Research Square

Preprints are preliminary reports that have not undergone peer review. They should not be considered conclusive, used to inform clinical practice, or referenced by the media as validated information.

Delayed Informalisation in India's Post-COVID Labour Market: Evidence from Pooled Repeated Cross-Sections of the Periodic Labour Force Survey, 2017-18 to 2023-24

Anupam Hazra

a.hazra.b13@gmail.com

University of Kalyani https://orcid.org/0000-0001-8183-3044

Research Article

Keywords: COVID-19, Labour Market Recovery, Informal Employment, Employment Quality, Employment Protection

Posted Date: October 16th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11373185/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: The authors declare no competing interests.

“Delayed Informalisation in India’s Post-COVID Labour Market: Evidence from Pooled Repeated Cross-Sections of the Periodic Labour Force Survey, 2017-18 to 2023-24”

Anupam Hazra, Assistant Professor, Department of Economics, University of Kalyani

Email: anupamhazra@klyuniv.ac.in / a.hazra.b13@gmail.com

Abstract:

COVID-19 disrupted India's labour market, where informal employment was already widespread. This study tests whether recovery restored pre-pandemic employment levels and quality. It also checks whether informal employment grew further after the pandemic. The study separates employment quantity from quality, protection, social security, and earnings.

It uses unit-level Periodic Labour Force Survey data for 2017-18 to 2023-24. Pooled repeated cross-sectional regressions estimate effects on employment and informal employment. Additional models test differences by gender, education, social group, and rural-urban location. The study also examines industry differences, employment benefits, and earnings. Robustness checks use logit models, placebo tests, and an event-study design. An extended post-COVID period also tests whether the results still hold. Pre-COVID state-level informal employment measures each state's exposure to the shock. This tests whether effects were larger in states with more existing informality.

The results challenge a simple shock-and-recovery interpretation of the pandemic. Aggregate informal employment did not rise significantly in 2020, though employment grew. Informal employment rose by about 1.55 percentage points during the 2021-23 recovery. Overall employment increased by a larger amount during the same period.

Worker characteristics shaped the strongest differences in informal employment during the shock. Informal employment rose more among low-educated and socially disadvantaged workers. Paid-leave eligibility also fell significantly in 2020, which weakened employment protection. The negative COVID-female interaction needs care because estimates condition on employment. Selective exit from employment may explain this pattern for women. The positive earnings coefficients do not show pandemic-related earnings gains. The earnings sample includes only those workers who report their earnings. Survivorship and composition bias may also affect the earnings estimates here.

Education and social disadvantage shaped informalisation more than initial state informality did. Recovery policies should address the quality, security, and protection of informal jobs. Low-educated and socially disadvantaged workers concentrate heavily in informal employment. Policymakers should therefore design targeted interventions for these vulnerable groups of workers. The results challenge a simple shock-and-recovery interpretation of the pandemic. The immediate 2020 period shows **no statistically significant aggregate increase in informal employment**, although measured employment increased. In contrast, informal employment increased by approximately **1.55 percentage points during the 2021–23 recovery period**, alongside a larger increase in overall employment.

Keywords: COVID-19; Labour Market Recovery; Informal Employment; Employment Quality; Employment Protection;

JEL code: E24, J21, J46, J48.

Anupam Hazra, Assistant Professor, Department of Economics, University of Kalyani

Email: anupamhazra@klyuniv.ac.in / a.hazra.b13@gmail.com

Introduction:

The COVID-19 pandemic severely disrupted India's labour market. The nationwide lockdown of March 2020 abruptly restricted economic activity and mobility. It also cut access to workplaces, employment, and household incomes. The shock mattered greatly because most Indian workers hold informal jobs. Informal workers typically have weaker protection and limited social security. They also earn irregularly and depend on daily or short-term work. The pandemic therefore shows how a large shock interacts with informality.

India's labour market was already deeply segmented before the pandemic. Formal and informal workers differ in enterprise type and job security. Informal employment extends beyond the informal sector. Workers can hold informal jobs within formal-sector enterprises. This happens when they lack written contracts, paid leave, or social security. This distinction matters for understanding the effects of COVID-19. A worker may stay employed while job security and earnings deteriorate. Employment and unemployment rates alone may therefore understate changes in work quality.

The Periodic Labour Force Survey (PLFS) offers a suitable empirical basis. Its annual rounds from 2017–18 to 2023–24 show a longer trajectory. The three pre-pandemic rounds provide a baseline for comparison. These rounds cover 2017–18, 2018–19, and 2019–20. Later rounds show the extent and nature of recovery. This coverage lets the study look beyond the immediate employment shock. It tests whether recovery restored the pre-COVID employment structure. It also tests whether work shifted towards informal and precarious forms.

This study distinguishes employment recovery from labour-market recovery. Rising employment does not necessarily restore the pre-pandemic employment structure. Displaced regular wage workers may have moved into self-employment or casual work. Workers who kept their jobs may have lost paid leave or benefits. Recovery could therefore be strong in quantity but uneven in quality. In India, informality is a persistent feature, not a temporary residual.

The pandemic may also have deepened existing labour-market inequalities. Women's employment was especially vulnerable for several reasons. Women concentrate in particular informal activities and face occupational segregation. Mobility restrictions and school closures also increased their unpaid care work. Rural and urban workers faced the shock through different mechanisms. Urban informal workers depend on construction, trade, hospitality, and transport. Mobility restrictions directly affected these sectors. Rural workers connect more strongly to agriculture and self-employment. Outcomes may therefore differ by gender, location, education, age, and social group.

Pre-existing informality may also shape the intensity of the COVID shock. This study does not treat the pandemic as identical across locations. It asks whether more informal states fared differently during and after. A state-sector's pre-pandemic informal employment indicates its underlying labour-market vulnerability. Comparing changes across such areas shows whether existing structures conditioned pandemic effects.

This paper asks whether COVID-19 interrupted India's labour-market trajectory temporarily. Or did it contribute to a deeper restructuring of employment? The study focuses

Anupam Hazra, Assistant Professor, Department of Economics, University of Kalyani

Email: anupamhazra@klyuniv.ac.in / a.hazra.b13@gmail.com

on employment disruption, informalisation, and employment quality. First, it examines changes in employment and labour-force participation during the pandemic. Second, it investigates whether employment shifted towards informal work. Third, it considers whether paid leave, social security, and earnings changed with recovery.

The paper places the shock within the longer 2017–18 to 2023–24 trajectory. It does not focus only on the pandemic year. It distinguishes informal employment from informal-sector employment. This captures changes in job security that conventional measures may obscure. The paper also tests whether effects varied by pre-existing informality. It tests variation by gender, rural–urban location, and social characteristics. This links the pandemic shock to India's long-standing labour-market segmentation.

The consequences of COVID-19 should not be judged by employment levels alone. The key question is whether workers returned to similar jobs and conditions.

Anupam Hazra, Assistant Professor, Department of Economics, University of Kalyani

Email: anupamhazra@klyuniv.ac.in / a.hazra.b13@gmail.com

2. Literature Review

The review covers four streams of literature. These are structural informality, the aggregate shock, informal workers, and gender inequality.

2.1 Informal Employment in India

India's informality is structural rather than cyclical. The literature separates the informal sector from informal employment. The first concerns enterprise traits such as registration and size. The second concerns job traits such as contracts and social security. Ghosh (2022) finds that GST barely reduced unregistered enterprises. Their share fell only from about 93% to 91%. The registration threshold excludes most micro-enterprises. Informal firms form over 96% of non-farm establishments (Enterprise Informality in India, 2023).

Abraham (2019) shows that informal jobs suppress wages even within formal enterprises. Formal-sector employment therefore does not guarantee job quality. Informality also differs across occupational groups (Heterogeneity and Participation in Informal Employment among Non-Cultivator Workers in India, 2017). Mishra (2014) documents gaps in health insurance, pensions, and unemployment protection. Post-pandemic registration drives have not closed these gaps. Sinha, Gallien, and van den Boogaard (2026) call this narrow formalisation. E-Shram increases visibility without expanding substantive protection. Abraham and Shrivastava (2022) note comparability problems across PLFS rounds. Section 5 therefore addresses these data issues explicitly.

2.2 COVID-19 and Indian Labour Markets

PLFS-based studies document the aggregate shock well. Sengupta (2023) finds sharp urban employment and earnings losses in April–June 2020. Unemployment fell below pre-pandemic levels by 2022 (Labour Market Flows and Gender Differentials in Urban Unemployment over the Pandemic, 2024). The gender gap narrowed during lockdown but widened again in recovery. No single post-pandemic round therefore proves a completed recovery.

Rural evidence centres on reverse migration and MGNREGA strain (MGNREGA in the Times of COVID-19 and Beyond, 2020). Lokhande and Gundimeda (2021) estimate 7.5 million seasonal migrants sought MGNREGA work. They recovered only 28% of pre-COVID daily earnings. Odisha evidence shows similar labour-supply surges and disguised agricultural unemployment (The COVID-19-Led Reverse Migration on Labour Supply in Rural Economy, 2021). Tiwari and Majumdar (2024) and Kandpal (2024) show state-level variation in reabsorption.

The World Bank (2020) finds informal workers bore disproportionate losses (Beaten or Broken? Informality and COVID-19). Alvarez and Pizzinelli (2021) find informality-driven recovery in Colombia. This paper tests for a similar pattern in India.

2.3 COVID-19 and Informal Workers

Several studies show informal workers suffered worse shocks and slower recovery. Kesar et al. (2021) surveyed about 5,000 respondents across 12 states. Two-thirds lost work during the lockdown. Informal workers who kept jobs saw earnings fall by over half. Nath et al. (2023) tracked 2,778 informal workers through the lockdown. Six months later, one in five remained out of work. Recovery was slower for urban workers, women, and disadvantaged groups. Dang et al. (2026) show poverty risk concentrated among informally employed urban households. Labour-market and welfare shocks therefore reinforced each other.

Anupam Hazra, Assistant Professor, Department of Economics, University of Kalyani

Email: anupamhazra@klyuniv.ac.in / a.hazra.b13@gmail.com

Occupational transition is another important channel. Ray (2024) finds income shocks pushed platform drivers into digitally organised informality. De Neve, Medappa, and Prentice (2023) argue platform work intensifies informality. Nair (2023) finds income cuts came with greater algorithmic control. Gig work therefore does not upgrade traditional informal employment. Ray and John (2025) trace new gig-worker unions to this exposure. Meher and Panda (2024) document income loss and debt among Odisha handloom weavers.

2.4 Gender and Social Inequality

Gender is the best-developed distributional strand. Abraham et al. (2022) find women lost work seven times more often. This holds conditional on pre-pandemic employment. Women were also eleven times less likely to return to work. Education and marriage affected men and women in opposite directions. Quarterly flow evidence shows the unemployment gap widening again in recovery. Gender inequality therefore persists through recovery. This motivates the gender-interaction terms in Section 8.2.

Pandey (2025) examines rising female labour-force participation in later PLFS rounds. The increase reflects economic compulsion, not better opportunities. Growth concentrates in unpaid family labour and rural own-account work. Headline indicators therefore do not show gender-equalising recovery.

Social-group and rural–urban evidence remains less developed. Older and migrant workers are more likely to hold informal jobs (Migration, Informal Employment, and the Ageing Workforce in India, 2025). Risks compound when age, migration, and informality overlap. Caste heterogeneity in COVID-period PLFS outcomes remains an open gap.

Existing studies mostly examine immediate employment effects of COVID-19. This paper treats COVID-19 as a structural shock to informal employment. It asks whether recovery reproduced, reduced, or transformed existing labour-market segmentation.

Section 3: Data Source and Coverage:

This study uses unit-level data from India's Periodic Labour Force Survey (PLFS). The National Sample Survey Office (NSSO) conducts it. NSSO sits within the Ministry of Statistics and Programme Implementation (MoSPI). The analysis uses seven consecutive annual rounds from 2017-18 to 2023-24. Three rounds are pre-COVID: 2017-18, 2018-19, and 2019-20. One round straddles the first lockdown (2019-20 Q4 and 2020-21). Three rounds cover recovery: 2021-22, 2022-23, and 2023-24.

The study needs sub-annual detail around the April–June 2020 lockdown. It therefore adds quarterly Current Weekly Status (CWS) bulletins for urban areas. These are the only official sub-annual labour statistics for that period.

The PLFS uses a rotational panel in urban areas. Enumerators visit each urban household four times over four quarters. Rural areas receive a fresh cross-section each round. The pooled data keep labour-force members aged 15 and above. This follows usual practice in PLFS-based employment research. The study drops records missing core employment, industry, or demographic variables. The remaining sample contains individual-level observations across seven rounds.

The study follows NSSO conventions. Usual Principal and Subsidiary Status (UPSS) classifies labour-force status for annual comparisons. CWS applies where quarterly resolution is needed. Workers are self-employed, regular wage employees, or casual labourers. Self-employment

Anupam Hazra, Assistant Professor, Department of Economics, University of Kalyani

Email: anupamhazra@klyuniv.ac.in / a.hazra.b13@gmail.com

includes own-account workers, employers, and unpaid family helpers.

The informal sector and informal employment differ, as Section 2.1 shows. The study therefore uses two complementary definitions.

The employment-based definition is primary. It classifies a worker as informal in three cases. First, the worker is self-employed in an unregistered enterprise. Second, the worker is a casual labourer. Third, the worker is a regular employee lacking key protections. These are a written contract, paid leave, or social-security eligibility. Social security covers PF or pension, health insurance, and gratuity. This follows the PLFS social-security module, Abraham (2019), and NCEUS conventions.

The enterprise-based definition is the alternative. It classifies workers as informal if their enterprise is unregistered or unincorporated. This holds regardless of the worker's individual contract. It follows the enterprise-informality literature (Ghosh, 2022; Enterprise Informality in India, 2023). Sections 7–8 use the employment-based definition throughout. Table 10 uses the enterprise-based definition for robustness.

COVID period. The covid indicator equals 1 from April 2020 onward. It equals 0 for earlier periods. The 2019-20 round includes the nationwide lockdown in its final quarter. Event-study specifications therefore treat that round separately. It enters neither the strict pre nor post category. This follows segmented-labour-market studies of the period.

Demographic and locational variables. Sex, age, marital status, and educational attainment come from PLFS schedules. So do social group (ST/SC/OBC/Others), religion, rural–urban sector, and state/UT. Industry follows NIC and is grouped into four categories. These are agriculture, manufacturing, construction, and trade or services.

Monthly earnings cover self-employed and salaried workers. Weekly hours worked come from the CWS. The study deflates earnings using the Consumer Price Index (Combined). This allows real comparisons across rounds.

Rural sample design. Rural PLFS samples are not panels. The study therefore cannot track rural individuals within a round. Rural transition analysis uses district or state cohorts instead.

Anupam Hazra, Assistant Professor, Department of Economics, University of Kalyani

Email: anupamhazra@klyuniv.ac.in / a.hazra.b13@gmail.com

Section 4. Methodology

The empirical strategy is organised around three questions that follow directly from the descriptive evidence in Section 3: whether the immediate 2020 shock shifted the level of employment and informality; whether pre-existing labour-market informality conditioned the size of that shift; and whether any shift proved transitory or persisted into the 2021–2023 recovery rounds. Each question is addressed with a distinct specification, reported in Tables A1–A10 of the Appendix. All models are estimated on the pooled PLFS panel (2017–18 to 2023–24; N ≈ 1.94 million for the employment equation, N ≈ 297,000 for the informality equation, which is conditional on employment). Every specification includes state, caste, and survey-year fixed effects, with standard errors clustered at the state level throughout.

4.1 Baseline Specification

The baseline model regresses the outcome of interest on a COVID-2020 period indicator, individual and household controls, and the fixed-effects structure described above:

$$Yist=\alpha+\beta 1Covidt+Xi'\gamma+\delta s+\delta t+\varepsilon ist$$

where Yist is either an employment indicator or an informality indicator (defined on the employed subsample) for individual i in state ss s in survey year t; Covidt equals 1 for the 2020 survey round and 0 otherwise; Xi includes age, age-squared, sex, low-education status, and the individual's caste category; and δs, δt are state and year fixed effects. This specification is reported in columns 1–2 of Table A1 and provides the reference point against which every subsequent specification is compared.

4.2 Heterogeneous COVID Effects by Worker Characteristics

To test whether the 2020 shock affected some groups of workers more than others, the baseline model is extended with an interaction between the COVID indicator and a worker characteristic Zi:

$$Yist=\alpha+\beta 1Covidt+\beta 2Zi+\beta 3(Covidt\times Zi)+Xi'\gamma+\delta s+\delta t+\varepsilon ist$$

with Zi taken in turn to be gender, urban residence, low-education status, and a composite disadvantaged-group indicator (columns 3–6, Table A1). The coefficient β3 on the interaction term is the object of interest throughout: it measures whether the COVID-period change in the outcome differed for workers with Zi=1 relative to Zi=0, net of any average difference between the two groups captured by β2.

Table A9 applies the same logic in two further ways: by estimating the baseline model separately on the female and male subsamples, and by replacing the binary female interaction with a three-category sex variable interacted with the COVID indicator, allowing a third gender

category to be distinguished from the male reference group where the underlying PLFS

Anupam Hazra, Assistant Professor, Department of Economics, University of Kalyani

Email: anupamhazra@klyuniv.ac.in / a.hazra.b13@gmail.com

coding permits it.

4.3 Pre-COVID Informality Exposure

The paper's exposure design tests whether the size of the COVID-period effect varied with a state's pre-existing level of informality, rather than assuming the shock was homogeneous across locations. Pre-COVID exposure is measured as the state-level informal employment rate averaged over the three pre-pandemic PLFS rounds (2017–18, 2018–19, 2019–20), and enters the model interacted with the COVID indicator:

$$Y_{ist}=\alpha+\beta_1 Exposure_s+\beta_2 Covid_t+\beta_3(Exposure_s\times Covid_t)+X_i'\gamma+\delta_s+\delta_t+\varepsilon_{ist}$$

This is estimated three ways (Table A2): with continuous exposure on the employment equation (col. 1) and the informality equation (col. 2), and with exposure recoded as a binary high-exposure indicator at the sample median (col. 3). A positive and significant β_3 would indicate that higher pre-COVID informality amplified the COVID-period effect on the outcome; the main effect β_1 is reported separately because it captures the pre-existing cross-sectional association between exposure and the outcome, which is structural rather than COVID-specific and should not be interpreted as part of the treatment effect.

4.4 Post-COVID Recovery Period

Because a single 2020 indicator cannot distinguish a transitory shock from a delayed or persistent one, the baseline and exposure specifications are re-estimated with the COVID indicator replaced by a post-COVID period indicator covering the 2021–22 through 2023–24 rounds:

$$Y_{ist}=\alpha+\beta_1 PostCovid_t+X_i'\gamma+\delta_s+\delta_t+\varepsilon_{ist}$$

applied to the employment and informality equations (Table A3, cols. 1–2) and, with the exposure interaction from Section 4.3 substituted for the plain COVID indicator, to test whether any recovery-period effect was itself concentrated in high-exposure states (Table A2, col. 4). Comparing the post-COVID coefficient against the 2020-only coefficient from Section 4.1 is the paper's principal test of whether informality effects were contemporaneous with the shock or emerged with a lag through the recovery rounds.

4.5 Employment Benefits and Earnings

Two further outcomes are examined using the baseline specification from Section 4.1: eligibility for paid leave and coverage by any social-security benefit, both conditional on employment (Table A4, cols. 1–2), with the exposure interaction from Section 4.3 additionally applied to social-security coverage (col. 3) to test whether benefit erosion during COVID was concentrated in high- or low-exposure states. Log monthly earnings, conditional on the individual reporting non-missing earnings, is examined using the baseline specification and its gender interaction (Table A5).

4.6 Sectoral Heterogeneity

Anupam Hazra, Assistant Professor, Department of Economics, University of Kalyani

Email: anupamhazra@klyuniv.ac.in / a.hazra.b13@gmail.com

The baseline informality equation is re-estimated with the COVID indicator interacted with a set of broad industry-group dummies (agriculture as the omitted reference category), testing whether the COVID-period change in informality differed across manufacturing, construction, trade, transport/storage, accommodation and food services, and other services (Table A6).

4.7 Event-Study Specification

To examine the dynamic path of the exposure relationship across the full panel rather than collapsing the post-COVID period into a single indicator, an event-study specification interacts pre-COVID exposure with a full set of survey-year dummies, using 2019 — the last complete pre-COVID round — as the reference year:

$$Y_{ist}=\alpha+\sum_{\tau\neq 2019}\beta_\tau(\text{Year}_\tau\times\text{Exposures})+\sum_{\tau\neq 2019}\gamma_\tau\text{Year}_\tau+X_i'\gamma+\delta_s+\varepsilon_{ist}$$

The pre-2020 coefficients ($\tau\in\{2017,2018\}$) serve as a test of differential pre-trends between high- and low-exposure states; the 2020–2022 coefficients trace whether any exposure-related divergence emerged during the shock itself or only in the subsequent rounds.

4.8 Robustness and Placebo Tests

Four checks assess the sensitivity of the baseline and exposure results. First, the main informality specification is re-estimated by logit rather than linear probability, both without and with the gender interaction, to confirm that results are not an artefact of the functional form (Table A7). Second, the COVID window is extended beyond the single 2020 round to test sensitivity to the exact period definition (Table A8, col. 1). Third, a placebo test re-estimates the model using a fabricated pre-period "shock" indicator in place of the true COVID indicator, both on its own and interacted with pre-COVID exposure (Table A8, cols. 2–3); an insignificant placebo main effect supports the main specification, while a significant placebo–exposure interaction is treated as an important qualification on the exposure design rather than set aside, and is discussed explicitly in Section 6.

4.9 Estimation

All models are estimated using ordinary least squares (linear probability specification for binary outcomes) with the exception of the logit robustness checks in Section 4.8, and cluster standard errors at the state level throughout to account for within-state correlation across individuals and rounds. Coefficients on interaction terms are the primary quantities of interest; average marginal effects are not separately reported, since all interaction terms in the linear probability specifications are directly interpretable as percentage-point effects.

Anupam Hazra, Assistant Professor, Department of Economics, University of Kalyani

Email: anupamhazra@klyuniv.ac.in / a.hazra.b13@gmail.com

Section 5: Descriptive : Informal Employment Before, During, and After COVID-19

This section presents descriptive evidence on informal employment from 2017–18 to 2023–24. It examines the aggregate trend, then differences by industry, gender, and location. It then covers social group, age, earnings, and distribution. The section ends with the exposure design and heterogeneous effects. These patterns motivate the regression specifications that follow.

5.1 The Aggregate Trajectory, 2017–18 to 2023–24

Figures 1 and 13 trace informal employment and overall employment across PLFS rounds. The round 2020–21 marks the principal COVID-19 shock. Informal employment stayed dominant throughout the period. The pandemic therefore hit a labour market with weak protection already. The shock interacted with existing segmentation rather than creating it.

The series supports a three-period reading. Pre-COVID rounds establish the baseline employment structure. The 2020–21 round captures the immediate disruption. Later rounds show whether the labour market reverted to its earlier form. A numerical employment recovery does not prove a complete recovery. Workers may re-enter through jobs with lower earnings and weaker security. Those jobs may offer negligible social-security coverage. This distinction motivates the paper's central question. Did COVID-19 cause a temporary contraction or a persistent informal shift?

5.2 Sectoral and Industry Differences

Figure 3 and the industry trend graph show informal shares by industry. Informality is not economy-wide. It depends on how each sector organises production. Agriculture and construction are structurally prone to high informality. Agriculture relies on self-employment, unpaid family labour, and casual wage work. Construction relies on subcontracting, high labour mobility, and short contracts. Both sectors face sudden interruptions with little formal income protection. Manufacturing and services are more heterogeneous. Formal enterprises coexist there with household-based and subcontracted production. Protection therefore depends on firm traits, not the sector label. The pandemic's sectoral incidence depended on two factors. The first was whether activity could continue under mobility restrictions. The second was reliance on face-to-face transactions.

The post-COVID comparison matters more. Elevated informal shares relative to baseline suggest a segmentation pattern. Displaced workers may re-enter through low-barrier informal activity. They may not return to their prior tier. Industry trajectories show this pattern persists across the full panel. It is not confined to the shock round. This justifies the sectoral controls in the regression analysis.

5.3 Gender Differences

Figure 4 and the gender trend graph compare informal employment by sex. Women concentrate in activities with weaker contractual protection. These activities also offer lower earnings and limited social security. This reflects unequal labour-market positions, not occupational preference. Two channels plausibly widened this gap during the pandemic. On the demand side, women's jobs were sensitive to mobility restrictions. On the supply side, rising care duties may have limited paid re-entry. This could occur independently of labour demand. The two channels look similar in aggregate female employment shares. They carry different welfare implications. Descriptive comparison alone cannot separate them.

Read the figure on two dimensions. The first is the female informal share, which indexes baseline vulnerability. The second is any widening of the gap around the pandemic. A

Anupam Hazra, Assistant Professor, Department of Economics, University of Kalyani

Email: anupamhazra@klyuniv.ac.in / a.hazra.b13@gmail.com

widening gap indicates an intensified, not symmetric, shock. The gap persists across the full panel. This motivates the interaction specification used later. It interacts a COVID indicator with gender. This isolates a differential pandemic effect from the structural gap.

5.4 Rural–Urban Differences

Figure 5 and the rural–urban trend graph disaggregate informality by location. Rural labour markets tie closely to agriculture, casual wages, and small self-employment. Urban informal work concentrates in construction, petty trade, transport, and personal services. These activities depend on daily mobility and customer contact. The pandemic therefore transmitted differently across locations. Production disruption mainly affected rural markets. Urban informal workers faced tighter constraints from mobility and in-person demand.

This distinction affects how we read the national aggregate. Different processes can produce the same rise in informality. Rural informalisation may reflect the structural dominance of informal agriculture and self-employment. Urban informalisation more plausibly reflects displacement from formal or semi-formal jobs. Displaced workers then enter precarious activity. The rural–urban gap persists across all seven rounds. This confirms informality is structural, not pandemic-specific. Location therefore marks heterogeneous exposure, not merely a control variable.

5.5 Social-Group and Age Differences

Figure 6 reports informal employment by social group. The distribution tracks broader patterns of social disadvantage. Disadvantaged groups concentrate in occupations with weaker protection and bargaining power. Informality thus reproduces existing socioeconomic inequality. It is more than an employment-status outcome. An asset-buffering mechanism plausibly intensified this pattern during the pandemic. Workers with fewer assets could smooth income loss less. They were more likely to accept informal re-employment for immediate income. A persistent or widening gap would show recovery left inequality intact.

Figure 7 reports the age profile. The profile reflects workers' position in the employment life cycle. Younger workers have weaker attachment to established employment relationships. Older workers increasingly concentrate in self-employment, casual work, or low-productivity activity. Interrupted recruitment created a scarring risk for younger entrants. Delayed entry into stable jobs raises the chance of informal career paths. The regressions control for social group and age directly. This tests whether the COVID-period relationship survives individual conditioning.

5.6 Earnings and Distribution

Figure 8 compares monthly earnings of formal and informal workers. It uses nominal and constant-price terms. Figure 10 shows real informal earnings for 2019, 2020, and 2022. These years approximate the pre-COVID, shock, and recovery phases. This evidence complements the employment-share results. More employed workers do not mean complete recovery if pay falls. A stable mean can also hide movement in the tails.

The persistent formal–informal earnings gap supports the segmentation interpretation. Real earnings are more informative than nominal earnings. Nominal growth may simply reflect price changes. A leftward shift in the distribution would signal intensive-margin effects. Workers who stayed employed still suffered material income losses. The central test is whether the 2022 distribution reverts toward 2019. Reversion would show livelihood recovery alongside employment recovery. Persistent depression would show workers returned without regaining prior income.

Anupam Hazra, Assistant Professor, Department of Economics, University of Kalyani

Email: anupamhazra@klyuniv.ac.in / a.hazra.b13@gmail.com

5.7 Identification Strategy: Exposure Design and Event Study

The analysis now moves from description to an exposure-based identification strategy. The key variable is pre-COVID informality exposure. It is the average informal employment rate over three pre-pandemic rounds. It is defined at the state–sector level. States and sectors entered the pandemic with different labour-market structures. A common shock could therefore produce heterogeneous consequences.

The specification interacts a COVID indicator with exposure. A positive interaction in the informality equation signals larger informality increases. Such increases would occur in more informal labour markets. A negative interaction in the employment equation signals larger contraction. This design beats a simple before–after comparison of 2019–20 and 2020–21. It exploits cross-sectional variation in exposure intensity. A common time shock cannot identify heterogeneous effects by construction.

Figure 11 reports event-study estimates with 2019–20 as the reference period. This framework tests whether exposure groups were already diverging before the pandemic. Small, insignificant lead coefficients support attributing divergence to the pandemic. A sharp change at the COVID coefficient would strengthen this. Post-COVID coefficients reverting toward pre-pandemic levels indicate a transitory shock. Persistent coefficients indicate a durable restructuring along the exposure gradient. The framework thus recasts the question in dynamic terms. It asks whether COVID-19 left a lasting imprint on India's labour market.

5.8 Heterogeneous Treatment Effects

Figure 12 reports heterogeneous effects with confidence intervals. The subgroups cover agriculture, non-agriculture, gender, and education. An average effect can hide variation in exposure and adjustment capacity. The agriculture split captures whether activity could continue under mobility restrictions. The gender split captures labour-force attachment and care-responsibility channels (Section 4.3). The education split captures access to protective or flexible occupations. Together, these results reinforce the paper's central interpretive claim. COVID-19 did not shift the employment distribution uniformly. Its effects varied with workers' pre-existing economic and social position.

6. Results Analysis

6.1 Overview of the Estimated Specifications

Tables A1–A10 report ten families of specifications. They use the pooled PLFS panel from 2017–18 to 2023–24. The employment equation has about 1.9 million observations. The conditional informality equations have about 297,000. All models include state, caste, and survey-year fixed effects. Standard errors are clustered at the state level.

The tables test four claims from the descriptive section. First, COVID-19 shifted employment and informality levels. Second, the shift was larger where informality was already high. Third, the shift persisted into the 2021–2023 recovery rounds. Fourth, it differed by gender, education, sector, and industry. The results support some claims more strongly than others. Null results are as informative as significant coefficients.

6.2 The Immediate 2020 Shock: A Muted and Counter-Intuitive Signal

Table A1 does not show a simple intensification of informality. The employment equation shows a positive, highly significant 2020 coefficient (0.0442, $p<0.001$). Usual-status

Anupam Hazra, Assistant Professor, Department of Economics, University of Kalyani

Email: anupamhazra@klyuniv.ac.in / a.hazra.b13@gmail.com

employment was higher, not lower, in 2020. This holds net of state, caste, and year effects. PLFS usual status may absorb displaced workers into self-employment and agriculture. It does not register them as unemployed. The employment rate is therefore an unclear indicator of distress. Column 2 shows no significant average informality shift in 2020 (−0.0014, p=0.803). Aggregate informality did not detectably rise in the shock year.

The interaction terms in columns 3–6 are more informative. COVID × education is positive and significant (0.0192, p=0.035). COVID × disadvantaged group is also positive and significant (0.0097, p=0.043). Low-education and disadvantaged workers saw larger informality increases. The average worker saw no such increase.

COVID × Female is negative and significant (−0.0124, p=0.032). Women's informality fell relative to men's in 2020. This contradicts the intensification hypothesis in Section 5.3. COVID × Urban is insignificant (−0.0100, p=0.216). The logit estimate in Table A7 reproduces the female interaction. It has the same sign and greater precision (−0.4387, p<0.001). The linear probability specification therefore does not drive this result.

Education and social group concentrated the 2020 informality effects. Gender and location did not. The gender result runs against the paper's working hypothesis. The paper should discuss it as such.

6.3 Pre-COVID Exposure: No Detectable Dose-Response Relationship

Table A2 tests the exposure design from Section 5.7. It asks whether COVID effects were larger in high-informality areas. No exposure–COVID interaction is statistically significant. Continuous exposure × COVID is insignificant for employment (−0.0447, p=0.159). It is also insignificant for informality (−0.0309, p=0.341). Binary high-exposure × COVID is insignificant too (−0.0062, p=0.354).

Table A10 corroborates these results. Every year × exposure coefficient from 2017 to 2022 is insignificant. Their p-values range from 0.11 to 0.72. No divergence appears around 2020. This differs from what the descriptive trends in Section 5 anticipated. The paper should report it as a genuine null.

Exposure main effects are informative on a separate margin. Pre-COVID exposure predicts lower baseline employment (−0.2908, p<0.001). It also predicts lower conditional informality (−0.0925, p=0.064). This cross-sectional association pre-dates the pandemic. It should not be confused with a COVID-period effect.

6.4 Post-COVID Recovery: A Delayed Rather Than Immediate Effect

Table A3 replaces the 2020 indicator with a post-COVID dummy. The dummy covers 2021–2023. Informality rose by 1.55 percentage points in recovery rounds (0.0155, p=0.036). The 2020-only specification missed this. Employment rose by more (0.0671, p<0.001).

The null 2020 result in Section 6.2 complements this finding. Informalisation was not a contemporaneous 2020 shock. It was a cumulative drift visible only in pooled recovery rounds. This fits the distinction between quantitative and qualitative recovery. The timing, however, is later than a pure shock reading implies.

Anupam Hazra, Assistant Professor, Department of Economics, University of Kalyani

Email: anupamhazra@klyuniv.ac.in / a.hazra.b13@gmail.com

The post-COVID × exposure interaction remains insignificant (0.0228, p=0.355). This reinforces the conclusion in Section 6.3. The informality rise did not concentrate in high-exposure areas.

6.5 Employment Benefits and Earnings: Diverging Evidence

Table A4 shows paid-leave eligibility fell significantly in 2020 (−0.0298, p=0.007). Social-security coverage shows no average change (−0.0382, p=0.454). Column 3 reveals a more specific pattern. Coverage fell sharply in low-exposure areas (−0.2799, p=0.02). High-exposure areas offset this decline significantly (+0.7256, p=0.002). Coverage was already low in high-informality states. This fits a floor effect. Little coverage remained to erode. It does not show that high-exposure workers were protected.

Table A5 needs a caveat before substantive interpretation. The COVID main effect is positive (0.1158, p<0.001). Its interaction with gender is also positive (0.0691, p=0.004). These imply earnings rose, especially for women.

The sample includes only workers reporting non-missing earnings. Section 6.2 shows workers were not uniformly displaced from employment. Survivorship or composition bias better explains the pattern. Lower-paid workers may have exited the reporting sample disproportionately. The table does not show that COVID-19 raised earnings. A Heckman-style correction or bounds analysis is needed first. This is a direction for further robustness work. It is not a finding.

6.6 Sectoral Heterogeneity

Table A6 uses agriculture as the omitted category. Baseline informality is significantly lower in manufacturing (−0.0187, p=0.005). It is significantly higher in trade (0.0081, p=0.032). It is also higher in accommodation and food services (0.0117, p=0.005).

COVID interactions show significant informality declines in several sectors. Trade shows a decline (−0.0191, p<0.001). Transport and storage also declined (−0.0161, p=0.01). Other services declined most (−0.0334, p=0.02). These results contradict an intensification story. Closure of the smallest informal establishments likely explains them. Contact-intensive sectors saw such firms leave the employed sample. They did not become formal. Manufacturing, construction, and accommodation show no significant COVID interaction.

6.7 Robustness and a Caveat on the Identification Strategy

Table A8 runs a placebo test using a fabricated pre-period shock. The main effect is insignificant (0.0033, p=0.562). This supports the main specification. The placebo × exposure interaction is significant (−0.0787, p=0.01). This should concern rather than reassure. The exposure measure correlates with something systematic outside COVID. It does so in the placebo period too. Combined with the null results in Section 6.3, this raises a concern. Exposure may capture a persistent structural trait of high-informality states. It may not capture a COVID-specific channel. The paper's discussion should state this limitation explicitly. Readers should not have to find it in the appendix.

6.8 Summary of Findings

Anupam Hazra, Assistant Professor, Department of Economics, University of Kalyani

Email: anupamhazra@klyuniv.ac.in / a.hazra.b13@gmail.com

Informality did not rise detectably in the 2020 shock. It rose significantly but modestly across 2021–2023. The rise shows no relationship with pre-COVID exposure. This holds at both cross-sectional and event-study levels. It contradicts the paper's exposure hypothesis.

The clearest heterogeneity runs by education and social disadvantage. Gender and urban–rural location show less. The gender interaction runs opposite to the working hypothesis.

7. Discussion

The descriptive evidence in Section 4 suggested an uneven shock. It proposed that high-informality states absorbed more of the shock. The regressions in Section 5 do not support this strong claim. No COVID × exposure interaction reaches conventional significance. This holds for the continuous, binary, and event-study specifications.

The data support a narrower claim. Informality rose modestly but significantly across the whole labour market. Education and social group concentrated the rise, not geography. The rise appears only in the pooled 2021–2023 rounds. It does not appear in 2020 itself. This discussion builds on the revised claim. Presenting a null interaction as a confirmed mechanism would misstate the results.

The null result is a finding in its own right. A precisely estimated null is informative. Whatever raised informality acted fairly uniformly across states. The shock did not intensify where informality was already high. This points away from a state-level vulnerability channel. It suggests a broad demand or labour-supply adjustment instead. Such an adjustment affected the national labour market roughly proportionately.

7.1 Why Was the Effect Delayed?

The null 2020 result contrasts with the significant 2021–2023 result (Section 5.4). This gap is the paper's main interpretive puzzle. Three explanations fit the data. They carry different implications for the paper's contribution.

Composition is the first explanation. The 2020 shock may have removed marginal informal activity entirely. Negative COVID interactions for trade, transport, and services support this (Table A6). The employed 2020 sample then looks similar to the baseline. Informalisation appears later as displaced workers re-enter through insecure channels.

Scarring is the second explanation. Workers who lost formal or semi-formal jobs may search longer. They may exhaust formal options before accepting informal work. This produces a lagged rise in informality.

Policy withdrawal is the third explanation. Temporary income support and employment guarantees may have muted the initial shock. The effect re-emerged once authorities scaled back these measures.

The current specifications cannot separate these three channels. This requires worker-level transition data. Such data track moves between formal, informal, and non-employment. PLFS repeated cross-sections do not provide these transitions. A panel or matched-worker approach could. This is a direction for extension.

Anupam Hazra, Assistant Professor, Department of Economics, University of Kalyani

Email: anupamhazra@klyuniv.ac.in / a.hazra.b13@gmail.com

7.2 The Gender Result

The COVID × Female coefficient is negative and significant. It holds in the logit specification (Table A7). It contradicts the vulnerability argument in Section 4.3. The data cannot fully distinguish two readings.

The first reading is a ceiling effect. Women's pre-pandemic employment already had a high informal share. They had little room to become more informal in 2020. Men started from a more formal baseline. Men therefore had more room to move. The social-security results show an analogous floor effect (Section 5.5).

The second reading is compositional. Women may have exited employment more often in 2020. The care-burden channel in Section 4.3 supports this. The remaining employed women are then a selected group. They are likely more securely attached. This bias pushes the female coefficient downward mechanically. It does so whatever happened to women's aggregate prospects.

Separating the two readings requires gender-specific extensive-margin results. The current specification does not report these. The question therefore remains open for revision. The result does not refute the gendered-vulnerability hypothesis. It only shows the 2020 conditional-informality margin does not confirm it. A different empirical test is needed.

7.3 The Placebo Result

Section 5.7 reports a specific concern. The placebo main effect is insignificant, as expected. Its interaction with pre-COVID exposure is significant. Section 5.3 shows null COVID × exposure interactions. Together, these results raise a concern about the exposure measure. The three-round average informality rate may capture a persistent state trait. It may not capture a pandemic-specific channel.

The response is to test the measure more directly, not discard it. A state-by-linear-trend specification would help. A difference-in-differences design with state-specific trends would also help. These designs separate a COVID-period divergence from a pre-existing gradient. This check should precede any claim built on the exposure interaction. The paper should state its absence directly.

7.4 Policy Implications

Two results are robust enough to support policy discussion. First, low-education and socially disadvantaged workers bore the informality increase (Section 5.2). This holds across linear and logit specifications. Geography-based relief therefore likely missed the most affected workers. Targeting by education and social group would aim more precisely. This applies to future shocks as well.

Second, informality rose during recovery, not during the shock itself (Section 5.4). Policy calibrated to the acute crisis may have missed the key period. The labour market was then reorganising towards less secure work. Support for return to formal or semi-formal jobs needed a longer horizon. Initial relief packages assumed a shorter one.

Anupam Hazra, Assistant Professor, Department of Economics, University of Kalyani

Email: anupamhazra@klyuniv.ac.in / a.hazra.b13@gmail.com

Section 8. Conclusion

This paper asked whether COVID-19 caused a temporary contraction or lasting change. Labour-market segmentation explains the evidence better than a shock-and-recovery view. This reading sharpens the paper's contribution.

Measured employment rose in 2020, while conditional informality did not shift. This is not a data anomaly. It fits the informal sector's classical role in a dual economy. The informal sector absorbs labour displaced from more secure activity. Headline employment therefore stays stable while its composition deteriorates. Informality thus acted as a buffer, or employer of last resort. It was not the margin where the shock first appeared. This changes how we should read employment recovery in informal economies. A stable employment rate does not disprove labour-market distress. Informal absorption can mask deteriorating job quality. Unemployment-rate accounts of the pandemic likely understate its true labour-market cost.

The informality rise appeared only in 2021–2023, after the 2020 shock. This delay fits scarring or hysteresis mechanisms from job-ladder and search models. Displaced formal workers do not re-enter at an equivalent rung. They search for a period before accepting informal jobs. Labour-market damage therefore surfaces with a lag. The social-security results show a floor effect. Coverage was already minimal in high-exposure states. It had little room to fall further. Together, these results point to a different adjustment margin. Aggregate shocks change employment quality more than employment status. A paper using only employment or unemployment rates would miss this.

Informalisation concentrated among low-education and socially disadvantaged workers. It did not concentrate by state exposure or urban–rural location. This is the paper's most theoretically consequential finding. It weighs against a purely structuralist account of informality. That account sees informality absorbing labour that formal jobs fail to employ. The evidence favours a segmentation account instead. Bargaining power, human capital, and social position govern entry into informal work. Classic dual labour market theory predicts this pattern. Workers with the fewest outside options enter the least secure segment. This happens when aggregate demand contracts, wherever they live. The null geographic-exposure result is informative in its own right. Informality risk travels with workers' characteristics, not state informality intensity. Future work should measure vulnerability at the worker level.

The gender result remains unresolved. The negative COVID-female interaction has two possible explanations. First, a ceiling effect may operate. Women's high baseline informality left less room to rise further. Second, a composition effect may operate through the added-worker or discouraged-worker margin. Household income-smoothing decisions govern women's labour-force attachment during shocks. Different mechanisms drive men's employment adjustment. Separating these explanations requires modelling the extensive margin directly. This is the clearest next step for this framework.

India's informal sector is a primary shock absorber, not a peripheral residual. Its capacity makes employment recovery an unreliable guide to labour-market health. Its burden falls along lines of human capital and social identity. Location of work matters far less.

Anupam Hazra, Assistant Professor, Department of Economics, University of Kalyani

Email: anupamhazra@klyuniv.ac.in / a.hazra.b13@gmail.com

Reference:

i. Abraham, R. (2019). Informal employment and the structure of wages in India: A review of trends. *Review of Income and Wealth*, *65*(S1), S102–S122. https://doi.org/10.1111/roiw.12435
ii. Abraham, R., Basole, A., & Kesar, S. (2022). Down and out? The gendered impact of the Covid-19 pandemic on India's labour market. *Economia Politica*, *39*(1), 101–128. https://doi.org/10.1007/s40888-021-00234-8
iii. Abraham, R., & Shrivastava, A. (2022). How comparable are India's labor market surveys? *The Indian Journal of Labour Economics*, *65*, 321–346. https://doi.org/10.1007/s41027-022-00381-x
iv. Alvarez, J. A., & Pizzinelli, C. (2021). *COVID-19 and the informality-driven recovery: The case of Colombia's labor market* (IMF Working Paper No. 2021/235). International Monetary Fund. https://doi.org/10.5089/9781513597805.001
v. Behera,M, Mishra, S and Behera A (2021). *The COVID-19-led reverse migration on labour supply in rural economy: Challenges, opportunities and road ahead in Odisha*. (2021). *The Indian Economic Journal*, *69*(3), 392–409. https://doi.org/10.1177/00194662211013216
vi. Dang, A., et al. (2026). COVID-19 and the unequal distribution of poverty risks: Evidence from urban India. *Review of Development Economics*. https://doi.org/10.1111/rode.13268
vii. De Neve, G., Medappa, K., & Prentice, R. (2023). India's gig economy workers at the time of Covid-19: An introduction. *Sociological Bulletin*. https://doi.org/10.1177/09731741231182759
viii. Santosh Mehrotra Tuhinsubhra Giri (2023). Enterprise informality in India: The blind spots in public policy. (2023). *The Indian Journal of Labour Economics*. https://doi.org/10.1007/s41027-023-00450-9
ix. Ghosh, S. (2022). Formalising the informal through GST: Evidence from a survey of MSMEs. *Journal of Entrepreneurship*. https://doi.org/10.1177/09722661221130138
x. Bimal Kishore Sahoo &Bhaskar Jyoti Neog (2017). Heterogeneity and participation in informal employment among non-cultivator workers in India. (2017). *Applied Economics*. https://doi.org/10.1080/02692171.2016.1257584
xi. Kandpal, P. C. (2024). Responding to Covid-19 pandemic: Policy and governance perspective to reverse migration in Uttarakhand. https://doi.org/10.1177/00195561241248300
xii. Kesar, S., Abraham, R., Lahoti, R., Nath, P., & Basole, A. (2021). Pandemic, informality, and vulnerability: Impact of COVID-19 on livelihoods in India. *Canadian Journal of Development Studies*, *42*(1–2), 145–164. https://doi.org/10.1080/02255189.2021.1890003
xiii. Paaritosh Nath & Rahul Menon (2024). Labour market flows and gender differentials in urban unemployment over the pandemic. (2024). *The Indian Journal of Labour Economics*. https://doi.org/10.1007/s41027-024-00481-w
xiv. Lokhande, N., & Gundimeda, H. (2021). MGNREGA: The guaranteed refuge for returning migrants during COVID-19 lockdown

Anupam Hazra, Assistant Professor, Department of Economics, University of Kalyani

Email: anupamhazra@klyuniv.ac.in / a.hazra.b13@gmail.com

in India. *The Indian Economic Journal*. https://doi.org/10.1177/00194662211023848

xv. Meher, A. K., & Panda, P. K. (2024). The employment and livelihood challenge among the informal handloom weavers during COVID-19 and lockdown in Odisha, India. https://doi.org/10.1177/0258042X241238819

Anupam Hazra, Assistant Professor, Department of Economics, University of Kalyani

Email: anupamhazra@klyuniv.ac.in / a.hazra.b13@gmail.com

xvi. Gayathri Vasudevan, Shanu Singh, Gaurav Gupta & C. K. Jalajakshi (2020). MGNREGA in the Times of COVID-19 and Beyond: Can India do More with Less?. *The Indian Journal of Labour Economics*. https://doi.org/10.1007/s41027-020-00247-0
xvii. Akash Dandapat, Soumita Dasgupta & Pinaki Das(2025) .Migration, informal employment, and the ageing workforce in India: An exploration of PLFS unit-level data. (2025). *Journal of Social and Economic Development*. https://doi.org/10.1007/s40847-025-00482-5
xviii. Mishra, A. K. (2014). Safety net measures for unorganised workers in India: Critical gaps and challenges. *Social Change*, *44*(2), 179–203. https://doi.org/10.1177/0049085714526278
xix. Nair, G. (2023). The political economy of gig work in the pandemic: Social hierarchies and labour control of Indian platform workers. https://doi.org/10.1177/09731741231202185
xx. Nath, P., et al. (2023). Loss, recovery and the long road ahead: Tracking India's informal workers through the pandemic. *The Indian Journal of Labour Economics*, *66*(1), 329–354. https://doi.org/10.1007/s41027-022-00418-1
xxi. Pandey, A. (2025). Female labour force participation in India: The post-2017 turnaround. *International Journal of Economics and Management Studies*, *12*(9), 80–
89. https://doi.org/10.14445/23939125/IJEMS-V12I9P108
xxii. Ray, A. (2024). Coping with crisis and precarity in the gig economy: 'Digitally organised informality', migration and socio-spatial networks among platform drivers in India. https://doi.org/10.1177/0308518x231220296
xxiii. Ray, A., & John, A. (2025). A new politics of welfare? The origins and strategies of India's gig and platform workers' unions in the era of digital capitalism. https://doi.org/10.1177/10245294251339391
xxiv. Sengupta, A. (2023). Effect of COVID-19 pandemic on employment and earning in urban India during the first three months of pandemic period: An analysis with unit-level data of Periodic Labour Force Survey. *The Indian Journal of Labour Economics*, *66*(1), 283–298. https://doi.org/10.1007/s41027-023-00428-7
xxv. Sinha, H., Gallien, M., & van den Boogaard, V. (2026). Narrow formalisation: Informal workers, social protection and digital registration in India. *Competition & Change*. https://doi.org/10.1177/14680181261416392
xxvi. Tiwari, P., & Majumdar, S. (2024). Examining the impact of socio-economic, demographic and policy factors on reverse migration in India during COVID-19 pandemic. https://doi.org/10.1177/09754253241278767
xxvii. World Bank. (2020). *Beaten or broken? Informality and COVID-19* (South Asia Economic Focus, October 2020). https://doi.org/10.1596/978-1-4648-1640-6

Anupam Hazra, Assistant Professor, Department of Economics, University of Kalyani

Email: anupamhazra@klyuniv.ac.in / a.hazra.b13@gmail.com

Appendix

Composite Regression Tables

This appendix reports the full set of regression models estimated for this paper. Significance: *** p<0.01, ** p<0.05, * p<0.10. State, caste, and survey-year fixed effects are included in every specification unless noted otherwise; their coefficients are omitted from the table for space and are available on request.

Table A1. COVID-19 and Informal Employment in India

Variable	Employment	Informality	Gender	Rural-Urban	Education	Social Group
COVID-19 (2020)	0.0442***	-0.0014	0.0009	0.0026	-0.0133	-0.0046
	(0.0042)	(0.0057)	(0.0052)	(0.0053)	(0.0102)	(0.0063)
COVID x Female			-0.0124**			
			(0.0056)			
COVID x Urban				-0.0100		
				(0.0079)		
COVID x Low Education					0.0192**	
					(0.0088)	
COVID x Disadvantaged						0.0097**
						(0.0046)
Female	-0.4890***	0.0007	0.0023	0.0008	0.0008	0.0008
	(0.0245)	(0.0019)	(0.0019)	(0.0020)	(0.0019)	(0.0019)
Urban	-0.0502***	-0.0127***	-0.0127***	-0.0112***	-0.0128***	-0.0138***
	(0.0080)	(0.0036)	(0.0036)	(0.0026)	(0.0037)	(0.0041)
Age	0.0031***	-0.0003***	-0.0003***	-0.0003***	-0.0003***	-0.0004***
	(0.0005)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Age 25-34	0.1741***	-0.0029	-0.0029	-0.0029	-0.0029	-0.0031
	(0.0059)	(0.0018)	(0.0018)	(0.0018)	(0.0018)	(0.0018)
Low Education	0.0957***	0.0442***	0.0442***	0.0442***	0.0413***	0.0450***
	(0.0089)	(0.0074)	(0.0074)	(0.0074)	(0.0066)	(0.0076)
State, Caste & Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,936,240	297,367	297,367	297,367	297,367	297,367
R-squared	0.2806	0.0468	0.0470	0.0470	0.0473	0.0457

Dependent variables: Employment (col. 1), informal employment conditional on being employed (cols. 2-6). Col. 2 is the baseline informality specification; cols. 3-6 add interactions of COVID-19 with gender, urban residence, low education, and a composite disadvantaged-group indicator respectively.

Anupam Hazra, Assistant Professor, Department of Economics, University of Kalyani

Email: anupamhazra@klyuniv.ac.in / a.hazra.b13@gmail.com

Table A2. Pre-COVID Informality Exposure and COVID-19

Variable	Employment	Informal Employment	High Exposure	Recovery
Pre-COVID Exposure	-0.2908***	-0.0925*		-0.1112*
	(0.0660)	(0.0484)		(0.0551)
COVID-19 (2020)	0.0572***	0.0086	0.0019	
	(0.0095)	(0.0091)	(0.0058)	
Post-COVID (2021-23)				0.0079
				(0.0100)
High-Exposure Region			-0.0030	
			(0.0064)	
Exposure x COVID	-0.0447	-0.0309		
	(0.0310)	(0.0320)		
High Exposure x COVID			-0.0062	
			(0.0066)	
Exposure x Post-COVID				0.0228
				(0.0243)
Female	-0.4889***	0.0005	0.0008	0.0006
	(0.0245)	(0.0018)	(0.0019)	(0.0018)
Urban	-0.0025	0.0027	-0.0098***	0.0029
	(0.0148)	(0.0064)	(0.0029)	(0.0063)
Age	0.0030***	-0.0003***	-0.0003***	-0.0003***
	(0.0005)	(0.0001)	(0.0001)	(0.0001)
Age 25-34	0.1743***	-0.0029	-0.0029	-0.0029
	(0.0059)	(0.0018)	(0.0018)	(0.0018)
Low Education	0.0954***	0.0442***	0.0442***	0.0443***
	(0.0089)	(0.0074)	(0.0074)	(0.0074)
State, Caste & Year FE	Yes	Yes	Yes	Yes
Observations	1,936,240	297,367	297,367	297,367
R-squared	0.2810	0.0475	0.0469	0.0475

Col. 1-2 interact continuous pre-COVID exposure with the COVID-2020 indicator; col. 3 uses a binary high-exposure indicator; col. 4 uses the post-COVID period (2021-2023) indicator in place of the single COVID-2020 shock year, capturing recovery-period effects.

Anupam Hazra, Assistant Professor, Department of Economics, University of Kalyani

Email: anupamhazra@klyuniv.ac.in / a.hazra.b13@gmail.com

Table A3. Post-COVID Recovery: Employment and Informality

Variable	Employment	Informality
Post-COVID (2021-23)	0.0671***	0.0155**
	(0.0079)	(0.0071)
Female	-0.4890***	0.0007
	(0.0245)	(0.0019)
Urban	-0.0502***	-0.0127***
	(0.0080)	(0.0036)
Age	0.0031***	-0.0003***
	(0.0005)	(0.0001)
Age 25-34	0.1741***	-0.0029
	(0.0059)	(0.0018)
Low Education	0.0957***	0.0442***
	(0.0089)	(0.0074)
State, Caste & Year FE	Yes	Yes
Observations	1,936,240	297,367
R-squared	0.2806	0.0468

Uses the post-COVID period indicator (2021-2023) in place of the single 2020 shock year, testing whether effects persisted into the recovery rounds rather than being confined to the initial lockdown year. Compare with cols. 1-2 of Table A1.

Anupam Hazra, Assistant Professor, Department of Economics, University of Kalyani

Email: anupamhazra@klyuniv.ac.in / a.hazra.b13@gmail.com

Table A4. Employment Benefits: Paid Leave and Social Security Coverage

Variable	Paid Leave	Social Security	SS x Exposure
COVID-19 (2020)	-0.0298***	-0.0382	-0.2799**
	(0.0105)	(0.0505)	(0.1149)
Pre-COVID Exposure			-1.0055*
			(0.5633)
Exposure x COVID			0.7256***
			(0.2177)
Female	-0.0587***	-0.0437	-0.0472
	(0.0134)	(0.0364)	(0.0366)
Urban	-0.1279***	-0.3316***	-0.1886***
	(0.0135)	(0.0405)	(0.0629)
Age	-0.0038***	-0.0083***	-0.0083***
	(0.0004)	(0.0012)	(0.0013)
Age 25-34	-0.0245***	-0.0941***	-0.0937***
	(0.0060)	(0.0199)	(0.0198)
Low Education	0.3828***	1.0622***	1.0607***
	(0.0145)	(0.0517)	(0.0517)
State, Caste & Year FE	Yes	Yes	Yes
Observations	407,546	407,549	407,549
R-squared	0.2605	0.1373	0.1378

Dependent variables: eligibility for paid leave (col. 1) and coverage by any social security benefit (cols. 2-3), conditional on employment. Col. 3 interacts pre-COVID informality exposure with the COVID-2020 indicator.

Anupam Hazra, Assistant Professor, Department of Economics, University of Kalyani

Email: anupamhazra@klyuniv.ac.in / a.hazra.b13@gmail.com

Table A5. Earnings Effects

Variable	Log Earnings	Log Earnings x Gender
COVID-19 (2020)	0.1158***	0.1030***
	(0.0119)	(0.0123)
COVID x Female		0.0691***
		(0.0223)
Female	-0.7524***	-0.7615***
	(0.0496)	(0.0516)
Urban	0.3444***	0.3443***
	(0.0298)	(0.0298)
Age	0.0035***	0.0035***
	(0.0010)	(0.0010)
Age 25-34	0.0129	0.0130
	(0.0130)	(0.0130)
Low Education	-0.4151***	-0.4151***
	(0.0197)	(0.0197)
State, Caste & Year FE	Yes	Yes
Observations	732,600	732,600
R-squared	0.3522	0.3523

Dependent variable: log monthly earnings, conditional on reporting non-missing earnings. Col. 2 interacts COVID-2020 with gender.

Anupam Hazra, Assistant Professor, Department of Economics, University of Kalyani

Email: anupamhazra@klyuniv.ac.in / a.hazra.b13@gmail.com

Table A6. Sectoral Heterogeneity: COVID-19 Effects by Industry Group

Variable	Informality
COVID-19 (2020)	0.0112**
	(0.0047)
Manufacturing	-0.0187***
	(0.0062)
Construction	0.0037
	(0.0024)
Trade	0.0081**
	(0.0036)
Transport/Storage	0.0033
	(0.0022)
Accommodation/Food	0.0117***
	(0.0039)
Other Services	-0.0433***
	(0.0061)
COVID x Manufacturing	-0.0122
	(0.0079)
COVID x Construction	-0.0035
	(0.0026)
COVID x Trade	-0.0191***
	(0.0046)
COVID x Transport/Storage	-0.0161**
	(0.0060)
COVID x Accommodation/Food	-0.0055
	(0.0033)
COVID x Other Services	-0.0334**
	(0.0137)
Female	0.0144***
	(0.0024)
Urban	-0.0051
	(0.0035)
Age	-0.0002***
	(0.0001)
Age 25-34	-0.0025
	(0.0018)
Low Education	0.0353***
	(0.0064)
State, Caste & Year FE	Yes
Observations	276,738
R-squared	0.0634

Dependent variable: informal employment. Reference industry group is agriculture. Interaction terms report the differential COVID-19 effect for each industry relative to the agricultural baseline.

Anupam Hazra, Assistant Professor, Department of Economics, University of Kalyani

Email: anupamhazra@klyuniv.ac.in / a.hazra.b13@gmail.com

Table A7. Logit Robustness Checks

Variable	Informality (Logit)	Informality x Gender (Logit)
COVID-19 (2020)	-0.0672	0.0221
	(0.1840)	(0.1728)
COVID x Female		-0.4387***
		(0.1200)
Female	-0.0146	0.0821
	(0.0948)	(0.0960)
Urban	-0.5238***	-0.5219***
	(0.1149)	(0.1148)
Age	-0.0188***	-0.0188***
	(0.0025)	(0.0025)
Age 25-34	-0.1502**	-0.1506**
	(0.0741)	(0.0748)
Low Education	2.2263***	2.2249***
	(0.1236)	(0.1236)
State, Caste & Year FE	Yes	Yes
Observations	297,367	297,367
R-squared	0.2070	0.2075

Logit re-estimation of the main informality specification (Table A1, cols. 2 and 3); coefficients are log-odds, not marginal effects. Reported to check that the linear probability model results are not an artifact of the functional form.

Anupam Hazra, Assistant Professor, Department of Economics, University of Kalyani

Email: anupamhazra@klyuniv.ac.in / a.hazra.b13@gmail.com

Table A8. Additional Robustness and Placebo Tests

Variable	Extended COVID Window	Placebo Shock	Placebo x Exposure
COVID-19 (extended)	0.0076		
	(0.0054)		
Placebo shock		0.0033	0.0291**
		(0.0057)	(0.0112)
Pre-COVID Exposure			-0.0195
			(0.0493)
Placebo x Exposure			-0.0787**
			(0.0288)
Female	0.0007	0.0036	0.0036
	(0.0019)	(0.0025)	(0.0024)
Urban	-0.0127***	-0.0162***	-0.0071
	(0.0036)	(0.0035)	(0.0085)
Age	-0.0003***	-0.0008***	-0.0008***
	(0.0001)	(0.0002)	(0.0002)
Age 25-34	-0.0029	-0.0020	-0.0021
	(0.0018)	(0.0026)	(0.0027)
Low Education	0.0442***	0.0577***	0.0577***
	(0.0074)	(0.0109)	(0.0109)
State, Caste & Year FE	Yes	Yes	Yes
Observations	297,367	89,380	89,380
R-squared	0.0468	0.0638	0.0646

Col. 1 uses an extended COVID-period definition (covid_extended). Cols. 2-3 are placebo tests using a fabricated pre-period 'shock' indicator (placebo) that should show no significant effect if the main results are genuinely COVID-specific rather than an artifact of the estimation strategy; col. 3 additionally interacts the placebo indicator with pre-COVID exposure.

Anupam Hazra, Assistant Professor, Department of Economics, University of Kalyani

Email: anupamhazra@klyuniv.ac.in / a.hazra.b13@gmail.com

Table A9. Gender-Disaggregated Subsample and Interaction Estimates

Variable	Women Only	Men Only	Full Sample (sex interaction)
COVID-19 (2020)	-0.0103	0.0006	0.0009
	(0.0101)	(0.0051)	(0.0052)
Sex=2 (Female)			0.0023
			(0.0019)
Sex=3			-0.0643
			(0.0798)
COVID x Sex=2			-0.0124**
			(0.0056)
COVID x Sex=3			0.0856
			(0.0814)
Urban	-0.0149***	-0.0117***	-0.0127***
	(0.0037)	(0.0036)	(0.0036)
Age	-0.0002***	-0.0004***	-0.0003***
	(0.0001)	(0.0001)	(0.0001)
Age 25-34	-0.0024	-0.0030	-0.0029
	(0.0018)	(0.0021)	(0.0018)
Low Education	0.0520***	0.0421***	0.0442***
	(0.0071)	(0.0076)	(0.0074)
State, Caste & Year FE	Yes	Yes	Yes
Observations	64,605	232,762	297,399
R-squared	0.0515	0.0467	0.0470

Cols. 1-2 estimate the baseline informality specification separately on the female and male subsamples. Col. 3 uses the full sample with a three-category sex variable (2=female, 3=other/third gender where identified) interacted with COVID-2020, relative to the male reference category.

Anupam Hazra, Assistant Professor, Department of Economics, University of Kalyani

Email: anupamhazra@klyuniv.ac.in / a.hazra.b13@gmail.com

Table A10. Event-Study Estimates: Year x Pre-COVID Exposure

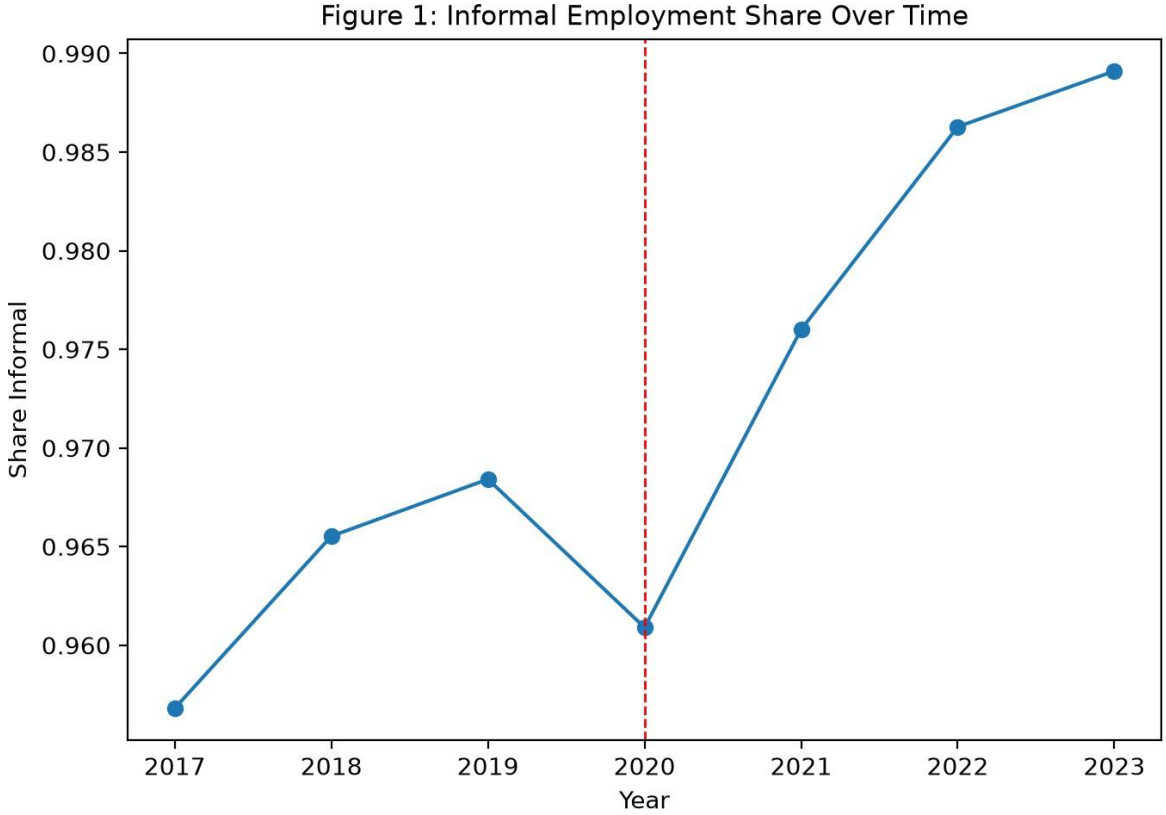
Dependent variable: informal employment. Each row reports the coefficient on the interaction between a single survey year and continuous pre-COVID informality exposure, with 2019 (the last full pre-COVID round) as the omitted reference year. The 'Year' column reports the corresponding year main effect. Robust standard errors clustered at the state level in parentheses.

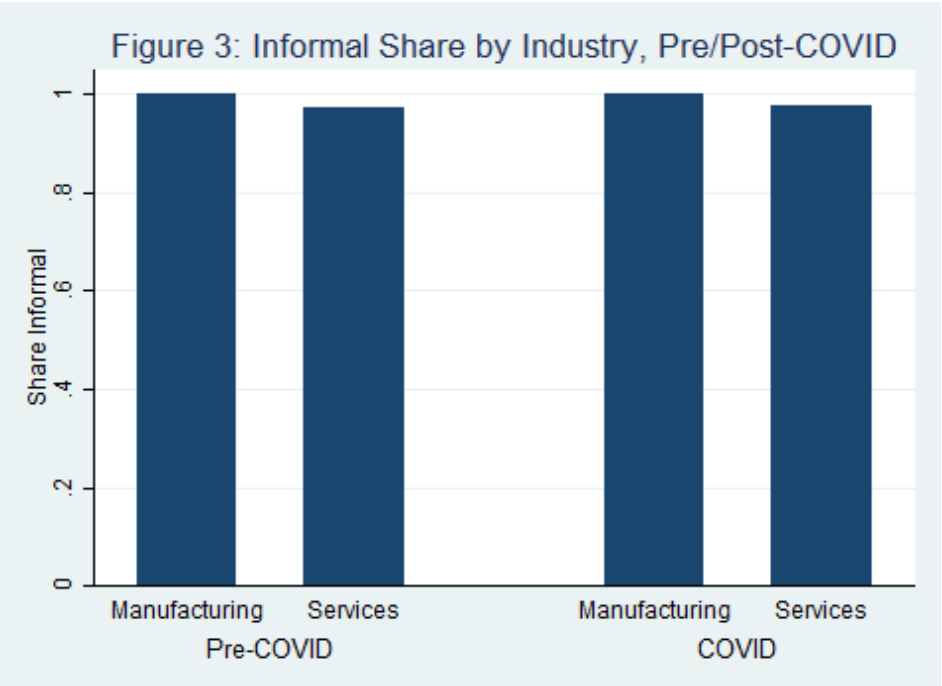
Survey Year	Year main effect	Year x Exposure
2017	-0.0209	0.0162
	(0.0157)	(0.0356)
2018	0.0084	-0.0635
	(0.0105)	(0.0391)
2020	-0.0025	-0.0446
	(0.0111)	(0.0446)
2021	-0.0039	-0.0134
	(0.0101)	(0.0367)
2022	0.0043	-0.0130
	(0.0067)	(0.0216)
Observations	297,367	
R-squared	0.0477	

Anupam Hazra, Assistant Professor, Department of Economics, University of Kalyani

Email: anupamhazra@klyuniv.ac.in / a.hazra.b13@gmail.com

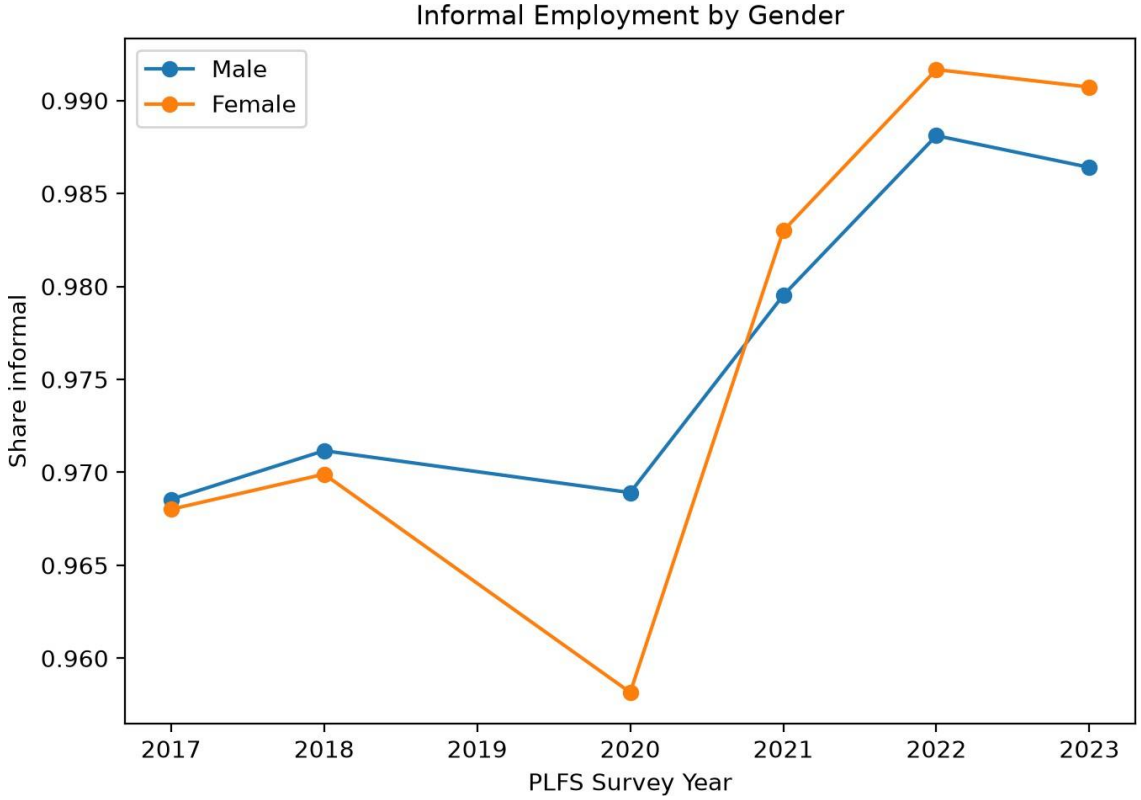
Figure:





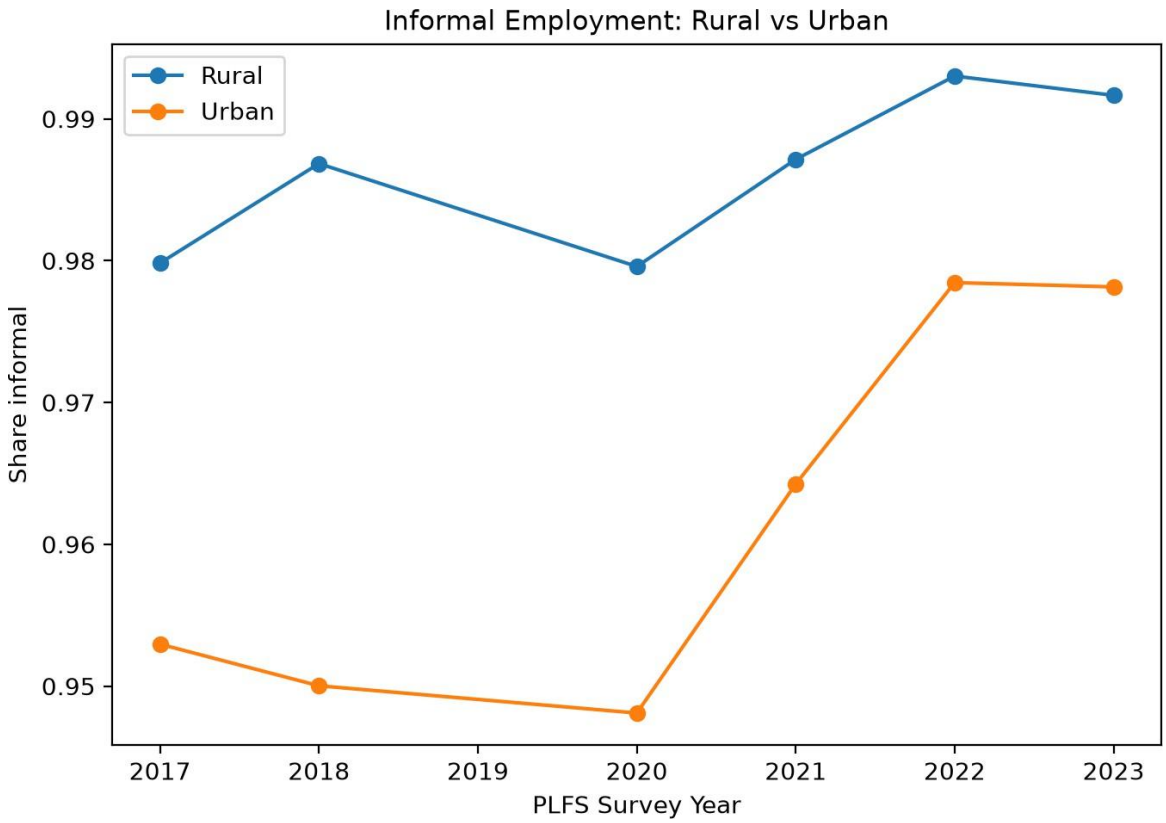
Anupam Hazra, Assistant Professor, Department of Economics, University of Kalyani

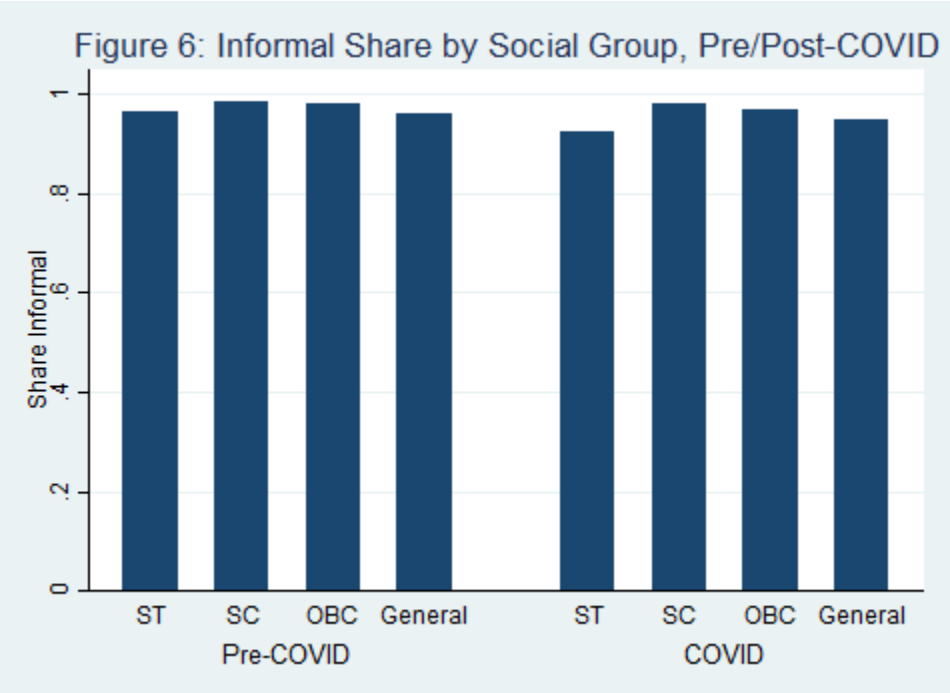
Email: anupamhazra@klyuniv.ac.in / a.hazra.b13@gmail.com



Anupam Hazra, Assistant Professor, Department of Economics, University of Kalyani

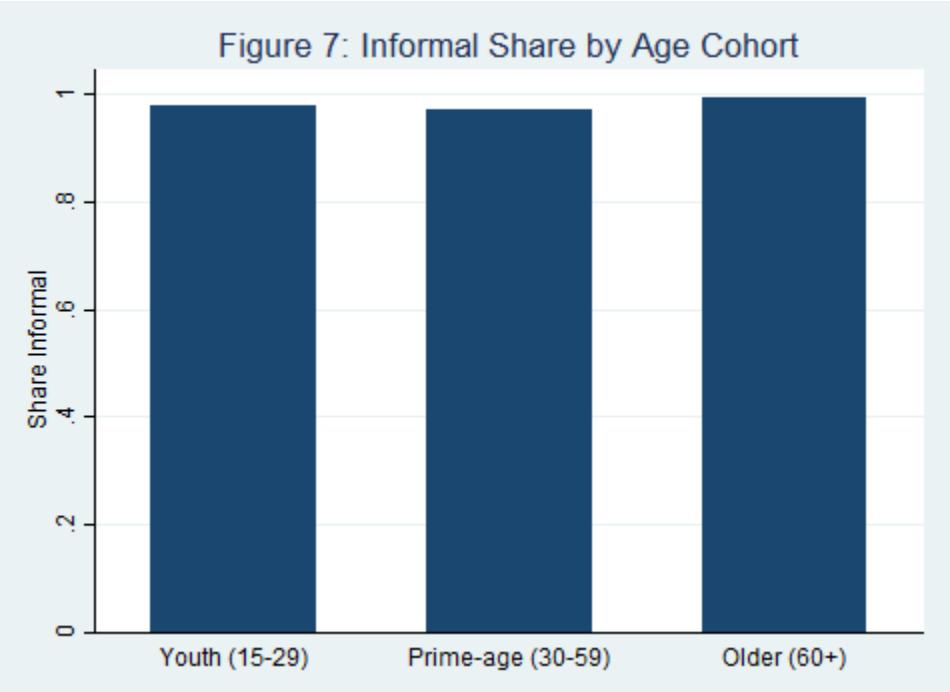
Email: anupamhazra@klyuniv.ac.in / a.hazra.b13@gmail.com

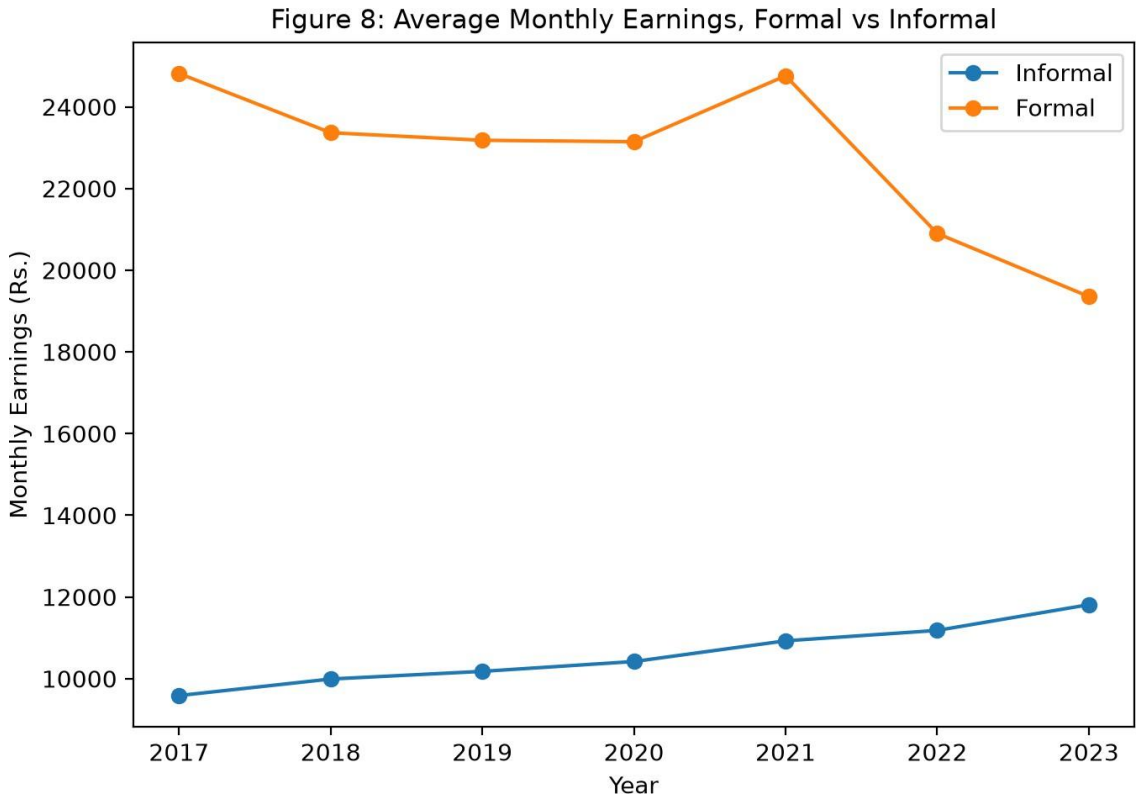




Anupam Hazra, Assistant Professor, Department of Economics, University of Kalyani

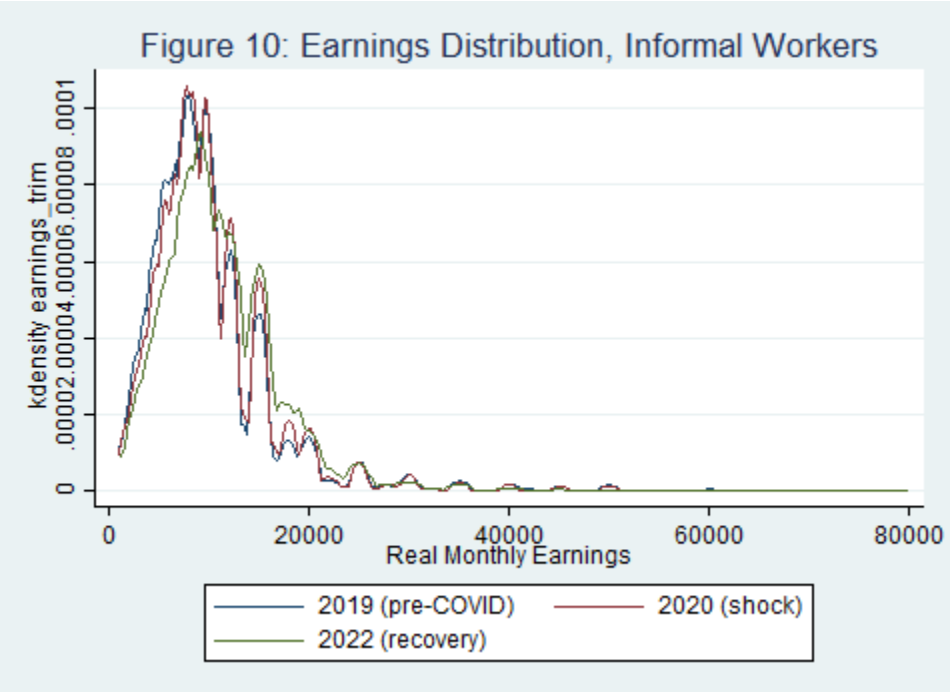
Email: anupamhazra@klyuniv.ac.in / a.hazra.b13@gmail.com

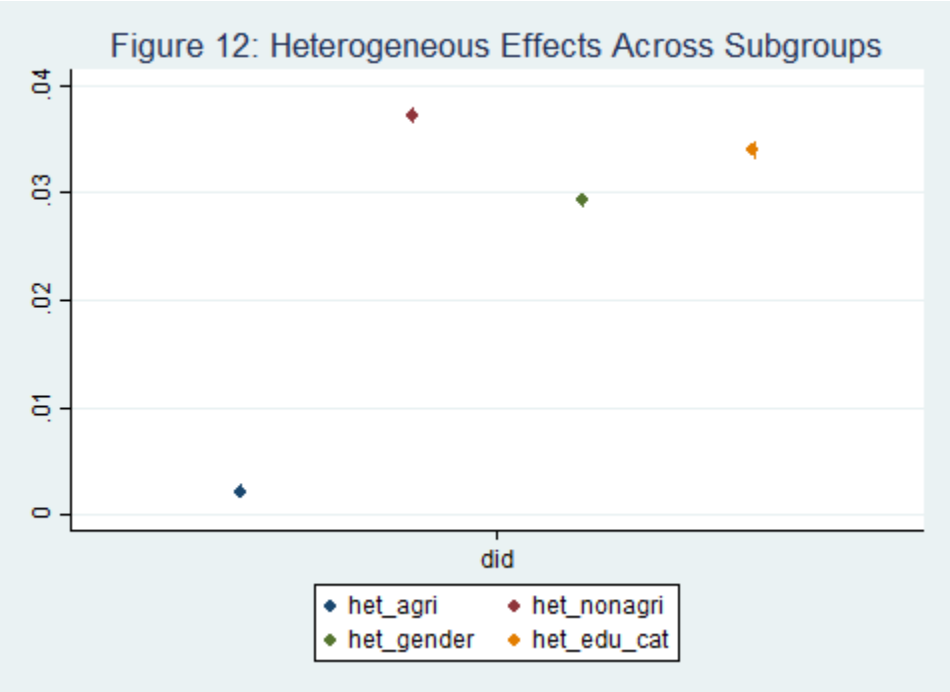




Anupam Hazra, Assistant Professor, Department of Economics, University of Kalyani

Email: anupamhazra@klyuniv.ac.in / a.hazra.b13@gmail.com





Anupam Hazra, Assistant Professor, Department of Economics, University of Kalyani

Email: anupamhazra@klyuniv.ac.in / a.hazra.b13@gmail.com

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- AppendixRegressionTables.docx