on savings, income (for example, income streams from employment or other sources), consumption, loans and consumer confidence<sup>14</sup>. For firms, we suggest collecting data on output, sales, access to loans, supplier credit, working capital, earnings, profit or loss, business optimism, number of employees and share price for listed companies. For banks we propose collecting data on equity, deposits, loans (i.e. loan to individuals, residential

---

<sup>14</sup> The University of Michigan Consumer Sentiment Index is a measure of how optimistic consumers feel about the current state of the economy and is available on a monthly basis. The index is derived from five questions asked to US consumers. These questions refer to changes in their financials during the last year, expectations on their future financials, expectations on the national financial situation over the next year and the next five years and the appropriateness of buying major households durables at present.

mortgages, agricultural loans etc), securities bought or sold, overdraft, interest rate on loans (approximating cost of capital), interest rate on deposits, gross income, new loans contracted, new financial products and share price for the banks that are listed in the stock exchange. A plethora of feasible projects can be done given that this kind of data is available to the research community. Some examples include investigating the portfolio behaviour of households during the pandemic, examining the current trends in household consumption with particular emphasis on the means of payment, assessing the flow of finance from banks to companies in an attempt to ease the economic effects of the lockdown or examining the impact of COVID 19 on the link between stock prices and exchange rates.

Our proposed data collection strategies would be of interest to a wide range of parties including academic researchers, policy makers, investors and scientific organisations. For example, prospective project sponsors may find commercial interest in the proposed database that we identify in this document. We also believe that the proposed study could be particularly relevant for international organisations. International organisations face increasing pressure to better monitor global developments and the rise of unstructured data requires updated strategies for collecting digital data. This can be tackled by integrating the potential of new data collection methods in their operations and mandates.

## REFERENCES

- AXA. 2020. *AXA Future Risks Report 2020*. Paris: AXA. <https://www.axa.com/en/magazine/2020-future-risks-report>
- Baker, S.R., Farrokhnia, R.A., Meyer, S., Pagel, M. and Yannelis, C., 2020. *How does household spending respond to an epidemic? consumption during the 2020 covid-19 pandemic* (No. w26949). National Bureau of Economic Research.
- Baltas, G. and Argouslidis, P.C., 2007. Consumer characteristics and demand for store brands, *International Journal of Retail & Distribution Management*, 35 (5), pp. 328-341.
- Biemer, P.P. and Lyberg, L.E., 2003. *Introduction to Survey Quality*. New York John Wiley & Sons.
- Brown, R. and Rocha, A., 2020. Entrepreneurial uncertainty during the Covid-19 crisis: Mapping the temporal dynamics of entrepreneurial finance. *Journal of Business Venturing Insights*, 14, p.e00174.
- Brown, R., Rocha, A. and Cowling, M., 2020. Financing entrepreneurship in times of crisis: Exploring the impact of COVID-19 on the market for entrepreneurial finance in the United Kingdom. *International Small Business Journal*, 38(5), pp.380-390.
- Bryman, A. and Bell, E., 2011. *Business Research Methods*. Oxford: Oxford University Press.
- Burkhauser, R.V. and Lillard, D.R., 2005. The contribution and potential of data harmonization for cross-national comparative research. *Journal of Comparative Policy Analysis*, 7(4), pp.313-330.
- Buskirk, T.D. and Andrus, C., 2012. Smart surveys for smart phones: Exploring various approaches for conducting online mobile surveys via smartphones. *Survey Practice*, 5(1), pp.1-11.
- Cavallo, A., 2013. Online and official price indexes: measuring Argentina's inflation. *Journal of Monetary Economics*, 60(2), pp.152-165.
- Chen, H., Qian, W. and Wen, Q., 2020. The impact of the COVID-19 pandemic on consumption: Learning from high frequency transaction data. *Available at SSRN 3568574*.
- Chetty, R., Friedman, J.N. and Rockoff, J.E., 2014. Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood. *American Economic Review*, 104(9), pp.2633-79.
- Corbin, J. and Morse, J.M., 2003. The unstructured interactive interview: Issues of reciprocity and risks when dealing with sensitive topics. *Qualitative Inquiry*, 9(3), pp.335-354.
- Culkin, R. and Das, S.R., 2017. Machine learning in finance: The case of deep learning for option pricing. *Journal of Investment Management*, 15(4), pp.92-100.
- De Leeuw, E.D., 2005. To mix or not to mix data collection modes in surveys. *Journal of Official Statistics*, 21(5), pp.233-255.
- De Leeuw, E. D., 2008. Choosing the method of data collection. In E. D. de Leeuw, J. J. Hox, & D. A. Dillman (Eds.), *International Handbook of Survey Methodology* (p. 113–135). London: Taylor & Francis Group/Lawrence Erlbaum Associates.
- De Prado, M.L., 2016. Invited Editorial Comment: Mathematics and Economics: A Reality Check.

- Delmas, M., & Pekovic, S., 2013. Environmental standards and labor productivity: Understanding the mechanisms that sustain sustainability. *Journal of Organizational Behavior*, 34(2), pp.230-252.
- Demirguc-Kunt, A., Lokshin, M. and Torre, I., 2020. The sooner, the better: The early economic impact of non-pharmaceutical interventions during the COVID-19 pandemic. *World Bank Policy Research Working Paper*, (9257).
- Dietrich, Alexander M., Keith Kuester, Gernot J. Müller, and Raphael S. Schoenle, 2020. News and Uncertainty about COVID-19: Survey Evidence and Short-Run Economic Impact. *Federal Reserve Bank of Cleveland, Working Paper No. 20-12*.
- Easterby-Smith, M., Thorpe, R. and Jackson, P.R., 2015. Designing management and business research. *Management and Business Research*, pp.66-106.
- Eckhardt, G.M., Belk, R. and Devinney, T.M., 2010. Why don't consumers consume ethically? *Journal of Consumer Behaviour*, 9(6), pp.426-436.
- Fang, Z. and George, K.M., 2017, November. Application of machine learning: An analysis of Asian options pricing using neural network. In *2017 IEEE 14th International Conference on e-Business Engineering (ICEBE)* (pp. 142-149). IEEE.
- Flögel, F. and Gärtner, S., 2020. The COVID-19 Pandemic and Relationship Banking in Germany: Will Regional Banks Cushion an Economic Decline or is A Banking Crisis Looming? *Tijdschrift voor economische en sociale geografie*, 111(3), pp.416-433.
- Fontana, A. and Frey, J.H., 2005. The interview: From neutral stance to political involvement.
- Fu, M. and Shen, H., 2020. COVID-19 and corporate performance in the energy industry. *Energy Research Letters*, 1(1), p.12967.
- Gan, L., Wang, H. and Yang, Z., 2020. Machine learning solutions to challenges in finance: An application to the pricing of financial products. *Technological Forecasting and Social Change*, 153, p.119928.
- Ghodduzi, H., Creamer, G.G. and Rafizadeh, N., 2019. Machine learning in energy economics and finance: A review. *Energy Economics*, 81, pp.709-727.
- GSMA Connected Women, 2020. The mobile gender gap report 2020.
- Gu, X., Ying, S., Zhang, W. and Tao, Y., 2020. How Do Firms Respond to COVID-19? First Evidence from Suzhou, China. *Emerging Markets Finance and Trade*, 56(10), pp.2181-2197.
- Hale, Thomas, Noam Angrist, Beatriz Kira, Anna Petherick, Toby Phillips, Samuel Webster, 2020. Variation in Government Responses to COVID-19 Version 6.0. *Blavatnik School of Government Working Paper*.
- Hofstede, G.H., 1980/1984. *Culture Consequences: International Differences in Work-related Values*. Sage Publications. London.
- Hutchinson, J.M., Lo, A.W. and Poggio, T., 1994. A nonparametric approach to pricing and hedging derivative securities via learning networks. *The Journal of Finance*, 49(3), pp.851-889.
- Iyke, B.N., 2020. COVID-19: The reaction of US oil and gas producers to the pandemic. *Energy Research Letters*, 1(2), p.13912.
- Janice A. Derr, Diane C. Mitchell, Diane Brannon, Helen Smiciklas-Wright, Lori Beth Dixon, Barbara M. Shannon, 1992. Time and Cost Analysis of a Computer-Assisted Telephone

- Interview System to Collect Dietary Recalls, *American Journal of Epidemiology*, 136 (11), pp. 1386–1392
- Kitzinger, J., 1994. The methodology of focus groups: the importance of interaction between research participants. *Sociology of health & illness*, 16(1), pp.103-121.
- Kober, R., Subraamanniam, T. and Watson, J., 2012. The impact of total quality management adoption on small and medium enterprises' financial performance. *Accounting & Finance*, 52(2), pp.421-438.
- Koschate-Fischer, N., Cramer, J., & Hoyer, W. D., 2014. Moderating Effects of the Relationship Between Private Label Share and Store Loyalty. *Journal of Marketing*, 78(2), 69–82.
- Kraus, S., Clauss, T., Breier, M., Gast, J., Zardini, A. and Tiberius, V., 2020. The economics of COVID-19: initial empirical evidence on how family firms in five European countries cope with the corona crisis. *International Journal of Entrepreneurial Behavior & Research*, in press.
- Krueger, R.A. and Casey, M.A., 2000. A practical guide for applied research. *A practical guide for applied research*.
- Lafuma, A., Laurendeau, C., Lamerain, E. and Berdeaux, G. 2009. Economics and Attitudes Regarding Spectacles in Daily Life: A European Perspective, *Ophthalmic Epidemiology*, 16(4), pp. 218-223.
- Lappalainen, J. and Niskanen, M., 2012. Financial performance of SMEs: impact of ownership structure and board composition. *Management research review*, 35(11), pp.1088–1108.
- Lau, C. Q., Cronberg, A., Marks, L. and Amaya, A. (2019) “In Search of the Optimal Mode for Mobile Phone Surveys in Developing Countries. A Comparison of IVR, SMS, and CATI in Nigeria”, *Survey Research Methods*, 13(3), pp. 305-318.
- L'Engle, K., Sefa, E., Adimazoya, E. A., Yartey, E., Lenzi, R., Tarpo, C., Heward-Mills, N. L., Lew, K., & Ampeh, Y. (2018). Survey research with a random digit dial national mobile phone sample in Ghana: Methods and sample quality. *PloS one*, 13(1), e0190902.
- Levy, S.J., 1981. Intepreting consumer mythology: a structural approach to consumer behavior. *Journal of marketing*, 45(3), pp.49-61.
- Lofland, J., & Lofland, L. H. (1984). Analyzing social settings: A guide to qualitative observations and analysis (2nd ed.). Belmont: Wadsworth.
- Morgan, D.L., 1996. Focus groups. *Annual Review of Sociology*, 22(1), pp.129-152.
- Nathan, G., 2001. Telesurvey Methodologies for Household Surveys-A Review and Some Thoughts for the Future?. *Survey Methodology*, 27(1), pp.7-32.
- Rizwan, M.S., Ahmad, G. and Ashraf, D., 2020. Systemic Risk: The Impact of COVID-19. *Available at SSRN 3615161*.
- Rundo, F., Trenta, F., Di Stallo, A.L. and Battiato, S., 2019. Advanced Markov-based machine learning framework for making adaptive trading system. *Computation*, 7(1), p.4.
- Saurabh Bhajibhakare, Ambika Chopra, Putul Gupta, and Mustufa Patel, 2020. Transitioning to CATI: Checklists and Resources, J-PAL South Asia, March 2020.
- Selvin, S., Vinayakumar, R., Gopalakrishnan, E.A., Menon, V.K. and Soman, K.P., 2017. Stock price prediction using LSTM, RNN and CNN-sliding window model. *2017 International*

*Conference on Advances in Computing, Communications and Informatics (icacci)*, September 2017. IEEE, pp. 1643-1647)

- Shen, H., Fu, M., Pan, H., Yu, Z. and Chen, Y., 2020. The Impact of the COVID-19 Pandemic on Firm Performance. *Emerging Markets Finance and Trade*, 56(10), pp.2213-2230.
- Siami-Namini, S., Tavakoli, N. and Namin, A.S., 2018. A comparison of ARIMA and LSTM in forecasting time series. *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*, December 2018. IEEE, pp. 1394-1401.
- Spitzmüller, C., Glenn, D.M., Barr, C.D., Rogelberg, S.G. and Daniel, P., 2006. “If you treat me right, I reciprocate”: Examining the role of exchange in organizational survey response. *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior*, 27(1), pp.19-35.
- Spradley, J., 1979. Asking descriptive questions. *The Ethnographic Interview*, 1, pp.44-61.
- Stern, M.J., Bilgen, I. and Dillman, D.A., 2014. The state of survey methodology: Challenges, dilemmas, and new frontiers in the era of the tailored design. *Field Methods*, 26(3), pp.284-301.
- Sun, S., Sun, Y., Wang, S. and Wei, Y., 2018. Interval decomposition ensemble approach for crude oil price forecasting. *Energy Economics*, 76, pp.274-287.
- Wilkinson, S., 2004. Focus groups. *Doing Social Psychology Research*, pp.344-376.
- Yu, L., Zhao, Y., Tang, L. and Yang, Z., 2019. Online big data-driven oil consumption forecasting with Google trends. *International Journal of Forecasting*, 35(1), pp.213-223.

Table 1: Reference and collection period for each wave of the “Business Impact of Coronavirus Survey (BICS)” undertaken by the Office for National Statistics.

<b>Wave</b>	<b>Reference Period</b>	<b>Collection Period</b>
1	9 to 22 March 2020	23 March to 5 April 2020
2	23 March to 5 April 2020	6 to 19 April 2020
3	6 to 19 April 2020	20 April to 3 May 2020
4	20 April to 3 May 2020	4 to 17 May 2020
5	4 to 17 May 2020	18 to 31 May 2020
6	18 to 31 May 2020	1 to 14 June 2020
7	1 to 14 June 2020	15 to 28 June 2020

Notes: Table presents the reference and collection period for the waves of the “Business Impact of Coronavirus Survey (BICS)” undertaken by the Office for National Statistics. Wave 1 was considered a pilot study with high-level questions. From Wave 2 onwards more detailed questions were added. For the comparativeness of the analysis, they use questions that have common response options from Waves 2 to 7. Source: Office for National Statistics – Business Impact of Coronavirus Survey.

Table 2: Total sample and response rates for each wave of the “Business Impact of Coronavirus Survey (BICS)” undertaken by the Office for National Statistics.

	<b>Wave 1</b>	<b>Wave 2</b>	<b>Wave 3</b>	<b>Wave 4</b>	<b>Wave 5</b>
	9 April 2020	23 April 2020	7 May 2020	4 June 2020	21 June 2020
Sample	17,786	17,786	17,623	18,506	20,566
Response	4,598	6,171	6,114	6,196	6,364
Rate	25.9%	34.7%	34.7%	33.5%	30.9%

Notes: Table presents total sample and response rates for each wave of BICS. Source: Office for National Statistics – Business Impact of Coronavirus Survey.

Table 3: Percentage of UK businesses continuing to trade, broken down by industry for the period 4 May to 17 May 2020

Industry	Turnover Decreased by more than 50%	Turnover Decreased between 20% and 50%	Turnover Decreased by up to 20%	Turnover Unaffected	Turnover Affected but within normal range	Turnover Increased by up to 20%	Turnover Increased between 20% and 50%	Turnover Increased by more than 50%	Not Sure
Manufacturing	29.1%	23.6%	11.5%	27.8%	1.4%	2.5%	*	*	2.8%
Water Supply, Sewerage, Waste Management and Remediation Activities	17.5%	36.5%	15.9%	27.0%	0.0%	1.6%	0.0%	0.0%	1.6%
Construction	40.2%	26.7%	10.0%	16.1%	*	0.0%	1.3%	0.0%	5.1%
Wholesale and Retail Trade; Repair of Motor Vehicles and Motorcycles	33.4%	23.3%	8.3%	21.0%	1.6%	3.9%	3.5%	1.0%	4.0%
Transportation and Storage	26.6%	29.2%	16.4%	18.2%	1.8%	2.2%	1.1%	*	3.6%
Accommodation and Food Service Activities	62.4%	9.9%	2.1%	14.9%	1.4%	2.1%	0.0%	*	6.4%
Information and Communication	8.2%	14.6%	14.9%	53.3%	2.6%	0.8%	0.3%	*	5.1%
Real Estate Activities	27.3%	19.7%	10.6%	33.3%	3.0%	1.5%	1.5%	0.0%	3.0%
Professional, Scientific and Technical Activities	11.2%	21.4%	20.6%	39.2%	1.9%	*	0.0%	*	5.4%
Administrative and Support Service Activities	27.0%	26.7%	14.2%	23.9%	1.7%	1.1%	0.5%	*	4.6%
Education	9.7%	12.6%	34.6%	36.1%	1.5%	0.0%	0.0%	0.0%	5.6%
Human Health and Social Work Activities	7.6%	14.4%	12.7%	53.4%	6.8%	0.0%	0.0%	0.0%	5.1%
Arts, Entertainment and Recreation	62.7%	13.4%	9.0%	11.9%	0.0%	0.0%	0.0%	0.0%	3.0%
<b>All Industries</b>	<b>25.7%</b>	<b>22.2%</b>	<b>14.0%</b>	<b>28.8%</b>	<b>1.7%</b>	<b>1.7%</b>	<b>1.1%</b>	<b>*</b>	<b>4.3%</b>

Notes: Table presents the percentages of UK businesses continuing to trade, broken down by industry for the period 4 May to 17 May 2020. Other services and Mining and Quarrying have been removed for disclosure purposes, but their totals are included in "All industries". Rows might not sum to 100% due to rounding. \* percentage less than 1% or industry count less than 10. Source: Office for National Statistics – Business Impact of Coronavirus Survey.

Table 4: Survey Demographics for the “COVID-19 Global Survey: Inside business-impacts and responses” conducted by ACCA – Top 20 responses

<b>Country</b>	<b>Responses</b>	<b>Country</b>	<b>Responses</b>
1. UK	2774	11. Nigeria	227
2. Malaysia	919	12. Canada	226
3. Republic of Ireland	493	13. UAE	209
4. Hong Kong SAR	481	14. Cyprus	195
5. Pakistan	353	15. Australia	185
6. Ghana	295	16. Trinidad & Tobago	143
7. Vietnam	279	17. Uganda	133
8. Mainland China	260	18. Zambia	133
9. Mauritius	236	19. USA	106
10. Singapore	234	20. Russia	94

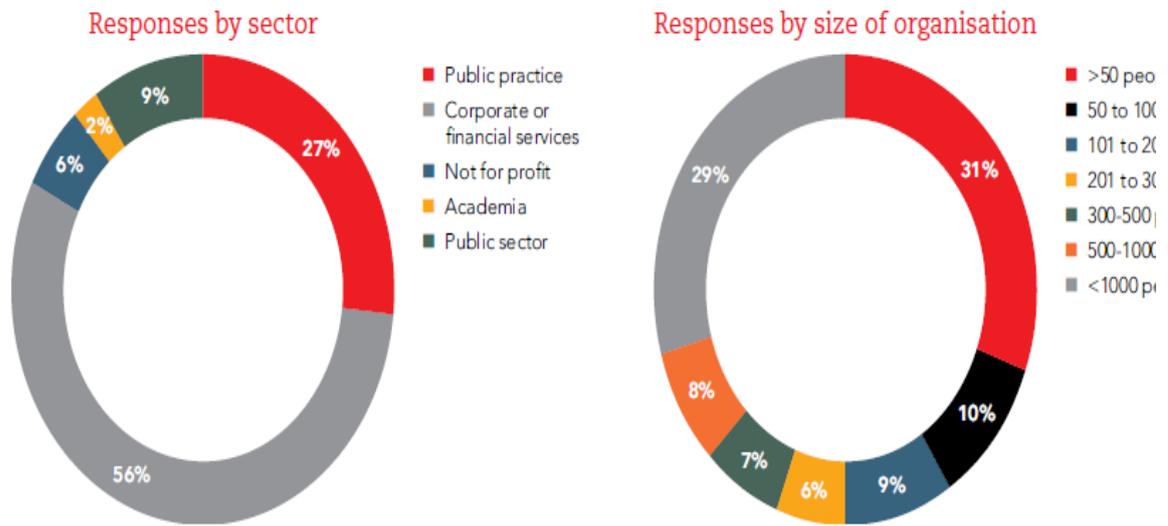
Notes: Figure shows top 20 responses for the “COVID-19 Global Survey: Inside business-impacts and responses” conducted by ACCA. Source: Association of Chartered Certified Accountants-COVID-19 Global Survey: Inside Business Impacts and Responses.

Table 5: Survey Respondent Characteristics for the “Consumers and COVID-19: A real-time survey” conducted by Federal Reserve Bank of Cleveland.

	Percentage		Percentage
<b>Age</b>		<b>Race</b>	
18-34	33.09%	Non-Hispanic white	68.44%
35-55	35.22%	Non-Hispanic black	13.31%
Older than 55	31.69%	Hispanic	8.16%
		Asian or other	10.09%
<b>Gender</b>		<b>Household Income</b>	
Female	50.00%	Less than \$50k	42.35%
Male	49.73%	\$50k-\$100k	39.65%
Other	0.27%	More than \$100k	18.00%
<b>Region</b>		<b>Education</b>	
Midwest	20.05%	Some college or less	58.01%
Norththeast	18.38%	Bachelors degree or more	41.99%
South	41.77%		
West	19.81%		N=3954

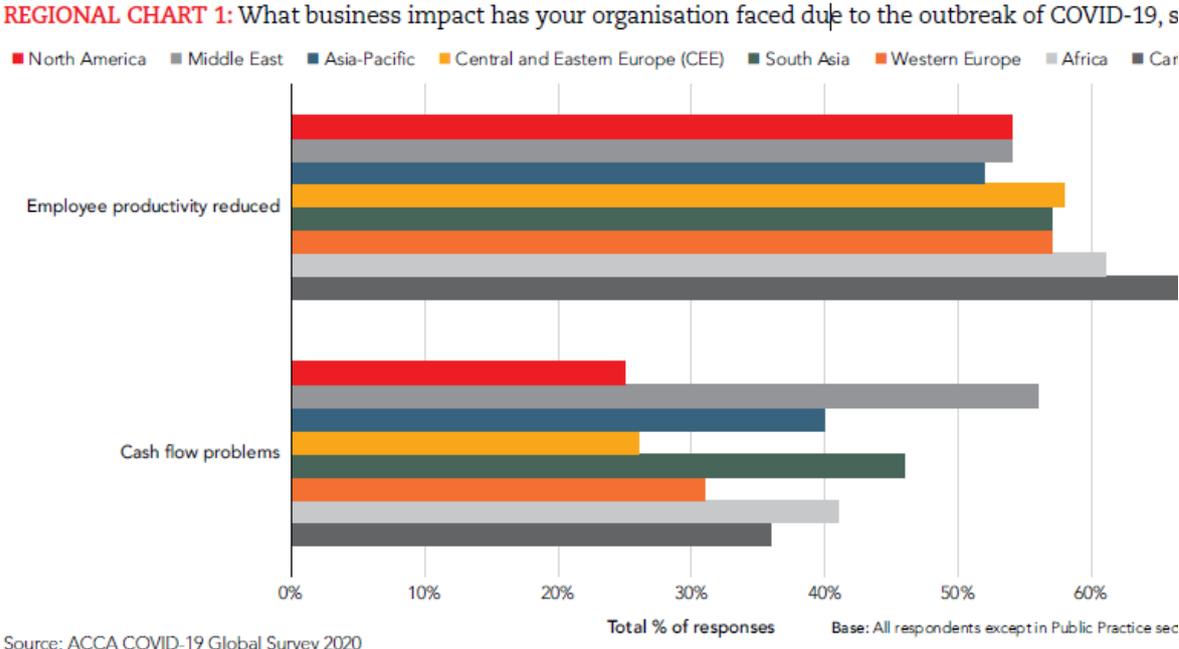
Notes: Table presents data on the characteristics of participants in the online survey conducted by Federal Reserve Bank of Cleveland. Source: Dietrich et al. (2020).

Figure 1: Survey Demographics for the “COVID-19 Global Survey: Inside business-impacts and responses” conducted by ACCA – Responses by sector and by size of organisation



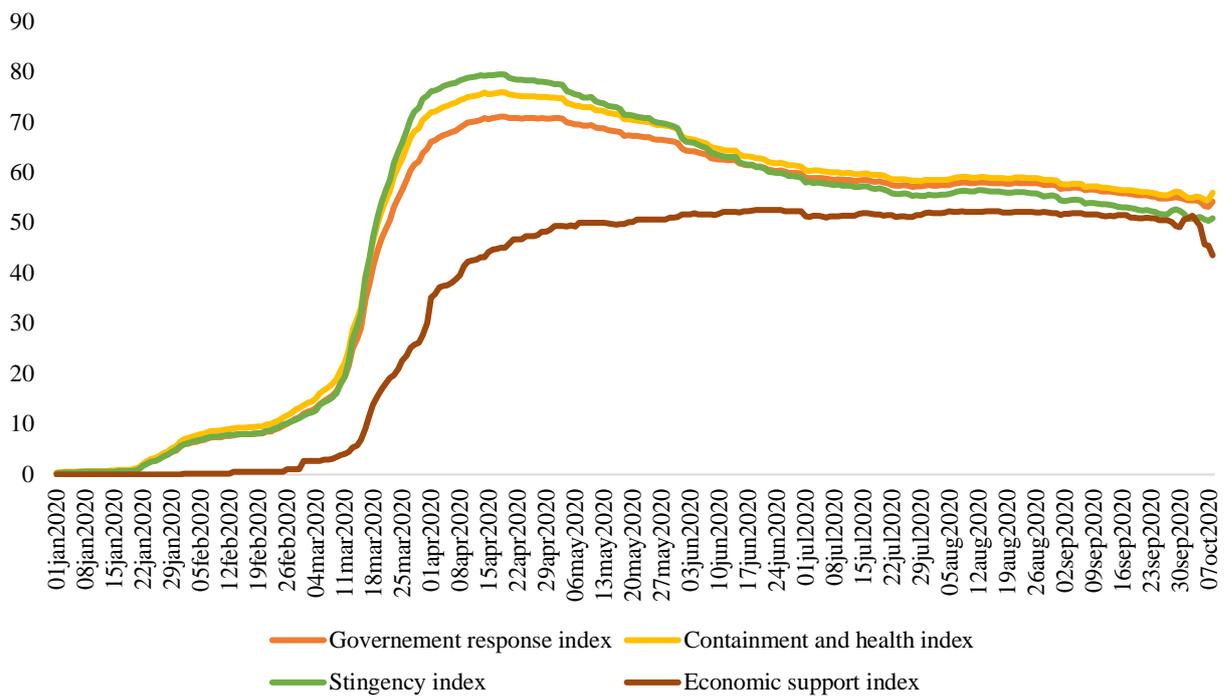
Notes: Figure shows Survey Demographics for the “COVID-19 Global Survey: Inside business-impacts and responses” conducted by ACCA by sector and size of the organisation. Source: Association of Chartered Certified Accountants- COVID-19 Global Survey: Inside Business Impacts and Responses.

Figure 2: Regional Comparisons for the “COVID-19 Global Survey: Inside business-impacts and responses” conducted by ACCA.



Notes: Figure shows regional comparisons for the “COVID-19 Global Survey: Inside business-impacts and responses” conducted by ACCA. Source: Association of Chartered Certified Accountants-COVID-19 Global Survey: Inside Business Impacts and Responses.

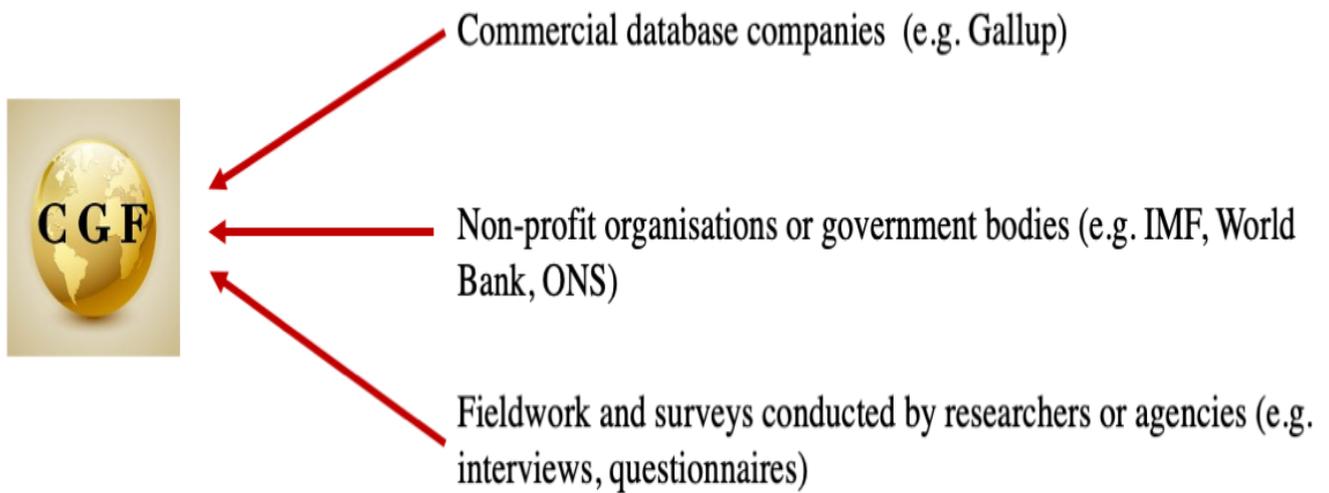
Figure 3: Global Mean Index Values for all countries over time



Notes: Figure reports mean values for the four indices over time for all countries from 1<sup>st</sup> January 2020 to 8<sup>th</sup> October 2020

Figure 4: Conventional approach to data collection

---

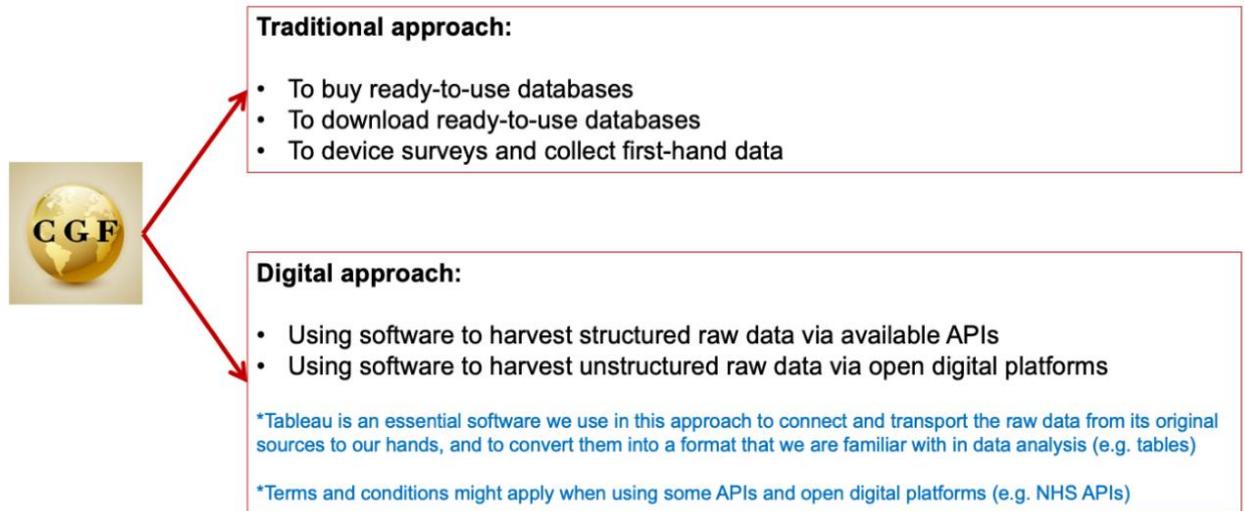


---

Notes: Figure shows conventional approaches to data collection which include buying packaged datasets from database vendors, using ready-to-use datasets by non-profit organisations and government bodies or conducting fieldwork and surveys to collect primary data.

Figure 5: A dynamic approach to data collection

---



---

Notes: Figure illustrates the traditional and digital approach to data collection. The digital approach includes using software packages to harvest structured raw data available via APIs and to harvest unstructured raw data via open digital platforms.

## Appendix

### Appendix 1: Example of closed question design in questionnaires

1) *Binary format:*

My organisation is a friendly place to work.

Agree            Disagree            Not sure

2) *Likert scale with verbal answers*

My organisation is a friendly place to work.

Strongly disagree, Disagree, Not sure, Agree, Strongly agree

3) *Likert scale with numerical answers*

My organisation is a friendly place to work.

-2, -1, 0, 1, 2 (where -2 means Strongly disagree and 2 means Strongly agree)

4) *Continuous scales with numerical answers*

My organisation is a friendly place to work

1 2 3 4 5 6 7 8 (where 1 means Strongly disagree and 8 means Strongly agree)

5) *Categorical question with a single choice*

How much do you spend on food every week?

- A. Below \$100
- B. \$100 - \$300
- C. \$301 - \$500
- D. Above \$500

6) *Categorical question with multiple choices*

Please select countries you have travelled to in the past 3 years.

- A. India
- B. United States
- C. Japan
- D. France
- E. Italy
- F. None above