JUE Insight: Inequality in India declined during COVID, but for different reasons in urban and rural areas

Arpit Gupta, Anup Malani, Philip Monagan, & Bartosz Woda*

September 27, 2022

Abstract

We use a large, representative panel data set from India with monthly data on household finances to examine the incidence of economic harms during the COVID pandemic. We observe a sharp spike in poverty and inequality, especially in urban areas, during India's sharp but short lockdown. However, there was a striking decrease in income inequality after the lockdown. Evidence supports two mechanisms for the decline in income inequality. In rural areas, the capital income of top-quartile earners covaries more with aggregate income. In urban areas, demand for labor fell more for higher income earners, due in part to their occupational specialization in services.

*Gupta: New York University; Malani: University of Chicago and NBER; Monagan: University of Chicago; Woda: Amazon. Correspondence: amalani@uchicago.edu. We thank Satej Soman, Sabareesh Ramachandran, Kaushik Krishna and Mahesh Vyas for help with Consumer Pyramids Household (CPHS) data. We thank Mushfiq Mobarak, Cynthia Kinnan, Matt Notowodigdo, Viral Acharya, Satej Soman, Sabareesh Ramachandran, and seminar participants at the University of Chicago Micro Workshop, IIM Calcutta-NYU Stern India Research Conference, CAFRAL, the CPHS Research Seminar and the University of Chicago Law School for comments. We thank Sabareesh Ramachandran and Satej Soman for assistance with COVID data. Malani acknowledges funding from the Becker-Friedman Institute at the University of Chicago to purchase a subscription to the Consumer Pyramids Household Survey and the support of the Barbara J. and B. Mark Fried Fund at the University of Chicago Law School.

There is growing evidence that the COVID pandemic was associated with greater poverty within low- and middle-income countries like India (e.g., Miguel and Mobarak, 2021; Egger et al., 2021; Gupta et al., 2021; Malani et al., 2020). A common inference is that COVID was therefore also regressive (Oxfam, 2021; Azim Premji University, 2021).

We directly examine inequality during the pandemic in India, within urban and rural areas, using a large, representative, panel data set of roughly 197,000 households (990 thousand members) with monthly data from January 2015–July 2021. Our data have advantages over other data typically used to study poverty and inequality in India or the U.S. Our data are longitudinal, allowing us to decompose income changes into movements across income quantiles and changes in income at particular quantiles. The data include income and consumption, so we can compare income to consumption inequality. Finally, a breadth of covariates permits us to explore the mechanisms behind shifts in inequality.

Our most notable finding is that the pandemic in India was associated with a decline in inequality in two senses. First, Indians from higher percentiles of the income distribution from 2015–2019 had, almost monotonically, larger relative reductions in income during the pandemic. Second, consumption inequality also fell. However, it was lower to start and contracted less than income inequality, likely due to consumption smoothing.[1]

We explore the mechanisms for the decline in income inequality at two levels: the nature of the inequality and the sources of income. On the first level, we find that the decline in inequality is driven less by changes in the share of income at different percentiles of the income distribution (i.e., the Gini coefficient) than by spikes in social mobility of households across percentiles (measured by Sharrock's index). This spike in social mobility was greater in rural India.

On the second level, we examine two components of income, corresponding loosely to business income and labor income. First, declining business income may explain the decline in inequality in rural India. The business profits of top-quartile individuals vary more with aggregate income, making high-income households more susceptible to the downturns during the pandemic. Second, lower labor demand can explain lower inequality in urban India. The income of higher-income individuals is more sensitive to demand for services (as opposed to agriculture or manufacturing) and demand for services fell more during the pandemic.[2]

[1] Our findings are aligned with Malani and Ramachandran (2021), which finds that excess deaths during COVID disproportionately affected high incomes.

[2] We also document that remote-amenable work (Dingel and Neiman, 2020) did not protect income of

Our objective in documenting the decline in inequality during the pandemic is not to minimize the burdens experienced by the poor. Indeed, we begin our analysis by showing the stark increase in poverty during India's lockdown, especially in urban areas. However, measuring poverty and inequality separately can help us understand the relationship between these two phenomena. Our results suggest that increases in poverty are not a sufficient statistic for inequality and, conversely, distributional effects may be an incomplete measure of genuine shifts in welfare.

Our analysis also has implications for income dynamics during lockdowns. Especially in urban areas, poverty and the Gini coefficient spiked sharply during India's lockdown, which well preceded the spikes in cases during India's two main initial COVID waves. The main reductions in income inequality took place outside the lockdowns, when voluntary distancing mediated economic activity.

An important caveat to our findings is that we only demonstrate the coincidence of the pandemic and changes in poverty and inequality, not a causal connection between the two. There was a stark change in outcomes during the lockdown, consistent with a causal interpretation based on the time series trend. However, there was a decline in income inequality that began in 2018, before the pandemic, suggesting a pre-trend. Nor can we say one way or another whether inequality will rise when the pandemic ends.

Our work contributes to three literatures. The first addresses inequality in India in general (Banerjee and Piketty, 2005; Chancel and Piketty, 2019; Anand and Thampi, 2016; Sarkar and Mehta, 2010). Our contribution is to employ data that is relatively novel to this literature and that enables a deeper analysis of trends in inequality.

The second literature examines poverty and inequality during the pandemic (Miguel and Mobarak, 2021). Several papers examine economic outcomes (Alstadsæter et al., 2020; Chen et al., 2020) and inequality (Egger et al., 2021) in other countries. Bertrand et al. (2020) and Deshpande (2020) use the same data we do to study inequality and hours worked, respectively, during India's lockdown. Several papers use different Indian data to study income during COVID (Dhingra and Machin, 2020; Lee et al., 2020; Pinto et al., 2020). Most of the papers above suggest COVID had regressive impacts. One exception is Choudhary et al. (2021) which finds, using nighttime light data, a greater impact of the pandemic on urban areas in Pakistan leading to growth convergence. We highlight these effects hold in India using household-level data, and describe some mechanisms to explain them. Another notable exception is Deaton (2021), which argues that inequality

the rich in India.

may fall across countries during the pandemic; we show that within-country inequality may have fallen, at least within India. Another exception is Scheidel (2018), which argues that plagues have historically reduced inequality by disproportionately impacting the rich.

The third literature examines the sensitivity of income to business cycle shocks across the income distribution (Guvenen et al., 2017, 2015; Parker and Vissing-Jorgensen, 2010). These papers focus on high-income countries, while ours on a developing country.

1 Data and methods

1.1 Background

India is a valuable setting to study the incidence of harms during the pandemic because of India's sizable population and the large magnitude of the pandemic shock in that country (Appendix Figure A1). Moreover, the economic consequences of the pandemic differed in rural and urban areas.

India's first COVID case was on 27 January 2020 (Andrews et al., 2020). After several weeks of gradual restrictions on travel, India declared a national lockdown on 24 March 2020, but lifted it by 1 June 2020 (Sheng et al., 2021). The lockdown and the concomitant decline in mobility was one of the most severe observed in the world at the time (Hale et al., 2020; Google LLC, 2021).

The national lockdown preceded the peak of the health shock. The first wave of cases did not peak until September–October 2020. A second wave, due to the Delta variant (Organization, 2021), peaked in April 2021. While local lockdowns returned and mobility declined during the second wave, the mobility decline was less severe compared to the national lockdown. Although official statistics suggest that 3.1 million were infected and 410,000 died with COVID, serological surveys and measures of excess deaths suggest perhaps 1.1 billion were infected with COVID and 5 million died by mid-July 2021 (Malani and Ramachandran, 2021). The period examined in this study does not include India's third wave, due to Omicron, which occured in January 2022.

1.2 Data

Our primary data source is the Consumer Pyramids Household Survey (CPHS), conducted by Center for Monitoring Indian Economy. CPHS covers nearly the whole of

India and employs stratified sampling to ensure representativeness down to a substate level (details in Appendix C). CPHS surveys each household every four month, and sampling is staggered so that a representative 25% of all households are sampled each month. At each survey, CPHS asks about income and consumption for each of the previous 4 months. Income is obtained at both the individual (individual wage income) and household level (non-attributed income for the whole household). Consumption expenditure is surveyed at the household level.

CPHS has several advantages relative to comparable surveys in the American context. First, CPHS is very large relative to panels such as the Panel Study of Income Dynamics (PSID) or the Survey of Consumer Finance (SCF), measured either in absolute numbers (197,000 Indian households compared to 6,000 households in the SCF) or in the percentage of the population (0.073% vs. 0.004%). Second, CPHS provides greater ability to conduct subnational analysis relative to the PSID. Third, CPHS is a true 6-year panel data set, not a repeated cross-section, rotating panel like the Consumer Expenditure Survey (CEX) or the Current Population Survey (CPS), or an infrequent panel like the SCF. Fourth, CPHS has data on both income and consumption, unlike credit report data; moreover income is broken down by source, unlike credit report data. Prior research has highlighted (Bairoliya et al., 2021) that the CMIE data is broadly representative of the levels and growth in consumption expenditure found in other Indian administrative surveys.

In Appendix C, we address known issues in the CPHS data, such as churn in the sample frame, falling response rate during India's COVID lockdown (Vyas, 2020a), and criticism over whether the sample adequately captures the very poor (Dreze and Somanchi, 2021).

1.3 Methods

Poverty. We measure poverty by calculating the fraction of households with income or consumption below the World Bank's $1.90 cutoff for extreme poverty. We use contemporaneous PPP-adjusted exchange rates from the IMF for currency conversions.

Per capita income and consumption indices. Each household member's income is obtained by dividing total household income by the household size. Consumption is defined similarly. We use this measure of income (rather than individual wage income plus per member non-attributed household income) to ensure all household members are assigned to the same income percentile, as we are not investigating intra-family inequality. We use inflation data from the Reserve Bank of India to convert nominal value to real 2012

rupees (Bhoi et al., 2020). We create income indices by dividing monthly income by average 2018 income for the same individual, and comparably define consumption indices. We chose 2018 because it is roughly in the middle of the panel but permits evaluation of inequality trends even prior to the pandemic.

Quantile assignment. To assign individuals to income quantiles, we use their average 2015–2019 per capita incomes. Quantiles are defined as a default at the state × urban status level.[3] (We define quantiles at the national × urban status level in robustness checks reported in the 2.) We calculate weighted-average income in these local income percentiles using CPHS individual-member weights.

Employment and wages. The employment rate is defined as the fraction of persons aged 15–65 reporting they were employed.[4] The number of hours at work is obtained from the CPHS time-use survey for the same persons. Wages are derived by dividing per-capita income by hours.

Pandemic-relevant periods. We define the pandemic period as February 2020–July 2021 (the last date for which data are currently available), inclusive. We define the lockdown period as March–May 2020. The post-lockdown period is June 2020–July 2021. India experienced two initial waves of cases during the pandemic. The first wave ranges from February 2020–December 2020, and the second wave spans January–July 2021.

Non-pandemic counterfactual. We use a regression of outcomes (relative to some pre-pandemic baseline) on month × quartile fixed effects to make inferences about how economic outcomes changed over time during the pandemic:

$$y_{i,t} - \hat{y}_{i,t} = \sum_{q} \sum_{s} \gamma_s^q \mathbb{1}(1 \in q) + e_{it} \tag{1}$$

where $y_{i,t}$ is an outcome such as income at t relative to the average income in 2018; $\hat{y}_{i,t}$ is predictions of the outcome in the absence of the pandemic (details below); q indexes quantiles; γ_s^q are quantile-specific time fixed effects, and e_{it} is a regression error that is clustered at the state × urban status level. This regression can be estimated with data just during the pandemic or also with pre-pandemic data. Observations are weighted to make them representative at the national level after accounting for non-response (see Appendix C).

[3]When defining quantiles, we weight individuals by their average sampling weight over 2015–2019 using the sampling weights that make individuals nationally representative.

[4]A small fraction of employed individuals report 0 hours worked. We still count them as employed because some jobs entail sporadic work. CPHS does not report employment for minors aged <15. Individuals above 65 are excluded because their employment rates are very low.

We consider multiple baselines, i.e., predictions $\hat{y}_{i,t}$ about outcome in the absence of a pandemic, when estimating the regression above. Our default baseline is $\hat{y}_{i,t} = 100$, which implies that the normalized outcome would remain at the average 2018 level. This baseline ensures that the monthly fixed effects are the equivalent of the weighted-average of monthly outcomes.

Two other alternative baselines are estimated from following prediction model with pre-pandemic period data:

$$y_{it} = \sum_{q} \zeta^{q} \mathbb{1}(i \in q) + \sum_{q} \delta^{q} \cdot t \cdot \mathbb{1}(i \in q) + \sum_{m=1}^{12} \sum_{q} \phi_{m}^{q} \mathbb{1}(t \in m, i \in q) + u_{it} \tag{2}$$

where ζ^{q} is a quantile fixed effect; δ^{q} is a quantile-specific time trend; m indexes months; and ϕ_{m}^{q} is a month-of-year fixed effect intended to capture seasonality. Our second baseline is the predicted y_{it} when equation (2) includes only a time trend δ^{q} and no seasonal controls ϕ_{m}^{q} in either equation above. A third baseline accounts for both the trend and seasonality.

We visualize changes in economic outcomes during the pandemic by plotting month $\times$ quartile fixed-effect estimates. We add 100 to fixed effects when plotting them to ease visualization.

Rural-urban differences. We define rural areas as agglomerations categorized as villages in the 2011 Indian Census. All other areas are urban areas. We describe the regression model we use to draw statistical inferences about the differences in outcomes across urban and rural areas in the Appendix D1.

Gini coefficient and Shorrock's index. We measure inequality primarily by comparing mean outcomes in the bottom and top quartiles of the income distribution. However, we also calculate monthly income and consumption Gini coefficients to measure distribution and a Shorrock's index to measure social mobility. We explain these calculations in the Appendix D2 and D3.

2 Incidence of economic shocks

2.1 Poverty

Extreme income poverty spiked during the lockdown (Figure 1A). Prior to pandemic, extreme income poverty was higher in rural areas (12.5% v. 1.7% in urban areas). During

the lockdown, extreme poverty spiked more in urban areas (by 27.3 percentage points (pp) to 29% v. by 25.5 pp to 37.9% in rural areas).[5] Extreme poverty declined after the lockdown, but remained somewhat elevated above pre-pandemic levels, especially in rural areas (18.6% v. 3.2% in urban areas). The differences in extreme poverty across periods are statistically significant. The urban-rural differences were statistically significant in the pre-pandemic and in the post-lockdown periods.[6]

2.2 Income inequality

Despite the increase in poverty, our central and striking finding is that income inequality fell during the pandemic. Figure 1B plots mean income for 4 income quartiles from 2018 to 2021. To directly quantify inequality, Figure 1B also plots the difference between the relative incomes of the top- and bottom-quartile groups. Note that a lower level of difference implies lower inequality.

Trends in inequality vary over time and space. Prior to the pandemic, income inequality was declining. The decline starting in 2018 in urban areas. The decline started later and was more modest in rural areas. During the lockdown, income inequality spiked in urban areas ($p<0.001$ v. rural), reversing nearly all the inequality reductions since 2018. There was a more modest spike in rural areas, reversing only about half of the pre-pandemic gains. After the lockdown, inequality resumed its pre-lockdown progressive trend in both regions. As of the end of 2021, both regions had seen about 40 pp declines in inequality.

The estimated reduction in inequality after the lockdown is robust to different methods of measuring baselines, sample weights, and to concerns about mean reversion. If we use a baseline that accounts for the pre-pandemic trend in income (rather than 2018 levels), we estimate a larger decline in inequality (Figure 2A), though the difference between baselines is not significant. The greater decline is because the top quartiles have greater pre-pandemic income growth rates than the bottom quartile. Our findings are also robust to different weighting schemes and to defining quartiles using the national rather than state income distribution (Figure 2B). Finally, our primary measures of inequality are already somewhat robust to mean reversion since we define quartiles based on long-term (2015–2019) income. When we go further and use month fixed effects in our prediction

[5] Consistent with the pattern in extreme poverty, mean (Figure 1A) and median income (Figure 1C) was rising in rural areas before the pandemic and declined more in urban areas during the lockdown.

[6] In the Appendix E1, we measure poverty using two India-specific minimum wage-based measures of poverty and observe similar patterns over time and across urban and rural spaces.

model to allow within-year mean reversion, we estimate similar declines in inequality (see Appendix E2, Figure A3).

Our findings also hold at the extremes of the income distribution, i.e., quantiles above the 75th- and below the 25th-percentiles (Figure A4). There was, if anything, slightly more compression in the right tail than the left tail of the income distribution. Yet, relative income declines during the pandemic were nearly monotonic in pre-pandemic income percentiles (Figure 3A).

The gains at the lower percentiles are particularly remarkable because India—unlike the US—had little fiscal stimulus in the form of income transfers (Figure A5A). While government transfers did spike during the lockdown, they were small in absolute value (Figure A5A) Income net of government transfers shows the same pattern of declining inequality as income with transfers during the pandemic (Figure A5B).

2.3 Consumption inequality

Consumption inequality follows a pattern similar to income inequality, except that changes are more muted (Figure 1C).[7] One important difference is that the decline in relative consumption is progressive even during the trough of the lockdown in urban areas. The decline of relative consumption is also monotonic in percentile of 2015–2019 income (Figure A6).

Consumption inequality may have been more muted because households were able to smooth consumption—even during the pandemic. To estimate consumption smoothing across quartiles, we use regressions of individual consumption on individual income, à la Cochrane (1991) and Townsend (1994):

$$\log c_{ikt} = \mu_i + \alpha \Delta d_{kt} + \gamma \log M_{ikt} + \pi(\log M_{ikt} \times \mathbb{1}(2020)) + \varepsilon_{ikt} \tag{3}$$

where c_{ikt} is consumption by household i in location k, μ_i are household fixed effects, d_{kt} is a measure of the aggregate shock (proxied by location average consumption as in Townsend (1994)); M_{it} is idiosyncratic household income, and $\mathbb{1}(2020)$ is an indicator for 2020. We add aggregate consumption to capture aggregate shocks at the homogeneous-region $\times$ community-type level. Here $\gamma = 0$ implies full consumption smoothing and π measures whether COVID affected the ability to smooth consumption.

[7] Using a baseline that accounts for the pre-pandemic trend in consumption yields virtually the same result (Figure A7A).

Pre-pandemic, a 10% income fall in income was associated with a 0.980% decline in consumption (Table A2, specification (1)). This is a somewhat smaller estimate of the marginal propensity to consume (MPC) than in Townsend (1994), but our sample benefits from a quarter-century of credit-market development and includes cities. During COVID, a 10% fall in income was associated with a 0.869% decline in consumption, suggesting even more smoothing during the pandemic. While this difference between smoothing before and during COVID is statistically significant, it is very small.

Remarkably, rural areas have one-third MPC than urban areas, suggesting better formal or informal smoothing technology in rural areas (specifications (3) v. (4)). The changes observed during the pandemic do not change this basic finding. That said, rural areas are also more susceptible to aggregate shocks. All together these findings are consistent with rural areas having better local and informal credit or insurance markets, but urban areas having better financial linkages with other parts of India (Munshi and Rosenzweig, 2016).

The ability to smooth consumption does not vary meaningfully across income quartiles, as some prior literature (e.g., Morduch, 1999) has suggested. We estimate a version of equation (3) interacted with indicators for income quartiles (Table A3). We do not find material differences in pre-COVID marginal propensity to consume (MPC) across the income distribution. The lowest quartile had an MPC of 8.23%, while the highest had an MPC of 9.67%. During COVID these MPCs are 7.96% and 9.22%.

A particular concern is that the pandemic led to greater food insecurity amongst the poor (e.g., Egger et al., 2021; Bottan et al., 2020). However, substituting food expenditures for total expenditures in Figure 1C, we find that inequality in food expenditures did not increase during the pandemic (Figure A8). Households across quartiles were able similarly to smooth food expenditures both before and during the pandemic (Table A4).

2.4 Gini coefficient and social mobility

Income Gini coefficients (Figure 3B, left) paint a somewhat less rosy picture of income inequality during the pandemic, at least in rural India. Pre-pandemic, the Gini coefficient is high, but lower and declining in urban areas ($\sim$0.4 v. 0.45 in rural India). During the lockdown, the Gini spikes sharply in rural and urban areas above 0.65, but the urban spike is relatively larger given the lower base. After the lockdown, the Gini returns to pre-pandemic trends in both regions, flat in rural and declining in urban areas. There is an upward blip during India's second wave in April-May 2021, but it is short-lived.

In contrast to the income Gini coefficient, the consumption Gini (Figure 3B, right)

shows less sensitivity to the pandemic, likely because of consumption smoothing. The consumption Gini starts lower ($\sim$0.3 v. 0.4 for the income Gini) and urban rather than rural areas have greater inequality in consumption. Finally, both regions see declines through the pandemic.

Our primary measure of inequality and the income Gini coefficients show somewhat different patterns, at least after the lockdown in rural areas, probably because the Gini fails to capture social mobility and social mobility increased after the lockdown. Income change can be thought of as the sum of two processes: (1) change in income holding relative position or percentile constant and (2) change in relative position. The Gini coefficient loosely tracks the first element.[8] The second element is often called social mobility. Our main measures of inequality track both measures combined.

Comparing the pre-pandemic to the pandemic period (aside from the lockdown), Lorenz curves for income show that there were no large changes in the (share of) income of each percentile (Figure A9).[9] However, there was a large spike in progressive social mobility across all quartiles, though the mobility was higher in rural areas. The Shorrock's Index, which reports transitions across quartiles, shows the bottom (top) 2 quartiles moved up (down) more from 2019–2020 than from 2018–2019, in rural and urban India (Figure 3C, left).[10] If we sum the bottom-left and top-right 2x2 sub-matrices of the Shorrock's Index as a rough measure of progressive mobility, we see that rural areas show greater mobility (23.48 v. 19.51 pp in urban). In other words, total inequality fell, it can be explained by social mobility rather than changes in the Gini coefficient, and social mobility was somewhat higher in rural India.

3 Mechanisms

3.1 Components of income

We examine the 5 components of income – business profits, labor, government transfers, pension, and other – to determine the mechanisms behind changes in inequality during

[8] We say loosely because Gini is normalized by the population's overall income. A simple change in income does not have to be normalized in that manner.

[9] Snapshot Lorenz curves show that inequality spiked during the lockdown (April 2020) and then returned to pre-lockdown levels by July 2020, even before the peak in cases.

[10] Comparing Shorrock's Index values before the pandemic and after the lockdown shows progressive social mobility increased by more than 10% in 2019–2020 relative to 2018–2019 (Table A6).

the pandemic.[11] Figure 4A shows trends in the amount of income attributable to these components for 1st (bottom) and 4th (top) quartile persons. Labor income, business profits and transfers show the largest changes during the pandemic.

To quantify the relative role each component plays in trends in inequality, we decompose the change in each category in income into two parts: changes in that category's share of income and changes in the amount of income from that category, holding the shares constant. Details of the decomposition are presented in Appendix F1.

Of the 39.7 pp reduction in inequality across India from the 2019 average to July 2021, we find that labor income and business profits play much bigger roles than government-transfer income and other components (Table A7). While labor income is a 5.4 pp larger (trend-weighted) share of bottom quartile income, (share-weighted) labor income falls 24.9 pp more in the top quartile. The net effect is a nearly 20 pp contraction of inequality due to that component. Business income plays the next most important role: both the weighted share effect (6.7 pp) the weighted change in income-component effect (9.4 pp) favor the bottom quartile. It contributes roughly 16 pp to the reduction in inequality. By contrast, transfer income is only a slightly larger share of bottom-quartile income (Figure A10), suggesting a problem with targeting of that support. And it changes very little over the course of the entire pandemic – though it does spike during the lockdown. Other components similarly have negligible roles in changes in inequality.

The relative role of business profits and wages in the decline of inequality differ across urban and rural India (Table A7). Changes in labor income play a larger role in urban India (total change 22.5 pp v. 17 pp in rural areas) while change in business profits play a larger role in rural India (18 pp v. 8.1 pp in urban India).

In the remainder of this section, we examine the two largest components of income—business profits and labor earnings—that explain the decline in income inequality during the pandemic in India.

3.2 Business profits

Aggregate income declined during the pandemic, especially during the lockdown. Payments to residual claimants may be the first on the cutting block in a downturn. If the rich

[11] We count "wage income" as labor income. "Business income" is non-passive capital income. We count non-pension income from dividends, interest, rent and sale of assets as passive capital income distinct from "pension income" and business income and include that in the other category as it is a small amount. There are sources like gambling that are also classified as other. Total income—and aggregate income—includes all components (labor income, business profits, pension income, government transfers and other.

are more likely to be residual claimants, their income may suffer more during a downturn. We find there is merit to this explanation in rural areas.

Following Guvenen et al. (2017), we regress changes in each component of an individual's income against changes in aggregate income (all components) interacted with quartile indicators:

$$\log y^{c,q}_{i,t+1} - \log y^{c}_{i,t} = \sum_{q} (\alpha^{c,q} + \beta^{c,q}(\log Y_{t+1} - \log Y_t)\mathbb{I}(i \in q) + e^{c,q}_{i,t+1} \tag{4}$$

where $y^{c}_{i,t}$ is person i's income from component c, $y^{c}_{i,t}$ narrows that to a person in quartile q, and Y_{t+1} is aggregate income from all components. Because income from a component may be 0, we add one to each component before taking logs. We estimate this equation separately for urban and rural areas. The coefficient $\beta^{c,q}$ gives the sensitivity ("beta") of income component c to aggregate income for individuals from quartile q.

The high beta of business income may explain the reduction in income inequality in rural areas during the pandemic. Figure 4B shows that the business income of the top quartile is more sensitive than the bottom quartile to aggregate income in rural areas, though not in urban areas. The difference is not significant, but it is nominally large.

In contrast, the beta of labor income is unlikely to explain the reduction in inequality. The labor income of lower quartiles are not nominally more sensitive to aggregate income in either rural or urban areas (Figure 4C).

3.3 Labor supply

Turning to labor income, we first examine labor supply here, then demand in the next section. Ultimately, we reject the hypothesis that the rich saw income declines because they disproportionately reduced their labor supply during the pandemic.

To test the labor supply explanation, we first measure labor supply by reservation wage (following a method from French and Taber (2011), explained in Appendix F2). We then see how that covaries with income quartile by estimating the following individual-level regression:

$$log\hat{V}_{i,j,j',k,t} = \alpha + \beta \mathbb{1}_k(> \text{median}) + \theta \mathbb{1}_k(t \geq p) + \pi[\mathbb{1}(> \text{median}) \times \mathbb{1}(t \geq p)] + u_{i,t} \tag{5}$$

where $\hat{V}_{i,j,j',k,t}$ is log predicted reservation wage, $\mathbb{1}_k(> \text{median})$ is an indicator for above median income in location k, $\mathbb{1}(t \geq p)$ is an indicator for the pandemic period, and $u_{i,t}$ is

a regression error. We report results using a sample from CPHS round 2 (May–August) in 2019 and 2020.[12]

We find that, while the well-off have higher reservations wages, the average person's reservation wage falls during the pandemic and, importantly, the reservation wages of the well-off fall more than average (Table A8, Column 5). Everyone appears to be willing to work for less during the pandemic, i.e., supply increases, but especially the rich. Thus, it is unlikely that labor supply explains the rise in inequality.

One might wonder whether factors that affected the behavior of the rich in high-income countries, such as COVID risk and access to remote work, also affected those in a lower-income country like India. We explore and rule out these factors in Appendix F3.

3.4 Labor demand

If not lower supply, then perhaps lower demand for top-quartiles' labor can explain the decline in inequality during the pandemic. We review indirect and then direct evidence for this mechanism. We find that falling demand for labor in the services sector combined with the greater reliance of top quartiles to labor income in the services sector can explain the decline in inequality in urban areas.

Indirect evidence. For indirect evidence, we examine the impact of the epidemic on the employment rate, hours worked when employed, and hourly wages. We focus first on the post-lockdown period. There are few urban-rural differences here (Figure A13). The main driver of reduced labor-income inequality is wages: wages of the poor recover more than wages of the rich. Employment fell somewhat more in urban areas, but the gaps across quartiles was similar. Finally, hours per day employed returned to nearly 8 hours for all quartiles.

These changes concern equilibrium outcomes, but they are consistent with a labor demand rather than supply shift. A supply contraction would be associated with higher wages but lower employment rate. By contrast, a reduction of demand for labor would be associated with both lower wages and lower employment. We observe the latter for top quartiles after the lockdown.

Lockdown was a different story. Employment rates fell in both regions, but much more so for the poor in urban areas, perhaps explaining the larger spike in urban inequality during the lockdown. Hour per day employed fell more in urban areas, but there do

[12]Our results are similar if we use quartiles of income and/or a sample that includes all of 2019 and 2020.

not appear to gaps across quartiles. Finally, while wages fell during lockdown in urban areas, they rose in rural areas. The poor experienced larger wage declines in both regions, but more so in urban areas. The different price responses in urban and rural areas suggest that falling demand (i.e., lower quantity and wage) explains the urban experience, while falling supply (i.e., lower quantity but higher price) explains the rural experience during lockdown.

Direct evidence. To obtain more direct evidence on the role of labor demand, we check if top-quartile incomes are more sensitive to expenditures in consumption categories that fell more during the pandemic. This is difficult to do precisely, as it is difficult to map how much labor from each income quartile is required to produce different consumer products. Our approach is therefore approximate. First, we map each of the 55 categories in which we have disaggregated household expenditure to 3 broad sectors: agriculture, manufacturing and services (Table A9). Second, we map 37 occupational categories to those 3 sectors (Table A10).

At the sectoral level, labor demand likely plays a role in the decline in inequality in urban areas more than rural ones. A larger fraction of top-quartile income is from the service sector (Figure 5B) and that sector experienced the largest drop in consumer expenditure during the pandemic (Figure 5C). This is true across regions, but the share of income from service is much larger in urban areas, so that labor demand effect on inequality is much larger in urban areas.

It is true that top-quartile income is less sensitive to manufacturing, and that manufacturing consumption fell more than agricultural consumption. However, manufacturing is a much smaller share than services of top-quartile income, and manufacturing consumption fell half as much as service consumption during the pandemic.

4 Discussion

We find that income inequality fell during the pandemic. The results are robust to alternative baselines, definitions of quartiles, and methods of weighting. Consumption inequality also fell, but to a lesser extent due to consumption smoothing. The fact that the Gini coefficient was the same post-lockdown suggests that most of the decline in inequality was due to progressive social mobility. In rural areas, some of the decline in income among the rich may be explained by higher sensitivity of business income to aggregate fluctuations. In urban areas, the decline in inequality is mainly explained a reduction in

labor demand in the occupations the rich occupy.

To benchmark these changes in inequality, we compare them to what happened in India during demonetization. Demonetization began in November 2016, when the Prime Minister declared that ₹500 and 1000 notes, which accounted for 86% of cash in circulation, would be banned to combat black market income. They would ultimately be replaced by new ₹500 and 2000 notes. In the short run, demonetization had the potential to generate a massive transaction cost or liquidity crisis because 78% of transaction in India were in cash (Chodorow-Reich et al., 2020). Plotting income and consumption by quartiles defined in 2015, we see that income inequality did not decline as much during demonetization as during the pandemic (Figures A14 and A15). Remarkably, income grew during demonetization. However, income grew *relatively* slower for top-quartiles.

The experience in India during the pandemic also contrasts with the experience of the United States. Higher income persons saw smaller reductions in income excluding government transfer than did lower income persons (Chetty et al., 2020). This was partly the result of better access to remote work (Adams-Prassl et al., 2020). Relative consumption by lower-income persons was protected by transfers from the government rather than earned or capital income (Han et al., 2020). Our results therefore highlight the importance of technological shifts and public insurance in estimating the incidence of economic shocks across countries.

Our work has important limitations. First, we do not show that the pandemic *caused* a decline in inequality. We only show that inequality declined during the pandemic. Indeed, Figure 1B suggests that the decline in income inequality started in late 2018. India's lockdown interrupted that decline, but it resumed afterwards during the bulk of the pandemic.

Second, while the CPHS data set has valuable features, it also has important flaws. Dreze and Somanchi (2021) criticized it for undersampling the poor in villages. We partly address this by showing declining inequality across percentiles, even in the bottom 10% or even 5% of the sample.

Third, much of the literature on inequality focuses on top earners, i.e., the top 5 or 1%, because they appear to be pulling ahead of even the remainder of the top quartile. We show inequality for the top 5% and 10% of the CPHS sample, but CPHS does not capture the top 0.1% of the actual income distribution.

Fifth, we provide evidence on inequality through July 2021, but the pandemic is not yet finished.

References

Abraham, R. and A. Shrivastava (2019). How comparable are india's labour market surveys? Report, Centre for Sustainable Employment, Azim Premji University.

Adams-Prassl, A., T. Boneva, M. Golin, and C. Rauh (2020). Inequality in the impact of the coronavirus shock: Evidence from real time surveys. Journal of Public Economics 189, 104245.

Alstadsæter, A., B. Bratsberg, G. Eielsen, W. Kopczuk, S. Markussen, O. Raaum, and K. Røed (2020). The first weeks of the coronavirus crisis: Who got hit, when and why? evidence from norway. Report 0898-2937, National Bureau of Economic Research.

Anand, I. and A. Thampi (2016). Recent trends in wealth inequality in india. Economic and Political Weekly 51(50), 59–67.

Andrews, M. A., B. Areekal, K. R. Rajesh, J. Krishnan, R. Suryakala, B. Krishnan, C. P. Muraly, and P. V. Santhosh (2020). First confirmed case of covid-19 infection in india: A case report. The Indian journal of medical research 151(5), 490–492.

Azim Premji University, C. f. S. E. (2021). State of working india 2021: One year of covid-19. Report.

Bairoliya, N., A. Chanda, and J. Fang (2021). Consumption smoothing and household savings in india: Role of demographics and durables. Technical report.

Banerjee, A. and T. Piketty (2005). Top indian incomes, 1922–2000. The World Bank Economic Review 19(1), 1–20.

Bertrand, M., K. Krishnan, and H. Schofield (2020). How are indian households coping under the covid-19 lockdown? 8 key findings. Rustandy Center for Social Sector Innovation Blog 11.

Bhoi, B., H. Shekhar, and I. Padhi (2020). Rural-urban inflation dynamics in india. RBI Bulletin December.

Bottan, N., B. Hoffmann, and D. Vera-Cossio (2020). The unequal impact of the coronavirus pandemic: Evidence from seventeen developing countries. PLOS ONE 15(10), e0239797.

Chancel, L. and T. Piketty (2019). Indian income inequality, 1922-2015: From british raj to billionaire raj? Review of Income and Wealth 65, S33–S62.

Chen, H., W. Qian, and Q. Wen (2020). The impact of the covid-19 pandemic on consumption: Learning from high frequency transaction data. Available at SSRN 3568574.

Chetty, R., J. N. Friedman, N. Hendren, M. Stepner, and T. O. I. Team (2020). The economic impacts of covid-19: Evidence from a new public database built using private sector data. National Bureau of Economic Research Working Paper Series No. 27431.

Chodorow-Reich, G., G. Gopinath, P. Mishra, and A. Narayanan (2020). Cash and the economy: Evidence from india's demonetization*. The Quarterly Journal of Economics 135(1), 57–103.

Choudhary, M. A., I. A. Haqqani, F. Lenzi, and N. Limodio (2021). Subnational income, convergence, and the covid-19 pandemic. Technical report.

Cochrane, J. H. (1991). A simple test of consumption insurance. Journal of political economy 99(5), 957–976.

Deaton, A. (2021). Covid-19 and global income inequality. National Bureau of Economic Research Working Paper Series No. 28392.

Deshpande, A. (2020). The covid-19 pandemic and lockdown: First effects on gender gaps in employment and domestic work in india? Report, Working Paper 30, Ashoka University, Department of Economics, revised June.

Dhingra, S. and S. J. Machin (2020). The crisis and job guarantees in urban india. Report, IZA DP No. 13760.

Dingel, J. I. and B. Neiman (2020). How many jobs can be done at home? Journal of Public Economics 189, 104235.

Dreze, J. and A. Somanchi (2021). View: New barometer of india's economy fails to reflect deprivations of poor households. The Economic Times June 21(June 21).

Egger, D., E. Miguel, S. S. Warren, A. Shenoy, E. Collins, D. Karlan, D. Parkerson, A. M. Mobarak, G. Fink, C. Udry, M. Walker, J. Haushofer, M. Larreboure, S. Athey, P. Lopez-Pena, S. Benhachmi, M. Humphreys, L. Lowe, N. F. Meriggi, A. Wabwire, C. A. Davis,

U. J. Pape, T. Graff, M. Voors, C. Nekesa, and C. Vernot (2021). Falling living standards during the covid-19 crisis: Quantitative evidence from nine developing countries. Science Advances 7(6), eabe0997.

French, E. and C. Taber (2011). Identification of models of the labor market, Volume 4, pp. 537–617. Elsevier.

Google LLC (2021). Google covid-19 community mobility reports.

Gupta, A., A. Malani, and B. Woda (2021, June). Explaining the income and consumption effects of covid in india. Working Paper 28935, National Bureau of Economic Research.

Guvenen, F., F. Karahan, S. Ozkan, and J. Song (2015). What do data on millions of us workers reveal about life-cycle earnings dynamics? FRB of New York Staff Report (710).

Guvenen, F., S. Schulhofer-Wohl, J. Song, and M. Yogo (2017). Worker betas: Five facts about systematic earnings risk. American Economic Review 107(5), 398–403.

Hale, T., S. Webster, A. Petherick, T. Phillips, and B. Kira (2020). Variation in government responses to covid-19. Report, Blavatnik school of government working paper.

Han, J., B. D. Meyer, and J. X. Sullivan (2020). Income and poverty in the covid-19 pandemic. National Bureau of Economic Research Working Paper Series No. 27729.

Lee, K., H. Sahai, P. Baylis, and M. Greenstone (2020). Job loss and behavioral change: The unprecedented effects of the india lockdown in delhi. University of Chicago, Becker Friedman Institute for Economics Working Paper (2020-65).

Malani, A. and S. Ramachandran (2021). Using household rosters from survey data to estimate all-cause mortality during covid in india. National Bureau of Economic Research Working Paper Series No. 29192.

Malani, A., D. Shah, G. Kang, G. N. Lobo, J. Shastri, M. Mohanan, R. Jain, S. Agrawal, S. Juneja, S. Imad, and U. Kolthur-Seetharam (2020). Seroprevalence of sars-cov-2 in slums versus non-slums in mumbai, india. The Lancet Global Health.

Miguel, E. and A. M. Mobarak (2021, October). The economics of the covid-19 pandemic in poor countries. Working Paper 29339, National Bureau of Economic Research.

Morduch, J. (1999). Between the state and the market: Can informal insurance patch the safety net? The World Bank Research Observer 14(2), 187–207.

Munshi, K. and M. Rosenzweig (2016). Networks and misallocation: Insurance, migration, and the rural-urban wage gap. The American Economic Review 106(1), 46–98.

Organization, W. H. (2021). Covid-19 weekly epidemiological update. 49(20 July 2021).

Oster, E. (2012). Hiv and sexual behavior change: Why not africa? Journal of Health Economics 31(1), 35–49.

Oxfam (2021). The inequality virus. Report.

Parker, J. A. and A. Vissing-Jorgensen (2010). The increase in income cyclicality of high-income households and its relation to the rise in top income shares. Technical report, National Bureau of Economic Research.

Pinto, A., A. Bhowmick, and R. Adlakha (2020). How did india's rural economy fare through the covid-19 lockdown and the re-opening? Report, World Bank.

Reserve Bank of India (2020). Annual report 2019–2020. Report, Reserve Bank of India.

Sarkar, S. and B. S. Mehta (2010). Income inequality in india: pre-and post-reform periods. Economic and Political Weekly, 45–55.

Scheidel, W. (2018). The Great Leveler. Princeton University Press.

Sheng, J., A. Malani, A. Goel, and P. Botla (2021). Does mobility explain why slums were hit harder by covid-19 in mumbai, india? Journal of Urban Economics.

Shorrocks, A. F. (1995). Revisiting the sen poverty index. Econometrica 63(5), 1225–1230.

Townsend, R. M. (1994). Risk and insurance in village india. Econometrica 62(3), 539–591.

Vyas, M. (2020a). Cphs execution during the lockdown of 2020. Report, Centre for Monitoring Indian Economy.

Vyas, M. (2020b). The questionnaire. Report, Centre for Monitoring Indian Economy.

Vyas, M. (2020c). Sample survival response rate. Report, Centre for Monitoring Indian Economy.

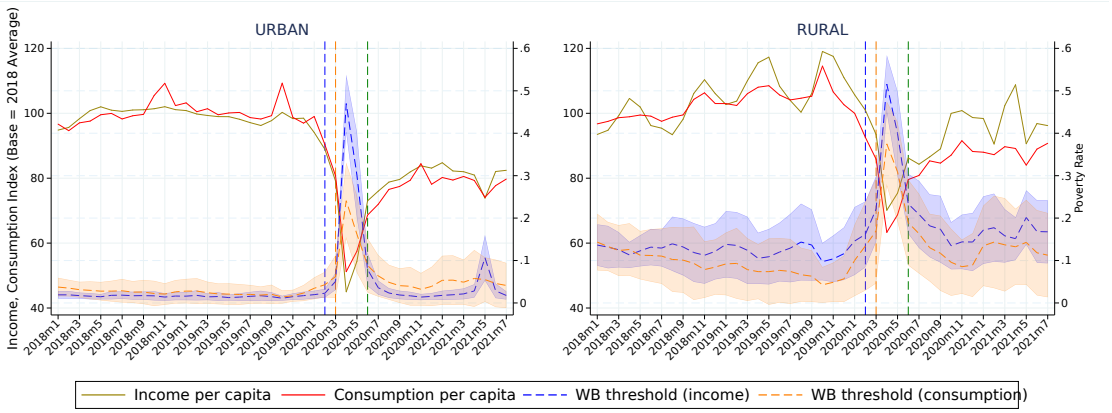
Vyas, M. (2020d). Survey design and sample. Report, Centre for Monitoring Indian Economy.

Vyas, M. (2021). View: There are practical limitations in cmie's cphs sampling, but no bias. The Economic Times June 23(June 23).

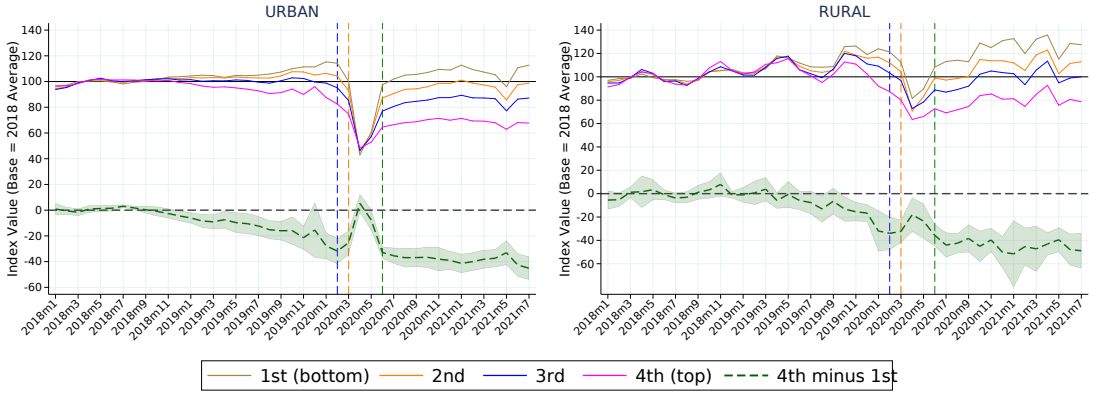
Exhibits

Figure 1: Time-series of poverty and income and consumption inequality

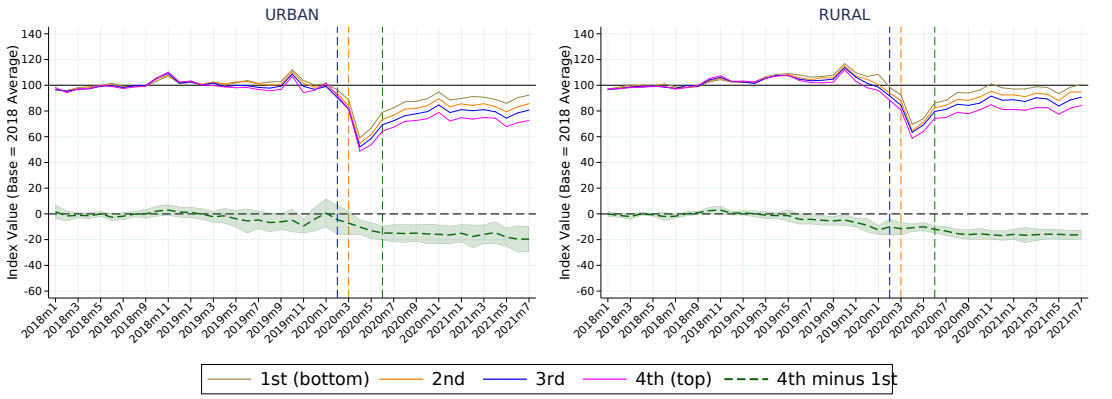
Panel A: Poverty measures



Panel B: Income inequality



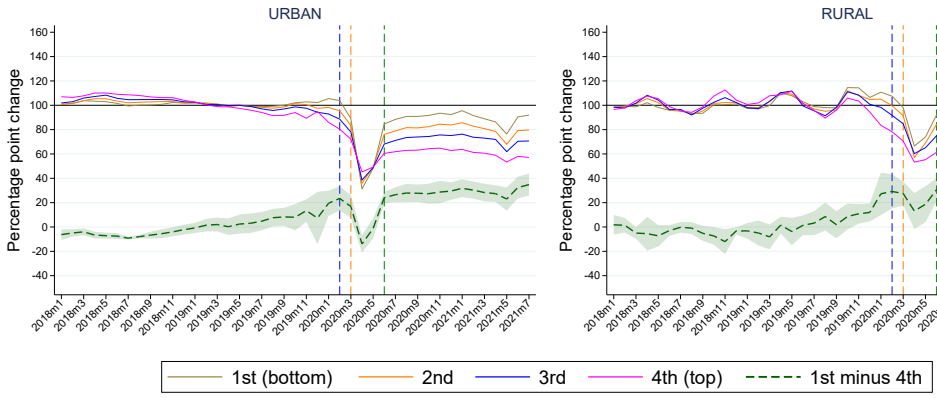
Panel C: Consumption inequality



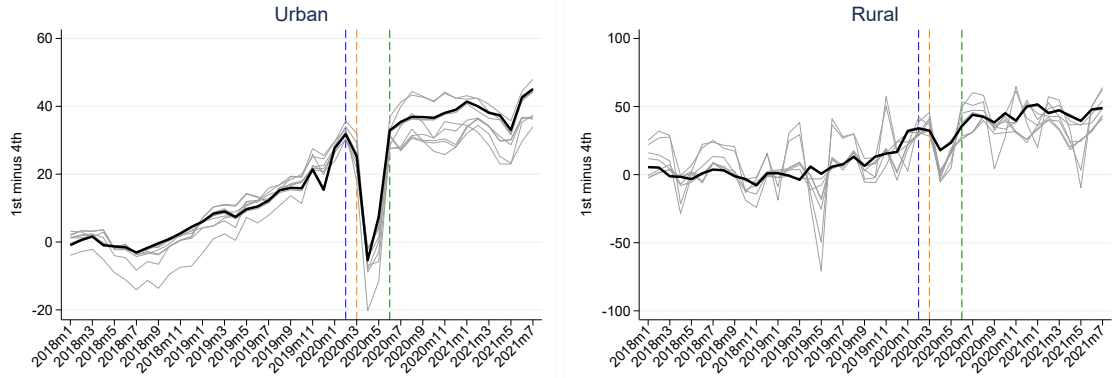
Note. Panel A: Mean per-capita income and consumption are each reported an index, relative to average 2018 values. We report shares of the population below the World Bank's $1.90 extreme poverty thresholds applied to income and consumption. Rupee conversions use the average exchange rate for each month from the IMF's International Financial Statistics. Panel B (C): The figure plots the fixed effects (lines) and 95% confidence intervals (shaded areas) for quartile x month fixed effects estimated using equation (1) with $\hat{y}_{i,t} = 100$. The lines are the equivalent of the weighted average of per-capita income (consumption) within income quartiles in each state $\times$ urban status location, using individual member weights. The units are an index where 100 is average 2018 income (consumption) of a person. The dashed line at the bottom indicates the difference between the first- and fourth-quartile index for income (consumption), measuring the decline in inequality in percentage points. All: Shaded area is the 95% confidence interval around a statistic. Dashed vertical lines in February 2020, March 2020 and June 2020 indicate the first month of the pandemic (blue), the month the national lockdown started (orange) and the month the national lockdown ended (green).

Figure 2: Robustness of trends in income inequality

Panel A: Linear trend baseline



Panel B: Different weighting schemes and national quartiles



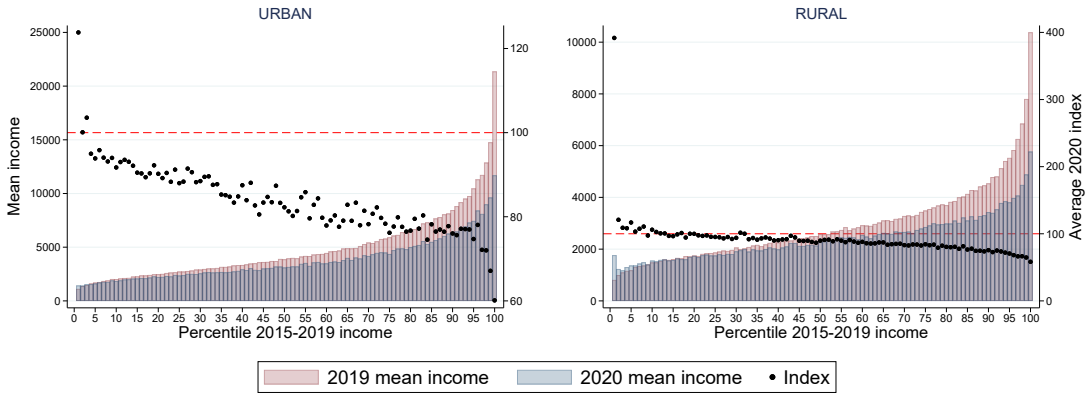
Note. Panel A: The units are an index where 100 is average 2018 income of that person. A trend was obtained from a regression of pre-pandemic income on time, separately for each income quartile, per equation (2). Deviations from the trend (plus 100) were then regressed on month × quartile dummies, per equation (1). Panel B: The figure was constructed by calculating the (weighted) average of per capita income within indicated income quantiles in each jurisdiction × urban status location and taking the difference between the 1st quartile values and the 4th quartile values. This difference measures the decline in inequality, i.e., positive values indicate a decline in inequality. The (weighted) means are obtained from the time fixed effects estimates in (1) using the default baseline (100, corresponding to a person's average 2018 income). The units are an index where 100 is average 2018 income of a person. The black line using individual weights that ensure the sample is nationally representative even after accounting for non-response, assuming that response is random. The grey lines present measures of inequality using the following, alternative weighting and quartile calculations:

> [noitemsep]using the last weight observed, instead of mean weight, using no weights at all, using a national base value instead of state x community specific with mean weights, and using a national base value instead of state x community specific with no weights.

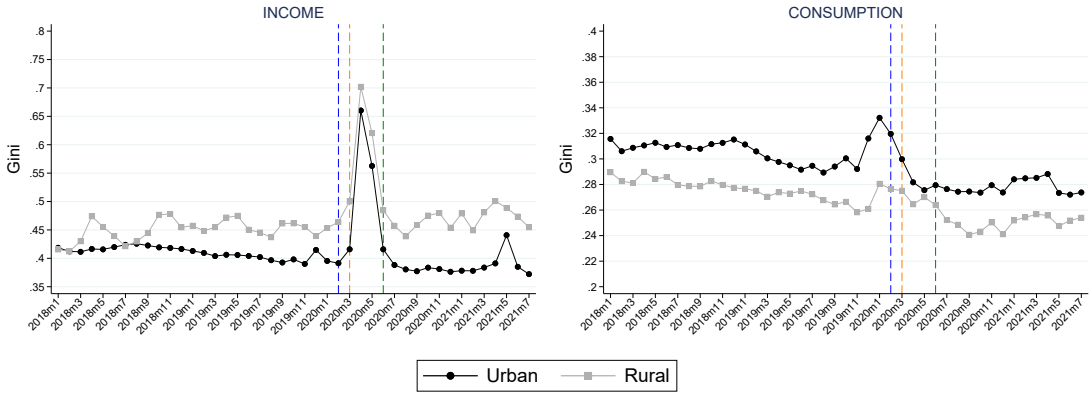
As an additional robustness check, all of the above are calculated using a constant sample of individuals who are observed in each month from January 2015 to July 2021. All: Dashed vertical lines in February 2020, March 2020 and June 2020 indicate the first month of the pandemic (blue), the month the national lockdown started (orange) and the month the national lockdown ended (green).

Figure 3: Aspects of income inequality

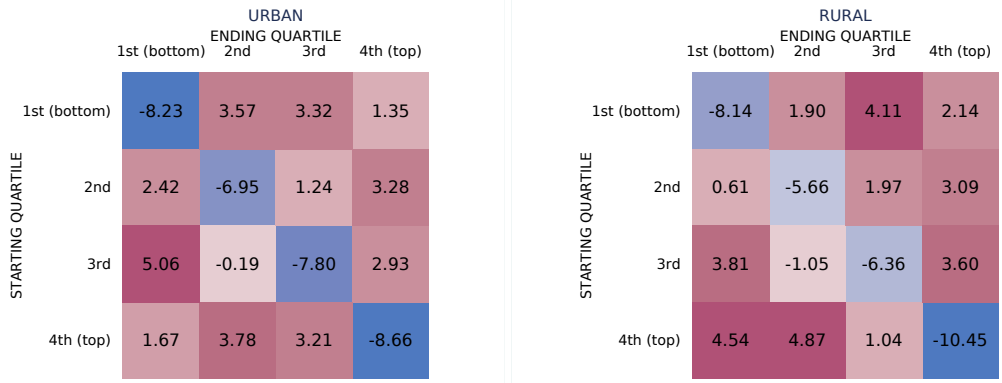
Panel A: Income distribution



Panel B: Gini coefficients



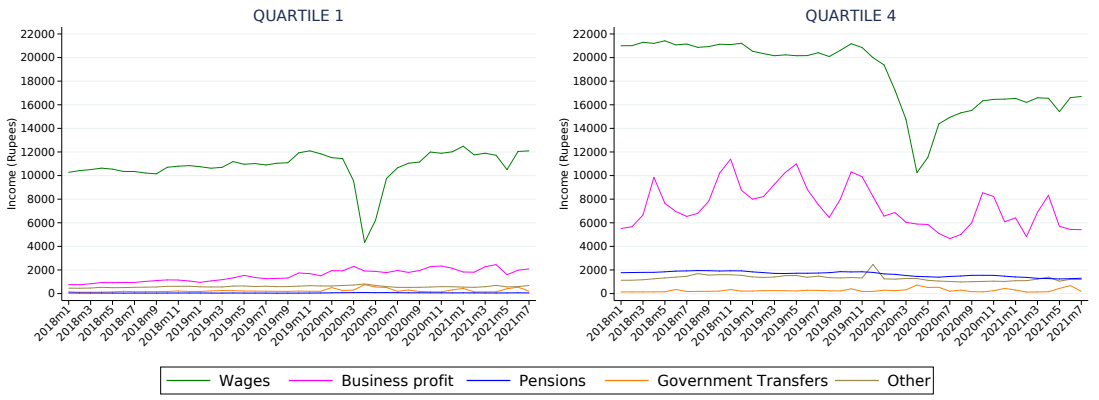
Panel C: Shorrock's Index and social mobility



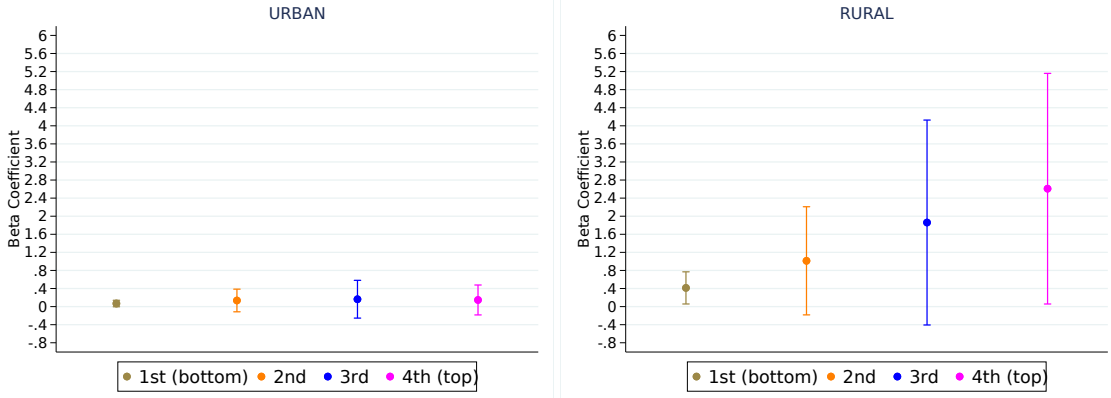
Note. Panel A: The red (blue) bars show the mean 2019 (2020) income in each percentile. Note that mean 2020 income divided by 2019 mean income is not equal to the exact value of the index because the base value is specific to each person's state × urban status. The dashed red line demarcates 100%. Panel B: Gini coefficients were obtained by sorting individuals on income, separately for urban and rural areas, and assigning each individual a place in the income distribution, based on their survey weight. We then calculated the individual's share of the population based on their survey weight, as well as their share of the total national income. We multiplied the population share by the cumulative share of income and summed across individuals in a given month to obtain the Gini coefficient for that month. Panel C: Each matrix is equal to the 2019-2020 income quartile transition matrix minus the 2018-2019 income quartile transition matrix. The 2018-2019 income transition matrix is the probability (× 100) that a person in the (row) income quartile in 2018 was in the (column) quartile in 2019. The 2019-2020 income transition matrix is calculated analogously. 2018, 2019 and 2020 income quartiles are defined separately for each year based on the distribution of average per-capita monthly income across all months.

Figure 4: Role of business income

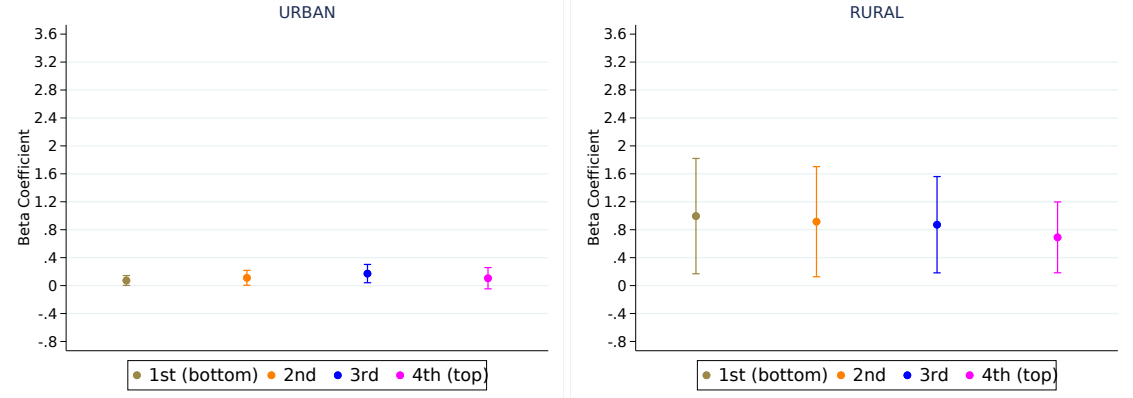
Panel A: Sources of income by quartile



Panel B: Sensitivity of business profit to aggregate income



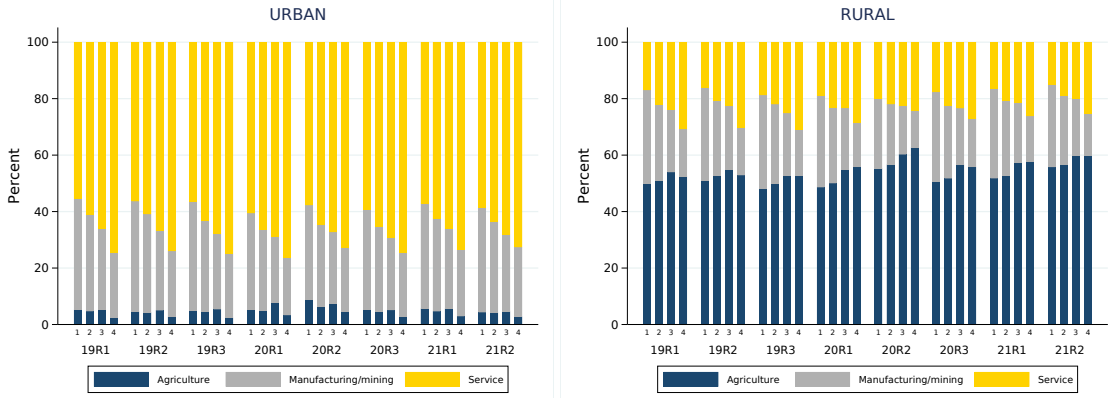
Panel C: Sensitivity of labor income to aggregate income



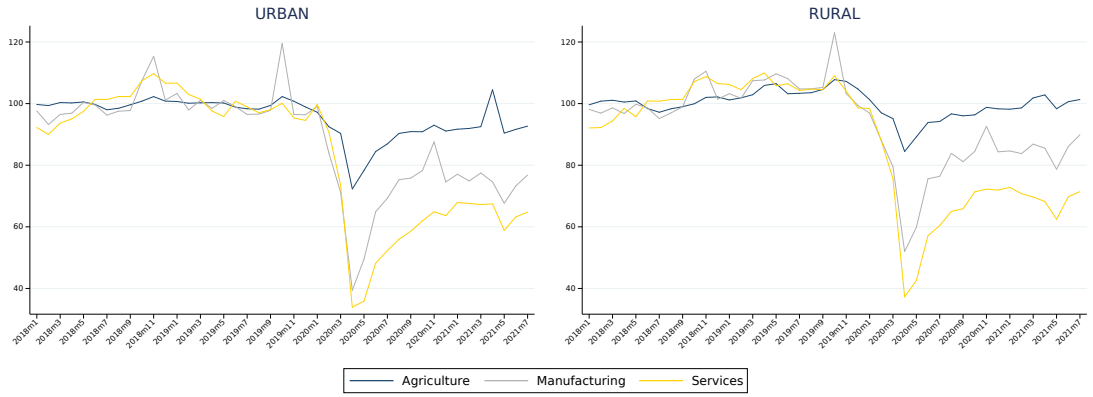
Note. Panel A: Amount of income from different income sources. Capital income includes dividends, interest, rent, sale of assets outside of pension accounts. Other income includes government transfers, private transfers, value of agricultural goods produced for self-consumption, lottery winnings, insurance payouts. Panel B: The subplots report coefficients from regressions of monthly differences in the log of (one plus) an individual's business (left plot) or wage (right plot) income against changes in aggregate income (all components) interacted with 2015-2019 quartile indicators (see (4)). The specific coefficient reported is the coefficient on aggregate income interacted with each quartile.

Figure 5: Role of labor demand

Panel A: Employment by sector



Panel B: Consumption by sector



Note. Panel A: Each bar reports the share of population in each quartile with occupations in each of 3 sector (agriculture, manufacturing and services) in each month. Table A10 reports the mapping of occupations to sectors. Panel C: This plot shows aggregate consumption of goods in 3 sectors by month relative to aggregate consumption in that sector in 2018. Table A9 reports the mapping of CPHS consumption categories to sectors.

Appendix

A Figures **3**

A1 Background . . . 3
A2 Poverty . . . 4
A3 Income inequality . . . 4
A4 Consumption inequality . . . 7
A5 Gini coefficient . . . 10
A6 Mechanisms . . . 11
A7 Labor supply . . . 12
A8 Labor demand . . . 13
A9 Discussion . . . 14
A10 Appendix: Data . . . 15
A11 Appendix: Mean reversion . . . 16

B Tables **17**

B1 Extreme poverty . . . 17
B2 Consumption inequality . . . 18
B3 Gini coefficient . . . 22
B4 Mechanisms . . . 23
B5 Labor supply . . . 24
B6 Labor demand . . . 25

C Data **27**

C1 Detailed description of CPHS . . . 27
C2 Issues with CPHS data . . . 27
C3 RBI inflation data . . . 30
C4 COVID cases, lockdown rules and mobility . . . 30

D Methods **30**

D1 Inferences about rural-urban differences . . . 30
D2 Gini coefficient . . . 31
D3 Social mobility . . . 31

E Incidence **32**
E1 Poverty 32
E2 Mean reversion 32

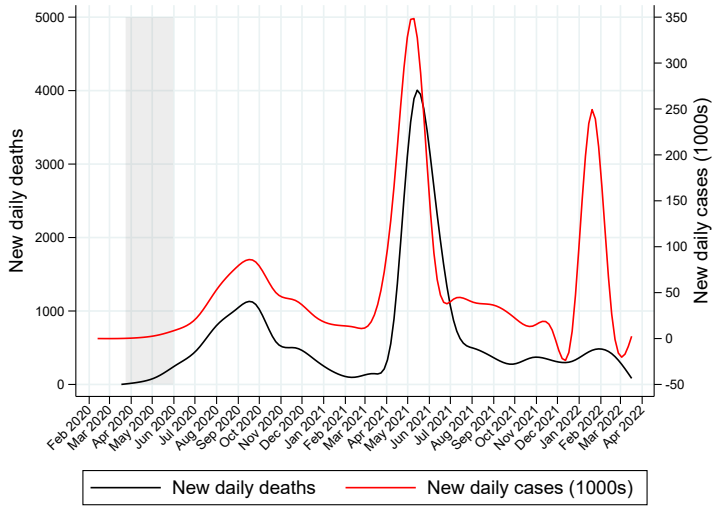
F Mechanisms **33**
F1 Decomposition of income inequality by source of income 33
F2 Labor supply: COVID and remote work 34
F3 Labor supply: COVID and remote work 34

A Figures

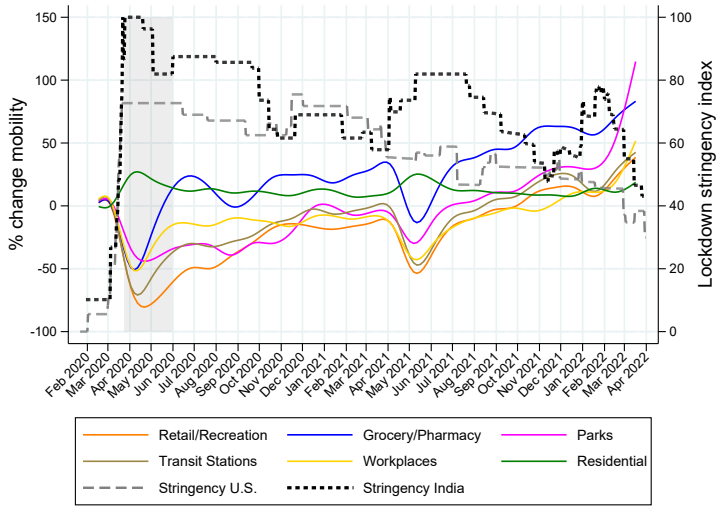
A1 Background

Figure A1: COVID Trajectory, Severity of Lockdown, and Mobility Changes

Panel A: Cases and deaths



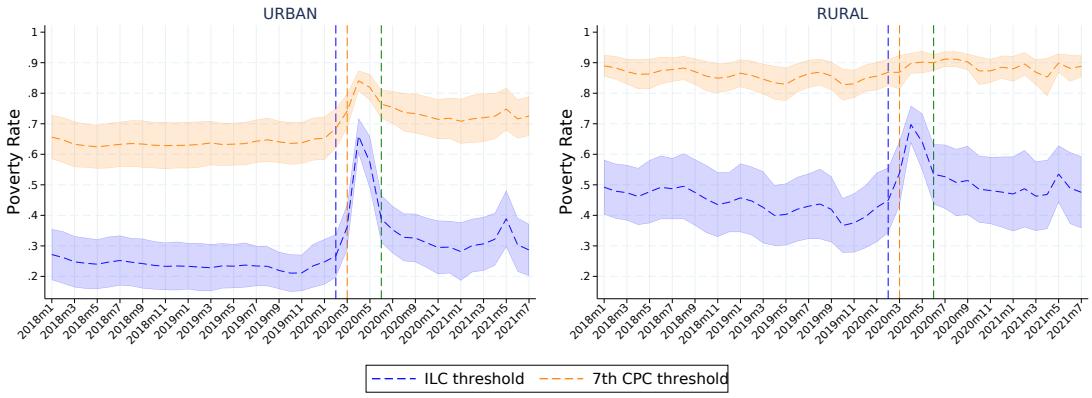
Panel B: Lockdown severity and mobility



Note. Case and death data are from `www.covid19India.org`. We show aggregated daily reported cases and deaths from the government. Shaded period marks the national lockdown. Lockdown severity data are from Oxford Hale et al. (2020). Mobility data are from Google mobility reports Google LLC (2021). Shaded period marks the national lockdown. Time periods cover February 2020–January 2021.

A2 Poverty

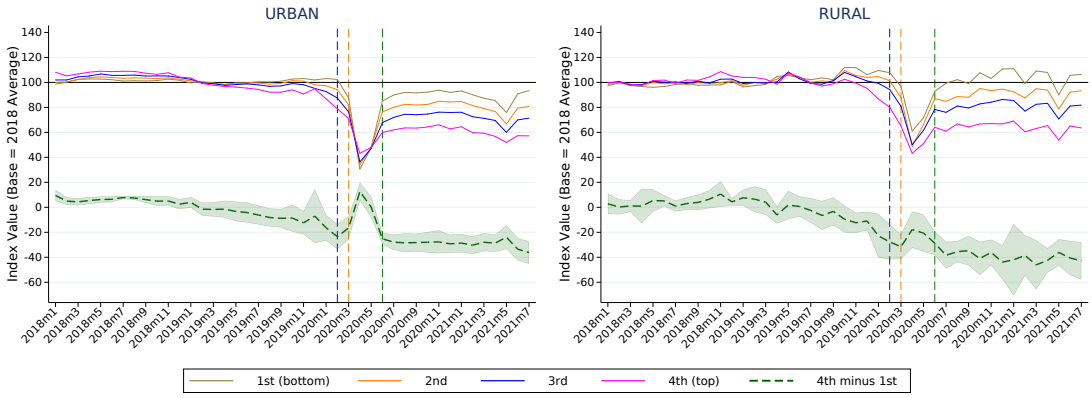
Figure A2: Time-series of India-specific poverty measures



Note. The figure plots shares of the population below 1) the National Minimum Wage, equal to ₹1909 (in 2012 terms) for rural areas and ₹2256 for urban areas per month and 2) the minimum wage recommended by the 7th Central Pay Commission, equal to ₹4660 per month. Shaded area is the 95% confidence interval around a statistic. Dashed vertical lines in February 2020, March 2020 and June 2020 indicate the first month of the pandemic (blue), the month the national lockdown started (orange) and the month the national lockdown ended (green).

A3 Income inequality

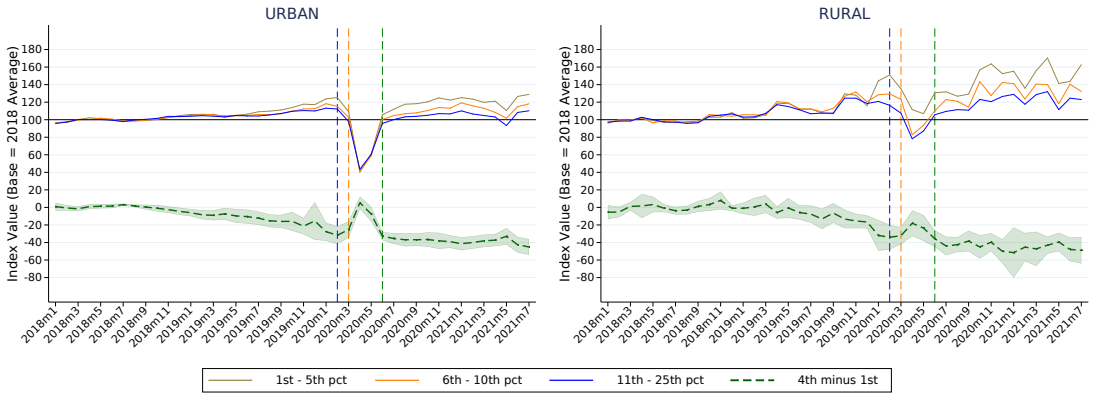
Figure A3: Income inequality with linear trends and seasonal adjustment.



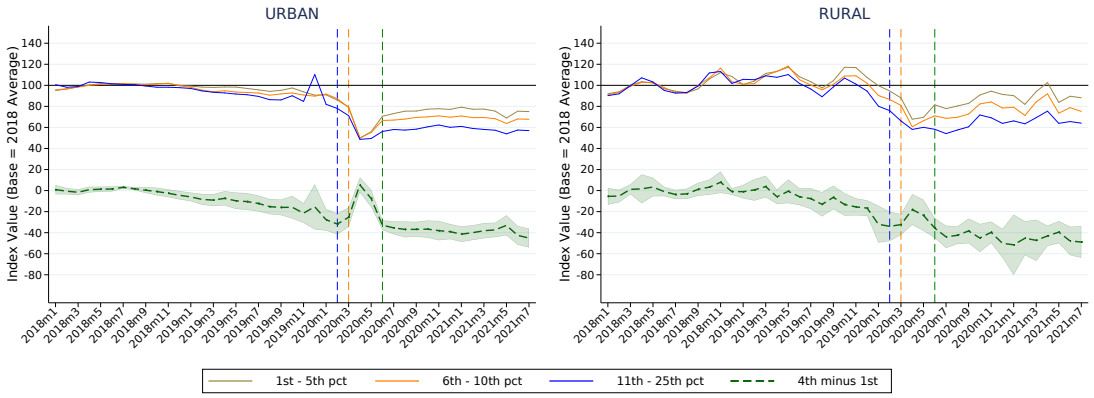
Note. The units are an index where 100 is average 2018 income of that person. This figure is constructed in 2 steps. First, a trend with seasonal adjustment was obtained from a regression of pre-pandemic income on a linear time trend and month fixed effects, separately for each income quartile, per equation (2). Second, deviations from the trend (plus 100) were then regressed on month-by-year × quartile fixed effects, per equation (1). The figure plots the coefficients on the month-by-year fixed effects separately for each quartile from the last regression.

Figure A4: Top and bottom earners.

Panel A: Bottom earners

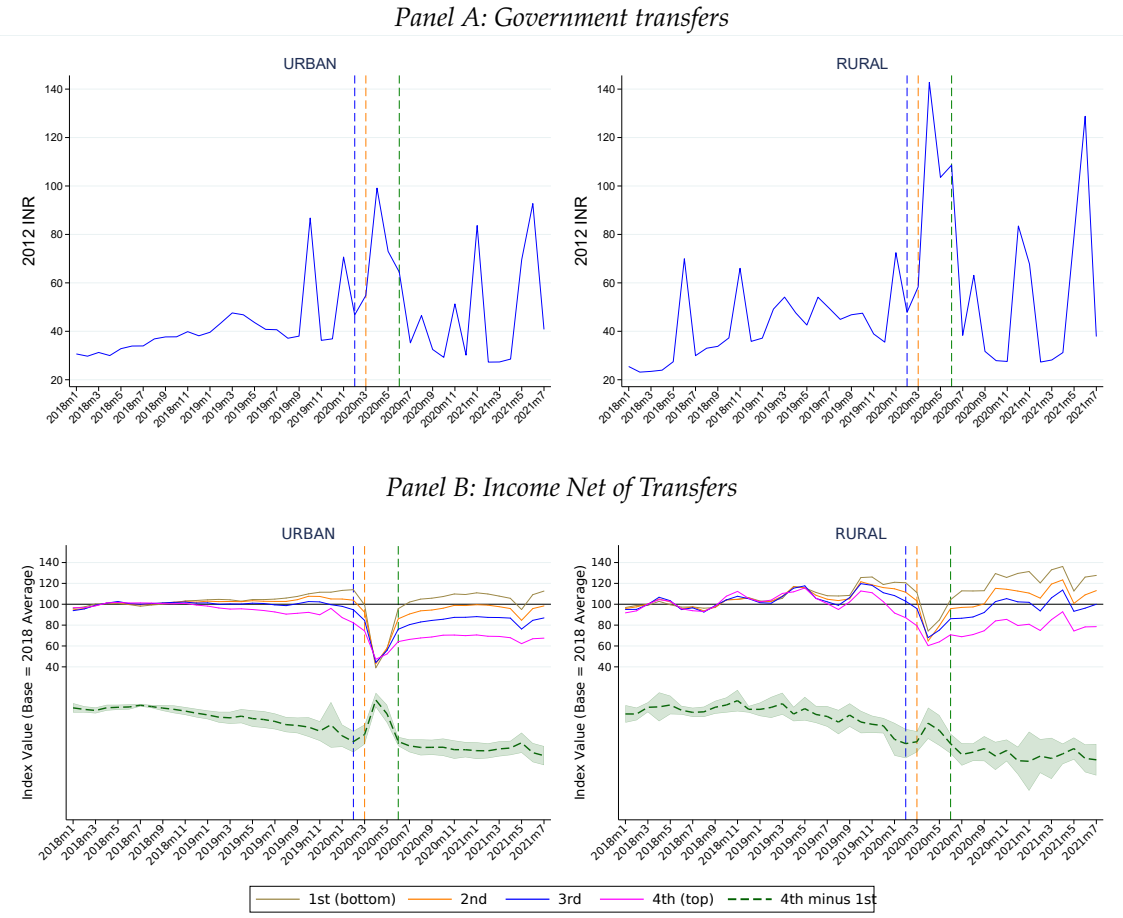


Panel B: Top earners



Note. The figure was constructed by calculating the weighted average of per capita income within indicated income quantiles in each state × urban status location, using individual member weights. These weighted means are obtained from the time fixed effects estimates in (1) using the default baseline. The units are an index where 100 is average 2018 income of a person. Dashed vertical lines in February 2020, March 2020 and June 2020 indicate the first month of the pandemic (blue), the month the national lockdown started (red) and the month the national lockdown ended (green).

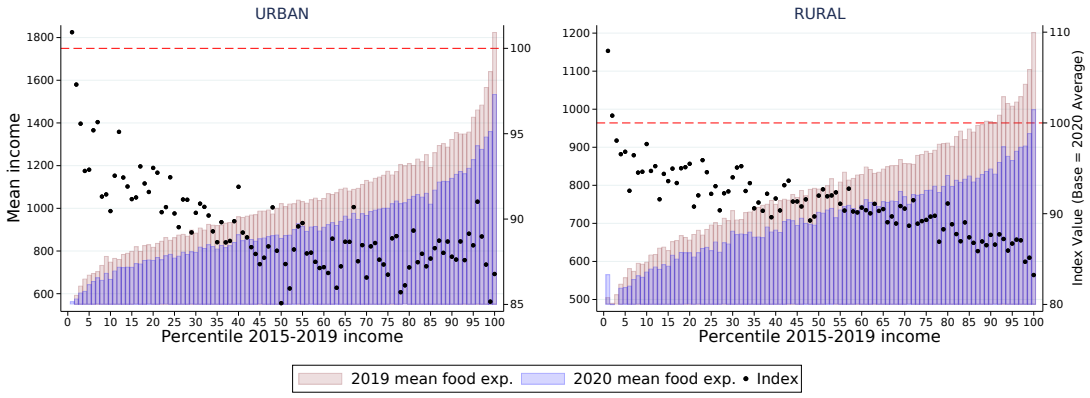
Figure A5: Role of Government Transfers in Income Inequality, 2019-2021



Note. Panel A presents the monthly average of government transfers per household in 2012 INR. Panel B presents a version of Figure 1B but with income excluding government transfers.

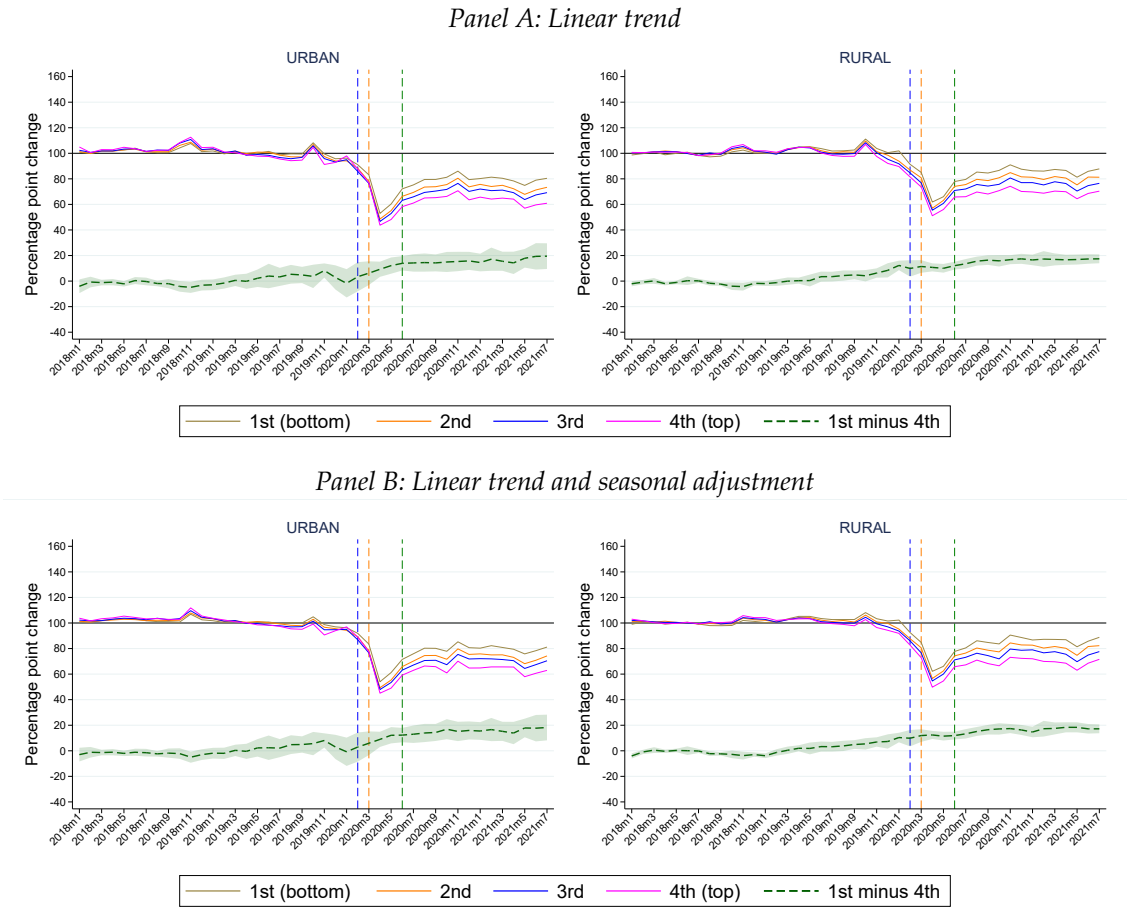
A4 Consumption inequality

Figure A6: Cross-sectional shifts in consumption



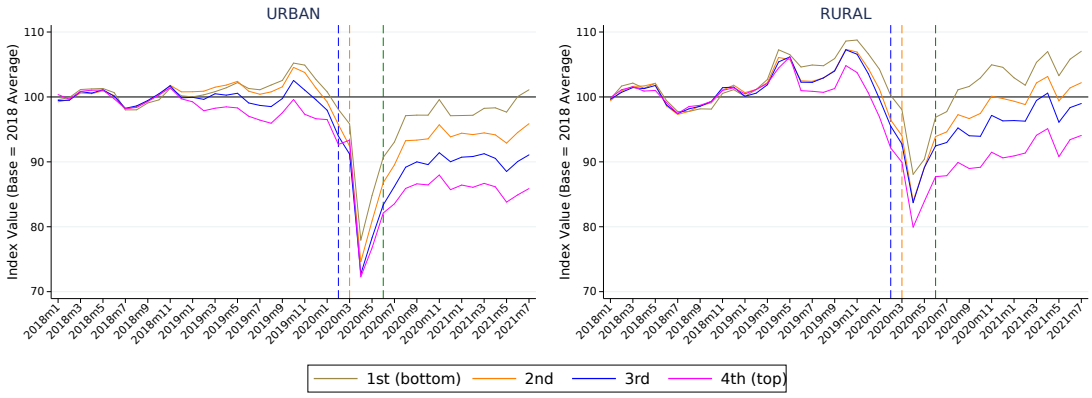
Note. First (second) column shows values for urban (rural) areas. The red (blue) bars show the mean 2019 (2020) consumption in each percentile. Note that mean 2020 consumption divided by 2019 mean consumption is not equal to the exact value of the index because the base value is specific to each person's state × urban status. The dashed red line demarcates 100%.

Figure A7: Consumption inequality with linear trends and seasonal adjustment.



Note. Panel A shows a consumption time series relative to a linear trend. The units are an index where 100 is average 2018 consumption of that person. A trend was obtained from a regression of pre-pandemic consumption on time, separately for each income quartile, per equation (2). Deviations from the trend (plus 100) were then regressed on month × quartile dummies, per equation (1). In Panel B, we add calendar month dummies to the trend regression to adjust for seasonal fluctuations.

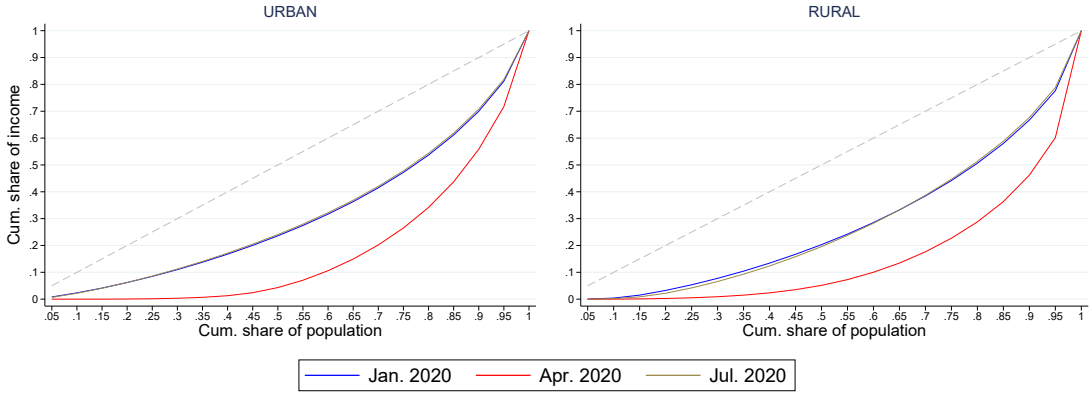
Figure A8: Relative Change in Food Expenditure by Income Quartiles



Note. Individuals were assigned to income quartiles at the state × community-type level calculated using individual's average 2015-2019 per capita incomes. This figure was constructed by first dividing the household food expenditure by the household size to calculate per capita food expenditure, then calculating the mean per capita food expenditure within income quartiles in their state × urban status locations, using individual member weights, and finally dividing by mean per capita local food expenditure in 2018 to create an index. Dashed vertical lines in February 2020, March 2020 and June 2020 indicate the first month of the pandemic (blue), the month the national lockdown started (red) and the month the national lockdown ended (green).

A5 Gini coefficient

Figure A9: Lorenz curves before and after the lockdown

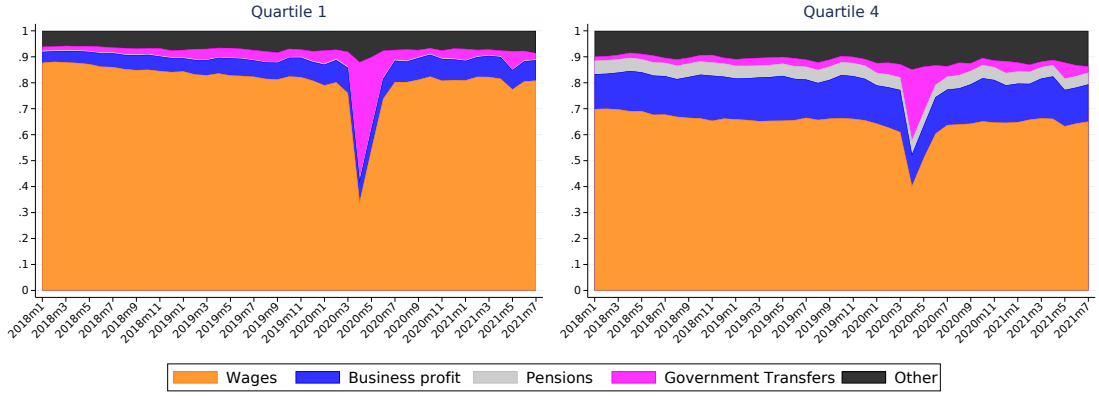


Note. Lorenz curves are obtained by sorting individuals on income, separately for rural and urban areas, and assigning each individual a place in the national income distribution, based on their survey weight. We then calculated the individual's share of the population based on their survey weight, as well as their share of the total national income. This is done separately for each reported month.

A6 Mechanisms

Figure A10: Share of different sources of income for bottom and top quartiles

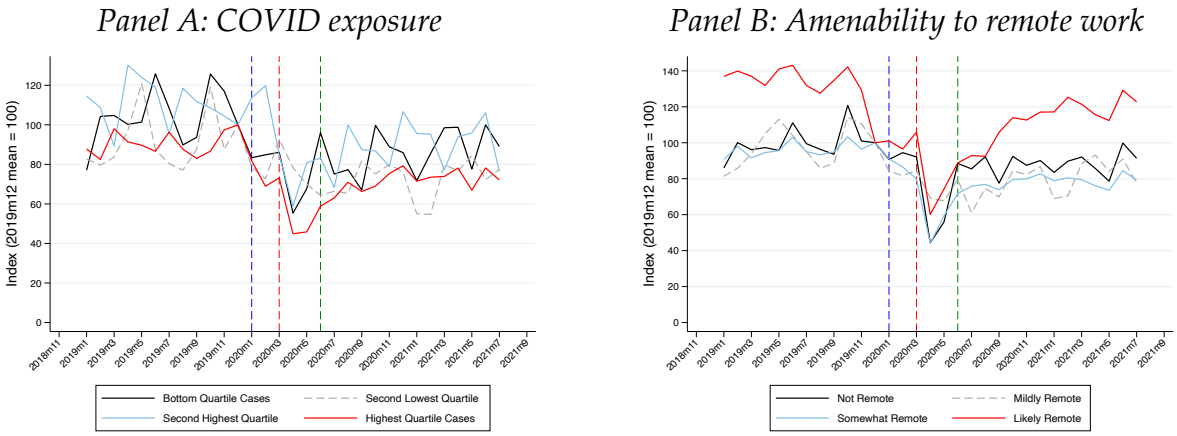
. [Can you match the categories here with those in Figure 4A, i.e., replace capital income with transfer income?]Done



Note. Share of income from different income sources. Capital income includes dividends, interest, rent, sale of assets outside of pension accounts. Other income includes government transfers, private transfers, value of agricultural goods produced for self-consumption, lottery winnings, insurance payouts.

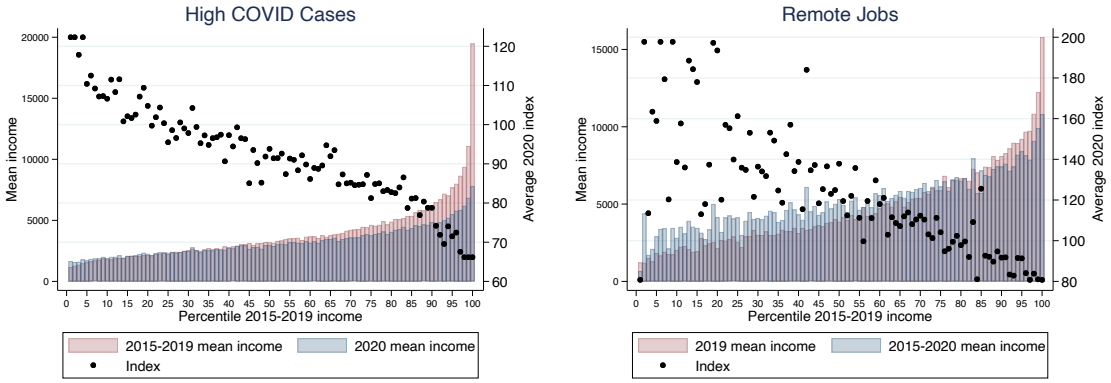
A7 Labor supply

Figure A11: Change in Income by Case Exposure and Ability to Work Remotely



Note. The left figure shows changes in labor income, normalized to the level in December 2019, for individuals divided into quartiles based on cumulative COVID case exposure, at the district level. The right figure divides occupations into four categories of exposure into remote work using the Dingel and Neiman (2020) measure of remote work, also normalized to December 2019. The lowest category has a Dingel and Neiman (2020) remote score of zero; the next category lies between zero and six percent; the next category is between six and 37 percent; and the highest category of remote work consists of occupations above 37 percent.

Figure A12: Distributional Consequences of Case Exposure and Remote Work

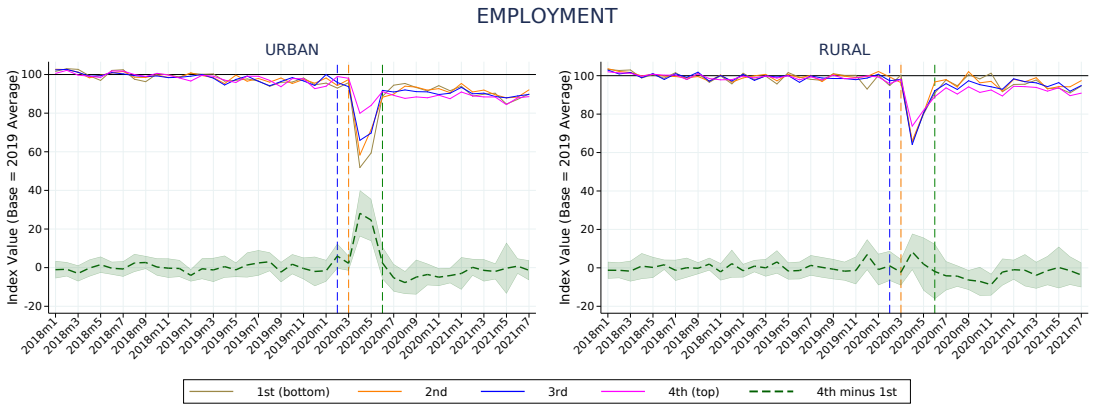


Note. The right y-axis shows the index value and the left y-axis shows average income. The black dots represent the index, the red bars show the mean 2015–2019 income in each percentile, the blue bars show the mean 2020 income in that percentile. Note that mean 2020 income divided by 2015–2019 mean income is not equal to the exact value of the index because the base value is specific to each person's state $\times$ urban status. The left panel restricts to individuals in the highest quantile of COVID confirmed case exposure; the right panel restricts to individuals in the highest category of possible remote work status.

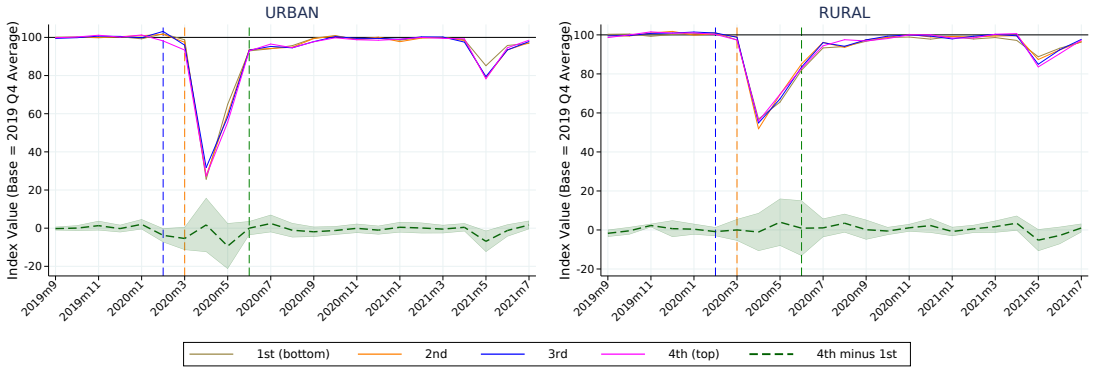
A8 Labor demand

Figure A13: Margins of income, by region

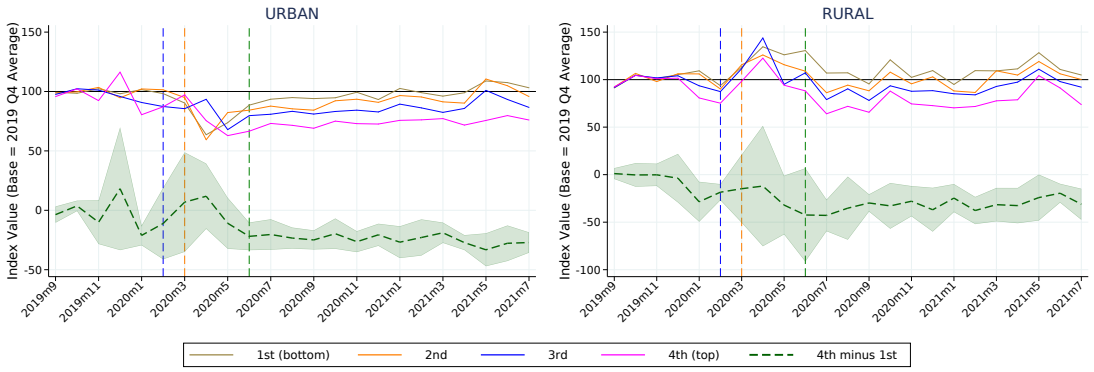
Panel A: Employment



Panel B: Hours per day employed



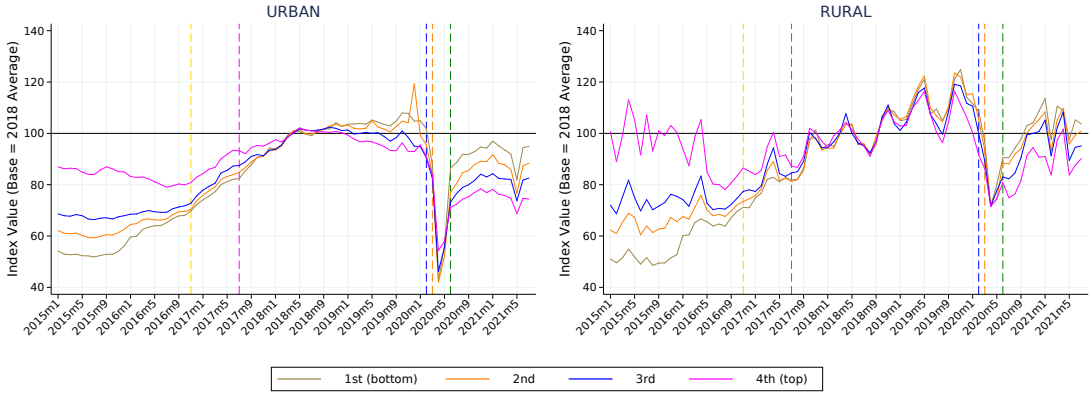
Panel C: Wages per hour



Note. We use the same income quartile categorization as before but only include individuals aged 15–65 for whom we have employment status. For hours and wage, we include individuals who report being employed but working 0 hours. Average 2019 employment rates are the baseline for employment rate. Average September-December 2019 hours and wage income per hour are the baseline for hours and wages because CPHS does not measure hours before September 2019. Green lines report differences between the 4th and 1st quartile; green shaded are 95% confidence intervals from a regression of the margin on month fixed effects, quartiles, and the interacted of month fixed effects and quartiles.

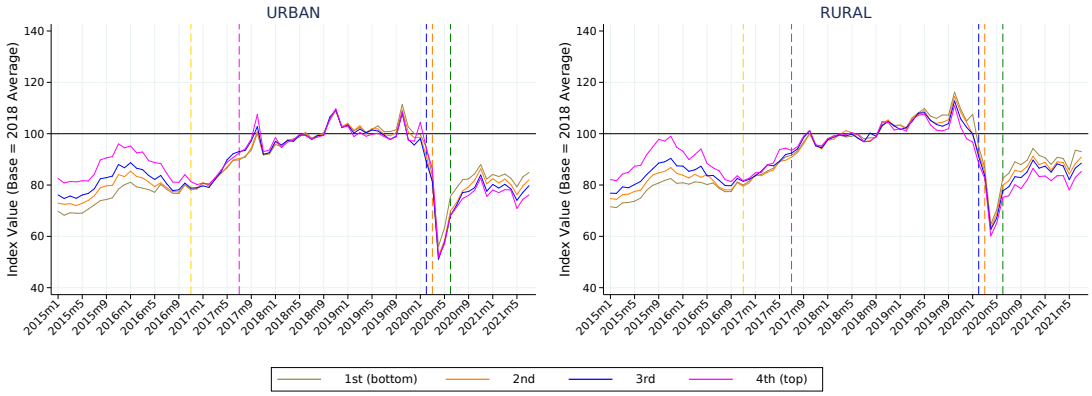
A9 Discussion

Figure A14: Relative Change in Income by 2015 Income Quartiles



Note. Individuals were assigned to income quartiles calculated using average 2015 per capita incomes. The figure was constructed by first dividing the household income by the household size to calculate per capita income, then calculating the mean within income quartiles in their state × urban status locations, using individual member weights, and finally dividing by mean per capita income in 2018 to create an index. Dashed vertical lines in November 2016, July 2017, February 2020, March 2020 and June 2020 indicate the month demonetization was announced (yellow), the month substantial amounts of new currency was introduced (magenta), the first month of the pandemic (blue), the month the national lockdown started (orange) and the month the national lockdown ended (green).

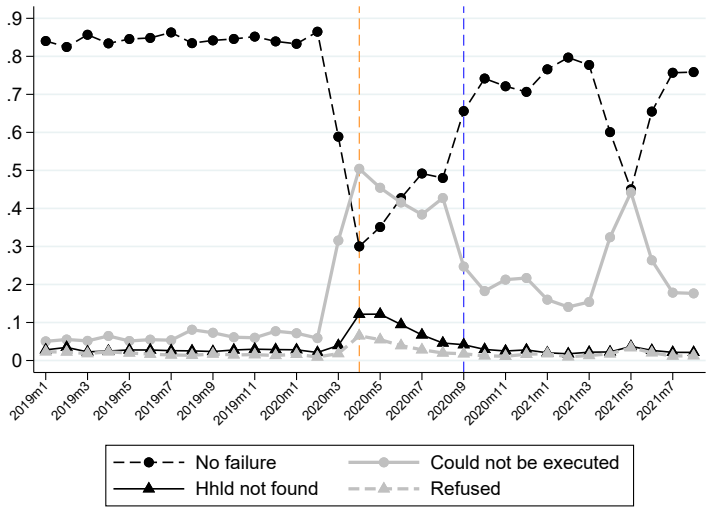
Figure A15: Relative Change in Consumption by 2015 Income Quartiles



Note. Individuals were assigned to income quartiles calculated using average 2015 per capita incomes. The figure was constructed by first dividing the household consumption by the household size to calculate per capita consumption, then calculating the mean per capita consumption within income quartiles in their state × urban status locations, using individual member weights, and finally dividing by mean per capita local consumption in 2018 to create an index. Dashed vertical lines in November 2016, July 2017, February 2020, March 2020 and June 2020 indicate the month demonetization was announced (yellow), the month substantial amounts of new currency was introduced (magenta), the first month of the pandemic (blue), the month the national lockdown started (orange) and the month the national lockdown ended (green).

A10 Appendix: Data

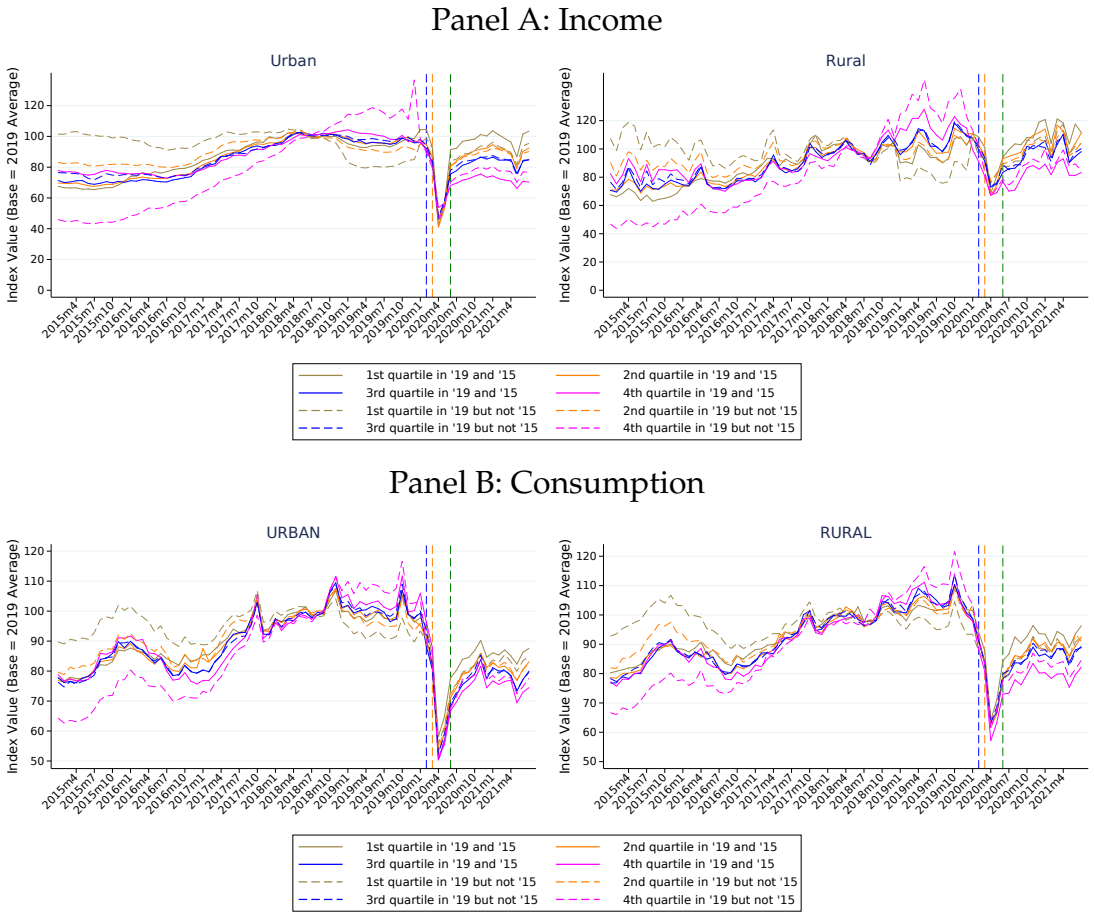
Figure A16: CPHS non-execution and non-response rates during 2020.



Note. Orange line indicates first month of phone surveys. Blue line indicates month that in-person surveys resumed. The sample includes all households. "Could not be executed" (non-execution) includes both CMIE's decision not to contact a household and its inability to speak to a household member because, e.g., no one answered the door ("door-lock"). "Household not found" means CMIE attempted to contact the household but surveyors were unable to locate the household.

A11 Appendix: Mean reversion

Figure A17: Change in Relative Income and Consumption among Persistent and Transient Members of Each 2019 Income Quartile, 2015-2021



Note. The income figures were constructed in five steps. First, we divide household monthly income by the household size to calculate per capita income for each member of the household. Second, we take the monthly average of a person's per capita income over 12 months (for 2015 and for 2019). Third, we weight individuals to be representative and then we sort weighted individuals into quartiles of average monthly per capita income for that person's state × community type (rural or urban) and that year. Fourth, we separate the set of all (weighted) individuals in a quartile in 2019 into two groups. One is individuals who were in that same quartile in 2015, using quartile definitions from that year. The other is people who were not in that same quartile in 2015. Fifth, to create an index, we divide (a) the average, weighted monthly income of all members of a quartile × persistence status group by (b) the average, weighted monthly income of all members of that quartile × persistence status group in 2018. For the consumption figures, we do the same thing except at the last step, we divide monthly consumption of the average member of a quartile × persistence status group by the average monthly consumption of that group in 2018. Dashed vertical lines in February 2020, March 2020 and June 2020 indicate the first month of the pandemic (blue), the month the national lockdown started (orange) and the month the national lockdown ended (green).

B Tables

B1 Extreme poverty

Table A1: Percentage of population under World Bank's extreme poverty measure, by region and period

Income			
	(1) Pre-Lockdown	(2) Lockdown	(3) Post-Lockdown
Urban	0.0168***	0.273***	0.0148***
	(0.00360)	(0.0322)	(0.00379)
Rural	0.108***	-0.0189	0.0461**
	(0.0258)	(0.0414)	(0.0173)
N	21830439	21830439	21830439
R^2	0.089	0.089	0.089
Consumption			
	(1) Pre-Lockdown	(2) Lockdown	(3) Post-Lockdown
Urban	0.0287**	0.140***	0.0261
	(0.0127)	(0.0313)	(0.0174)
Rural	0.0448	0.0520	0.00616
	(0.0287)	(0.0435)	(0.0250)
N	21830439	21830439	21830439
R^2	0.028	0.028	0.028

Notes. The extreme poverty measure is nominal income (top panel) and consumption (bottom panel) below PP-adjusted INR equivalent of $1.90 per day. The table presents the average level of extreme poverty in each of the periods in columns for the regions listed in rows. Standard errors are presented parenthesis. Significance levels: * 10% ** 5% *** 1%.

B2 Consumption inequality

Table A2: Test of consumption smoothing (of total expenditure).

	(1)	(2)	(3) Rural	(4) Urban
Ln(aggregate consumption)	0.6533*** (0.0578)		0.8790*** (0.0254)	0.5302*** (0.0560)
Ln(income)	0.0980*** (0.0146)	0.1297*** (0.0163)	0.0506*** (0.0081)	0.1585*** (0.0170)
Ln(income) × Pandemic	-0.0052*** (0.0014)	-0.0233*** (0.0016)	-0.0012 (0.0009)	-0.0069*** (0.0017)
N	3773651	3773651	1237376	2536275
R^2	0.715	0.691	0.700	0.721

Notes. The regressions covers 2019-July 2021. Aggregate consumption is equal to mean household consumption at the state × community type level. Household fixed effects included. Standard errors clustered at the state × urban level. Significance levels: * 10% ** 5% *** 1%.

Table A3: Test of smoothing (of total expenditure), by income quartiles.

	(1)	(2)	(3) Rural	(4) Urban
Ln(aggregate consumption)	0.5576***		0.8621***	0.4436***
	(0.0748)		(0.0218)	(0.0572)
2nd quartile	0.0921***		0.0229	0.0843***
	(0.0253)		(0.0148)	(0.0198)
3rd quartile	0.1653***		0.0205	0.1660***
	(0.0476)		(0.0233)	(0.0386)
4th quartile	0.2520***		0.0037	0.2591***
	(0.0789)		(0.0302)	(0.0690)
Ln(income)	0.0953***	0.1345***	0.0497***	0.1317***
	(0.0122)	(0.0132)	(0.0077)	(0.0095)
2nd quartile	-0.0025	-0.0032	-0.0009	0.0058
	(0.0039)	(0.0050)	(0.0032)	(0.0072)
3rd quartile	-0.0052	-0.0110	-0.0022	0.0145
	(0.0063)	(0.0085)	(0.0049)	(0.0110)
4th quartile	0.0056	-0.0045	0.0054	0.0664***
	(0.0090)	(0.0116)	(0.0055)	(0.0118)
Ln(income) × Pandemic	-0.0047***	-0.0218***	0.0014	-0.0068***
	(0.0013)	(0.0013)	(0.0010)	(0.0014)
2nd quartile	0.0003	-0.0017***	-0.0017**	0.0005
	(0.0005)	(0.0005)	(0.0007)	(0.0006)
3rd quartile	0.0007	-0.0030***	-0.0029***	0.0015
	(0.0009)	(0.0008)	(0.0010)	(0.0011)
4th quartile	0.0005	-0.0051***	-0.0058***	0.0031*
	(0.0016)	(0.0012)	(0.0018)	(0.0017)
N	2937692	2937692	977542	1960150
R^2	0.725	0.700	0.713	0.730

Notes: The regressions covers 2019-July 2021. Aggregate consumption is equal to mean household consumption at the state × community type level. Household fixed effects included. Standard errors clustered at the state × urban level. Specification (1) drops our measure of aggregate consumption, (3) only includes observations on rural households, and (4) only includes observations on urban households. Significance levels: * 10% ** 5% *** 1%.

Table A4: Test of consumption smoothing (of food expenditure), by income quartiles.

	(1)	(2)	(3) Rural	(4) Urban
Ln(aggregate consumption)	0.2674***		0.4274***	0.2038***
	(0.0404)		(0.0294)	(0.0323)
2nd quartile	0.0543**		0.0318*	0.0447*
	(0.0210)		(0.0174)	(0.0219)
3rd quartile	0.0763**		0.0066	0.0715**
	(0.0353)		(0.0213)	(0.0337)
4th quartile	0.1124*		-0.0180	0.1176*
	(0.0573)		(0.0330)	(0.0588)
Ln(income)	0.0703***	0.0891***	0.0430***	0.0923***
	(0.0078)	(0.0078)	(0.0085)	(0.0067)
2nd quartile	-0.0038	-0.0036	-0.0038	0.0021
	(0.0034)	(0.0038)	(0.0037)	(0.0044)
3rd quartile	-0.0090	-0.0119*	-0.0086	0.0048
	(0.0056)	(0.0066)	(0.0056)	(0.0074)
4th quartile	-0.0082	-0.0133	-0.0063	0.0247**
	(0.0070)	(0.0084)	(0.0069)	(0.0092)
Ln(income) $\times$ 2020	-0.0018	-0.0100***	0.0015	-0.0031*
	(0.0015)	(0.0013)	(0.0017)	(0.0018)
2nd quartile	0.0003	-0.0009*	-0.0007	0.0005
	(0.0006)	(0.0005)	(0.0010)	(0.0007)
3rd quartile	0.0008	-0.0009	-0.0012	0.0012
	(0.0009)	(0.0007)	(0.0017)	(0.0009)
4th quartile	0.0013	-0.0012	-0.0036	0.0035**
	(0.0014)	(0.0011)	(0.0023)	(0.0014)
N	2937654	2937654	977536	1960118
R^2	0.705	0.695	0.691	0.709

Notes: The regressions cover 2019-July 2021. Aggregate consumption is equal to mean household consumption at the state $\times$ community type level. Household fixed effects included. Standard errors clustered at the state $\times$ urban level. Specification (1) drops our measure of aggregate consumption, (3) only includes observations on rural households, and (4) only includes observations on urban households. Significance levels: * 10% ** 5% *** 1%.

Table A5: Annualised death rates by 2018 income tercile

	Annualised death rates (%)					
	(1)	(2)	(3)	(4)	(5)	(6)
Pandemic	0.113		0.237		0.0823	
	(0.0689)		(0.122)		(0.0801)	
Pandemic × 2nd tercile	0.132		-0.0223		0.177	
	(0.0867)		(0.121)		(0.109)	
Pandemic × 3rd tercile	0.357***		0.235		0.387**	
	(0.0986)		(0.129)		(0.146)	
Wave 1		-0.0247		0.0416		-0.0411
		(0.0767)		(0.135)		(0.0893)
Wave 1 × 2nd tercile		0.239*		-0.00842		0.338**
		(0.102)		(0.141)		(0.129)
Wave 1 × 3rd tercile		0.367**		0.154		0.569**
		(0.114)		(0.149)		(0.181)
Wave 2		0.371***		0.591**		0.315**
		(0.102)		(0.190)		(0.118)
Wave 2 × 2nd tercile		-0.0682		-0.0459		-0.124
		(0.122)		(0.174)		(0.151)
Wave 2 × 3rd tercile		0.328*		0.372		0.0501
		(0.147)		(0.194)		(0.193)
2019 mean	1.160***	1.160***	1.170***	1.170***	1.158***	1.158***
	(0.0463)	(0.0463)	(0.0746)	(0.0746)	(0.0541)	(0.0541)
2nd tercile	-0.172**	-0.172**	-0.150*	-0.150*	-0.184**	-0.184**
	(0.0540)	(0.0540)	(0.0715)	(0.0715)	(0.0671)	(0.0671)
3rd tercile	-0.232***	-0.232***	-0.232**	-0.232**	-0.242**	-0.242**
	(0.0567)	(0.0567)	(0.0726)	(0.0726)	(0.0795)	(0.0795)
Sample	All	All	Urban	Urban	Rural	Rural
N	2816368	2816368	1852668	1852668	963700	963700

Note. Estimates are from a regression model based on equation (1) in Malani and Ramachandran (2021), with the addition of a income tercile indicator and income tercile indicator interacted with the pandemic or wave indicator. For each individual we calculate the income per capita in 2018. We compute the household's income percentile in their homogeneous region and region type (urban/rural). Households between 33 and 67 percentile are in income tercile 2 and households between 67 and 100 percentile are in income tercile 3. Columns 1 and 2 includes all data, columns 3 and 4 include only urban regions and columns 5 and 6 include only rural regions. Sample includes only consecutive observations and is weighted to be nationally representative. Standard errors clustered at the village/ward × month level are reported in parentheses. $p < 0.05/0.01/0.001$.

B3 Gini coefficient

Table A6: Income quartile transition matrix from 2018 to 2019

	2019 quartile			
2018 quartile	1st quartile	2nd quartile	3rd quartile	4th quartile
1st quartile	62.62	26.57	8.31	2.50
2nd quartile	24.58	43.28	24.73	7.41
3rd quartile	8.15	24.81	43.79	23.25
4th quartile	2.46	6.64	23.06	67.84

Note. This table presents the probability ($\times$ 100) that a person in the (row) income quartile in 2018 was in the (column) quartile in 2019. 2018 and 2019 income quartiles are defined separately for each year based on the distribution of average per capita monthly income across all months.

B4 Mechanisms

Table A7: Attribution of changes in inequality during the pandemic to different components of household income.

Components of Income	Change in Inequality Due to:					
	Trend-weighted Change in *Share*	Share-Weighted Change in *Amount*	Trend-Weighted Change in *Share*	Share-Weighted Change in *Amount*	Trend-Weighted Change in *Share*	Whare-weighted Change in *Amount*
	TOTAL		URBAN		RURAL	
Total Income		-39.741		-36.347		-41.744
Labor Income	5.412	-24.929	4.865	-27.325	5.753	-22.738
Business Income	-6.701	-9.378	-3.784	-4.462	-9.127	-12.853
Transfer Income	0.185	-0.333	0.292	-0.777	0.132	-0.040
Pension Income	-1.938	-0.283	0.292	-0.777	-0.713	-0.128
Other Income	-0.212	-1.564	-0.525	-0.789	-0.044	-1.987

Note. Changes are from 2019 average to July 2021. Units are percentage points. Trend-weighted change in share is $[S_j^1 - S_j^4](d \log C_j^1 / dt)$, where S_j^q is average share of quartile q income attributable to component j and c_j^q is average amount of income from component j in quartile q. Change in share weighted amount is $S_j^4[(logC_j^1/dt) - (logC_j^4/dt)]$. Source is Consumer Pyramids Household Survey.

B5 Labor supply

Table A8: Household Labor Supply over the Pandemic

Panel A: Urban Areas

	(1)	(2)	(3)	(4)	(5)	(6)
New Confirmed Cases	2.8203		2.1552	2.8203		2.1552
	(2.1466)		(1.6069)	(2.1465)		(1.6068)
Top Half 2019 Income		0.1498***	0.1499***		0.1967***	0.1967***
		(0.0080)	(0.0081)		(0.0146)	(0.0146)
Cases × High Income			-0.6684			-0.6684
			(2.9631)			(2.9629)
Pandemic Period				-0.5448***	-0.5180***	-0.5187***
				(0.0095)	(0.0090)	(0.0090)
Pandemic × High Income					-0.0469***	-0.0468***
					(0.0121)	(0.0123)
N 111755	111755	111755	642695	642695	642695	
R^2 0.000	0.007	0.007	0.017	0.021	0.021	

Panel B: Rural Areas

	(1)	(2)	(3)	(4)	(5)	(6)
New Confirmed Cases	-0.6017		-5.4169**	-0.6017		-5.4169**
	(1.7175)		(2.3411)	(1.7172)		(2.3407)
Top Half 2019 Income		0.0808***	0.0794***		0.1200***	0.1200***
		(0.0121)	(0.0122)		(0.0195)	(0.0195)
Cases × High Income			5.9592**			5.9592**
			(2.6112)			(2.6108)
Pandemic Period				-0.3176***	-0.3007***	-0.2994***
				(0.0114)	(0.0119)	(0.0119)
Pandemic × High Income					-0.0392***	-0.0407***
					(0.0126)	(0.0126)
N	64375	64375	64375	368404	368404	368404
R^2	0.000	0.004	0.004	0.010	0.012	0.012

Notes: This table regresses log reservation wages, drawn from a Roy Model, against income percentiles and confirmed COVID cases. Columns (1)–(3) focus on the pandemic period (2020 Q2) and regress reservation wages against confirmed cases, the income, and an interaction of cases and income. Columns (4)–(6) include both the pre-pandemic period of 2019 Q2 as well as 2020 Q2, with an interaction with the pandemic period. Reservation wages are measured as rupee income, and confirmed cases are divided by 1×10^6 for ease of interpretation. Panel A subsets on urban areas, and Panel B focuses on rural areas. Significance levels: * 10% ** 5% *** 1%.

B6 Labor demand

Table A9: Mapping expenditure categories to agriculture, manufacturing and service sectors.

Sector	Expenditure category
Agriculture	Whole-grain cereal, pulses, edible oils, dry spices, dry fruit, noodles and flakes, sugar, vegetables and fruit, potatoes and onions, milk and milk products, bread, biscuits, salty snacks, jam, ketchup, pickles, meat, eggs, fish, ghee, mithai, cigarettes and tobacco.
Manufacturing	Processed cereal, health supplements, tea and coffee, beverages and bottled water, chocolate, cakes, ice-cream, ready-to-eat food, baby food, other food, liquor, clothes and footwear, cosmetics and toiletries, appliances, electronic storage, toys, power and fuel, lighting, furniture, utensils.
Services	Restaurants, entertainment, bills and rent, transport, communication and information, education, health expenses, loan installments, domestic help, car repair, remittances, social obligations, religious obligations, general insurance, vacation, pocket money, pets, professional services, renovations.

Table A10: Mapping employment categories to SIC codes and agriculture, manufacturing and service sectors.

Sector	SIC code	Employment category
Agriculture.	Agriculture, forestry and fishing.	Agriculture-allied activities; crop cultivation; fishing; forestry, including wood cutting; fruits and vegetable farming; plantation crop cultivation; poultry farming, animal husbandry and vermiculture.
Manufacturing.	Mining and quarrying.	Mines.
	Manufacturing.	Automobiles and other transport equipment manufacturers; cement, tiles, bricks, ceramics, glass and other construction materials; chemical industries; food industries; footwear and other leather industries; gems and jewelry; handicraft industries; machinery manufacturers; metal industries; pharmaceutical manufacturers; soaps, detergents, cosmetics, toiletries; textile industries.
	Construction and real estate.	Real estate and construction.
Services.	Utilities.	Utilities.
	Retail and wholesale trade.	Retail trade; wholesale trade.
	Transportation and storage.	Communication, post and courier.
	Accomodation and food services.	Hotels and restaurants.
	Information and communication.	IT and ITES; media and publishing.
	Financial services.	Financial services.
	Professional services.	Personal professional services.
	Non-professional services.	Personal non-professional services.
	Public administration and defense.	Defense services, public administrative services.
	Education.	Education.
	Health care.	Health care.
	Entertainment, sports, tourism.	Entertainment and sports; travel and tourism.

C Data

C1 Detailed description of CPHS

The CPHS sample covers nearly all states in India except for a few states in the north-east (e.g., Nagaland, Mizoram, Arunchal Pradesh) that are difficult to sample because of instability (Vyas, 2020d).

CPHS divides each state into homogeneous regions, clusters of districts with similar features.[13] Each region is divided into rural and urban strata, where rural regions are villages as defined by the Indian Census. The urban strata is further subdivided into four sub-strata defined by town size. The primary sampling units are villages and towns. Thirty villages were randomly selected from rural strata of each region. For urban strata, a random subsample of towns in each sub-strata are selected. The ultimate sampling units are households. In each selected village 16 households were selected by systematic random sampling (every n^{th} household on a street, where n is a random number between 5 and 15). In cities, 21 Census Enumeration Blocks (CEB) were randomly selected. In each CEB, 16 households were selected via systematic random sampling. Sample households were selected to be representative at the level of urban and rural areas of regions.

The CPHS started surveying in January 2014. We use the data starting January 2015, the point CPHS suggests that they had stabilized their survey quality.

C2 Issues with CPHS data

Sample churn and non-response

The sample in the CMIE data churns somewhat over time Vyas (2020c). On average 2.1% of households are lost in each four month wave and 2.4% are added in each wave to replace lost households and to grow the sample over time. Weights are included to ensure that the sample remains representative of its region. Prior to the COVID epidemic, response rates were roughly 84%. Non-response was due more to inability to reach all households in the allotted 4 months for each sampling wave more than refusal to be surveyed. Separate weights are included in an attempt to correct for non-response.

During the pandemic, response rates fell (Vyas, 2020a). CPHS is ordinarily an in-person survey. The household response rate (responding households/sample house-

[13] The features are similar agro-climactic conditions, urbanization levels, female literacy, and average household size.

holds) prior to the pandemic was typically around 85%. When India declared a national lockdown on March 24, 2020, in person surveys had to cease. CPHS switched temporarily to a phone survey. Moreover, survey managers, rather than surveyors, conducted phone surveys to keep up the quality of those surveys. Because there are fewer managers than surveyors, CPHS decided to call only half the households in each strata (defined as homogeneous region $\times$ community type). As a result, overall response rates fell. Figure A16 shows that the response rate of the subset of households that were contacted fell to roughly 60% and responding households constituted roughly 35% of the full sample at the height of the lockdown in April and May 2020. When CPHS finished its second round in August 202, it returned to in-person surveys. However, the response rate only rose to 75%.[14]

Despite the drop in response rates, CMIE was able to maintain the distribution of surveys across two dimensions the same as prior to the pandemic. First, the ratio of rural to urban households was roughly 35:65 pre- and post-lockdown, with only a two week deviation to 43:57 when lockdown was declared. The distribution of households across states shifted a bit in favor of rural states, but was roughly the same as pre-pandemic. Second, the distribution across income also remained roughly the same. The fraction earning between ₹150,000–300,000 per annum was 45% before and after lockdown. However, sampling at the extremes of the distribution did change. The share earning $\geq$ ₹500,000 fell from 12.9% to 9%; the share earning $\leq$ ₹150,000 increased from 23% to 29.1%. Of course, some of this change may be a reduction in income due to COVID[15]

The low response rates during round 2 in 2020 are partly addressed by using the responding sample as the denominator in our estimate of rates of various economic outcomes. We address non-random response that makes the same differently representative of the population by using CMIE-supplied weights intended to ensure the responding population has the same demographics as the population in each strata.

Representativeness of sample

Dreze and Somanchi (2021) argues that CPHS undersamples the poor based on evidence that it yields both higher levels of literacy and faster improvement in literacy than government surveys. Of course, the fact of difference between CPHS and government surveys of literacy is not dispositive of whether CPHS is biased since it is possible that the gov-

[14]There was also a drop in response rates during wave 2, but there was no switch to telephone interviews and the dip was not as severe as during the lockdown.

[15]CMIE also reports a shift in the fraction of households across occupational groups.

ernment surveys are the ones that are off. After all, the government has taken steps to suppress data (e.g., the 2017-18 consumption survey by the National Statistical Survey Office) that it finds unflattering. Moreover, government surveys are known to give different estimates of items like slums populations, with the differences driven by the policy aim of the survey.[16]

Dreze and Somanchi suggest that CPHS is the one likely to be wrong because its frame samples more from the main streets of villages than from outskirts, where the poor tend to live. CMIE has responded that its sampling does get to outskirts and that the bias has not changed over time because that sampling frame is largely fixed and that it method for selection (of new households) has been constant (Vyas, 2021).

We address the problem of representativenes by conditioning on where in the income distribution a person falls in 2019. If bias is due to where the individual resides, this approach even addresses the problem that social mobility may cause a change in a person's income percentile over time. To the extent that Dreze and Somanchi's argument suggests that our calculation of 25th percentile is wrong (i.e., too high) because CPHS undersamples the very poor, we provide evidence that lower percentiles in CPHS do relatively better than even the 25th percentile in CPHS (Figure A4).

Survey content and comparisons

The survey is conducted at the household level but measures both individual and household level variables. It measures employment status, time use, and occupation for each member of the household once every four months. It measures income for each household member and the overall household and consumption for the household every month by asking members to recall income and consumption each of the last four months. The survey is conducted on a smart device and gathers data on up to 12 members of each household.

The CPHS has analogues in Indian NSSO surveys on labor statistics and on consumer expenditures. It is difficult to compare the CPHS to the consumer expenditure survey because the NSSO's 2017–2018 consumer expenditure survey was rejected by the government and thus not released (Vyas, 2020b). The previous one was from 2011–2012. Nonwithshanding Dreze and Somanchi (2021), Abraham and Shrivastava (2019) show that the CPHS and NSSO produce similar results for male workers.

[16]For example, public health officials in Mumbai in private conversations have noted that surveys of slum populations by the public health department tend to generate higher estimates of slum population because higher numbers in slums are more likely to generate large appropriations for the department.

C3 RBI inflation data

We obtain data on overall and constituents[17] of inflation from the Reserve Bank of India (Reserve Bank of India, 2020). The data are available at the monthly level for rural and urban areas of each state; the base year is 2012. We obtain relative prices from constituents price indices using the rural and urban constituent weights reported by the RBI (Bhoi et al., 2020).[18]

C4 COVID cases, lockdown rules and mobility

We obtain data on COVID cases and deaths from `www.covid19India.org`, which compiles reports from government sources across the country. We obtain data on national lockdown severity from the Oxford Covid-19 Government Response Tracker (Hale et al., 2020) and district-level mobility from Google Mobility Reports (Google LLC, 2021).

D Methods

D1 Inferences about rural-urban differences

We make statistical inferences about the differences in outcomes across urban and rural areas during the pre-pandemic (prior to February 2020), lockdown (March-May 2020 inclusive), and post-lockdown period (June 2020 and after) using the following regression model:

$$\begin{aligned} y_{it} = \beta_0 + \beta_{L0}\mathbb{1}(t \in lock) + \beta_{P0}\mathbb{1}(t \in postlock) + \beta_1\mathbb{1}(i \in rural) \\ + \beta_{L1}\mathbb{1}(t \in lock,\ i \in rural) + \beta_{P1}\mathbb{1}(t \in postlock,\ i \in rural) + u \end{aligned} \tag{6}$$

where y_{it} is some outcome (not quartile specific), *lock* (*postlock*) is the set of time periods during lockdown (after lockdown), and u is a regression error term. To make statistical inferences about differences in inequality across urban and rural India, we use a variant

[17] Food, clothing, fuel and light, transport and communication, intoxicants, housing, household goods and services, health and education, and others.

[18] CPI for items other than constituent g ($\neg g$) calculated as $\text{CPI}_{\neg g} = (\sum_{k \neq g} w_k \text{CPI}_k)/(\sum_{k \neq g} w_k)$, where k indexes constituents.

of this last regression that includes interactions with a 4th quartile indicator:

$$\begin{aligned} y_{it} = &[\beta_0 + \beta_{L0}\mathbb{1}(t \in lock) + \beta_{P0}\mathbb{1}(t \in postlock) + \beta_1\mathbb{1}(i \in rural) \\ &+ \beta_{L1}\mathbb{1}(t \in lock,\ i \in rural) + \beta_{P1}\mathbb{1}(t \in postlock,\ i \in rural) \\ &+ [\gamma_0 + \gamma_{L0}\mathbb{1}(t \in lock) + \gamma_{P0}\mathbb{1}(t \in postlock) + \gamma_1\mathbb{1}(i \in rural) \\ &+ \gamma_{L1}\mathbb{1}(t \in lock,\ i \in rural) + \gamma_{P1}\mathbb{1}(t \in postlock,\ i \in rural)] \times \mathbb{1}(i \in q4) + u \end{aligned} \tag{7}$$

where $q4$ is the set of people in the 4th quartile.

D2 Gini coefficient

To obtain Gini coefficients separately across urban and rural regions, we first sort households by income (consumption) within-urban/rural region status at a given point in time. We then scale each household's share of population and national income (consumption) by the household's survey weight. Third, the area under the Lorenz curve is approximated by multiplying the scaled population share by the scaled cumulative national income (consumption) share and summing across households in the relevant month. Finally, we calculate the Gini coefficient by subtracting the area of the Lorenz curve from 1/2. We do this separately for each month to generate a time-series of Gini coefficients.

D3 Social mobility

To measure social mobility, we create a version of Shorrock's index (Shorrocks, 1995). First, we sort households by yearly income using monthly reported income, averaged separately for the years 2018, 2019, and 2020. Households are assigned a quartile for each year according to their ranking in that year. Urban and rural households are sorted separately for purposes of assignment to quartiles. Second, we calculate a transition matrix (Sharrock's index) for 2018-2019 and then for 2019-2020 to show the probability that a person in quartile q in year t is in quartile q' in year $t+1$. Finally, we calculate how much social mobility changed during the pandemic by subtracting the transition matrix for 2018-2-19 from the matrix for 2019-2020. Our implicit assumption is that, but for the pandemic, the transition matrix for 2019-2020 would have been the same as that for 2018-2019.

E Incidence

E1 Poverty

Here we consider two other measure of poverty from Azim Premji University (2021) that show slightly different degrees of poverty increases (Figure A2). Using India's National Minimum Wage (income of ₹1909 and ₹2256 per capita month in rural and urban areas, in 2012 terms), poverty spiked sharply from 40% to 70% in rural areas and 25% to 65% in urban areas during the lockdown. It only partly recovered, hovering at 50 and 30%, respectively, by the end of the year. By contrast, the poverty line defined by the Indian government's 7th Central Pay Commission (income of ₹4,660 per capita per month) only increased gradually in rural areas and more moderately (and $< 10\%$) than the Minimum Wage poverty measure in urban areas (from 65 to 85%) during the lockdown.

E2 Mean reversion

We have documented that the income of top quartiles fell relatively more than those of lower quartiles. One might wonder whether this is actual evidence of a reduction in inequality or mistaking mean reversion for a reduction in inequality. Mean reversion might be expected to yield similar patterns as reduced inequality if incomes of the rich, which were rising prior to 2020, see mean reversals downward in 2020, while incomes of the poor experienced opposite changes.

To some extent our base results are robust to mean reversions. Critically, our default definition of quartiles is based on average 2015-2019 income rather than just 2019 income. This means our top quartiles persons are not *recently* top quartile person. The latter will be more subject to reversion than people who are consistently top quartile.

It is possible, however, that among those individually who on average in a quartile, especially a middle quartile, there are those who have experienced mean reversion during the 2015-2019 period. To address that concern, we compare the performance of (a) those who are in, say, the top quartile of 2019 incomes but not in the top quartile of 2015 income to (b) those in the top quartile of both 2015 and 2019 income (Figure A17).[19] The latter group tends to net out mean reversion because quartile members do not move around the distribution; the former group highlights mean reversion for the opposite reason. We

[19]For this exercise, we define quartiles cutoffs for each year so that, e.g., a person was in the top quartile in 2015 if she was in the top 25% of the income distribution in 2015.

validate this by looking at the period before 2019, when new arrivals grow faster (slower) than persistent members of the top quartile (bottom).

Consistent with mean reversion, recent arrivals to the top (bottom) quartile experience the sharpest declines (recoveries) in income during the pandemic. However, persistent members of these groups experience similar changes in relative income. The decline in regressivity is somewhat smaller when we adjust for mean reversion in this way, but the differences across transient and persistent quartile members are not stark. The pattern for consumption is the same, except that the differences between persistent quartile members and transient ones is even smaller during COVID.

An alternative approach to measuring changes to inequality after filtering out mean reversion is to compare relative income during the pandemic against a baseline that attempts to account for mean reversion. We do this by comparing actual income to a baseline that accounts for mean reversion with the periodicity of seasons, i.e., we include both trend and seasonality in our model for predicted income. We find (Figure A3) a significant decline in inequality even against this benchmark.

A limitation of all these approaches is that, with only data from January 2015 - July 2021, we cannot measure mean reversion that has a cycle longer than half our covered time period, i.e., 3.25 years. Any reversion longer than that would occur outside our sample period. That said, mean reversion with a cycle that long would not alter our conclusion that inequality fell during the pandemic because the pattern we observe is during the short 1.5 years of the pandemic.

F Mechanisms

F1 Decomposition of income inequality by source of income

Our decomposition proceeds in two steps. First, we write the percentage change in income for the average person in quartile q over time t as:

$$\frac{d\log y^q(t)}{dt} = \sum_c \frac{y_j^{c,q}(t)}{y^q(t)}\left(\frac{1}{y^{c,q}(t)}\frac{dy^{c,q}(t)}{dt}\right) = \sum_c s^{c,q}(t)\frac{d\log y^{c,q}(t)}{dt}, \tag{8}$$

where $y^{c,q}(t)$ is the average of the c-th component of income and $s^{c,q}(t)$ is the average share of income attributable to that component. Second, we subtract the 4th quartile equation from the 1st quartile equation. Finally, we organize this difference into the sum

of differences attributable to each component, which in turn we decomposed into:

$$s^{c,1}(t)\frac{d\log y^{c,1}}{dt} - s^{c,4}(t)\frac{d\log y^{c,4}}{dt} = [s^{c,1}(t) - s^{c,4}(t)]\frac{d\log y, c, 1}{dt} + s^{c,4}\left[\frac{d\log y^{c,1}}{dt} - \frac{d\log y^{c,4}}{dt}\right] \tag{9}$$

which yield the two parts – differences in component share and differences in growth rate of that component – we seek. In our calculations, we convert t from a continuous to discrete variable.

F2 Labor supply: COVID and remote work

We predict individual-level reservation wages by first estimate a Roy model with two sectors, employed and unemployed, to address selection. We assume a wage equation for employment of the form:

$$Y_{ijkt} = \beta_{jt}X_{ijkt} + \gamma_{jt}\mathbb{1}_{ik,t}(j) + e_{ijkt} \tag{10}$$

The controls (X_{ijkt}) are demographics, education, and location (k), and $\mathbb{1}_{ik,t}(j)$ is an indicator for whether the person just switched into the relevant sector. The indicator, which differs in the employed and unemployed wage equations, is our exclusion restriction. (The coefficient γ_{jt} in the employed equation captures the lower wage that just-hired individuals may earn because they lack experience in employment.) We assume the errors for the two sectors are bivariate normal. Following French and Taber (2011), we estimate a probit first stage, obtain consistent estimates for the wage equation for the employed sector, and then back out consistent estimates for the coefficients of (unobserved) wage equation for the unemployed wage sector from the combination of the two equations. We then predict each individual's reservation as the value of switching from employment status j to unemployment status j':

$$V_{i,j,j',kt} = \hat{\beta}_{j'k}X_{ij'kt} + \hat{\gamma}_{j'kt} - \hat{\beta}_{jk}X_{ijkt}. \tag{11}$$

F3 Labor supply: COVID and remote work

One might still wonder whether factors that affected the relative behavior of the rich in high-income countries—COVID risk and access to remote work—also affected those in a

lower-income country like India. Labor supply, particularly of the rich, may have been especially sensitive to disease risk to the extent that the rich have higher risk aversion, and hence greater interest in avoiding disease exposure (Oster, 2012). Moreover, greater ability to engage in remote work may have made the rich more willing and able to supply labor despite pandemic risk Adams-Prassl et al. (2020).

We investigate the plausibility of these two channels in Figure A11, which plots relative income over the course of the pandemic for households divided into exposure quantiles of confirmed COVID cases, based on district-level exposure, as well as different categories of remote work based on the Dingel and Neiman (2020) measure of remote work applied to the occupations in our sample.

While areas exposed to the highest level of COVID cases (Panel A) see lower income in the post-lockdown period, income subsequently recovers towards the end of 2020 and does not see a monotonic relationship between case exposure and income changes. We explore further the role that COVID exposure may have had on reservation wages by regressing log reservations wages on the average of daily new COVID cases at the month $\times$ district level, an indicator for above median income, and the interaction of the two (Table A8, column (3)). This regression only include data from the pandemic period, specifically, the second quarter of 2020. The reservation wages of the rich are not unusually sensitive to the pandemic period, or are they sensitive to cases in urban areas (Panel A of Table A8).

We also see scant evidence that remote work (Figure A11B) is protective of incomes in India. While workers in our highest level of remote work classification see their incomes recover more by the end of our sample; these workers still see quite large decreases in income during the lockdown period, when remote work might be expected to shelter the employment of individuals. Additionally, workers in this highest remote work classification see large pre-trends in income declining even prior to the beginning of the pandemic, and never recover the level of income they saw in the beginning of 2019. Additional caveats around this measure include the fact that we do not measure whether an individual actually worked remotely; we only impute, based on their occupational status, whether a comparable occupation in the United States would have had the potential to work remotely. It seems likely that many workers, who have latent ability to work remotely in the United States, face much stronger barriers to doing so in India due to reasons of organizational capacity, broadband access, etc. As a result, we are limited in our ability to draw strong inferences in this sample, but overall see little reason to think that remote work was protective for labor supply in India, in ways that it might have been in

other countries.

Finally, we explore the distributional consequences COVID case exposure and amenability to remote work in Figure A12. Even areas with high COVID case counts see income protected more among the rich than the poor. By contrast, we observe substantial income decreases, even among the rich, among occupations that have the highest latent ability to work remotely in our sample. Overall, we see little evidence that labor supply is changing differentially for the risk as a function of these two cross-sectional predictors.