Final report

International Growth Centre

Subnational income, convergence, and the COVID-19 pandemic

M. Ali Choudhary
Ijlal A. Haqqani
Federico Lenzi
Nicola Limodio

August 2021

When citing this paper, please use the title and the following reference number:
S-20204-PAK-1

DIRECTED BY
LSE
UNIVERSITY OF OXFORD

FUNDED BY
UKaid
from the British people

Subnational Income, Convergence and the Covid19 Pandemic[§]

M. Ali Choudhary[¶] Ijlal A. Haqqani[‖] Federico Lenzi[**]

Nicola Limodio[††]

August 2021

Abstract

The Covid19 pandemic severely affected the growth and subnational distribution of income in low-income countries. To understand these effects, we produce monthly aggregates on gross national income (GNI) for 147 Pakistani districts through real-time data and machine learning between 2012 and 2021. Three findings emerge from our analysis. First, urban districts experienced a sizeable decline in income during the pandemic, as their monthly growth rate dropped by 0.21%. Second, districts with a higher Covid19 incidence experienced larger income drops. Third, Covid19 accelerated the within-country convergence, as districts with high pre-pandemic income experienced a stronger negative growth during the pandemic.

Keywords: Covid19 Pandemic, Growth, Convergence, Satellite Data

Word count: 4486

[§]We would like to thank the members of the Monetary Policy Committee of the State Bank of Pakistan namely: Governor Reza Baqir, Murtaza Syed, Jameel Ahmed, Asad Zaman, Naved Hamid, Azam Faruqee, Hanid Mukhtar and Tariq Hassan, for encouraging us to find innovative data sources to measure economic growth during the Covid-19 pandemic. We also express our gratitude to Minister Asad Umar, Ijaz Nabi, Brigadier Saeed, Lieutenant Colonel Adnan, Major Sami and his team, Imtiaz Ahmed, Kashif Sahazad and Nadeem Hanif. We would also like to acknowledge the useful suggestions of Johannes Boehm, Nicola Gennaioli, Federico Rossi, Chris Roth, Nicolas Serrano Velarde, Tom Schmitz and seminar participants at various conferences, seminars and workshops. We are grateful for the financial support of International Growth Center and the State Bank of Pakistan. We are responsible for all errors.

[¶]ali.choudhary@sbp.org.pk, State Bank of Pakistan, I.I. Chundrigar Road, Karachi, Pakistan, and Centre for Economic Performance, 32 Lincoln's Inn Fields, WC2A 3PH, London, UK.

[‖]Ijlal.Ahmad@sbp.org.pk, State Bank of Pakistan, I.I. Chundrigar Road, Karachi, Pakistan.

[**]federico.lenzi@kellogg.northwestern.edu, Northwestern University, Kellogg School of Management, 2211 Campus Drive Evanston, IL 60208

[††]nicola.limodio@unibocconi.it, www.nicolalimodio.com, Corresponding Author, Bocconi University, Department of Finance, BAFFI CAREFIN and IGIER, Via Roentgen 1, 20136 Milan, Italy.

1 Introduction

The Covid19 pandemic is producing dramatic changes around the globe, as the effects of the virus and government containment policies have been disrupting our societies and economies since February 2020. While there is ample knowledge on how high-income countries and their subnational units reacted (Chetty et al. (2020), Woloszko (2020), Chen et al. (2020), Monache et al. (2020)), the same level of analysis and evidence is lacking in low-income countries because of the scarce availability of recent data.

This paper examines the effect of the Covid19 pandemic on subnational income in a low-income country, Pakistan and its districts. We develop a machine-learning algorithm to nowcast the gross national income (GNI) covering 147 districts in Pakistan from 2012 to 2021, by combining traditional administrative data with night-lights and other real-time data. Our approach builds on the frontier in this literature by relying on multiple satellite products (Asher et al. (2021)), integrating our prediction exercise with a robust empirical inference (Athey (2017)) and connecting micro with macro data (Vavra (2021)).

Three novel findings contribute to the current debate on the effects of the Covid19 pandemic and containiment policies on the economy. First, we observe that Pakistani income slowed down during the pandemic and this is primarily due to urban districts, whose monthly income growth declined by 0.21% bringing the overall monthly growth rate near zero, and below, in many instances. Second, there exists a robust and negative correlation between the incidence of Covid19 (cases, deaths, recoveries) and income growth. Third, we observe that the previous two effects lead to an acceleration of within-country convergence toward an overall lower level of income. This result is due to higher-income districts experiencing a higher negative income growth during the pandemic and creating a "convergence to the bottom".

Our findings bring a perspective to three separate, yet interrelated literature. First, we contribute to the novel and emerging cross-country convergence literature (Patel et al. (2021), Kremer et al. (2021), Pande and Enevoldsen (2021), Acemoglu and Molina (2021)) by offering a within-country perspective. Our monthly dataset uncovers that conditional convergence has been taking place in Pakistan. In fact, districts with a lower level of income in 2012 have been growing faster than districts with higher income before the pandemic. We also observe an acceleration of convergence during the Covid19 period. However, a key distinction exists between the pre-pandemic and pandemic-induced convergence. While the former is due to a higher growth across low-income districts and may represent a permanent move toward a new steady state, the pandemic acceleration in convergence is due to high-income and urban districts slowing down the most and may only be a temporary shock - an effective convergence to the bottom. Our work is also connected to the literature on growth in regions and regional convergence (Gennaioli et al. (2013), Gennaioli et al. (2014), Hsieh and Moretti (2015), Ganong and Shoag (2017), Giannone (2017), Lessmann and Seidel (2017)) showing two key determinants of the recent pandemic-induced recession: urbanization and Covid19 incidence. While we highlight the differential effect of the pandemic across urban and rural districts, we are unable to investigate at this stage which elements of urbanisation drive our results (population, density, sectoral differences et cetera). Second, our results are in line with the work of Saez and Zucman (2016) showing that inequality declines during recessions, though this specific case may be due to a decline in contact-intensive activities in urban centers (Koren and Peto (2020)) rather than financial returns. In this respect, our results are aligned with the findings of Deaton (2021) on the lower pandemic-induced within-country inequality and the recent World Bank report suggesting that years of poverty eradication vanished in few months.[1] Finally, this paper contributes to an emerging literature in macro-development assessing the effects and costs of Covid19 on low and middle income countries (Alfaro et al. (2020), Alon et al. (2020), Gottlieb et al. (2021b), Gottlieb et al. (2021a)).

The remainder of the paper is organized as follows: the next section introduces some key papers in this literature, we illustrate the data gathering procedures and the methodology in Section 3, present the main results in Section 4 and leave a technical guide on the employed algorithms in the Online Appendix. Finally, Section 5 offers some concluding remarks.

2 Related Literature

As stressed by the Bank for International Settlements (Tissot and de Beer (2020)), the current crisis has called into question the traditional statistical aggregates. The constant mutations of

[1] Refer to "Updated estimates of the impact of COVID-19 on global poverty: Turning the corner on the pandemic in 2021?" by Daniel Gerszon Mahler, Nishant Yonzan, Christoph Lakner, R. Andres Castaneda Aguilar and Haoyu Wu, published on June 24th, 2021, on the World Bank Data Blog and available at https://blogs.worldbank.org/opendata/updated-estimates-impact-covid-19-global-poverty-turning-corner-pandemic-2021

the virus result in a rapidly escalating framework, where the economic impact varies heterogeneously among sectors and geographic areas. Standard statistical aggregates are often available at the national level and with several months of delays. For this reason, a literature exploring novel sources of data is rapidly expanding.

Chetty et al. (2020) is an important contribution in this field. Exploiting real-time and granular data of American companies, it tracks the crisis impact on consumption and labor market. Through a different approach, Woloszko (2020) proxies them from Google Trend and nowcasts the national GDP for 46 OECD and G20 countries. A wider approach is proposed by Chen et al. (2020), integrating search queries with electricity and unemployment data. Following a similar approach, Monache et al. (2020) builds a weekly economics index for Italy through granular administrative data.

Similar studies are not reproducible in emerging markets with a structural deficiency of administrative data and low Internet penetration. To overcome this obstacle, a growing number of researchers is recurring to satellite data (see Donaldson and Storeygard (2016) and Nagaraj and Stern (2020)). This novel source of information is available at a very granular level for the entire globe and almost in real-time. Following this literature, Beyer et al. (2020) combines VIIRS night-lights and electricity consumption to monitor the pandemic impact in India. This study shows that the drop in habitual activities persists after the restrictions' lifting. It also suggests as the pandemic particularly affects the manufacturing and in-migration areas, while the out-migration states seem to experience a reduced decline. Also the work of Roberts (2021) obtains similar results, using night-lights to study Covid19's impact on Morocco.

In this literature, the work of Henderson et al. (2012) has popularized the use of night-lights as a popular proxy for economic development in emerging markets. Although, the recent findings of Asher et al. (2021) cast some shadow on their effectiveness in time-series analysis. Their elasticity with the local output varies according to the level of aggregation and the context. In other words, night-lights can be correlated to several development indicators and discerning what they are proxying in different regions results difficult. Some papers overcome this issue by adopting different and more detailed proxies for local economic output. Engstrom et al. (2017) proves as the extraction of daytime features from satellite data explains 60% of average log consumption in emerging markets. Jain (2020) shows as all the satellite data hide implicit biases (for example clouds, saturation, non-random misclassification, meteorological variables) in the realization process. Whereas Burke et al. (2019) specifies as the errors attributed to these models tend to be overestimated and related to the low-quality administrative data used as reference.

Our work includes insights from this literature: we use satellite lights in line with Henderson et al. (2012), but consider different moments of these series and include other real-time datasets as discussed by Asher et al. (2021). In addition to these datasets, we also partnered with local electricity providers to build a district-level electricity dataset, in line with Beyer et al. (2020), and add a finer split between electricity used for domestic, commercial and industrial use.

3 Data and Methodology

The present sections provide an overview of the data and the methods used in the analysis. While Appendix A offers schematic information on the employed database, Online Appendix A digs in the details of the data extraction methodologies.

3.1 Data

Pakistan is one of the few developing countries offering detailed and extensive administrative data. Most of these resources are available in traditional wide economic macro-aggregate, but micro-aggregates are often available on request. The principal statistical publications are released by the State Bank of Pakistan, the Ministry of Finance and the Pakistani Bureau of Statistics. From the latter, we use some granular variables contained in the annual Pakistan Economic Survey (cities wages, doctors consulting fees, import/export of cargos in the main terminals). We also employ additional indicators coming from the Monthly Bulletins of Statistics released by the Pakistan Bureau of Statistics. This publication reports price indexes for over four hundred items at the city-monthly level. At the same time, the National Electric Power Regulatory Authority produces granular data for electricity consumption at the tehsil-monthly level, a subnational unit more granular than the district. This information also distinguishes among the different destination uses (commercial, domestic, industrial, others). Statistics on Covid19 are produced by the provincial authorities and released at the district level upon our request, reporting basic indicators (number of cases, deaths and recoveries).

The aggregates for the gross national income in Pakistan at the province level come from the United Nations' Sustainable Development Goals dataset, which combines country-level information with the periodical household survey. This work is realized by the Global Data Lab, hosted by the Nijmegen Center for Economics (NiCE) at Radboud University in the Netherlands. Our study relies on the yearly gross national income per capita (2011 $ PPP) between 2010 to 2018. In addition, we decompose this aggregate at the district level using the share of provincial population per district from the WorldPop platform by the United Nations with unconstrained data. In our work we also include accurate weather data. This is very important given that the agriculture sector still plays a primary role in national accounts. For this reason, we obtain weather data for 140 cities from the State Bank of Pakistan. The information is subsequently assigned to all the administrative units exploiting a proximity criterion.

In line with the literature on satellite imaginery, we also employ the VIIRS night-lights to track local economic activities. The Earth Observations Group releases this data at the yearly and monthly levels in conventional formats. However, there are often long gaps before their publications. To monitor this rapidly escalating framework, we opt for the high-frequency and less pre-processed VNP46A1 - VIIRS/NPP Daily Gridded Day Night Band 500m Linear product. This daily time series is updated after few hours, constituting an unvaluable tool for emergencies response. An automatic algorithm removes outliers, substituting them with the most recent moonlight-adjusted observations. This mechanism permits to phase out the impact of snow, clouds and other artifacts. We build our time series by choosing the median day of each month. When not possible, given calibration problems or prolonged bad weather conditions, we replace the missing observation with the nearest date. The selected dates are, then, manually cleaned from local imperfections: we sample the background noises in some areas of the country (Gilgit-Baltistan and Balochistan), defining a threshold for separating them from human-made luminous emissions. This approach differs from Beyer et al. (2020) for two main reasons:

- the aggregation of night-lights can delete tiny settlements around metropolitan areas;
- the use of an older reference model can exclude novel neighborhoods.

Considering the results of Asher et al. (2021), we proxy the economic activities also with other satellite data. The most important is the yearly landcover map produced by the European

Space Agency (ESA) under the Climate Change Initiative (CCI). This product divides the world into quadratic grids of 300mt, estimating the primary destination use of the soil. As a reference, it adopts the twenty-two classes defined by the United Nations Food and Agriculture Organization's (UN FAO) Land Cover Classification System (LCCS). This product records a variation only if present for two years in a row. Thus, it is available only from 1992 onward. Mapping the evolution of urban and cultivated lands provides valuable insights into the Pakistani economy and growth. A notable precedent is provided by Vogel et al. (2018), exploiting the MODIS land cover data to detect urban markets in India.

Other features come from the Nasa Earth Observations platforms. This service offers monthly pre-processed rasters with a high resolution of 0.1 degrees. The available data spans from observations on meteorological events (land surface temperature, clouds cover and density, water vapor, ...) to agriculture (vegetation index, leaf area index, night-fires,...) and pollution (carbon monoxide, aerosol thickness, ozone,...). As already mentioned, the first two sets of variables are essential for monitoring the rural dimension of this economy. On top of this, the literature finds pollution a good proxy for human activities. The seminal contribution of Grossman and Krueger (1995) demonstrates as environmental deterioration strictly soars during the first phases of economic development, improving after a certain level of per capita income. The Nasa Earth Observations also release valuable insights on the fishing sector, monitoring water temperatures and chlorophyll concentration.

Lastly, we use another set of novel satellite products: the VIIRS Boat Detection (VBD) and the Global Gas Flares Observed from Space. The first exploits the high resolution of the VIIRS sensor to monitor the marine traffic and the extraction platforms in real-time. The second uses the same technologies to track the quantity and temperature of upstreams/downstream flares. In this way, it monitors the entire oil industry from fields to refineries.

3.2 Methodology

Combining all the data gathered in the previous section, we obtain a balanced monthly panel for all the 554 Pakistani tehsils (these are third-level administrative units, after provinces and the districts). Our final dataset spans from January 2012 to March 2021. We do not include years before 2012, as the VIIRS sensors are only operative from such date onward.

The reference variable for the prediction exercise is the gross national income at the tehsil level, produced in the previous section. Most of the employed data are pre-processed and do not present missing/outliers. Notwithstanding, we use a min-max normalization to rescale them between 0 and 1. The heterogeneous ranges of our data make this step compulsory. Otherwise, some variables may assume a predominant role only because of their scale. We opt for this normalization over standardization, as we do not have a prior over their distribution. To implement the machine-learning algorithms, we also convert categorical variables into larger set of dummy variables.

The data on the yearly gross national income is available until 2018. Thus, we assign three-quarters of this sample to the training set and the remaining to the test one. We develop a model consisting of series of machine learning algorithms (elastic net, random forest, bagging, boosting, support vector machine) on the yearly aggregates of satellite and economic data. This model also includes a categorical variable for the location. Our final algorithm can directly compute the monthly income from the observed values. Among the several models tested, the bagging obtains the best result with an overall mean square error of 0.01506. Appendix B reports additional elements on the methodology for further consultation, including the mean square error of all models and the ten most important predictors.

As well as generating the lowest mean square error, our algorithm also produces economically meaningful estimates. The right panel of Figure 3 compares our aggregates produced through the bagging method with the official World Bank statistics on GNI using 2017 PPP and current dollars. While these numbers are close, there is one crucial distinction. Our estimated GNI is higher in level and grows more slowly than the World Bank estimates. Both of these effects may be due to an adequate accounting of the informal economy, which the predictors in Figure 4 may reflect. At the same time, we acknowledge that our estimates display a relatively contained country-wide volatility. This may be a result of the limited availability of district-level or tehsil-level aggregates with a high frequency. While more refinements on statistical capacity are left for future research and policy, we believe that these are the most comprehensive estimates, especially among low and middle income countries.

Because our dataset for Covid19 incidence is available only at the district level, we sum the gross national income from the tehsil to the district level to proceed with the analysis. Figure

5 in Appendix C shows the evolution of Covid19 in Pakistan, indicating that May 2020 was the first month in which the pandemic touched the country and never left. For this reason, we define the arrival of the pandemic in Pakistan from this month onward.

To understand the geographic extent of the pandemic, Figure 6 in Appendix C provides three maps indicating the incidence of Covid19. These figures plot the total number of Covid cases, deaths and recoveries from March 2020 to March 2021. It is evident that the most densely populated districts are the most affected by the contagion. Table 1 reports the summary statistics for the main variables used in the empirical analysis. Table 5 in Appendix D reports additional statistics.

Table 1: Summary Statistics

Variables	(1) Observations	(2) Mean	(3) St. Deviation	(4) 50th P.tile	(5) 5th P.tile	(6) 95th P.tile
Log Income	5,733	21.93	1.368	22.03	19.91	23.94
Growth Income	5,733	0.297	2.556	0.195	-2.505	3.324
Dummy Covid	5,733	0.282	0.450	0	0	1
Covid Cases	5,733	785.3	6,565	0	0	2,073
Covid Deaths	5,733	19.91	158.4	0	0	56
Covid Recoveries	5,733	681.2	5,946	0	0	1,721

Notes: This table reports the summary statistics of all the variables considered in this study. The variable "Log Income" represents the logarithm of the district gross income, while "Income Growth" is the percentage variations between months. The "Dummy Covid" assumes the value 1 from May 2020 to March 2021. "Covid Cases", "Covid Deaths" and "Covid Recoveries" are set to zero for the months preceding the pandemic. The dataset follows 147 districts from January 2018 to March 2021.

4 Results

Figure 1 reports the results of our estimation aggregated at the country level on the left and divided between the average urban district (in blue and dot) and rural district (in red, with dash and square) on the right. The left panel of Figure 1 shows a steady growth path from early 2018 to August 2020. The persistence of the pandemic arrests this path in September 2020, when a steady declines begins. Figure 7 in Appendix C shows the same picture but from 2012 to 2021.

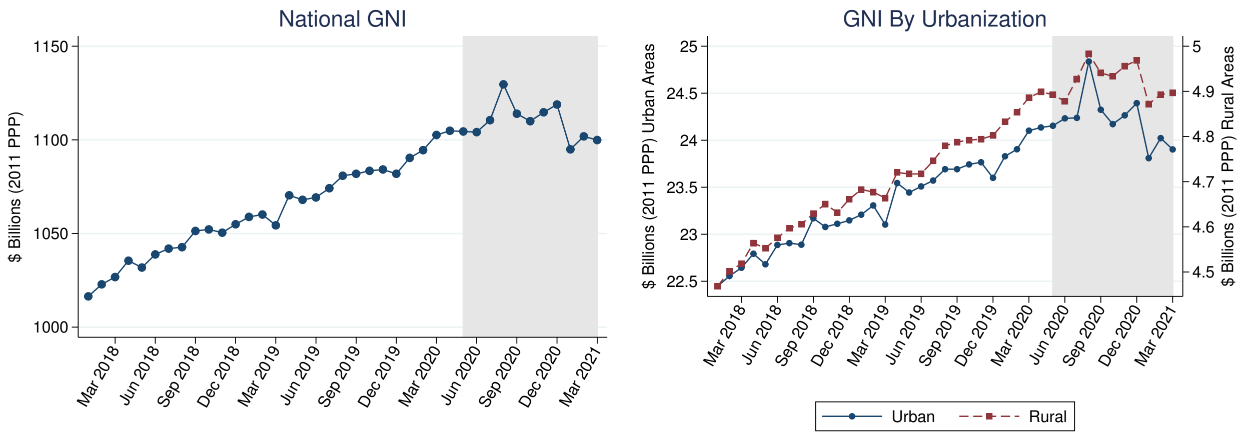
To understand the drivers of this decline, we define a district as being "urban" if it contains one of the top 20 cities by size, as defined by the Pakistani Bureau of Statistics in its 2017 census.[2] As a result, 20 districts are classified as "urban" and the remaining 127 as "rural".

[2]The list of principal cities established with the 2017 census, available at https://www.pbs.gov.pk/content/provisional-summary-results-6th-population-and-housing-census-2017-0 and also at https://en.wikipedia.org/wiki/List_of_cities_in_Pakistan_by_population.

The right panel of Figure 1 highlights that the GNI of the average urban district (on the left y-axis) is four times as large as the average rural district (on the right y-axis). It is important to note that while urban district exhibits a steep decline in income as the pandemic begins, this decline is much milder in rural districts. These might have experienced a lower loss in economic activity due to the suspension of high-interaction activities following the lockdown and other mobility restrictions.

Figure 8 in Appendix C offers a version of Figure 1 in which urban districts are classified by only using the top 10 cities, on the right, or the top 70 cities on the left (these are all the ones made available by the census) and the results are qualitatively similar, with urban districts exhibiting a key decline and rural districts seeing a mild slowdown.

Figure 1: Gross National Income 2018-2021

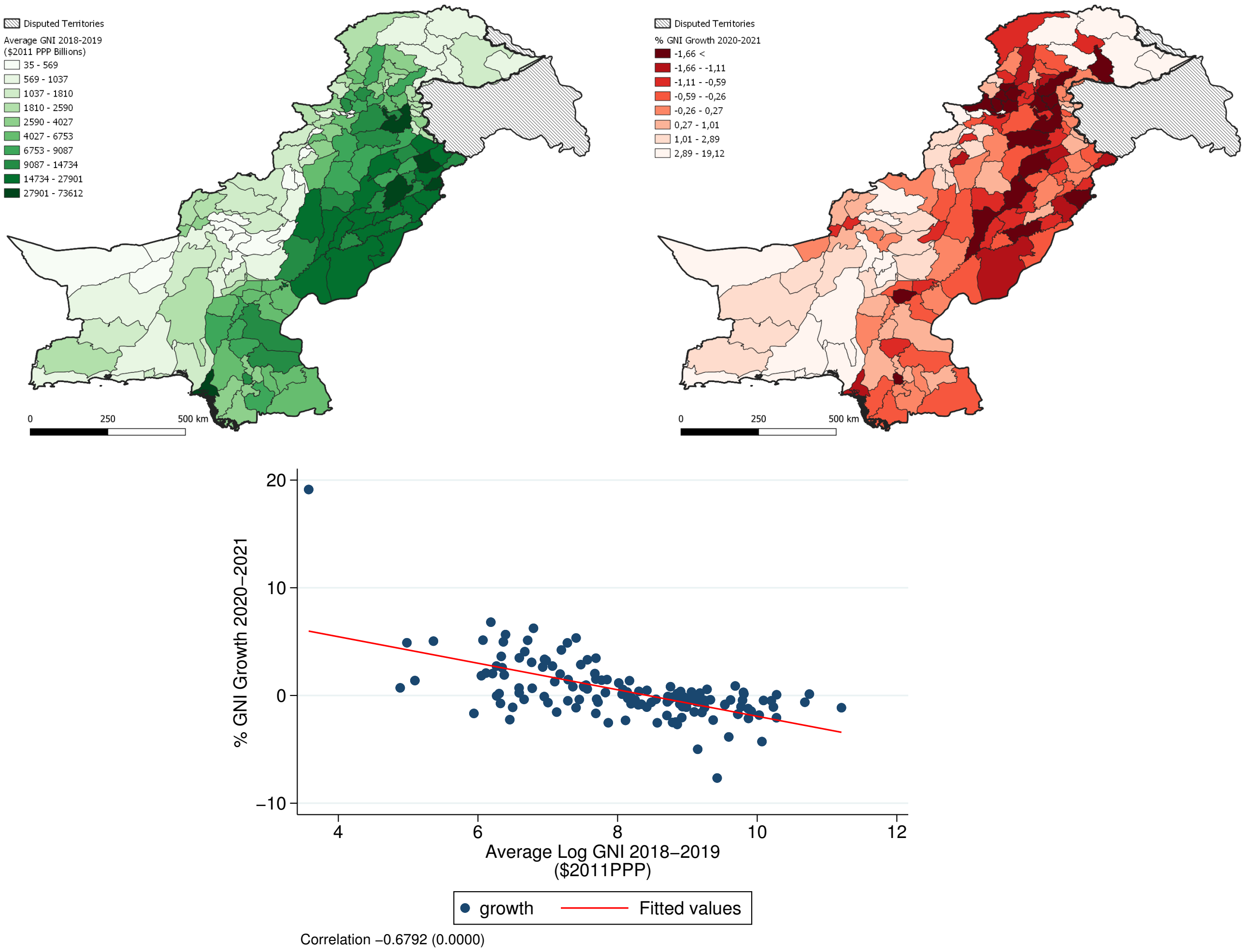


Notes: The first graph reports the Pakistani Gross National Income from January 2018 to March 2021. The second plot shows the mean income of rural districts in red (RHS vertical-axis) and urbanized districts in blue (LHS vertical-axis). All the values are expressed in billions of 2011 PPP dollars.

Figure 2 reports three pictures. The left panel presents a map with the average district income between 2018 and 2019. In this map, high-income districts are indicated with dark green colors and it is notable that concentration of economic activities takes place along the Indus River and metropolitan areas (Islamabad, Karachi, Lahore, Peshawar, Quetta). The arid and sparsely populated lands of Balochistan appear to be the poorest, followed at a considerable distance by the mountain regions of Gilgit-Baltistan and Khyber Pakhtunkhwa. The center panel shows the average income growth between 2020 and 2021. In this case, darker colors indicate a stronger decline in growth. We document a marked economic decline in the densely populated Punjab districts and in major urban areas (Peshawar, Quetta, Sargodha, Hyderabad, Abbottabad, Rawalpindi). By comparing these two maps, it is already clear that districts with high-income before the pandemic were the most hit during the pandemic.

The panel at the bottom of Figure 2 shows exactly this negative correlation between GNI growth during 2020-2021 on the y-axis and the log level of GNI in the previous year on the x-axis. This correlation is -0.67 and statistically different from zero below the 1% significance threshold. Figure 9 in Appendix C reports the same descriptive evidence, but in terms of income per capita and the results are similar, including the negative and significant correlation between the pandemic growth of per capita income and the pre-pandemic level of income. We prefer to present the analysis in terms of levels of income, rather than per capita, given that the numbers for the population may not be adjusted based on the incremental Covid19 mortality. Figure 10 in Appendix C shows the overall income growth during the pre-pandemic period from 2012 to 2019.

Figure 2: GNI and GNI Growth by District



Notes: The first map reports the average district-wise income between 2018 and 2019, while the second displays its percentual variation between 2020 and 2021. For 2021, we consider only the first three months for which data is available. All the values are expressed in billions of 2011 PPP dollars. The graph below reports the district average gross national income in 2018-2019 on the horizontal axis and the percentage variation between 2020-2021 on the vertical one. The linear relation between these variables is reported in red, while the correlation is noted at the bottom.

After these descriptive figures, we explore two important elements: 1) the relation between income growth and Covid19 incidence; 2) the differential effect of Covid19 between Urban and Rural districts. To do so, we explore the following empirical model

$$growth_{dmy} = \alpha_m + \gamma_y + \beta\, Covid19_{dmy} + \epsilon_{dmy} \tag{1}$$

in which the monthly income growth of district d in month m of year y is regressed on four measures of Covid19 incidence and both month and year fixed effects. The standard errors are clustered at the district level. The four measures of incidence are: 1) a dummy variable which takes unit value from May 2020 onward for all districts in Pakistan, $Covid19_{my}$; 2) the natural logarithm of the number of Covid19 cases in district d during the month m of year y, $Cases_{dmy}$; 3) the natural logarithm of the number of Covid19 deaths in district d during the month m of year y, $Deaths_{dmy}$; 4) the natural logarithm of the number of Covid19 recoveries in district d during the month m of year y, $Recoveries_{dmy}$.

In addition to this, we augment equation (1) by including an interaction for a dummy describing urban districts, $Urban_d$, as districts in which the 20 largest Pakistani cities are placed. Adding this variable is valuable to understand a key determinant of the Covid19 impact on the Pakistani economy. However, it is important to note that Covid19 incidence is very high in urban districts, compared to rural ones. Table 6 in Appendix D shows that urban districts present an incidence of cases, deaths and recoveries related to Covid19 more than 100% higher than rural districts.

Table 2 presents the empirical results of estimating equation (1). Column (1) of Panel A shows that the Covid19 dummy exhibits neither a negative nor a statistically significant effect on growth between 2018 and 2021, therefore we conclude that the overall district growth during Covid19 does not differ from zero across Pakistani districts. However, the remaining 3 columns of Panel A show that districts exhibiting a higher incidence of Covid19 cases, deaths or recoveries experience a lower income growth during this period. Column (2) shows that a 100% increase in Covid19 cases implies a 0.0167 decline income growth, with this number being 0.046 for deaths in column (3) and 0.035 for recoveries in column (4).

Panel B of Table 2 further investigates whether and how the results of Panel A differ across urban and rural districts by adding a series of interactions with a dummy for urban district,

$Urban_d$. Column (1) shows a key result for our analysis. The Covid19 dummy does not have an effect on growth on average, as we cannot reject that its coefficient is statistically different from zero. However, its interaction with the Urban dummy shows that the growth rate of urban districts declines by 0.21%, which made the overall growth of these districts zero in most cases and negative in some. Columns (2), (3) and (4) extend the results of Panel A with the urban dummy interaction. This shows that the elasticity of growth to cases is statistically different from zero and negative only for urban districts (and is not for rural districts), while the effect of deaths and recoveries on growth does not appear to be statistically different between rural and urban districts.

A battery of additional tables verifies the robustness of this result across specifications. Table 7 in Appendix D replicates Table 2 and also includes district fixed effects: while those results have a slightly different interpretation (deviations in growth in Table 7 versus growth in Table 2), the magnitudes, signs and significance are very close. Table 8 in Appendix D replicates the results of Table 2 but replaces the growth of income with the natural logarithm of the level of income and the results are qualitatively aligned. Table 9 replicates Table 2 but replaces the urban dummy with two alternative definitions and the results are again in line with the previous ones. Finally, Table 10 and Table 11 present the same results as those in Table 2 but using the logarithm of income per capita and its growth with results being consistent with the ones discussed.

To inspect whether and how Covid19 shaped the convergence of Pakistani districts, we explore the following empirical model

$$growth_{dmy} = a\,Income_{d2012} + b\,Covid_{my} + c\,Income_{d2012} \times Covid_{my} + u_{dmy} \qquad (2)$$

in which we regress the monthly growth of district d in month m of year y on the level of income of district d in 2012 (the first year in our sample) $Income_{d2012}$, the dummy describing the months during which Covid19 begun from May 2020 onward $Covid_{my}$ and an interaction between these two variables. As in equation (1), we cluster standard errors at the district level.

Table 3 presents five versions of equation (2), with the first two being presented without any fixed effect. The first column gives evidence of conditional convergence and shows that districts with a 1 standard deviation lower income in 2012 are growing by 0.023 points more between 2012 and 2021. Column (2) introduces the Covid19 dummy and its interaction with the standardized level of income in 2012. In this case, the coefficient on the first term declines in point estimate and becomes insignificantly different from zero, yet negative. The Covid19 dummy is negative and statistically different from zero below the 1% conventional threshold. The interaction between the level of 2012 income and Covid19 dummy is particularly interesting, because it shows the presence of a pandemic-induced convergence taking place across Pakistani districts. This result is quite stable across the four specifications of Table 3: regardless of the presence of district, month and year fixed effects, the point estimate and significance do not change.

To further validate these findings, Appendix D offers two additional tables. First, Table 12 replaces the definition of income: instead of the standardized income in Table 3, it employs the natural logarithm of the mean income per district in 2012. The results are qualitatively unchanged, especially the interaction between income and Covid19. Finally, Table 13 replicates Table 3 but replaces the measure of Income with three dummies which map each district into an income tercile in 2012. This test is particularly useful because it shows that the convergence during the Covid19 pandemic is due to districts that were in the second and third tercile in 2012, which exhibited the highest negative growth during the pandemic period.

Table 2: Covid19 and GNI Growth, 2018 - 2021

Panel A - Overall

	(1)	(2)	(3)	(4)
Variables		Income Growth		
$Covid_{my}$	0.0625			
	(0.0636)			
$Cases_{dmy}$		-0.0167**		
		(0.00763)		
$Deaths_{dmy}$			-0.0460***	
			(0.00934)	
$Recoveries_{dmy}$				-0.0358***
				(0.00695)
Month FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Obs.	5733	5733	5733	5733
Adj. R sq.	0.0173	0.0175	0.0182	0.0185
Mean Dep. Var.	0.297	0.297	0.297	0.297
S.D. Dep. Var.	2.556	2.556	2.556	2.556

Panel B - Urban

	(1)	(2)	(3)	(4)
Variables		Income Growth		
$Covid_{my}$	0.0918			
	(0.0661)			
$Covid_{my} \times Urban_d$	-0.216***			
	(0.0596)			
$Cases_{dmy}$		-0.00984		
		(0.00812)		
$Cases_{dmy} \times Urban_d$		-0.0131**		
		(0.00659)		
$Deaths_{dmy}$			-0.0414***	
			(0.0110)	
$Deaths_{dmy} \times Urban_d$			-0.00690	
			(0.0104)	
$Recoveries_{dmy}$				-0.0332***
				(0.00760)
$Recoveries_{dmy} \times Urban_d$				-0.00543
				(0.00791)
$Urban_d$	-0.0339	-0.0709**	-0.0511	-0.0524
	(0.0362)	(0.0356)	(0.0343)	(0.0364)
Month FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Obs.	5733	5733	5733	5733
Adj. R sq.	0.0173	0.0173	0.0179	0.0182
Mean Dep. Var.	0.297	0.297	0.297	0.297
S.D. Dep. Var.	2.556	2.556	2.556	2.556

Notes: Panel A estimates the impact on the income's growth rate of the pandemic months (column 1), the log cases (column 2), the log deaths (column 3) and the log recoveries (column 4). Panel B repeats this analysis decomposing the impact between rural and urban areas. The study follows all the 147 Pakistani districts from January 2018 to March 2021 We control for the year and month fixed effects. Standard errors are clustered at the district level. The number of observations and adjusted R2, indicated as Adj. R sq., of each regression are reported at the end of the table. The last row presents the mean of the dependent variable (Mean Dep. Var.). ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Table 3: Covid19 and Convergence, 2012 - 2021

	(1)	(2)	(3)	(4)	(5)
Variables			Income Growth		
$Income\ 2012_d$	-0.0303***	-0.0164**	-0.0164**		
	(0.00885)	(0.00731)	(0.00731)		
$Covid_{my}$		-0.137***	0.00250	-0.137***	0.00250
		(0.0380)	(0.0635)	(0.0380)	(0.0635)
$Income\ 2012_d \times Covid_{my}$		-0.139***	-0.139***	-0.139***	-0.139***
		(0.0449)	(0.0449)	(0.0449)	(0.0449)
District FE	No	No	No	Yes	Yes
Month FE	No	No	Yes	No	Yes
Year FE	No	No	Yes	No	Yes
Obs.	16170	16170	16170	16170	16170
Adj. R sq.	7.38e-05	0.000458	0.0118	-0.00627	0.00509
Mean Dep. Var.	0.324	0.324	0.324	0.324	0.324
S.D. Dep. Var.	2.600	2.600	2.600	2.600	2.600

Notes: Column (1) estimates the impact on the income growth rate of the standardized mean income in 2012, without controlling for fixed effects. The remaining columns examine how the standardized mean income in 2012, the pandemic period and the interaction between these terms influence the growth rate. The second column control for no fixed effect, the third for year and month fixed effects, the fourth for district fixed effect, the fifth for istrict, month and year fixed effects. Standard errors are clustered at the district level. The study follows all the 147 Pakistani districts from January 2012 to March 2021. The number of observations and adjusted R2, indicated as Adj. R sq., of each regression are reported at the end of the table. The last row presents the mean of the dependent variable (Mean Dep. Var.). ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

5 Conclusions

In this paper, we apply a method at the frontier to calculate monthly aggregates on gross national income (GNI) for 147 Pakistani districts between 2012 and 2021 using machine learning, real-time and satellite data. Our work shows that urban districts presented a negative and large income growth during the Covid19 pandemic, as the average monthly growth rate dropped by 0.21%. We verify that the incidence of Covid19 measured through cases, deaths and recoveries appears to have a negative and sizeable effect on income. Finally, we show that Covid19 induced a sizeable within-country convergence, as districts with high pre-pandemic income experienced negative and strong growth during the pandemic. While, on the one hand, this may reduce inequality and the prominence of urban centers, on the other hand, this process may lower the long term prospects of the most dynamic Pakistani districts and harm long-term growth.

References

Acemoglu, D. and Molina, C. A.: 2021, Converging to convergence? a comment, *National Bureau of Economic Research* .

Alfaro, L., Becerra, O. and Eslava, M.: 2020, Emes and covid-19 shutting down in a world of informal and tiny firms, *National Bureau of Economic Research* .

Alon, T., Kim, M., Kim, M., Lagakos, D. and VanVuren, M.: 2020, How should policy responses to the covid-19 pandemic differ in the developing world?, *National Bureau of Economic Research* .

Asher, S., Lunt, T., Matsuura, R., Novosad, P. and Novosad, P.: 2021, Development research at high geographic resolution: An analysis of night-lights, firms, and poverty in india using the shrug open data platform, *The World Bank Economic Review* .

Athey, S.: 2017, Beyond prediction: Using big data for policy problems., *Science* .

Beyer, R., Franco-Bedoya, S. and Galdo, V.: 2020, Examining the economic impact of covid-19 in india through daily electricity consumption and nighttime light intensity, *World Development* .

Burke, M., Burke, M., Driscoll, A., Lobell, D. B., Lobell, D. B. and Ermon, S.: 2019, Using satellite imagery to understand and promote sustainable development, *Science* .

Chen, S., Igan, D., Pierri, N. and Presbitero, A. F.: 2020, Tracking the economic impact of covid-19 and mitigation policies in europe and the united states, *IMF Working Papers* .

Chetty, R., Friedman, J. N., Friedman, J. N., Hendren, N., Stepner, M. and Stepner, M.: 2020, The economic impacts of covid-19: Evidence from a new public database built using private sector data, *National Bureau of Economic Research* .

Deaton, A.: 2021, Covid-19 and global income inequality, *National Bureau of Economic Research* .

Donaldson, D. and Storeygard, A.: 2016, The view from above: Applications of satellite data in economics, *Journal of Economic Perspectives* .

Engstrom, R., Hersh, J. and Newhouse, D.: 2017, Poverty from space : using high-resolution satellite imagery for estimating economic well-being, *World Bank Policy Research Working Paper* .

Ganong, P. and Shoag, D.: 2017, Why has regional income convergence in the u.s. declined, *Journal of Urban Economics* .

Gennaioli, N., Porta, R. L., de Silanes, F. L. and Shleifer, A.: 2013, Human capital and regional development, *Quarterly Journal of Economics* .

Gennaioli, N., Porta, R. L., de Silanes, F. L. and Shleifer, A.: 2014, Growth in regions, *Journal of Economic Growth* .

Giannone, E.: 2017, Skill-biased technical change and regional convergence, *Research Papers in Economics* .

Gottlieb, C., Grobovsek, J., Poschke, M. and Saltiel, F.: 2021a, Lockdown accounting, *B E Journal of Macroeconomics* .

Gottlieb, C., Grobovsek, J., Poschke, M. and Saltiel, F.: 2021b, Working from home in developing countries, *European Economic Review* .

Grossman, G. M. and Krueger, A. B.: 1995, Economic growth and the environment, *Quarterly Journal of Economics* .

Henderson, J. V., Storeygard, A. and Weil, D. N.: 2012, Measuring economic growth from outer space, *The American Economic Review* .

Hsieh, C.-T. and Moretti, E.: 2015, Why do cities matter? local growth and aggregate growth, *Research Papers in Economics* .

Jain, M.: 2020, The benefits and pitfalls of using satellite data for causal inference, *Review of Environmental Economics and Policy* .

Koren, M. and Peto, R.: 2020, Business disruptions from social distancing, *arXiv: General Economics* .

Kremer, M., Willis, J. and You, Y.: 2021, Converging to convergence, *National Bureau of Economic Research* .

Lessmann, C. and Seidel, A.: 2017, Regional inequality, convergence, and its determinants - a view from outer space, *European Economic Review* .

Monache, D. D., Emiliozzi, S. and Nobili, A.: 2020, Tracking economic growth during the covid-19: a weekly indicator for italy., *Bank of Italy Mimeo* .

Nagaraj, A. and Stern, S.: 2020, The economics of maps, *Journal of Economic Perspectives* .

Pande, R. and Enevoldsen, N. T.: 2021, Growing pains? a comment on "converging to convergence", *National Bureau of Economic Research* .

Patel, D., Sandefur, J. and Subramanian, A.: 2021, The new era of unconditional convergence, *Journal of Development Economics* .

Roberts, M.: 2021, Tracking economic activity in response to the covid 19 crisis using nighttime lights the case of morocco, *Development Engineering* .

Saez, E. and Zucman, G.: 2016, Wealth inequality in the united states since 1913: Evidence from capitalized income tax data, *Quarterly Journal of Economics* .

Tissot, B. and de Beer, B.: 2020, Implications of covid-19 for official statistics: a central banking perspective, *Research Papers in Economics* .

Vavra, J.: 2021, Tracking the pandemic in real time: Administrative micro data in business cycles enters the spotlight, *Journal of Economic Perspectives* .

Vogel, K. B., Goldblatt, R., Hanson, G. H., Hanson, G. H. and Khandelwal, A. K.: 2018, Detecting urban markets with satellite imagery: An application to india, *National Bureau of Economic Research* .

Woloszko, N.: 2020, Tracking activity in real time with google trends, *Working Paper* .

Online Appendix

A Data

.

- **DATASET** Covid 19 Data
- **AGENCY** Provincial Authorities
- **FREQUENCY** Monthly
- **LEVEL** District
- **TIME March** 2020-May 2021
- **URL** Available on request
- **VARIABLES** Cases, Deaths, Recoveries

- **DATASET** Earth Observations Group
- **AGENCY** Nasa - National Aeronautics and Space Administration
- **FREQUENCY** Monthly
- **LEVEL** Tehsil
- **TIME** January 2012-March 2021
- **URL** https://neo.sci.gsfc.nasa.gov/
- **VARIABLES** Aerosol Thickness, Clouds Fraction, Clouds Optical Thickness, Clouds Particle Radius, Clouds Water Content, Night Fires, No2, Ozone, Sea Chlorophyll, Sea Surface Temperature, Snow on the Ground, Surface Temperature Anomaly Day, Surface Temperature Anomaly Night, Surface Temperature Day, Surface Temperature Night, Vegetation Index (NDVI), Water Vapor

- **DATASET** Electricity
- **AGENCY** National Electric Power Regulatory Authority and K-Electric
- **FREQUENCY** Monthly
- **LEVEL** Tehsil
- **TIME** July 2011-March 2021
- **URL** Available on request
- **VARIABLES** Commercial Consumption, Domestic Consumption, Industrial Consumption, Other Consumption

- **DATASET** Global Gas Flares Observed from Space
- **AGENCY** Nasa - National Aeronautics and Space Administration
- **FREQUENCY** Yearly
- **LEVEL** Tehsil
- **TIME** 2012-2020
- **URL** https://eogdata.mines.edu/download_global_flare.html
- **VARIABLES** Average Temperature, Clear Observations, Detection Frequency, Total Volume (for both upstream and downstream flames)

- **DATASET** Gross National Income
- **AGENCY** Global Data Lab
- **FREQUENCY** Yearly
- **LEVEL** Province

- **TIME** 2010-2018
- **URL** https://globaldatalab.org/areadata/view/gnic/PAK/?levels=1%2B2%2B3%2B5%2B4&interpolation=0&extrapolation=0&nearest_real=0
- **VARIABLES** Gross National Income Per Capita ($ 2011 PPP)

- **DATASET** Land Cover
- **AGENCY** Esa - European Space Agency
- **FREQUENCY** Yearly
- **LEVEL** Tehsil
- **TIME** 1992-2020
- **URL** https://cds.climate.copernicus.eu/cdsapp#!/dataset/satellite-land-cover?tab=overview
- **VARIABLES** Bare Areas, Crop Irrigated, Crop Rainfed, Grassland, Lichens and Mosses, Mosaic Crop, Mosaic Herbaceous Cover, Mosaic Tree and Shrubs, Mosaic Vegetation, Shrubland, Shrubs or Herbaceous Flooded, Sparse Vegetation, Tree Broadleaved Deciduous, Tree Broadleaved Evergreen, Tree Cover Flooded Fresh, Tree Cover Flooded Saline, Tree Mixed Leaf Type, Tree Needleaved Deciduous, Tree Needleaved Evergreen, Permanent Snow and Ice Urban Areas, Water Bodies

- **DATASET** Monthly Bullettins of Statistics
- **AGENCY** Pakistan Bureau of Statistics
- **FREQUENCY** Monthly
- **LEVEL** City
- **TIME** January 2012- March 2021

- **URL** https://www.pbs.gov.pk/publications
- **VARIABLES** Price in Pakistani Rupees for 486 Items

- **DATASET** Pakistan Economic Survey
- **AGENCY** Government of Pakistan - Finance Division
- **FREQUENCY** Yearly
- **LEVEL** City
- **TIME** 2006-2021
- **URL** https://www.finance.gov.pk/survey_2021.html
- **VARIABLES** Container Imported, Container Exported, Doctors Consulting Fees, Total Container in Terminals

- **DATASET** VIIRS Boat Detection (VBD)
- **AGENCY** Nasa - National Aeronautics and Space Administration
- **FREQUENCY** Monthly
- **LEVEL** Territorial waters
- **TIME** July 2016-May 2021
- **URL** https://eogdata.mines.edu/map_selector/
- **VARIABLES** Number of Blurry Lights, Number of Boats, Number of Gas Flares, Number of Glow Lights, Number of Platform lights, Number of Recurring Lights, Number of Weak Lights, Number of Weak and Blurry Lights

- **DATASET** VIIRS-Nightlights
- **AGENCY** Nasa - National Aeronautics and Space Administration
- **FREQUENCY** Monthly
- **LEVEL** Tehsil
- **TIME** January 2012-June 2021
- **URL** https://ladsweb.modaps.eosdis.nasa.gov/search/order/2/VNP46A1--5000
- **VARIABLES** Night-lights

- **DATASET** Weather
- **AGENCY** State Bank of Pakistan
- **FREQUENCY** Daily
- **LEVEL** City
- **TIME** January 2012-March 2021
- **URL** Available on request
- **VARIABLES** Cloud Cover, Dew Point, Feels Like Temperature, Heat Index, Humidity, Max Temperarure, Min Temperature, Moon Illumination, Pressure, Sun Hours, Total Snow, Total Precipitations, UV Index, Visibility, Wind Chill Temperature, Wind Degree, Wind Gust Speed, Wind Speed

- **DATASET** Population
- **AGENCY** WorldPop
- **FREQUENCY** Yearly

- **LEVEL** Tehsil
- **TIME** 2000-2020
- **URL** https://www.worldpop.org/geodata/listing?id=30
- **VARIABLES** Population by Age and Sex

SAMPLED DAYS VIIRS NIGHT-LIGHTS VNP46A1

20 January 2012; 22 February 2012; 15 March 2012; 22 April 2012; 20 May 2012; 20 June 2012; 22 July 2012; 26 August 2012; 15 September 2012; 21 October 2012; 15 November 2012; 15 December 2012; 15 January 2013; 17 February 2013; 7 March 2013; 16 April 2013; 16 May 2013; 8 June 2013; 10 July 2013; 4 August 2013; 14 September 2013; 9 October 2013; 28 November 2013; 28 December 2013; 25 January 2014; 24 February 2014; 29 March 2014; 23 April 2014; 23 May 2014; 29 June 2014; 21 July 2014; 15 August 2014; 15 September 2014; 21 October 2014; 15 November 2014; 15 December 2014; 15 January 2015; 21 February 2015; 18 March 2015; 20 April 2015; 15 May 2015; 18 June 2015; 15 July 2015; 15 August 2015; 17 September 2015; 20 October 2015; 15 November 2015; 18 December 2015; 10 January 2016; 14 February 2016; 9 March 2016; 13 April 2016; 12 May 2016; 6 June 2016; 10 July 2016; 12 August 2016; 24 September 2016; 25 October 2016; 26 November 2016; 23 December 2016; 29 January 2017; 15 February 2017; 23 March 2017; 26 April 2017; 22 May 2017; 18 June 2017; 24 July 2017; 26 August 2017; 15 September 2017; 15 October 2017; 15 November 2017; 21 December 2017; 15 January 2018; 15 February 2018; 21 March 2018; 21 April 2018; 15 May 2018; 15 June 2018; 20 July 2018; 10 August 2018; 15 September 2018; 17 October 2018; 16 November 2018; 17 December 2018; 14 January 2019; 8 February 2019; 14 March 2019; 28 April 2019; 27 May 2019; 27 June 2019; 2 July 2019; 4 August 2019; 23 September 2019; 24 October 2019; 24 November 2019; 21 December 2019; 18 January 2020; 24 February 2020; 15 March 2020; 23 April 2020; 15 May 2020; 15 June 2020; 18 July 2020; 21 August 2020; 15 September 2020; 15 October 2020; 15 November 2020; 18 December 2020; 20 January 2021; 16 February 2021; 8 March 2021; 8 April 2021; 21 May 2021;

B Machine Learning and Satellite Data

B.1 Additional elements on machine learning and income

Global Data Lab provides data on the per capita income for the Pakistani provinces. Starting from this aggregate, we reconstruct the provincial gross national income redistributing this measure at the tehsil level. We perform this exercise through the population data described in the Data and Methodology section.

Data on per-capita income is available from 2010 to 2018, but our satellite observations start from 2012 (when VIIRS sensors for night-lights became its transmissions). We, thus, produce a final dataset on local income at the tehsil and yearly level from 2012 to 2018. To finalize our training dataset, we also aggregate in the same way all the potential predictors:

- Electricity; [3]

- Landcover; [4]

- Local Prices; [5]

- Natural Observations; [6]

- Night-lights;

- Population;

- Weather; [7]

To handle the correlated predictors, we decide to maintain only the most influent on local income. We also include a set of dummies for tehsil, districts, and provinces to capture time-invariant geographical features and absorbing the effect of different practices in data production. It is worth noting as we do not include an indicator for years: we want our algorithm to learn directly from the data without following a specific temporal path.

[3] Commercial Consumption, Domestic Consumption, Industrial Consump- tion, Other Consumption

[4] Bare Areas, Crop Irrigated, Crop Rainfed, Grassland, Lichens and Mosses, Mosaic Crop, Mosaic Herbaceous Cover, Mosaic Tree and Shrubs, Mosaic Vegetation, Shrubland, Shrubs or Herbaceous Flooded, Sparse Vegetation, Tree Broadleaved Decidu- ous, Tree Broadleaved Evergreen, Tree Cover Flooded Fresh, Tree Cover Flooded Saline, Tree Mixed Leaf Type, Tree Needleaved Deciduous, Tree Needleaved Evergreen, Perman- ent Snow and Ice Urban Areas, Water Bodies

[5] Price in Pakistani Rupees for 486 Items

[6] Aerosol Thickness, Clouds Fraction, Clouds Optical Thickness, Clouds Particle Radius, Clouds Water Content, Night Fires, No2, Ozone, Sea Chlorophyll, Sea Surface Temperature, Snow on the Ground, Surface Temperature Anomaly Day, Surface Temperature Anomaly Night, Surface Temperature Day, Surface Temperature Night, Ve- getation Index (NDVI), Water Vapor

[7] Cloud Cover, Dew Point, Feels Like Temperature, Heat Index, Humidity, Max Temperarure, Min Temperature, Moon Illumination, Pressure, Sun Hours, Total Snow, Total Precipitations, UV Index, Visibility, Wind Chill Temperature, Wind Degree, Wind Gust Speed, Wind Speed

Then, we rescale our data between 0 and 1 through a min-max normalization. Such an operation is a mandatory step given the heterogeneous scale of the observed variables, combined with an absence of priors on their final distributions. We also convert to dummies our categorical variables (district, tehsil, province, electric company).

In the prediction exercise, we compare the performance of the following machine learning algorithms:

• elastic-net: it adopts a classical linear regression, coping with the presence of multiple predictors through a shrinkage parameter. While the lasso and the ridge refer to two opposite reduction formulas, this model finds the best combinations among these criteria. After this step, it uses another cross-validation to identify the optimal weight for the final predictors;

• random forest: based on trees, this model assumes a series of sequential decisions using the most predictive variable at each step. The random forest proposes only a set of the predictors in every decision node, performing its task on the entire dataset;

• bagging: this is another model based on decision trees, but it differs from the random forest by proposing all the predictors and a bootstrap of the original data at every decision node;

• boosting: it is the last model based on decision trees. In every node, it chooses among all the predictors and attributes a greater weight in the sample constitution to the observations wrongly classified in precedent steps;

• support vector machine: it draws hypothetical hyperplanes in variables' space to categorize the observations into final predictions. The criteria behind the hyperplane realization define its type: the linear opts for a parametric approach, the polynomial switches to a non-parametric one, and the radial basis function opts for a k-nearest neighborhood algorithm.

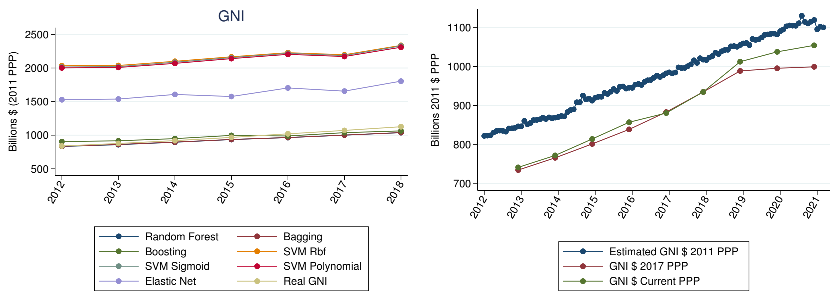
We do not implement more complex algorithms to permit an easy and fast replication of our policy toolkit. Besides, we tune every algorithm through a randomized grid search. Table 4 reports the mean square error of our models, while Figure 3 plots their estimates and our real data.

Table 4: Mean Square Error

	(1)
Models	Mean Square Error
bagging	0.01506
boosting	0.01532
elastic net	0.01606
random forest	0.01509
svm - sigmoid kernel	0.01566
svm - polynomial kernel	0.01562
svm - rbf kernel	0.01568

Notes: This table reports the mean square error for the algorithms employed in our study. The abbreviation "svm" stands for support vector machine, while "rbf" means radial basis function.

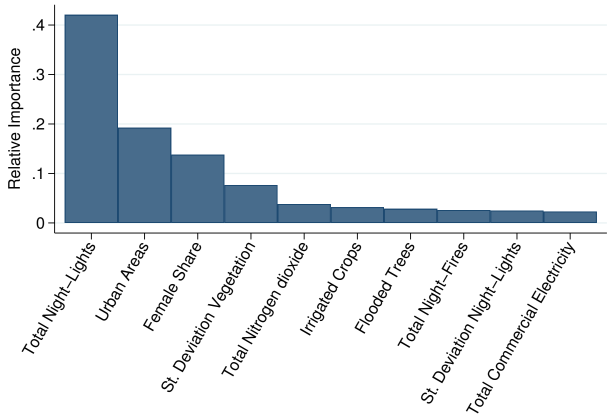
Figure 3: Algorithms' Comparison



Notes: The first graph (starting from the left) plots the estimated yearly GNI and the reference values as "Real GNI". The abbreviation SVM stands for "Support vector machine", while Rbf indicates "Radial Basis Function". The second graph compares the GNI estimated with our best performing model (bagging) with the yearly data provided by the World Bank. For these time series, we report the value on the last month of the year.

The bagging obtains the lowest mean square errors, replicating well the original data. Therefore, we adopt it as the final algorithm of our study. We make new predictions on the monthly dataset, having trained the model only on raw economic and satellite observations. Our algorithm has not pre-imposed temporal and seasonal trends. Figure 4 displays the ten most influential predictors in explaining local income.

Figure 4: Main Predictors



Notes: The figure proposes the relative importance of the ten main predictors of our bagging model. Their influence is rescaled on a 0 to 1 basis, for providing a better graphical visualization.

B.2 An introduction to satellite data

Satellites are ever more present in academic and non-academic research as a massive amount of high-resolution and high-frequency data become available. Once in orbit, satellites can automatically collect and pre-process data through automated algorithms. This procedure permits bypassing the physical barriers of the traditional data-gathering operations (high-costs, inaccessible areas, continuity in the project). Although, their innovative nature still poses some obstacles to their mainstream use. This short appendix provides the basic concepts behind this technology.

Almost all the satellite products come in the raster format: simple images enclosing the globe in a georeferenced grid. Every quadratic cell of this grid is composed of pixels. The resolution is expressed as the side length of every cell. The smaller is the side, the more are the pixels contained for each meter (the resolution). There is not a common standard for this unit of measure. Some companies report it in meters, while others in degrees. A good conversion rule is assuming $0.1° = 11.1$ km. Going a step deeper into the subject, we must understand how the pixel can reproduce a figure on our screen. This passage is essential for the storage and extraction of the collected data.

Each pixel contains three phosphors emitting red, blue and green light. Being able to work under different intensities and combinations, these latter can recreate all the existing palette of colors. When all of the phosphors are off, we visualize a black pixel. Oppositely, when they are all at the maximum intensity, we observe a white element. The satellite recreates a picture of the planet in pixels, assigning to each the measured values in the form of pixel luminosity. The highest observations will correspond to white areas, while the null to black ones. All the values left in the middle will assume several shades of gray.

Our Python algorithms exploit this technology to retrieve back the data. Indeed, they overlaid the satellite rasters with the shapefile of the studied areas to produce summary statistics on the underlying pixels. This procedure is the standard for the continuous variables, but the satellite can also assign arbitrary values to map categorical variables (like the landcover). In this framework, the Pyhton algorithm counts the recurrence of a determined value in a given shapefile.

In addition, a common practice consists of cutting off the studied area from the entire raster. This operation has two main goals:

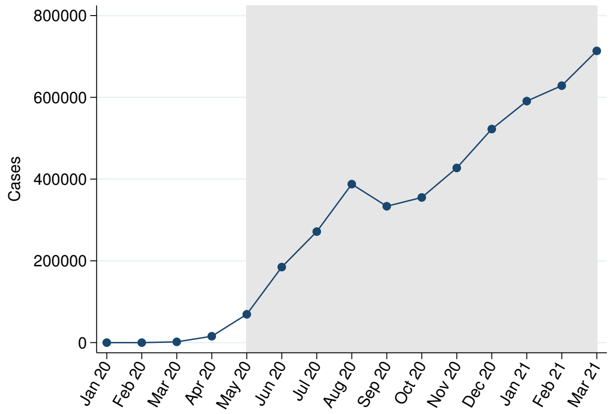
- producing illustrative material;
- reducing the computation time and the storing space;

The most popular format for raster files is the "GeoTIFF" extension. Although, some satellite products use "NetCDF" files containing several rasters at once. Another popular format is the new "HDF5", offering a wide range of enriched functions. It is also worth noting as some rasters represent only a fraction of the planet. This strategy allows better handling of this massive amount of data.

Stata has no functions to compute these resources. This shortage can be related to the high-processing performances required by these files. Thus, we must resort to more advanced open-source software. Among them, the most employed by economists are Python, R and Julia. It is also possible to work with QGIS or ArcGis. This software runs on Python, offering a more user-friendly interface. Unfortunately, it is more complex in organizing loops and storing the commands in reproducible code.

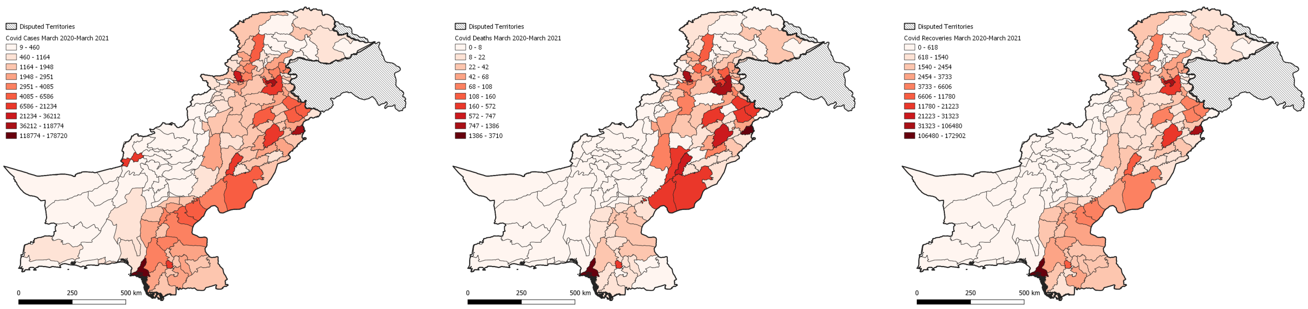
C Additional Figures

Figure 5: National Epidemic Curve



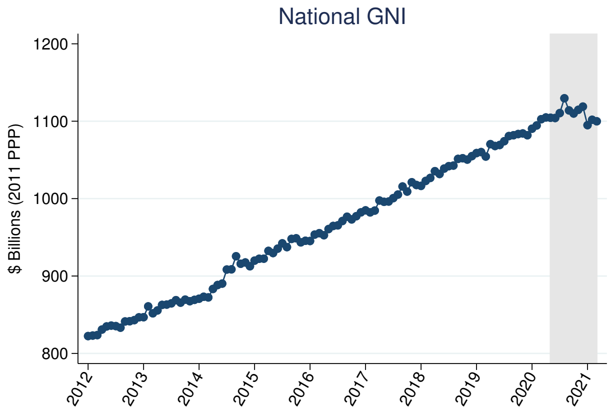
Notes: The present graph displays the monthly cases of Covid19 reported in Pakistani districts, between January 2020 and March 2021. The gray area indicates the temporal framework covered by our "Dummy Covid": May 2020 - March 2021.

Figure 6: Covid Statistics by Districts



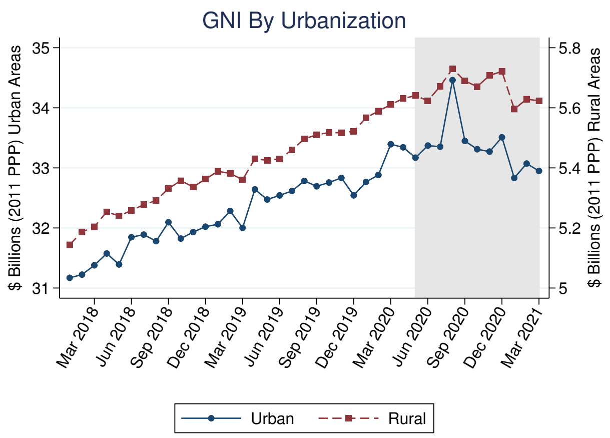
Notes: These three maps report the total Covid19 cases (left), deaths (central) and recoveries (right) for the Pakistani districts between March 2020 and March 2021.

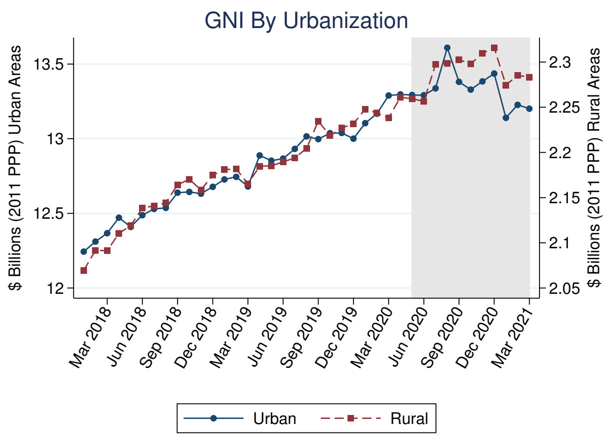
Figure 7: Gross National Income 2012-2021



Notes: This plot reports the Pakistani Gross National Income from January 2012 to March 2021. All the values are expressed in billions of 2011 PPP dollars. The gray area indicates the temporal framework covered by our "Dummy Covid": May 2020 - March 2021.

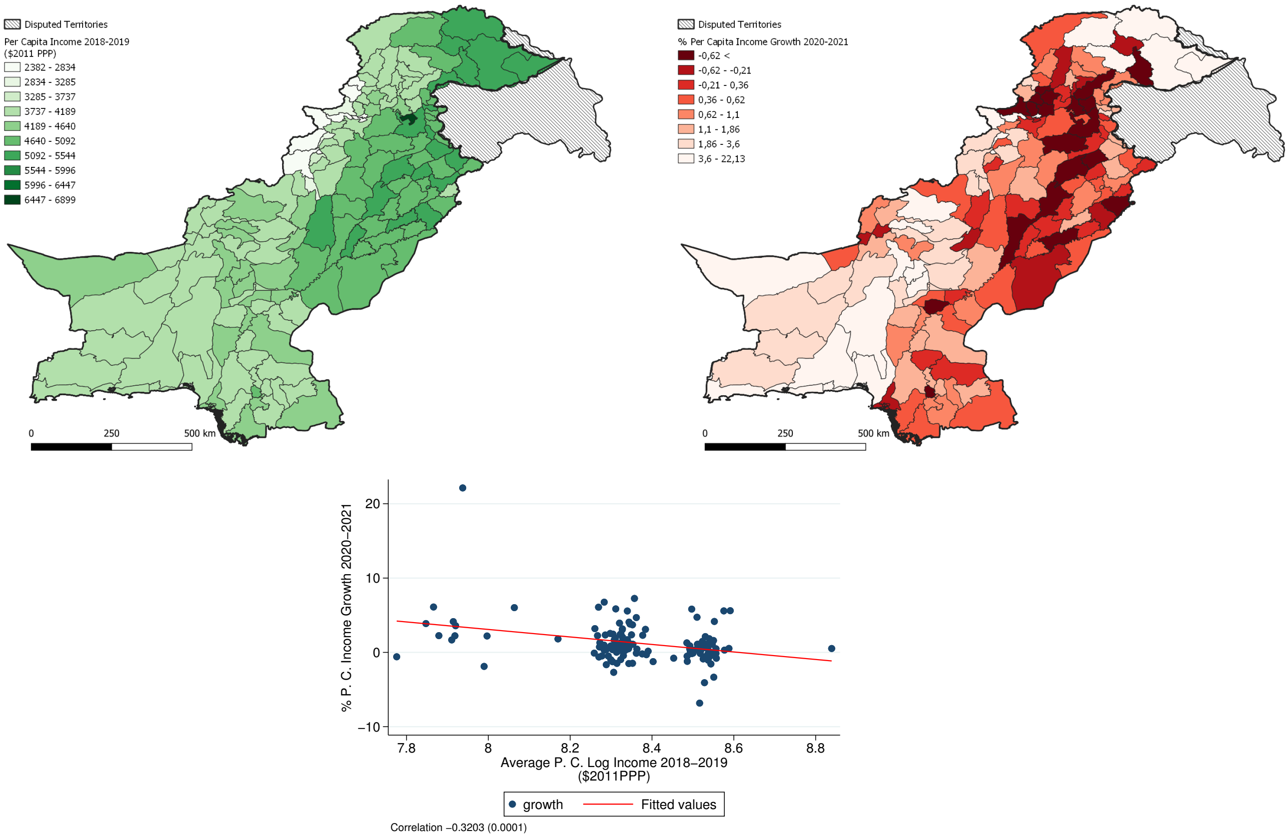
Figure 8: Urban vs Rural GNI: Different Aggregates





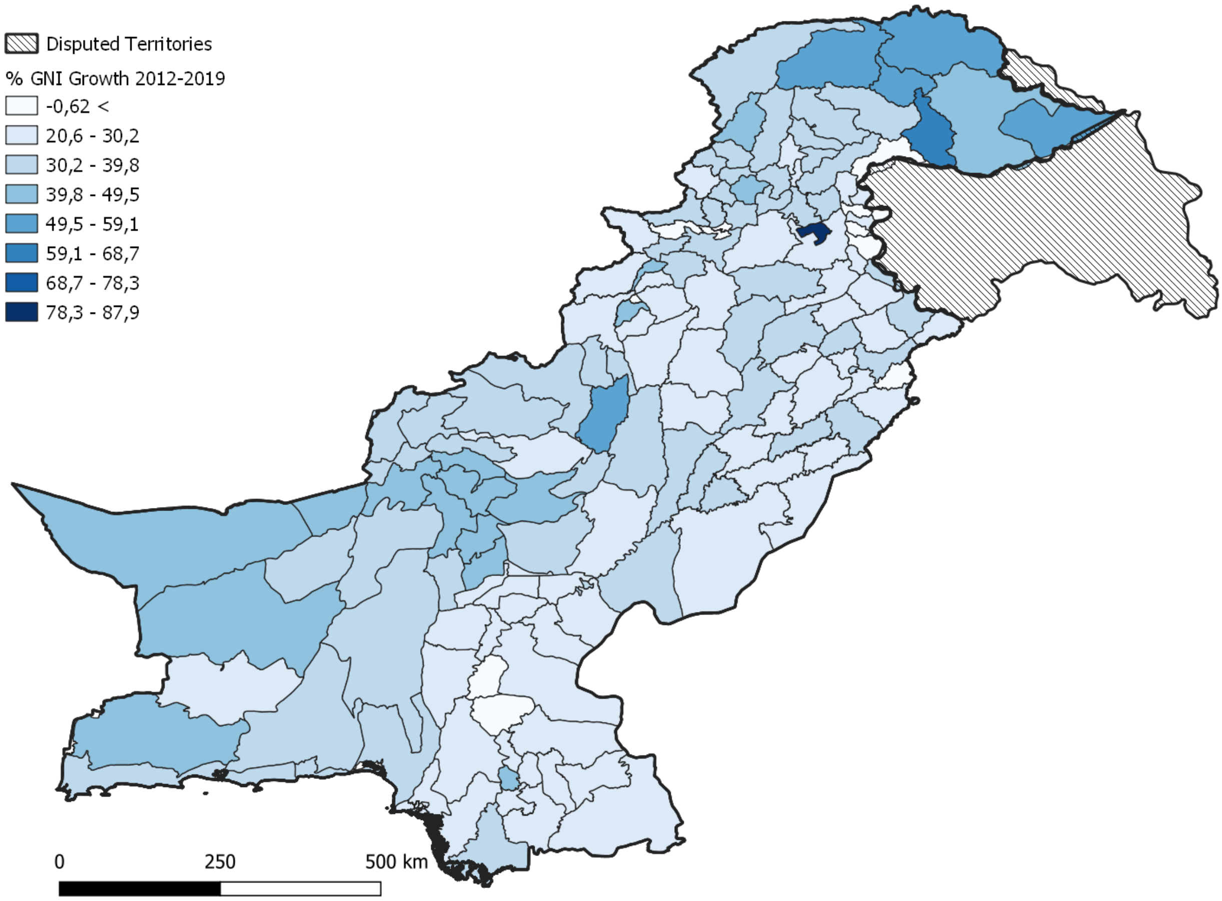
Notes: These plots decompose the average gross national income between rural districts in red (RHS vertical-axis) and urbanized districts in blue (LHS vertical-axis). The left-hand-side graph classifies as urban the ten most urbanized districts, while the right-hand-side extend this classification to seventy administrative units. All the values are expressed in billions of 2011 PPP dollars.

Figure 9: Average Income Per Capita and Income Per Capita Growth by District



Notes: The first map reports the average district-wise per-capita income between 2018 and 2019, while the second displays its percentual variation between 2020 and 2021. For 2021, we consider only the first three months for which data is available. All the values are expressed in billions of 2011 PPP dollars. The graph reports the district per capita income in 2018-2019 on the horizontal axis and the percentage variation between 2020-2021 on the vertical one. The linear relation between these variables is reported in red, while the correlation is noted at the bottom.

Figure 10: Growth by District 2012-2019



Notes: This map displays the GNI's growth rate between 2012 and 2019, for the 147 Pakistani districts. The yearly GNI is computed from the monthly average.

D Additional Tables

Table 5: Summary Statistics

Variables	(1) Observations	(2) Mean	(3) St. Deviation	(4) 50th P.tile	(5) 5th P.tile	(6) 95th P.tile
Log Income	5,733	21.93	1.368	22.03	19.91	23.94
Growth Income	5,733	0.297	2.556	0.195	-2.505	3.324
Standardize Income 2012	5,733	-4.34e-10	1.000	-0.358	-0.674	1.625
Log Income 2012	5,733	21.66	1.389	21.77	19.63	23.64
Log Income Per Capita	5,733	8.374	0.170	8.358	7.949	8.571
Growth Income Per Capita	5,733	0.175	2.536	0.0306	-2.490	3.220
Log Domestic Electricity	5,733	11.54	7.577	14.90	-2.303	18.25
Growth Domestic Electricity	4,498	2,783	73,767	-0.108	-49.29	110.0
Log Commercial Electricity	5,733	10.11	6.923	13.24	-2.303	16.49
Growth Commercial Electricity	4,477	79.72	3,264	-0.495	-41.28	60.37
Log Industrial Electricity	5,733	10.61	7.326	13.92	-2.303	17.75
Growth Industrial Electricity	4,467	44.19	744.2	-0.546	-52.21	96.79
Log Other Electricity	5,733	11.26	7.180	14.48	-2.303	17.40
Growth Other Electricity	4,549	208.2	3,912	1.194	-70.38	271.3
Log Night-Lights Sum	5,733	9.413	2.712	10.08	4.949	12.47
Growth Night-Lighs Sum	5,609	179.6	3,660	-1.288	-84.73	329.8
Log Night-Lights St. Deviations	5,733	1.654	1.528	1.894	-1.048	3.878
Growth Night-Lights St. Deviations	5,609	60.81	961.0	-1.254	-77.40	222.2
Log Night-Lights Mean	5,733	0.213	1.745	0.295	-2.245	3.178
Growth Nigth-Lights Mean	5,609	191.2	3,638	-0.632	-84.90	342.6
Log Night-Lights Max	5,733	5.235	1.806	5.412	2.715	7.745
Growth Night-Lights Max	5,609	95.95	1,442	0	-81.88	330
Dummy Covid	5,733	0.282	0.450	0	0	1
Covid Cases	5,733	785.3	6,565	0	0	2,073
Covid Deaths	5,733	19.91	158.4	0	0	56
Covid Recoveries	5,733	681.2	5,946	0	0	1,721
Log Cases	5,733	0.144	3.909	-2.303	-2.303	7.637
Log Deaths	5,733	-1.159	2.270	-2.303	-2.303	4.027
Log Recoveries	5,733	-0.189	3.798	-2.303	-2.303	7.451
Year	5,733	2,019	0.948	2,019	2,018	2,021
Month	5,733	6.154	3.534	6	1	12
District	5,733	74	42.44	74	8	140
Dummy Urban	5,733	0.136	0.343	0	0	1

Notes: This table reports the summary statistics of all the variables considered in this study. The variable "Log Income" represents the logarithm of the district gross income, while "Income Growth" is the percentage variations between months. A similar approach is also adopted for reporting electricity consumption of domestic, commercial, industrial, and others users. The electricity growth has fewer observations, being constant to zero in districts subsequently connected to the grid. For the night-lights, we indicate the logarithm and the growth of several statistical aggregates observed at the district-month level (mean, sum, max and standard deviation). The "Dummy Covid" assumes the value 1 from May 2020 to March 2021. "Covid Cases", "Covid Deaths" and "Covid Recoveries" are set to zero for the months preceding the pandemic. We also report the logarithm of these variables. We use "Dummy Urban" for indicating the districts containing the first twenty metropolitan areas of the country (Karachi, Lahore, Faisalabad, Gujranwala, Rawalpindi, Peshawar, Multan, Hyderabad, Sialkot, Bahawalpur, Islamabad, Quetta, Rahim Yar Khan, Sheikhupura, Sargodha, Attock, Sukkur, Larkana, Swat, Muzaffargarh). The dataset follows 147 districts from January 2018 to March 2021.

Table 6: Covid and Urban Areas

Variables	(1) Cases	(2) Deaths	(3) Recoveries
$Urban_d$	1.038***	1.091***	1.146***
	(0.113)	(0.155)	(0.209)
District FE	No	No	No
Year FE	No	No	No
Month FE	No	No	No
Obs.	5733	5733	5733
Adj. R sq.	0.00811	0.0270	0.0105
Mean Dep. Var.	0.144	-1.159	-0.189
S.D. Dep. Var.	3.909	2.270	3.798

Notes: In this table, we estimate the difference in log cases (column 1), log deaths (column 2) and log recoveries (column 3) between rural and urban areas. The study follows all the 147 Pakistani districts from January 2018 to March 2021. We do not control for fixed effects. Standard errors are clustered at the district level. The number of observations and adjusted R2, indicated as Adj. R sq., of each regression are reported at the end of the table. The last row presents the mean of the dependent variable (Mean Dep. Var.). ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Table 7: Covid Impact on GNI Growth - Robustness Check with District FE

Panel A - Overall

	(1)	(2)	(3)	(4)
Variables		Income Growth		
$Covid_{my}$	0.0625			
	(0.0636)			
$Cases_{dmy}$		-0.0109		
		(0.00851)		
$Deaths_{dmy}$			-0.0409***	
			(0.0130)	
$Recoveries_{dmy}$				-0.0349***
				(0.00919)
District FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
Obs.	5733	5733	5733	5733
Adj. R sq.	-0.000889	-0.000852	-0.000369	-8.48e-06
Mean Dep. Var.	0.297	0.297	0.297	0.297
S.D. Dep. Var.	2.556	2.556	2.556	2.556

Panel B - Urban

	(1)	(2)	(3)	(4)
Variables		Income Growth		
$Covid_{my}$	0.0918			
	(0.0661)			
$Covid_{my} \times Urban_d$	-0.216***			
	(0.0596)			
$Cases_{dmy}$		-0.00562		
		(0.00901)		
$Cases_{dmy} \times Urban_d$		-0.0147**		
		(0.00687)		
$Deaths_{dmy}$			-0.0363**	
			(0.0155)	
$Deaths_{dmy} \times Urban_d$			-0.0111	
			(0.0125)	
$Recoveries_{dmy}$				-0.0330***
				(0.0101)
$Recoveries_{dmy} \times Urban_d$				-0.00610
				(0.00872)
District FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
Obs.	5733	5733	5733	5733
Adj. R sq.	-0.000894	-0.000947	-0.000530	-0.000174
Mean Dep. Var.	0.297	0.297	0.297	0.297
S.D. Dep. Var.	2.556	2.556	2.556	2.556

Notes: Panel A estimates the impact on the income's growth rate of the pandemic months (column 1), the log cases (column 2), the log deaths (column 3) and the log recoveries (column 4). Panel B repeats this analysis decomposing the impact between rural and urban areas. The study follows all the 147 Pakistani districts from January 2018 to March 2021 We control for the district, year and month fixed effects. Standard errors are clustered at the district level. The number of observations and adjusted R2, indicated as Adj. R sq., of each regression are reported at the end of the table. The last row presents the mean of the dependent variable (Mean Dep. Var.). ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Table 8: Covid Impact on GNI, 2018 - 2021

Panel A - Overall

	(1)	(2)	(3)	(4)
Variables		Log Income		
$Covid_{my}$	0.000332			
	(0.00270)			
$Cases_{dmy}$		-0.000360		
		(0.000324)		
$Deaths_{dmy}$			-0.00165***	
			(0.000542)	
$Recoveries_{dmy}$				-0.00115***
				(0.000336)
District FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
Obs.	5733	5733	5733	5733
Adj. R sq.	1	1	1	1
Mean Dep. Var.	21.93	21.93	21.93	21.93
S.D. Dep. Var.	1.368	1.368	1.368	1.368

Panel B - Urban

	(1)	(2)	(3)	(4)
Variables		Log Income		
$Covid_{my}$	0.00209			
	(0.00300)			
$Covid_{my} \times Urban_d$	-0.0129**			
	(0.00569)			
$Cases_{dmy}$		0.0000848		
		(0.000364)		
$Cases_{dmy} \times Urban_d$		-0.00124**		
		(0.000564)		
$Deaths_{dmy}$			-0.00123*	
			(0.000642)	
$Deaths_{dmy} \times Urban_d$			-0.00102	
			(0.000881)	
$Recoveries_{dmy}$				-0.000926**
				(0.000410)
$Recoveries_{dmy} \times Urban_d$				-0.000721
				(0.000592)
District FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
Obs.	5733	5733	5733	5733
Adj. R sq.	1	1	1	1
Mean Dep. Var.	21.93	21.93	21.93	21.93
S.D. Dep. Var.	1.368	1.368	1.368	1.368

Notes: Panel A estimates the impact on the logarithm income of the pandemic months (column 1), the log cases (column 2), the log deaths (column 3) and the log recoveries (column 4). Panel B repeats this analysis decomposing the impact between rural and urban areas. The study follows all the 147 Pakistani districts from January 2018 to March 2021 We control for the district, year and month fixed effects. Standard errors are clustered at the district level. The number of observations and adjusted R2, indicated as Adj. R sq., of each regression are reported at the end of the table. The last row presents the mean of the dependent variable (Mean Dep. Var.). ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Table 9: Covid Impact on GNI - Alternative Definitions of Urban Districts

Panel A - Principal Ten Urban Areas

	(1)	(2)	(3)	(4)
Variables	Log Income			
$Covid_{my}$	0.00145			
	(0.00287)			
$Covid_{my} \times Urban_d$	-0.0164**			
	(0.00675)			
$Cases_{dmy}$		-0.0000602		
		(0.000347)		
$Cases_{dmy} \times Urban_d$		-0.00137**		
		(0.000651)		
$Deaths_{dmy}$			-0.00129**	
			(0.000608)	
$Deaths_{dmy} \times Urban_d$			-0.00126	
			(0.000908)	
$Recoveries_{dmy}$				-0.000957**
				(0.000379)
$Recoveries_{dmy} \times Urban_d$				-0.000978
				(0.000627)
District FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
Obs.	5733	5733	5733	5733
Adj. R sq.	1	1	1	1
Mean Dep. Var.	21.93	21.93	21.93	21.93
S.D. Dep. Var.	1.368	1.368	1.368	1.368

Panel B - Principal Seventy Urban Areas

	(1)	(2)	(3)	(4)
Variables	Log Income			
$Covid_{my}$	0.00715			
	(0.00470)			
$Covid_{my} \times Urban_d$	-0.0143**			
	(0.00556)			
$Cases_{dmy}$		0.000889		
		(0.000617)		
$Cases_{dmy} \times Urban_d$		-0.00166**		
		(0.000705)		
$Deaths_{dmy}$			-0.000461	
			(0.000901)	
$Deaths_{dmy} \times Urban_d$			-0.00162	
			(0.00115)	
$Recoveries_{dmy}$				-0.000214
				(0.000756)
$Recoveries_{dmy} \times Urban_d$				-0.00134
				(0.000832)
District FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
Obs.	5733	5733	5733	5733
Adj. R sq.	1	1	1	1
Mean Dep. Var.	21.93	21.93	21.93	21.93
S.D. Dep. Var.	1.368	1.368	1.368	1.368

Notes: In this table, we estimate the impact on the logarithm income of the pandemic months (column 1), the log cases (column 2), the log deaths (column 3) and the log recoveries (column 4). Panel A proposes an interaction for estimating the differential impact in the ten most urbanized districts, while Panel B repeats this analysis enlarging this classification to seventy districts. The study follows all the 147 Pakistani districts from January 2018 to March 2021. We control for the district, year and month fixed effects. Standard errors are clustered at the district level. The number of observations and adjusted R2, indicated as Adj. R sq., of each regression are reported at the end of the table. The last row presents the mean of the dependent variable (Mean Dep. Var.). ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Table 10: Covid Impact on Per Capita Income, 2018 - 2021

Panel A - Overall

	(1)	(2)	(3)	(4)
Variables	Log Income Per Capita			
$Covid_{my}$	0.0000670			
	(0.00264)			
$Cases_{dmy}$		-0.00103***		
		(0.000342)		
$Deaths_{dmy}$			-0.00310***	
			(0.000583)	
$Recoveries_{dmy}$				-0.00233***
				(0.000367)
District FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
Obs.	5733	5733	5733	5733
Adj. R sq.	0.971	0.971	0.971	0.972
Mean Dep. Var.	8.374	8.374	8.374	8.374
S.D. Dep. Var.	0.170	0.170	0.170	0.170

Panel B - Urban

	(1)	(2)	(3)	(4)
Variables	Log Income Per Capita			
$Covid_{my}$	0.00263			
	(0.00294)			
$Covid_{my} \times Urban_d$	-0.0188***			
	(0.00579)			
$Cases_{dmy}$		-0.000472		
		(0.000371)		
$Cases_{dmy} \times Urban_d$		-0.00157***		
		(0.000560)		
$Deaths_{dmy}$			-0.00267***	
			(0.000689)	
$Deaths_{dmy} \times Urban_d$			-0.00104	
			(0.000873)	
$Recoveries_{dmy}$				-0.00209***
				(0.000433)
$Recoveries_{dmy} \times Urban_d$				-0.000774
				(0.000582)
District FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
Obs.	5733	5733	5733	5733
Adj. R sq.	0.971	0.971	0.971	0.972
Mean Dep. Var.	8.374	8.374	8.374	8.374
S.D. Dep. Var.	0.170	0.170	0.170	0.170

Notes: Panel A estimates the impact on the per capita income of the pandemic months (column 1), the log cases (column 2), the log deaths (column 3) and the log recoveries (column 4). Panel B repeats this analysis decomposing the impact between rural and urban areas. The study follows all the 147 Pakistani districts from January 2018 to March 2021 We control for the district, year and month fixed effects. Standard errors are clustered at the district level. The number of observations and adjusted R2, indicated as Adj. R sq., of each regression are reported at the end of the table. The last row presents the mean of the dependent variable (Mean Dep. Var.). ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Table 11: Covid Impact on Per Capita Income Growth, 2018 - 2021

Panel A - Overall

Variables	(1)	(2)	(3)	(4)
	Per Capita Income Growth			
$Covid_{my}$	0.0645			
	(0.0633)			
$Cases_{dmy}$		-0.00600		
		(0.00842)		
$Deaths_{dmy}$			-0.0294**	
			(0.0128)	
$Recoveries_{dmy}$				-0.0260***
				(0.00917)
District FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
Obs.	5733	5733	5733	5733
Adj. R sq.	0.0000628	0.0000525	0.000322	0.000547
Mean Dep. Var.	0.175	0.175	0.175	0.175
S.D. Dep. Var.	2.536	2.536	2.536	2.536

Panel B - Urban

Variables	(1)	(2)	(3)	(4)
	Per Capita Income Growth			
$Covid_{my}$	0.0885			
	(0.0659)			
$Covid_{my} \times Urban_d$	-0.177***			
	(0.0600)			
$Cases_{dmy}$		-0.00174		
		(0.00895)		
$Cases_{dmy} \times Urban_d$		-0.0119*		
		(0.00688)		
$Deaths_{dmy}$			-0.0245	
			(0.0151)	
$Deaths_{dmy} \times Urban_d$			-0.0121	
			(0.0121)	
$Recoveries_{dmy}$				-0.0242**
				(0.0101)
$Recoveries_{dmy} \times Urban_d$				-0.00612
				(0.00859)
District FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
Obs.	5733	5733	5733	5733
Adj. R sq.	0.00000218	-0.0000709	0.000165	0.000382
Mean Dep. Var.	0.175	0.175	0.175	0.175
S.D. Dep. Var.	2.536	2.536	2.536	2.536

Notes: Panel A estimates the impact on the per capita income's growth of the pandemic months (column 1), the log cases (column 2), the log deaths (column 3) and the log recoveries (column 4). Panel B repeats this analysis decomposing the impact between rural and urban areas. The study follows all the 147 Pakistani districts from January 2018 to March 2021 We control for the district, year and month fixed effects. Standard errors are clustered at the district level. The number of observations and adjusted R2, indicated as Adj. R sq., of each regression are reported at the end of the table. The last row presents the mean of the dependent variable (Mean Dep. Var.). ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Table 12: Convergence and Covid19 - Log of Income

Variables	(1)	(2)	(3)	(4)	(5)
	Income Growth				
$Income\ 2012_d$	-0.0406***	-0.0210*	-0.0210*		
	(0.0114)	(0.0123)	(0.0123)		
$Covid_{my}$		4.114***	4.253***	4.114***	4.253***
		(1.199)	(1.197)	(1.199)	(1.197)
$Income\ 2012_d \times Covid_{my}$		-0.196***	-0.196***	-0.196***	-0.196***
		(0.0542)	(0.0543)	(0.0542)	(0.0543)
District FE	No	No	No	Yes	Yes
Month FE	No	No	Yes	No	Yes
Year FE	No	No	Yes	No	Yes
Obs.	16170	16170	16170	16170	16170
Adj. R sq.	0.000410	0.00153	0.0129	-0.00553	0.00583
Mean Dep. Var.	0.324	0.324	0.324	0.324	0.324
S.D. Dep. Var.	2.600	2.600	2.600	2.600	2.600

Notes: Column (1) estimates the impact on the income growth rate of the logarithm mean income in 2012, without controlling for fixed effects. The remaining columns examine how the logarithm mean income in 2012, the pandemic period and the interaction between these terms influence the growth rate. The second column control for no fixed effect, the third for year and month fixed effects, the fourth for district fixed effect, the fifth for the district, month and year fixed effects. Standard errors are clustered at the district level. The study follows all the 147 Pakistani districts from January 2012 to March 2021. The number of observations and adjusted R2, indicated as Adj. R sq., of each regression are reported at the end of the table. The last row presents the mean of the dependent variable (Mean Dep. Var.). ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Table 13: Convergence and Covid19 - 2012 Income Terciles

Variables	(1)	(2)	(3)	(4)	(5)
	Income Growth				
$Tercile2\ Income\ 2012_d$	-0.130***	-0.0874***	-0.0874***		
	(0.0250)	(0.0252)	(0.0253)		
$Tercile3\ Income\ 2012_d$	-0.0945***	-0.0354	-0.0354		
	(0.0277)	(0.0281)	(0.0281)		
$Covid_{my}$		0.204**	0.343***	0.204**	0.343***
		(0.0945)	(0.103)	(0.0945)	(0.103)
$Tercile2\ Income\ 2012_d \times Covid_{my}$		-0.431***	-0.431***	-0.431***	-0.431***
		(0.0973)	(0.0974)	(0.0973)	(0.0973)
$Tercile3\ Income\ 2012_d \times Covid_{my}$		-0.591***	-0.591***	-0.591***	-0.591***
		(0.0991)	(0.0992)	(0.0991)	(0.0992)
District FE	No	No	No	Yes	Yes
Year FE	No	No	Yes	No	Yes
Month FE	No	No	Yes	No	Yes
Year FE#Month FE	No	No	No	No	No
Obs.	16170	16170	16170	16170	16170
Adj. R sq.	0.000324	0.00122	0.0125	-0.00575	0.00561
Mean Dep. Var.	0.324	0.324	0.324	0.324	0.324
S.D. Dep. Var.	2.600	2.600	2.600	2.600	2.600

Notes: Column (1) estimates the impact on the income growth rate of being in the second and third tercile of income during 2012, without controlling for fixed effects. The remaining columns examine how these dummies, the pandemic period and the interaction between them influence the growth rate. The second column control for no fixed effect, the third for year and month fixed effects, the fourth for district fixed effect, the fifth for the district, month and year fixed effects. Standard errors are clustered at the district level. The study follows all the 147 Pakistani districts from January 2012 to March 2021. The number of observations and adjusted R2, indicated as Adj. R sq., of each regression are reported at the end of the table. The last row presents the mean of the dependent variable (Mean Dep. Var.). ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

The International Growth Centre (IGC) aims to promote sustainable growth in developing countries by providing demand-led policy advice based on frontier research.

Find out more about our work on our website
www.theigc.org

For media or communications enquiries, please contact
mail@theigc.org

Subscribe to our newsletter and topic updates
www.theigc.org/newsletter -signup

Follow us on Twitter
@the_igc

Contact us
International Growth Centre, London School of Economic and Political Science, Houghton Street, London WC2A 2AE

IGC
International Growth Centre

DIRECTED BY

LSE UNIVERSITY OF OXFORD

FUNDED BY

UKaid from the British people

Designed by soapbox.co.uk