

# Data Matching: Construction of COVID-19- related Variables for Young Lives Peru

Alessandra Hidalgo, Alan Sánchez, and Santos Zhu



# Data Matching: Construction of COVID-19-related Variables for Young Lives Peru

Alessandra Hidalgo, Alan Sánchez, and Santos Zhu

First published by Young Lives in March 2024

© Young Lives 2024

Printed on FSC-certified paper from traceable and sustainable sources.

## About Young Lives

Young Lives is an international study of poverty and inequality, following the lives of 12,000 children in four countries (Ethiopia, India, Peru and Vietnam) since 2001. [www.younglives.org.uk](http://www.younglives.org.uk)

Young Lives is funded by UK aid from the UK Government.

The views expressed are those of the author(s). They are not necessarily those of, or endorsed by, Young Lives, the University of Oxford, the UK Foreign, Commonwealth & Development Office or other funders.



**Young Lives**, Oxford Department of International Development (ODID), University of Oxford,  
Queen Elizabeth House, 3 Mansfield Road, Oxford OX1 3TB, UK

Tel: +44 (0)1865 281751 • Email: [younglives@qeh.ox.ac.uk](mailto:younglives@qeh.ox.ac.uk)

# Contents

<b>The authors</b>	<b>4</b>
<b>Acknowledgements</b>	<b>4</b>
<b>Summary</b>	<b>4</b>
<b>1. Introduction</b>	<b>5</b>
<b>2. Context</b>	<b>6</b>
<b>3. Data sources</b>	<b>9</b>
<b>4. Estimating COVID-19 related variables</b>	<b>12</b>
<b>5. Matched dataset</b>	<b>16</b>
<b>6. Conclusions</b>	<b>18</b>
<b>References</b>	<b>18</b>
<b>Appendix</b>	<b>19</b>

## The authors

**Alessandra Hidalgo** is a PhD candidate in economics at Lancaster University. Her academic journey includes an MSc in Economics from Lancaster University as part of her integrated PhD program and an MA in Economics from Universidad de Piura, Peru. She previously worked for Young Lives as a Research Consultant during the summer of 2023 and 2022. Additionally, she has worked from 2017 to 2020 as a Research Assistant with the Young Lives Peru team at GRADE in Lima, Peru.

**Alan Sánchez** is the Senior Quantitative Researcher at Young Lives in Oxford. He has been the Principal Investigator of Young Lives Peru since 2012, now sharing this responsibility with Antonio Campos. He is also a Senior Researcher at GRADE.

**Santos Zhu** is an Economic Analyst at the Ministry of Education of Peru. He previously worked for Young Lives as a Research Assistant with the Young Lives Peru team at GRADE during 2020 and 2021. His academic journey includes a bachelor's degree from the Universidad Nacional Mayor de San Marcos, Peru.

## Acknowledgements

This study was funded by the UK's Foreign, Commonwealth and Development Office (FCDO) (Grant no. GB-GOV-1-301108). We are grateful to Monica Lizama (Young Lives Peru Data Manager) for her support in curating and accessing the Young Lives geographic data and for her feedback, and to Marta Favara and participants at the 'Data matching day' in Oxford in 2023 for providing very useful comments. We are also grateful to Sofia Madrid (Young Lives Peru Field Manager), the team of enumerators that administered the Young Lives phone survey between 2020 and 2021. Thanks also to Adam Houlbrook for copyediting, Garth Stewart for design, and Julia Tilford for oversight of this publication.

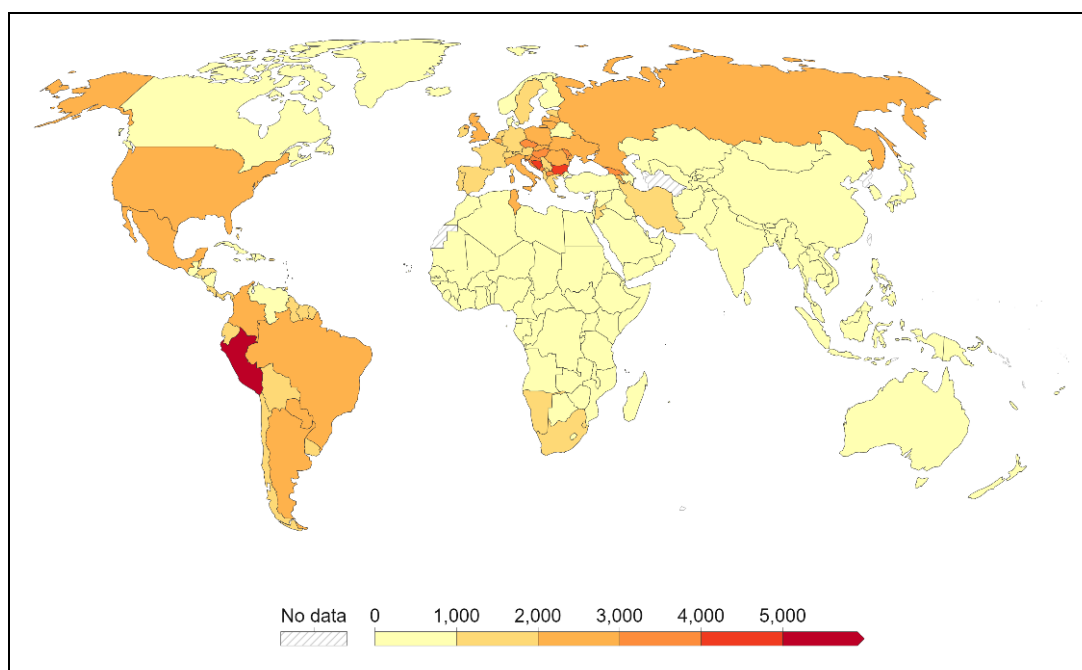
## Summary

The COVID-19 pandemic significantly affected many countries. Peru had one of the highest rates of COVID-19 deaths per capita worldwide, despite implementing some of the most restrictive lockdown policies. To understand the differentiated impacts that occurred in a geographically heterogeneous country such as Peru, we matched publicly available COVID-19 data with comprehensive longitudinal data for Young Lives Peru, including the five COVID-19 phone surveys between August 2020 and December 2021. The final dataset illustrates the spread of COVID-19 cases, fatalities, the timing, duration and geographical coverage of the lockdown policies implemented by the Peruvian government, and the vaccine distribution in the areas where the Young Lives participants were living at the time of the phone calls. This technical note describes how the COVID-19 data was matched to the Young Lives Peru sample and how each of the variables included in the final dataset were defined.

# 1. Introduction

Peru is classified as an upper-middle-income country by the World Bank and is the third-largest country in South America, with a population of over 32 million people. During the COVID-19 pandemic the Peruvian population faced unique challenges. In the early stages of the pandemic in mid-March 2020, the Peruvian government implemented some of the most restrictive lockdown measures globally (Hale et al. 2021) and closed its borders swiftly to stop the spread of the virus. However, the country's health care system struggled to cope with the surge in cases, leading to one of the highest rates of COVID-19 deaths per capita in the world by the end of 2021 (Figure 1). Peru's economy saw a significant fall of 11 per cent in GDP in 2020, as well as an increase in poverty rates to 30 per cent. Problems with access to health care and testing in remote and precarious regions introduced additional obstacles. Despite these difficulties, Peru made efforts to provide financial assistance through various cash transfer programmes, secure vaccines and expand health facilities to improve its response as the pandemic progressed.

**Figure 1.** *Cumulative confirmed COVID-19 deaths per million people up to 30 December 2021*



Source: WHO COVID-19 Dashboard (WHO 2023).  
Elaboration: Our World in Data.

The health impact of the pandemic in Peru – as reflected by the number of cases and related deaths – varied significantly by geographical location and over time. Similarly, although lockdown measures initially encompassed the entire nation, they later shifted to specific regions or provinces. Thus, any attempt to assess the overall impact of the pandemic on people's lives in Peru might benefit from integrating COVID-19 administrative data into the analysis.

Young Lives follows the lives of 12,000 young people from Ethiopia, India (the states of Telangana and Andhra Pradesh), Peru and Vietnam, who were initially surveyed in 2002 and most recently surveyed in person in 2016. In Peru, the study collected data on two cohorts: the Younger Cohort born around 2000 and the Older Cohort, who were around 8 years old at that time. While not nationally representative, the sample covers a great wealth spectrum and includes participants in urban and rural areas, and in the three climatic regions (coast, highlands, and Amazonian jungle). In 2020, the COVID-19 outbreak interrupted the planned sixth round of in-person data collection, when participants were between 18–19 and 25–26 years old. As a result of the worldwide pandemic, the Young Lives team switched from in-person surveys to phone interviews, the 'Listening to Young Lives at Work: COVID-19 Phone Survey'. The surveys were conducted between August 2020 and December 2021 and aimed to assess the short-term impacts of COVID-19 on the health, well-being, transition to the labour market and education trajectories of the young people in the Young Lives study (Sánchez, et al. 2020).

This technical note lists the steps followed to construct a group of variables to measure the number of COVID-19 cases and deaths observed in the districts where the Young Lives participants lived at the time of the interview, and the approximate number of days in lockdown endured by them. To do this, we used administrative data and merged this with the Young Lives Call 2 (August to October 2020), Call 3 (November to December 2020), Call 4 (August to September 2021) and Call 5 (November to December 2021) datasets, according to the interview date and the respondent's geographical location.<sup>1</sup>

The rest of the note is organised as follows: Section 2 provides the context of the pandemic experience in Peru, government measures and emerging impacts; Section 3 describes the data from administrative open sources that are matched to the Young Lives Peru data; Section 4 discusses the methodology used to calculate the variables of interest; while Section 5 describes the structure and content of the publicly archived dataset.

## 2. Context

Peru had one of the highest per capita rates of COVID-19 cases and deaths globally (Figure 1), despite enforcing lengthy and strict stay-at-home measures. Soon after detecting its first COVID-19 cases, Peru declared a national emergency, closed its borders and imposed mandatory social isolation, starting on 16 March 2020. Initially planned for 15 days, this lockdown restricted movement to essential activities such as buying groceries and medicines and participating in critical economic activities. It excluded physical attendance at schools and universities. Subsequently, more restrictions were introduced, including limitations on children and senior citizens' movement, Sunday bans, and an evening curfew from 4pm to 5am.

The lockdown was extended multiple times, lasting a total of 107 days, and the duration of the daily curfew was gradually reduced. In July, Peru transitioned to local lockdowns in areas with surging COVID-19 cases, which persisted until 30 September 2020. Educational institutions

---

<sup>1</sup> These datasets are available to download at the UK Data Archive: <https://www.data-archive.ac.uk>.

remained closed throughout the year, and senior citizens and children faced mandatory social isolation until October and November 2020, respectively. As of 1 January 2021, an evening curfew remained in place from 11pm to 4am. Peru's response to the pandemic was highly stringent, reaching a peak of 96.3 out of a possible 100 on the COVID-19 Government Stringency Index in May 2020, surpassing measures taken by many other countries (Hale et al. 2021).<sup>2</sup>

The first confirmed case of COVID-19 in Peru was announced on 6 March 2020. Approximately one month later, and with the progressive increase in the number of cases detected, the first wave of infections in the country officially began. During this first period, the average case fatality rate was 15 per cent (Araujo-Castillo 2022), and the most vulnerable populations were elderly people and people with comorbidities. This first wave of infections would begin to disappear at the beginning of September 2020 and reached a stable level between October and December of 2020.

In January 2021, the second wave began, with some regions returning to full lockdowns. Similar to the first wave, the second wave showed the vulnerability of the Peruvian health system to contain the increase in cases and treat affected persons, due to a lack of infrastructure and specialised care equipment. However, unlike with the first wave, the vaccination process with began shortly after. In the middle of the year, the official figures for deaths due to COVID-19 were updated by the government,<sup>3</sup> showing that by June 2021, 180,000 people had died from the virus, meaning that Peru had the highest mortality rate in the world per 100,000 inhabitants. A year later, on 4 January 2022, the third wave began, which had much higher peaks of infections than the two previous waves. However, the number of deaths did not significantly change, due to the increase in the vaccinated population around the country.

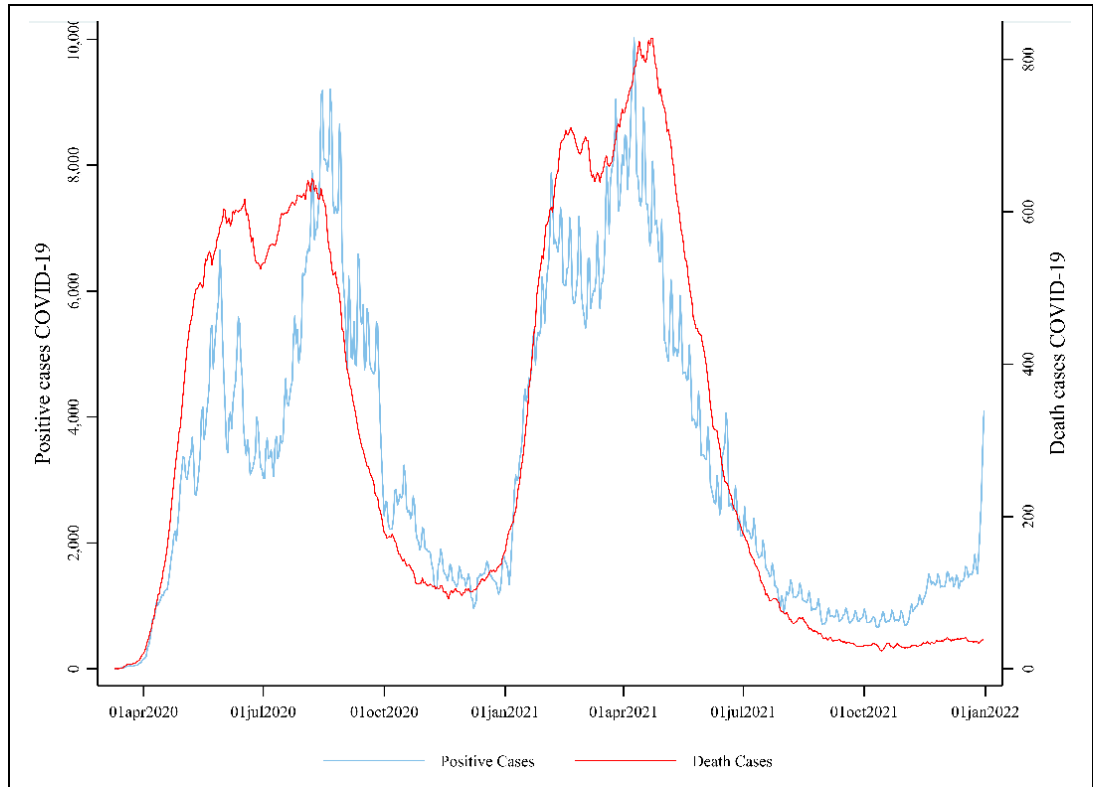
Figure 2 reports the evolution of COVID-19 cases using a five-day moving average of the number of positive cases in the country, obtained from collapsing the total number of cases observed across districts per day. As part of the double-entry graph, we include the moving average of deaths, where we can clearly observe the first (April 2020), second (January 2020) and the very beginning of the third wave (December 2021).

---

2 Table A1 in the Appendix provides detailed information on the emergency laws, dates and geographical coverages of each lockdown.

3 In June 2021, Peru was one of the first countries to correct their official number of deaths and cases based on advice provided by a group of public health experts. For more details, see <https://elperuano.pe/noticia/121843-covid-19-peru-es-el-primer-pais-del-mundo-en-sincerar-sus-cifras-de-fallecidos>.

**Figure 2.** Daily COVID-19 cases and deaths in Peru (until 31 December 2021), five-day rolling average

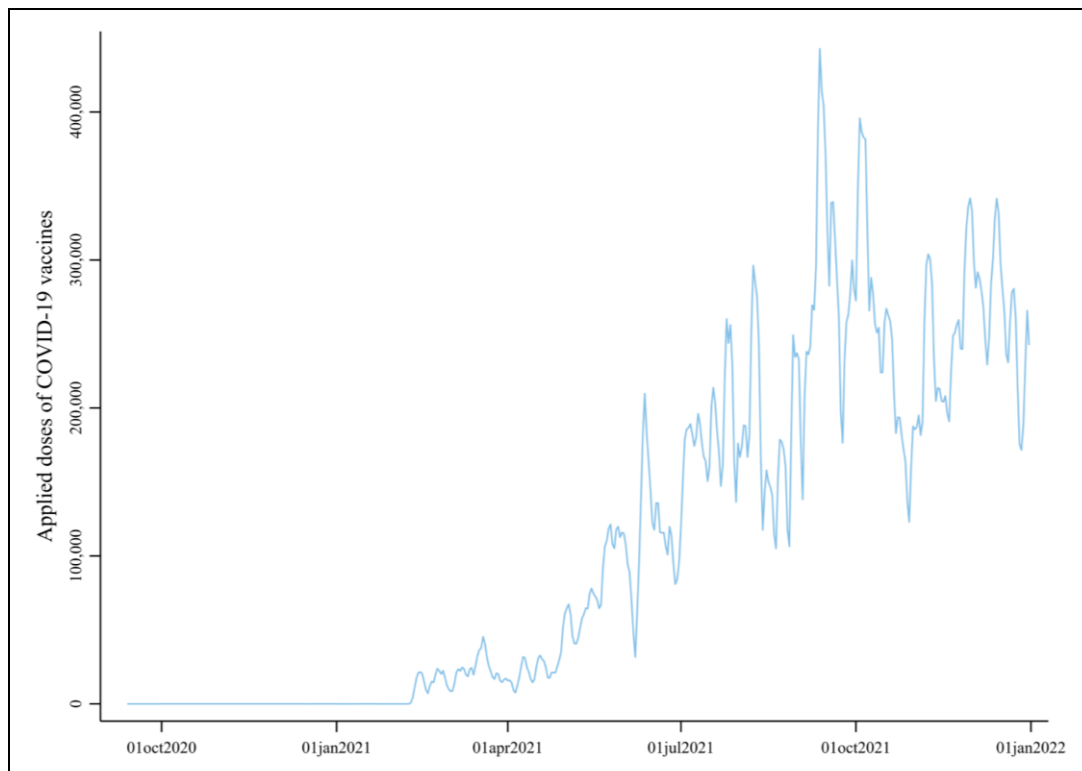


Source: Peru Ministry of Health.

The COVID-19 vaccination campaign started in February 2021, with frontline health care workers prioritised. Subsequently, the first vaccination phase expanded to encompass all health care professionals, firefighters, armed forces members – including the police – and other key groups. By March, elderly people – initially those aged 60 and above – were incorporated into this phase. As additional vaccine doses became available, the vaccination effort extended to include younger age groups. Figure 3 shows the evolution of COVID-19 vaccination doses using a five-day moving average in the country up to the end of 2021, by collapsing the total number of doses applied in each district daily.



**Figure 3.** COVID-19 vaccination doses in Peru (until 31 December 2021), five-day rolling average



Source: Peru Ministry of Health.

## 3. Data sources

### 3.1 Administrative data sources

The following administrative data were matched with the Young Lives data: (1) number of COVID-19 positive cases; (2) number of COVID-19 confirmed deaths; (3) lockdown extension (in days); and (4) number of COVID-19 vaccine doses applied. Data about the number of COVID-19 positive cases<sup>4</sup> and confirmed deaths are provided by the Peru Ministry of Health on its COVID-19 website, which contains information about the number of positive cases and deaths reported on a daily basis in each district since early March.<sup>5</sup> The Ministry of Health also provided the information on vaccinations. The 'lockdown days' variable was calculated using information from the emergency laws introduced by the Peruvian government since mid-March 2020 (see Table A1 in the Appendix). Between July and September 2020,

<sup>4</sup> Positive cases are those according to three types of tests: 'PCR', 'Prueba rápida' and 'Prueba antigenico'.

<sup>5</sup> The dataset can be downloaded at <https://www.datosabiertos.gob.pe/dataset/fallecidos-por-covid-19-ministerio-de-salud-minsa>. For a summary, see [https://covid19.minsa.gob.pe/sala\\_situacional.asp](https://covid19.minsa.gob.pe/sala_situacional.asp).

lockdowns were defined at the region and province levels, extending the lockdown in areas that were deemed to have a high number of COVID cases.

We also have information on the economic activities that were exempted from the lockdown regulations, as they were considered essential activities. In the initial phase, starting on 16 March 2020, only four sectors did not stop working (Table 1). The national government then proposed to restart the economy through the gradual opening and expansion of economic activities. There were four phases in this reactivation process (see Table A2 in the Appendix for more details).

Finally, to rescale variables in per capita terms, we obtained information about the population at the district level from the 2017 Peruvian National Census.<sup>6</sup>

**Table 1.** *Economic sectors and specific activities that were not affected by the first COVID-19 lockdown from 16 March 2020*

<b>Economic activity</b>	<b>Specific activity</b>
Mining and industry	Production, storage, transportation, distribution and sale of fuel
Services and tourism	Health facilities and services
	Services in financial institutions, insurance and pensions
	Media services and call centres
Commerce	Retail establishments of food, beverages, products and essential goods
	Pharmaceutical, medical, optical, orthopaedic and hygienic products
	Fuel sales establishments
Agriculture/farming	Food production and supply

### 3.2 Young Lives phone surveys

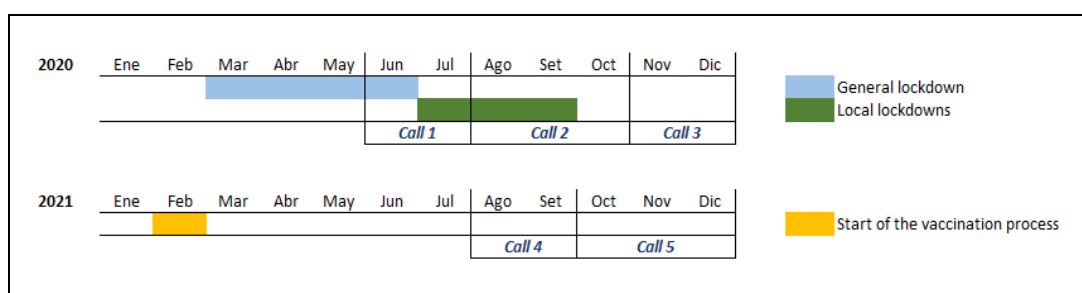
The Young Lives longitudinal study tracks 12,000 young people in Ethiopia, India (in the states of Telangana and Andhra Pradesh), Peru and Vietnam. This technical note focuses on the Peruvian sample. Young Lives has followed two cohorts, the Older Cohort and Younger Cohort, in each of the countries since 2002. In Peru, the Older Cohort consists of approximately 714 children who were born in 1994–95, while the Younger Cohort consists of approximately 2,052 children who were born in 2001–02. Participants have been visited in person on five occasions: 2001, 2006, 2009, 2012 and 2016. Overall, Young Lives has shown relatively low attrition rates compared to other longitudinal studies in similar contexts (Sánchez and Escobal 2020). Young Lives followed a multi-stage sentinel site approach, where 20 sentinel sites were randomly selected from all of the districts in Peru, excluding the wealthiest 5 per cent. While the study was not designed to be nationally representative, Escobal and Flores (2008) found that the Young Lives sample includes households from a vast range of economic backgrounds.

In March 2020, the COVID-19 outbreak interrupted the planned sixth round of in-person data collection, when participants were between 18–19 and 25–26 years old. As a result, the Young Lives team shifted from in-person surveys to phone interviews. A three-part phone

<sup>6</sup> The dataset can be downloaded at <http://censos2017.inei.gob.pe/redatam>.

survey was devised, conducted between August 2020 and December 2021 to assess the short-term impacts of COVID-19 on the health, well-being, transition to the labour market and education trajectories of the young people in the study (Sánchez et al. 2020a; Sánchez et al. 2020b). Figure 4 gives the timeline of the phone surveys as well as the main COVID-19-related events. The Call 1 phone survey took place during the transition period when the national lockdown ended and local lockdowns started. Call 2 took place when the local lockdowns were already established. The sample of participants corresponds to those who could be initially tracked (by phone or in-person) by the end of 2019, before the in-person fieldwork was postponed.

**Figure 4.** *Timeline of COVID-19 pandemic and Young Lives phone surveys*



The Young Lives phone survey covers a wide range of health, well-being, household wealth, labour and educational variables in order to understand the impact of the COVID-19 pandemic on young people’s lives throughout 2020 and 2021.<sup>7</sup> Table 2 presents a range of examples of variables that can be found in the different phone surveys. Moreover, the availability of five previous rounds of surveys in Young Lives makes it possible to understand not only the short-term impact of the pandemic but also how it has shaped life trajectories. During the phone surveys the Young Lives team registered the geographical location of participants at the region/province/district levels – which is considered personal data – and asked for consent to use this information to understand the impact of the pandemic on young people’s lives at the population level, and for tracking purposes for future Young Lives rounds.

<sup>7</sup> For more details on the phone survey questionnaires, see <https://www.younglives.org.uk/2020-phone-survey>.

**Table 2.** *Examples of COVID-19 related survey questions asked in the COVID-19 Young Lives phone survey*

Dimension	Examples of COVID-19 related survey questions	Calls
Knowledge and information about COVID-19	From whom did you receive information about social distancing or self-isolation? Do you know the signs and symptoms of COVID-19?	Call 1
Effect on health	Has any member of your household has been infected or do you think you think they have been infected with coronavirus? How is that person now?	Calls 1, 2, 4 and 5
Preventive measures and risk perception	Have you done anything to avoid getting the COVID-19 infection? For example, change your behaviour, eat certain food or drink, or avoid certain things? In the last 7 days, on how many of these days did you stay at home all day, without going out at all?	Calls 1, 2 and 5
Vaccinations and attitude towards vaccines	Has someone in your household been vaccinated against COVID-19? To what extent do you agree or disagree with the statement: 'If a vaccine for COVID-19 was available, I would get it'?	Calls 4 and 5
Effects on labour market	Has someone in your household lost or considerably reduced their source of income from their own business/cultivation due to the current pandemic? Who? Have you started working remotely from home (teleworking) during lockdown? What is your main economic activity?	Calls 1, 3 and 5
Effects on education	Are you currently studying full time? Did someone in the household have their studies interrupted at school/university/institute due to the pandemic? Who?	Calls 1, 3 and 5
Effects on caring responsibilities	Who was the main household member who took who took care of individuals requiring special care within the household during lockdown?	Call 1
Effects on food security	Did you or others in your household worry about not having enough food to eat due to lack of money or other resources?	Calls 2, 3 and 5
Government support	Did your household receive government support to cope with the crisis? What type of support?	Call 1
Self-reported anthropometric measures	What is your current weight? What is your current height?	Call 5

## 4. Estimating COVID-19 related variables

### 4.1 Confirmed cases and deaths

We used information provided by the Peru Ministry of Health in the National Open Data Platform,<sup>8</sup> which has been updated weekly since the start of the pandemic. Based on this, a time-series panel dataset was created, which contains the number of cases reported each

<sup>8</sup> For more details, see <https://www.datosabiertos.gob.pe/dataset/casos-positivos-por-covid-19-ministerio-de-salud-minsa>

day in each district since early March 2020.<sup>9</sup> Considering the focus on the Young Lives phone surveys, the following variables were created (Table 3).

**Table 3.** *Lockdown cases for interrupted days and uninterrupted days*

Type of variables	Description
Cumulative count normalised by the population size	<ul style="list-style-type: none"> <li>Total number of COVID-19 cases per 1,000 inhabitants, by district since start of the pandemic until interview date for each call.</li> <li>Total number of COVID-19 deaths per 1,000 inhabitants, by district since start of the pandemic until interview date for each call.</li> </ul>
Difference in count between the end date of national lockdown and Calls 2 and 3	<ul style="list-style-type: none"> <li>Total number of COVID-19 deaths per 1,000 inhabitants, by district, between the date of the end of the national lockdown and the interview date for Call 2 and Call 3.</li> </ul>

The variables cases per 1,000 inhabitants and deaths per 1,000 inhabitants are created with the number of cases and deaths multiplied by 1,000 and divided by the population of the district according to the 2017 Census, as follows:

$$CCasesx1000inhabitant_{n,m} = \frac{\text{Total number of COVID cases per district}_n * 1000}{\text{Population of the district}_m} \dots \quad (1)$$

$$CDeathsx1000inhabitant_{n,m} = \frac{\text{Total number of COVID deaths per district}_n * 1000}{\text{Population of the district}_m} \dots \quad (2)$$

To construct the variable capturing the difference in the total number of COVID-19 deaths in the district (between the date of the end of the national lockdown and the interview date for Calls 2 and 3) expressed per 1,000 inhabitants, two input variables are created. The first calculates the accumulated COVID-19 deaths until 30 June 2020 (the date that the national lockdown ended in Peru), labelled as 'X'; the second calculates the accumulated COVID-19 deaths until the beginning of the second (August 2020) and third (November 2020) calls, labelled as 'Y'. For the creation of the variables of interest, the following formula is used:

$$diff\_30jun\_Ylcall\_deathx1000inhabitant_{n,m} = \frac{(Y-X)_n * 1000}{\text{Population of the district}_m} \dots \quad (3)$$

## 4.2 Lockdown variables

From the emergency laws introduced by the government it is possible to identify for how many days a given province remained in lockdown at a given moment in time. When Call 2 and Call 3 took place in Peru, all participants had already endured at least 107 days of national lockdown from 16 March to 30 June 2020. A participant might also have spent additional days in lockdown if he/she lived in a province that was selected for local (targeted) lockdowns in July, August and September 2020 (Table 4). Therefore, the variation in the number of lockdown days comes exclusively from differences in local lockdowns in July, August and September.

<sup>9</sup> The code for each district is called UBIGEO, which represents the district administrative unit, the level below the province and region (department) in Peru.

**Table 4.** *Lockdowns in Peru by month*

	Extension started	Extension ended	Duration (days)	Number of provinces in lockdown
Lockdown (March to June)	16 March		107	196
July extension <sup>10</sup>	1 July	31 July	31	66
August extension <sup>11</sup>	1 August	31 August	31	62
September extension <sup>12</sup>	1 September	30 September	30	79

Table 5 identifies the different permutations of cases.<sup>13</sup> For instance, ‘No/No/Yes’ means a certain province was in lockdown in July and August, but not in September. The various permutations are included in the dataset (see variables ‘lock\_1’, ‘lock\_2’, ... ‘lock\_8’). We also created a dummy variable to establish whether a given province was in lockdown for more than 107 days without interruption – this corresponds to cases 1, 4 and 7 in Table 5, where the lockdown continued during July (only), July and August (only), or July, August and September, respectively.

**Table 5.** *Lockdown types of cases observed*

Cases lockdown	Month in lockdown (2020)				Total interrupted days of lockdown	Total uninterrupted days of lockdown	Name of variable in STATA
	Mid-March – June	July	August	September			
Case 1	Yes	Yes	No	No	138	138	lock_1
Case 2	Yes	No	Yes	No	138	107	lock_2
Case 3	Yes	No	No	Yes	137	107	lock_3
Case 4	Yes	Yes	Yes	No	169	169	lock_4
Case 5	Yes	No	Yes	Yes	168	107	lock_5
Case 6	Yes	Yes	No	Yes	168	138	lock_6
Case 7	Yes	Yes	Yes	Yes	199	199	lock_7
Case 8	Yes	No	No	No	107	107	lock_8

Based on this information, we created the following variables:

- Number of lockdown days (‘dint\_lockdown’)
- Lockdown – uninterrupted days (‘lock\_ABC’)
- Number of lockdown days – uninterrupted days only (‘dint\_cdays’).

The variable ‘dint\_lockdown’ captures the number of days a given province was in lockdown at a given moment in time. Variable ‘lock\_ABC’ is a dummy that takes the value of 1 if the province corresponds to cases 1, 4 or 7, and 0 otherwise. In other words, it takes the value of

10 Decreto Supremo N° 116-2020-PCM, introduced in July 2020.

11 Decreto Supremo N° 135-2020-PCM, introduced in August 2020.

12 Decreto Supremo N° 146-2020-PCM, introduced in September 2020.

13 These permutations could vary, depending on how many months of focused lockdown each country had. In Peru, there were only three months of focused lockdown, generating eight possible permutations.

1 if the province remained in lockdown beyond June. Finally, 'dint\_cdays' is the number of uninterrupted lockdown days.

### 4.3 COVID-19 vaccines

The COVID-19 vaccination campaign in Peru started in February 2021, with frontline health care workers initially prioritised, before the first vaccination phase was expanded to encompass all health care professionals, firefighters, armed forces members – including the police – and other key groups. By March, people over 60 years old were incorporated into this phase. As additional doses became available, the vaccination effort was extended to younger age categories. Three phases guided the vaccination schedule for Peru; however, the implementation period was subject to the availability of vaccines. Table 6 details the vaccination schedule and target population groups followed by the Ministry of Health.<sup>14</sup>

**Table 6.** *COVID-19 vaccine phases and target population*

COVID-19 phases	Target population
Phase I: Protecting the integrity of the healthcare system and continuity of basic services	<ul style="list-style-type: none"> <li>All individuals who provide services, regardless of the nature of their employment or contractual relationship, in various health care settings.</li> <li>Armed forces and police personnel.</li> <li>Firefighters, Red Cross personnel.</li> <li>Security personnel, municipal security officers, brigade members, and cleaning staff.</li> <li>Health care students.</li> <li>Members of electoral boards.</li> </ul>
Phase II: Reducing severe morbidity and mortality in high-risk populations	<ul style="list-style-type: none"> <li>Adults aged 60 and older.</li> <li>People with prioritised comorbidities.</li> <li>Population in native or indigenous communities.</li> <li>Staff of the National Penitentiary Institute (INPE), incarcerated individuals.</li> </ul>
Phase III: Reducing infection transmission and achieving herd immunity	<ul style="list-style-type: none"> <li>People aged 18 to 59 years.</li> </ul>

To construct the vaccination dose variables, we used information from the Open Data Platform for Peru.<sup>15</sup> For each day, there are data on the number of doses applied in each district in Peru. With these statistics, we constructed a cumulative number of doses applied in each district where the participant lived during Calls 4 or 5 and established a cutoff point according to both interview dates (August and November 2021), expressed in per capita terms. The following variables are created:

- COVID-19 doses applied per capita by the district until Young Lives Call 4 and 5 ('doses\_pc\_4call' and 'doses\_pc\_5call')

The calculation is as follows:

$$CVaccinesYLCall_{n,m} = \frac{\text{Total number of COVID doses applied per district}_n}{\text{Population of the district}_m} \dots \quad (4)$$

<sup>14</sup> Law Resolución Magisterial N° 161-2021/MINSA, introduced in February 2021.

<sup>15</sup> The data were obtained from <https://www.minsa.gob.pe/reunis/data/vacunas-covid19.asp>.

When calculating this variable is worth considering that the input variable takes into account all doses applied in the corresponding district. As such, more than one dose could be applied to the same person since some health care workers were already entitled to receive a second dose of COVID-19 vaccine by that time. Also, because of how the vaccination scheme was established in Peru, it is possible that those districts with a higher number of health centres also have a higher number of vaccines applied by construction.

## 5. Matched dataset

Following the methodology described in the previous section, the matched dataset contains variables related to total accumulated COVID-19 cases, deaths, vaccination doses applied and lockdown duration for all Young Lives participants that were part of the phone surveys. Table 7 reports the key variables created, with the name of the variable in STATA in the first column and description in the second column.

**Table 7** *Data dictionary for archived dataset*

Variable	Description	Type
childid	ID variable for each participant	String
yc	Indicates if the participant belongs to the Younger Cohort (yc=1) or the Older Cohort (yc=0).	Numeric, dichotomic
cases_per_1000_30jun	Cumulative number of COVID-19 cases per 1,000 inhabitants according to participant's (residing) district until 30 June 2020	Numeric, continuous
cases_per_1000_2call	Cumulative number of COVID-19 cases per 1,000 inhabitants according to participant's (residing) district until beginning of Call 2 data collection (1 August 2020)	Numeric, continuous
cases_per_1000_3call	Cumulative number of COVID-19 cases per 1,000 inhabitants according to participant's (residing) district until beginning of Call 3 data collection (1 November 2020)	Numeric, continuous
cases_per_1000_4call	Cumulative number of COVID-19 cases per 1,000 inhabitants according to participant's (residing) district until beginning of Call 4 data collection (1 August 2021)	Numeric, continuous
cases_per_1000_5call	Cumulative number of COVID-19 cases per 1,000 inhabitants according to participant's (residing) district until beginning of Call 5 data collection (1 November 2021)	Numeric, continuous
deaths_per_1000_30jun	Cumulative number of COVID-19 deaths per 1,000 inhabitants according to participant's (residing) district until 30 June 2020	Numeric, continuous
deaths_per_1000_2call	Cumulative number of COVID-19 deaths per 1,000 inhabitants according to participant's (residing) district until beginning of Call 2 data collection (1 August 2020)	Numeric, continuous
deaths_per_1000_3call	Cumulative number of COVID-19 deaths per 1,000 inhabitants according to participant's (residing) district until beginning of Call 3 data collection (1 November 2020)	Numeric, continuous
deaths_per_1000_4call	Cumulative number of COVID-19 deaths per 1,000 inhabitants according to participant's (residing) district until beginning of Call 4 data collection (1 August 2021)	Numeric, continuous
deaths_per_1000_5call	Cumulative number of COVID-19 deaths per 1,000 inhabitants according to participant's (residing) district until beginning of Call 5 data collection (1 November 2021)	Numeric, continuous
doses_pc_4call	Per capita number of COVID-19 doses (vaccine jabs) applied each day in participant's (residing) district until beginning of Call 4 data collection (1 August 2021)	Numeric, continuous



Variable	Description	Type
doses_pc_5call	Per capita number of COVID-19 doses (vaccine jabs) applied each day in participant's (residing) district until beginning of Call 5 data collection (1 November 2021)	Numeric, continuous
dint_lockdown	Number of lockdown days for a given province at a corresponding moment in time	Numeric, continuous
dint_cdays	Lockdown extension measured in uninterrupted days for a given province at a corresponding moment in time	Numeric, continuous
lock_1	Participant was living in a province under lockdown restrictions in July 2020, but not in August 2020 nor September 2020	Numeric, dichotomic
lock_2	Participant was living in a province under lockdown restrictions in August 2020, but not July 2020	Numeric, dichotomic
lock_3	Participant was living in a province under lockdown restrictions in September 2020, but not July 2020 nor August 2020	Numeric, dichotomic
lock_4	Participant was living in a province under lockdown restrictions in July and August 2020, but not in September 2020	Numeric, dichotomic
lock_5	Participant was living in a province under lockdown restrictions in August and September 2020, but not in July 2020	Numeric, dichotomic
lock_6	Participant was living in a province under lockdown restrictions in July and September 2020, but not in August 2020	Numeric, dichotomic
lock_7	Participant was living in a province under lockdown restrictions in July, August and September 2020	Numeric, dichotomic
lock_8	Participant was living in a province that was not under any lockdown restrictions in July, August and September 2020	Numeric, dichotomic
diff_30jun_2call_death	Difference in the cumulative number of COVID-19 deaths per 1,000 inhabitants according to participant's (residing) district until beginning of Call 2 data collection (1 August 2020) and 30 June 2020	Numeric, continuous
diff_30jun_3call_death	Difference in the cumulative number of COVID-19 deaths per 1,000 inhabitants according to participant's (residing) district until beginning of Call 3 data collection (1 November 2020) and 30 June 2020	Numeric, continuous
migrant_c3c5	Participant lived in a different district in Call 5 compared to Call 3	Numeric, dichotomic

In the Appendix, we present analysis using the archived dataset of an example of the distribution of the COVID-19 related variables. The histogram of the variable 'cases\_per\_1000' is reported in Figures A1, A2, A3 and A4 for Calls 2 to 5; the histogram of the variable 'dint\_lockdown' is reported in Figures A5 and A6 for Calls 2 and 3; and the analogous information for the variables 'percap\_doses\_4call' and 'percap\_doses\_5call' is reported in Figures A7 and A8.

## 6. Conclusions

This technical notes outlines the COVID-19 matched dataset focused on the Young Lives cohorts in Peru that was created by the Young Lives team. The matched dataset combines COVID-19 administrative data with the extensive and thorough longitudinal dataset from the Young Lives study in Peru. We document the steps taken to create a set of variables that measure the number of COVID-19 cases, deaths and vaccine doses applied in the districts where participants reside, along with the duration of lockdown periods they endured. To achieve this, we collected administrative data and integrated it with the Young Lives datasets from Call 2 (August to October 2020), Call 3 (November to December 2020), Call 4 (August to September 2021) and Call 5 (November to December 2021). This data-matching process improves the versatility of the already extensive Young Lives demographic dataset, enabling the assessment of how the COVID-19 pandemic had an impact on children and young adults' development outcomes, such as their health, well-being, transition to the labour market and education trajectories.

## References

- Araujo-Castillo, R. (2022) 'Dos años de pandemia, una batalla que aún no termina', *Acta Médica Peruana* 39.1: 3–6.
- Escobal, J., and E. Flores (2008), 'An Assessment of the Young Lives Sampling Approach in Peru'. Young Lives Technical Note N 3.
- Hale, T., N. Angrist, R. Goldszmidt, B. Kira, A. Petherick, T. Phillips, S. Webster, E. Cameron-Blake, L. Hallas, S. Majumdar, and H. Tatlow (2021) 'A Global Panel Database of Pandemic Policies (Oxford COVID-19 Government Response Tracker)', *Nature Human Behaviour* 5.4: 529–538.
- Sánchez, A., and J. Escobal (2020) 'Survey Attrition After 15 years of Tracking Children in Four Developing Countries: The Young Lives Study', *Review of Development Economics* 24.4: 1196–1216.
- Sánchez, A., S. Cueto, M. Penny, and A. Hidalgo (2020a) 'Listening to Young Lives at Work in Peru: First Call: Young Lives COVID-19 Phone Survey Headlines Report', [https://www.younglives.org.uk/sites/default/files/2023-11/202009\\_Headlines\\_FirstPhoneSurvey\\_Peru\\_Accessible\\_2311.pdf](https://www.younglives.org.uk/sites/default/files/2023-11/202009_Headlines_FirstPhoneSurvey_Peru_Accessible_2311.pdf) (accessed 9 October 2023).
- Sánchez, A., S. Cueto, M. Penny, J. Lopez, and N. Lago (2020b) 'Listening to Young Lives at Work in Peru: Second Call. Young Lives COVID-19 Phone Survey Headlines Report', <https://www.younglives.org.uk/sites/default/files/migrated/YOL-Peru-Headlines-PhoneSurvey2-Nov20-2%2023rd%20.pdf> (accessed 9 October 2023).
- WHO (World Health Organization) (2023) 'WHO COVID-19 Dashboard', <https://covid19.who.int> (accessed 9 October 2023).

# Appendix

**Table A1.** *Emergency laws on COVID-19 introduced by the Peruvian government in 2020*

Law	Period covered	Geographical coverage
044-2020-PCM	Mandatory social isolation for 15 days from 16 March	National level
051-2020-PCM	Extension for 13 days from 31 March	National level
064-2020-PCM	Extension for 14 days from 13 April	National level
075-2020-PCM	Extension for 14 days from 27 April	National level
083-2020-PCM	Extension for 14 days from 11 May	National level
094-2020-PCM	Extension from 25 May to 30 June	National level
116-2020-PCM	During July, mandatory social isolation in specific regions	Regions: Arequipa, Ica, Junín, Huánuco, San Martín, Madre de Dios, Áncash
A.146-2020-PCM	During August, mandatory social isolation in specific regions and provinces	Regions: Arequipa, Ica, Junín, Huánuco, San Martín Provinces: Tambopata (Madre de Dios), Santa, Casma, Huaraz, Huarmey (Ancash), Mariscal Nieto, Ilo (Moquegua), Tacna (Tacna), Cusco, La Convención (Cusco), San Román, Puno (Puno), Huancavelica (Huancavelica), Cajamarca, Jaén, San Ignacio (Cajamarca), Bagua, Condorcanqui, Utcubamba (Amazonas), Abancay, Andahuaylas (Apurímac).
146-2020-PCM	During September, mandatory social isolation in specific regions and provinces.	Regions: Ica, Junín, Huánuco y San Martín, Provinces: Tambopata (Madre de Dios); Santa, Casma, Huaraz (Ancash); Mariscal Nieto, Ilo (Moquegua); Tacna (Tacna); Cusco, La Convención (Cusco); San Román, Puno (Puno); Huancavelica (Huancavelica); Cajamarca, Jaén, San Ignacio (Cajamarca); Bagua, Condorcanqui, Utcubamba (Amazonas); Abancay, Andahuaylas (Apurímac); Camaná, Islay, Cailloma, Castilla (Arequipa); Huamanga, Huanta, Lucanas, Parinacochas (Ayacucho); Huánuco, Leoncio Prado, Puerto Inca, Humalies (Huánuco); Ica, Pisco, Nasca, Palpa (Ica); Huancayo, Satipo, Chanchamayo (Junín); Trujillo, Pacasmayo, Chepén, Ascope, Sánchez Carrión, Virú (La Libertad); Barranca, Cañete, Huaura, Huaral (Lima); Pasco, Oxapampa (Pasco).

**Table A2.** *Economic sectors and specific activities that were exempt from lockdown regulations*

Economic activity	Specific activities	General lockdown			Focused lockdown	
		PHASE 0	PHASE 1	PHASE 2	PHASE 3	PHASE 4
		(16 March)	(4 May)	(22 June)	(1 July)	(1 October)
<b>Mining and industry</b>	Production, storage, transportation, distribution and sale of fuel	x	x	x	x	x
	Exploitation, benefit, storage, transport and closure of large mines, and projects in construction of national interest and hydrocarbons		x	x	x	x
	Inputs for agricultural activity		x	x	x	x
	Industrial fishing (indirect human consumption)		x	x	x	x
	Temporary production: purchase orders (exports) expired and about to expire		x	x	x	x
	Glass, forestry (timber or non-timber), paper and cardboard, plastics and ice industries. Expansion of textiles and clothing, machinery and equipment		x	x	x	x
	Metalworking industry		x	x	x	x
	Basic chemical substances and fertiliser and complementary services to agriculture (for essential activities)		x	x	x	x
	Exploration of large and medium mining stratum			x	x	x
	Exploitation, beneficiation, storage, transportation and closure of mines			x	x	x
	All activities corresponding to the energy and mines sector not included in Phases 1 and 2 of resumption of activities.				x	x
<b>Construction</b>	Projects of the National Infrastructure Plan for competitiveness (PNIC)		x	x	x	x
	Authority for Reconstruction with Changes (ARCC) projects		x	x	x	x
	Fifty-six projects in the transport and communications sector		x	x	x	x
	Thirty-six sanitation works		x	x	x	x
	Agricultural infrastructure activities (including irrigation, maintenance, drainage rehabilitation)		x	x	x	x
	Prioritised real estate projects (excavation phase, structures and finishes, and housing in rural areas)		x	x	x	x
	Agricultural products (rental/sale of machinery)		x	x	x	x
	Investments for optimisation, marginal expansion, rehabilitation and replacement (IOARR), access to water and sewerage in police stations, hospitals and schools		x	x	x	x
	Industries and services related to construction		x	x	x	x
	Public investment projects, private investment projects, public-private associations, active projects, IOARR and the PIRCC			x	x	x
	Projects in general.				x	x
<b>Services and tourism</b>	Health facilities and services	x	x	x	x	x
	Services in financial institutions, insurance and pensions	x	x	x	x	x
	Media services and call centres	x	x	x	x	x
	Transportation service reduced to 50% by land and river (freight and merchandise transport)		x	x	x	x
	Restaurants and the like authorised for home delivery (with the establishment's own logistics and security protocol and local pick-up)		x	x	x	x
	Categorised hotels and tourist transport for essential activities		x	x	x	x
	Telecommunications-related services		x	x	x	x
	Complementary services to agriculture		x	x	x	x
	Services provided to companies (IT support and professional services, knowledge service exports)		x	x	x	x
	Notary services		x	x	x	x
	Recycling services		x	x	x	x
	Equipment maintenance services related to buildings and homes (pumps, hot springs, elevators, plumbing, electrician, carpentry, etc.)		x	x	x	x
	Storage services for fertilisers and agricultural raw materials, plastic articles, glass, paper, cardboard, wood, ice for activities in general		x	x	x	x
	Services professionals, scientists and technical staff			x	x	x
	Accommodation (apart from hotels)			x	x	x
	Protection and security services			x	x	x

DATA MATCHING: CONSTRUCTION OF COVID-19-RELATED VARIABLES FOR YOUNG LIVES PERU

Economic activity	Specific activities	General lockdown			Focused lockdown	
		PHASE 0	PHASE 1	PHASE 2	PHASE 3	PHASE 4
		(16 March)	(4 May)	(22 June)	(1 July)	(1 October)
	Car rental			x	x	x
	Rental of machinery and equipment and tangible goods			x	x	x
	Infrastructure in transport and telecommunications			x	x	x
	License plate issuance			x	x	x
	Routine and periodic maintenance, improvement and conservation of national, departmental and local roads			x	x	x
	Water supply: sewerage, waste management and sanitation			x	x	x
	Transport services			x	x	x
	Restaurants and related services, except bars with a capacity of 40%				x	x
	Issuance of certificates				x	x
	Marketing services				x	x
	Rail passenger transport services, except urban				x	x
	Ground transportation services for people at the national, regional and provincial levels				x	x
	Passenger and cargo transportation services by inland waterways (rivers and lakes)				x	x
	Air transport (only domestic flights)				x	x
	Services related to air transport				x	x
	Special tourist air transport services				x	x
	General aviation activities				x	x
	Moving, household goods or related activities				x	x
	General accommodation services				x	x
	Book publishing services, directories and other publishing services				x	x
	Production, post-production and distribution services for films, videotapes and tv programmes				x	x
	Sound recording and music editing services				x	x
	Accounting, auditing and business management consulting services				x	x
	Technical analysis, scientific research and development services				x	x
	Market research services and public opinion polls				x	x
	Specialised services of design, photography and other professional, scientific, technical activities				x	x
	Rental and leasing services for recreational and sports equipment, video tapes, records, other personal property and household items				x	x
	Leasing services for intellectual property and similar products				x	x
	Services related to employment agencies				x	x
	Travel agency and tour operator services				x	x
	Cleaning services, building support and garden maintenance				x	x
	Photocopying services, document preparation and other specialised office support activities				x	x
	Specialised support services for administrative offices and other business activities				x	x
	Other art, entertainment and leisure services (in health destinations and with a capacity of 50%)				x	x
	Repair of computers and communication equipment				x	x
	Repair of consumer electrical appliances, furniture and other personal property				x	x
	Restaurants and related services, except bars with a capacity of 50%					x
	Passenger maritime transport services: maritime transport and passenger cabotage (with a capacity of 50% and not including tourist transport)					x
	Passenger and cargo transport services by inland waterways by rivers and lakes: passenger transport by inland waterways (with capacity at 50%)					x
	Air transport: international flights to health destinations from 5 October.					x
	Travel agency and tour operator services (other reservation services and related activities)					x
	Other art, entertainment and leisure services (in health destinations and with a capacity of 60%)					x
	Libraries and archives activities, including university libraries.					x

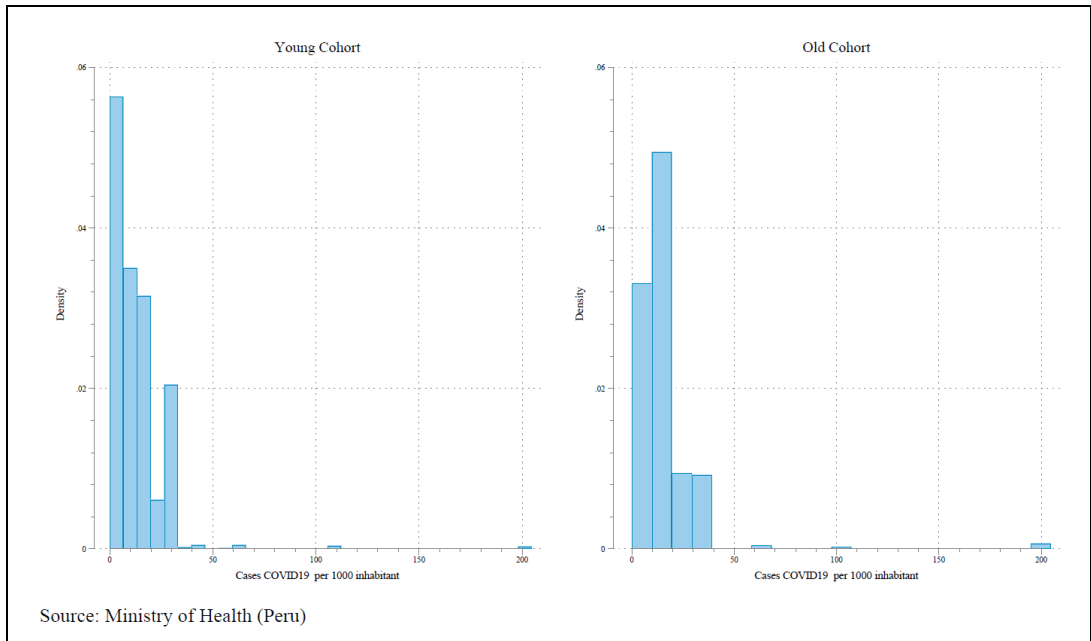
DATA MATCHING: CONSTRUCTION OF COVID-19-RELATED VARIABLES FOR YOUNG LIVES PERU

Economic activity	Specific activities	General lockdown			Focused lockdown	
		PHASE 0	PHASE 1	PHASE 2	PHASE 3	PHASE 4
		(16 March)	(4 May)	(22 June)	(1 July)	(1 October)
	Activities and management of museums, pre-Hispanic archaeological monuments, historical places and buildings, cultural centres (not including screening of films, plays and shows) and galleries					x
	Botanical and zoological gardens and national parks activities, including natural areas					x
	Theme park activities					x
	Other art, entertainment and leisure services (with a capacity of 50%). Activities of sports clubs and associations (individual or couples activities carried out in the open air). Sport fishing and nautical sports, as well as their management of reservations and support activities. Mountain guide activities.					x
<b>Commerce</b>	Retail establishments of food, beverages, products and essential goods	x	x	x	x	x
	Pharmaceutical, medical, optical, orthopaedic and hygienic products	x	x	x	x	x
	Fuel sales establishments	x	x	x	x	x
	Marketing of agricultural products		x	x	x	x
	Electronic commerce of household goods and the like		x	x	x	x
	Sale, maintenance and repair of motor vehicles and motorcycles			x	x	x
	Sale, maintenance, and repair of non-motorised vehicles, sale of spare parts and supplies			x	x	x
	Shops in general: with a capacity of 50%				x	x
	Shops in general: with a capacity of 60%.					x
<b>Agriculture/ farming</b>	Food production and supply	x	x	x	x	x
	Enabling titles and administrative acts			x	x	x
	Environmental impact studies and environmental management instruments			x	x	x
	Land titling and rural cadastre			x	x	x
	Implementation of business plans			x	x	x
	All activities corresponding to the agricultural sector not included in Phases 1 and 2 of resumption of activities.				x	x
<b>Manufacture</b>	Preparation of prepared food for animals			x	x	x
	Malted and malted beverage making			x	x	x
	Distillation, rectification and mixing of alcoholic beverages, wine making, tobacco product making			x	x	x
	Manufacture of footwear			x	x	x
	Printing activities and printing-related service activities			x	x	x
	Repair of transport equipment, except motor vehicles			x	x	x
	Manufacture of medical and dental instruments and supplies. Repair and maintenance of other equipment			x	x	x
	Manufacture of weapons and ammunition.				x	x
	Manufacture of aircraft, spacecraft and related machinery				x	x
Manufacture of military combat vehicles				x	x	

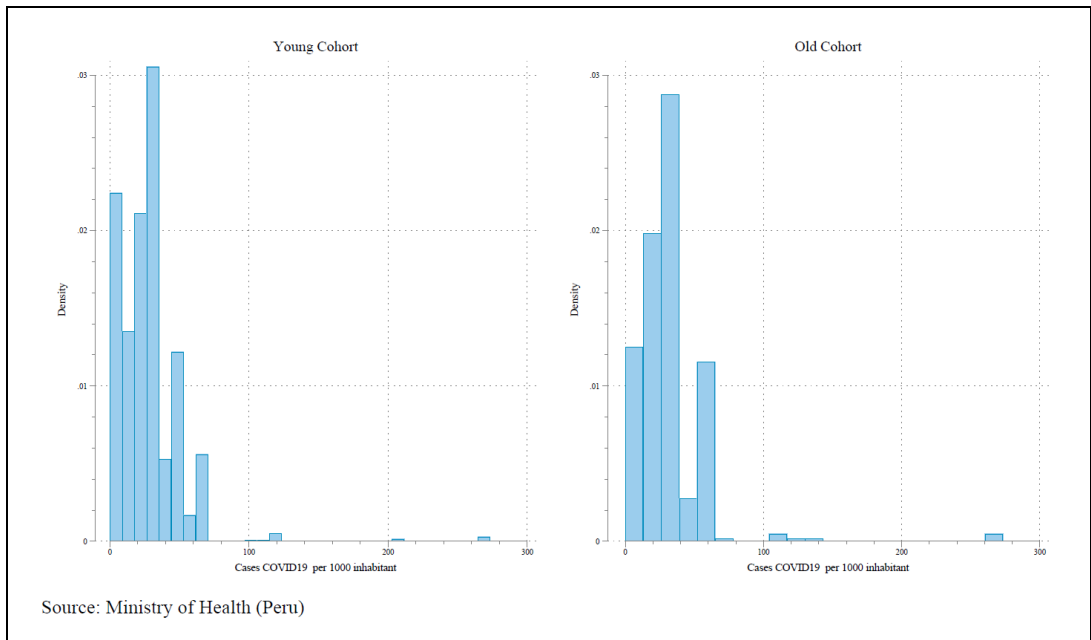
Notes: Table shows the services and economic activities that were gradually reactivated during lockdown. X means that the economic activity was in operation or functioning normally, or was exempt from the lockdown regulations.

### COVID-19 cases distribution by cohort and Young Lives survey call

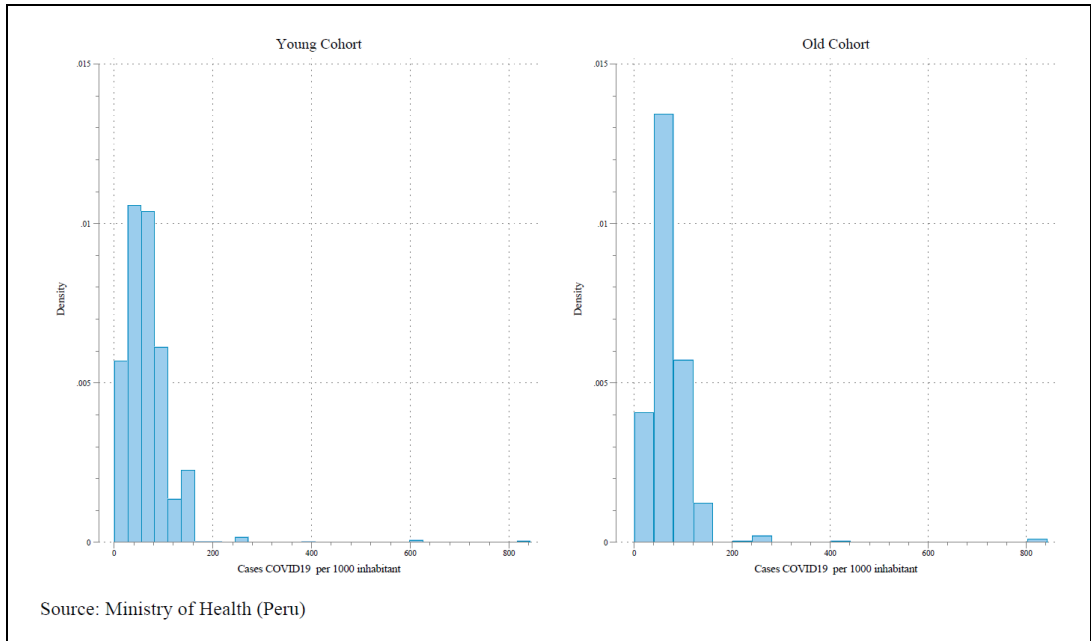
**Figure A1.** *Young Lives participants and COVID-19 cases by cohort (Call 2)*



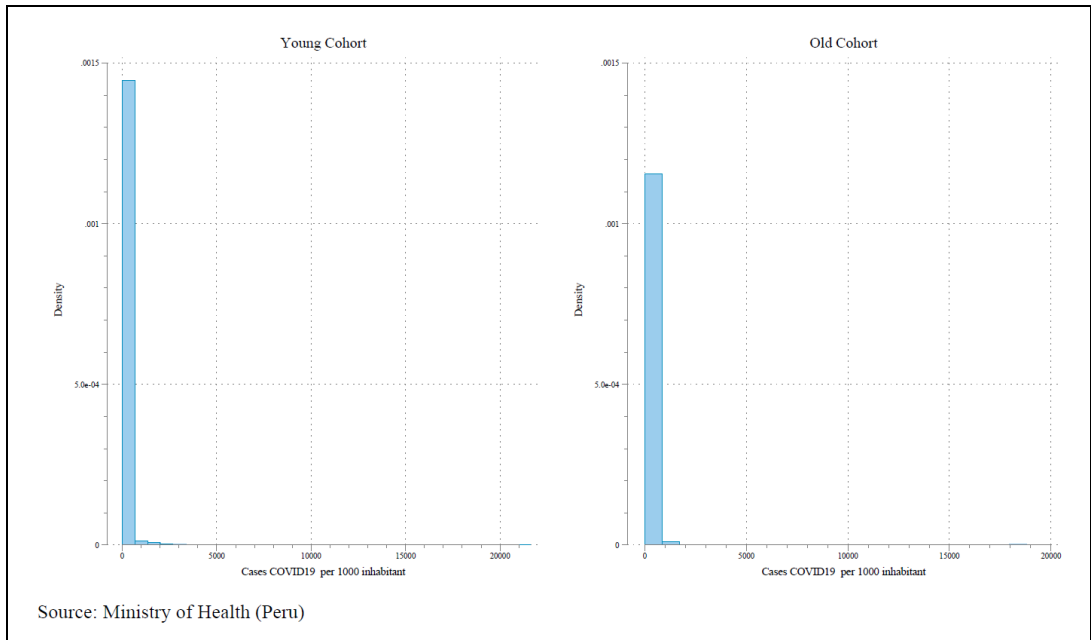
**Figure A2.** *Young Lives participants and COVID-19 cases by cohort (Call 3)*



**Figure A3.** *Young Lives participants and COVID-19 cases by cohort (Call 4)*



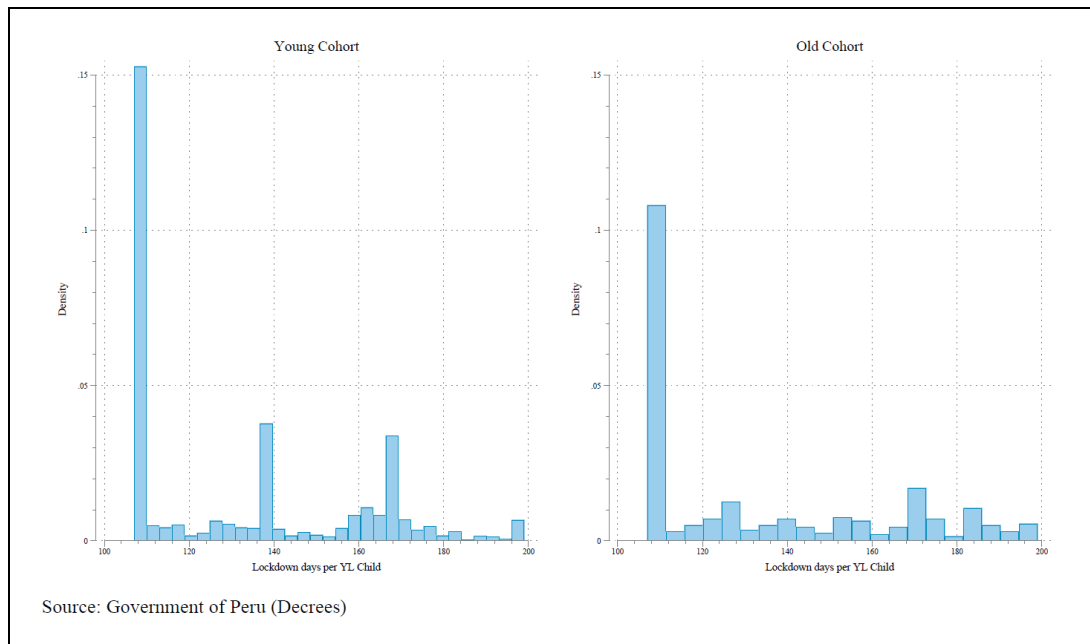
**Figure A4.** *Young Lives participants and COVID-19 cases by cohort (Call 5)*



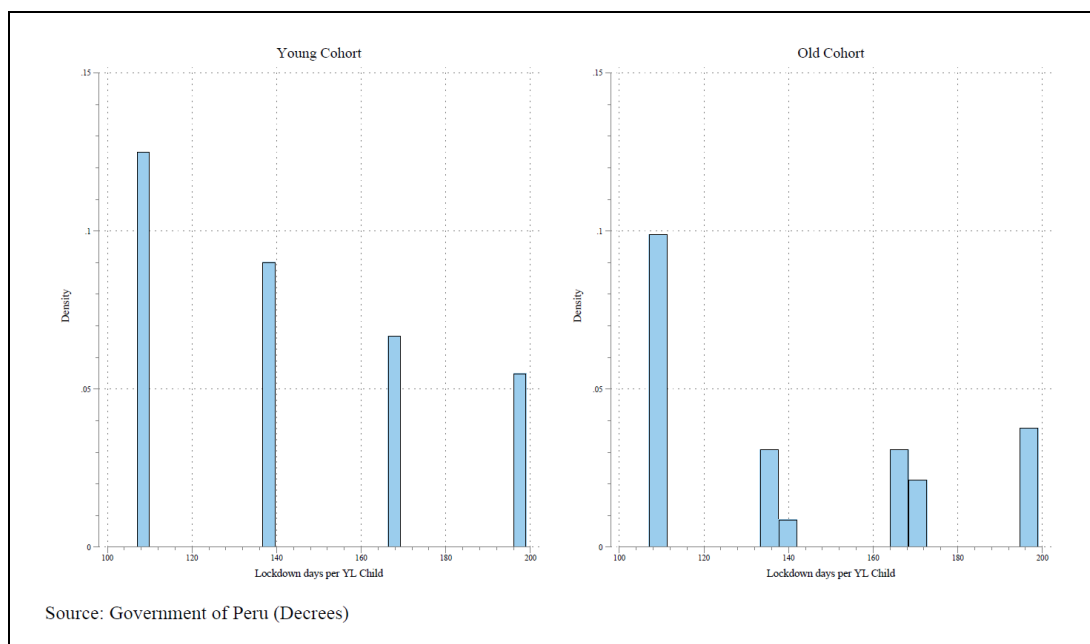


### COVID-19 lockdown duration distribution by cohort and Young Lives survey call

**Figure A5.** *Young Lives participants and number of COVID-19 lockdown days by cohort (Call 2)*

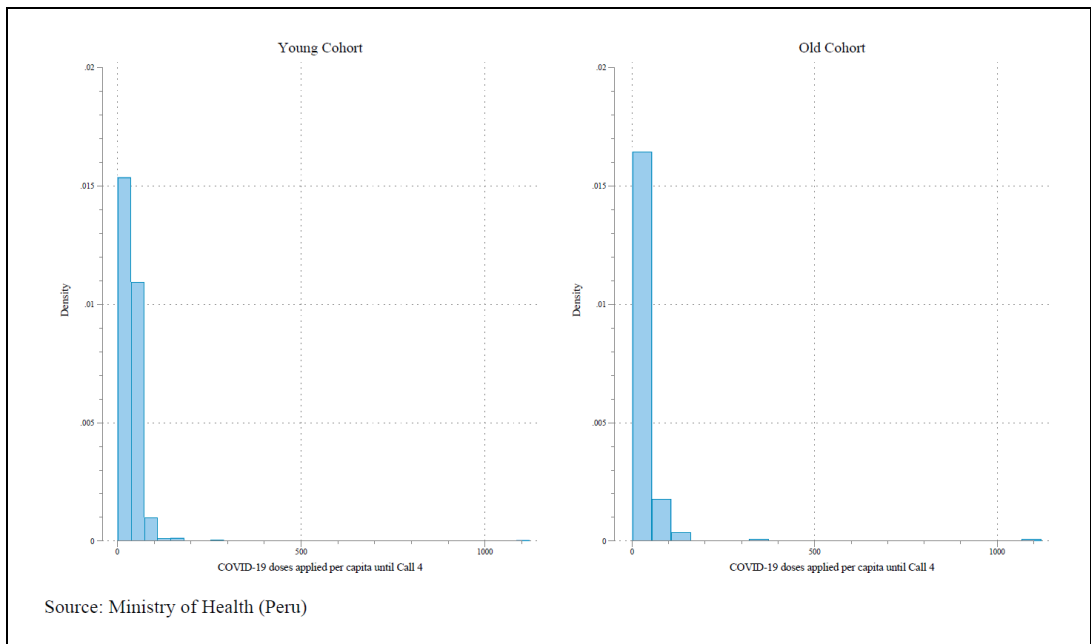


**Figure A6.** *Young Lives participants and number of COVID-19 lockdown days by cohort (Call 3)*

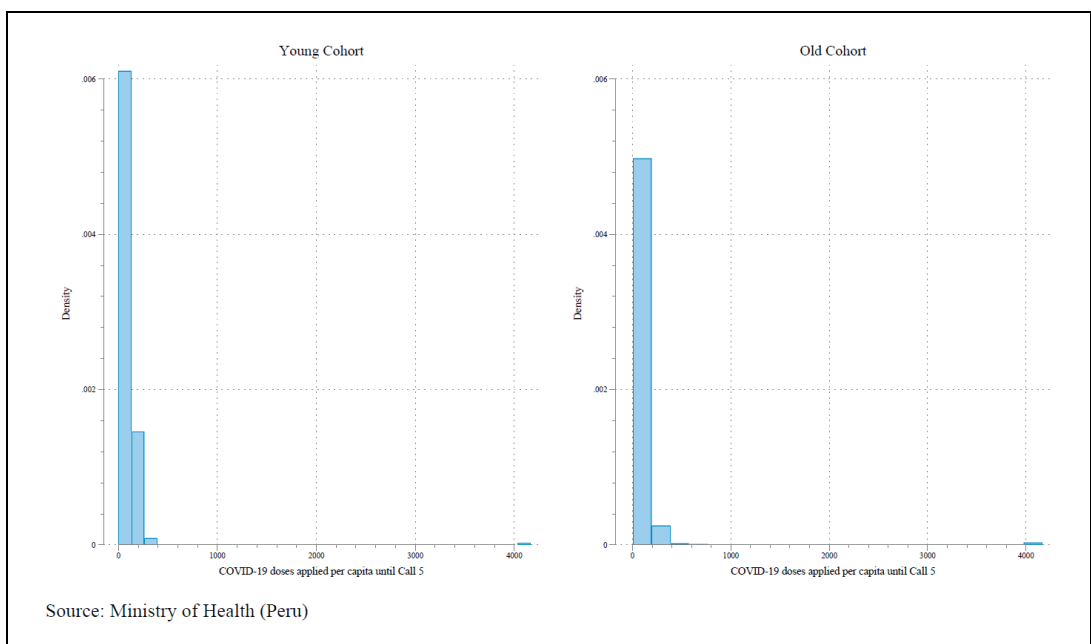


### COVID-19 vaccination doses per capita distribution by cohort and Young Lives survey call

**Figure A7.** *Young Lives participants and COVID-19 vaccine doses applied per capita until Call 4, by cohort*



**Figure A8.** *Young Lives participants and COVID-19 vaccine doses applied per capita until Call 5, by cohort*







An International Study of Childhood Poverty

### About Young Lives

Young Lives is an international study of poverty and inequality, following the lives of 12,000 children in four countries (Ethiopia, India, Peru and Vietnam). Young Lives is a collaborative research programme led by a team in the Department of International Development at the University of Oxford in association with research and policy partners in the four study countries.

Through researching different aspects of children's lives across time, we seek to improve policies and programmes for children and young people.

---

### Young Lives Research and Policy Partners

#### Ethiopia

- *Policy Studies Institute*
- *Pankhurst Development Research and Consulting plc*

#### India (Andhra Pradesh and Telangana)

- *Centre for Economic and Social Studies, Hyderabad (CESS)*
- *Sri Padmavati Mahila Visvavidyalam (Women's University), Tirupati (SPMVV)*

#### Peru

- *Grupo de Análisis para el Desarrollo (GRADE)*
- *Instituto de Investigación Nutricional (IIN)*

#### Vietnam

- *Centre for Analysis and Forecast, Viet Nam Academy of Social Sciences (CAF-VASS)*
  - *General Statistics Office of Viet Nam (GSO)*
- 

### Contact:

#### Young Lives

Oxford Department of  
International Development,  
University of Oxford,  
3 Mansfield Road,  
Oxford OX1 3TB, UK  
Tel: +44 (0)1865 281751  
Email: [younglives@qeh.ox.ac.uk](mailto:younglives@qeh.ox.ac.uk)  
Website: [www.younglives.org.uk](http://www.younglives.org.uk)

