STATE OF WORKING INDIA 2023

Social Identities and Labour Market Outcomes

STATE OF WORKING INDIA 2023

Social Identities and Labour Market Outcomes

Disclaimer: The analysis and opinions expressed in this publication are those of the author(s) and may not represent the view of the University.

Suggested citation:

State of Working India (2023) Social identities and labour market outcomes, Centre for Sustainable Employment, Azim Premji University.

About Azim Premji University's Work on Sustainable Employment

Azim Premji University was established in 2010, by the Azim Premji Foundation, with a clear social purpose of working towards a just, equitable, humane, and sustainable society. All of the University's programmes, teaching, research, and practice, work towards this purpose.

To contribute to the critical matter of India creating just and sustainable employment, the University has set up the Centre for Sustainable Employment (CSE), which conducts and supports research in areas of work, labour, and employment. The University is attempting to provide empirically grounded, analytical reflections on the state of work and workers in India, as well as to evaluate and propose policies that aim to create sustainable jobs. To this end the University also gives grants to create new knowledge in the above areas. It also hosts a working paper series to which contributions are invited from researchers, policy-makers, civil society actors, and journalists. The University's CSE website is an important part of this agenda. In addition to research papers and policy briefs, it hosts government reports, as well as data and statistics on the Indian labour market.

Website: https://cse.azimpremjiuniversity.edu.in/
Twitter: @working_india
Facebook: https://www.facebook.com/centreforsustainableemployment
LinkedIn: https://www.linkedin.com/company/centre-for-sustainable-employment/
Email: cse@apu.edu.in

About IWWAGE

The Institute for What Works to Advance Gender Equality (IWWAGE) aims to build
on existing research and generate new evidence to inform and facilitate the agenda of
women's economic empowerment. IWWAGE is an initiative of LEAD, an action-oriented
research centre of IFMR Society (a not-for-profit society registered under the Societies
Act). LEAD has strategic oversight and brand support from Krea University (sponsored by
IFMR Society) to enable synergies between academia and the research centre.

Website: https://iwwage.org/

Report Team

Authors

Amit Basole, Rosa Abraham, Angarika Rakshit, R. Vijayamba, Anand Shrivastava,
Tamoghna Halder

Research Assistance

Akshit Arora, M.K. Shravan, Shritha Sampath, Hami Ameen, Hammad Farooqui,
Subhapriya Chakraborty

Interns

Ayan Sadhukan, Vaishnavi Krishna Mohan, Ananya Kumar, Nidhi C

Design

Silja Bansriyar, Vipul S Kumar, Nanit BS, Palak Sharma, Kshiraja Krishnan

Outreach

Shashwat DC, Akanksha Ojha, Sachin Mulay, K2 Communications

India Working Survey Team

Principal investigators

Rosa Abraham, Girish Bahal, Amit Basole, Deepti Goel, Rahul Lahoti, Anand Shrivastava,
Hema Swaminathan

Research Assistance

Paaritosh Nath, Atma Jyoti Mahapatra, Sravan P, Mridhula Mohan, Zaeen D'Souza,
Priyanshu Krishnamoorthi, Kasturi Kandalam, Aru Bhardwaj, Aaqib Qayoom

Field Team

Debmalya Nandy, Neha Joshi, Adithi Rao, Imran Rajpoot, Shivprasad Amblimath, Suresh
Chandra, Suchitra Yegna Narayan, Banu P., T. R. Ramakrishna

Voices Team

Priyanshu Krishnamoorthi, Anumeha Yadav, Shreya Ghosh, Parijatha GP, Deepa C,
Arpitha R, Rashmi Bala, Mansa

Acknowledgements

Table of Contents

List of Figures

List of Tables

List of Boxes

Abbreviations

ASI	Annual Survey of Industries
CAGR	Compounded Annual Growth Rate
CMIE-CPHS	Centre for Monitoring Indian Economy-Consumer Pyramid Household Survey
CPI	Consumer Price Index
CWS	Current Weekly Status
GDP	Gross Domestic Product
GoI	Government of India
GST	Goods and Services Tax
IHDS	India Human Development Survey
ILO	International Labour Organisation
IWS	India Working Survey
IWWAGE	Institute for What Works to Advance Gender Equality
LEAD	Leveraging Evidence for Access and Development
LFPR	Labour Force Participation Rate
MOSPI	Ministry of Statistics and Programme Implementation
MSME	Micro, Small and Medium Enterprise
NSS-EUS	National Sample Survey-Employment Unemployment Survey
NSO	National Statistical Office
NSSO	National Sample Survey Office
OBC	Other Backward Caste
PLFS	Periodic Labour Force Survey
RBI	Reserve Bank of India
SC	Scheduled Caste
ST	Scheduled Tribe
SWI	State of Working India
UPS	Usual Principal Status
UPSS	Usual Principal and Subsidiary Status
UR	Unemployment Rate
WPR	Workforce Participation Rate

State Abbreviations

AP	Andhra Pradesh
AR	Arunachal Pradesh
AS	Assam
BR	Bihar
CG	Chhattisgarh
DL	Delhi
GA	Goa
GJ	Gujarat
HP	Himachal Pradesh
HR	Haryana
JH	Jharkhand
JK	Jammu and Kashmir
KA	Karnataka
KL	Kerala
MH	Maharashtra
ML	Meghalaya
MN	Manipur
MP	Madhya Pradesh
MZ	Mizoram
NL	Nagaland
OR	Odisha
PB	Punjab
RJ	Rajasthan
SK	Sikkim
TS	Telangana
TN	Tamil Nadu
TR	Tripura
UK	Uttarakhand
UP	Uttar Pradesh
WB	West Bengal

Foreword

Good news in a report on jobs is rare anywhere in the world today. Even more so after the pandemic with its very uneven and lingering effects. So, that the State of Working India 2023 (SWI) is full of good things that have happened in India on jobs and livelihood is not only heartening but also energizing. This SWI is, as always, rigorously researched, and the basic story that it tells is that of progress on many dimensions of jobs, their quality and equity and justice.

Significant increase in proportion of those who are employed with regular wage or salaried work is directly reflective of the improvement in the quality of jobs. There is clear evidence that upward mobility has increased, caste-based segregation has reduced, and gender disparities have also reduced. These are all remarkable indicators of progress which are not there in public discourse. For India, this should be even more heartening because some of these inequities have been the sharpest matters of injustice in our country.

As we must expect, the report also points out the challenges that the country continues to face including that at the absolute level, many of the inequities remain while significant progress has been made. It also points out a phenomenon that has been seen across the world, that the connection between economic growth and jobs is weak.

We must do a lot more to address the challenges that continue. But the clear progress that we see tells us that we can overcome all these challenges. With this sense of confidence and hopefulness, we should also confront the emerging challenges, for example, from the effect of Artificial Intelligence on jobs, from the urgent need to transition to a green economy, and the structural realignment of the global political economy.

We are glad to share with you this volume of SWI, which is clear-eyed about the task ahead and at the same time summarizes the achievements which must give us the energy to move forward.

Indu Prasad
Vice Chancellor, Azim Premji University
September 9th, 2023

Executive Summary

The Indian story of economic growth and structural transformation has been one of significant achievements as well as continuing challenges. On the one hand, the economy has grown rapidly since the 1980s, drawing millions of workers out of agriculture. And the proportion of salaried or regular wage workers has risen while that of casual workers has fallen. On the other hand, manufacturing has failed to expand its share of GDP or employment significantly. Instead construction and informal services have been the main job creators. Further, the connection between growth and good jobs continues to be weak.

When we speak of new opportunities, another important set of questions arises. Who is able to take advantage of them, and who is not? Has growth created faster improvements for marginalised groups, enabling them to catch up with more advantaged groups? This year's report takes a detailed look at the impact that growth and structural change have had on some long-running social disparities. We show that significant progress has been made on all fronts since the 1980s, but also that there is a long road ahead.

The report makes use of official datasets such as the NSO's Employment-Unemployment Surveys, the Periodic Labour Force Surveys, the National Family Health Surveys, Annual Survey of Industries, and the Economic and Population Censuses. We also make use of a unique primary survey carried out in rural Karnataka and Rajasthan, the India Working Survey. This year's report goes further than our earlier three editions and makes extensive use of regression analysis to offer more precise estimates of the impacts of structural change on employment conditions and outcome gaps.

Highlights from the report:

- **Faster structural change:** After stagnating since the 1980s, the share of workers with regular wage or salaried work started increasing in 2004, going from 18% to 25% for men and 10% to 25% for women. Between 2004 and 2017, around 3 million regular wage jobs were created annually. Between 2017 and 2019 this jumped to 5 million per year. Since 2019, the pace of regular wage jobs creation has decreased due to the growth slowdown and the pandemic.
- **Upward mobility has increased:** In 2004 over 80% of sons of casual wage workers were themselves in casual employment. This was the case for both SC/ST workers and other castes. For non-SC/ST castes, this fell from 83% to 53% by 2018 and incidence of better quality work such as regular salaried jobs increased. It fell for SC/ST castes as well, but to a lesser extent (86% to 76%).
- **Caste-based segregation has reduced:** In the early 1980s Scheduled Caste workers were more than 5 times over-represented in waste-related work and over 4 times in leather-related work. This has declined rapidly over time, though it is not completely eliminated as of 2021-22. In the leather industry, the representation index declined sharply to 1.4 in 2021. In waste management and sewerage, over-representation of SCs decreased to 1.6 times in 2011 before increasing slightly again.

- **Gender-based earnings disparities have reduced:** In 2004, salaried women workers earned 70% of what men earned. By 2017 the gap had reduced and women earned 76% of what men did. Since then the gap has remained constant till 2021-22.

- **Connection between growth and good jobs remains weak:** Since the 1990s year-on-year non-farm GDP growth and non-farm employment growth are uncorrelated with each other suggesting that policies promoting faster growth need not promote faster job creation. However, between 2004 and 2019, on average growth translated to decent employment. This was interrupted by the pandemic which caused larger growth in distress employment.

- **Unemployment is falling but remains high:** Post-Covid the unemployment rate is lower than it was pre-Covid, for all education levels. But it remains above 15% for graduates and more worryingly it touches a huge 42% for graduates under 25 years.

- **After falling for years, women's WPR is rising, but not for the right reasons:** After falling or being stagnant since 2004, female employment rates have risen since 2019 due to a distress-led increase in self-employment. Before Covid, 50% of women were self-employed. After Covid this rose to 60%. As a result earnings from self-employment declined in real terms over this period. Even two years after the 2020 lockdown, self-employment earnings were only 85% of what they were in the April-June 2019 quarter.

- **Gender norms continue to be significant for women's employment:** As husband's income rises, women are less likely to work. In urban areas, after the husband's income crosses ₹40,000 per month, the chance of the wife working increases again (i.e. there is a U-shaped relationship). There is also a strong intergenerational effect of gender norms. Compared to households where there is no mother-in-law present, married women living in households where the mother-in-law is present but not employed are 20% (rural) to 30% (urban) less likely to be employed. However, if the mother-in-law is employed herself, daughters-in-law are 50% (rural) to 70% (urban) more likely to be employed.

- **Marginalised caste entrepreneurs are still rare:** We find that even in the smallest firm sizes, SC and ST owners are under-represented compared to their share in the overall workforce. But even more significantly, SC and ST owners are barely represented among firms employing more than 20 workers. Correspondingly, General caste overrepresentation increases with firm size.

In this executive summary we elaborate on the above points. The notes below a figure or a table point the reader to the relevant portion of the report for more information and context.

1. Growth pulled people out of agriculture and the share of regular salaried workers rose. But women left the workforce and informality levels remain a concern.

a. Although share of non–agricultural employment rose, it was not matched by a similar increase in the share in regular wage employment or employment in the organised sector

Share of workers in non–agricultural employment rose much faster than the share in regular wage employment or organised sector
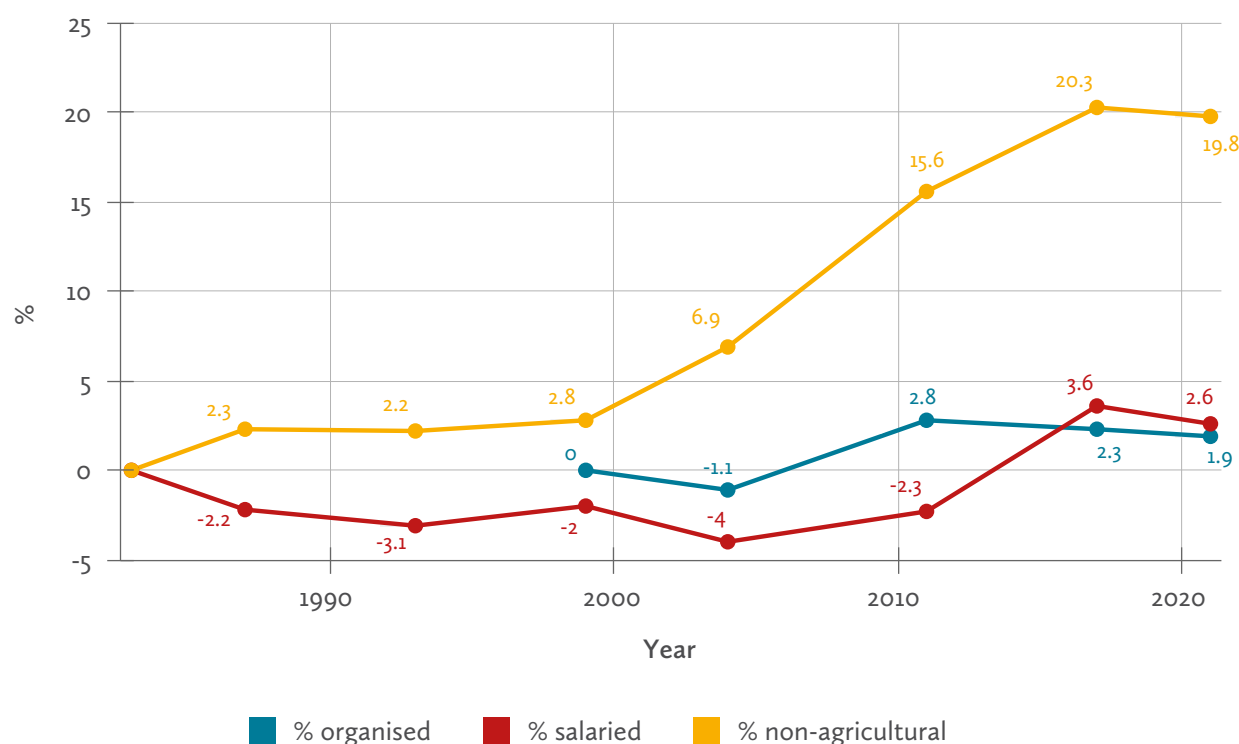
Sources and notes: NSS EUS and PLFS various rounds. The numbers are indexed to their 1983 values. % salaried implies the proportion of regular wage workers in the non-agricultural sector. This is Figure 3.7 in the report.

One sign of a successful structural transformation is a decline in agricultural employment share, accompanied by a rise in the share of regular wage or salaried workers. Between 1983 and 2019 (eve of the pandemic) the share of the non-farm sector in employment rose 20 pct pts, but the majority of such jobs were of the informal variety. On net, the share of regular wage work increased less than 3 pct pts and that of the organised sector less than 2 pct pts. The period since 2004 saw a more rapid increase in salaried work, going from 15% of the workforce to 25% by 2018 before falling due to the pandemic.

b. A large share of women exited the workforce but for those who remained in employment, the share in regular wage work increased.

Workforce participation rate over the long run for men and women in rural and urban areas
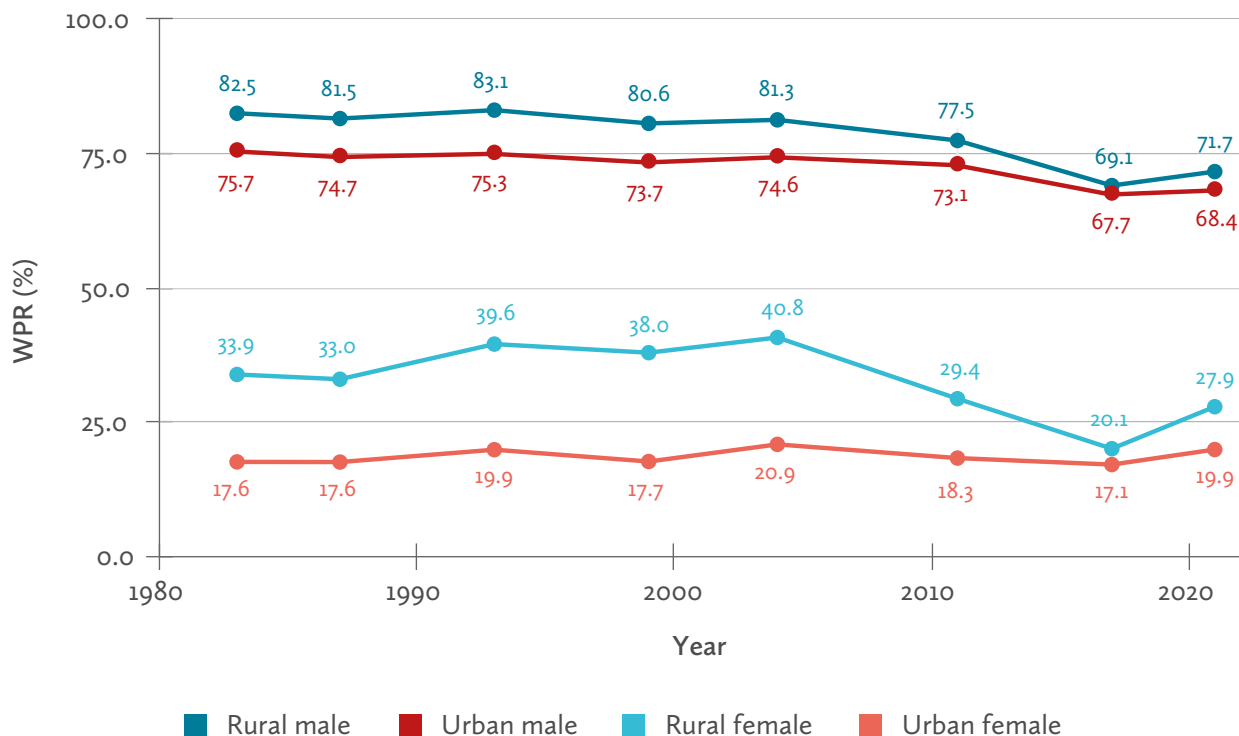
Sources and notes: NSS EUS and PLFS various rounds.*This is Figure 2.1 in the report.*

For men, exit from agriculture meant a large increase in the share of construction while for women it meant an exit from the workforce. The figure shows two main challenges with respect to women's employment - the decline in the rural female workforce participation rate (WPR) and the stagnant, low urban rate. The recent rise in women's WPR since 2019 is discussed in Point 8. The period between 2004 and 2018 (just prior to the slowdown and the pandemic) saw a dramatic change in the composition of the female workforce. Older, less educated women working in agriculture exited while younger more educated women entered. As a result the large decline in the employment rate for rural women was accompanied by a significant increase in the proportion of women workers who had regular wage jobs - going from just under 10% in 2004 to 25% in 2018 before falling to 20% due to the pandemic.

c. For men, a doubling of GDP reduced agricultural share of employment much more than informal share. For women who remained in the workforce, a decline in agriculture share was accompanied by a decrease in informal share.

Combining employment data with State Domestic Product (SDP) data, we measure how effective growth was in pulling workers out of agriculture as well as in creating regular wage jobs at the state level. We find significant variation across states. The figures show the pct pt change in agriculture share (X axis) and regular wage share (Y axis) with a doubling of SDP per capita. For men, the decline in agricultural share was much faster as compared to the rise in regular wage share (left). For example, in Bihar, between 1983 and 2018, a doubling of SDP per capita reduced agricultural share of male employment by 20 pct pts but there was no

change in the proportion of regular wage male workers. But for women, the story is different (right). Leaving aside women who left the workforce (see earlier point), for those who remained, the share of regular wage workers increased rapidly in almost all states.

Bringing together the Lewis and Kuznets processes: extent of decline in agricultural share and increase in regular wage share with a doubling of SDP across states.
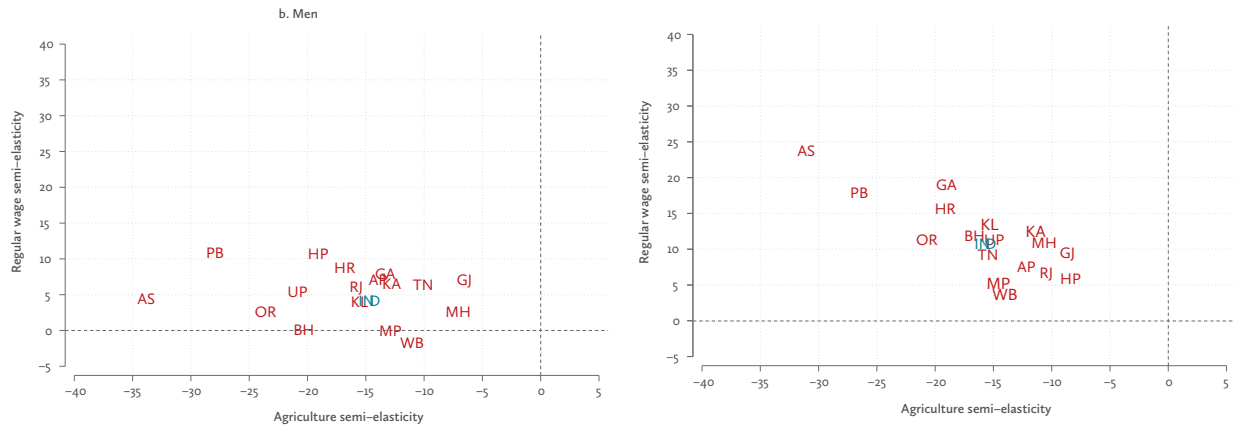
Sources and notes: NSS EUS and PLFS various rounds. This is Figure 3.23 in the report.

d.   Growth created regular wage jobs for all castes but at very different rates.

Changing composition of employment type – caste
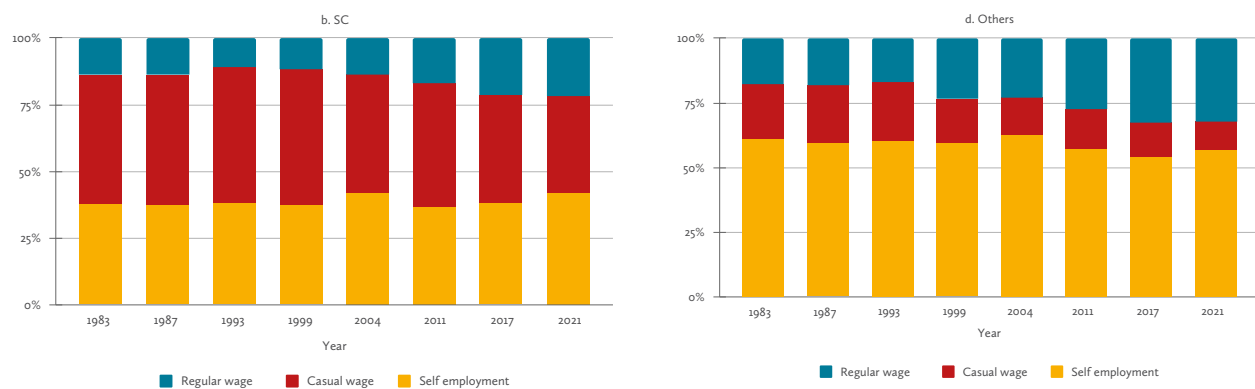
Sources and notes: NSS EUS and PLFS various rounds. Others - General category (excld. SC, ST, and OBC). This is Figure 3.13 in the report.

The figure shows the proportion of SC (left) and Other (right) workers found in various types of employment. As can be seen, SC workers are far more likely to be in casual employment as compared to Others. But encouragingly, the proportion of regular wage SC workers has risen since 2011. As of 2021-22, around 22% of SC workers were regular wage as compared to 32% of Others. But 40% of SC workers were in casual employment as compared to only 13% for Others.Jati-level analysis within the SC category using Census 1991 and 2011 data shows that significant variation exists in movement out of agriculture across jatis but strikingly, movement into construction (a principal source of casual work) observed at the aggregate SC level is seen for almost all SC jatis across states.

The employment structure varies far more across gender and caste identities than it does across religion. But, it is worth noting that Muslims are less likely to hold regular wage jobs and more likely to be in own-account or casual wage work over the entire four decade period after controlling for education, household size, state and other relevant factors. The persistent under-representation in regular wage work was noted in the Sachar Committee Report of 2006 as well and continues to be a matter of concern.

2. Intergenerational mobility has increased but less so for marginalised castes.

An analysis of father-son pairs in the NSSO employment surveys shows that, over the last 15 years, upward mobility has increased in terms of the type of work performed. In 2004 over 80% of sons of casual wage workers were themselves in casual employment (bottom right coloured cell). This was the case for both SC/ST workers and other castes. For non-SC/ST castes, by 2018, this proportion had fallen to 53% with the rest engaged in regular wage work (mostly of the informal kind). It fell for SC/ST castes as well, but to a lesser extent (75.6%). Another sign of improving job conditions is that the percentage of sons who had regular wage jobs like their fathers rose significantly between 2004 and 2018 (top left cell in table).

Intergenerational mobility matrices over time disaggregated by caste

SC/ST

Year 2018

Son's Employment	Father's employment				
	Formal RW	Semi-formal RW	Informal RW	Self-employed	Casual
Formal RW	38.6	0.2	2.2	3.0	2.0
Semi-formal RW	10.5	43.3	6.1	3.8	2.7
Informal RW	20.9	22.2	64.1	11.0	14.1
Self-employed	16.7	26.5	11.8	62.9	5.6
Casual	13.3	7.9	16.0	19.2	75.6
Col Sum	100	100	100	100	100

SC/ST

Year 2004

Son's Employment	Father's employment				
	Formal RW	Semi-formal RW	Informal RW	Self-employed	Casual
Formal RW	14.8	2.7	0.2	1.6	0.2
Semi-formal RW	7.0	14.4	0.8	1.4	1.0
Informal RW	23.4	18.8	70.4	5.0	5.6
Self-employed	37.8	45.1	19.7	74.7	6.7
Casual	17.0	18.7	8.8	17.4	86.5
Col Sum	100	100	100	100	100

Others

Year 2018

Son's Employment	Father's employment				
	Formal RW	Semi-formal RW	Informal RW	Self-employed	Casual
Formal RW	34.4	2.5	16.8	6.3	4.7
Semi-formal RW	17.6	50.0	12.5	9.1	7.6
Informal RW	15.7	21.7	50.4	14.3	22.6
Self-employed	28.5	25.0	15.1	64.0	12.1
Casual	3.9	0.9	5.2	6.5	53.0
Col Sum	100	100	100	100	100

Others

Year 2004

Son's Employment	Father's employment				
	Formal RW	Semi-formal RW	Informal RW	Self-employed	Casual
Formal RW	17.3	1.3	4.2	2.7	0.0
Semi-formal RW	9.7	15.5	0.2	2.2	1.2
Informal RW	11.2	19.0	57.6	6.8	3.5
Self-employed	52.9	58.2	19.8	80.9	12.1
Casual	9.0	6.2	18.2	7.5	83.2
Col Sum	100	100	100	100	100

Sources and notes: NSSO EUS and PLFS various rounds. Numbers are percentages adding to 100 along the columns. Warmer colour indicate higher values. RW- Ragular wage .This is Table 5.3 in the report.

3. Job creation continues to be India's main challenge

a. The relationship between growth and employment became weaker over time

Despite the improvements noted above, economic growth remains weakly connected to job creation. Since the 1980s, non-farm output consistently grew much faster than non-farm employment resulting in a steady fall in the employment elasticity (output growth divided by employment growth). India's employment elasticity is far lower than the developing country average. The most recent period (2017-2021) is an exception to the trend. Growth slowed down significantly while employment growth quickened and the elasticity increased. This period is discussed in Point 8. Over the long-run GDP growth and employment growth have been uncorrelated in India suggesting that policies oriented towards achieving faster GDP growth will not necessarily speed up job creation.

Period	Non-agricultural value-added growth (%)	Non-agricultural employment growth (%)	Non-agricultural employment elasticity
1983 to 1993	6.5	4.2	0.6
1993 to 2004	6.9	3.2	0.5
2004 to 2011	7.7	3.4	0.4
2011 to 2017	7.4	1.7	0.2
2017 to 2021	3.4	3.4	1

Sources and notes - RBI Database of the Indian Economy, RBI KLEMS, NSSO Employment-Unemployment Surveys, Periodic Labour Force Survey. This is adapted from Table 3.1 in the report.

b.   Open unemployment remains high among educated youth

Encouragingly, unemployment is lower post-Covid for all education levels. But it remains above 15% for graduates and more worryingly it touches a huge 42% for young graduates. At the other extreme, among older, less educated workers, it is in the range of 2-3%.

Unemployment is concentrated among educated youth

	Less than 25 years	25–29 years	30–34 years	35–39years	40 years and above
Graduate & above	42.3	22.8	9.8	4.5	1.6
Higher secondary	21.4	10.6	5.0	3.1	2.1
Secondary	18.1	7.5	4.6	2.4	1.7
Primary or middle	15.0	5.4	3.0	2.4	2.2
Literate but below primary	10.6	3.3	1.5	2.4	2.2
Illiterate	13.5	4.3	4.0	3.4	2.4

Sources and notes: PLFS 2021-22. This is Figure 2.3 in the report.

There is large variation in the rate of unemployment even within the higher educated group. The unemployment rate falls from over 40% for educated youth under 25 years of age to less than 5% for graduates who are 35 years and above. This indicates that eventually graduates do find jobs but the key questions are, what is the nature of jobs they find and do these match their skills and aspirations? More research is needed on this important topic.

4. The majority of women still remain outside the workforce due to supply and demand side challenges.

a. Indian states are consistently lower than comparable developing countries in terms of the female LFPR

It is well-known that the female LFPR varies a lot across states of India. But despite this variation there is an overall "India effect". The figure plots the log SDP per capita in constant US dollars against the female LFPR for all Indian states (red) and GDP per capita for all the countries in the World Development Indicators database (blue). The data is for the year 2018. Note that all the states with the exception of Himachal Pradesh (51%), Meghalaya (49%)and Sikkim (46.5%) lie well below the line of best fit. This means that they have lower rates of female LFPR than predicted for their level of per capita GDP.

Indian states are consistently lower than comparable developing countries in terms of the female LFPR
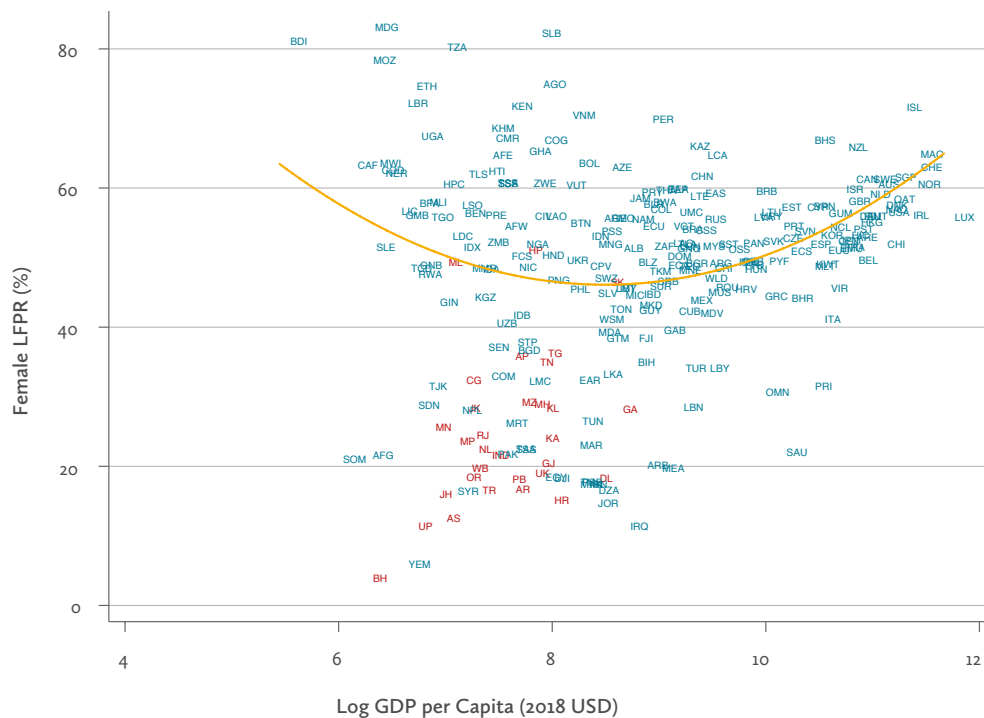
Sources and notes: WDI and PLFS 2018-19. *This is Figure 2.11 in the report.*

b. Employment rates vary significantly across social groups.

The decline in women's employment since 2004 has been witnessed across social groups, but the levels are very different pointing to the possible importance of gender norms on the supply-side (such as barriers to mobility) as well as the demand side (such as discrimination by employers).

Long-run female workforce participation rates by caste and religion

Sources and notes: *NSS EUS and PLFS various rounds. This is Figure 4.1 in the report.*

c. As the husband's earnings rise, the wife's probability of being employed first falls and then rises.

A common gender norm is the "male breadwinner" norm, i.e. husbands are considered the primary earners with wives contributing to household income only if necessary. The PLFS data shows that in households where the husband's earnings are high, the probability of the wife being employed is low, controlling for individual and household factors. The fall slows down as husband's incomes increase in rural India, while for urban areas, there is a reversal beyond a certain level of income. As husband's earnings keep rising, from approximately ₹40,000 per month onwards, there is an increased likelihood of wives being employed. This U-shaped pattern could result from a change in norms with rising incomes or it could also be due to the fact that such husbands are matched with higher educated wives who have preferences as well as opportunities to access better paid work.

Impact of husband's earnings on probability of being employed for women
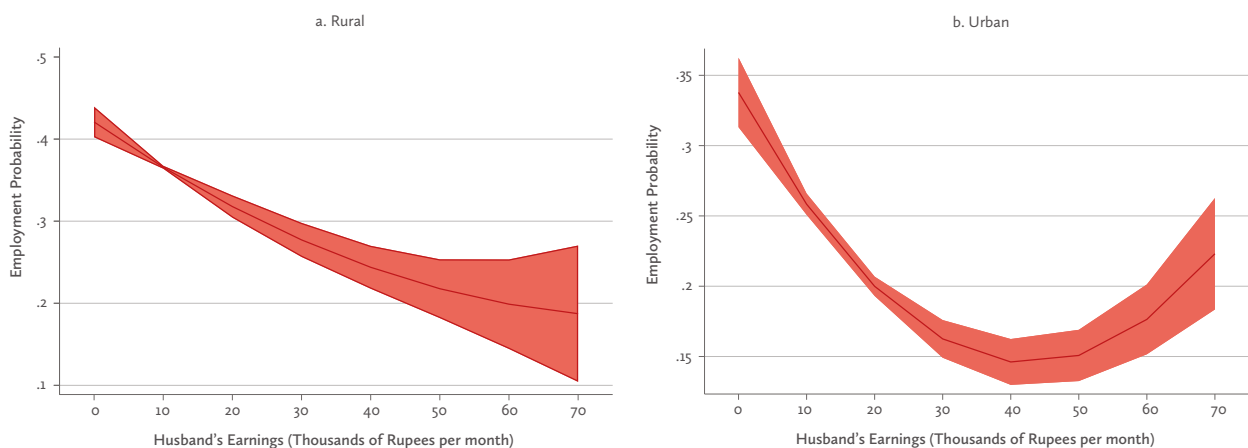
Sources and notes: *PLFS 2021-22. The marginal effects for each level of income are shown along with the confidence bands. See Chapter 4 and Methods Appendix for details. This is Figure 4.4 in the report.*

We also find a strong intergenerational effect of gender norms. Compared to households where there is no mother-in-law present, married women living in households where the mother-in-law is present but not employed are 20% (urban) to 30% (rural) less likely to be employed. However, if the mother-in-law was

present and employed herself, this is associated with a higher likelihood of women's participation to the extent of 50% (rural) to 70% (urban) more than households with no mother-in-law present.

d.   Districts where women experience less restrictions are also the ones where women are likely to undertake paid work but there is evidence for a backlash effect as well.

The National Family Health Survey (NFHS) collects data on women's ownership of assets and their ability to make decisions in the household with regard to their own mobility as well as household purchases. It also collects data on justifications offered for domestic violence as well as actual incidence of violence.

Progressive norms on women's autonomy correlate with higher likelihood of employment, but there is evidence for male backlash also
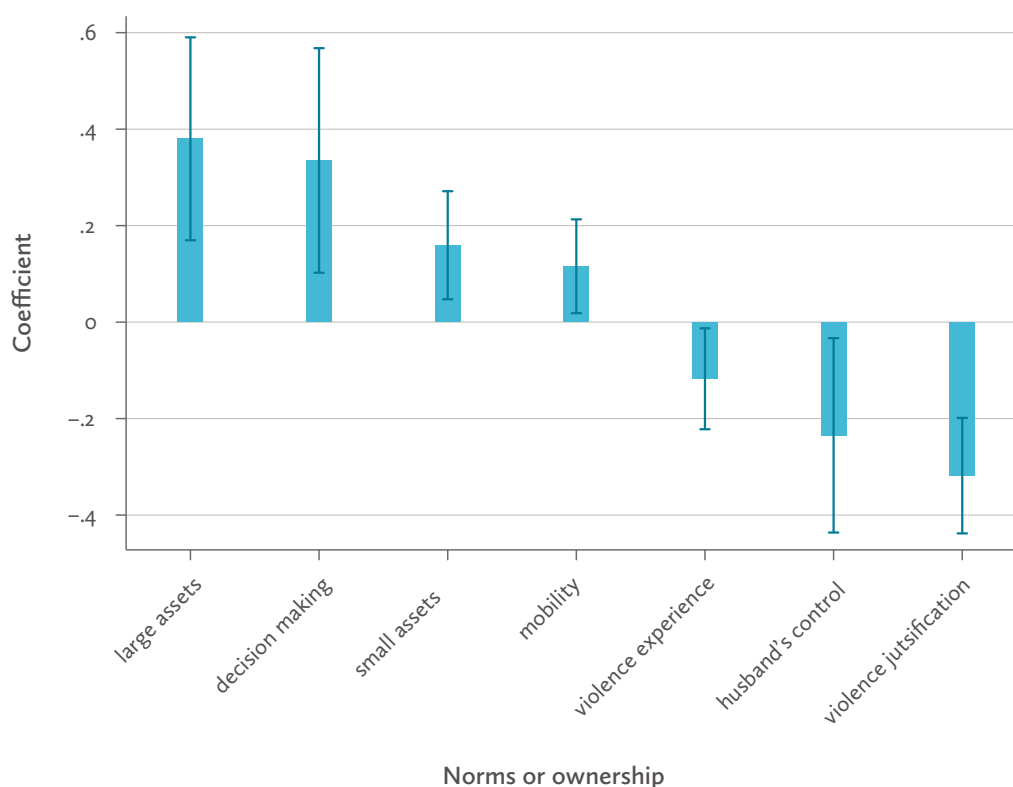
Sources and notes:NFHS 2015-16. The plot shows the coefficients from a regression of women's employment status on district level norms. This is Figure 4.10 in the report.

We find that district-level norms are significantly correlated with the probability of a woman being employed. A 10% increase in the district-level proportion of women who report the ability to make their own decisions on seeking healthcare or meeting relatives, or owning large assets is associated with a 4% increase in the probability of a woman in that district working.

But women who are more likely to do paid work reside in districts where domestic violence is more prevalent. This could be because of the backlash effect. The suggested mechanism is that working women are more likely to face partner violence due to challenging of established gender norms.

e.   Marriage increases the likelihood of being employed for women in rural Karnataka and Rajasthan

As part of the India Working Survey (IWS), we collected long-term, retrospective data on men's and women's life events and occupational history from the time they were 15 years of age. Women experienced a sharp jump in workforce participation from 26% in the year preceding marriage to 49% in the first five years of marriage. This employment was largely as contributing family workers or self-employment in agricultural work. Thus, for women in the informal economy, rather than a marriage or motherhood penalty for employment, we find the reverse. But the absence of a penalty may not be a positive outcome since it likely reflects a compulsion to work for subsistence.

Marriage increases the likelihood of being employed for women in rural Karnataka and Rajasthan
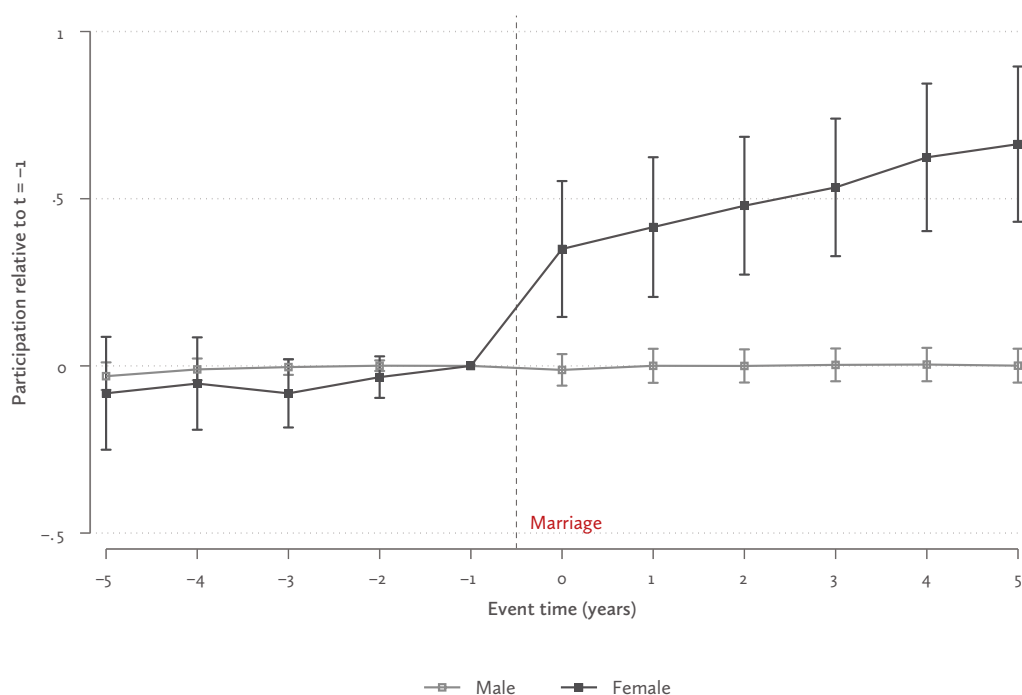
Sources and notes: India Working Survey 2020. These are results from a Life History Calendar exercise. Dashed line indicates time of marriage. 0 indicates the first year of marriage. This is Figure 4.6 in the report.

f.   Are there enough jobs for women workers?

Even if gender norms change and are no longer a barrier for women to undertake paid work, there still remains the question - are there enough jobs? We find evidence that lack of labour demand is also an important factor in explaining low levels of women's work participation. Using Census 2011 data we find that the proportion of women employed outside home is significantly negatively associated with the distance men travel for work. If we take long commuting distances as an indication of lack of local work opportunities, this finding also strengthens the case for a demand-side explanation for low levels of female participation. Combining the Population Census data with the Economic Census we find that women are more likely to work outside the home in those districts where the proportion of large firms (employing more than 10 workers) is higher.

g. Measurement of employment in surveys can be improved by asking detailed questions and relying on self-reports rather than proxies.

Could the low rates of female workforce participation be a result of measurement error? The India Working Survey shows that measurement of women's WPR is affected by how questions are asked and to whom they are asked. We find that "calling out" activities (e.g. did you engage in any own-account work?, did you do salaried work?, etc) is better than asking individuals one overarching question asking if they did paid work. We also find that asking women directly about their employment status increases female WPR by over 5 percentage points. Thus, while measurement problems cannot explain the falling trend of female WPR, they can provide part of the explanation of low levels.

Differences between self and proxy reported employment status

	Women			**Men**		
	Self	Proxy	Proxy-Self	Self	Proxy	Proxy-Self
Labour force participation rate	69.5	63.8	-5.7***	79.7	81.5	1.8
Workforce participation rate	63.2	57.9	-5.4***	76.9	78.7	1.8
Unemployment rate	9	6	-3	3.5	2.8	-0.7

*Sources and notes - India Working Survey 2020, *** - p<0.01. This is Table 7.2 in the report.*

5. Raw gender earnings gaps have narrowed over time, but remain much higher than caste or religion-based gaps.

a. The disparity in earnings is widest in self-employment and for SC/ST women workers.

Comparing caste, gender and religion-based earnings gaps for regular wage workers
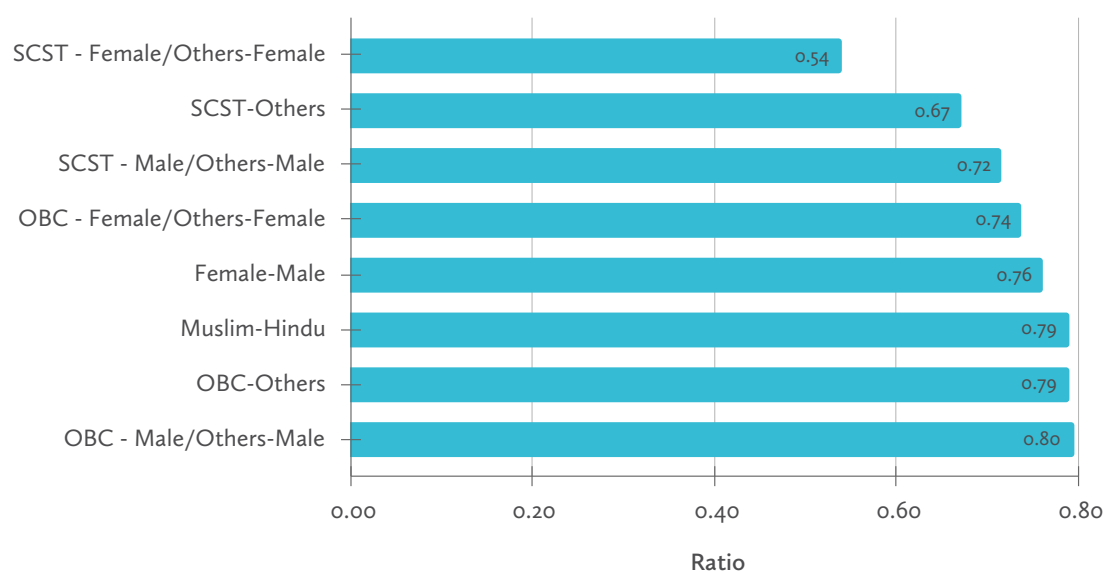
Sources and notes: PLFS 2021-22. This is Figure 6.3 in the report.

On average women earn 76% of what men earn in salaried work, which drops to only 40% for the self-employed. The salaried earnings gap for caste (SC/ST - Others) is similar at 76%, while the religion gap is narrower. At the intersection of caste and gender, more severe disparities arise. Women SC/ST workers earn only 54% of what General caste women earn in salaried work.

b. Raw gap between earnings for male and female salaried workers declined across the distribution since 2004.

The gender earnings gap has decreased over time among regular wage workers
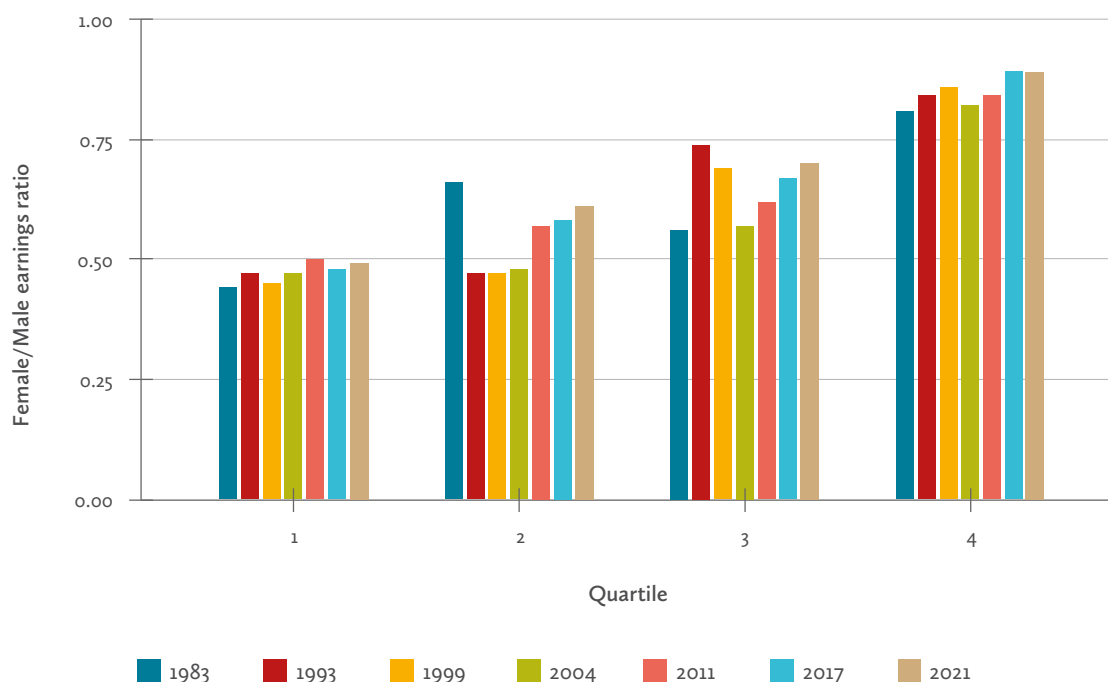
Sources and notes: NSS EUS and PLFS various years. *This is Figure 6.4 in the report.*

The salaried earnings gap varies across the earnings distribution. On average, at the upper end, there is greater parity in wages. Women workers in the top quartile of the salaried earnings distribution earn 90% of what men earn. In the bottom quartile this drops to 50%. The gender earnings gap has narrowed since 2004, particularly in the top quartile.

c. Controlling for individual and household characteristics as well as industry and occupation, gender penalties remain much higher than caste and religion-based

The raw earnings gaps reported above can be the result of differences in levels of education or experience as well as the type of industry or occupation. Regression analysis allows us to control for these factors and measure the residual or unexplained gap. This penalty is much higher for gender (34.6%) as compared to caste (4%) or religion (6%). This means that the earnings disparities observed for the latter two are almost entirely explained by differences in education, household economic status or industry and occupation of work. But for women, a large part of the disparities is not so explained. We note that while a SC/ST penalty

(unexplained gap) has always been present since the 1980s, a statistically significant earnings penalty for Muslim salaried workers has only been observed since 2017 and only in urban areas (around 4 to 5%).

Blinder–Oaxaca decomposition shows a much higher unexplained gap for gender than caste

	Caste		Gender	
Year	Explained	Unexplained	Explained	Unexplained
1983	69.6	30.4	36	64
1993	88.7	11.3	41.4	58.6
1999	96.6	3.4	41.8	58.2
2004	85.3	14.7	42.3	57.7
2011	82.6	17.4	31.5	68.5
2017	82.3	17.7	16.5	83.5
2021	85.5	14.5	18.9	81.1

Sources and notes: NSS EUS and PLFS various years. The table shows results of a Blinder-Oaxaca decomposition analysis. This is adapted from Table 6.4 of the report.

6. Industrial segregation has declined over time but segregation by occupation and industry remains important in explaining disparities of caste, religion and gender.

a. Caste-based industrial segregation decreased but gender-based segregation increased between 1983 and 2021

We calculate a Representation Index (RI) for each industry by caste and gender. The RI takes the value 1 if the share of a social group in a particular industry is the same as the share of that group in the entire workforce. At the beginning of the analysis period, in the early 1980s there was strong caste-based segregation with SC workers more than 5 times over-represented in waste-related work and over 4 times

Industries that show consistent over-representation for disadvantaged identities
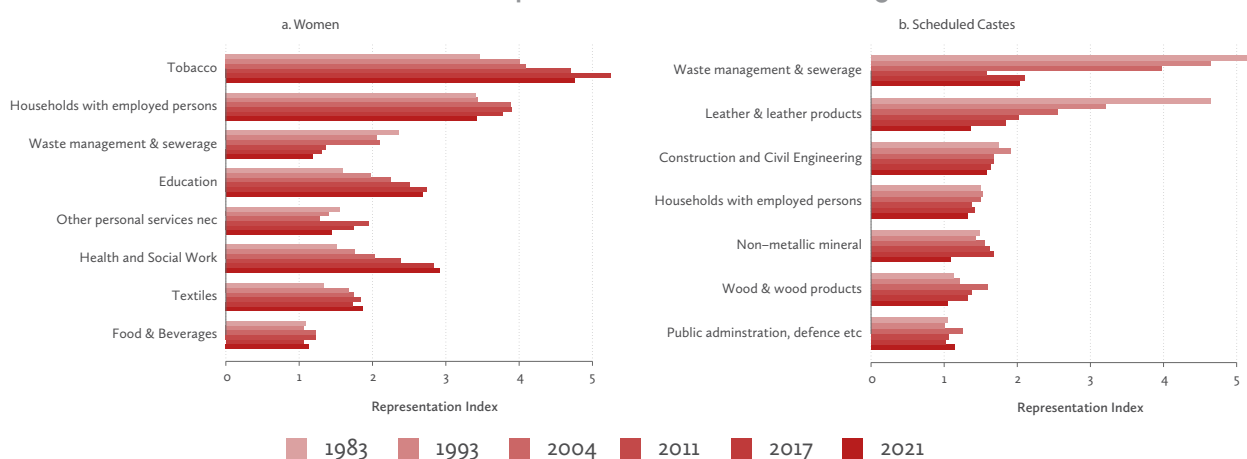
Sources and notes: NSSO EUS and PLFS various rounds. Shown are Representation Indices. This is Figure 6.8 in the report.

in leather-related work. This has declined rapidly over time though they remain over-represented as of 2021-22. In the case of the leather industry, the representation index declined from 4.6 in 1983 to 1.4 in 2021. In case of waste management and sewerage, over-representation of SCs decreased from over 5 times their workforce share in 1983 to 1.6 times in 2011 before increasing slightly. On the other hand, industrial segregation along gender lines has worsened in this period.

b. Marginalised identities continue to be overrepresented in low-paying occupations

The PLFS data for 2021-22 clearly shows a statistically significant inverse relationship between the RI and monthly earnings for various occupations. For example, the lowest paid occupation of personal care workers, with monthly average real earnings of less than ₹10,000, has four times as many women as there are in the workforce as a whole. The same is the case for Scheduled Caste workers. Naturally, the opposite is the case for General caste workers, whose representation grows stronger with increasing earnings.

Women and Scheduled Castes are over-represented in low paying occupations
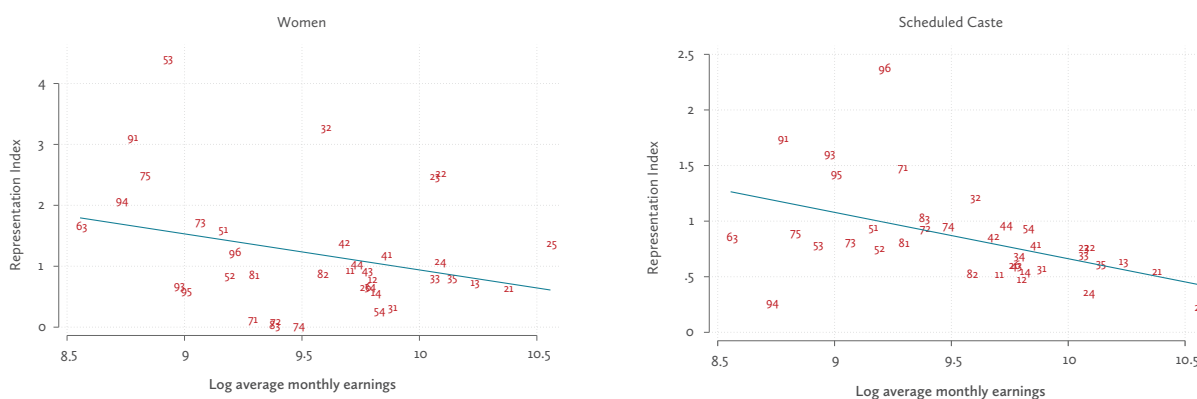
Sources and notes: PLFS 2021-22. The X-axis shows log earnings. Actual earnings range between ₹5000 and ₹36000 per month. These are adapted from Figures 6.9 and 6.11 in the report.

7. Caste-based marginalisation is clearly visible in firm ownership patterns.

We construct a Representation Index for various caste groups based on firm ownership data from the 6th Economic Census data (2013). We find that even in the smallest firm sizes, SC and ST workers are under-represented compared to their share in the overall workforce. But even more significantly, the RI decreases steadily as firm size increases implying that in larger firms, SC/ST owners are more of a rarity. SC and ST owners are barely represented among firms employing more than 20 workers. Correspondingly, General caste overrepresentation increases with firm size. The OBCs are found between the two extremes. In 2013 (the year of the last Economic Census) we estimate that this under-representation cost SC, ST, OBC owners ₹42,000 crores (in 2013 prices). That is, around one-fourth of all private proprietary GVA in the manufacturing sector would be under the control of marginalised social groups if there was no bias in caste representation in ownership of enterprises.

Privileged castes are over-represented to a greater degree in larger enterprises
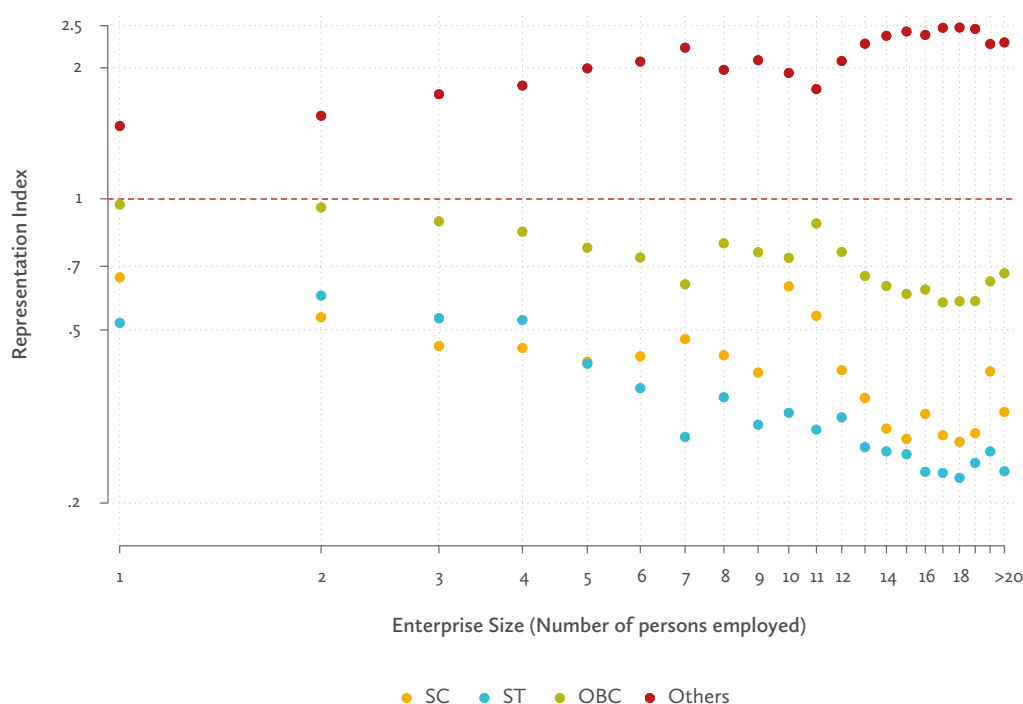
Sources and notes - 6th Economic Census, 2013. This is Figure 6.13 in the report.

8. The last few years saw the creation of more formal salaried jobs than earlier. But women were compelled to enter self-employment due to distress caused by the growth slowdown and the pandemic.

Between 2017 and 2021, there was a slowdown in overall regular wage job creation but formal jobs (with a written contract and benefits) as a share of all regular wage work rose from 25% to 35%. In 2020-21 (pandemic year) regular wage employment fell by 2.2 million. But this net change hides an increase in formal employment by 3 million and a loss of about 5.2 million of semi and informal regular wage employment. While half of the lost employment is accounted for by women, only a third of the increase in formal employment accrued to women. So in net terms, women lost out on formal employment in this period. Not only that, there was a shift towards self-employment due to distress.

a. Workforce participation rate for women rose during the growth slowdown but most of the increase was in self-employment

After falling or being stagnant since 2004, female employment rates have risen since 2019. Why did employment rates rise at a time when growth was slowing down? The explanation is that it was mainly self-employment that rose, led by distress. Compared to the April-June 2018 quarter, salaried employment for women was down a cumulative of 8 pct pts as of the most recent available data (Oct-Dec 2022) while

self-employment was up 14 pct pts. The figure shows this cumulative quarter-by-quarter change compared to the base (April-June 2018) quarter. As a result earnings from self-employment declined in real terms over this period. Even two years after the 2020 lockdown (in April-June 2022), self-employment earnings were only 85% of what they were in the April-June 2019 quarter.

This is a cautionary note against placing too much emphasis merely on rising work participation rates for women. If participation rises due to economic growth and rising labour demand, this has very different implications than if it rises due to falling household incomes which force women into self-employment.

Women largely entered self-employment and moved away from wage work, reducing earnings from self-employment
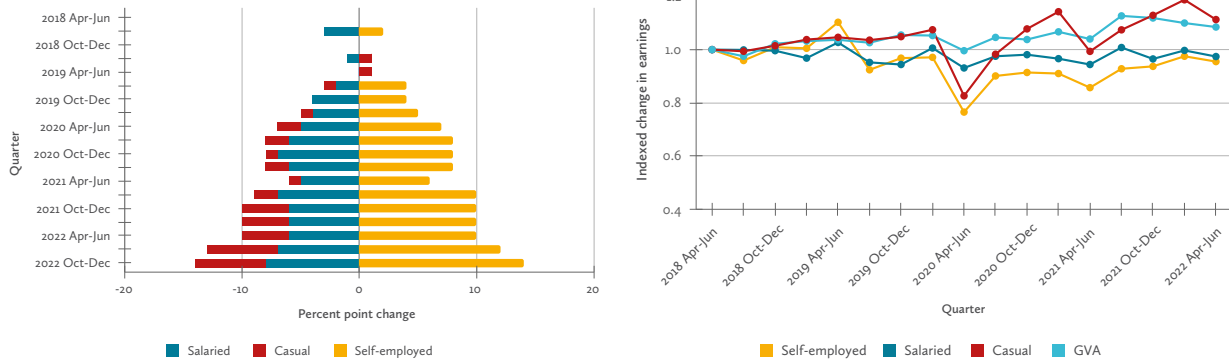
Sources and notes: PLFS, various rounds. *These are Figures 2.6 and 2.8 in the report*.

b.   Covid-induced reversal of structural change – women impacted longer

The loss of jobs during the pandemic forced workers to fall back on agriculture or on self-employment in order to survive. As a result the share of employment accounted for by these two sectors rose sharply in the lockdown quarter (April-June 2020). Two years later, it had declined to pre-pandemic levels for men but continued to remain elevated for women.

Quarter	Agriculture share (%)		Self-employment share (%)	
	Male	Female	Male	Female
April to June 2018	36.8	47.9	51.9	50.6
April to June 2019	33.9	44.2	50.5	50.2
April to June 2020	42.3	57.7	58.2	57.8
April to June 2021	38.5	58.2	55.2	56.7
April to June 2022	34.3	55	52.1	61

Sources and notes: PLFS, various rounds. *This is Table 2.3 in the report.*

The foregoing highlights some important takeaways from this year's report. We hope that the analysis and findings presented in the report will be of use to researchers, policy-makers, students, journalists and the general public.

1

Introduction

1

Introduction

In 75 years since its independence, India has made significant progress in almost every socio-economic indicator. The task that lies ahead is to build on this progress and bring about a society that is not only prosperous but also just, equitable and sustainable. The creation of an adequate number of productive and decent employment opportunities plays a key role in achieving this goal. Unfortunately, despite years of rapid economic growth, India's structural transformation remains slow. The lack of adequate good jobs is a key sign of this. The problems of insufficient employment generation and persistence of informality are compounded by the fact that one's social identity continues to play a crucial role in determining one's opportunities. This year's State of Working India (SWI) report takes a long-run view of India's structural transformation experience and its implications for three key social identities: caste, gender, and religion.

Our three previous reports, SWI 2018, 2019, and 2021 have provided a detailed analysis of a variety of aspects of the employment and labour situation in India. SWI 2018, our first report, took stock of India's changing sectoral structure, sources of job creation, quality of jobs created as well as gender and caste-based disparities in labour market outcomes. It thus set the stage for more detailed explorations to follow. SWI 2019 focused on a few key policy ideas to promote employment generation, including proposals for an urban employment guarantee programme and a universal basic services programme. SWI 2021 was produced in the middle of the Covid-19 pandemic and focused exclusively on the impact of the pandemic. This year's report builds on all the previous reports, but especially on SWI 2018, by examining the bi-directional relationship between structural change and social identities.

1.1 Social identities and structural change

This report takes a long view of the relationship between India's structural transformation and key social identities. Though we focus mainly on caste, gender and religion, we also realise that there are other crucial identities that matter, including ethnicity, race, linguistic background, sexual orientation and others. Our focus is limited to the three mentioned in part due to data limitations and in part to make the scope of the report manageable.

Previous reports have elaborated on the theme of structural transformation. Briefly, as economies grow, their sectoral structure changes. The importance of agriculture as a source of livelihood declines and that of manufacturing and services grows. Alongside, the importance of traditional or home-based economic activity declines and the capitalist sector grows in size. We refer to these processes as the Kuznets Process and the Lewis Process respectively, after two economists, Simon Kuznets and Arthur Lewis who wrote extensively on the phenomenon of structural change.

The relationship between growth and structural change is not automatic. That is, an economy may experience several years of economic growth without much structural change. India presents one such case. Further, the relationship between structural change and social disparities is also not straightforward. Structural transformation or the emergence of the modern capitalist economy creates new opportunities. But who will be able to take advantage of these new opportunities depends on existing social relations, norms and

power hierarchies. Conversely, these very norms and hierarchies play a role in determining the nature of the structural transformation being experienced by a country. There is thus a dual relationship at play.

To take an example, patriarchal norms may prevent women from taking advantage of new jobs being created, by preventing them from leaving the home for employment. But equally, the presence of gender norms that restrict women's mobility changes the supply of labour to the modern sector and has implications for the relative price of labour and capital, and in turn for the type of industries that grow. Similarly, caste-based norms prevent marginalised castes from accessing good quality education and keep them tied to traditional livelihoods. New opportunities are thus availed of by dominant castes. But conversely, the existence of caste-based norms that prevent dominant castes from taking up work associated with marginalised castes, particularly manual work, can create lopsided demand for office or service sector work and very few takers for manufacturing jobs. These two examples illustrate that norms not only mediate the relationship between structural change and labour market outcomes, they also play a role in the type of transformation that will take place. Finally, economic growth and structural change in turn change norms by loosening traditional ties and networks, completing the circle.

It is not possible to investigate all the above interconnections in one report. We focus mainly on the first part, how effective structural change has been in reducing existing social disparities based on gender, caste and religion. Occasionally, we discuss the other aspects of the relationship outlined above.

1.2 Data sources, time periods and methods

This report uses a wide range of datasets at the state, national and international levels. These are described in the relevant sections of the report, but we present a list here for quick reference. The analysis commences in the 1980s in most cases. The most recent data used is the Periodic Labour Force Survey data from the Oct-Dec 2022 quarter. So the first few months of 2023 are not part of our analysis. A note on survey nomenclature - we generally refer to an NSO survey either by the two calendar years over which it takes place (so PLFS 2021-22) or by the first calendar year (so PLFS 2021).

The following data sources have been used in this report:
- Surveys conducted by the National Statistical Organisation (previously the National Sample Survey Organisation or NSSO)
  - Employment-Unemployment Surveys, quinquennial - 1983-84, 1987-88, 1993-94, 1999-2000, 2004-05, 2011-12
  - Periodic Labour Force Surveys, quarterly and annual - 2017-18, 2018-19, 2019-20, 2020-21, 2021-22
  - Unincorporated Enterprises Surveys, 2005-06, 2011-12, 2015-16
  - Annual Survey of Industries, 2004-05, 2011-12, 2015-16
- Census of India - 1991, 2001, 2011
- Economic Census - 1998, 2005, 2013
- National Family Health Survey - 2004-05, 2015-16, 2019-21
- India Working Survey -2020
- World Development Indicators - 1990 -2019
- Groningen Economic Transformations Database - 1990 - 2019

This report takes advantage of a wider range of statistical techniques than has been the practice in the past. We have made extensive use of the technique of regression analysis. For researchers and interested readers the details of the regression models as well as the full results are available in the Appendices. The Methods Appendix is available at the end of this report. The Results Appendix is available online.

For our analysis of caste disparities, we are mostly forced to rely on large administrative categories that are usually reported in surveys - Scheduled Caste, Scheduled Tribe, Other Backward Classes and Others (residual). We realise that many significant social and economic outcomes vary across jatis (sub-castes) and that the administrative categories are too broad to detect these changes. We make some efforts to analyse jati-level outcomes in a few places drawing on Population Census data (only for Scheduled Castes). We are also limited in being able to analyse caste disparities within non-Hindu castes, once again due to data limitations. Here too, we recognise the importance of caste among Muslim and other minority religions and make some attempts to examine intersectionalities of caste and religion.

Though the major portion of the report relies on secondary datasets and on quantitative analysis, we bring in qualitative data in the form of in-depth interviews in a few places.

1.3 Organisation of the report

The report consists of five data chapters, a chapter discussing measurement and data issues, and a concluding chapter. In **Chapter Two** we start with a discussion of the recent labour market trends, covering the pre-pandemic growth slowdown and the pandemic with its aftermath. **Chapter Three** starts the long-run analysis by first examining the relationship between growth and structural change at the national and state level since the 1980s. We examine the pace of both the Lewis and the Kuznets processes, overall and for key identities. **Chapter Four** focuses on women's employment. We approach the issue from both a labour supply and a labour demand angle. From the supply side we investigate the role of gender norms in allowing women to access paid work, especially outside the home. On the demand side, the data is more limited, but we show that controlling for supply

side factors such as norms and education, labour demand (availability of work opportunities) plays an important role in explaining whether women do paid work or not. **Chapter Five** shows that gender, caste and religion-based identities continue to be strongly correlated with the type of employment (causal, salaried, self etc) as well as the quality of jobs (informal, semi-formal, formal). It also shows that intergenerational mobility has increased over time in India. **Chapter Six** examines disparities in earnings and measures the changing occupational and industrial segregation over time. It also looks at the relationship between caste and entrepreneurship. **Chapter Seven** discusses several issues related to measurement of employment for women as well as the necessity for data going down to the jati level and data on religion-caste intersections. **Chapter Eight** concludes.

Finally a word on the academic and policy literature in which this report is situated. There is a large amount of writing - books, articles, policy papers - on questions of gender, caste, religion and labour. It is not our intention here to review this vast literature. Where appropriate we have pointed readers to further readings on various topics, but the resulting bibliography has no claims to being exhaustive.

The analysis presented here also leaves many important and interesting questions unanswered. We have no doubt that future research will take on these questions and produce knowledge that is of practical use for the design of better policy.

2

Rising employment, stagnant earnings: Recent trends in the Indian labour market

2

Rising employment, stagnant earnings: Recent trends in the Indian labour market

This chapter analyses recent trends in the Indian labour market. The past few years have been very eventful ones for the economy. Even prior to the unprecedented shock delivered by the Covid-19 pandemic, the Indian economy was experiencing its most prolonged growth slowdown in recent decades. Economic growth slowed for 10 quarters between 2017-18 and 2019-20. Our previous report (State of Working India 2021) began with a brief analysis of this slowdown and we do not go into its details here. The slowdown, severe though it was, was completely overshadowed by the impact of the 2020 nationwide lockdown during the pandemic. In the lockdown quarter (April-June 2020) the Indian economy shrank by 21 percent. A detailed analysis of the labour market impacts of the lockdown and subsequent year of the pandemic is available in State of Working India 2021. However, the effects of the pre-Covid slowdown and the pandemic are still with us. In this chapter we document the cumulative effect on the labour market, particularly focusing on women workers. We also take stock of employment, unemployment and earnings at the national and state-level as of 2022.

2.1 Rising employment rates and stagnant labour incomes

The regular availability of data from the Periodic Labour Force Survey (PLFS) since 2017-18 has made it possible to keep track of labour market indicators on an annual as well as a quarterly basis. We start with the broad indicators, viz. the working age population (WAP), the labour force (LF), the workforce (WF) and the level of unemployment (Table 2.1).[1] Three things are immediately obvious. The workforce has grown in size, participation rates have risen, and unemployment has fallen. As of 2021-22, the rate of open unemployment had dropped from a high of 8.7 percent in 2017-18 to 6.6 percent. The fall is registered both in rural and urban areas among men and women (Table 2.2).

The rise in the female Labour Force Participation Rate (LFPR) and Workforce Participation Rate (WPR) is especially interesting, as it comes against the backdrop of a steady fall in these rates since 2004 which have been the subject of much discussion and debate (Figure 2.1). But it is worth noting that the rise in workforce size between 2017-18 and 2018-19 was more or less at par with the rise in the working age population, thereby keeping the female WPR (ratio of female workforce to female working age population) constant. However, in the subsequent year, the rise in employment outstripped the rise in the working age population, pushing the WPR up significantly from around 20 percent to 26 percent in rural areas and 17 percent to 19 percent in urban areas (Table 2.2). With the exception of the lockdown quarters (April-June 2020 and April - June 2021) the rise has been sustained all the way through the pandemic and until the most recent period for which numbers are available. Further, it is worth

Table 2.1: Key labour market indicators for the past five years

Indicator/Year	2017-18	2018-19	2019-20	2020-21	2021-22
Absolute numbers (millions)					
Working age population	968	985	1000	1022	1021
Labour force	468	478	512	525	528
Workforce	427	436	467	486	493
Unemployed	41	42	45	39	35
Out of labour force	500	508	488	489	493
Rates (%)					
Labour force participation rate	48.4	48.5	51.2	51.4	51.7
Workforce participation rate	44.1	44.3	46.7	47.5	48.3
Unemployment rate	8.7	8.7	8.8	7.5	6.6
Out of labour force	51.6	51.5	48.8	47.9	48.3
Monthly earnings (₹)					
Self-employment	12,318	12,988	11,560	11,411	12,089
Casual wage	6,959	7,209	7,324	7,431	7,856
Regular wage	19,450	19,690	18,907	19,074	19,456

Sources and notes: Periodic Labour Force Survey (PLFS), various rounds. Employment is defined as per Current Weekly Status (CWS). Absolute numbers are calculated by multiplying survey ratios with Census population projections. See Methods Appendix for details. Rupee values are in 2022 rupees.

noting that the rise is limited to older women (25 years and above). Among younger women, it is the rate of educational enrolment that has increased (data not shown).

Ordinarily, rising employment rates and falling unemployment rates would be positive signs indicating an improving labour market. However, because this period has been characterised by negative shocks to economic growth, the numbers need to be interpreted with caution. One indicator suggesting that all is not well in the labour market, is the trend in labour earnings. Table 2.1 shows that earnings from self-employment have declined in real terms over this period while those for regular wage or salaried workers have been largely stagnant.

Earnings of casual wage workers showed a small increase in real terms. Obviously, the Covid-19 pandemic is part of the explanation. Indeed, the earnings series clearly shows the Covid effect for the year 2020-21. But, as we show later in the chapter, self-employment earnings, particularly for women, were declining even prior to the pandemic. The important point to note is that stagnant or very weakly growing real labour earnings indicate a continued softness in the labour market post-Covid.

Stagnant or very weakly growing real labour earnings indicate a continued softness in the labour market post-Covid.

Table 2.2: Key indicators by region and gender

Indicator/Year	2017-18	2018-19	2019-20	2020-21	2021-22
LFPR (%)					
Male Rural	75.6	75.5	76.7	75.9	76.7
Male Urban	74.1	73.7	73.8	73.4	74.2
Female Rural	21.7	22.5	28.3	29.8	29.2
Female Urban	19.6	19.7	22.1	21.5	22.1
WPR (%)					
Male Rural	69.1	68.9	70	70.5	71.7
Male Urban	67.7	67.2	66	66.5	68.4
Female Rural	20.1	20.9	26.7	28.4	27.9
Female Urban	17.1	17.4	19.4	18.9	19.9
UR (%)					
Male Rural	8.7	8.7	8.7	7.2	6.5
Male Urban	8.7	8.8	10.5	9.4	7.8
Female Rural	7.5	7.3	5.5	4.8	4.5
Female Urban	12.7	12.1	12.4	12.2	9.9

Sources and notes: PLFS, various rounds. Employment is defined as per Current Weekly Status (CWS). LFPR -Labour Force Participation Rate, WPR - Workforce Participation Rate, UR - Unemployment Rate.

Figure 2.1:   Workforce participation rate over the long run for men and women in rural and urban areas
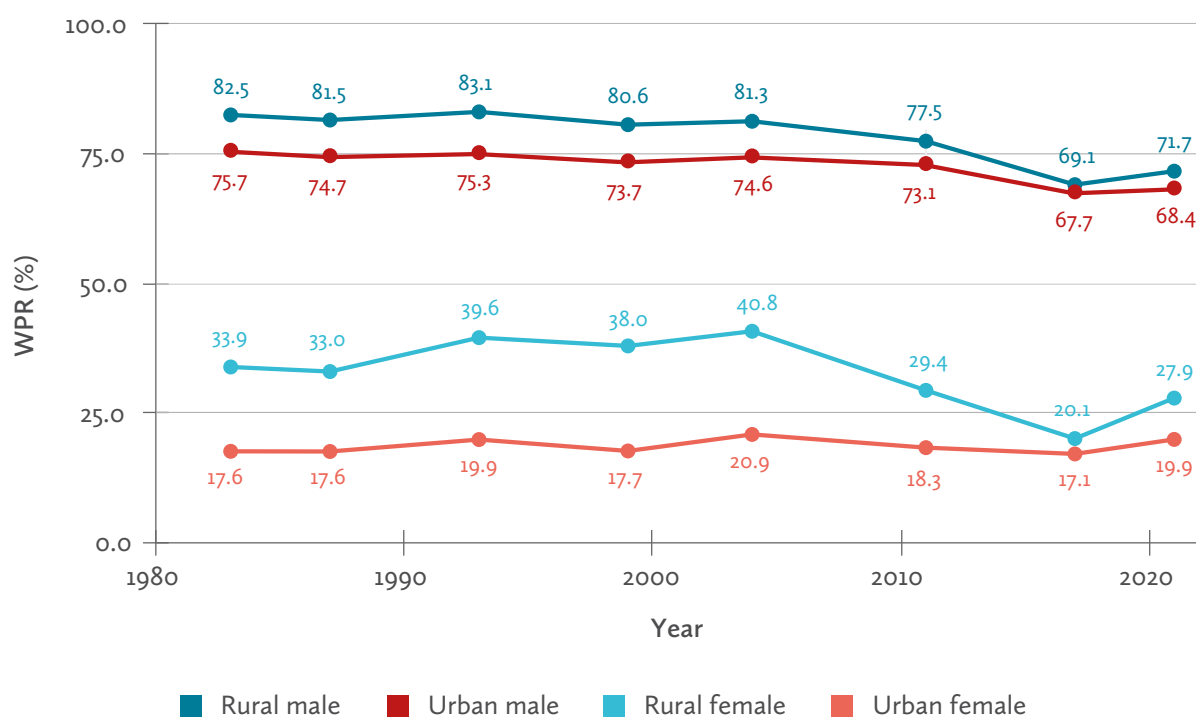
Sources and notes: NSSO EUS and PLFS, various years. Employment is defined based on Current Weekly Status.

Figure 2.2: Quarterly rate of unemployment by level of education
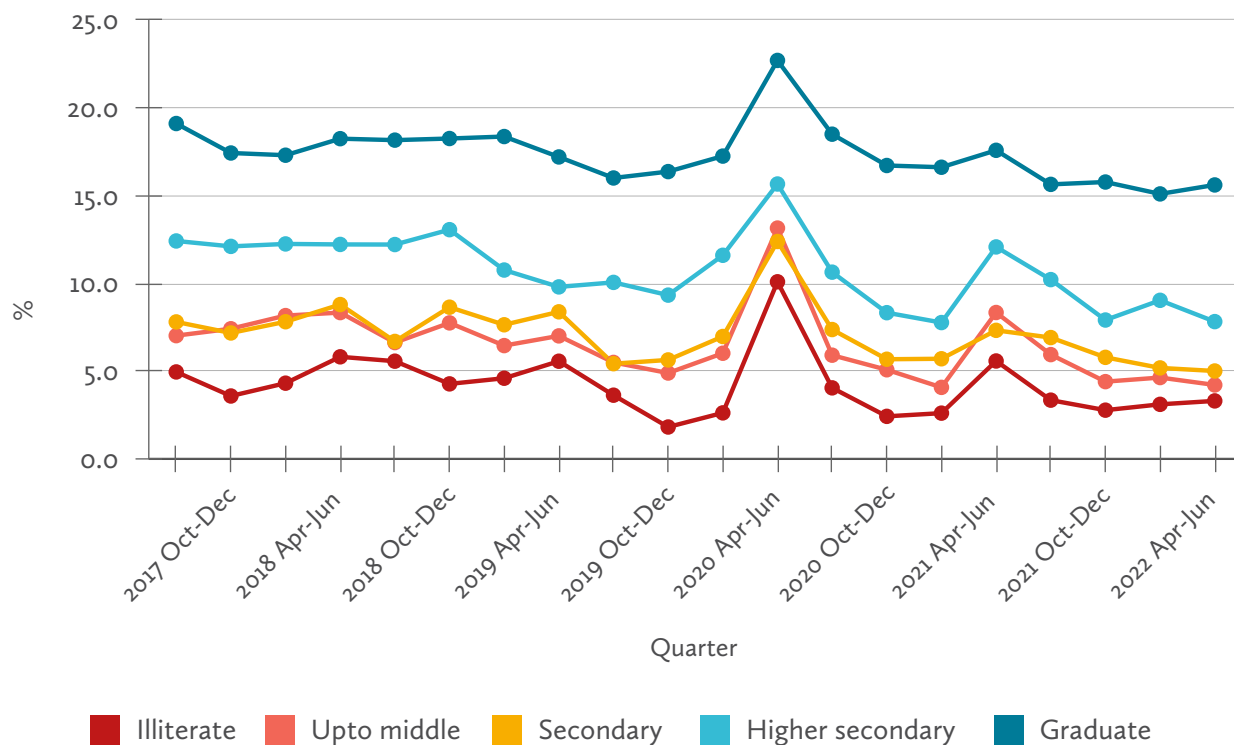
Sources and notes: PLFS, various quarters. Unemployment is defined based on Current Weekly Status.

Figure 2.3: Unemployment is concentrated among educated youth

	Less than 25 years	25–29 years	30–34 years	35–39 years	40 years and above
Graduate & above	42.3	22.8	9.8	4.5	1.6
Higher secondary	21.4	10.6	5.0	3.1	2.1
Secondary	18.1	7.5	4.6	2.4	1.7
Primary or middle	15.0	5.4	3.0	2.4	2.2
Literate but below primary	10.6	3.3	1.5	2.4	2.2
Illiterate	13.5	4.3	4.0	3.4	2.4

Sources and notes: PLFS 2021-22. Employment is defined based on Current Weekly Status. Numbers in the boxes are unemployment rates (%). Colour shading is for ease of interpretation.

Lastly, let us cast a glance at one of India's long running problems - unemployment among educated youth. As is well-known, the aggregate unemployment rate hides substantial variation by age and level of education. Figure 2.2 shows the quarterly rate of unemployment by level of education.[2] It is encouraging to note that unemployment is lower for every category post-Covid. But two caveats are necessary. First, it remains above 15 percent for graduates as of the most recent available data. And more worryingly it touches a huge 42 percent for young graduates (Figure 2.3). At the other extreme, among older, less educated workers, it is in the range of 2-3 percent (Figure 2.3). The striking thing is the large variation in the rate of unemployment even within the higher educated group across age. The unemployment rate falls from over 40 percent for educated youth under 25 years of age to less than 5 percent for graduates who are 35 years and above. This indicates that, on average, eventually graduates do find jobs, by their late 20s or early 30s, but the key questions are, what is the nature of jobs they find and whether or not these match their skills and aspirations. More research is needed on this important topic.

The second caveat to keep in mind is that low unemployment is matched by stagnant earnings indicating that labour demand continues to be weak and that the effects of continued weak aggregate demand in the economy are being passed on to workers. Indeed, as per the PLFS, total household income from labour only grew at a compounded annual rate of 1.7 percent in real terms since 2017-18.

2.2 Understanding the recent rise in the female employment rate

We now turn to the rise in the female employment rate. As noted earlier, for more than a year prior to the rise, the Indian economy had been slowing down. The quarter-on-quarter rate of GDP growth

fell from 7.1 percent in the April-June quarter of 2018 to 4.2 percent in the July-September quarter of 2019. At first glance it may seem paradoxical that employment rates should rise at a time when growth is slowing down. The explanation lies in the type of employment that was being generated. As Dhamija and Chawla (2023) report, it was mainly self-employment (and, within that, unpaid work) that rose in this period. Let us take a look at this phenomenon.

Why did female employment rates rise at a time when growth is slowing down? The explanation is that it was mainly self-employment that rose, led by distress.

First, taking advantage of the quarterly frequency of PLFS data, we present an analysis of the employment numbers from the Jul-Sep 2017 quarter (the first PLFS quarter) to Oct-Dec 2022 (last quarter for which PLFS data are available). The higher frequency of analysis enables us to locate the rise in female WPR more precisely in time as well as to separate the direct effect of the Covid-19 lockdowns (in April-June 2020 and April-June 2021) from the rest of the year. As expected, there is a seasonal pattern to female workforce participation in rural areas. Women tend to enter the workforce during the peak seasons and withdraw during the lean seasons. But in addition, Figure 2.4 clearly shows the uptick in female WPR in rural areas in the 2019 July-September quarter. The rise in the urban female WPR is less dramatic but also present.

When we examine the quarterly changes in the employment status of women, we see a clear pattern of a shift from engagement in domestic duties to participation in the workforce. Figure 2.5 shows the cumulative percentage point change in the share of women (rural and urban combined) reporting various activities in each quarter compared to the base quarter of April-June 2018. The base quarter remains the same for all subsequent quarterly graphs in this chapter and is chosen to

Figure 2.4:   Workforce participation rate for women rose during the growth slowdown
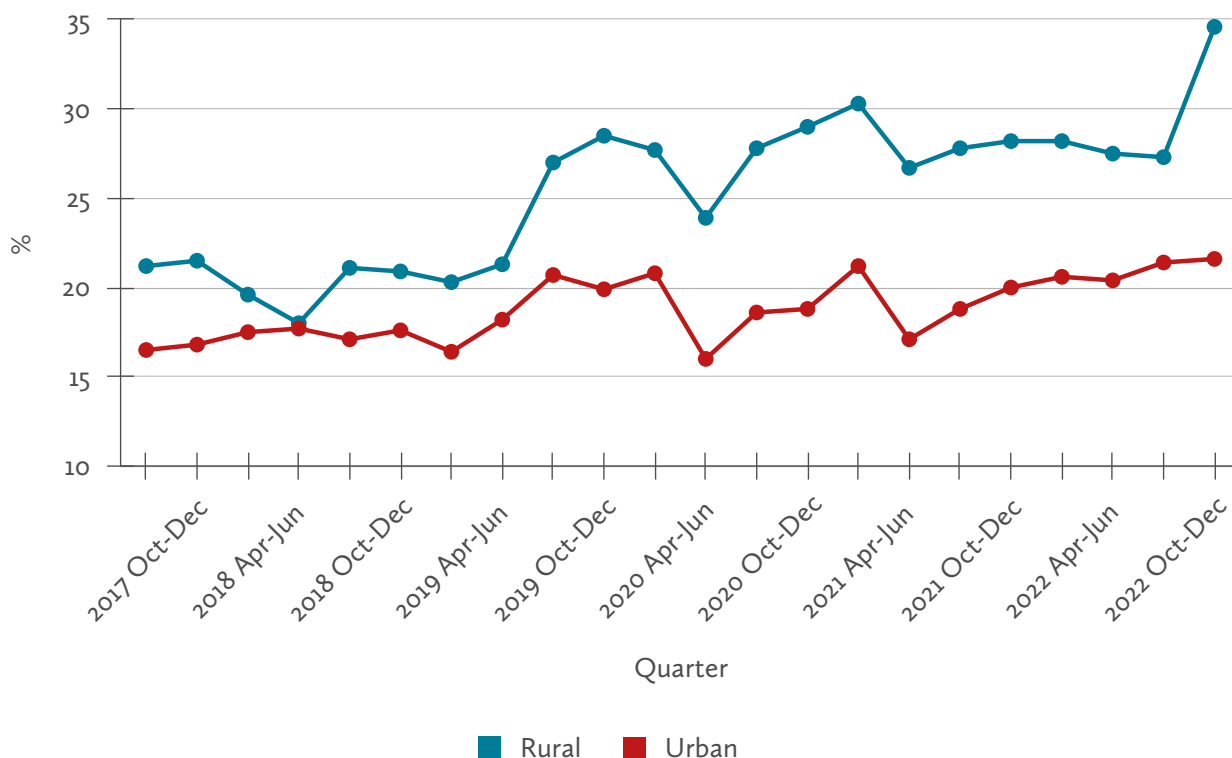
Sources and notes: PLFS various years. Employment is defined based on Current Weekly Status.

Figure 2.5:   Shift from domestic duties to workforce participation for women
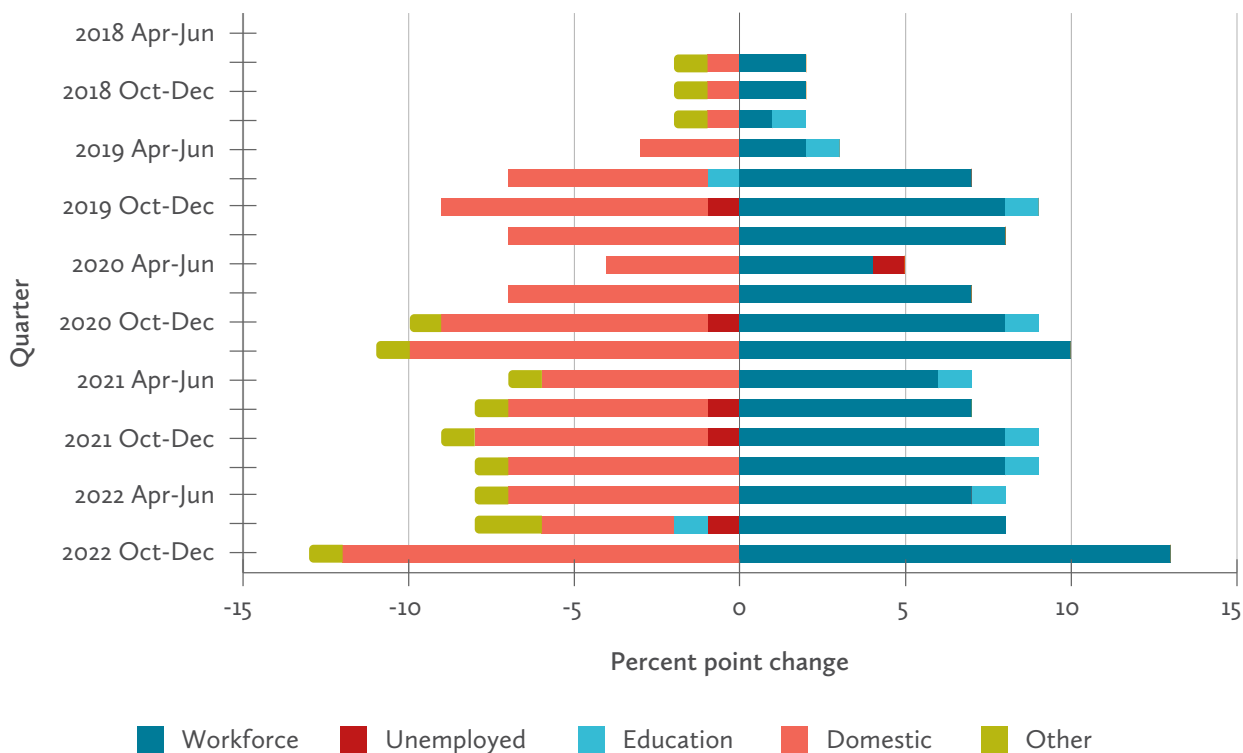
Sources and notes: PLFS various years. Employment is defined based on Current Weekly Status. Each bar indicates the cumulative percent point change as compared to the base quarter of April-June 2018.

provide a direct comparison to the Covid lockdown quarter (April-June 2020).[3] In the graph, numbers less than zero indicate a fall in the share of a particular activity while positive numbers indicate a rise.

Notice that the share of women in the workforce jumps suddenly in the 2019 July-September quarter. This jump comes three quarters prior to the onset of the pandemic and a few quarters after the start of the growth slowdown in 2018. It is important to emphasise that a shift in current weekly status from domestic duties to employment need not mean a reduction in hours spent in care work and other domestic activities. In the India Working Survey we find that increased workforce participation may be accompanied by constant or even increased time spent in housework for women. This is discussed later in the report.

What did these women who entered the workforce do? A similar quarter-wise analysis of the type of employment is shown in Figure 2.6. Note the steady

rise in the share of self-employed women. This is accompanied by a fall in the share of both casual and regular wage work. No such clear pattern is seen for male workers (data not shown). And, as noted by Dhamija and Chawla (2023), even within the category of self-employment, it is the share of unpaid workers (also known as contributing family workers) who have grown faster. Figure 2.7 provides a different way of appreciating the large shift in the structure of the female workforce that has occurred in this period. It disaggregates self-employment into own-account work and unpaid work in family enterprises and shows the growth in various types of employment (regular, casual, own-account, and unpaid) indexed to the April-June 2018 quarter. Thus, for example, a value of 1.5 indicates a 50 percent increase in the number of women reporting a certain type of employment. Also shown here is the change in the total female workforce. The fact that own-account work and unpaid work have consistently grown faster than the total workforce indicates a significant shift in the structure of employment for

Figure 2.6:   Women largely entered self–employment and moved away from wage work
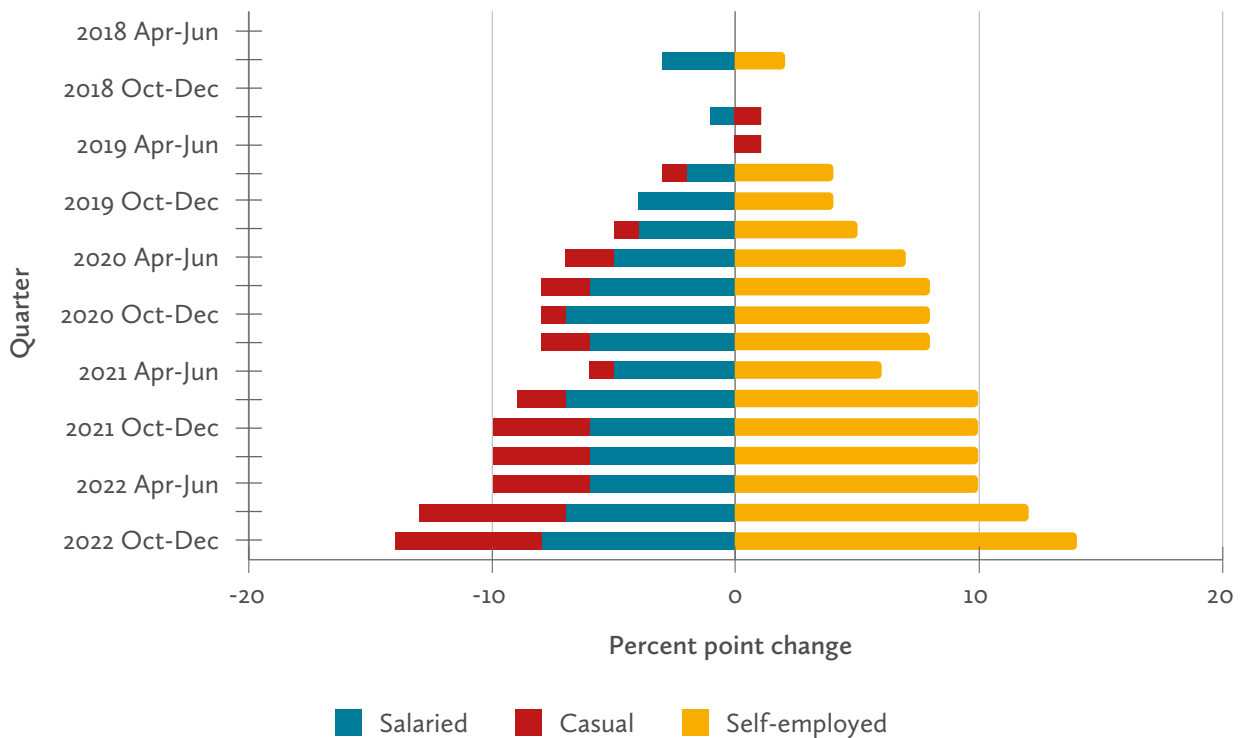
Sources and notes: PLFS various years. Employment is defined based on Current Weekly Status. Each bar indicates the cumulative percent point change as compared to the base quarter of April-June 2018.

Figure 2.7:   Rapid growth of self–employment compared to wage work resulted in a change in structure of women's employment
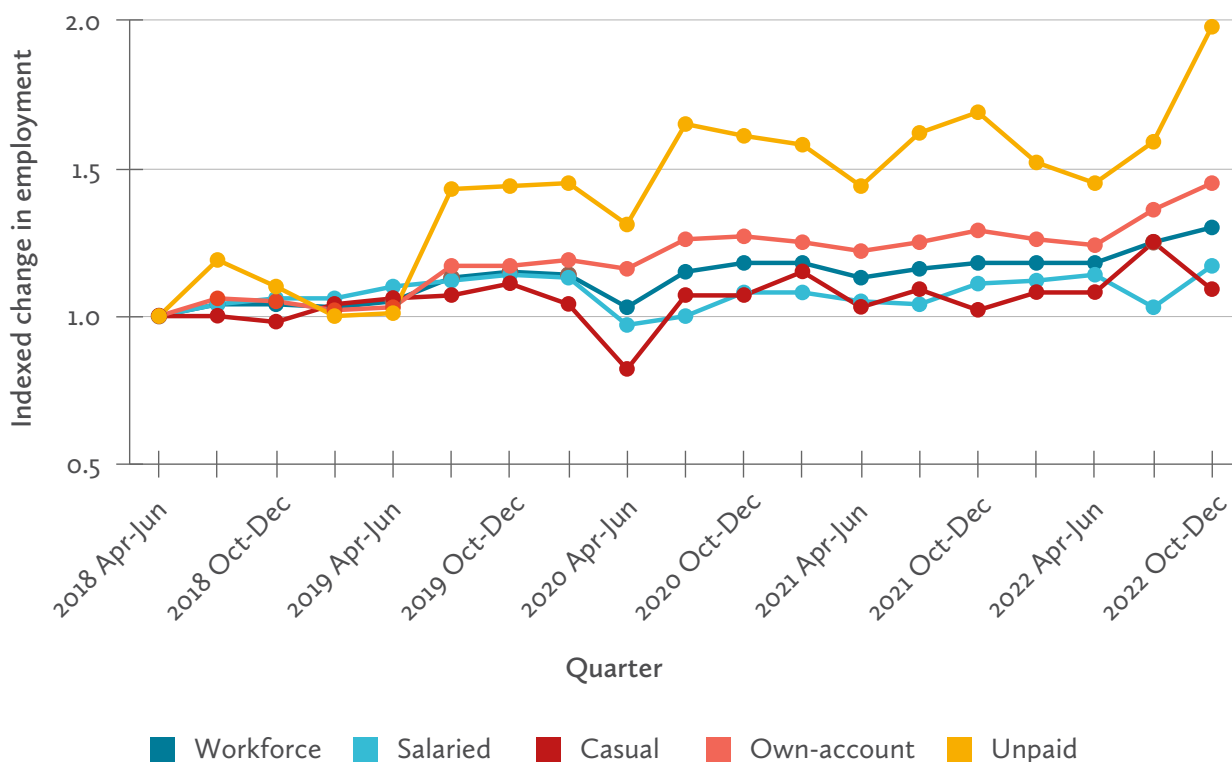
Sources and notes: PLFS various years. Employment is defined based on Current Weekly Status. Values are indexed to the base quarter of April-June 2018.

women.

On top of the increase in self-employment that was occurring prior to the pandemic came the pandemic-induced structural changes to the economy. Table 2.3 shows the share of male and female workers who report being self-employed for each April-June quarter from 2018 to 2022. Also shown are the

shares of the male and female workforce engaged in agriculture. Note that the April-June quarter is one in which agricultural employment tends to be the lowest in the year. For men, the Covid-induced increase in agricultural employment had corrected itself by April-June 2022 while for women it remained elevated well above the pre-Covid level. Similarly, while male and female workers start with roughly

Table 2.3: Covid–induced reversal of structural change

Quarter	Share in agriculture (%)		Share in self employment (%)	
	Male	**Female**	**Male**	**Female**
April to June 2018	36.8	47.9	51.9	50.6
April to June 2019	33.9	44.2	50.5	50.2
April to June 2020	42.3	57.7	58.2	57.8
April to June 2021	38.5	58.2	55.2	56.7
April to June 2022	34.3	55	52.1	61

Sources and notes: PLFS, various rounds. Employment is defined as per Current Weekly Status (CWS). The April-June 2020 quarter was the nationwide lockdown quarter.

the same proportion in self-employment in 2017, this diverges over time. The increased incidence of self-employment caused by the pandemic was back to pre-pandemic levels for men by 2022 but remained elevated for women.

Since both self-employment and agriculture constitute fallback options for workers who have lost work and cannot afford to remain unemployed, we can interpret these numbers as indicating a rise in distress. Taken together, the numbers paint a picture of an economy that was slowing down and was then hit by a massive shock, with the result that there was a distress-led participation of women in paid work as well as a reversal of structural change.

This view is further supported by looking at what has happened to labour earnings over this period. A crowding into the self-employment sector by women who were previously out of the workforce is expected to increase competition among the self-

employed. Given a constant level of demand in the product market, this will lead to lower per person earnings from self-employment. And given that growth was generally slowing down as well as given the pandemic shock, demand in the product market actually collapsed even as supply of labour to the self-employed sector grew. This impacted earnings severely. Figure 2.8 shows that compared to earnings from casual work (which registered an increase in real terms) or salaried or regular wage work (which remained stagnant), earnings from self-employment fell further and remained depressed far longer during this period. For comparison, Figure 2.8 also provides the trend in overall Gross Value Added over the same period.

In sum, this is a cautionary note against placing too much emphasis merely on rising participation rates for women. If work participation rises due to economic growth and rising labour demand, this has very different implications than if it rises due to

Figure 2.8: Earnings from self-employment have declined since 2019

Sources and notes: PLFS various years. Employment is defined based on Current Weekly Status. Values are indexed to the base quarter of April-June 2018.

falling household incomes which force women into self-employment. What we would like to see is an increase in the WPR alongside constant or increasing earnings in the subsistence sector. A key thing to watch out for in the coming quarters is a return to the pre-slowdown sectoral structure of employment for women accompanied by rising earnings in self-employment, signalling a return to a healthier labour market.

2.3 The state-level story

We end this chapter with a more detailed look at the state of the labour market across Indian states in 2021-22, the most recent year for which PLFS data is available. The spatial pattern of the female workforce participation rate across states in urban and rural areas is shown in Figure 2.9. While broadly the female WPR continues to be higher in the southern, western and northeastern Indian states as compared to the northern states, there are significant departures from this known

stylised fact. Focusing on the urban rate, note the relatively high rates in Tamil Nadu, Chhattisgarh and Himachal Pradesh, stretching across the north-south extent of the country. On the other hand, as expected, Rajasthan, Uttarakhand, Uttar Pradesh, Bihar and Jharkhand emerge as the states with the lowest rates of participation for urban women. The state-level variation in the rate of unemployment among educated men and women is shown in Figure 2.10. Unemployment tends to be higher for women compared to men across most states. But note that some of the same states which display low rates of female WPR, also register high rates of female unemployment (e.g. Rajasthan and Bihar). This indicates a generally unfavourable environment for women to be employed in these states.

It is a truism that there is large heterogeneity in almost every indicator across India, at the state as well as lower levels. Hence, problems need to be analysed and their solutions proposed at the appropriate level of aggregation. But where barriers to women working are concerned, despite large

Figure 2.9:   Female workforce participation rate in rural and urban areas across Indian states in 2021–22
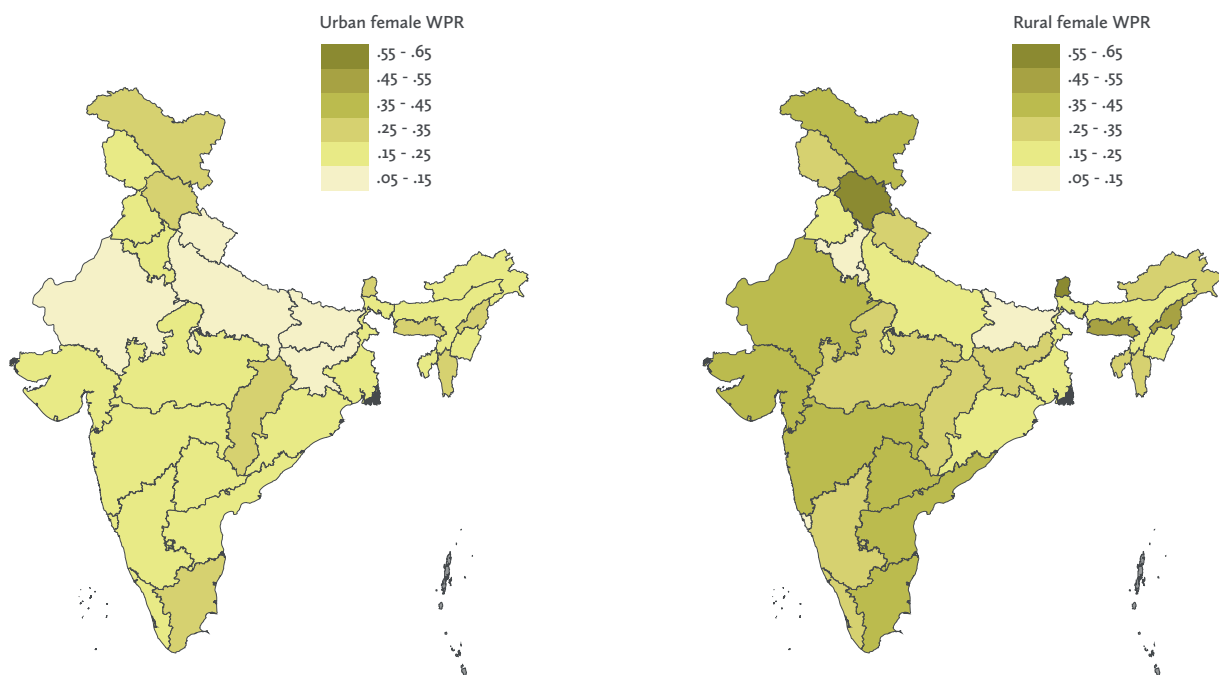
Sources and notes: PLFS 2021-22. Employment is defined based on Current Weekly Status. The scale is the same in both maps to illustrate the overall difference between rural and urban rates.

Figure 2.10: Unemployment rate among educated men and women across Indian states in 2021–22

Male highly educated 15-30 age UR

.55 - .65
.45 - .55
.35 - .45
.25 - .35
.15 - .25
.05 - .15

Female highly educated 15-30 age UR

.55 - .65
.45 - .55
.35 - .45
.25 - .35
.15 - .25
.05 - .15
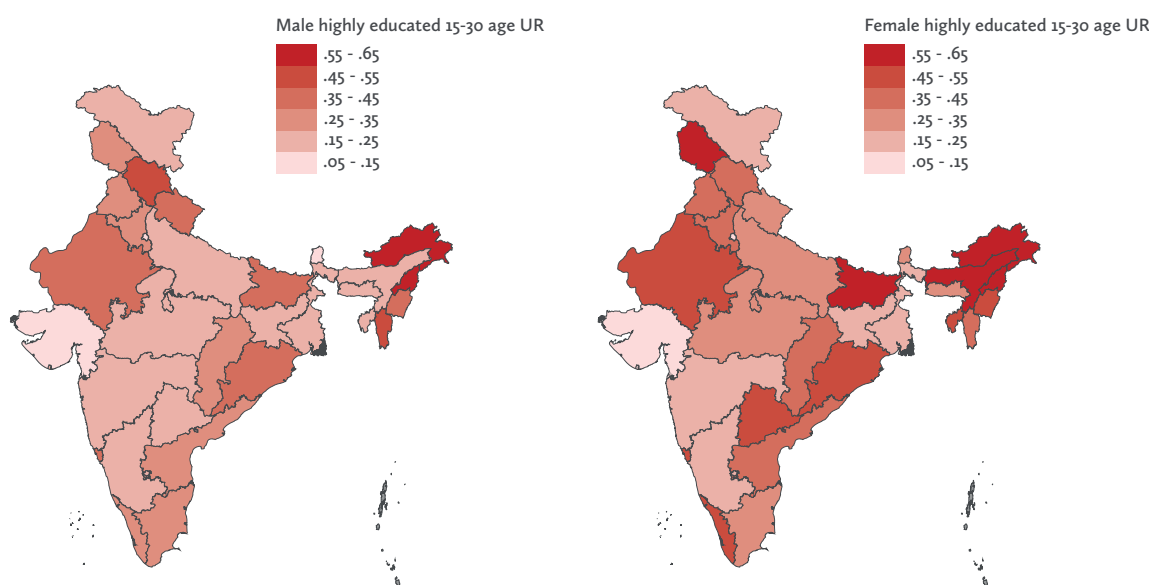
Sources and notes: PLFS 2021-22. UR-Unemployment is defined based on Current Weekly Status. The scale is the same in both maps to make for easier comparison across gender.

cross-state variation there is an overall "India effect". This becomes clear when we place the state-level variation in a cross-country perspective. Figure 2.11 plots the log GDP per capita in 2018 US dollars against the female LFPR for all Indian states (red) and all the countries in the World Development Indicators database (blue). There is a statistically significant U-shaped relationship between GDP and the female LFPR. That is, female LFPR tends to be higher in poorer and richer countries compared to those in the middle. We discuss the reasons behind this in Chapter Four. However, there is also a huge amount of dispersion, especially for poorer countries indicating that at lower levels of GDP, factors other than income are important in determining the level of women's participation in the labour force.[4]

With only three exceptions, all other states have lower rates of female LFPR than is predicted for their level of per capita SDP.

Significantly, for our purposes here, all the states with the exception of Himachal Pradesh (51 percent), Meghalaya (49 percent) and Sikkim (46.5 percent) lie

well below the line of best fit. This means that they have lower rates of female LFPR than is predicted for their level of per capita SDP. For example, the Philippines with a GDP per capita of $3200 has a female LFPR of 45 percent while Kerala and Haryana with similar per capita SDP have rates of 28 percent and 15 percent respectively. At a similar level of income, Vietnam with a GDP per capita of around $3250 is a positive outlier with a female LFPR of 70 percent. Sub-Saharan African (SSA) countries like the Congo, Kenya, Ghana, Nigeria, and Zambia have a per capita GDP within 10-15 percent of the Indian states of Punjab, Maharashtra and Tamil Nadu but while the SSA countries have a female LFPR in the range of 50-70 percent, that of these Indian states is 18, 29 and 35 percent respectively.

How much of this underperformance on part of the Indian states is due to measurement errors in levels of women's employment, how much due to cultural barriers to women's participation in the workforce and how much due to a lack of labour demand is an important question. We take these issues up in future chapters.

Figure 2.11: Indian states are consistently lower than comparable developing countries in terms of the female LFPR
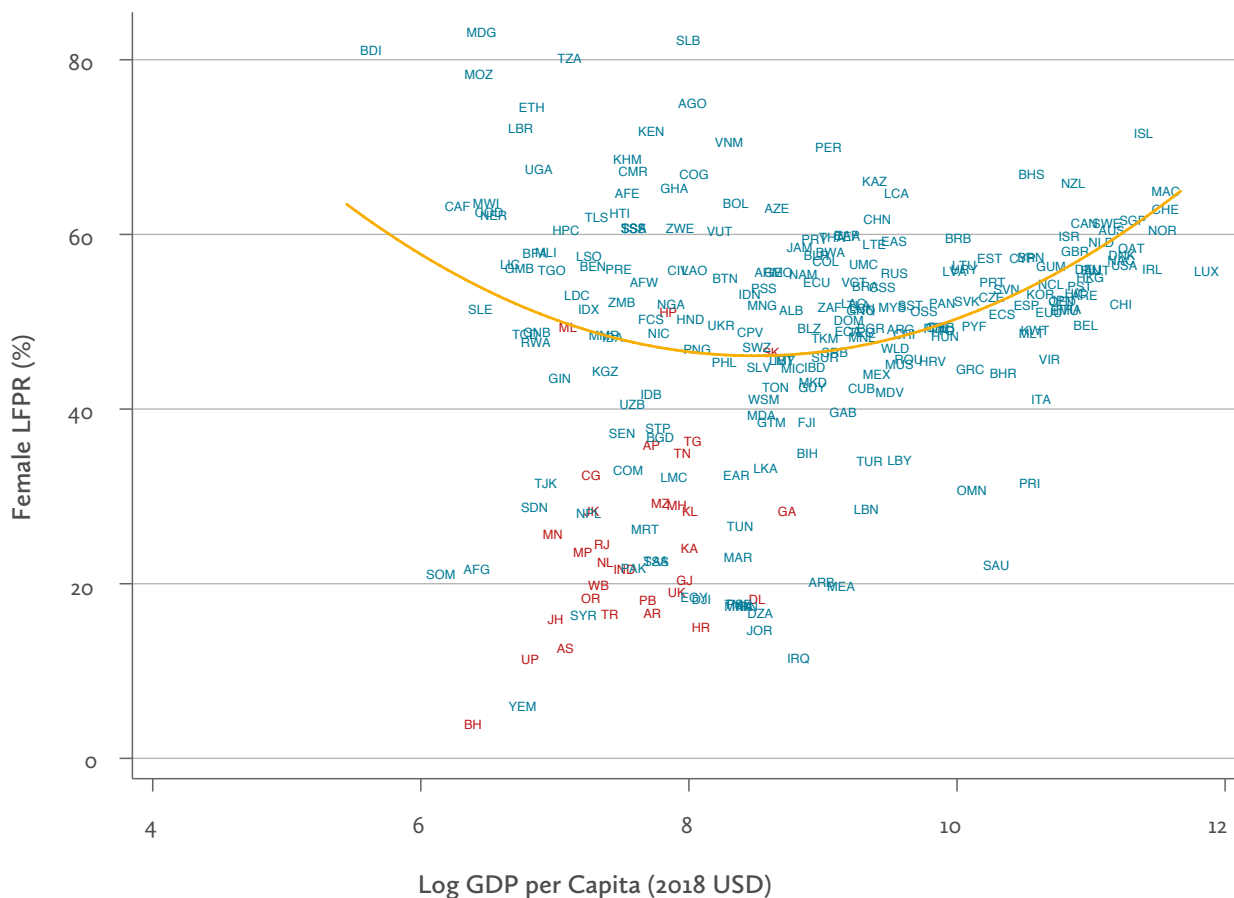
Sources and notes: PLFS 2018-19 and World Development Indicators. State Domestic Product data from the RBI is converted to USD using the prevailing exchange rate.

A key theme of this report is India's experience of structural transformation and how various social identities have fared in this respect. As a prelude to this analysis (presented in the next chapter), here we examine the state-level variation in 2021-22 in two key indicators of structural transformation- the proportion of workforce engaged in non-farm work and the proportion employed as regular wage or salaried workers (Figure 2.12 and 2.13). Firstly, as expected, men tend to be in non-farm work and regular wage work far more than women, across states. And also as expected, Kerala and Tamil Nadu stand out as states with relatively high levels of male and female employment in non-farm work. An interesting case, however, is that of Punjab, which also shows high levels of non-farm employment share for men and women and even more strikingly, a high proportion of women regular wage workers compared to most other states.

Finally, we take a look at average earnings at state level for various types of employment (Table 2.4). The data are ordered in increasing levels of total labour income ranging from under ₹10,000 per month in the case of Chhattisgarh to more than ₹20,000 per month in the case of Delhi and Goa. Absolute levels are likely to be underestimated, but it is worth focusing on the ratios. The richest to poorest ratio for self-employment is much higher than that for regular wage work (3.5 versus 2.2).

Punjab is an interesting case showing a high proportion of women in regular wage work, compared to most other states.

This is to be expected since local infrastructure and institutions that are likely to vary a great deal across states matter far more for earnings from self-

Figure 2.12: Share of women and men engaged in non-agricultural employment across Indian states in 2021–22
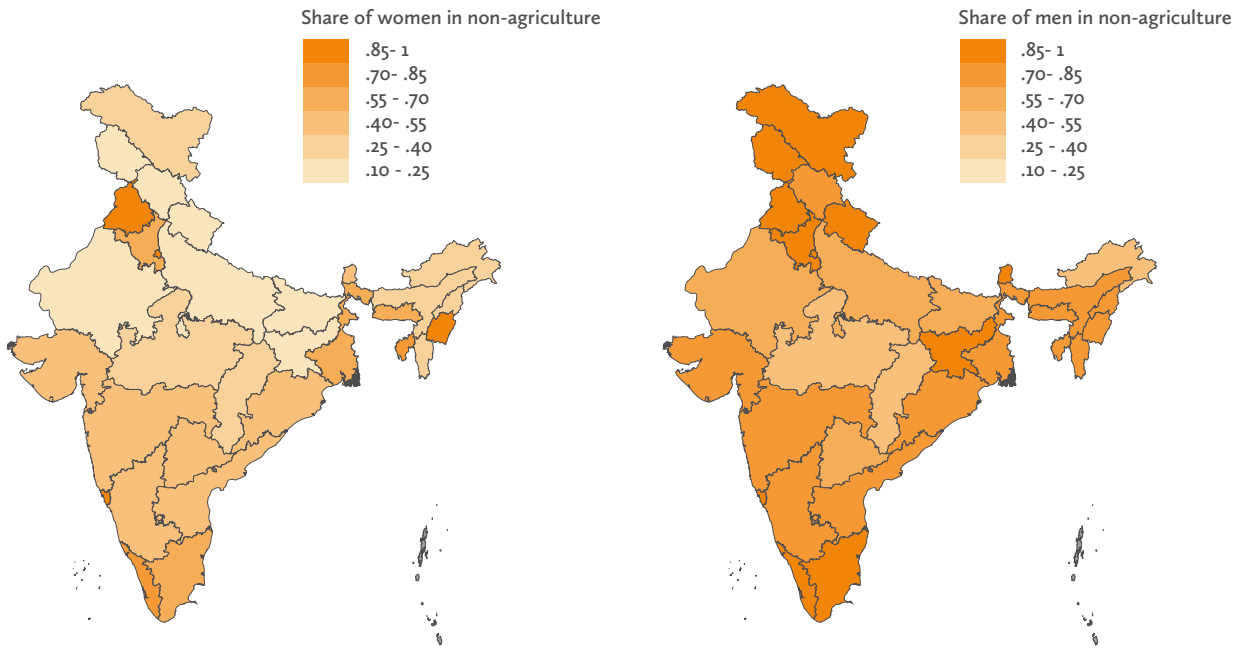
Sources and notes: PLFS 2021-22. Employment is defined based on Current Weekly Status. The scale is the same in both maps to make for easier comparison across gender.

Figure 2.13: Share of women and men engaged in salaried or regular wage work across Indian states in 2021–22
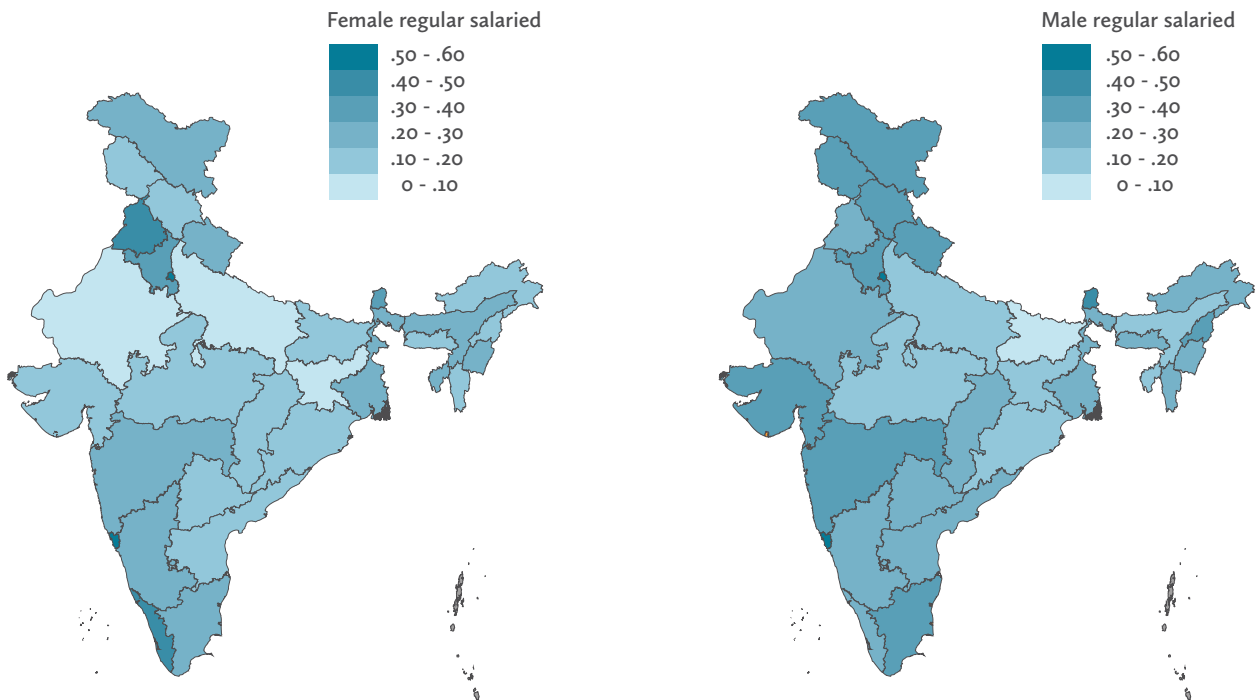
Sources and notes: PLFS 2021-22. Employment is defined based on Current Weekly Status. The scale is the same in both maps to make for easier comparison across gender.

Table 2.4: Average monthly labour earnings across states (2021–22)

State	Casual	Regular wage	Self employed	All types
Chhattisgarh	5,643	17,122	7,705	9,716
West Bengal	6,850	14,540	9,016	9,867
Orissa	5,870	17,357	9,676	10,104
Uttar Pradesh	7,313	16,110	9,422	10,156
Jharkhand	6,354	18,079	9,299	10,254
Madhya Pradesh	4,951	17,520	11,331	10,862
Bihar	9,449	17,325	10,962	11,172
Meghalaya	7,496	20,252	9,281	11,201
Assam	9,390	15,281	10,424	11,335
Tripura	7,344	18,382	12,090	11,797
Himachal Pradesh	9,285	20,979	7,862	11,921
Andhra Pradesh	7,468	19,131	12,557	12,873
Gujarat	6,779	15,761	13,321	13,266
Tamil Nadu	7,784	18,772	13,833	13,593
Rajasthan	7,854	23,017	11,875	13,990
Maharashtra	6,794	20,896	14,120	14,471
Punjab	8,601	15,333	18,462	14,856
Karnataka	7,588	21,939	14,805	14,938
Uttarakhand	10,252	21,329	12,238	15,168
Telangana	6,954	22,899	14,360	15,174
Jammu & Kashmir	10,082	24,931	12,311	15,631
Manipur	13,097	23,386	12,464	15,788
Nagaland	10,335	25,031	10,411	15,976
Kerala	11,395	22,287	13,369	16,043
Sikkim	14,513	22,166	10,300	16,175
Haryana	8,901	19,849	15,900	16,246
Arunachal Pradesh	10,606	31,858	11,834	17,407
Ladakh	10,203	31,900	10,904	17,977
Mizoram	9,709	31,109	16,149	20,205
Goa	11,801	21,209	24,967	21,849
Delhi	13,040	22,790	26,548	23,580

Sources and notes: PLFS 2021-22. 2022 constant rupees. States have been ordered in increasing levels of total labour income.

employment than for salaried work where wages are likely to be set taking regional or even national and global conditions into account. This is an argument for enabling a more rapid transition to regular wage work.

2.4 Conclusion

The growth slowdown of 2018-2020 followed by the Covid-19 pandemic of 2020-2021 have left a strong impact on the labour market. As of the end of 2022, the structure of the workforce is skewed more towards self-employment than it was before the slowdown, and earnings remain depressed. The impact is strongly gendered with women being more affected. In addition, there are legacy problems such as persistently low female labour force participation rates and persistently high rates of unemployment. However, it is also true that when we view developments in a more long-run perspective, the Indian economy has made important strides in its structural transformation journey. We now turn to these changes.

Endnotes

1   Absolute numbers are obtained by multiplying survey ratios using the Current Weekly Status definition of employment with population projections based on the 2011 Census (Government of India 2019). A different definition of employment or another source of population projections will alter the exact rates and absolute numbers but will not change the main conclusions. See Nath and Basole (2020) for issues around estimation of absolute numbers from survey rate using population projections.

2   The PLFS allows us to analyse employment trends on a quarterly basis. This is possible because both in rural and urban areas the sampling is done in such a way that representative estimates can be calculated for each quarter. In rural areas, samples for the stratum are randomly drawn in the form of sub-samples which are equally allocated among the four quarters. In each quarter of the survey period, twenty-five percent of the first-stage units of annual allocation are covered. In urban areas, in addition to the above, there are re-visits that provide a panel. But we use only the first visit data for both rural and urban households. The NSSO also

provides quarterly numbers for both rural and urban areas in its reports. For e.g. see Tables 42-47 of the Appendix, pgs A-268 to A-359 of PLFS 2019-20 Annual Report. See Methods Appendix for more details.

3   Note that this means that a given quarter is not being compared to the equivalent quarter in an earlier or later year but rather to the April-June 2018 quarter. Since our focus is on secular changes over this period and not seasonal ones, we have chosen this method of analysis.

4   There is a debate on whether a U-shape relationship actually exists, whether across countries at a point in time or within a country over time. See Goldin (1994), Lahoti and Swaminathan (2016), Heintz, Kabeer and Mahmud (2018), Deshpande and Kabeer (2019).

3

Growth, structural change and social identities – India and the States

3

Growth, structural change and social identities – India and the States

In the previous chapter we analysed significant changes in the Indian labour market in the past few years. We now review the experience of structural transformation over the last four decades. We analyse the effectiveness of economic growth in enabling structural change at the all-India and state-levels and delve into the differential nature of structural change with respect to the key social identities of gender, caste and religion.

In broad terms, the Indian economy presents an interesting combination of jobless growth along with (slow) structural transformation. By "jobless growth" we refer to the phenomenon of low or declining employment elasticity of growth or a lack of correlation between output growth rate and employment growth rate (Kannan and Raveendran 2009; Tejani 2016; State of Working India 2018). We define structural transformation as both a rise in the proportion of workers engaged in non-agricultural activities (the Kuznets process) and the proportion engaged in regular wage or salaried work, as opposed to casual wage work or self-employment (the Lewis process).

We analyse these processes at cross-country, national and state levels. The period of analysis is from 1983-84, the first year that detailed unit-level data is available from the NSSO employment surveys till 2021-22, the last available PLFS round. Though the recent growth slowdown as well as the Covid-19 pandemic are part of the overall period, we do not examine the impact of these in this chapter since we have discussed both in Chapter Two.

3.1 India's changing economic structure at-a-glance

Before delving into the time trends we examine the approximate number of individuals in broad sectors of the economy as of 2021-22, the most recent PLFS round (Figure 3.1).[1] The entire working age (15+ years) population of around 1021 million can be divided into those in the workforce (493 million) and those outside it. The workforce constitutes less than 50 percent of the working age population, a phenomenon largely explained by low rates of female labour force participation.

From a structural change perspective, two further divisions are important to note: the number of workers in agriculture (211 million) versus the non-agricultural economy (282 million), and within the non-agricultural sector, the "subsistence" sector consisting of those engaged in self-employment or casual wage work (170 million) versus the "modern" sector consisting of those engaged in regular wage or salaried work (112 million). Here it is important to clarify our terms. Box 3.1 elaborates on definitions of "subsistence" versus "modern" and the related concepts of informal and formal.

The 2021-22 picture just discussed is a result of a growth process. Before we analyse this process in some detail, it is worth casting a glance at the evolution of sectoral shares in output (GDP or value-added) and employment (Figure 3.2). As is

Figure 3.1: Major components of the workforce with absolute numbers
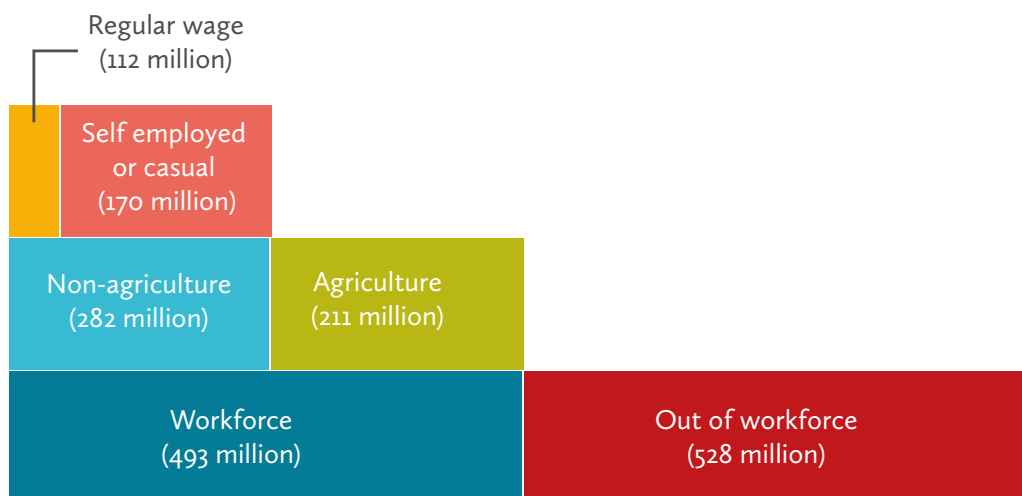
Sources and notes: PLFS (2021-22) and Government of India (2019)

Box 3.1:   Clarifying the terms: subsistence–modern, formal–informal

When analysing structural change the analytical distinction between the subsistence economy and the modern economy is more relevant for us than the difference between formal and informal employment, important as the latter is from a welfare point of view. The subsistence sector comprises self-employment (own-account work and unpaid family helpers) as well as casual labour but not regular wage work. The hallmark of the first two is that the demand for labour (workers) adjusts to the supply of workers (i.e. all those who are seeking work). So open unemployment is rare and underemployment is prevalent (see Ghose 2016 for details). In contrast, firms who hire regular wage labour (typically on weekly or monthly wages or salaries) determine their demand for labour based on profitability. Here labour demand may fall short of labour supply and open unemployment (or fallback into subsistence sector which is a residual sector) is also a possibility. Labour productivity is generally far higher in the modern sector and hence the transfer of labour from

the subsistence sector to the modern sector is a desirable aspect of structural change.

Data on the subsistence economy is not directly available. Hence, to gauge the pace of the Lewis process, we need to resort to proxies. We consider two such proxies, the share of regular wage employment in the non-agricultural sector and the share of non-agricultural employment accounted for by establishments that employ 10 or more workers (the "organised sector"). The first is a looser definition of the "modern" sector because it includes regular wage employment in all enterprises of any size, some of which may be very small and part of the unorganised sector (i.e. employing less than 10 workers). The latter is a stricter definition. But it is important to note that neither proxy takes into account job security or benefits. We deliberately set aside the issue of security of tenure or social security because our interest here lies in the ability of the economy to create labour demand for regular wage work of any kind.

Admittedly this is a low bar. Indeed, our policy focus should be on reducing the proportion of regular wage workers who do not have written contracts or benefits (generally considered informal). While most regular wage workers in the unorganised sector are likely to be informally employed by this definition, over the years,

the organised sector has expanded its share of informal workers via third-party contracts (see Chapter Four in State of Working India 2018, especially Table 4.1). We take this issue up in Chapter Five.

well-known, India's sectoral structure has evolved over time away from agriculture and towards services with a significant role also being played by construction. Manufacturing has failed to expand either its share of output or its share of employment significantly over this period.

Table 3.1 shows a few key indicators for the periods shown in Figure 3.2. In the table the first two sub-periods have been combined into one decade-long period. The periods correspond roughly to the pre-reform (1983 to 1993), early reform (1993 to 2004) and later reform (2004 to 2011 and 2011 to 2018)

years. A few points are worth noting. The Indian economy registered a healthy rate of overall as well as non-agricultural output growth, the latter being above 6 percent per annum over the entire period, increasing to more than 7 percent in the most recent sub-periods, prior to the slowdown and the Covid crisis. However, except for the last sub-period, non-agricultural employment grew slower than output and the growth elasticity of employment (percent change in non-farm GDP divided by percent change in non-farm employment) fell from 0.6 in the first period (1983-1993) to 0.2 in the last but one period (2011-2017).

Figure 3.2: Sectoral shares in GDP and employment since the 1980s

Legend: 1983-84, 1987-88, 1993-94, 1999-00, 2004-05, 2011-12, 2017-18, 2021-22
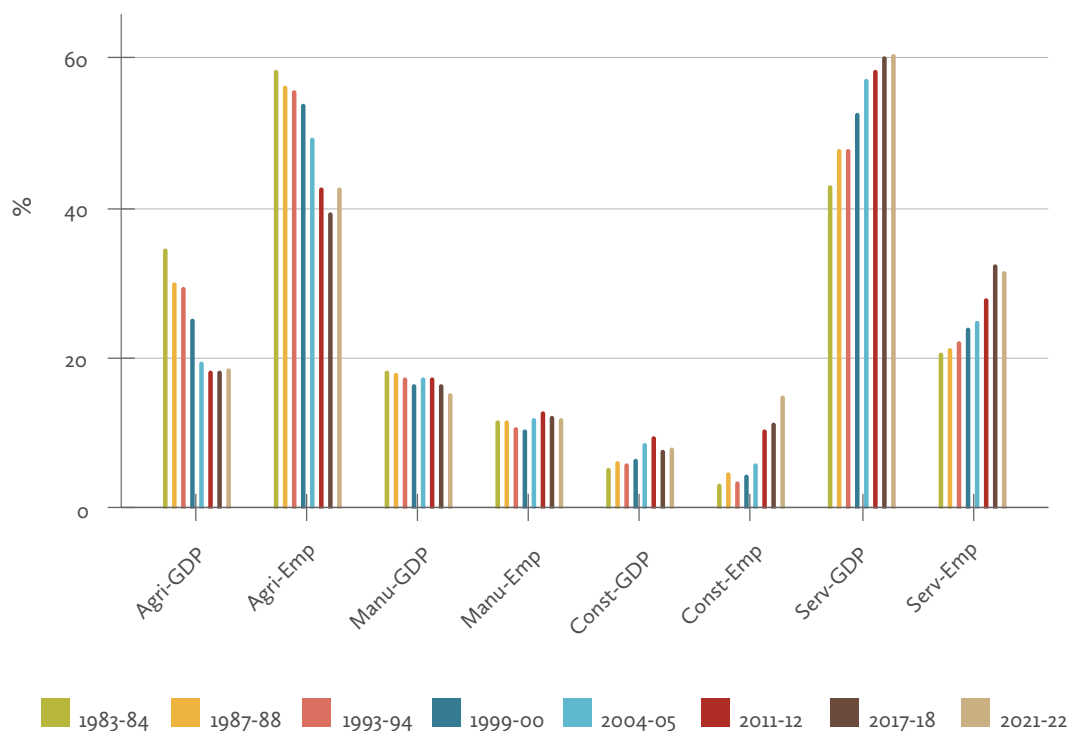
Sources and notes: NSSO EUS and PLFS various years and RBI-DBIE. Employment is defined based on Current Weekly Status.

Table 3.1: Key indicators since the 1980s

Indicator	1983-1993	1993-2004	2004-2011	2011-2017	2017-2021
GDP growth (%)	5.1	5.7	6.9	6.7	3.5
Non-agri value-added growth (%)	6.5	6.9	7.7	7.4	3.4
WAP growth (%)	3.2	2.3	2.5	2.1	1.3
WF growth (%)	3.6	2.2	0.8	-0.1	3.7
Non-agri emp growth (%)	4.2	3.2	3.4	1.7	3.4
Non-agri salaried growth (%)	3.3	3	4.2	4.1	2.8
Non-agricultural employment elasticity	0.6	0.5	0.4	0.2	1.0

Sources and notes - RBI Database of the Indian Economy, RBI KLEMS, NSSO EUS and PLFS various rounds.

Coming to the employment indicators, note that between 2004 and 2017, the workforce grew much slower than the working age population (i.e. the workforce participation rate dropped). In part this is due to a movement of younger individuals into education, a desirable outcome of the development process. In part it is also due to a decline in the number of agricultural workers not being compensated for by a corresponding increase in non-agricultural employment. Most strikingly, in the 2011 to 2017 sub-period, the decrease in agricultural employment was barely met with a commensurate increase in the non-agricultural workforce, resulting in near zero growth of the workforce. This is a gendered process related to the widely discussed withdrawal of women from the labour force (Lahoti and Swaminathan 2016, Mehrotra and Parida 2017, Deshpande and Kabeer 2019; State of Working India 2021). While male workers moved from agriculture to construction and services, women agricultural workers largely moved out of the workforce in this period.

The last sub-period which includes the Covid crisis presents a completely different picture. We examined this period in Chapter Two and do not discuss it here except to reiterate that it illustrates the danger of relying only on aggregate measures such as WPR to gauge the health of the labour market. For example, note that between 2004

to 2017 regular wage employment grew faster than overall non-agricultural employment but the opposite was true in the 2017 to 2021 period. That is, the share of salaried workers in the non-agricultural workforce increased till 2017 and then declined. As we have seen, this was due to the rise in self-employment during the growth slowdown and the Covid crisis.

3.2 Jobless growth and structural transformation

3.2.1 India in cross-country perspective[2]

We now move to a detailed analysis of India's performance compared to other developing countries. We start with a commonly used indicator of the tightness of the relationship between economic growth and employment, viz. the growth elasticity of employment or the percent change in total employment divided by the percent change in GDP over a period. While the elasticities for India have been widely reported both in our previous reports and elsewhere (Papola and Sahu 2012; Misra and Suresh 2014; State of Working India 2018), here we go beyond this analysis to place India

in a cross-country perspective. We focus on the relationship between non-agricultural output and non-agricultural employment.

We draw on a cross-country database (the Groningen Economic Transformations Database or ETD) which allows us to estimate the elasticity for a large number of developing countries. Figure 3.3a shows the distribution of non-agricultural employment elasticities across 50 countries in the database for the period 1990 to 2019 (just prior to the pandemic). The value for India is indicated by the red line. Note that it is significantly lower than the mean.[3] This means that, even though most developing countries have been facing headwinds that make employment generation difficult, the pattern of growth in India has been less job-intensive than the developing country average.[4]

The elasticity numbers answer the question of what happens to the level of employment when the level of GDP changes by a certain percentage amount. From a policy as well as analytical perspective, there is another relationship that is important to examine - that between output growth and employment growth as opposed to the level changes captured by the elasticity numbers. In economic policy making, it is commonly assumed that higher GDP growth will lead to higher employment growth. Thus the

question is, does the Indian economy tend to create jobs at a more rapid pace in periods when it grows faster?

Most developing countries have been facing headwinds that make employment generation difficult but the pattern of growth in India has been less job-intensive than average.

We estimate the contemporaneous relationship between output growth rate and employment growth rate in a cross-country regression framework using the same dataset as was used for the elasticity exercise. Figure 3.3b shows the distribution of the coefficients obtained from this regression.[5] Strikingly, there is almost no relationship (and if anything a weak negative relationship) between output or GDP growth and employment growth for India (coefficient = -0.11). What this means is that far from employment growing faster when GDP grows faster, years of fast GDP growth have, on the contrary, tended to be years of slow employment growth. Note that the developing country average is 0.3, which means that on average there is a positive relationship between GDP growth and employment growth. But this is not the case for India.[6] In order to understand this lack of relationship better, we need to look at India-specific data sources, which we

Figure 3.3: India's jobless growth in cross–country perspective

a. Distribution of growth elasticities of non-agricultural employment

b. Distribution of coefficients - GDP growth versus employment growth
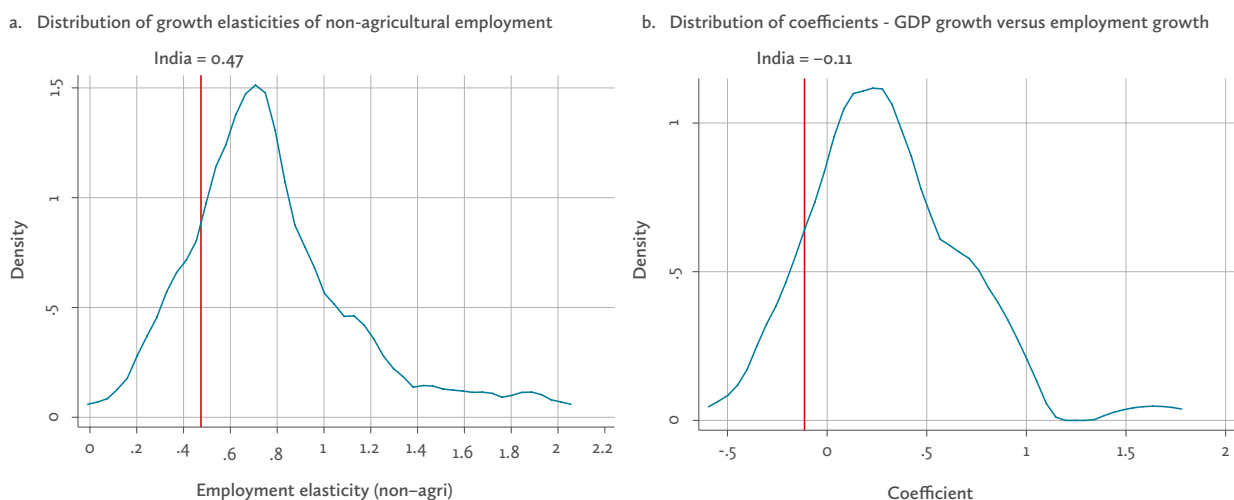
Sources and notes: Groningen ETD. See Methods Appendix for details on the regressions from which coefficients are obtained.

do in the next section. We have discussed possible solutions to India's employment challenge in the State of Working India 2021. In the final chapter of this report we return to this concern.

Does this jobless growth mean that the Indian economy has been unable to structurally transform itself? Indeed, the common understanding is that India still has "too many" workers in agriculture and a larger than expected informal sector. Once again, it is useful to place India's performance, in the years prior to the pandemic, in a cross-country perspective.

Basole (2022) examines the cross-country relationship between sectoral shares of GDP or employment and the level of GDP per capita. We

expect the relationship to be negative for the share of agriculture in GDP or employment, and positive for the share of industry or services. Similarly we once again expect a negative relationship between the informal share of employment and GDP per capita. That is, richer countries tend to have a smaller share of their workforce in agriculture as well as a small share of informality. But the more significant question is, where does India lie with respect to the developing country average?

Figures 3.4, 3.5 and 3.6 are a series of cross-country scatter plots where each point is a country (given by its three letter acronym). The Y-axis shows the proportion of the workforce in a sector (0.5 means 50 percent). We see that India (shown in red) lies much closer to the regression line for the agricultural

Figure 3.4: India's employment shares in cross-country perspective
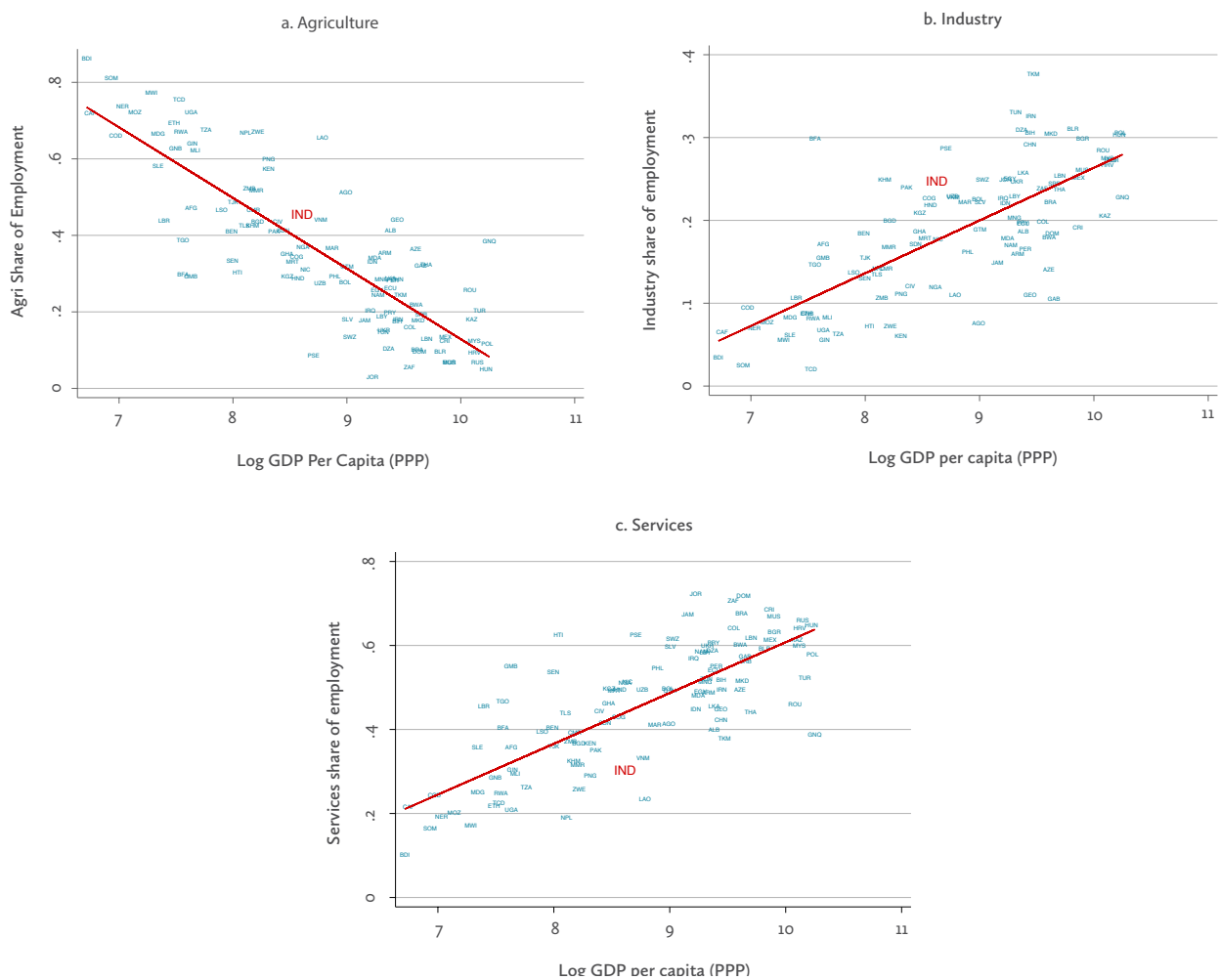
a. Agriculture

b. Industry

c. Services

Sources and notes: World Development Indicators (WDI). The Y axis shows proportions (out of 1) and not percentages.

Figure 3.5: Industrial employment shares in cross-country perspective
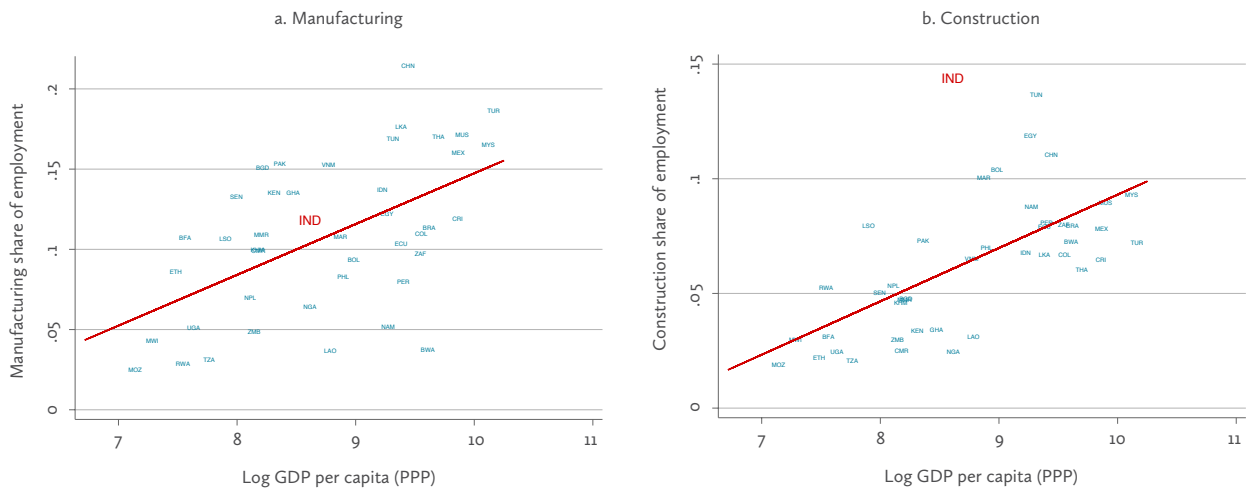
a. Manufacturing

b. Construction

Sources and notes: World Development Indicators (WDI). The Y axis shows proportions (out of 1) and not percentages.

share (Figure 3.4a) as compared to the informal share where it is a significant outlier (Figure 3.6). Let us examine this in more detail.

The proportion of the workforce in agriculture (0.43 or 43 percent) was around 8.8 percent points higher than expected from the average relationship across all the countries in the dataset as of 2018. What might perhaps come as a surprise is that India lies above the regression line in industry share of employment and below it in services share (Figure 3.4b and c). India's industrial share of employment is 6.4 percentage points higher than predicted, while its services share is 15 percentage points lower than predicted for its level of GDP per capita.

The answer to the puzzle lies in the performance of the construction sector. Figure 3.5 shows similar plots as before for the two main components of the industrial sector, viz. manufacturing and construction. The data are noisier and only available for half the number of countries. But clearly (and perhaps surprisingly), for manufacturing India lies close to the regression line. While for construction, India is a large outlier with its employment share being 9 percentage points higher than predicted. Note that construction tends to have a much smaller

share of the total workforce than agriculture (say 5-10 percent rather than 40-50 percent), so a 9 percentage point difference from average is very large indeed.

Given its level of GDP per capita, India has a higher share of workers in agriculture, but not extraordinarily so. It is a strong outlier when it comes to the share of the workforce in construction and the informal sector.

The above analysis broadly confirms well-known stylized facts regarding the Kuznets process for India but also throws up a few surprises. Narrowly defined as movement of workers out of agriculture to the non-farm sector, the process is taking place more or less as expected. It is true that given its level of GDP per capita, India has a higher share of workers in agriculture, but not extraordinarily so. More importantly, instead of moving into higher value-added sectors such as manufacturing or modern services, the bulk of the transition has been into low value-added construction activities (Nayyar 2019a, b). India is indeed a strong outlier when it comes to the share of the workforce engaged in construction.

Now, we move from the Kuznets Process to the Lewis Process. Here, data is much more scarce. Reliable and comparable cross-country measures of the size of the informal sector are difficult to come by. The WDI database gives only one indicator, viz. the share of self-employment in total employment. However, the ILO database from which this indicator is drawn, has some definitional problems and inconsistencies. It seems safe to assume that this indicator in fact gives the share of the workforce engaged in both self-employment and casual wage work.[7] Hence we refer to it as the share of informal employment in total employment. Figure 3.6 gives the by now familiar scatter plot for informal share versus GDP per capita. At 0.765 or 76.5 percent India is an outlier having a much higher informal share of employment for its level of GDP per capita with its share being 23 percent points higher than predicted.

3.2.2 Output growth versus employment growth over the long-run

The disconnect between the Lewis and Kuznets processes can be substantiated by looking at data from the quinquennial NSSO employment surveys as well the more recent annual Periodic Labour Force Surveys. Figure 3.7 shows that the share of agriculture in total employment has been steadily declining in India, picking up pace since the early 2000s. The share of non-agricultural employment rose from 37 percent in 1983-84 to 60 percent just prior to the pandemic in 2018-19. However, the pace of rise in non-agricultural employment was not matched by the rise in the share of regular wage employment within the non-agricultural sector. Only in the most recent high growth period do we see a small increase in the share of regular wage workers, as well as a small increase in the share of

Figure 3.6: India's informal employment share in cross-country perspective
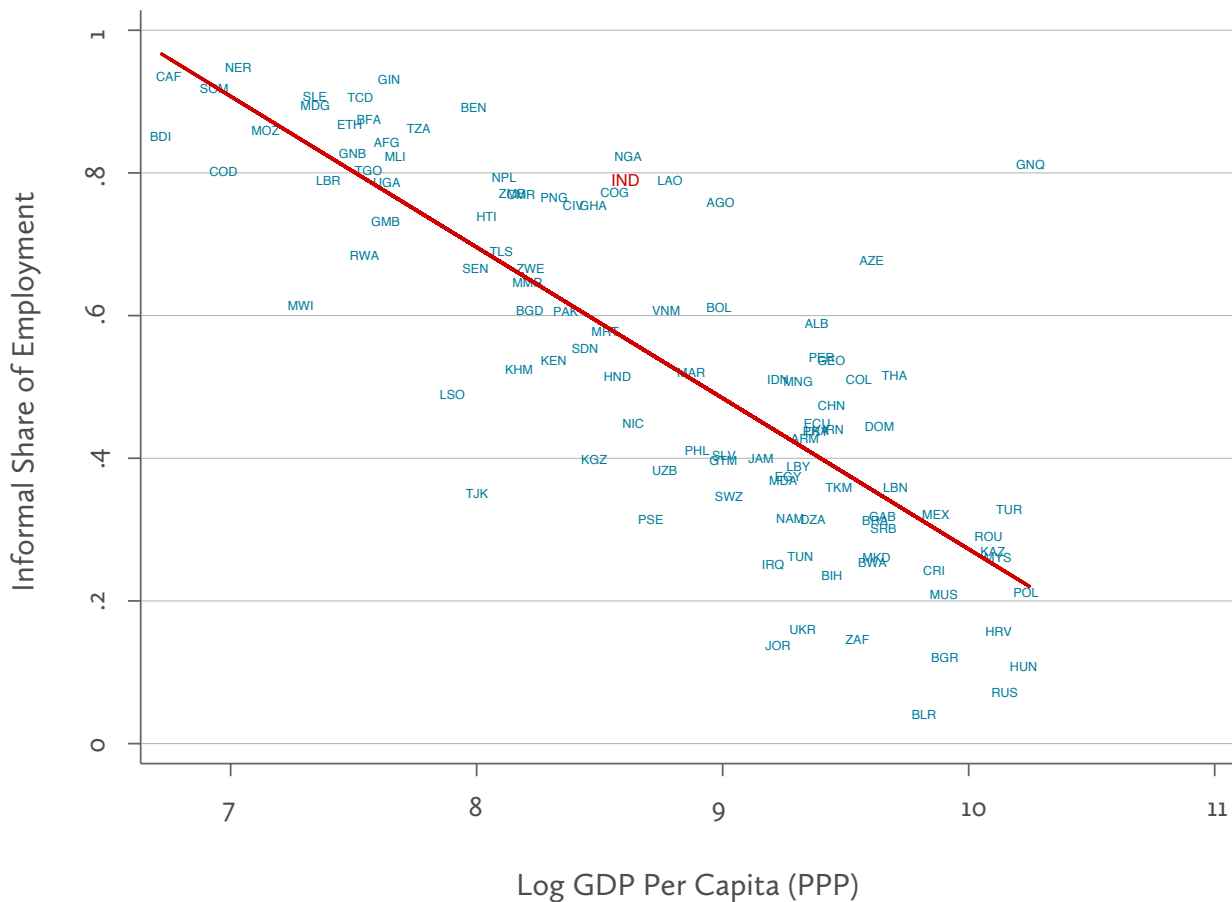
Sources and notes: World Development Indicators (WDI). The Y axis shows proportions (out of 1) and not percentages.

Figure 3.7: Share of workers in non-agricultural employment rose much faster than the share in regular wage employment or organised sector
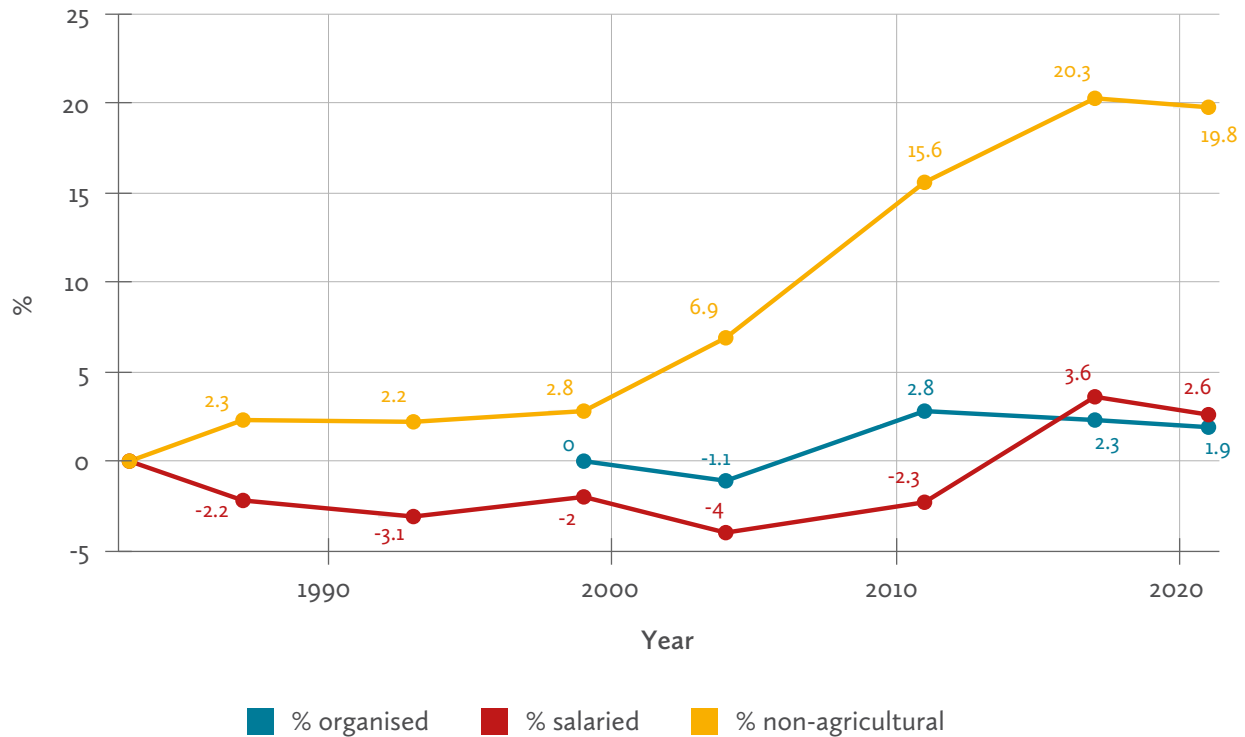
Sources and notes: NSSO EUS-PLFS various years. The numbers are indexed to their 1983 values. % salaried implies the proportion of regular wage workers in the non-agricultural sector.

workers employed in establishments having 10 or more workers. This means that workers leaving agriculture were much more likely to continue in the subsistence sector albeit doing non-agricultural work.

We noted in Section 3.2 that the relationship between GDP growth and employment growth has been weak in India. We now examine the dynamics of the workforce more closely to better understand this lack of relationship. First, note that non-agricultural employment growth and non-agricultural output growth move roughly out of phase with each other over the entire period from 1983-84 to 2018-19 (Figure 3.8a) as expected from the low correlation seen in cross-country data.

Further there is no clear relationship between non-agricultural output growth and the growth of salaried or organised sector employment either (Figure 3.8b). Observe that employment growth in the organised sector correlated positively with

output growth during the first part of the high growth period, rising from 4.2 percent to 5.8 percent CAGR between 1999-2004 and 2004-2011. But subsequently it plummeted to less than 1 percent in the 2011-2017 period rising again in the last sub-period when growth took a hit. Admittedly as noted in Box 3.1, "regular wage" is a liberal proxy for the modern sector. Stricter definitions such as the presence of social security, protection under labour laws etc, will change the nature of the trends somewhat, but for our purposes the key analytical difference is between employment that is created via labour demand in the modern economy and does not display income sharing or work sharing properties, versus employment that is created via supply and has income and work sharing.

Taken together, these trends present a far more ambiguous picture of the Lewis process than what we saw for the Kuznets process. There is some movement away from casual to regular wage work, indicating an overall improvement in the quality of

Figure 3.8: Non-agricultural output growth versus employment growth

Sources and notes: RBI KLEMS and NSSO EUS-PLFS various years. Compounded Annual Growth Rates between survey years..

employment but it is not accompanied by a rise in the scale of production in the economy, as seen in the stagnant share of the organised sector in total employment. These divergent trends also suggest that the rise in regular wage work must have occurred even in microenterprises that are part of the unorganised sector. Indeed, the share of regular wage workers in the unorganised sector rose from 21 percent in 2011-12 to 25 percent in 2018-19. While in the organised sector it rose from 72 to 80 percent. Encouragingly the share of casual workers only rose marginally in the unorganised sector and fell sharply in the organised sector in this period. Overall, this represents a small but significant improvement in employment conditions in the Indian economy.

The lack of relationship between output growth and employment growth documented in the foregoing pages can be due to two distinct reasons. First due to a rise in labour productivity that cancels out employment effects and second, the labour supply dynamics or movements into and out of the workforce or into and out of the modern economy. The slow pace of structural transformation in India seems to owe itself to both these factors. We focus here on the latter. The labour productivity channel is discussed in more detail in Dasgupta and Basole (forthcoming).

3.2.3 Labour force dynamics and structural transformation – the role of gender

We start by looking at how the working age population has moved into either the workforce (employment), education or domestic duties (care work and other household work). The following three figures (3.9 to 3.11) show what fraction of the change in the working age population or the workforce between two survey years is accounted for by different components mentioned. So for example, if the working age population increased by 100 individuals between two time points and 25 of them went into education, the value of education will be 0.25. It should be kept in mind that sometimes there may be a decrease in a particular component even as the denominator increases.

We focus on the differences between male and female workers rather than looking at the aggregate numbers which hide substantial gender-based heterogeneity. Several features are worth noting. As expected, domestic (unpaid) work forms a large destination for new women entering the working age. In contrast, this is negligible for men. Pertinent to the foregoing discussion, note that during the high growth years (2004 to 2017), the share of women in the workforce turned negative (Figure 3.9). This is the much-discussed fall in women's workforce participation and it entails an absolute fall in the size of the female workforce in

Figure 3.9: Composition of the change in working age population
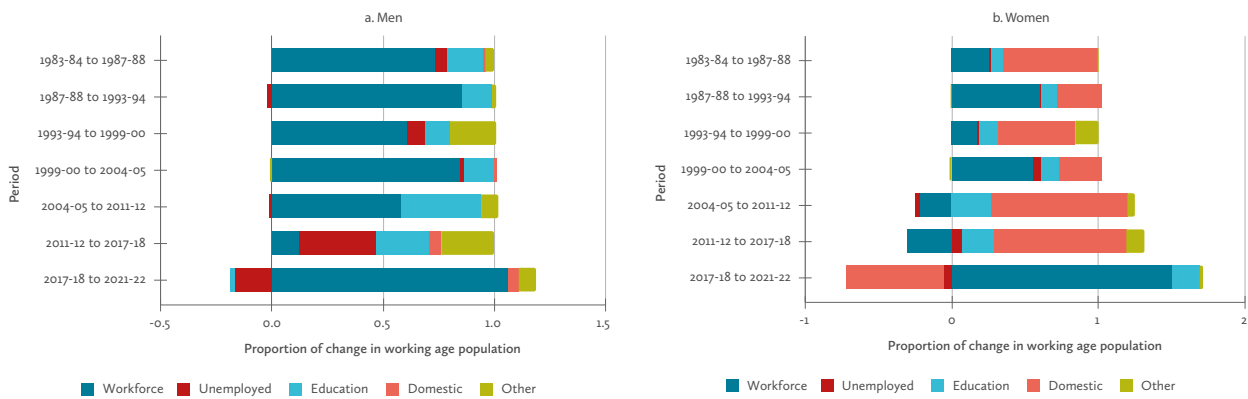
a. Men

b. Women

Sources and notes: NSSO EUS-PLFS various years. X-axis shows the share of each category in the total increase in working age population in each period.

this period. No such decline is seen for men. The post-2017 period clearly stands out as discussed in Chapter Two. Thus it seems that at least some of the lack of relationship between output growth and employment growth could be explained by falling women's employment during the high growth period and rising employment in the low growth period.

Much has been written about India's declining rural female labour force participation rate with two major factors identified for the falling LFPR - a displacement of women workers due to mechanisation and rising household income combined with social norms that value women's unpaid domestic work more than paid work (Lahoti and Swaminathan 2016; Mehrotra and Parida 2017; Deshpande and Kabeer 2019; Afridi,

Bishnu and Mahajan 2023). We discuss the factors affecting women's employment in more detail in Chapter Four. Here our interest lies in exploring the implications of movements into and out of the labour force as well as within the labour force across different sectors, for the aggregate relationship between output growth and employment growth.

The most striking decline in female employment has been in agriculture. Figure 3.10 shows what part of the change in the workforce between each survey year is accounted for by changes in agricultural versus non-agricultural employment. For men, agricultural employment fell in absolute terms between 2004 and 2017, while non-agricultural employment rose. The rise in the latter more than compensated for the fall in the former. There is a

Figure 3.10: Composition of change in workforce
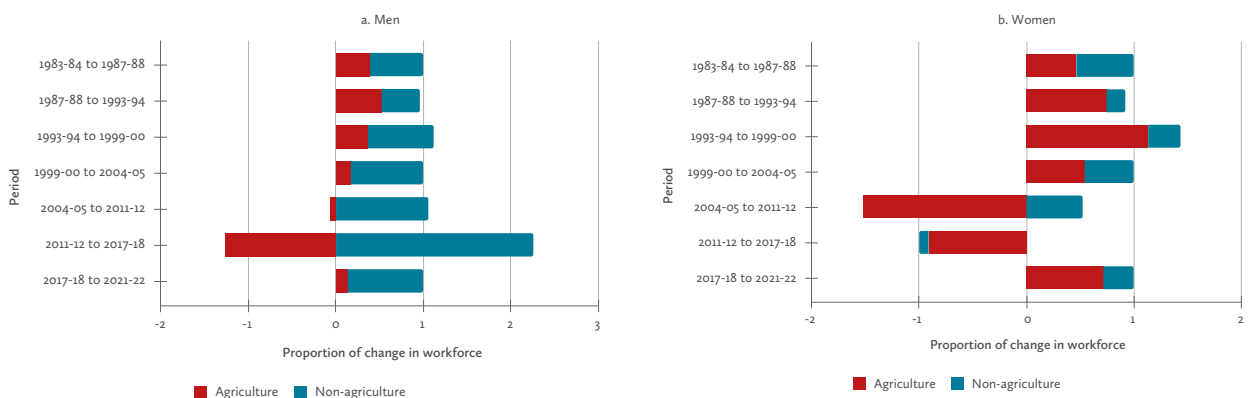
a. Men

b. Women

Sources and notes: NSSO EUS-PLFS various years. X-axis shows the share of each category in the total increase in workforce in each period.

Figure 3.11: Composition of change in the non-agricultural workforce
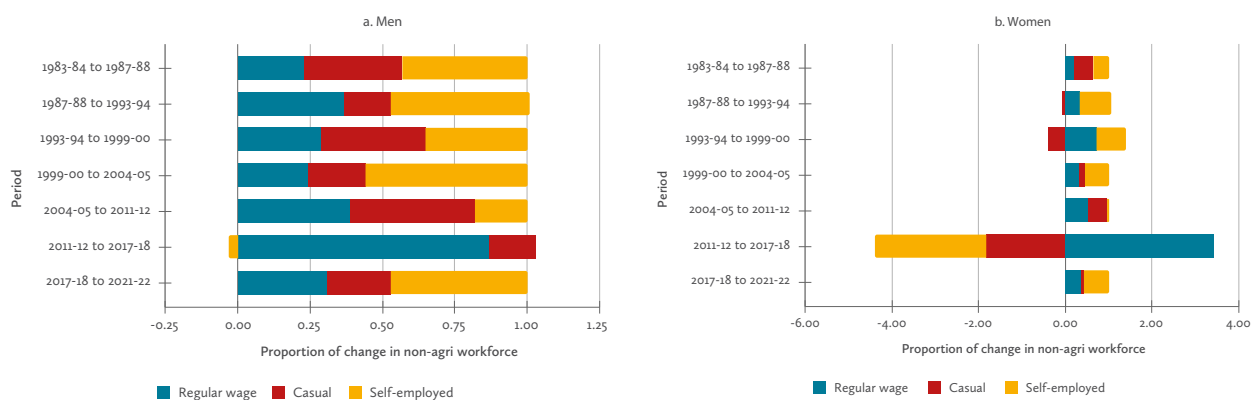
a. Men

b. Women

Sources and notes: NSSO EUS-PLFS various years. X-axis shows the share of each category in the total increase in non-agricultural workforce in each period.

rise in agricultural employment in the last period due to the pandemic.

But for women, between 2004 and 2011 the rise in non-farm work did not compensate for the decline in farm work and in the 2011-2017 period there was an absolute decline in both kinds of employment. Whether this was mainly a result of social norms that prevent a diversification of women's work (say from farms to factories) or due to a lack of labour demand (i.e. employment opportunities) for women, is a topic of current debate. We discuss this in Chapter Four. Note that in the most recent period, post 2017, we see a rise in women's farm as well non-farm employment. As a result, as discussed in Chapter Two, there has been a recovery in the female labour force participation rate.

These movements, especially of women workers, from agriculture to out of the workforce during the high growth period can explain the inverse relationship between growth and aggregate employment, but it cannot explain the lack of relationship between non-farm output and non-farm employment. For this, we look at the change in just the non-agricultural part of the labour force(Figure 3.11). Here, if the Lewis process is operating as expected, we hope to see a decline in the share of self-employed and casual wage work and a rise in regular wage work.

We see this only for the period 2011-2017 and largely for women workers. Strikingly, in this period, even though overall non-farm employment for women fell, we see a significant increase in salaried or regular wage non-farm employment. However, it is more than compensated for by a fall in self-employment and casual wage work, resulting in an overall decline in non-farm work for women in this period. Thus it seems that the income effect dominates the pace of employment generation in the formal sector. However, this increase in salaried work for women, taken together with the fact that the same period saw the largest increase in salaried work for men as well, meant that the proportion of regular wage or salaried workers in the workforce increased sharply during this period (see Figure 3.7).

3.3 The extent of structural change differs for different identities

The principal motivation for this year's theme of Social Identities and Labour Market Outcomes is the well known fact that one's gender, caste, religious or ethnic identity plays a crucial role in shaping one's employment opportunities, types of work,

earnings and so on. Hence it is not a surprise that the process of structural change varies a great deal across identities. In this section we examine the sectoral and employment structure of the labour force disaggregated by gender, social group (broad caste group) and religion. We disaggregate the services sector into traditional services such as retail and domestic service and modern services such as education, healthcare, information technology etc. See Methods Appendix for the details.

It is worth keeping in mind that, for the female workforce, these changes in sectoral or employment structure only pertain to women who are in the workforce at any point in time. Thus, the backdrop to these changes is the exit from the workforce of women between the period 2004 to 2017.

3.3.1 Sectoral shares

Coming to gender first, the decline in agricultural employment has been far more pronounced for men as compared to women. The share of male workers in agriculture and allied activities declined from 61 percent in 1983 to 37 percent in 2021, while the comparable change in women was from 75.5 percent to 59 percent (Figure 3.12). Thus not only was the share of men in agriculture lower at the start of our period of analysis, their pace of exit was also faster than women.

That is, agriculture has become feminised over time as men move into other sectors. Men still

constituted the majority (65 percent) of agricultural workers in 2021, but this was down from 69 percent in 1983. Figure 3.12 also shows that within the non-agricultural sector, while construction plays a crucial role for men (up from 3.6 percent in 1983 to 15.2 percent in 2021), it is far less important for women rising from 2 percent in 1983 to a peak of 5.7 percent in 2011, and then down to 4.3 percent in 2021. The rest of the difference between men and women is accounted for by traditional services such as retail with the proportion being identical for both sexes in manufacturing and modern services.

Share of male workers in agriculture declined from 61% in 1983 to 37% in 2021. Change for women was from 75.5% to 59%.

Figure 3.13 presents the same sectoral breakdown over time for the four major social or caste groups - Scheduled Tribes, Scheduled Castes, Other Backward Classes and Others (proxy for General caste). Others before 1999 included OBC's but since 1999 onwards others are total population - SC - ST - OBC. The predominance of agriculture and related activities as a source of livelihood for STs is clear, though there has been diversification over time. Another point to note here is the difference between SCs and OBCs. As of 2021, 44 percent of OBCs as opposed to 40 percent of SCs were in Agriculture while the share of SCs in construction was double that of OBcs (20 percent as opposed to 11). This is the major difference between the two

Figure 3.12: Changing sectoral structure of employment – gender
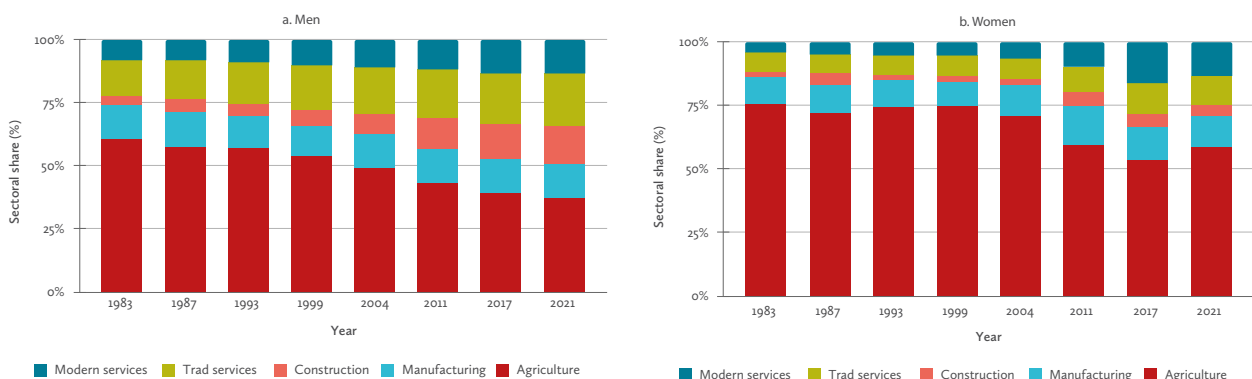
Sources and notes: NSSO EUS-PLFS various years.

groups with manufacturing and modern services being very similar for both. As expected, the Others show the most diversified sectoral structure and the largest pace of structural change. The drop in share of agriculture is the largest for this group from 60 percent in 1983 to 34 percent in 2021. They are also twice as likely to be in modern services compared to OBCs and SCs and three times more likely as compared to STs.

General castes show the most diversified sectoral structure and the largest pace of structural change.

Finally, coming to religion (Figure 3.14), we see a large difference in the sectoral breakdown between Hindus and Muslims. Muslims have tended to be far less in agriculture all through the time period under consideration and their pace of exit from this sector is similar to that of Hindus. The difference is made up for by manufacturing (12.4 versus 19.8 percent), traditional services (17 versus 26 percent)

and to a lesser extent by construction (12 versus 14.6 percent). Correspondingly, the share of modern services is higher for Hindus. The higher tendency of Muslims to be in manufacturing (which is likely to be mainly household and small-scale workshops) and services such as retail compared to Hindus has been the case for a long time. But their relative exclusion from modern services seems to be a more recent phenomenon. Until 2004 just under 10 percent of Hindus as well as Muslims were in modern services. But since then the distance has widened. While the share remained similar for Muslims (around 11 percent), it went up to 14 percent for Hindus before declining slightly during Covid.

3.3.2 Employment structure

A successful structural transformation process entails a shift of the workforce away from agriculture as well as from forms of employment such as own-account work, unpaid work in family enterprises and casual wage work. Though identity

Figure 3.13: Changing sectoral structure of employment – caste
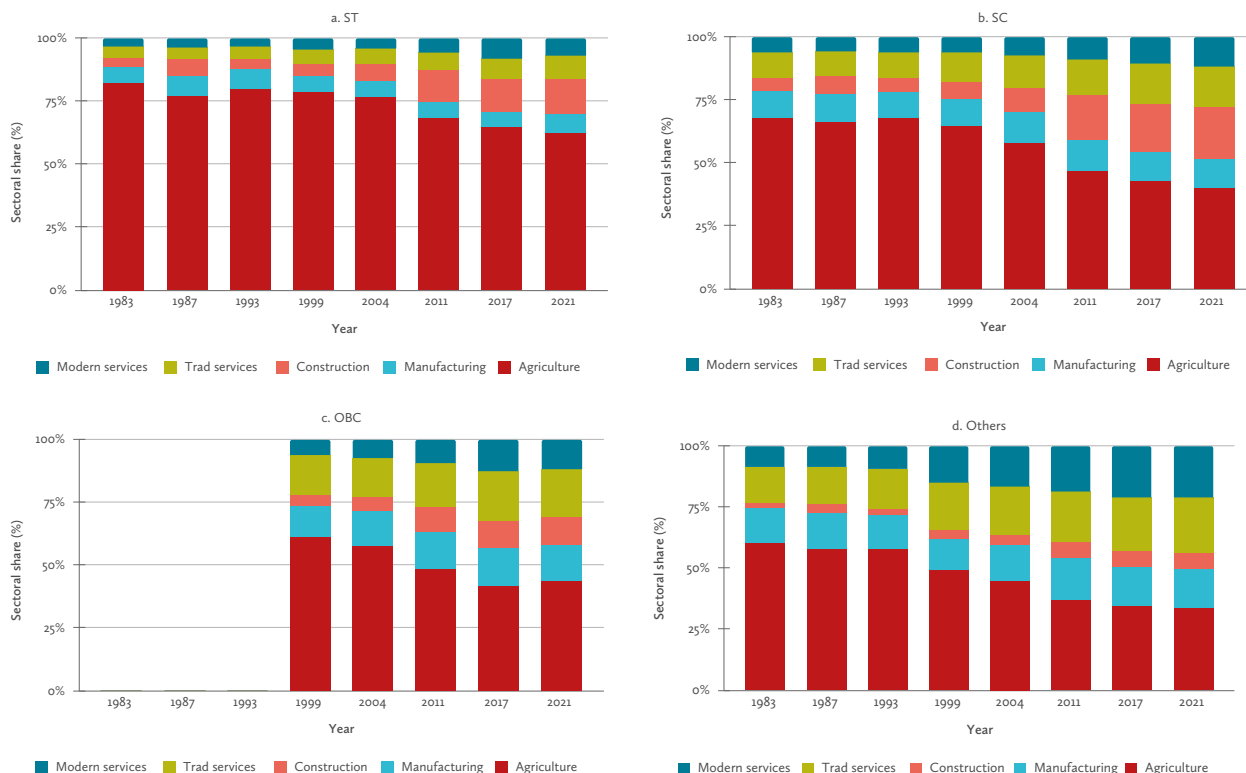
Sources and notes: NSSO EUS-PLFS various years.

Figure 3.14: Changing sectoral structure of employment – religion
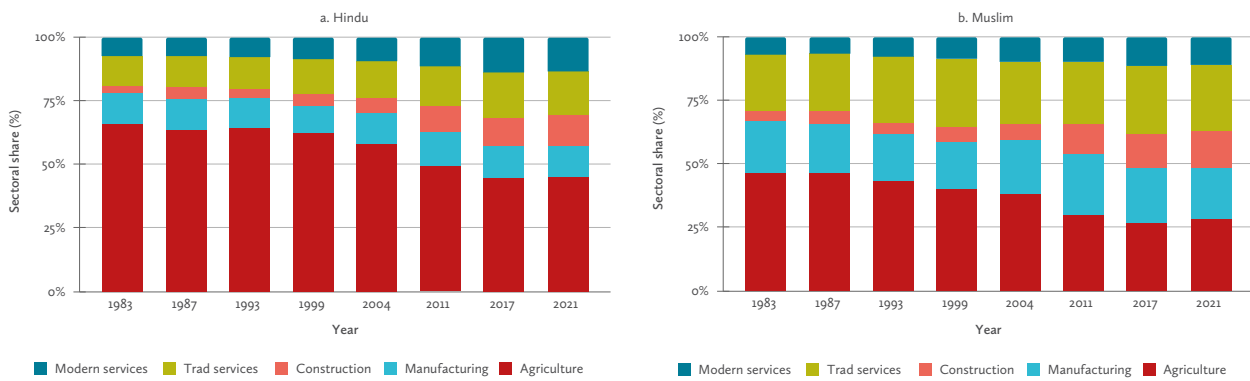
Sources and notes: NSSO EUS-PLFS various years.

plays a role in the second type of transition as well, as we shall see shortly, the first observation that jumps out from Figure 3.15 is that the pace of this change has been far slower than the movement out of agriculture, across all identities.

For example, the share of self-employment for men declined from around 57 percent in 1983 to 51 percent in 2011 before rising slightly to 53 percent by 2021. The corresponding figures for women are 56 percent in 1983 to 49 percent in 2017 before rising sharply (for reasons discussed in Chapter Two) to 60 percent in 2021. The interesting difference between men and women emerges when we look at the pace of increase of regular wage or salaried work, which shot up for the latter from 8 percent in 1983 to 24 percent in 2017. The corresponding numbers for men are from 19 percent to 24 percent, a much

smaller increase. Alongside this dramatic increase in the share of regular wage work, there has been a decline in casual wage work for women. Of course, it is worth recalling here that the composition of the female workforce itself changed during this period with the exit of older, less educated women who were mainly engaged in agriculture. Thus it appears that there have been two parallel movements as far as women workers are concerned - an exit of older, less educated women from self-employed work in agriculture and an entry of younger more educated women into regular wage work, mainly in modern services. We go deeper into this issue in Chapters Four and Five.

As with gender, so with caste groups, the decline in the proportion of self-employed workers has been slow and most of the redistribution in shares

Figure 3.15: Changing composition of employment type – gender
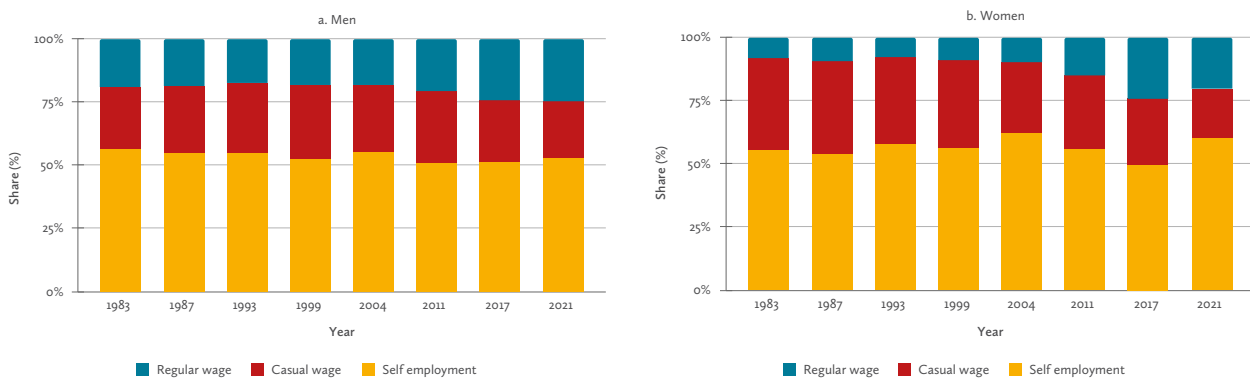
Sources and notes: NSSO EUS-PLFS various years.

Figure 3.16: Changing composition of employment type – caste
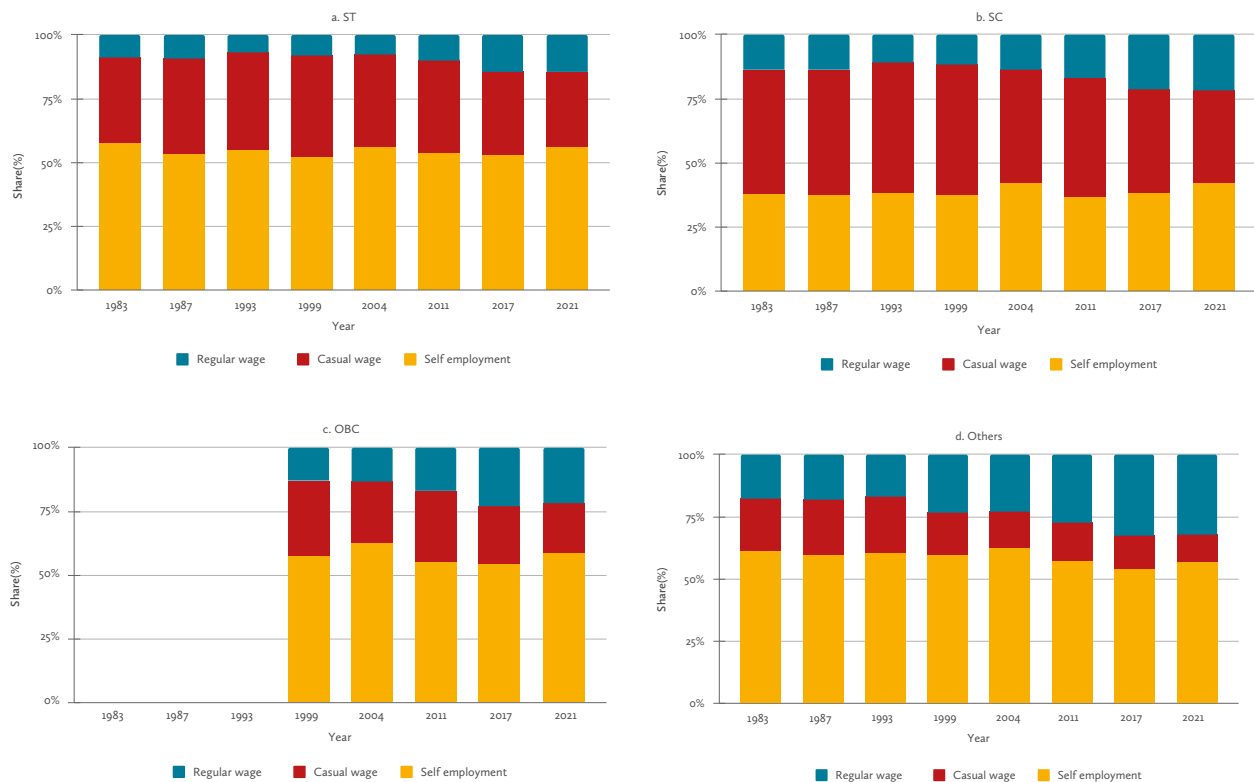
Sources and notes: NSSO EUS-PLFS various years.

is seen between casual and regular wage work with the former declining and latter increasing (Figure 3.16). Of course, differences do exist along expected lines between caste groups. For example, while 32 percent of Other caste workers are in regular wage employment, only 22 percent of OBC and SC workers and 14 percent of ST workers are in this type of work.

The time trends analysed in this section leave out one crucial factor - the extent of economic growth that occurred in any given period. For the perspective of economic policy, it is desirable to know the effectiveness of growth in promoting structural change. We now turn to this.

3.4 Effectiveness of economic growth in driving structural change

3.4.1 The India story

From Section 3.2 where we examined the Indian experience from a cross-country perspective, we learned that India's sectoral structure was largely in keeping with its GDP per capita, but the same could not be said for its employment structure. That is, the share of the workforce engaged in agriculture was close to that expected from the average relationship between GDP and agricultural share across all developing countries. But the share of the workforce engaged in self-employed and casual wage work was far more than expected. This suggests that the pattern of economic growth that India has experienced since the 1980s has created non-farm opportunities mainly of an informal nature.

Of course, this broad generalisation needs to be nuanced with the varied experiences of men versus women as well as different social groups. Further, the experiences of the states also need to be brought into the story. We now elaborate on each of these in turn.

The parameter that captures this relationship is the growth semi-elasticity of structural change. This is the percent point change in the share of the workforce engaged in a particular sector for a given amount of GDP growth. The semi-elasticity can be estimated in a simple regression framework. Table 3.2 shows these semi-elasticities at the all-India level for the period 1983-84 to 2018-19. The ability of economic growth to create non-farm employment opportunities can be quantified as follows.

A doubling of GDP per capita resulted in a drop in agricultural share of employment by 15.5 pct pts and a drop in share of informal work by just 6 pct pts.

The pace of exit from agriculture was such that a doubling of GDP would result in a drop in the agricultural share of employment by 15.5 percentage points. And an increase in the share of construction in total employment by 6.45 percentage points. In contrast, growth barely resulted in any increase in the manufacturing share in this period.

By themselves, these numbers are hard to interpret. How do we know if a 15 percentage point decline is good or bad? To answer this question, we need to have a comparison point. The contrast is provided by the semi-elasticity for the Lewis Process. Table 3.2 shows that a doubling of GDP would have reduced the share of self-employment and casual wage work by just under 6 percentage points (-3.06 for self-employment share and -2.7 for casual wage share). This provides another way to understand the disconnect between the Kuznets and Lewis processes, which has been the overarching theme in this chapter.

However, when we start looking at the experience of structural change disaggregated by gender and caste, some interesting differences emerge. Table 3.2 shows that the agricultural semi-elasticities are similar for men and women, but the semi-elasticities for construction and modern services are very different. The modern services elasticity for women is 7.5 as opposed to only 2.3 for men. Correspondingly, the same amount of growth caused a far larger increase in salaried share for

Table 3.2: Semi-elasticities of structural change at the all-India level

Sector	Overall	Male	Female	SC	Others	Hindu	Muslim
Agriculture	-15.55	-14.97	-15.85	-20.12	-14.69	-15.33	-15.30
Manufacturing	0.99	0.25	3.21	0.65	1.22	0.73	2.31
Construction	6.45	7.56	2.62	11.37	4.96	6.24	7.85
Traditional Services	5.00	5.23	3.47	5.20	5.09	5.13	2.88
Modern Services	3.58	2.27	7.46	3.34	3.80	3.67	2.92

Employment type	Overall	Male	Female	SC	Others	Hindu	Muslim
Salaried	5.75	3.95	10.57	5.92	6.05	5.95	5.06
Casual	-2.68	-0.41	-8.82	-5.81	-2.14	-2.84	-2.10
Self-employed	-3.06	-3.55	-1.75	-0.11	-3.91	-3.11	-2.96

Sources and notes- The numbers in the table are obtained from regressing sectoral or employment shares on log GDP per capita. They are to be interpreted as the percentage point change in the share of a sector in total employment with a doubling of GDP per capita. See Methods Appendix for details of the regression model.

women as compared to men (10.6 percentage points as compared to 4). And connectedly, women saw a far larger drop in share of casual labour as compared to men.

How do we square these results with the general impression that the Indian labour market has not been welcoming for women? The answer lies in the fact that economic growth resulted in a large withdrawal of some women from employment, mostly in the agricultural sector. These were women engaged in self-employment and casual labour. This withdrawal changed the composition of the female workforce along the lines described above. As we saw in the previous section, the share of women in salaried or regular wage work expanded rapidly in the high growth period.

Next we come to caste. Here we analyse caste in binary terms, Scheduled Caste versus the rest (i.e. General caste and OBCs). Scheduled Tribes are excluded from this analysis. The main point to note here from Table 3.2 is the larger negative semi-elasticity for agriculture for SCs and a corresponding larger positive elasticity for construction. Broadly, this captures the movement of SC workers from doing agricultural labour to doing construction labour. Though this generalisation does hide substantial variation with the SC group at the jati level, an issue we take up in the next section.

Finally, coming to religion, the differences here are more subtle than those observed for gender and caste. This suggests that the latter two may be the more important identities in explaining outcomes than religion per se. Intersectional differences (such as religion-gender or religion-caste) are not always possible to analyse with secondary data due to sample size limitation. But we do analyse this, to the extent possible, in future chapters.

3.4.2 The story of the States

Indian states are large and diverse economies in their own right and differ considerably in both average levels of GDP per capita as well as the structure of their economies. Though state-level growth experiences have been analysed for convergence or divergence, comparatively less attention has been focused on heterogeneity in structural change experiences. We now turn to this issue.

Our dataset includes all major Indian states. Since it is a long-run analysis going back to the 1980s, states created since 2000 have been merged with the state they used to be a part of (e.g. Jharkhand with Bihar). The northeast states and Jammu and Kashmir have been left out for reasons of inadequate sample size. This leaves us with 18 states and seven time points (1983-84, 1987-88, 1993-94, 1999-2000, 2004-05, 2011-12 and 2018-19). Though data is available at an annual frequency since 2017-18, we do not include all available PLFS rounds since this would weight the sample excessively towards the latter period.

Before we present the state-level semi-elasticities as we did for India as a whole, we look at the GDP-structural change relationship at the state-level graphically. Figure 3.17 shows the relationship between GDP per capita and agricultural share as well as the regular wage share of employment for all the states in our sample. The plot allows us to see the starting and endpoint variation in GDP per capita across states as well as the effectiveness of GDP growth in reducing the agricultural share and increasing the regular wage share of employment. Goa stands out in this and many other plots as a state that is both much richer than the rest and has a much larger formal economy. Of course, it is much smaller than the average Indian state and has a very different history, so the lessons to be drawn may be limited. But even setting the outlier states aside, it is clear that the lines have different slopes. It is this slope that is measured by the semi-elasticity.

We saw earlier that the rate at which growth has led to movement out of agriculture has been nearly the same for male and female workers. Figure 3.18a shows this to be broadly the case across states. In this figure, each point represents a state-year

Figure 3.17: Relationship between log per capita output and structural change across Indian states

a. Decline in agricultural share of employment

b. Increase in share of regular wage employment

Sources and notes: NSSO EUS-PLFS various years and RBI DBIE

combination. Red dots represent male and blue dots female workers (it is essentially the gender-disaggregated version of the plot in Figure 3.17 without the connecting lines to avoid clutter). The two lines have different intercepts but similar slopes, indicating that on average a larger share of women are in agriculture than men through the period but both are moving out at similar rates. When we examine the same relationship for caste groups, however, differences emerge. Figure 3.18b shows that the rate of exit from agriculture has been faster for OBC workers compared to workers from

other caste groups, such that we see convergence between them and the General castes. At the beginning of the period, in 1999 on average across states, the share of OBC workers who were engaged in agriculture was close to the share of ST workers so engaged (nearly 80 percent). By the end of the period in 2018, this had declined dramatically to only 20 percent on average while for ST workers it remained much higher at over 40 percent. This shows that the pace of the Kuznets process differs widely across caste groups.

Figure 3.18: Decline in agricultural share of employment with economic growth across Indian states

a. Gender

b. Caste
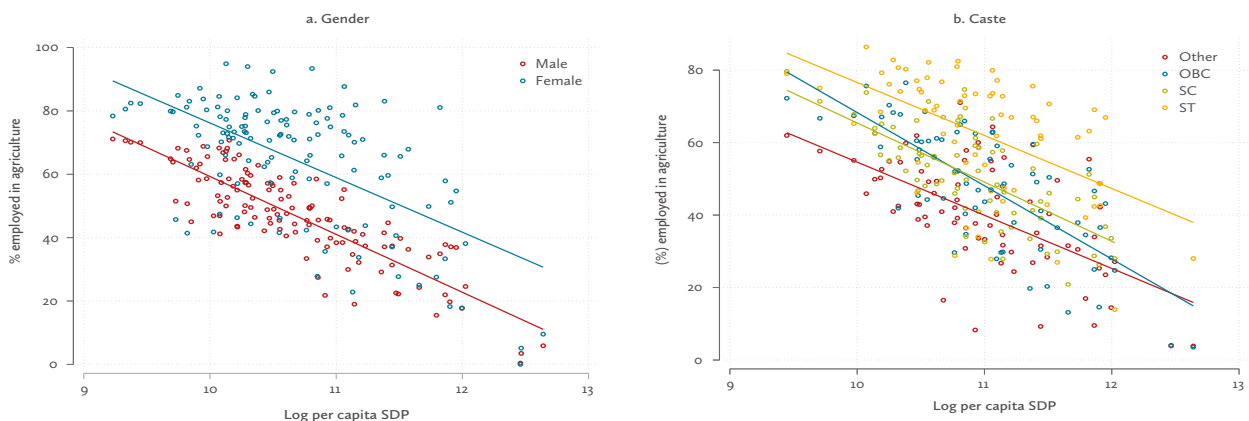
Sources and notes: NSSO EUS-PLFS various years and RBI DBIE

Figure 3.19: Increase in share of regular wage employment with economic growth across Indian states
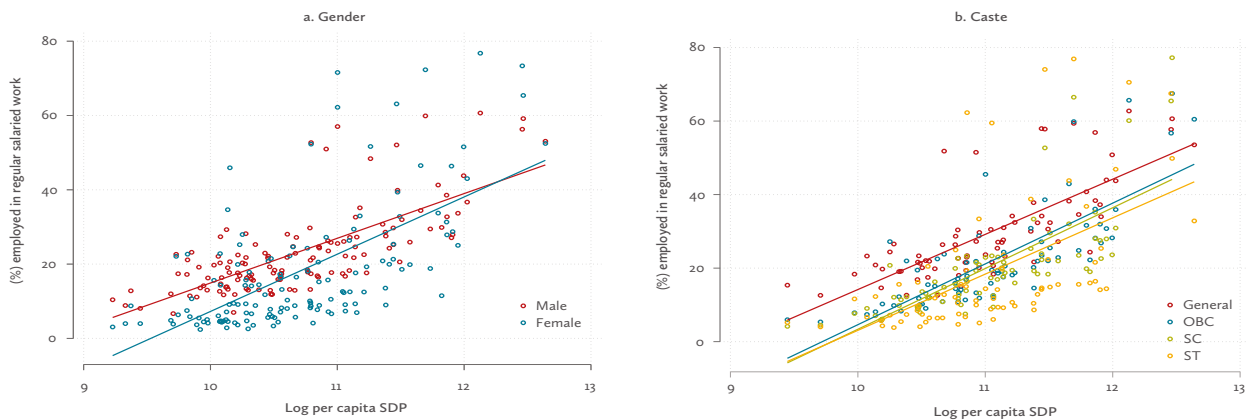
Sources and notes: NSSO EUS-PLFS various years and RBI DBIE

Coming to the Lewis process, we would like to examine the rate of increase in regular wage work at the state level. Once again, we plot each state-year combination disaggregated by gender and caste group (Figures 3.19). Note that for gender, women start lower than men and then cross over. We have seen this phenomenon before at the all-India level. This represents the rapid increase in regular wage work among women workers since 2004. For caste groups, something similar, albeit to a lesser extent, is seen for Scheduled Castes. The green line (SCs) starts below the blue line (OBCs) but then catches up to it. The modest convergence that we see here between SC and OBC workers is a welcome development, but at the same time, the convergence

with the General caste group is much weaker.

The cross-state regression framework (explained in more detail in the Methods Appendix) allows us to estimate the semi-elasticities for structural change for each state and for different social identities. Here we highlight a few notable results from this exercise.

Exceptions to the trend of declining workforce participation for women are Punjab and WB.

First, before we examine the changing structure of the workforce in terms of the Kuznets and Lewis processes, it is important to acknowledge the large

Figure 3.20: State-level differences in entry into and exit from the workforce with economic growth
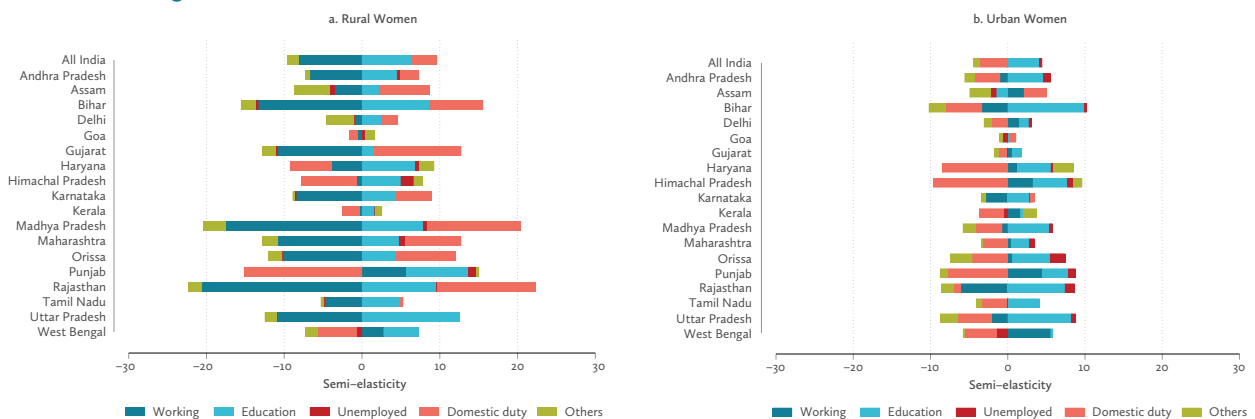
Sources and notes: NSSO EUS-PLFS various years and RBI DBIE The figure shows coefficients from a state panel regression. See text and Methods Appendix for details.

difference across gender in movements into and out of the workforce. Across states, women exited the workforce on average and entered education as well as were more likely to be doing housework (Figure 3.20).

In this and subsequent such graphs, the X-axis shows semi-elasticity. Thus a value of -20 for Rajasthan says that over the entire period the pace of withdrawal for women workers was such that a doubling of GDP would have led to a 20 percentage point decline in workforce participation rate.

This decline was compensated roughly to an equal extent by rising education enrolment and domestic duties. Interesting exceptions to the trend of declining rural workforce participation for women are the states of Punjab and West Bengal. These states saw a decline in housework and a rise in work participation. The West Bengal exception has been discussed in an earlier report (State of Working India 2018). The case of Punjab is more of a surprise and needs further investigation.

We can now examine the semi-elasticities for the Kuznets and Lewis processes across states for various identities. Figure 3.21 (b,c) does this for gender and the Kuznets process. The striking thing to note here is the large increase in modern services for women which is comparatively muted for men. For men, it is construction and traditional services which dominate. Of course, as we have been emphasising through this chapter, the compositional shift in the female workforce must be seen in the background of a large exit of older, lesser educated women from the workforce, leaving behind a very different demographic profile. A second feature worth noting is that states display considerable heterogeneity in terms of the value of the semi-elasticities ranging from a low of -7 for Gujarat to -28 for Punjab and even higher for Assam. Of course, one reason for this heterogeneity can be the starting level of share in agriculture, but clearly this is not the

only factor. Rather, the nature of economic growth has differed across states resulting in very different outcomes.

Coming to caste, Figure 3.21d and e shows the differences observed across SC and non-SC (General caste plus OBC) groups. We have not performed a state-wise analysis of ST groups here due to sample size limitations. The pace of exit from agriculture is higher for SC workers across states, which is likely to be a base effect. But the differences in terms of where these workers go are more important. While construction emerges as the main entry sector across states for SC workers, the experience is more diversified for other caste groups. But again there are some interesting exceptions we should note. Both Maharashtra and Gujarat, which have lower rates of exit from agriculture, show hardly any increase in construction for SC workers. Rather the movement is into traditional (and to a lesser extent modern) services.

We now come to the Lewis process. Figure 3.22 shows the semi-elasticities for type of employment in a by now hopefully familiar format. First note the overall muted nature of the process. Figure 3.21 and 3.22 have the same scale to illustrate the difference in magnitudes visually. Clearly, across states, the movement out of informal work is far slower than movement out of agricultural work. States such as Bihar have hardly shown any change in the employment structure, especially for male workers. The story is different for women though. Across states we see a decline in the share of casual work and a rise in regular wage work to a much greater extent than we see for men. The contrast for Bihar is especially striking, though it must be kept in mind that the female workforce participation rate for Bihar is in single digits. Punjab again offers an interesting variation from the general trend and shows a greater decline in self-employment compared to many other states. For caste, the story is one of a decline in self-employment and rise in

Figure 3.21: How much did a doubling of GDP per capita change sectoral structure of the workforce between 1983–84 and 2018–19?
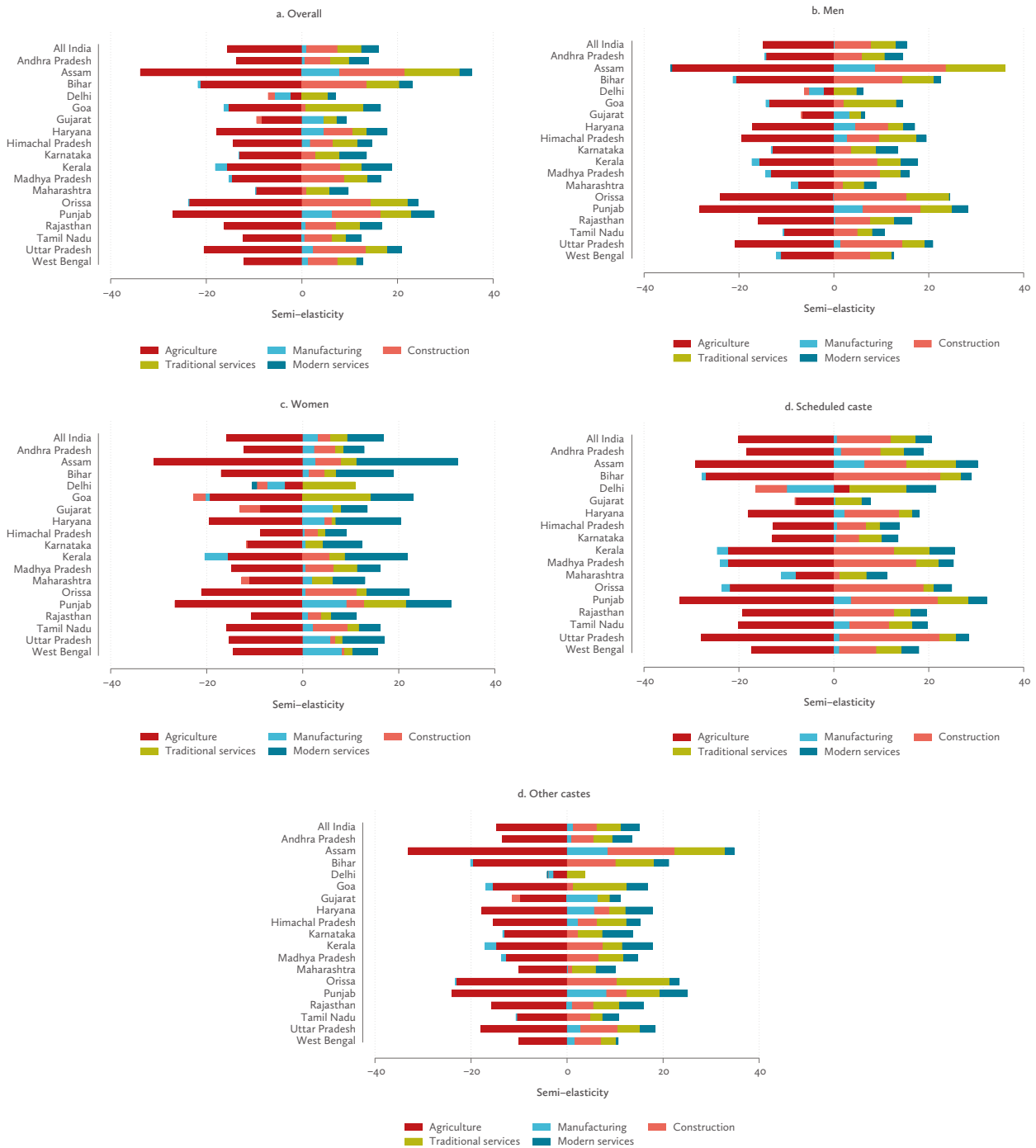
Sources and notes: NSSO EUS-PLFS various years and RBI DBIE The figure shows coefficients from a state panel regression. See text and Methods Appendix for details.

regular wage work for "other" caste groups and a decline in casual wage work and rise in regular wage work for SC workers.

We summarise the state-level contrast between the Kuznets and Lewis processes in Figure 3.23. These are scatter plots which show the two semi-elasticities for each state for all workers (a) and then

separately for male and female workers (b and c). The information here is the same as in the preceding bar graphs. It is being represented differently to appreciate the differences at a glance. Ideally, we would like to see all the states clustered high up in the top left indicating a robust decline in agricultural share and a rise in the regular salaried share. But the pace of change varies a great deal across states

Figure 3.22: How much did a doubling of GDP per capita change the employment structure of the workforce between 1983–84 and 2018–19?
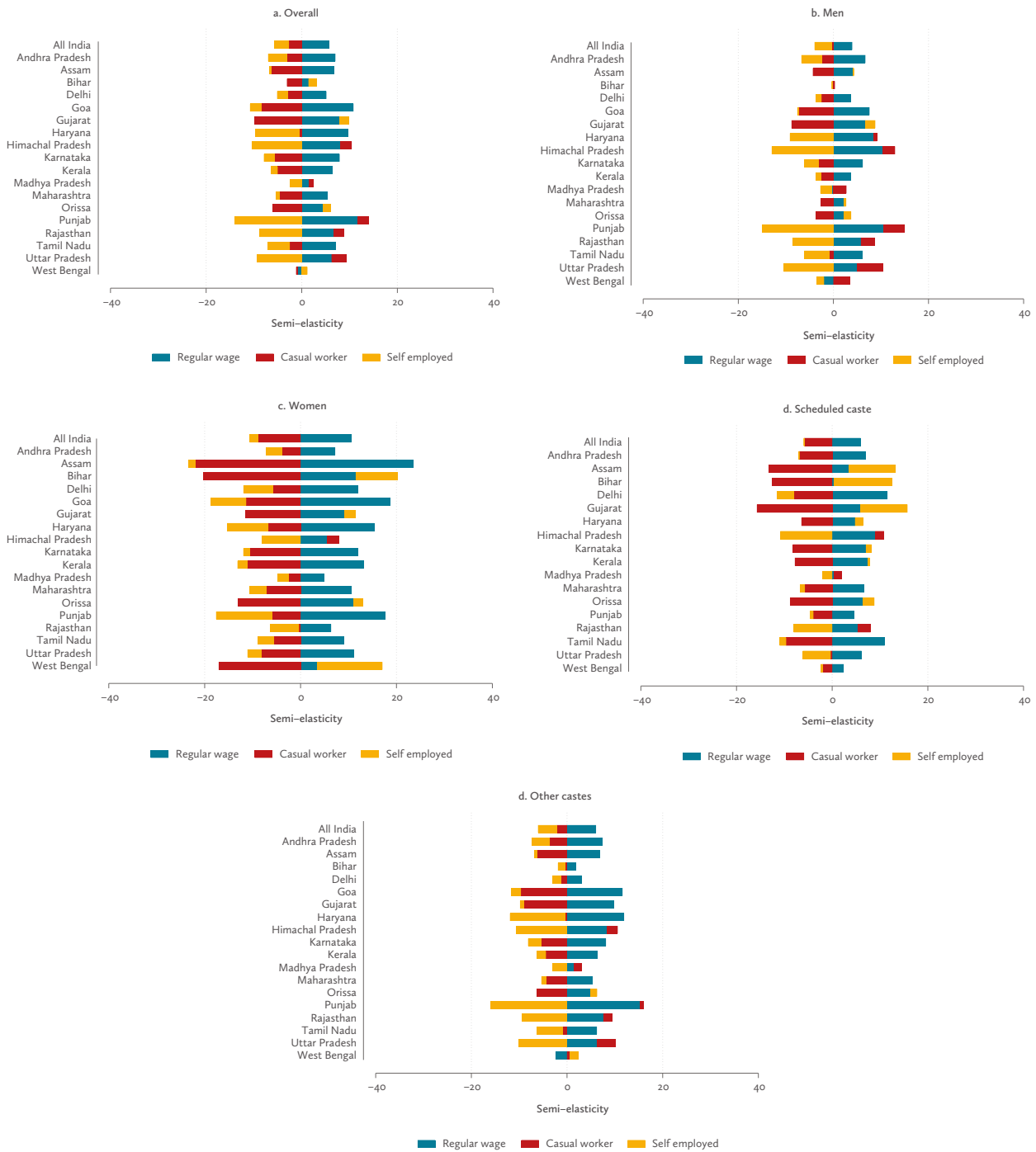
a. Overall

b. Men

c. Women

d. Scheduled caste

d. Other castes

Legend: Regular wage, Casual worker, Self employed

Sources and notes: NSSO EUS-PLFS various years and RBI DBIE The figure shows coefficients from a state panel regression. See text and Methods Appendix for details.

and as expected from the preceding discussion, the points are more or less flat for men but upwardly sloped for women.

3.4.3 Structural transformation at the jati level

In our discussion thus far on the experiences of structural change for caste groups, we have been forced to remain at the level of broad administrative categories. However, within-caste group variation at the level of jatis is significant and has its own story to tell. Unfortunately, as we discuss in greater detail in a separate chapter on Measurement, we generally

Figure 3.23: Bringing together the Lewis and Kuznets processes: extent of decline in agricultural share and increase in regular wage share with a doubling of GDP per capita across states.
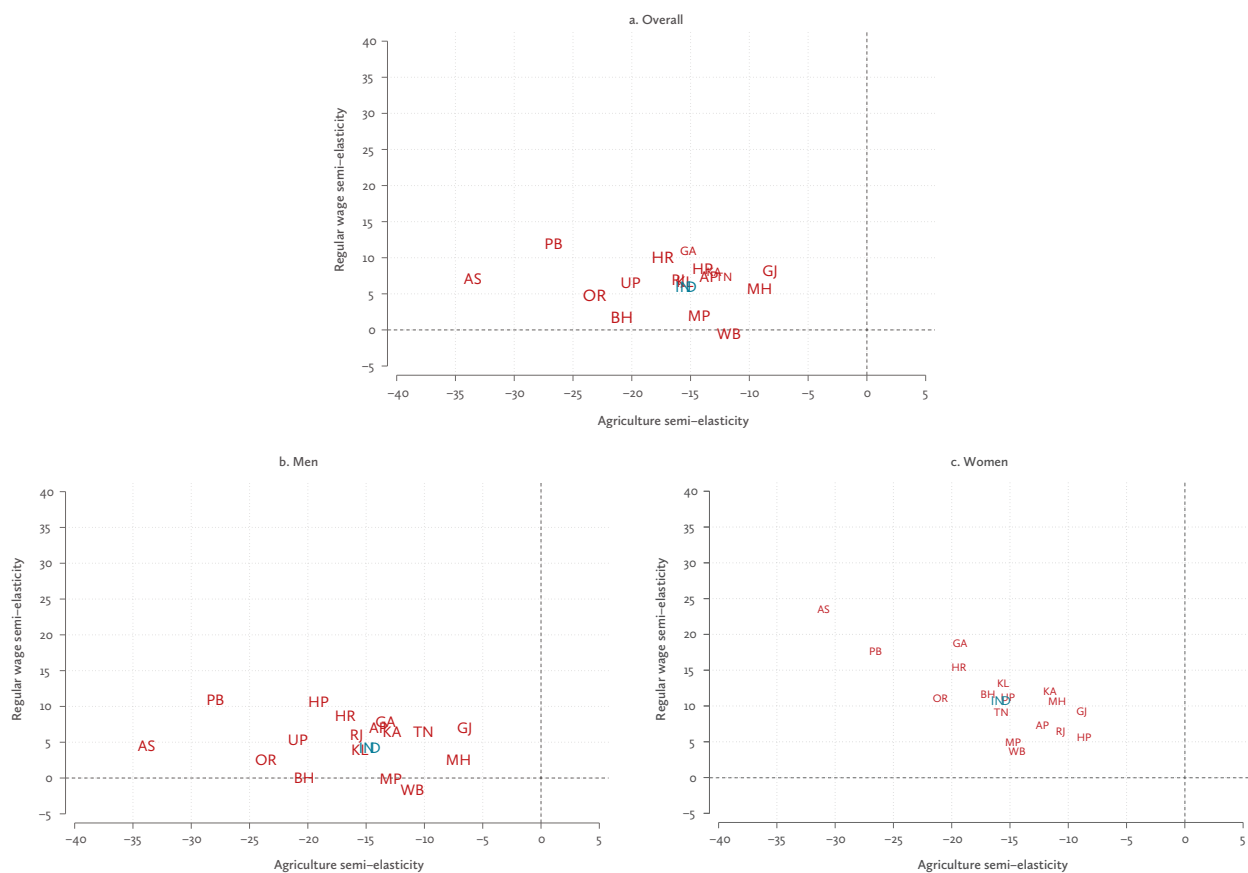
Sources and notes: NSSO EUS-PLFS various years and RBI DBIE The figure shows coefficients from a state panel regression. See text and Methods Appendix for details.

lack adequate data at this disaggregated level. In this section, we take advantage of the fact that the Census of India collects information on jati identity for individuals belonging to Scheduled Castes (SC) and Tribes (ST). Comparable information on both jati name and industrial category is available for the 1991 and 2011 Censuses. We use this data here (see Methods Appendix for details).

We restrict our attention to SC jatis and do not perform any sub-tribe level analysis for Scheduled Tribes. While sub-groups are important within the tribal identity, the nature of horizontal and vertical stratification for tribes differs in significant ways from that of castes. A further restriction is that we limit the analysis to jatis that constitute at least 0.5 percent of the SC population in a state. Finally, for most of our analysis, we focus on 12 states: Andhra

Pradesh, Bihar, Gujarat, Karnataka, Kerala, Madhya Pradesh, Maharashtra, Punjab, Rajasthan, Tamil Nadu, Uttar Pradesh and West Bengal. We do not comment on all India patterns. This is because of two practical difficulties. One, matching jati names across states is a difficult task prone to errors, and two, even for jatis that are easy to identify across states, such aggregation may not be desired from a social standpoint. That is, the same jati may experience very different trajectories across states owing to a multitude of factors. For example, the experience of Chamars in Tamil Nadu would be significantly different from that in Bihar, West Bengal, or Maharashtra.

We utilise information on six industrial categories of main workers[8]. These are: cultivation, agricultural labour, household manufacturing (HHI), non-

Table 3.3: Structural change at the jati level compared to Scheduled castes as a whole as well as other castes

a. Movement out of agriculture

State	Cultivators			Agricultural labourers		
	Median jati	SC	Others	Median jati	SC	Others
Andhra Pradesh	-5	-5	-10	-7	-8	0
Bihar	-10	-6	-18	-10	-10	2
Gujarat	-4	-4	-10	-9	-7	-1
Karnataka	-3	-5	-9	-10	-12	-7
Kerala	-1	-1	-7	-29	-33	-14
Madhya Pradesh	-12	-15	-15	2	4	6
Maharashtra	-1	-4	-6	-3	-3	-2
Punjab	-1	0	-12	-27	-28	-7
Rajasthan	-12	-13	-12	-3	-1	1
Tamil Nadu	-4	-7	-12	-12	-17	-8
Uttar Pradesh	-14	-16	-19	0	-2	3
West Bengal	-10	-10	-12	-3	-6	-1

Sources and notes - Population Census of India 1991 and 2011. The numbers indicates the percent point difference in shares of workers engaged in agriculture between the two time points. See text for explanation of the term "median jati".

b. Movement into non-agricultural sectors

State	Non HHI manufacturing			Construction			Services		
	Median jati	SC	Others	Median jati	SC	Others	Median jati	SC	Others
Andhra Pradesh	1	1	1	5	4	4	7	7	7
Bihar	2	2	2	5	5	3	10	9	10
Gujarat	6	5	3	6	7	3	2	2	5
Karnataka	1	2	1	4	5	3	8	9	10
Kerala	3	4	1	19	20	10	8	10	11
Madhya Pradesh	0	0	0	5	6	3	7	7	7
Maharashtra	-5	-1	-1	3	4	2	6	7	8
Punjab	3	4	1	12	13	5	11	10	10
Rajasthan	2	2	1	7	8	3	6	5	6
Tamil Nadu	5	6	3	6	6	5	10	11	10
Uttar Pradesh	1	2	1	4	5	3	9	10	10
West Bengal	1	1	-2	5	6	4	9	9	10

Sources and notes - Population Census of India 1991 and 2011. The numbers indicates the percent point difference in shares of workers engaged in indicated sectors between the two time points. See text for explanation of the term "median jati". HHI - Household Industry.

household manufacturing (non-HHI), construction, and services. Since manufacturing constitutes a very small share of SC workers, most of the analysis deals with the other categories. Additionally, for each state, we compute the distribution of main workers across these six categories for SCs as a whole as well as the Others (that is, for non-SC/STs). Note that while there has been movement out of agriculture, the Censuses of 1991 and 2011 record hardly any movement out of household-based manufacturing. We do not present this data here.

The first point to note is that, given the historically unequal distribution in land ownership and the resulting connection between SC jatis and agricultural labour, "movement out of agriculture" implies qualitatively different trajectories across caste groups: out of cultivation for non-SC/STs and out of agricultural labour for SCs. Table 3.3a shows that SCs move out of agricultural labour at a higher

rate than non-SC/STs and the trend is reversed for movement out of cultivation. The term 'median jati' in Table 3.3 indicates the jati that lies in the middle in terms of its share in a given sector. Note that over time the median jati may change. For example, in 1991, 63.4 percent of the main workers from jati Pallan in Tamil Nadu worked as agricultural labourers, and this jati was positioned at the median of the distribution. In 2011 this changed to 42.5 percent and the median jati was Adi Dravida.

Figure 3.24 is a set of box-and-whisker plots at the state-level. These plots show both differences across states as well as the variation at the jati level for two points in time, 1991 and 2011. We see that some states have higher rates of transition out of agriculture than others - e.g. Kerala, Punjab, Tamil Nadu, Karnataka and Gujarat. However, even within these states, there is substantial variation across jatis. For example, in both Karnataka and Tamil

Figure 3.24: Jati–level changes in the sectoral structure of employment between 1991 and 2011 for 12 Indian states
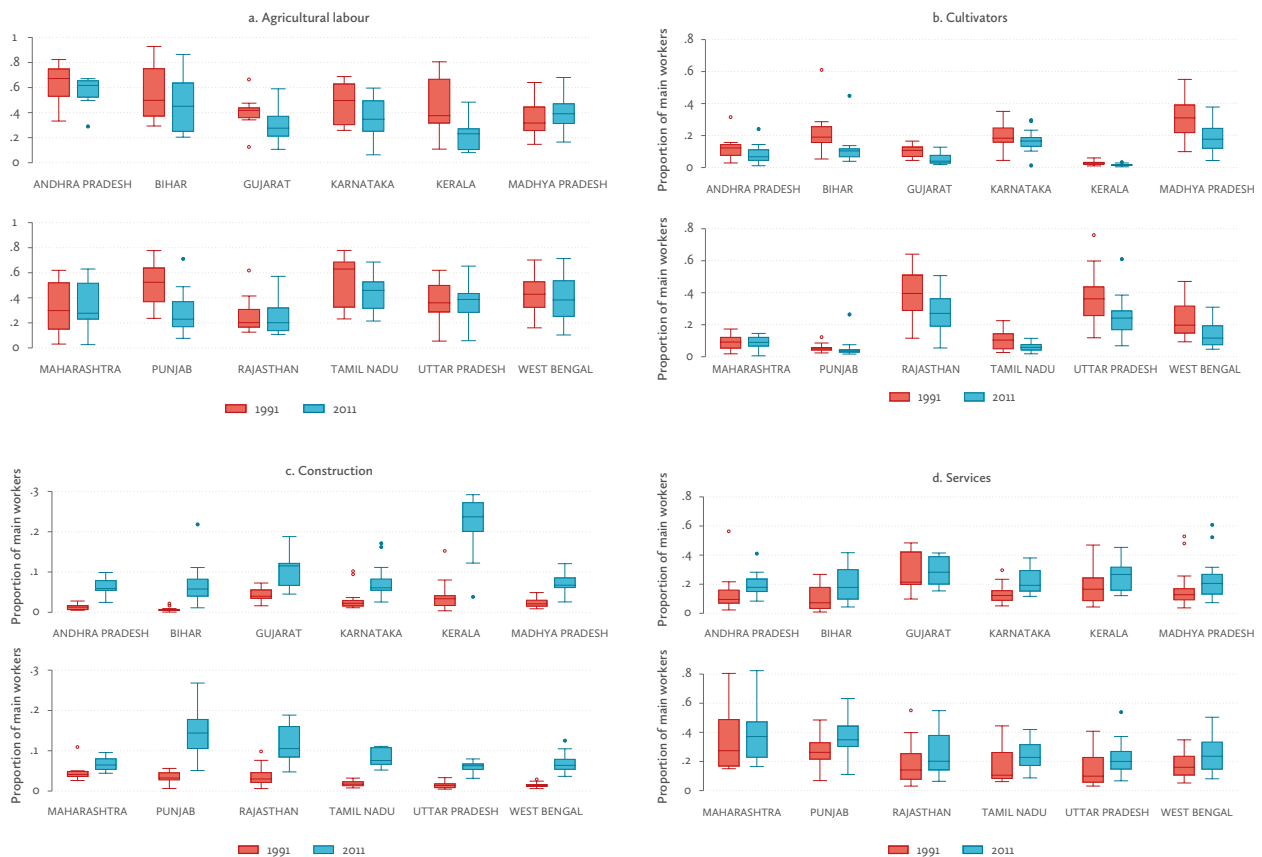
Sources and notes: Population Census of India 1991 and 2011. See text and Methods Appendix for details.

Nadu, the median share of workers in agricultural labour shifts downward at a relatively comparable rate (9 percent and 12 percent respectively). But in the case of Tamil Nadu the heterogeneity of experience across jatis, as captured by the height of the boxplot, is also decreasing, unlike the case of Karnataka where the overall heterogeneity remains the same. Figure 3.24b, when complemented with Figure 3.24a, shows that in Rajasthan, almost all SC jatis are moving out of cultivation and not from agricultural labour, a unique experience when compared to even the neighbouring states of Gujarat and Punjab.

The plots are useful but do not reveal information on any specific jati. We note, only in passing, that the same jati may experience different trajectories in different states. For example, Meghs or Meghwals experience very different outcomes across three neighbouring states. In Rajasthan and Gujarat, they had higher share among cultivators in 1991 and transitioned out of cultivation at a relatively higher rate, as opposed to in Punjab, where they had smaller share among cultivators, and did not experience any significant transition out of cultivation at all. Thus, the regional norms, land relations and social standing of a jati may interact with each other contributing to the unique trajectories of each jati, and therefore the heterogeneity of jati level experiences of structural transformation.

As with the overall picture discussed earlier in this chapter, however, a movement out of agriculture does not mean a movement into the formal sector. For all selected states, Table 3.3b reveals three key trends in terms of movement into non-agricultural work. First, the rate of movement into non-household manufacturing is slow for everyone, whether for the median Jati within the SCs, or for the larger groups such as SCs and non-SC/STs. Second, as one may expect, the construction sector draws a significantly higher share of SCs, and from all jatis within the SCs, while the non-SC/STs hardly exhibit any significant movement into construction

(Figure 3.24c). Third, there is significant movement into services for both SCs and non-SC/STs, however, unlike the case of construction, the experience of SCs as a group, does not translate to all SC jatis. That is, while the median jati may experience an upward movement, the distribution as a whole, does not necessarily shift upward as it does in the case of construction (Figure 3.24d).

In almost all the states, the non-SC/STs, and SCs as a whole and the median jati within SCs, exhibit more or less similar rates of movement into services (Table 3.4). However, there is a significant difference in terms of how this experience translates within the jatis. As Figure 3.24d reveals, for states such as Andhra Pradesh, Karnataka, Bihar, Madhya Pradesh, Punjab, and Rajasthan, we observe a relative upward shift in the distribution of SC jatis in services, however, we also observe that in some of these states (Bihar, Karnataka, MP, Rajasthan) the variance within the distribution has also increased, implying an increased heterogeneity across jatis. Such an observation is likely to be driven by some jatis moving into services at a higher rate than other jatis, and speaks to a possibly unequal experience of structural transformation within SCs.

States such as Gujarat, Maharashtra, TN, and UP, however, exhibit peculiar patterns. In all these cases, the median of the within SC distribution shifts, but the distribution as a whole, either does not shift upward at all points, or the variation within the distribution decreases. Taken together, this speaks to relative convergence in services for jatis within SCs.

One caveat in terms of movement into the service sector, is that it includes a wide range of work within it, and in the presence of caste-based rigidity, certain castes might be moving into services that are allied with, or linked to, their traditional caste occupations. Unfortunately, even though we can differentiate between types of modern and traditional services for the 2011 census, the 1991 census does not allow a similar exercise. Therefore, even if all jatis experience

an increase in employment within the service sector, we are unable to comment on the nature of services in which such increases happen, and if there is a significant association of these services with caste-based traditional occupations.

Finally, we note that although the movement into non-household manufacturing is slow for all groups, Gujarat and Tamil Nadu do show relatively the greatest increase, both in terms of the median (an increase of 5-6 percent), as well as at all points of the distribution. Of course, in both these states, the non-SC/STs also experience a 3 percent rise in non-HHI manufacturing (Table 3.4), unlike other states. Tamil Nadu and Gujarat are relatively more industrialised states, and the availability of manufacturing jobs may be drawing in workers from all castes alike (and the SCs at a relatively higher rate). While this evidence alone is not enough to conclude that a higher share of manufacturing industries will necessarily improve the rate of structural transformation for Scheduled Castes, it does call for an investigation of industrial policies of states and types of employment generated by these industries.

In terms of structural transformation, most of the discussion above, speaks more to the Kuznets process, that is movement out of agricultural work as opposed to the Lewis process (movement out of subsistence production). The latter is usually always difficult to capture with available datasets, but specifically so in case of jatis since we do not have nationally or regionally representative surveys on jati level characteristics of employment (salaried/casual etc). Thus, even the typical proxies are not available for understanding the Lewis process of structural change at the level of jatis.

3.5 Conclusion

In this chapter we have seen that India presents an interesting case of jobless growth coupled with weak structural transformation. The movement out

of agriculture has proceeded as expected overall and for various social groups. But the same is not the case for transition out of informality. The exception here is women workers. Due to the exit of older, less educated women workers mostly from agriculture, there has been a compositional shift in the female workforce and a rise in the share of regular wage work. In Chapter Five we examine which industries created the regular wage jobs that women entered. We also take a look at the determinants of regular wage work. The caste story is more muted, but also important. The convergence to General caste is seen for OBCs but not for STs. SCs present an intermediate case of a group which has exited agriculture but largely gone into construction. The last point to note is that the jobless nature of growth is clearly acting as a strong constraint on the creation of an adequate number of regular wage jobs. This remains true across states and identities. The next few chapters delve deeper into gender, caste and religion differences in determinants of type of work, industrial and occupational segregation and earnings disparities.

Endnotes

1   Absolute numbers are derived by multiplying survey derived ratios with official population projections based on the 2011 Census. Employment is defined based on the Current Weekly Status criterion. We have left out the openly unemployed who constitute a small fraction of the working age population. See Chapter Two and Methods Appendix for more details.

2   This section draws extensively on Dasgupta and Basole (forthcoming) and Basole (2022).

3   These are coefficients obtained from a cross-country fixed effects log-log regression between aggregate output (value-added) and aggregate employment as well as non-agricultural output and non-agricultural employment. The full regression results are provided in the online Appendix. For more details see Basole (2022). The results are not substantially different in the World Development Indicators database. We present analysis using ETD here since it gives more detailed sector-wise data on output and employment than does the WDI database. We have also verified that the aggregate and non-agricultural employment elasticities obtained here match those obtained from the India-specific KLEMS

database.

4   Aggregate employment elasticity (which we do not show in the figure) is lower at 0.2 than the elasticity of non-agricultural employment at 0.47. This is expected since agriculture has been shedding labour over a significant part of the period under analysis. From a structural change perspective, it is the non-agricultural elasticity that is more important.

5   As before the full regression results are displayed in the Appendix. And once again we have verified that the coefficients obtained in ETD are the same as those obtained in KLEMS.

6   A corollary of this lack of relationship between GDP growth and output growth is a one-for-one relationship between output growth and labour productivity growth. Note that labour productivity is defined as output divided by employment. So, the weaker the relationship between output and employment, the stronger is the relationship between output and labour productivity. Thus we find that the coefficient on labour productivity (the Kaldor-Verdoorn coefficient) is slightly greater than one. Note that generally the Kaldor-Verdoorn coefficient is expected to be around 0.7. That is, productivity growth does not rise one-for-one with output growth, allowing some correlation between the former and employment growth. But this has clearly not been the case for India over the period in question.

7   See https://cse.azimpremjiuniversity.edu.in/on-the-misclassification-of-employment-status-data-in-the-ilostat-database/ for details.

8   Main workers are defined in the Census as those who work for at least six months in the year.

4

What determines women's employment in India?

4

What determines women's employment in India?

In the preceding chapter we took a long view of the process of structural transformation in India as it has unfolded at the aggregate level as well as for key social identities. We saw that the dynamics of women's entry and exit from the labour market played a crucial role in understanding employment trends in the Indian economy over the past four decades. In this chapter we take a closer look at what determines whether women participate in paid work by bringing together evidence from primary surveys and secondary data.

4.1 Role of norms in shaping female labour supply: a review of the literature

4.1.1 Supply-side and demand-side considerations

Before we present our analysis, we briefly examine some conceptual issues around women's participation in paid work. The key thing to keep in mind is that the observed level of employment is a result of both the level of labour demand and the level of labour supply. In the case of women, both the demand for labour (i.e. employment opportunities) and the supply of labour (i.e. availability of women for employment) are impacted by social norms and beliefs. On the demand side, beliefs regarding appropriate types of work can shape employer hiring preferences and therefore demand for women's labour. Even

if direct employer-side discrimination is absent, previous period's segregation can propagate over time due to exemplar effects. Further, the male breadwinner norm means that a fixed amount of labour demand is first met with the supply of male labour, and only if men are not available, are women hired for the job. In the context of jobless growth, such as we discussed in detail in Chapter Three, the consequences for women's employment are easy to see. It pushes women out of the workforce, raises female unemployment rates, or traps them in low-productivity informal work. We see all three outcomes in the Indian case.

On the supply side, norms around responsibility for household work, marriage and motherhood norms, and norms around decision-making within the household as well as mobility outside it, profoundly influence availability of women for work, particularly work outside the home. This is why home-based employment, whether of the own-account, piece rate or contributing family worker variety is so prevalent among women. Another way in which norms can impact labour supply is via the threat of sexual violence and beliefs around "family honour" which can lead households to limit women's mobility. Note that even if a woman does not face mobility restrictions from within the family or community, she may still be unfree because of larger concerns around violence.

The demand and supply side factors discussed above can result in low *levels* of female workforce participation at a given point in time. To this we must add a set of factors that can affect the change in levels over time. The male breadwinner norm

ensures that rising male incomes are accompanied by falling female work participation. That is, women work only as long as a certain minimum living standard is not met by male earnings. This 'income effect' says that household incomes are crucial in determining women's participation (Kapsos, Bourmpoula, and Silberman 2014; Abraham 2013; Neff, Sen, and Kling 2012; Rangarajan, Kaul, and Seema 2011). When household incomes are low, women enter the market and exit when they rise. Further, in India, women working outside the home is often seen as a sign of lower social standing. As household incomes increase, norms against women may become more rigid and families may withdraw women from the workforce (Bussolo et al. 2022; Fletcher, Pande, and Moore 2017; Srinivas 1977). Though the two effects (income effect and norms effect) are analytically distinct, in practice it is hard to separate them.

U-shaped relationship is observed between women's employment and economic growth. Structural transformation initially pushes women out of the workforce as demand for women's labour in agriculture falls. As growth continues and female education levels rise, new employment opportunities and higher wages pull women into the workforce.

Several studies have established a consistently negative relation between household earnings and women's participation in paid work (Mehrotra and Parida 2017; Klasen and Janneke 2013). But the above phenomenon is not monotonic. Rather a U-shaped relationship between women's employment and economic growth has been observed (Goldin 1994). When agricultural activities predominate and household incomes are low , women are likely to be engaged in farm activities either as wage labourers or contributing family workers. However, as an economy grows, the sectoral composition of activities begins to shift from agriculture to manufacturing and services. This process of structural transformation initially pushes women out of the workforce. Here both

the income and norms effects described earlier as well as demand-side considerations are relevant as the demand for women's labour in agriculture falls due to mechanisation of farm activities. As growth continues and female education levels rise, new employment opportunities and higher wages pull women into the workforce. In microeconomic terms, this is when the substitution effect starts to dominate over the income effect. That is, women increasingly substitute paid work for unpaid work as the opportunity cost of not being employed rises.

4.1.2 Penalty for marriage and motherhood

Gender norms can be highly persistent over time. This is clear when we consider the "motherhood penalty" in developed countries, where female labour force participation rates have been much higher than in India for several decades. Norms that make women the primary caregiver in the household have been slow to change with the result that women incur a penalty after motherhood in terms of access to employment as well as earnings (Kleven et al. 2019; Lundborg, Plug, and Rasmussen 2017; Angelov, Johansson, and Lindahl 2016). Paradoxically, evidence for a motherhood penalty from developing countries is more mixed (Kleven et al. 2019; Aaronson et al. 2017) . Aguero et al. (2020) use Demographic and Health Survey (DHS) data from 21 developing countries to show that the penalty is larger in middle-income countries than in lower-income countries. In lower-income countries, there is a greater concentration of women in agricultural activities or in the informal sector, which allows them to combine childcare responsibilities with employment. With economic development, as labour markets become formal without accompanying changes in supporting infrastructure such as childcare facilities, combining childcare with employment becomes more difficult.

There are few studies that examine the motherhood penalty in the Indian context. Based on cross-sectional data, Das and Zumbyte (2017) find that

women who live in households with younger children tend to participate less in paid work. But this is offset to a certain extent when there are older women in the household, presumably due to a sharing of care-giving responsibilities. These findings are indirectly reinforced by Khanna and Pandey (2021) using panel data from the India Human Development Survey (IHDS) collected in 2004–05 and 2011–12. Their results show that the death of a co-resident mother-in-law negatively impacts women's labour supply, particularly for women with four or more children. Using the same IHDS data, Mukherjee and Sarkhel (2021) find a negative association between young children and women's labour market wages and working hours.[1]

Gautham (2022) uses Indian Time Use Survey data to examine the difference in patterns of employment and time use among mothers and non-mothers. Comparing between married childless women and married women with one child, she finds that employment is lower among women with children by nine percentage points in urban areas and two percentage points in rural areas. The urban-rural difference can be explained by differences in the type of work - urban women are more likely to be in wage employment that does not allow for childcare responsibilities to happen concurrently.

Deshpande and Singh (2021) using panel data from the CMIE-CPHS (2016-2019) find that women show frequent transitions in and out of the labour market. Using entropy balancing, the authors compare between individuals who have had a new child in the period of the survey and a similar demographic of individuals who have not. They find that although the former, in general, have a lower employment rate, there is no effect around the time of child birth. Instead, women, perhaps in anticipation of child birth, leave the workforce early, and remain out. Therefore, there is no post-childbirth penalty in terms of women's employment.

Clearly, norms-based constraints on women's work operate very differently across the income and class spectrum. Among highly educated women, norms seem to loosen. However, for poorer households and less educated women in these households, the need to contribute to the family budget implies that employment is not a choice (Abraham 2013). The imperatives of employment imply that withdrawal from work during motherhood and around childcare is often a luxury. Later in this chapter we show the results of a pseudo-panel exercise created using retrospective data as a part of the India Working Survey. We explore how imperatives of marriage, motherhood and employment play out for women in two states in rural India.

4.1.3 Employment and male backlash

Finally, we note that participation in paid work can interact with existing norms to produce undesirable outcomes for working women. In Rwanda, Finnoff (2012) found that employed women experienced partner violence more often than non-working women. Using data from the National Family Health Survey (NFHS -3) a part of the cross-country DHS surveys, Paul (2016) found the same result in India. These studies support the theory of "male backlash", higher violence that working women face as a backlash to the relative disempowerment of men in the household (Jayaraman 2023; Dhanaraj and Mahambare 2022) . Biswas and Thampi (2021) find a positive relationship between women's employment and spousal violence for all except the top wealth quintile using NFHS-4. Mondal and Paul (2021) find that working women with low levels of education and poor economic backgrounds were more vulnerable to experiencing violence.

The foregoing brief review of existing work serves to illustrate the complex interplay of gender and caste norms, economic growth and labour demand that shapes the extent and pattern of women's

employment. One aspect missing from the foregoing is the role of public policy in shaping the nature of growth, the type of labour demand and norms themselves. This requires a detailed evaluation outside the scope of the report.

4.2 Factors determining the supply of women's labour

We now present evidence on the ways in which social norms impact women's labour supply. We start with aggregate trends showing that caste and religious identities matter for female WPR. We then proceed to the results of a linear probability model that identifies the determinants of women's workforce participation, especially highlighting the role of caste, religion, education, husband's earnings and the employment status of a co-resident mother-in-law. Next, we present the results from the India Working Survey that shed light on the effect of marriage and childbirth on women's employment. We find that, in the context of rural Karnataka and Rajasthan, these events raise the employment rate instead of lowering it. We also show the importance of women's autonomy in determining their employment status, again drawing on findings from the India Working Survey. Next we analyse nationally representative National Family Health Survey (NFHS) data for 2015-16, to show

the importance of norms in determining women's employment status controlling for labour demand and other relevant variables.

4.2.1 Trends in women's employment by caste, religion, and education

Figure 2.1 (see Chapter Two) illustrates two distinct challenges with respect to women's employment - a decline in the rural WPR and a stagnant, low urban WPR. That is, economic growth in India has been accompanied by a withdrawal of women from agriculture but has not succeeded in pulling women into the non-farm workforce to the same extent.[2]

The decline in women's employment since 2004 was witnessed across caste and religious groups.

The first indication that social norms may play an important role in determining women's employment outcomes comes from looking at trends disaggregated by caste and religion (Figure 4.1). In general, women from SC, ST, and OBC castes have higher employment rates compared to Other caste women. As of 2021-22, about 40 percent and 25 percent of ST and SC women, respectively, were in employment compared to only 21 percent of others.

However, it is worth noting that the decline in women's employment since 2004 has been

Figure 4.1: Long-run female workforce participation rates by caste and religion
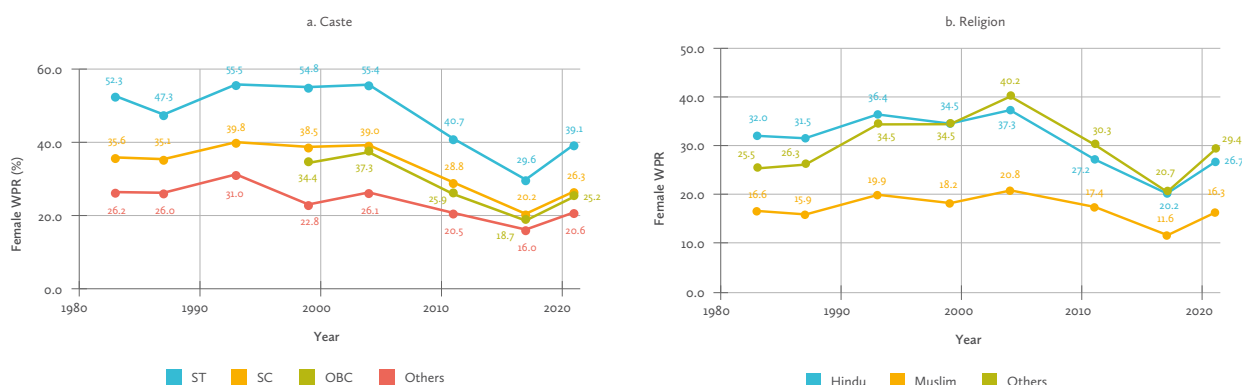
Sources and notes: NSSO EUS and PLFS various years. Employment is defined as per Current Weekly Status.

witnessed across caste and religious groups. For instance, for ST women, employment rates fell from 55 percent in 2004 to 30 percent in 2017 and subsequently increased to about 40 percent in 2021. For non-SC/ST/OBC women, the corresponding fall was from 26 percent to 16 percent and then an increase to 21 percent. All through, the relative employment rate of STs compared to others has remained unchanged. We note a similar parallel trend in the case of Hindus, Muslims and women from other religions. Muslim women have lower WPR than Hindu and Other women over the entire period.

Of course, the level differences seen across caste or religious identity need not be solely due to gender norm-related differences within a community. Historical disadvantage or exclusion as well as active discrimination on part of dominant communities can affect the ability of a household to educate girls as

well as the ability of women to find work. And once again, there is a U-shaped relationship observed between education and employment. As education levels increase, the likelihood of women being employed first decreases, until the diploma or degree level, after which it rises (Mehrotra and Parida 2017; Klasen and Pieters 2015).

Figure 4.2 shows that there exists a U-shaped relationship between the level of education and the employment rate (WPR). This is consistent with the earlier discussion on income and substitution effects. Though the figure only shows this for the latest PLFS round, the U-shaped relationship exists for all survey years starting from 1983. With respect to time trends, we should also note that behind an overall stagnant WPR for urban women, there lie opposing trends, with the WPR increasing for literate and primary/middle school educated women and falling for more educated women (data not shown).

Figure 4.2: Female workforce participation rates by education for 2021–22
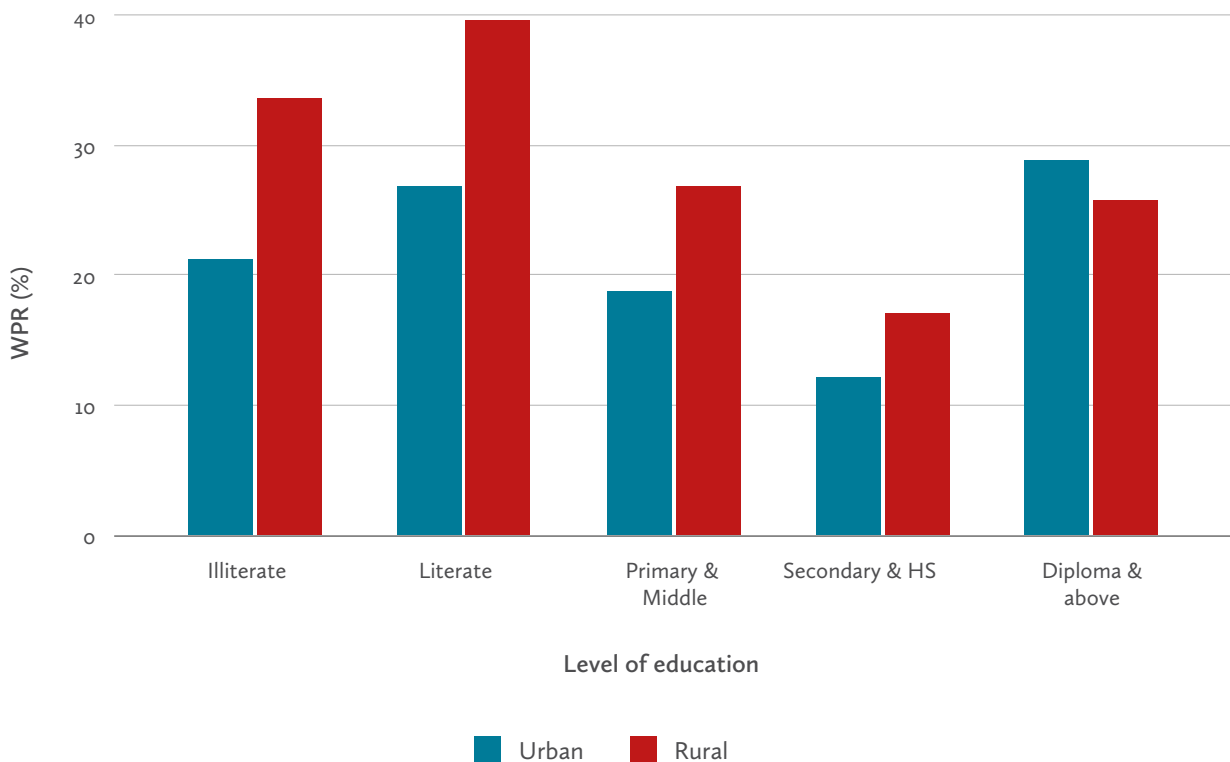
Sources and notes: PLFS 2021-22. Employment is defined as per Current Weekly Status.

4.2.2 Correlates of women's employment -an analysis of PLFS 2021-22 data

The caste, religion and education differences shown in the previous section only provide an initial indication that norms may be important. But income and wealth also vary across these dimensions as do other important variables such as fertility rates, all of which also determine women's employment outcomes. To disentangle these various effects, we need to adopt a regression framework.

We estimate a linear probability model of employment participation of women (with the dependent variable taking the value of 1 if the woman is employed, zero otherwise) with a set of controls for individual, spousal and household characteristics.[3] We consider all types of paid work including own-account work, unpaid work as a

contributing family worker, casual wage work and regular wage work.

Caste, religion and education

Figure 4.3 shows the marginal effects of selected characteristics on the likelihood of the woman being employed for the most recent available PLFS round (2021-22). The complete regression results are provided in the Results Appendix.

The numbers on the graph show how much more or less likely to be employed a woman is compared to a woman in the base category - e.g. a married woman in urban India is 10 percent less likely to be employed compared to an unmarried woman (coefficient = -0.1). A marginalised caste Hindu woman is more likely to be employed compared to an dominant caste one in both rural and urban areas. While for Muslims in rural areas, neither

Figure 4.3: Impact of marital status, caste, religion and education on probability of being employed for women
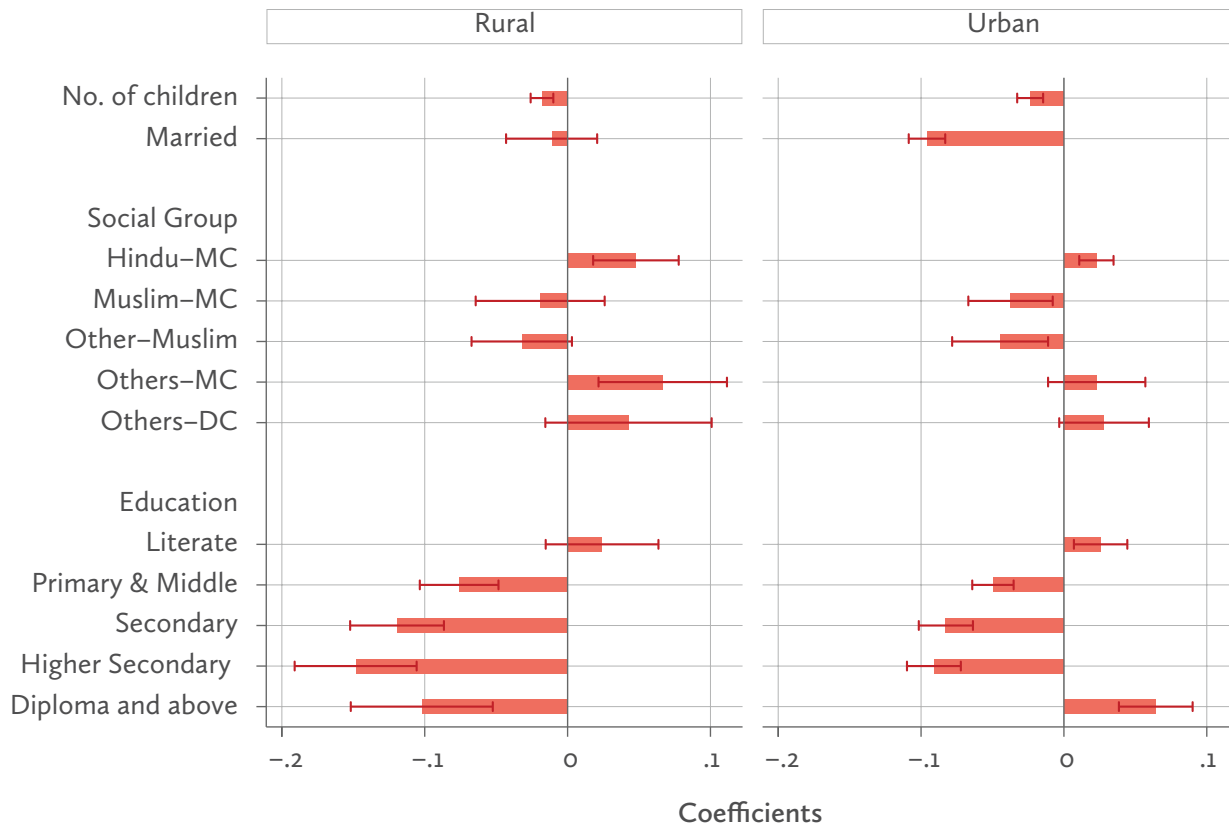
Sources and notes: PLFS 2021-22. Each bar gives the magnitude of the marginal effect of a variable on the probability of being employed. Values where the confidence interval bars do not intersect with the horizontal zero line are statistically significant.

marginalised caste nor dominant caste women are significantly different as compared to dominant caste Hindu women.[4] However in urban areas, we do see significant differences - marginalised and dominant caste Muslim women are less likely to be employed compared to the base category of Hindu dominant-caste women. And this difference is even more compelling if we compare only marginalised caste Hindu and Muslim women in urban areas. That is, controlling for caste, level of education, and other relevant characteristics, urban Muslim women are much less likely to be employed than Hindu women. Of course, recall that employment is a result of both supply and demand side factors. Thus, from this analysis alone, we cannot say whether the difference observed is due to gender norm differences affecting labour supply or discrimination that reduces the demand for Muslim women's labour.

Coming to education, the U-shaped relationship alluded to earlier is clearly seen in urban areas, even after controlling for various demographic variables as well as for the education level of the head of the household (a proxy for the household's economic status). The base category here is illiterate women. As the education level rises, at first the probability that a woman is employed falls with respect to the base category. Thus, for instance, in rural areas, women with higher secondary education were nearly 15 percentage points less likely to participate in employment compared to illiterate women (coefficient = -0.15). But, once we pass the higher secondary level, the relationship changes in urban areas, where diploma or degree holding women have higher employment rates than illiterate women (the base category).[5]

Husband's earnings and mother-in-law status

For most women in India, decisions around employment are seldom their own, but rather made by or in conjunction with household members - parents, spouses and/or in-laws. Household structures directly and indirectly affect women's mobility and decisions around employment

(Jayachandran 2020; Dhanaraj and Mahambare 2019). Using PLFS 2021-22, we estimate a similar women's employment participation regression model as above, but restricted to only married women to examine how intra-household dynamics affect women's work. We include variables to capture the husband's earnings and the presence of a mother-in-law in the household.

As noted in Section 4.1, a rise in husband's earnings is likely to reduce the likelihood of women being in paid work. However, since norms operate differently across the income spectrum, this relationship may not be linear. We introduce both husband's earnings as well as a squared term to capture possible non-linearity. Figure 4.4 shows the effect of husband's earnings on the probability of women's employment for rural (a) and urban (b) areas estimated from our regression (i.e. controlling for age, education, region, state and other factors). A few points are worth noting. First, the overall probability of women working is higher in rural areas as we expect (note the difference in scales between the two plots).

Presence of a mother-in-law who is employed is associated with a higher likelihood of the woman (daughter-in-law) participating in paid work

Second, initially, with rising husband's earnings the probability of the wife being employed decreases in both rural and urban areas. Third, the fall slows down but does not reverse in rural India, while for urban areas, there is a clear U-shaped pattern. As husband's earnings keep rising, from approximately ₹40,000 per month onwards, there is an increased likelihood of wives being employed. This U-shaped pattern could result from a change in norms with rising incomes or it could also be due to the fact that such husbands are matched with higher educated wives who have preferences for as well as opportunities to access better paid work.

The presence of a mother-in-law in the household may influence a woman's employment status in

Figure 4.4: Relationship between husband's earnings and probability of women's employment
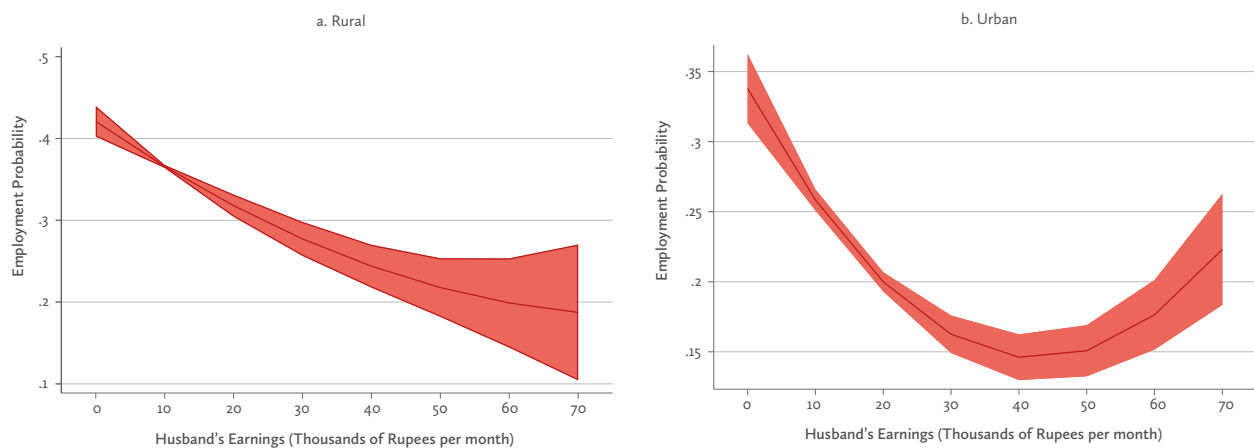
Sources and notes: PLFS 2021-22. Shown here are predicted values of the employment probability for each level of monthly income along with the confidence interval band.

two different ways. On the one hand, it could mean sharing of childcare responsibilities allowing women to pursue paid employment. On the other hand, it may imply the presence of greater restrictions around mobility and autonomy including employment (Khanna and Pandey 2021; Anukriti et al. 2020). We explore this issue by comparing employment outcomes for three different kinds of women - those who live in households without a mother-in-law present, those who live in households where the mother-in-law is present and also employed, and those where she is present but not employed.

We find that compared to households where there was no mother-in-law present, married women living in households where the mother-in-law was present but not employed were less likely to be employed themselves. However, if the mother-in-law was present and employed herself, this was associated with a higher likelihood of the woman herself being employed (Figure 4.5). Note that, here too, both income and norms effects may be playing a role. Low income households require additional earnings mandating all adults (men and women) to participate in employment, and norms around women's employment may be looser when another senior female member is also employed (see Box 4.1).

The inclusion of husband's earnings in the regression model means that the income effect is accounted for to some extent.

A few key takeaways from the above analysis are worth repeating. We find clear evidence that caste as well as religion continue to matter in determining whether a woman is employed. Further, there is strong evidence for the income effect in the form of the U-shaped curve that we find for husband's earnings, particularly in urban areas. Lastly, the fact that the employment status of the mother-in-law plays a significant role in the daughter-in-law's status suggests the norms reproduce over generations and that therefore, changing employment outcomes today can have strong effects on future outcomes (Schmitz and Spiess 2022; Farré and Vella 2013)

4.2.3 Marriage, motherhood and employment : Evidence from a Life History Calendar[6]

The results presented above are suggestive of the importance of household structures and norms. But they present a static picture at a point in time. Ideally, to understand long term transitions of women in the labour market, in particular, the impact of marriage and childbirth, we require

Figure 4.5: Impact of mother-in-law status on probability of being employed for women
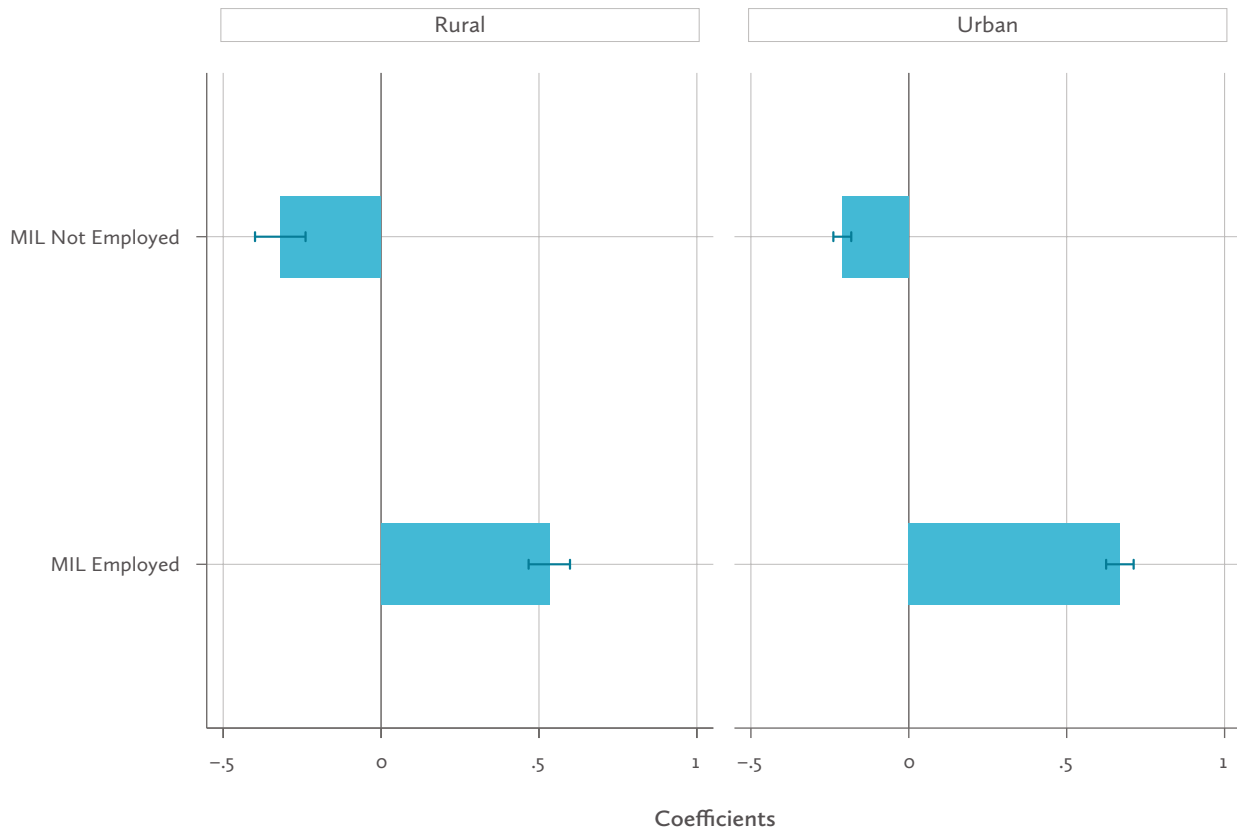
Sources and notes: PLFS 2021-22. Each bar gives the magnitude of the marginal effect of a variable on the probability of being employed. Values where the confidence interval bars do not intersect with the horizontal zero line are statistically significant.

long-term panel surveys that track the same individual over years. In India, and in most low and middle income countries, such data is generally not available.[7] An alternative method is to collect retrospective data from respondents by asking them to recall important life events. The Life History Calendar is such a method where respondents provide autobiographical information across various domains for a specified period that is determined by the research question (Morselli, Le Goff, and Gauthier 2019).[8]

Using this approach, as part of the India Working Survey (IWS), we collected long-term, retrospective data on men's and women's life events and occupational history from the time they were 15 years of age. It was administered only to respondents who were between 18 to 47 years of age, thus giving us information for up to a maximum of 32 years (from 15 to 46 years) of an individual's

life.[9] The sample is predominantly rural, with more than 80 per cent of respondents in both states from rural areas owing to the premature termination of the IWS survey during the Covid-19 pandemic.

The LHC allows the construction of a yearly panel with information on the respondent's education, year of marriage and childbirth(s). It also collects information on employment status, type of employment, spouse's employment, household structure (nuclear or joint, making a distinction between different kinds of co-residents), location of residence and exposure to income/other shocks. If an individual is employed, information is collected on the type of employment. This information is available for each individual in the sample for every year from 15 years to their current age.

We start with some basic descriptive statistics that are along expected lines. Women get married and

Table 4.1: Life history calendar – key indicators

	Men			Women		
	Mean	Median	SD	Mean	Median	SD
Age at time of interview	35	35	6.5	33	32	7.2
Age at marriage	23	22	4.8	18	17	3.3
Age at first childbirth	25	25	4.5	20	20	3.3
Gap between marriage and childbirth	3	2	2.4	2	2	2.3
Age at entry in workforce	18	15	4.6	19	17	5.4
Age at entry in paid work	19	17	5.4	20	18	6.3
Age at entry as contributing worker	17	15	4.0	19	18	5.2

Sources and notes: India Working Survey (IWS). The table presents age of the respondent at various key life events in the rural sample. The average age for the events are conditional on that respondent experiencing the event. For example, the average age of entry into workforce is calculated only among those individuals who ever enter the workforce.

have children early in life in India, but their entry into the workforce is later than men (Table 4.1). About half of the women in the sample got married by 18 years of age whereas the median age at marriage for men is 22 years. On average, women experience first childbirth when they are 20 years of age, while

Figure 4.6: Marriage increases the likelihood of being employed for women in rural Karnataka and Rajasthan
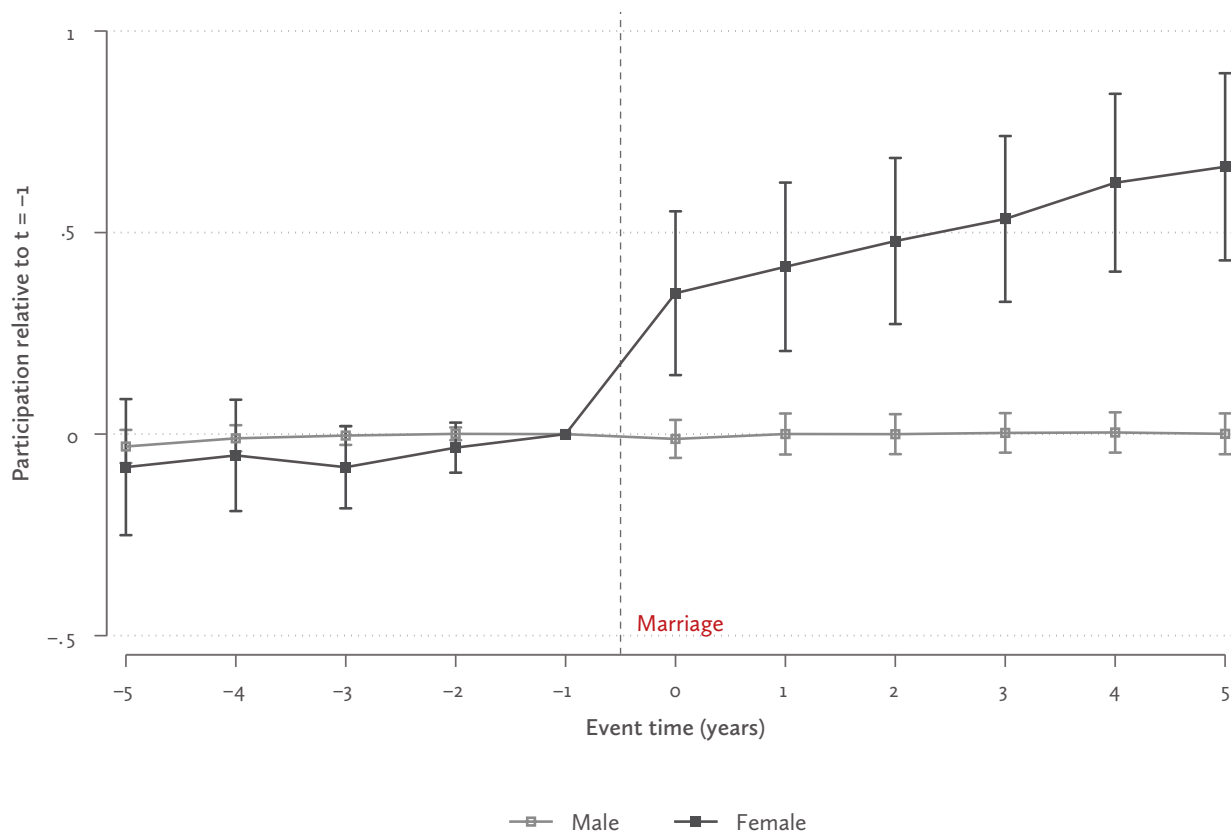
Sources and notes: India Working Survey 2020. These are results from a Life History Calendar exercise. Dashed line indicates time of marriage. Zero indicates the first year of marriage. See text and Methods Appendix for details.

men are 25 years of age at the time of birth of the first child. The average gap between marriage and first childbirth is two years for men and women. Conditional on ever participating in the workforce, men on average are likely to join the workforce just before turning 18 while women on average join after they turn 19.

In our sample from rural Karnataka and Rajasthan, women experienced a sharp jump in workforce participation from 26 percent in the year preceding marriage, to an average of 49 percent in the first five years of marriage.[10] This employment was largely as contributing family workers or self-employment in agricultural work. For men, the corresponding change was far smaller from 88 to 94 percent. Women also experienced a jump in work force participation after first childbirth, but it was smaller in magnitude compared to the increase seen after marriage. Workforce participation for women

increased from 45 percent one year before childbirth to an average of 51 percent in the first five years after childbirth.

Most women in India experience both events closely spaced, with marriage usually preceding child-birth. The median gap between marriage and childbirth is two years. We used an event study analysis to understand the causal impact of the event of marriage and childbirth on women's participation in the workforce. We included year and age fixed effects to control for calendar-year specific events and life cycle effects, respectively. We capture the effect of the event when controlling for age and year as there is variation in the age at which women were married or had their first child.

Figure 4.6 shows the trajectories of participation in the workforce with respect to time of marriage for men and women. The marriage event is depicted by

Figure 4.7: Increase in employment after marriage among rural women is largely in paid work
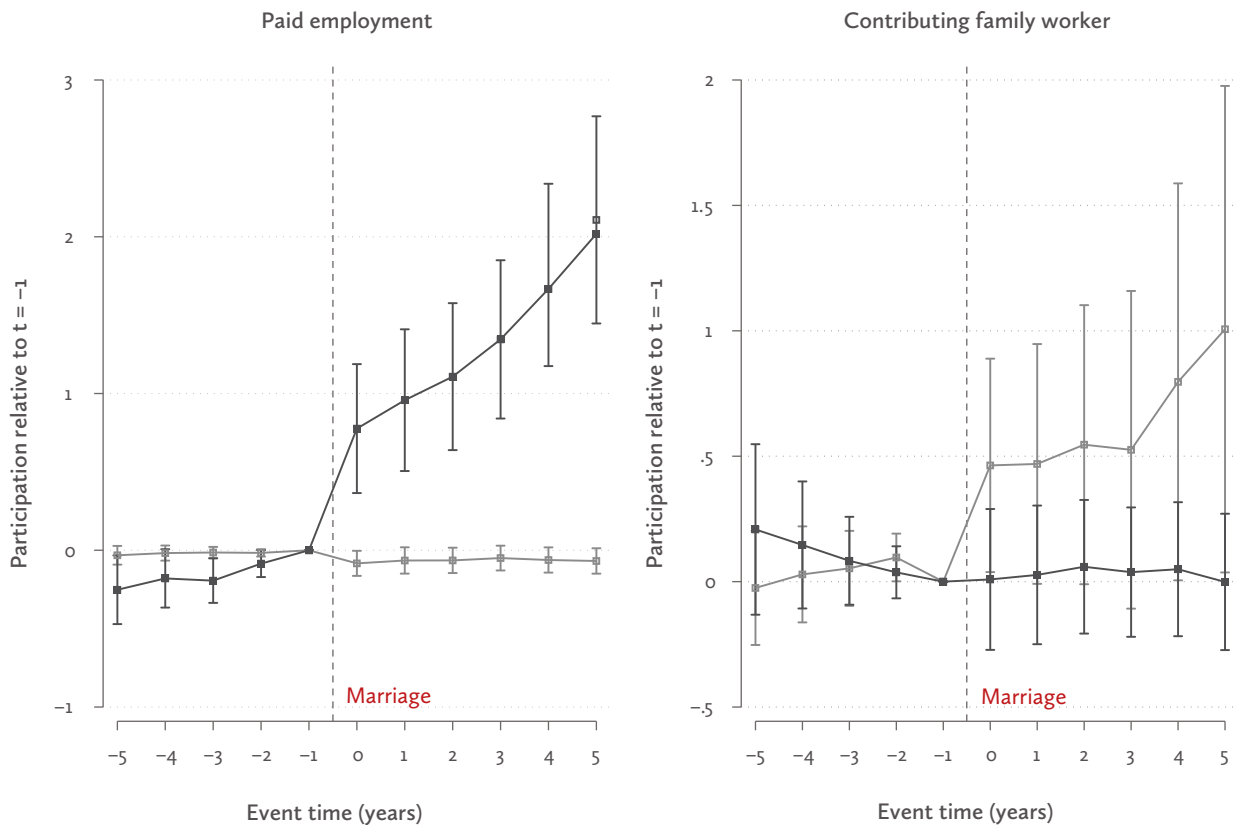
Paid employment   Contributing family worker

Sources and notes: India Working Survey 2020. These are results from a Life History Calendar exercise. Dashed line indicates time of marriage. Zero indicates the first year of marriage. See text and Methods Appendix for details.

the dashed line between -1 (a year prior to marriage) and o (the year starting immediately after marriage). The Y-axis values are indexed to the value of the labour force participation rate of each group at the time of marriage (at year zero). This implies that in the years prior to marriage, men's employment rate was not significantly different from the rates at the time of marriage. We find a similar pattern for women too. At marriage, the trajectories diverge significantly. Women experience an immediate increase in participation in work by 35 percentage points, while men experience no significant change in their participation rates. In the years following the initial increase, the levels do not fall back and instead there is a gradual increase from this initial jump. In the fifth year of marriage, the participation rate for women is 66 percentage points higher than what it was one year before marriage. This increase in labour force participation among women was driven by an increase in only paid work participation. Participation in contributing family work showed no

significant change (Figure 4.7).

The increase in paid work participation can be driven by an increase in work in the informal sector or in the formal sector. Figure 4.8 shows the evolution of informal and formal employment rate among men and women.

Contrary to the "motherhood penalty", the year of childbirth saw a five percentage point increase in work participation rates among rural women.

There was no significant difference in the trends in informal and formal work pre-marriage for both men and women. In the year of marriage, women's participation in informal work increased significantly by 84 percentage points as compared to the year before marriage, but formal work participation showed no significant change. Within informal work, self-employment and casual work showed a sharp

Figure 4.8: Increase in likelihood of doing paid work post-marriage is limited mostly to informal work
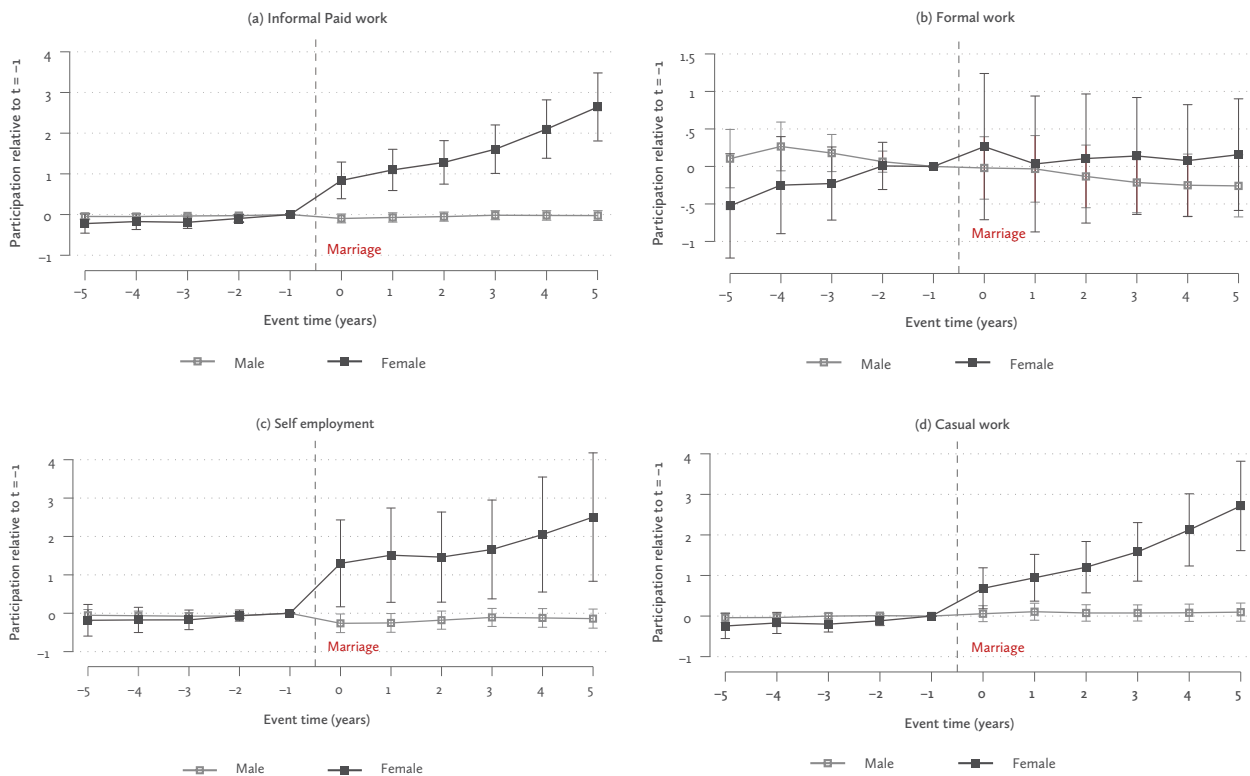
Sources and notes: India Working Survey 2020. These are results from a Life History Calendar exercise. Dashed line indicates time of marriage. o indicates the first year of marriage. See text and Methods Appendix for details.

Figure 4.9: Increase in likelihood of employment post–childbirth in rural Karnataka and Rajasthan
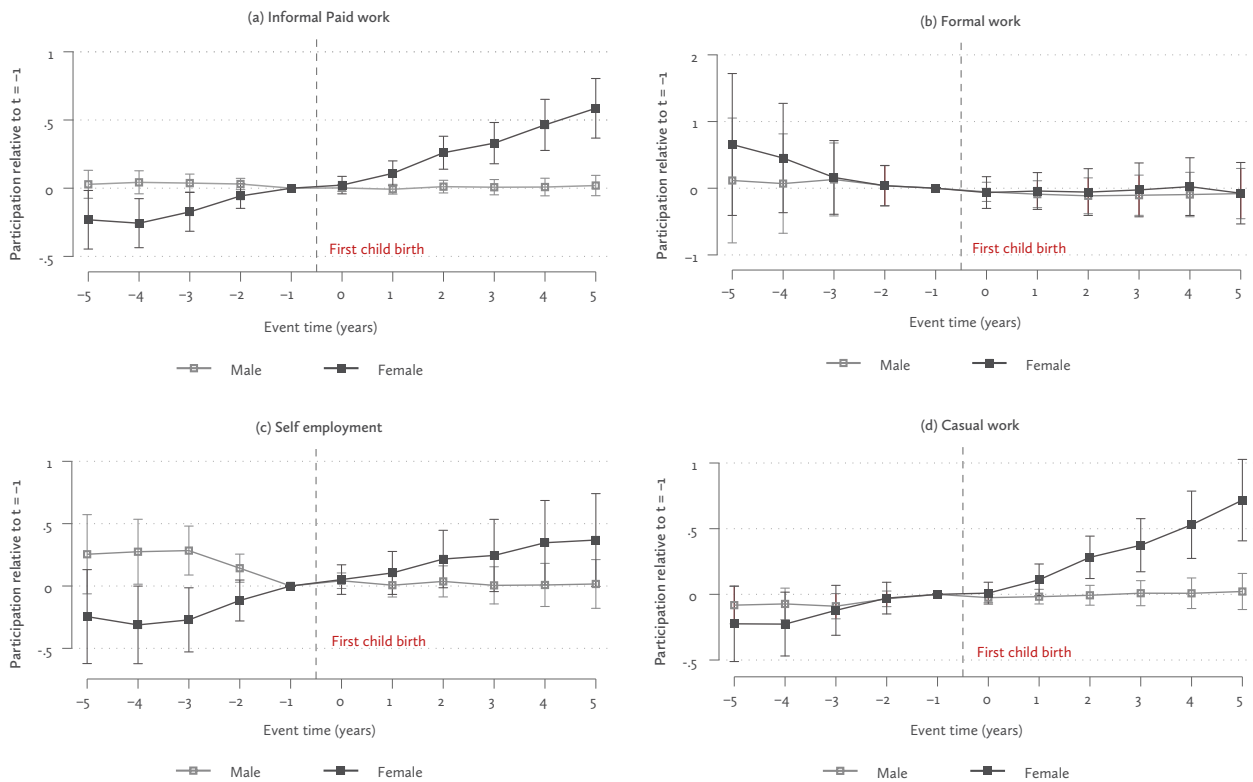
Sources and notes: India Working Survey 2020. These are results from a Life History Calendar exercise. Dashed line indicates time of childbirth 0 indicates the year of childbirth.

increase in the year of marriage and continued in the five years after marriage.

Coming to the second event, childbirth, as with marriage, men and women's labour force participation rates evolved similarly prior to this event. Contrary to what one would expect based on standard understanding of the "motherhood penalty", the year of childbirth saw a five percentage point increase in work participation rates, and five years after childbirth, women's work participation was 30 percentage points higher than the year prior to marriage (Figure 4.9). The magnitude of impact for women after childbirth was far smaller than experienced upon marriage. In contrast, men do not experience any significant change in their work participation rates in the years following first childbirth. Like in the case of marriage, this increase is mostly driven by participation in paid work, largely casual wage work in agriculture. Notably, as in the case of marriage, there is no significant change in

the women's employment in salaried work.

Our findings contradict standard predictions regarding women's employment and marriage and childbirth. Why is this so? First, the age of marriage in India is very low at 18 years. For most women, there is limited possibility for employment pre-marriage as they have either just concluded their education or have been engaged in domestic work within the home. Second, for many young girls, pre-marriage, the norms around mobility and interaction with outsiders can be particularly restrictive. An experiment using fake portfolios of women on online marriage portals found that women who were employed before marriage received far fewer interests from potential suitors compared to women who were not employed (Dhar 2021). The post-marriage situation may be different with the expectation to contribute to the family farm or business.

The presence of a large informal labour market also facilitates the employment of women, although in precarious, often tedious and low-paid work. The costs of entry and exit from such types of jobs are low. For example, women can choose to participate in daily wage work in other's farms from a few days during harvest season. Agricultural or home-based work, as Gautham (2022) finds, gives women temporal and spatial flexibility that allows them to continue working even soon after childbirth. Though flexibility in terms of hours and entry-exit afford an ability for women to manage work and household responsibilities, these are not desirable jobs in many respects and women often take it up only due to lack of options.

Thus, in sum, the absence of a penalty, in itself may not necessarily imply a positive outcome. For many new mothers, post childbirth, there is a compulsion to continue working. Therefore, the lack of the penalty in reality is a burden for many mothers who have to endeavour to fit their productive work around their work for social reproduction.

4.2.4 Balancing employment and household work – the role of social norms

We have seen thus far in the chapter that caste and religious identities as well as household composition and life-events like marriage and childbirth, all play a crucial role in determining women's employment. Though gender norms have come up many times, we have not attempted to measure their impact directly. The problem is that it is difficult to gather information on norms in a survey. Ordinarily labour force surveys do not collect any information on variables that capture intra-household relations and gender norms. This limits our ability to estimate their impact on employment.

We approach this problem in two ways. First, in this section, we draw on data from the India Working Survey where we did attempt to measure norms. The next section uses data from the NFHS to do the same.

Besides collecting detailed information on employment activities, the IWS asked respondents about their participation in household work and in decision-making across multiple dimensions within the household. We asked respondents about the time they spent in the last week on a number of household activities including (a) domestic chores like cooking, cleaning, gathering firewood or water, (b) maintaining kitchen gardens and taking care of livestock that produce eggs/milk for household consumption, and (c) taking care of children and elderly. Women were asked if they were involved in decisions made in their family on (a) their employment, (b) purchases of expensive items, and immoveable property, and (c) their mobility outside

Table 4.2: Women who report higher autonomy inside the home are also more likely to be employed

	Employment		Hours worked per week	
	(1)	(2)	(3)	(4)
Higher autonomy	0.095***	0.074***	3.825***	2.522**
	(0.017)	(0.017)	(0.966)	(1.014)
Controls	No	Yes	No	Yes
R squared	0.009	0.088	0.008	0.061
Observations	3354	3354	1874	1874

*Sources and notes - India Working Survey 2020. See text and Methods Appendix for details of the regression. *** - p<0.01*

their home (to a shop, to a health centre, to visit family or friends, or travel a distance by train). We also asked if they had access to their own money.[11] Using women's responses to these questions, we created an index of autonomy, where women who were involved in more decisions and had access to money would have higher values.

We find that women with higher autonomy levels have an employment rate that is on average almost 10 percentage points higher than women with lower autonomy (Table 4.2, col 1). Similarly, conditional on working for at least one hour, women with higher autonomy work around 4 hours, or half a day, more per week (col 3). It could be that these differences are driven by other factors, like education, where more educated women have higher autonomy and are also more likely to work. Hence we control for such factors - age, education, parents education, caste group, marital status, household size and household wealth measured by an asset index - in a regression framework, and find that the differences

remain significant. Women with higher levels of autonomy are 7 percentage points more likely to be employed than women with similar demographics but lower autonomy (Table 4.2, col 2). Similarly, women with higher autonomy work 2.5 hours more per week than those with lower autonomy but similar characteristics (col 4).

We should note that the results discussed above are statistically significant correlations and a causal interpretation cannot be drawn just from this exercise. This is because both norm-related questions and employment questions are being asked of the same individual. The causal relationship could go both ways - from decision-making norms to employment as well as from employment to ability to participate in decision making. In the next section, we address this problem to some extent using NFHS data.

Autonomy also matters when it comes to the well-known double burden, or the observation

Table 4.3: The double burden in rural Karnataka and Rajasthan – women who are employed also do more housework

| | Hours of household work | | | | |
	(1)	(2)	(3)	(4)	(5)
Hours employed	0.165*** (0.017)	0.134*** (0.016)	0.139*** (0.016)	0.139*** (0.016)	0.277*** (0.034)
Higher autonomy			-3.314*** (0.755)	-3.310*** (0.756)	-1.339 (1.154)
Social Capital				-0.251 (0.826)	0.086 (1.261)
Restrictive norms					2.927*** (1.120)
Controls	No	Yes	Yes	Yes	Yes
R squared	0.028	0.174	0.178	0.179	0.186
Observations	3354	3354	3354	3354	1399

*Sources and notes - India Working Survey 2020. See text and Methods Appendix for details of the regression.*** - p<0.01*

that participation in paid work does not reduce the burden of unpaid housework for women. First, we note that the IWS data on time spent by women on household chores clearly shows the multiple and simultaneous burdens of care and income-earning that women have to bear. We found that, on average, women reported having spent 55 hours doing household work per week, whereas men reported only 21 hours. Further, women who reported higher employment hours also reported higher hours spent in household work. Women who do 10 additional hours of remunerative work in a week, also do around an hour and a half of additional household work (Table 4.3, col 1). This relationship

remains even after controlling for demographics variables though the magnitude reduces somewhat (col 2).

But encouragingly, employed women with higher autonomy inside the home (as measured by the questions mentioned earlier) spend less time on housework compared to those reporting less autonomy (note the negative sign on the coefficient for autonomy in Table 4.3, cols 3 and 4). However, the positive correlation between employment and housework still holds. In the last column of Table 4.3, we restrict the sample to only women who reported being out of the labour force for the majority of

Box 4.1: Voices from the India Working Survey Project

As part of the India Working Survey, we carried out some qualitative interviews of working women in urban centres of Bangalore and Delhi. These were largely women working in informal occupations such as cooks or maids, but also women with enough education to get jobs as data entry operators, call-centre workers or other low-end modern service sector jobs. These interviews not only serve to introduce us to the people behind the numbers, they also give us a concrete sense of the trade-offs that women face and how they are sometimes faced with impossible choices. Here we present a few excerpts from their interviews that illustrate the themes of this chapter.

A common theme we heard from women across different occupations was one of having to juggle childcare alongside paid employment. A particularly poignant story is that of a single mother who works as a domestic worker in Delhi.

> I was worrying about how I would bring up three children, so I left the oldest son with my parents in Agra. My sister-in-law (brother's wife) used to trouble my son a lot...My sister

had gone there and she saw this and then she told me that our sister-in-law was not treating my son well, so I should get my son to come back. So, after coming here, I got these children admitted to a government school. I used to go for work to Vasant Kunj, so I used to get up at 3:00 am in the morning, I used to prepare food, clean the house, I used to then bathe and finish my prayers, prepare tiffin for my children and I would tell the neighbours that once my children wake up, please tell them to get ready, take tiffin and go to school. The neighbours supported me a lot, they used to keep an eye on my children.

We have discussed the possible roles played by mothers-in-law either by setting norms or supporting childcare. For example, a 26 year old woman who helps out with her family shop and is going to be married soon noted:

> My to-be mother-in-law teaches tailoring to students in an academy, which is run by an NGO. So she has told me that I can join her. She has never told me not to work, in fact she has told me that it would be better if both of us work together in the academy. My mother-in-law has been doing this work for quite

some time, she has been training students in tailoring for several years now. So she has told me that if I want to do something, say, if I want to learn computer or teach computers to students or work in an anganwadi (government-run creche), then I can tell her and they would help me fill the form.

Interestingly the same woman spoke about negotiating her post-marriage work life with her fiance:

My fiance is insisting that I shouldn't do a job. He says that he can get married to me tomorrow itself but I shouldn't be doing a job. But I also haven't got a very comfortable job yet. If I get one, I would do it and he would also agree to it, I think. But since I haven't got one yet, the issue doesn't come up often.

So far he has been telling me not to do a job. But when he gets to know that I am getting a good job with a good salary, then why would he say no?

Another woman working in a garment factory in Bangalore said with regard to the decision to work or not to work outside the home:

Simply, what will I do sitting at home. To look after my kid, my mother-in-law is there. If I sit at home there will be no use. If I go outside [work] it will help my house members, to run the house.

The above testimonials illustrate both the constraints that women operate under as well as the agency they exhibit in meeting their objectives under the constraints.

the past year. These women were asked about the attitudes of their families about them working and the reasons for not working. If they reported that the family did not approve of them working, then they were classified as having restrictive norms. We find that women who reported facing restrictive norms did a higher amount of household work.

Taken together, the results from the India Working Survey point to the role of women's autonomy in affecting both their employment status and demands made upon them for household work.

4.2.5 Social norms and employment participation – An analysis of NFHS data

The findings from the IWS reported in the previous section point to the importance of gender norms surrounding mobility and decision-making for determining employment outcomes. But this finding comes from a largely rural sample. Rural women have traditionally been engaged in agricultural work and hence display higher rates of employment

historically. The family farm is seen as an extended "home" and farm work is considered a part of a woman's daily activities. On the other hand, a particularly salient aspect of the female LFPR or WPR in India is the persistently low levels seen among urban women. What accounts for the largely stagnant and low urban rates?

One explanation lies in the fact that, leaving aside home-based work, participation in employment in urban areas requires women to have the freedom to leave the home for extended periods of time. Gender norms are potentially even more binding here than in the rural context. Other obstacles include concerns around safety, inadequate public transportation as well as the absence of affordable childcare facilities, whether provided by the government, the employer or obtained privately. Indeed, managing the care of young children while being employed full-time appears repeatedly in our interviews of urban working women in Bangalore and Delhi as part of the 20 Voices project (see Box 4.1).

We deal with the question of public transport and distances travelled to work in the next section. In this section, we investigate the relationship between gender norms and employment among urban married women using a nationally representative survey, the National Family Health Survey. The NFHS collects data from eligible women aged 15-49 on many aspects of family planning, reproduction, work, and household relations. We use these questions to construct norms indices along these dimensions: decision making, mobility, justification of domestic violence and husband's control over wife. In addition, we also use data on wife's ownership of large and small assets as well as the actual experience of domestic violence.[12]

Before proceeding to the results, we note that the NFHS has been used extensively to study gender norms. Gupta and Yesudian (2006) and Kishor and Gupta (2004) constructed indices of household autonomy or decision making, freedom of mobility, acceptance of unequal roles between spouses, son preference and attitude to domestic

violence using NFHS-1 (1992-93) and NFHS-2 (1998-99). Sinha, Jha, and Negi (2012) constructed three indices of empowerment using NFHS-3 (2005-06) related to women's decision making, mobility, and restrictions placed on them. Singh et al. (2019) use characteristics such as educational attainments, participation in labour markets alongside indicators of women's ownership of assets, and participation in decision-making as indicators of empowerment. Singh et al. (2022) constructed a patriarchy index using NFHS-4 (2015-16), an adaptation from Gruber and Szoltysek (2016) in the European context. Nandwani and Roychowdhury (2023) use the same data to look at the long-run impact of colonialism on gender norms.

We draw on the above work to ask whether progressive social norms enable greater participation of women in paid work controlling for women's individual characteristics such as age and education as well as household characteristics such as caste, religion, number of members, and wealth. As discussed briefly in the previous section, using

Table 4.4: Some examples of gender norm questions from the National Family Health Survey, 2015–16

Index	Questions
Mobility	Are you usually allowed to go to the market alone, only with someone else, or not at all?
	Are you usually allowed to go to places outside village/community alone, only with someone else, or not at all?
	Are you usually allowed to go to the health facility alone, only with someone else, or not at all?
Decision making	Who usually makes decisions about health care for yourself: mainly you, mainly your husband, you and your husband jointly, or someone else?
	Who usually makes decisions about visits to your family or relatives?
	Who decides how the money you earn (woman) will be used?
	Who decides how your husband's earnings will be used?
Justification of domestic violence	In your opinion, is a husband justified in hitting or beating his wife if he suspects her of being unfaithful?
	In your opinion, is a husband justified in hitting or beating his wife if she shows disrespect for in-laws?
	In your opinion, is a husband justified in hitting or beating his wife if she goes out without telling him?

individual level employment outcomes and norm variables has a potential problem of reverse causality. That is, progressive norms may be positively correlated with employment because women who live in households with progressive norms find it easier to work or because women who work are able to secure a more autonomous space for themselves in the household. This problem can be partially resolved if we estimate the relationship between an individual woman's employment status and the average norms prevalent in the district in which she resides. Any positive correlations obtained here cannot simply be attributed to the fact if a woman is employed she is more likely to have greater autonomy in her household.

The analysis is carried out on NFHS-4 (2015-16) data.[13] Gender norms are measured in two different ways. One set of variables is based on answers to individual questions. These are binary variables that take the value 0 for a regressive response and 1 for a progressive response. For example, take the question: is it justified for a husband to beat his wife if she disrespects her in-laws? The possible responses are yes, no, and don't know. If the woman responds "yes" it is coded as 0, "no" as 1, and "don't know" as missing. The district-level average of this variable is then calculated. Table 4.4 gives a few more examples of such questions in three dimensions - mobility, decision-making and justification of domestic violence.

For questions where the options are not simply binary (e.g. the decision-making questions in the table), we constructed both "strict" and "lenient" versions of the variable. For example, The strict variable takes the value 1 if the woman is the sole decision maker and the value 0 if she takes the decision with someone else or is not involved in the decision. The exception to this are questions where it is expected that both partners will together make a decision (e.g. large household purchases or use of husband's earnings). We calculate the district level average of each question and this becomes the main

explanatory variable in the regression. Once again note that all these questions are from the woman's questionnaire.

Next we combine a set of questions that have a dimension in common, into an index. Thus the mobility index is a combination of questions related to mobility. The index is calculated for each district using district-level averages of individual questions. Every index ranges from 0 to 1. The closer the index value is to 1, the more progressive the district is.

Our main independent variables are thus the district-level average of either a single norm-related question (such as the example given above) or the district-level average of an index (e.g.mobility index). The Methods Appendix provides more details on the construction of these variables.

The probability of a woman being employed is not just influenced by supply-side factors such as gender norms. The demand for labour plays a central role here. If jobs are not available, it does not matter if gender norms are progressive or not. We capture the level of labour demand in the district via the employment rate for husbands' constructed from the information on husbands' occupation reported in the NFHS woman's questionnaire.

If a woman resides in a district with relatively more progressive norms as regards mobility, decision-making, and ownership of assets, she is more likely to be employed.

A woman's employment status for the last seven days is the dependent variable. It takes the value 1 if the woman did any other work apart from housework in the last seven days. This includes self-employment (including unpaid work in a family enterprise) as well as wage work (regular and casual).

First off, we note that the proxy for district labour demand (employment rate for husbands in the non-

Figure 4.10: Progressive norms on women's autonomy correlate with higher likelihood of employment, but there is evidence for male backlash also
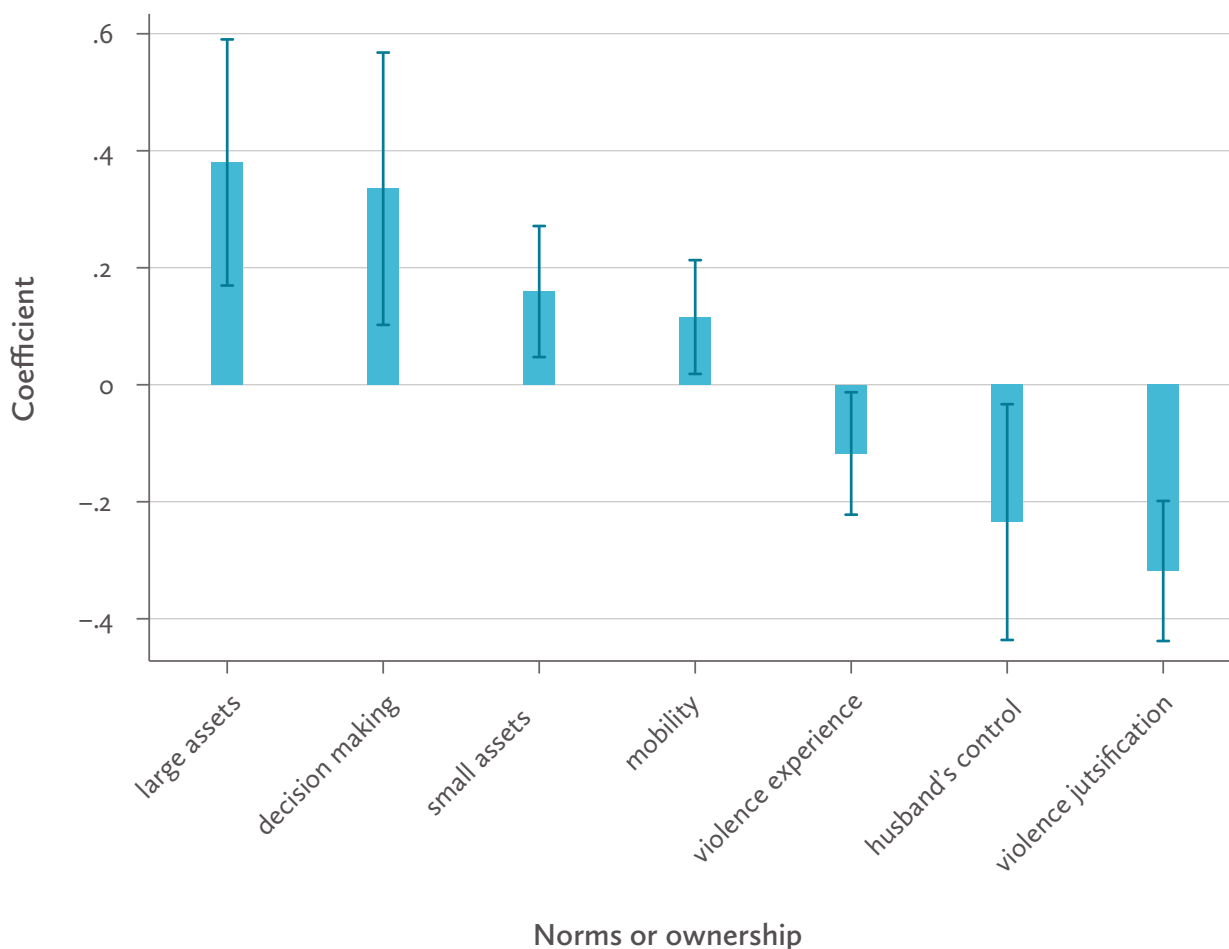
Sources and notes: NFHS 2015-16. The plot shows the coefficients from a regression of women's employment status on district level norms. See Methods Appendix for details on variables and Results Appendix for the full set of coefficients.

agricultural sector), is positive and significant in all the models. That is, in districts where husbands are employed to a greater extent, so are wives.

Coming to the norms variables, indices of mobility, decision making, and ownership of assets are positive and significant (Figure 4.10). That is, if a woman resides in a district with relatively more progressive norms as regards mobility, decision-making, and ownership of assets, she is more likely to be employed. This is after controlling for the household's religion, caste, wealth status and state of residence.[14] (see Results Appendix for all the coefficients)

On the other hand, the coefficients for indices that capture the incidence of domestic violence, women's

justification of violence by their husbands and the extent to which husband's seek to control their wives are negative and significant. Recall that the way the indices are constructed, higher numbers indicate more progressive norms and beliefs. Thus a negative coefficient means that women are more likely to be employed in those districts where average social norms are less progressive. The seeming paradox is resolved when we recall that there is significant evidence in the literature for a "male backlash effect", i.e. that working women either face more violence at home or tend to justify violence. This was discussed in Section 4.1.

Finally, we use the NFHS 2015-16 data along with Economic Census 2013 data and find the same result. We estimate a district level regression

where the dependent variable is the ratio of female hired workers to total hired workers in a district (obtained from the Economic Census) and the main independent variable is the NFHS-derived district level norms index. We find that districts where women constitute a larger proportion of total workers are also the same districts with more progressive social norms around decision-making and mobility (see Results Appendix). We also find a strong labour demand effect as seen in the positive and significant coefficient on the proportion of firms in the district employing more than 10 workers. That is, districts which have a larger proportion of large firms tend to be ones where women are a larger fraction of the workforce.

A closer look at the spatial pattern of our dependent variable is also instructive. Notice that this is not the female workforce participation since the denominator is not all women of working age but rather all workers. In this sense it is a proxy of the demand for female labour. Figure 4.11 shows the district level average for this ratio of female hired workers to total hired workers. We find large district-level variation from a low of only 6 percent to a high of 50 percent. That is, there are districts

where women constitute half of the workforce and others where they constitute only 6 percent. Moreover there are also some noteworthy within-state variations such as that between north and south Karnataka which we return to in the next section when we look at distances travelled to work by women workers.

Taken together this section has explored the relationship between gender norms and women's employment in a variety of different contexts using a range of datasets. While none of the analyses can claim to have established firm causal relationships between the two, we find a lot of suggestive evidence that confirms the continued importance of norms controlling for labour demand. At the same time, we find evidence that labour demand also plays a crucial role. In sum, the policy approach to improving the participation of women in the Indian workforce has to work simultaneously on both fronts - creating opportunities for women and enabling an environment in which they can take advantage of those opportunities. We return to this theme in the last chapter.

Figure 4.11: District level variation in the share of women employed per 100 hired workers

Legend:
- 6.0 - 18.8
- 18.8 - 22.5
- 22.5 - 26.2
- 26.2 - 32.1
- 32.1 - 50.0
- No data
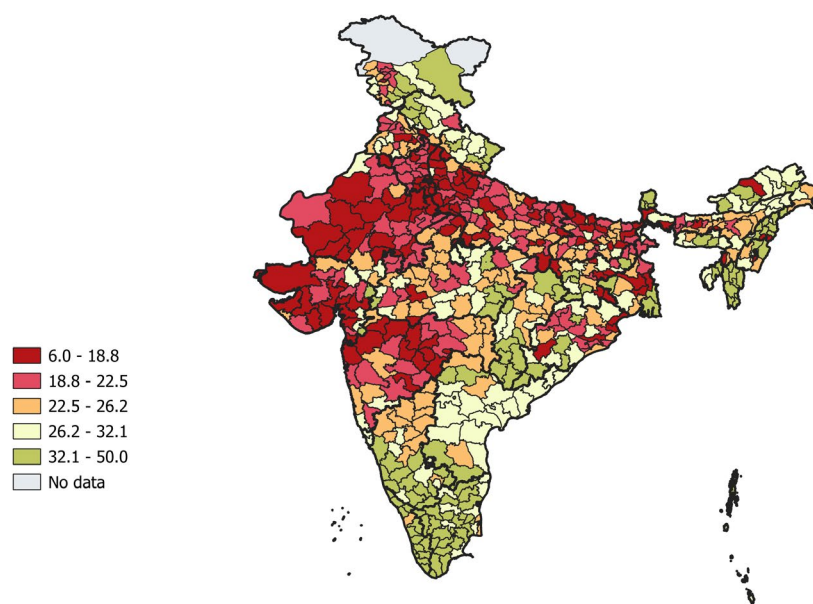
Sources and notes: Sixth Economic Census 2013.

4.3 The role of public infrastructure and local demand

The analysis presented thus far was largely focused on factors determining the supply of female labour, though occasional references were made to the role played by labour demand. We now address the demand question. The lack of aspirational or good jobs has likely affected women's willingness to participate in paid work, especially given increasing educational attainments (Deshpande and Kabeer 2021; Chaudhary and Verick 2014). But it is not easy to isolate the role of labour demand in women's employment. Using district level Economic Census data alongside Population Census data, demand and supply side information can be combined. In this section, we see how demand side factors (number and type of firms present in a district and typical commute distance to employment for men) interact with supply side factors and availability of public transport infrastructure to determine if women travel outside the home for employment.

Across the world, multiple studies have reiterated the role of public infrastructure in enhancing socio-economic outcomes for women including employment and empowerment (Small and Van Der Meulen Rodgers 2023). Using IHDS data Lei, Desai, and Vanneman (2019) find that frequent bus services as well as access to roads increased the odds of non-farm employment of men and women, with the increases being more pronounced in regions with more egalitarian gender norms. Recently, the role of public transport infrastructure on women's employment and mobility has received renewed attention with the announcement of schemes by several state governments. While such schemes have improved women's ridership on public buses, the employment impacts of this remain unexplored.

4.3.1 Gender differences in distances travelled and modes of transport

We begin by examining the Census 2011 data on patterns of women's mobility. Despite the dated nature of this data, it is useful to inform some general conclusions as we outline below. The Census asks questions on distance and mode of travel to all main and marginal workers in 'other' employment. Main workers are workers who work for at least six months in the year, whereas marginal workers are those who worked for less than six months in the year. 'Other' activities refers to all employment that is not agricultural labour, cultivation or household industries.

The Census asks these selected individuals about the one-way distance from their place of residence to place of work. Responses are recorded in kilometre bins ranging from 0-1, 2-5, 6-10, 11-20, 21-30, 31-50 and 50 and above. Besides distances travelled, information on the mode of transport is also collected. The modes of transport include by foot, bicycle, scooter/moped, car/jeep, bus, train or water transport, and others. For each distance bin and mode of travel, information on the number of individuals is aggregated to the district level.

As per Census 2011, about 43 percent of working age women in rural areas were employed. The share was lower in urban areas at 20 percent. For men, employment rates between urban and rural areas were more or less similar. Within the urban workforce, the predominant share, for both men and women, was accounted for by the non-farm category. In rural areas, this share was significantly lower, accounting for about 18 percent of rural female workforce and 28 percent of the male workforce. Therefore, as a share of the workforce, non-agricultural workers, who are the individuals for whom we have distance to work information, account for 30.7 percent in urban areas and 10.2 percent in rural areas.

In rural areas, more than half of women employed in non-farm industries did not travel outside their homes (Table 4.5). Another 18 percent travel less than 1.5 kilometres from their home. Among urban

Table 4.5: Percent of individuals travelling for work for specified distances

	Rural		Urban	
	Men	Women	Men	Women
% travelling > 1.5 km	53.5	27.6	60.1	45.3
% travelling 0-1 km	13.2	17.6	18.1	18.9
% not travel	33.2	54.8	21.8	35.8

Sources and notes: Census 2011.

women, the pattern is substantially different. Only 36 percent of urban employed women stayed within the home. And, about 46 percent travelled more than a kilometre to their place of work. Coming to the median distance travelled for work, interestingly, rural women, urban men and urban women travelled similar distances of around 11 kilometres. However, in rural areas, men travelled much longer distances than their urban and female counterparts at 16 kilometres.[15]

There are significant inter-state differences in women's travel patterns. Southern states like Kerala, Tamil Nadu, Maharashtra and Goa had nearly 60 percent of their employed women travelling more than 1.5 kilometres to work, whereas in northern states like Bihar, UP the corresponding share was only about 30 percent (Figure 4.12). Since the Census data is at the district level, it also enables us to measure within-state variation in the proportion of women travelling for work. Figure 4.13 shows district-level maps for women workers in rural and urban areas. Some states such as UP and Bihar in the

north and Kerala and Tamil Nadu in the south do not show much intra-stata variation, being uniformly low or high respectively. But Karnataka is an interesting case with a clear north-south difference within the state. A greater proportion of women travel for work outside the home in the south, especially in urban areas. Of course, the majority of women who travel outside the home for work do so for short distances. For example only around 12 percent of women who work in non-agricultural activities travel more than 10 kilometres.

As commute increases, women rely more than men on public transport. For a commute of 6 to 10 kms, 37% of women use public transport, compared to only 22% of men. For 11-20 kms women's share increases by nearly 25 percentage points to 62%, whereas for men it is 40%.

Figure 4.14 shows the mode of transport used within each distance bin. When commutes are short (less than 10 kms), most men travel by motorised

Figure 4.12: Share of employed women who work outside their homes

a. Rural

b. Urban
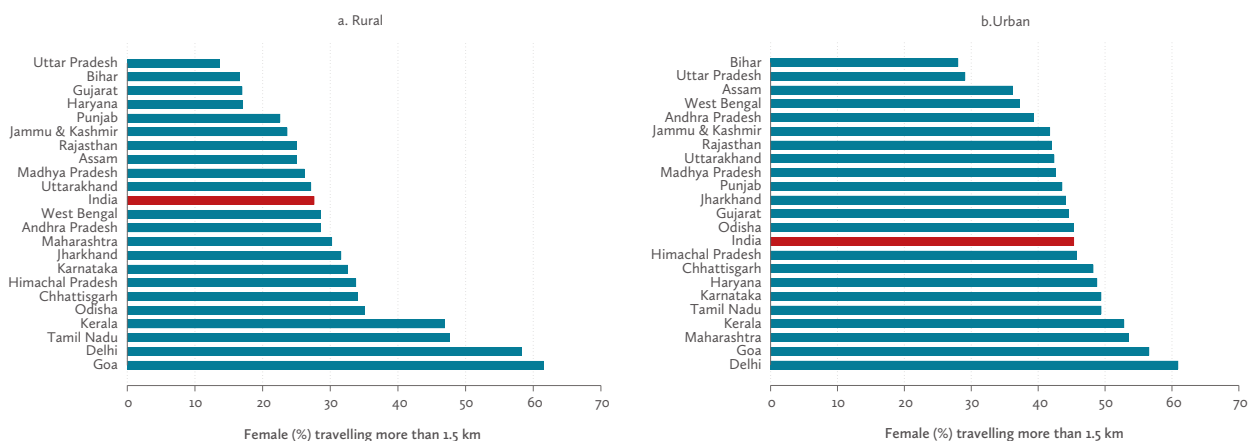
Sources and notes: Population Census 2011

Figure 4.13: Percentage of women in the non-agricultural sector who travelled at least 1.5 kms to work

a. Rural

b. Urban

No data 00–20 20–40 40–60 60–80 80–100

No data 00–20 20–40 40–60 60–80 80–100
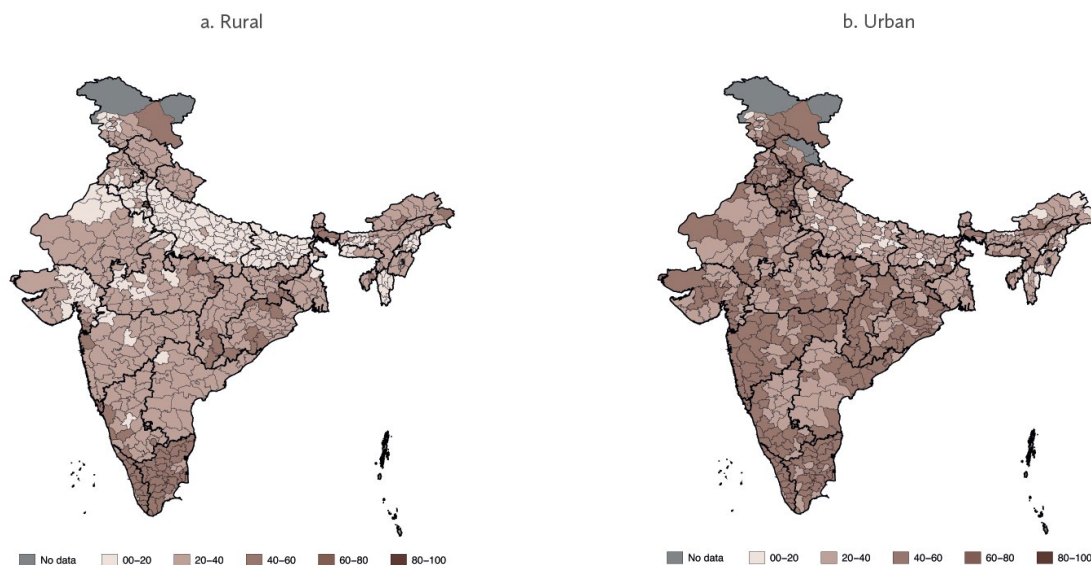
Sources and notes: Population Census 2011

vehicle.[16] Other than this, the choice of commute is as expected. As distances increase, a larger share of men begin to rely on public transport. For women, the starkest contrast with men is the share that commute by foot.

At distances between 1 to 5 kilometres, about 43 percent of women walk to work. Even at longer distances, 6-10 kilometres, about 44 percent of women are still walking to work whereas for men we observe a fall in the share commuting by foot. Unlike in the case of men, there is no substantial shift

towards the use of motorised vehicles. Rather, for women, as commute distance increases, there is an increased reliance on public transport. For example, between the distance of 6 to 10 kilometres, about 37 percent of women use public transport, compared to only 22 percent of men. And, when distance increases further (11-20), women's share increases by nearly 25 percentage points to 62 percent, whereas for men it is still about 40 percent. Therefore, from the second distance bin onwards, women have a very high dependence on public transport for their commute to work.

Figure 4.14a: How do workers travel? Mode of transport for each distance bin in rural India for men (left) and women (right)

On Foot   Bicycle   Motorised Vehicle   Public Transport

On Foot   Bicycle   Motorised Vehicle   Public Transport
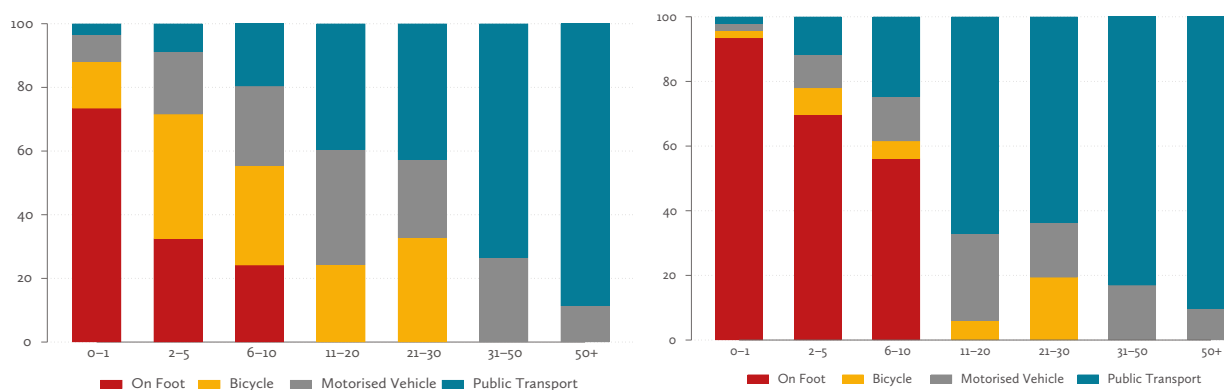
Sources and notes: Population Census 2011

Figure 4.14b: How do workers travel? Mode of transport for each distance bin in urban India for men (left) and women (right)
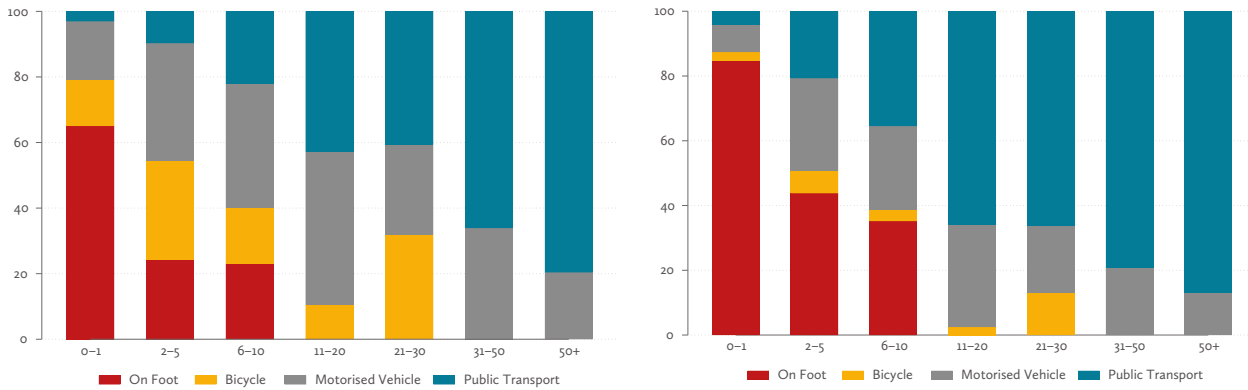
Sources and notes: Population Census 2011

4.3.2 Labour demand and women working outside the home

The descriptive analysis so far reveals two important features. First, there is significant spatial variation in the share of women working outside home, both with regards to rural-urban as well as at the state level. Second, as the distance to the workplace increases, while the share of working women decreases, their reliance on public transport increases significantly compared to their male counterparts.

This brings us to the continuing theme of this chapter-: what are the factors that influence women's workforce participation? In particular, this data allows us to sharpen the question to women working outside the home. One can imagine a number of socio-economic factors, such as labour demand, structure of the local economy, dominant social norms of the region, transport infrastructure, economic condition of the household, caste identity etc., to play a role. While it is difficult to isolate the causal impact of such factors, we implement a simple district-level OLS regression to comment on the strength and the nature of association of these variables with the proportion of women working outside the home.

To capture labour demand in a district we use data from the 6th Economic Census (2013) on the number

of residential-commercial and purely commercial enterprises hiring at least one worker. We estimate a regression model for rural and urban areas separately, taking proportion of women working outside the home in non-agricultural activities as the dependent variable with the following independent variables - district level number of firms in construction, manufacturing, traditional and modern services (proxying for labour demand), proportion of female graduates in a district (as an indicator of both labour supply and local gender norms), percentage of SC/ST population in the district (as an indicator of identity based privilege/vulnerability), median distance travelled by men (as another indicator of labour demand), percentage of men travelling by bus (as an indicator of availability of public transport) and finally, the working age population in the district (indicating the overall labour supply of the region). We also control for state level administrative policies and other factors that do not vary within a state with the help of state controls.

Districts with higher numbers of firms in the modern services sector were also districts with a larger proportion of women working outside the home, while the opposite was the case for traditional services.

The regression results are presented in Table 4.6. First, note that for urban areas, districts with higher

Table 4.6: Labour demand in modern services is significant in raising women's employment outside the home

Variable	Rural	Urban
Proportion of female graduates	1.691***	1.318***
	(0.31)	(0.09)
Proportion of SC/ST	0.115***	0.08**
	(0.03)	(0.03)
Median male distance travelled for work (km)	-0.004***	-0.002***
	(0.00)	(0.00)
Proportion of men travelling through bus	-0.101**	-0.012
	(0.04)	(0.04)
Log number of manufacturing firms hiring atleast 1 worker	0.013**	-0.013**
	(0.01)	(0.01)
Log number of construction firms hiring atleast 1 worker	-0.004	0.005
	(0.00)	(0.00)
Log number of traditional services firms hiring atleast 1 worker	0.011	-0.022**
	(0.01)	(0.01)
Log number of modern services firms hiring atleast 1 worker	0.019*	0.023**
	(0.01)	(0.01)
Log working age population	-0.035***	0.022***
	(0.01)	(0.01)
State controls	Included	Included
N	526	538

*Sources and notes: Census 2011 and Economic Census 2013. Dependent variable - Proportion of women working outside the home in the district. *p<0.10, **p<0.05, ***p<0.01. Standard errors in parentheses. Standard errors are clustered at district level.*

numbers of firms in the modern services sector were also districts with a larger proportion of women working outside the home, while the opposite was the case for traditional services. Even in rural areas, the number of firms in the modern services sector correlates positively with the proportion of women working out of the home. Manufacturing shows the opposite effect in rural versus urban areas, being positively associated with women working outside in rural and negatively in urban areas.

A second proxy for labour demand in our analysis is the distance that men travel for work. We find that the proportion of women employed outside home is significantly negatively associated with the distance men travel for work. If we take long commuting distances as an indication of lack of local work opportunities, this finding also strengthens the case for a demand-side explanation for low levels of female participation.

This observation on distance to the workplace, brings us to the next question: to what extent the availability of public transport, especially buses (on which women depend more than men) matter for women's workforce participation outside home. We included a variable that captures the availability of public transport in the form of the proportion of men travelling by bus. Interestingly, when we exclude state controls, we see a positive and

significant coefficient for this variable (see Results Appendix). This suggests that better transport infrastructure may encourage more women to work outside the home. But with the inclusion of state controls, this effect goes away. This means that most of the variation in this variable is coming from across states rather than within the state. Nevertheless, the point remains that transportation infrastructure is likely to be an important determinant of whether women work outside the home. We return to this point when we discuss public policy in Chapter Six.

Finally, the proportion of female graduates in a district is positively associated with the proportion of women employed outside home. The strength of the association remains consistent across all the specifications. There can be two intuitive explanations for this. First, districts with relatively higher share of female graduates, will most likely have a higher labour supply of women. Thus, employment outside home is more likely and common in such areas. Second, it is likely that districts with a higher share of female graduates are economically more developed and thus offer more employment opportunities for women. It could also be that if women in these districts are already travelling outside home to access education the norms around women's mobility might be less strict as compared to other districts. At the same time, it is also possible that norms governing access to education for women are different from those governing employment since education is independently important for marital prospects.

Similar to the share of female graduates, we observe that in all our specifications, the share of SC or ST population matters, and is positively associated with the proportion of women working outside home. It is well documented that both with regards to vulnerability and norms, the SC and ST groups differ significantly from others. The socio-economic vulnerability of such households may often result in both men and women from the same household working simultaneously at a workplace away from home. In addition, as noted at the beginning of this chapter, gender norms operate differently across caste hierarchy and thus, norms around mobility of women could be quite different in SC and ST households as compared to General caste households, even at the face of equal levels of economic vulnerability.

4.4 Conclusion

Persistently low female employment rates remain a concern in India. Constraints on women's freedoms and autonomy and therefore female labour supply continue to play an important part in the story. However, the role of weak labour demand, discussed in Chapter Three, cannot be discounted either. In this chapter, we presented evidence for the continuing importance of gender norms in determining employment outcomes for women. We start with the observation that female workforce participation rates vary significantly by caste and religion. Next we presented the results of a linear probability model that showed the significant role of level of education as well as husband's earnings and mother-in-law employment status. The Life History Calendar approach allowed us to estimate the impact of marriage and childbirth on rural women's employment in Karnataka and Rajasthan. Contrary to the motherhood penalty, we saw that these events raise employment rates for women. This underlines the importance of the socio-economic context in determining the impact of such life events. We made use of the NFHS data on gender norms to show that women's autonomy positively correlated with employment. And finally, we used

the Economic Census and Population Census data to estimate the impact of labour demand as well as public transport on women's employment.

In the next chapter, we take a closer look at how social identities can play a role in determining the type of employment as well as the industry and occupation of work and earnings gaps.

Endnotes

1   In addition to a traditional variable that captures the presence of a young child, the authors use the difference between current number of children and desired number of children as a motherhood proxy. The authors argue that 'extra children' is a closer estimate of the motherhood burden as the desired number of children could be endogenous to women's labour market outcomes, i.e. women internalize their ideal family size when making labour supply choices.

2   Poor measurement of women's employment has also been cited as a reason for India's low levels of women's workforce participation. We discuss this in detail in Chapter Six. However, while inadequate measurement may explain low levels to some extent, it cannot explain the steady decline in employment rates witnessed in the last few decades given the definitions used have remained more or less unchanged.

3   The controls include age, education, religion, caste, household size, marital status, education of the head of the household and state of residence.

4   We have constructed a combined religion-caste variable for this analysis to capture the fact that caste divisions are to be found across religions in India. See the chapter on Measurements for more details.

5   We also examined previous NSSO survey rounds to see if the U-shaped relationship has changed over time. In urban India, we see a small decrease in the extent to which higher educated women exceed the likelihood of being employed compared to illiterate women (see Results Appendix).

6   This section is based on Lahoti et al. (2021)

7   The Consumer Pyramids Household Survey of the CMIE is one such panel survey but it has been in existence only since 2014.

8   Typically, a chronological time frame is presented graphically to the respondent and information is collected for this timeframe using specific personal events such as childbirth, death, and marriage or

other major public events as anchors(Glasner and van der vaart 2009). See Methods Appendix for details including a picture of our survey instrument.

9   The LHC was administered to 3,078 individuals in 2,065 households. Of these, 1,766 were women (1,010 from Karnataka and 756 from Rajasthan) and 1,312 were men (698 in Karnataka and 614 in Rajasthan).

10   Note that the levels of female workforce participation rate as measured in our India Working Survey are much higher than those recorded in the NSSO surveys. We discuss some possible measurement-related reasons for this in Chapter 6.

11   This module was based on the NFHS questionnaire.

12   The National Family Health Survey (NFHS) collects data from eligible women aged 15-49 under various heads (reproduction, marriage and cohabitation, contraception, pregnancy, post natal care, child nutrition, family planning, maternal and child health, sexual activity, fertility preferences, HIV, woman's work, household relations, domestic violence). The five rounds were carried out in 1992-93, 1998-99, 2005-06, 2015-16, and 2019-20. To construct the social norms, data collected in the sections on woman's work, household relations, and domestic violence is used. The NFHS has different questionnaires asked to men, women, household. This study is based on data asked to women only.

13   NFHS-5 (2019-20) also has sections of the questionnaire asked at the district level. But the data is collected during pre and post covid periods and has not been used for the regression. The full model along with more details of the methods used are available in the Methods Appendix.

14   To check the robustness of the results, we also ran regressions with individual questions as opposed to the combined indices. The results are consistent with those already shown. For example, all three variables that make up the mobility index have positive coefficients, and are significant. If a woman resides in a district in which more women could go out on their own to different places such as a market, a health facility, outside a village/community, it increases her chances of being employed. See Results Appendix for the numbers.

15   Median distance is calculated by using the midpoint of a distance bins and taking the weighted average of this mid-point, across all distance bins, with weights being the population shares in each bin. For the calculation of median distance, the midpoint for 50+ distance bin is taken as 75. However the distance travelled in relative terms remains unchanged even if some other midpoint is chosen. For instance the rural men still travel the

highest distance of 15.12 km if 60 is considered as the midpoint followed by urban men travelling 11.32 km. The urban and rural women travel 10.38 km and 10.09 km respectively using the above mentioned midpoint for median calculations.

16 There is an unusual jump in the share of men using bicycles within the distance of 21 to 30 kms which we cannot explain.

5

Good jobs, intergenerational mobility and educational convergence

5 Good jobs, intergenerational mobility and educational convergence

So far we have examined India's structural transformation experience through the lens of social identities and have taken a detailed look at women's employment. In this chapter and the next, we undertake an in-depth examination of how the process of economic growth and structural transformation has impacted labour market outcomes for our three key identities -gender, caste, and religion. A running theme in this chapter is, who is able to access better quality jobs and who is not? Here we extend the analysis presented in our first report, State of Working India (2018) (Chapter Five, Who does the work?).

We start with an analysis of the individual and household characteristics that correlate with regular wage work and then proceed to identify the determinants of better quality regular wage work. We use the terms "regular wage" and "salaried" interchangeably. Next we provide some estimates of the extent of intergenerational mobility observed in recent years, with regard to the type of employment as well as the sector of work. Finally, we present some data on educational convergence across castes and jatis.

5.1 Caste, gender, religion and regular wage work

Self-employment, casual wage work and regular wage work constitute the three major types of employment in the Indian economy. In Chapter Three we examined the long-run structure of the economy in terms of these three categories. We saw that, compared to the transition out of agriculture, the transition out of informal work (self-employment and casual wage work) was much slower. But we also saw that in recent years the share of regular wage work has increased and the share of casual wage work has fallen. This shift is particularly salient for women. In this section we examine the determinants of each employment type.

Between 1983 and 2004, the share of regular wage workers in the workforce did not change very much, hovering between 14 and 16 percent. By 2011 it jumped to 19 percent and by 2017 to 24 percent of the workforce. Between 2004 and 2017, even as India's workforce grew by about 12 million (from 415 million to 427 million), the absolute number of salaried workers increased by 35 million (indicating that the absolute number of workers in casual wage work and self employment fell during this period). However, in more recent years, there has been a slowdown in the creation of regular wage jobs.

Figure 5.1: Conditional on being employed, the likelihood of women being in regular wage employment has increased over time
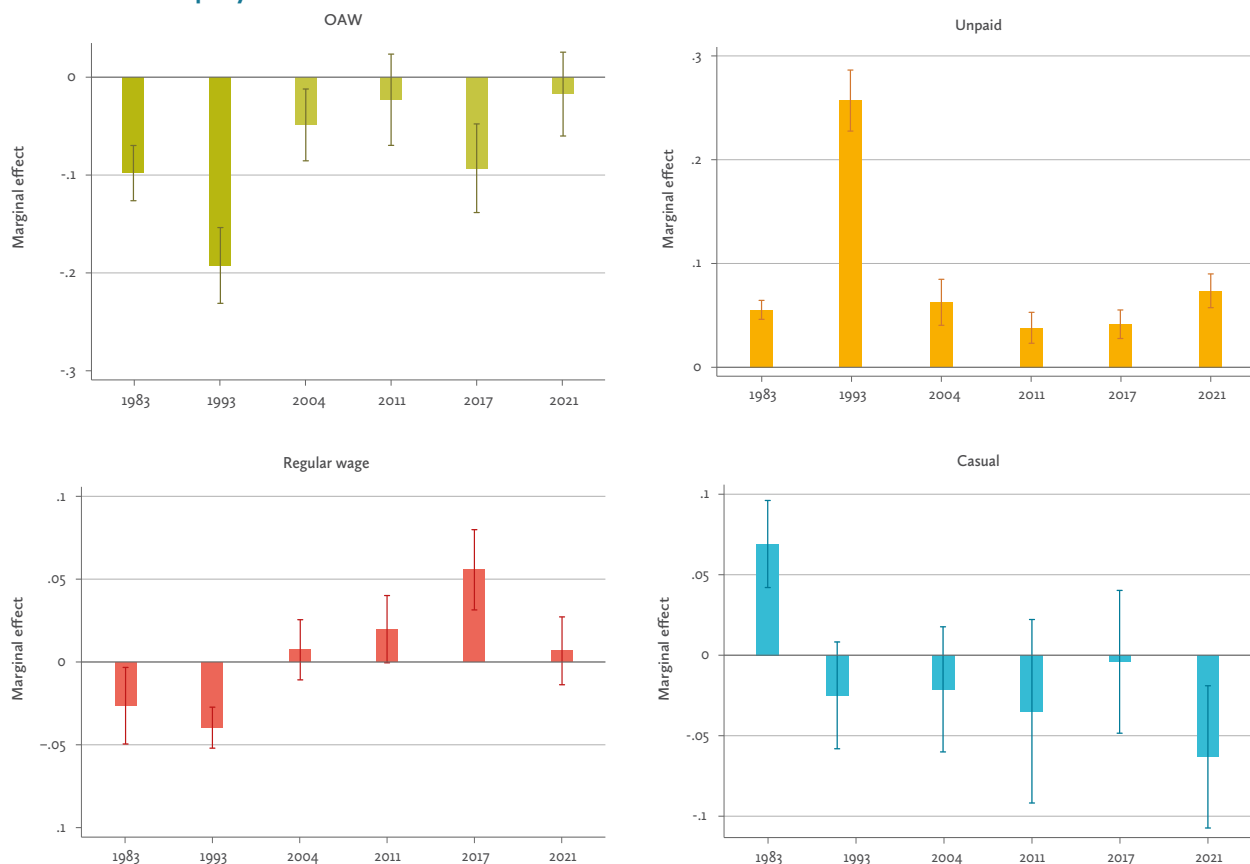
Sources and notes: NSS EUS and PLFS various years. Figures show marginal effects estimated from a multinomial logit model. See text and Methods Appendix for details. OAW = own-account worker

Between 2017 and 2021, although India's workforce grew by nearly 65 million, the increase in number of salaried workers was only 13 million and their overall share in the workforce stagnated at 23 percent.

5.1.1 Who is able to access regular wage work?

It is of interest to know who was able to access the new regular wage opportunities being created in the economy. To answer this question, the approach we follow is that of a multinomial logit model. This allows us to quantify how belonging to a social group affects one's chances of securing a particular kind of job and how these change over time.

The model has four outcomes: own-account work, unpaid work (contributing helpers in household-based enterprises), regular wage work (or salaried employment) and casual wage work.[1] In addition to

an individual's gender, caste, religion and educational background we also control for their age (and square of age), marital status, rural-urban location, household size, state of residence, and education level of the head of the household (if the individual is not the head). This last variable is a proxy for the socio-economic status of the household. Details of the multinomial logit model are provided in the Methods Appendix and complete results are provided in the Results Appendix. Here we focus on some major takeaways from the analysis on the worker's gender, caste, religion and educational attainment.

For social identity, we analyse three binaries, women compared to men, SCs/STs to others, and Muslims compared to Hindus. Figure 5.1 shows the results for gender. The way to read this and the subsequent few graphs is as follows. We plot the marginal effect of a change from the base category

Figure 5.2: SC/ST workers are more likely to be in casual or regular wage work and less likely to be self-employed
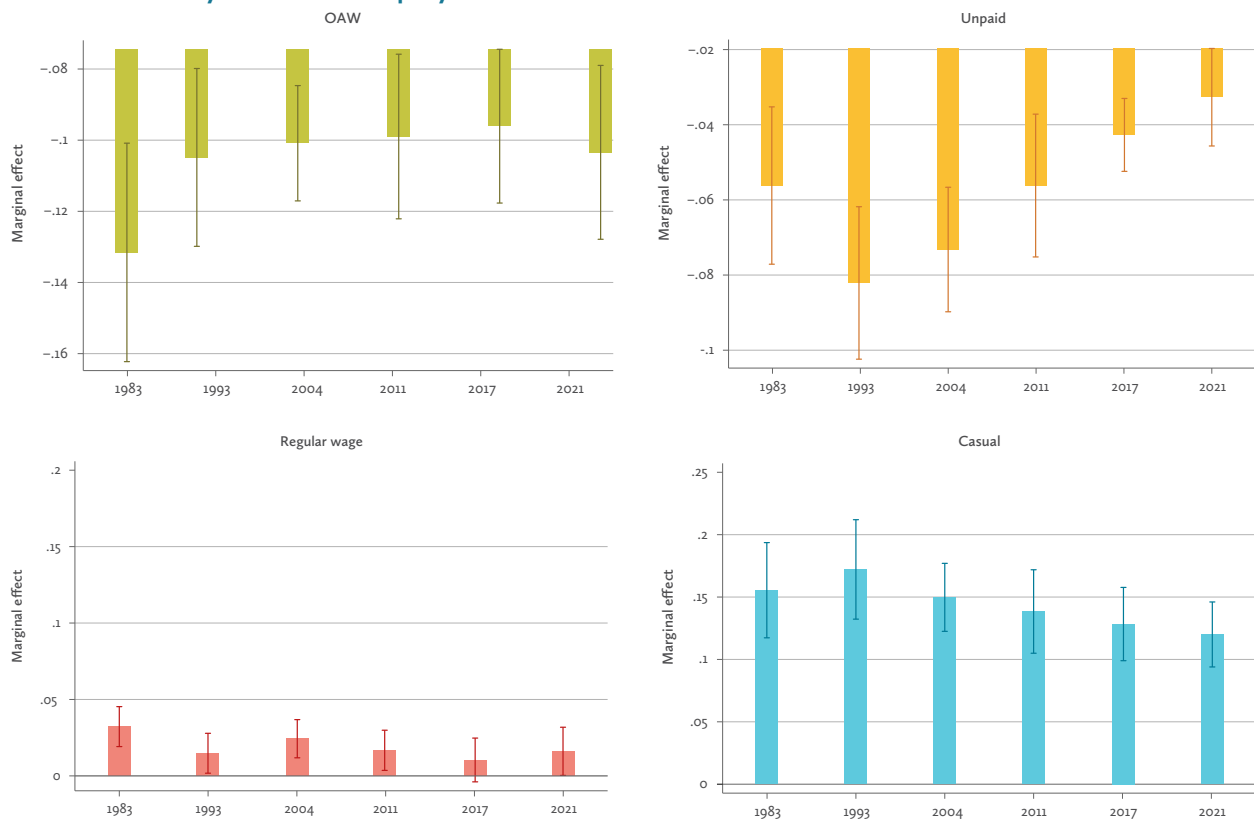
Sources and notes: NSS EUS and PLFS various years. Figures show marginal effects estimated from a multinomial logit model. See text and *Methods Appendix* for details.

on the probability of being in a certain kind of work. For example, take the panel showing the gender effect for regular wage work. It shows that in 1983, the probability of women being in regular wage jobs was less than men. Even in the 1990s, women were less likely to be in salaried work compared to men but this changed in the 2000s. In 2004, the effect was positive but not statistically significant (the confidence intervals overlap with the zero line) indicating that women were not any more or less likely than men to be in regular wage work. But by 2017, women were ~5 percent more likely than men (since the coefficient size is approximately 0.05) to be in regular wage work. Correspondingly, women were less likely than men to be in own-account work or casual wage work and more likely to be in unpaid work. With the coming of the pandemic, the picture changed dramatically and women became even more likely to be engaged in unpaid work as we discussed in Chapter Two.

For caste, the picture is different. The comparison provided in Figure 5.2 is between SC/ST workers and everyone else. Though there are likely to be differences in employment outcomes of SC and ST workers, sample size restrictions do not allow us to separate the two for this analysis. Disaggregated data for OBCs is not available going back to the 1980s. We see that SC/ST workers are far less likely to own enterprises and far more likely to be casual workers compared to Other and OBC workers though the picture did change somewhat over the decades. In the early 1980s, SC/ST workers were 13 percent less likely to be own-account workers and 15 percent more likely to be casual workers compared to everyone else. Over time their over-representation in casual work declined but they remain 12 percent more likely than other castes to be in casual work. Box 5.1 draws on recent studies using both secondary data and ethnographic work to document how social identities interact

Figure 5.3: Muslim workers are more likely to be in own account or casual wage work and less likely to be in regular wage employment.

Sources and notes: NSS EUS and PLFS various years. Figures show marginal effects estimated from a multinomial logit model. See text and Methods Appendix for details.

with informal work in India to produce persistent hardship for workers.

The marginal effects for regular wage work are much smaller (note the difference in the Y axis scale) but it is worth noting that in the 1980s, SC/ST workers were a little more likely to be in regular wage work as compared to others (around 3 percent more likely). But over time this fell and in more recent years, they are no more or less likely to be in regular wage work. This is possibly due to a fall in the share of public employment in regular wage work, where caste-based reservations improve the representation of marginalised groups.[2]

For religion, the effects are, in general, much smaller as compared to caste (Figure 5.3). This is consistent with what we saw in Chapter Two -that the employment structure varies far more across gender and caste identities than it does across religion. But

that said, it is worth noting that Muslims are less likely to hold regular wage jobs and more likely to be in own-account or casual wage work over the entire four decade period. Note that this result is obtained after controlling for education, household size, state and other relevant factors. The persistent under-representation of Muslims in regular wage work was noted in the Sachar Committee Report of 2006 as well and continues to be a matter of concern as it indicates continuing exclusion despite significant economic growth.

The persistent under-representation of Muslims in regular wage work was noted in the Sachar Committee Report of 2006 and continues to be a matter of concern.

Education, expectedly, plays a crucial role in determining the type of work an individual is

Box 5.1: Marginalised identities and the experience of work

In order for growth to be inclusive, it must result in better socio-economic conditions for marginalised sections of the society. Kohli (2023) in her SWI background paper examines caste-class associations during the high growth period covering 1999-2012. She argues that though some associations are weakening, the overall caste-class inequalities have continued to persist during this period. The Scheduled Tribes and Scheduled Castes continue to be overrepresented in worse-off classes. The Scheduled Castes (SC), Scheduled Tribes (ST), and Other Backward Classes (OBC) were more likely to work as rural labour than the Others. The SCs and STs were less likely to be self-employed. Even with improving educational outcomes, the lack of access to resources, credit, and infrastructure continues to impede social mobility..

De (2020), in another background paper on Adivasi migrant workers in the city of Ahmedabad, argues that there is a close link between the status of tribal workers as marginalised within society, and their status as displaced and marginalised in their living areas and workplace. These workers face identity based discrimination in the city as well as structural exclusion from the governance apparatus as migrants. Therefore, tribal migrant workers do not earn enough to subsist and are highly dependent on non-remunerated services of their family and the social security net provided by their village community. The paper

presents a granular picture of what the rapid rise in construction work (discussed in Chapter Three) actually means for those who do this labour.

In the backdrop of a 140 percent increase in heat related mortality in India between 1960 and 2009, Shah et al. (2023) study the association between caste identity and occupational exposure to stressful heat conditions. Over the past decade, researchers across academic disciplines have emphasised how social and economic inequalities are reproduced and exacerbated due to climate change. In the Indian context, however, while anecdotal evidence and journalistic works exist, there is a lack of large scale studies that look at the impact of climate change on existing social inequalities, especially regarding its differential impact across caste groups. The authors construct an Environmental Stress Index (ESI) at the district level for the summer of 2022 and combine this with the PLFS 2018-19 round to first identify workers who work outside home. The district level analysis shows that workers from marginalised caste groups, specifically OBCs and SCs experience about 35 percent higher elasticities of heat exposure because of occupational constraints, as opposed to their General caste peers. Authors argue that policymakers must account for caste and other processes of social stratification while drafting any action plan on climate change or public health.

engaged in. The effect of education has not changed significantly over time, so to avoid clutter, Figure 5.4 only shows the marginal effects for the most recent year, 2021-22. Shown in the figure are probabilities of being in a certain kind of work for a person with a given level of education (X-axis) as compared to an illiterate person. As the level of education

increases we see a large decline in the probability of a person being in casual employment and a large increase in the probability of being in regular wage employment. Own-account work shows an inverted U-shaped pattern. That is, at low levels of education individuals are no more or less likely to be in own-account work as compared to illiterate individuals.

Figure 5.4: Higher education attainment raises the likelihood of being in regular wage work
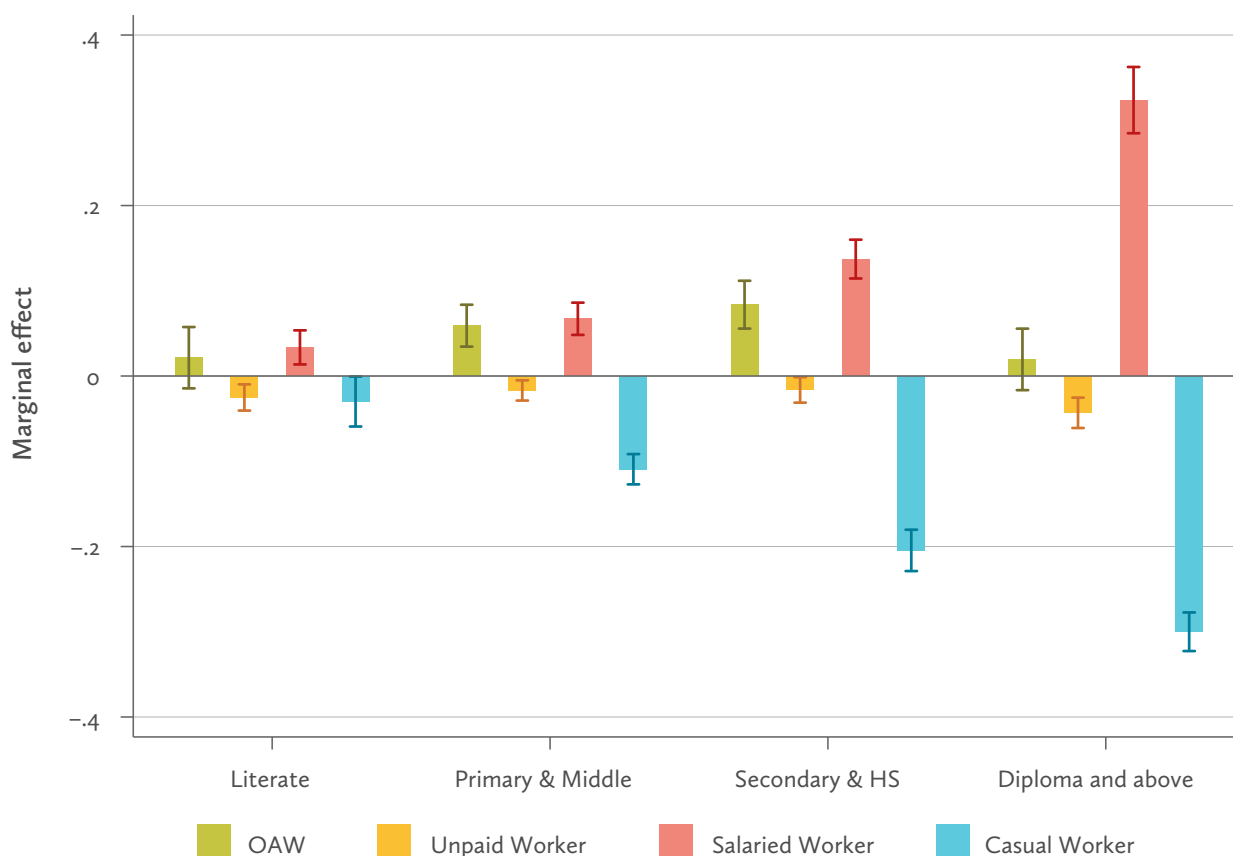
Sources and notes: PLFS 2021-22. Figures show marginal effects estimated from a multinomial logit model. See text and Methods Appendix for details.

As education increases, they are more likely to be in own-account work and with even more education this probability once again declines (because they are more likely to be in regular wage work).

Taken together, the above results allow us to construct a picture of who is likely to be a regular wage worker in India and how this has changed over time. There are expected and unexpected aspects to this. Expectedly, being a non-SC/ST, Hindu, well-educated man strongly improves the chances of being in a regular wage job. But less expectedly, the recent increase in prevalence of regular wage work due to high economic growth has had two opposing effects - it has increased the probability of women being in such work compared to men, and decreased the probability of SC/ST workers having such jobs compared to other castes.

To go deeper into the intersectional effects of

gender and caste or gender and religion we perform an analysis only for women workers (separating out rural and urban regions). This is done only for the most recent year, 2021-22. We use the combined religion-caste variable that we introduced in Chapter Four. This variable has the following levels - Hindu-SC/ST/OBC, Hindu - all Others, Muslim-SC/ST/OBC, Muslim - all Others, Other religions - SC/ST/OBC, Other religions - all Others.

Figure 5.5 shows the results for rural (a) and urban (b) areas. Though the rural results are shown for completeness, the effects are not strong, so we focus on the urban results here. The base category here, which is not on the graph, is Hindu Other caste women. As before, the effects are to be interpreted as an increase or decrease in probability of being in a certain type of work compared to this base category. Thus we see that compared to urban Hindu women (of whichever caste), urban Muslim women are far

Figure 5.5: Caste-religion intersections and type of employment for women
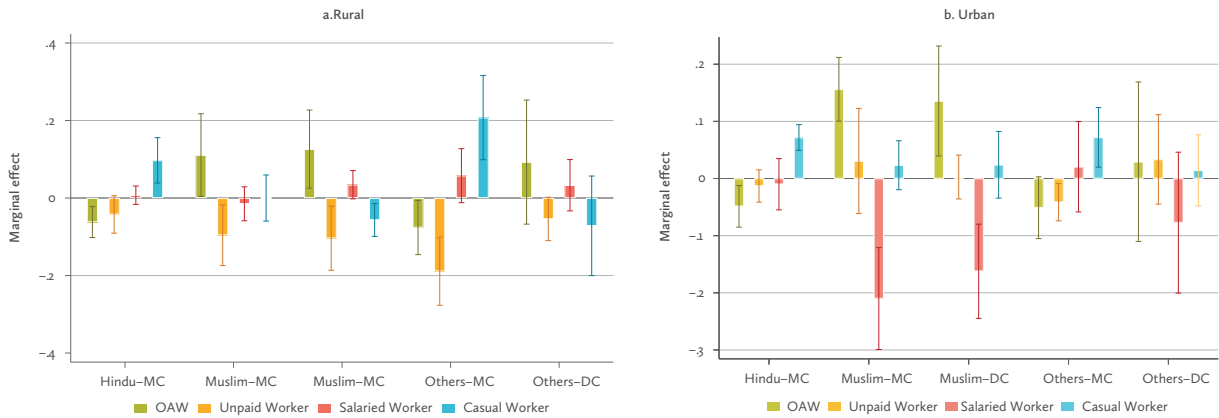
Sources and notes: PLFS 2021-22. Figures show marginal effects estimated from a multinomial logit model. See text and Methods Appendix for details. MC - marginalised caste. DC - dominant caste.

less likely to be in regular wage work, controlling for age, education, state of residence and other factors. Note that, where regular wage work is concerned, the caste difference within Muslim women or within Hindu women matters much less than the difference between Hindu and Muslim women as a whole.

Compared to Hindu women, Muslim women are far less likely to be in regular wage work, controlling for age, education, state of residence and other factors.

Women from other religions are not significantly different from the base category, as seen by the confidence intervals overlapping the zero line.

Finally, coming to the role of educational

attainment, Figure 5.6 shows that in both rural (a) and urban (b) areas, women having diplomas or degrees look very different from all other women in terms of the type of employment they are likely to be in. They are far more likely to be in regular wage work than in any other kind of work (all the other bars are negative). In rural areas they are nearly 60 percent more likely to be in regular wage work compared to the base category (illiterate) and in urban areas 37 percent more likely.

Of course the comparison to illiterate women may not be very meaningful here. But even if we compare these highly educated women to those little less educated (completed higher secondary school) we see a big jump in the probability of being in regular wage work; from 24.5 percent to 59 percent in rural

Figure 5.6: Educational attainment and employment type for women workers

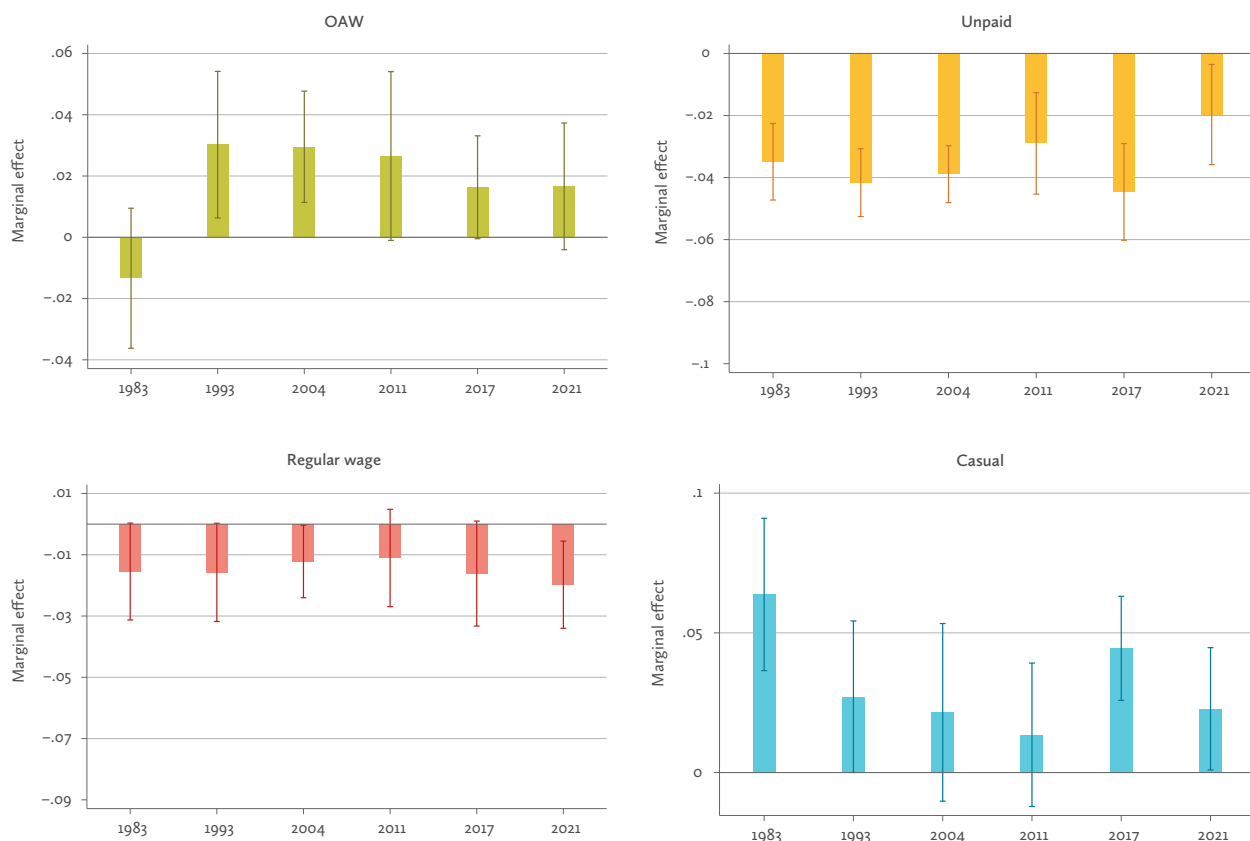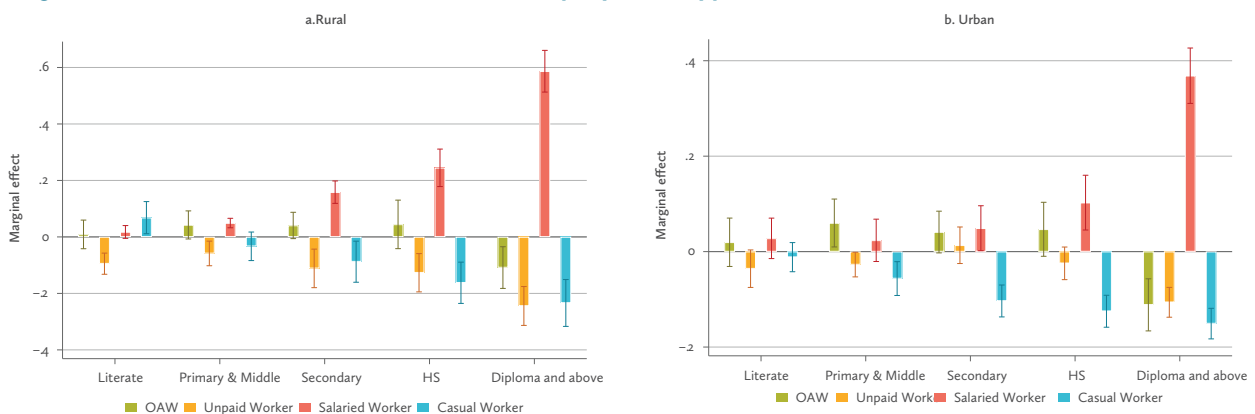*Sources and notes: PLFS 2021-22. Figures show marginal effects estimated from a multinomial logit model. See text and Methods Appendix for details.*

and from 10 percent to 37 percent in urban areas. The crucial role played by educational attainment in securing good jobs is, of course, a well known aspect of not just the Indian, but the global economy. As should be clear from the discussion on the educated unemployed from Chapter Two, this does not mean that education is a sufficient condition for securing such jobs, only that it is a necessary condition. There are two ways of looking at this result. We can see it as a clear indication of the value of college education in securing good employment. But it can also be seen as an inability of the economy to create good jobs for those beyond the minority of the higher educated. At the end of this chapter we return to the theme of education when we analyse the extent of educational convergence since the 1990s across caste.

5.1.2 A shift in the type of regular wage jobs being created

We now delve deeper into the relatively recent rise in regular wage work which we alluded to at the beginning of this chapter. Along with the structural shift away from casual wage to regular wage work since 2004, we also observe a change in the nature of such jobs during this period. We create an ordering of regular wage employment beginning with "formal" where the job comes with a written contract and benefits, followed by "semi-formal" jobs with either of the two (written contract or benefits) but not both, and finally, "informal" - regular wage jobs with neither a written contract nor any benefits. The nomenclature is for ease of reference. It should be kept in mind that in the analytical sense, when we discuss the Lewis process, informal regular wage jobs are also part of the "modern" sector.

Table 5.1 shows that between 2004 and 2017, the decline in casual wage work and self-employment was made up by an increase in semi-formal and informal regular wage work. The proportion of formal regular wage workers stagnated in this period. If we look at the composition of the regular wage workforce by itself (setting aside casual wage workers and the self-employed) the change can be seen more easily. In 2004, the formal category accounted for 36 percent of all regular wage jobs. By 2017-18, its share had fallen to 25 percent, and the share of semi-formal jobs had risen from 21 to 30 percent. Of the 35 million increase in regular wage jobs during the 2004 to 2017 period, only 2 million were in the formal category, about 16 million offered either a written contract or benefits and the remaining 17 million were in the informal category. Semi-formal and informal regular wage jobs are frequently found in labour intensive manufacturing industries such as garments. Box 5.2 takes a closer look at one such job.

Of the 35 million increase in regular wage workers between 2004 and 2017, 26 million were women and 9 million were men.

As we alluded to in Chapter Three, there is a clear gender dimension to the creation of new regular wage jobs. The increase between 2004 and 2017 is larger for women compared to men. In 2004, about 10 percent of employed women were in regular work and the corresponding share for men was 18 percent.

Table 5.1: Regular wage work has become more informalised over time

Share in total workforce (%)	2004	2011	2017
Formal RW	5.1	5.3	5.2
Semi-formal RW	2.8	4.1	5.6
Informal RW	7.2	9.6	12.3
Casual	30.5	30.5	25.3
Self + Unpaid	54.5	50.5	51.6

Sources and notes - NSSO EUS and PLFS various years. See text for definitions. RW - regular wage.

However, by 2017, the share of women in regular wage work had more than doubled to 25 percent. For men too the share had increased, but by a smaller extent to 24 percent. Strikingly, of the 35 million

Box 5.2: Adivasi women from Jharkhand – Stories behind regular wage work

Large secondary datasets such as the PLFS provide important representative information on the evolving employment structure. As we have seen, the rise in proportion of regular wage work, especially for women, is a key finding from this data. But behind the category of "regular wage" lies a lot of variation in conditions and experiences of work. To understand this variation and to get a concrete sense of the conditions of work in various sectors, detailed surveys and qualitative interviews are more useful. Anumeha Yadav conducted a series of interviews with young women in Jharkhand who were part of various skilling programmes including the PM Kaushal Vikas Yojana (PMKVY). In these interviews the women, mostly in their early 20s, and from an Adivasi background, speak about their experiences in the training centres as well as in the regular wage factory jobs they were able to secure after training.

This is an excerpt from one of the interviews of a woman who had returned to her village in Jharkhand after working in a garment factory in Bengaluru.

> I am not sure if I wish to go back to the factory. When you work in a place like that, you have to be prepared to be yelled at. At the company, you are given production targets. If you don't manage to complete those, then you have to be prepared to hear a few abuses. The target used to be 80 pieces an hour. But I managed to make only 50 or so an hour. Then that means abuse in Hindi and Tamil, both. They don't beat or hit the workers. But they may make a gesture to your face. It feels worthless then. I don't think I want to go back

there. Because they make one do a lot of night work there. Night work means working at 1 am, 3:30 am, like that. Then you have to start again the next day. Sometimes there was no day off even on Sundays.

At the company, I earned Rs 7,500 a month. If I did night work, then it would go up to Rs 10,000 a month. Then after some time, all of a sudden, when orders came, there was a pressure to work many nights at a stretch. Because there was pressure to work nights regularly then, I left the company and came back home. Though initially I got some stitching, but they change roles suddenly. I became a Helper. I had to fold garments. That meant stand all the time, be on your feet continuously and fold garments. You have to stand from 8 am till 730 pm.

Another interviewee pointed out lacunae in the way training centres operated.

> Even after attending (PMKVY) training once, I am training a second time because I feel that training was incomplete. In the PMKVY centre, in the three month course, the first one and a half month we were taught about machines in textbooks. When we were due to leave for the factory in a day or two, for the city, we were made to attend practical classes. In these, we had to cut paper in the shape of clothes. Can one really learn this way? We were given cloth to practise stitching only once, to stitch an undergarment, a tiny piece of cloth. Now I am learning and training a second time investing my own funds. That training was free of cost. I am now paying a woman ₹2,000 a month to teach me stitching.

Box 5.3: Investigating the recent formalisation of regular wage jobs: a gendered story

Between 2017 and 2021, the period that saw a slowdown in overall regular wage job creation, we observe a reversal of trends - there is a shift towards more formal jobs. Jobs with a written contract and benefits as a share of all regular wage work rose from 25 percent in 2017 to 35 percent in 2021, close to the 2004 level. To understand this shift better, we examined the period between 2017 and 2021 more closely. Of the 13 million increase in regular wage employment seen during this period, 11 million occurred during the period between 2017 and 2019. Notably, nearly half of this increase accrued to women, which is sizable given that women typically account for only 20 percent of overall regular wage employment. Of this increase in regular wage jobs, 34 percent were of the formal kind in case of women, while a much larger 72 percent were formal for men. Therefore, although 2017 to 2019 saw an expansion in regular wage employment for women, the majority of this employment was of the informal and semi formal type. This is the opposite of what

occurred in the preceding ten years.

In 2020-21, the pandemic year that saw an overall contraction of the workforce, regular wage employment too fell by 2.2 million. Interestingly though, this period saw an increase in formal employment by 3 million and a loss of about 5.2 million of semi and informal regular wage employment. While half of the lost employment is accounted for by women, only a third of the increase in formal employment accrued to women. By 2021-22, as the economy recovered from the lockdown, an additional 4.5 million regular wage jobs were created. Women in formal regular wage employment increased during this period by 0.9 million, however their numbers in semi and informal employment declined by 1.5 million thereby resulting in a net contraction of women in regular wage employment by 0.6 million. Therefore, the majority of regular wage job creation in the recent past has accrued to men rather than women.

increase in regular wage workers during this period, 26 million were women and 9 million were men.

Even more interestingly, of the increase in regular wage work among women between 2004 and 2017, 18 percent was in jobs with contracts and benefits (and about 40 percent each was in the other two types), while for men only two percent of the new jobs were in this category. For men, the majority (52 percent) of the new regular wage jobs were informal and had neither benefits nor a written contract. Conversely between 2017 and 2021, both men and women have seen a relatively higher increase in the number of formal jobs. We discuss this phenomenon in Box 5.3.

In 2004, the largest provider of the best kind of regular wage jobs was the public sector, accounting for a quarter of such jobs. By 2017, this sector has become the second largest employer, with the education sector accounting for 27 percent of these jobs. The two million increase in formal jobs seen during this period came largely from education, financial services, insurance and real estate, and health services. Notably, the absolute number of salaried jobs in the public sector contracted during this period. The contraction in public sector employment largely explains the relatively higher fall in men's secure salaried employment since this is the predominant source of good jobs for men, whereas for women it is typically in the education and health sectors.

Box 5.4: Employment precarity and social exclusion: A short-run analysis of labour market transitions using panel data

By Rosa Abraham and Surbhi Kesar

Even though the employment structure of the Indian economy has been slow to change at the aggregate level, recent studies have identified significant churn or transitions of the working population across different employment types or arrangements (Kesar 2023). In this work, we push this line of enquiry further by using high frequency panel data between 2017 and 2019, across eight time points (covering more than two and a half years), to map the patterns of labour market transitions across employment arrangements- namely, out of work-force, daily wage work, temporary salaried work, self-employment and permanent salaried work, and understand some of the dynamics of the labor market.

We use the Centre for Monitoring Indian Economy (CMIE)'s Consumer Pyramid Household Survey (CPHS) data. The CMIE-CPHS interviews households three times in a year, with an interval of four months giving us a unique high frequency panel. During the interview, information on the employment status of all individuals in the household is collected, including the type of employment as well as the corresponding earnings. While concerns have been raised on CMIE-CPHS's sampling, data related to employment and earnings remains broadly comparable to national estimates with some caveats (Jha and Basole 2022; Abraham and Shrivastava 2022; Drèze and Somanchi 2022).

For those in casual wage work during the first time point of our analysis, only about 56 percent continue in this employment arrangement by the

Figure 1: Employment transitions over 8 waves in CPHS
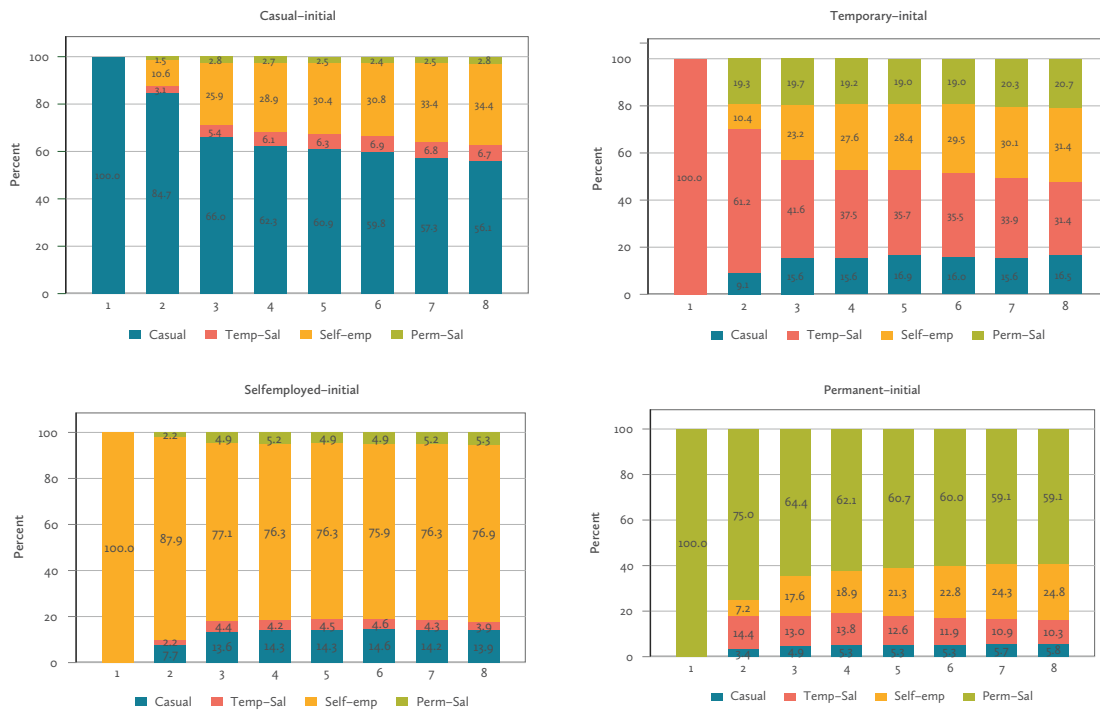
Sources and notes: CMIE Consumer Pyramids Household Survey

last time point, while the other 44 percent have transitioned to other employment arrangements over this period. Similar transitions are seen for other employment arrangements as well (Figure 1). Notably, self-employed shows the most stability, with 77 percent of self-employed in period one continuing to be self-employed by the eighth time point.

This suggests that the unchanged patterns in the employment structure of the overall labour market hides the massive churn within and across employment arrangements that is "netted out" to keep the overall structure unchanged. In order to systematically capture this voluminous amount of transitions in the economy, we undertake a trajectory analysis, following Nagin (2005) and Virtanen et al.(2011). Trajectory analysis, also called "group-based modelling of development", measures the course (or trajectory) of an outcome over a time period, and allows us to identify the distinct clusters of individual trajectories. A trajectory is defined by a polynomial function. In the case of this analysis,

this is assumed to be a multinomial function and a maximum likelihood strategy is employed for the estimation of parameters.

We first order the different outcomes, in our case the employment arrangements, based on the monthly earnings for each employment arrangement as a measure of 'formality' and 'less precarity'. The ordering of employment is as follows: - out of workforce (0) , daily wage workers (1), temporary salaried workers (2), self-employed (3), permanent salaried (4). This also denotes the Y axis in Figure 3 below. We employ the trajectory analysis to identify the distinct clusters or trajectory groups and estimate the probability of individuals to belong to each of those trajectory groups. Each individual is assigned to that trajectory group for which the likelihood is the highest.

We identify seven dominant trajectories in the Indian labour market (Figure 2). The percentages next to each trajectory line in the figure indicates the share of population that belong to this group.

Figure 2: Major employment trajectories
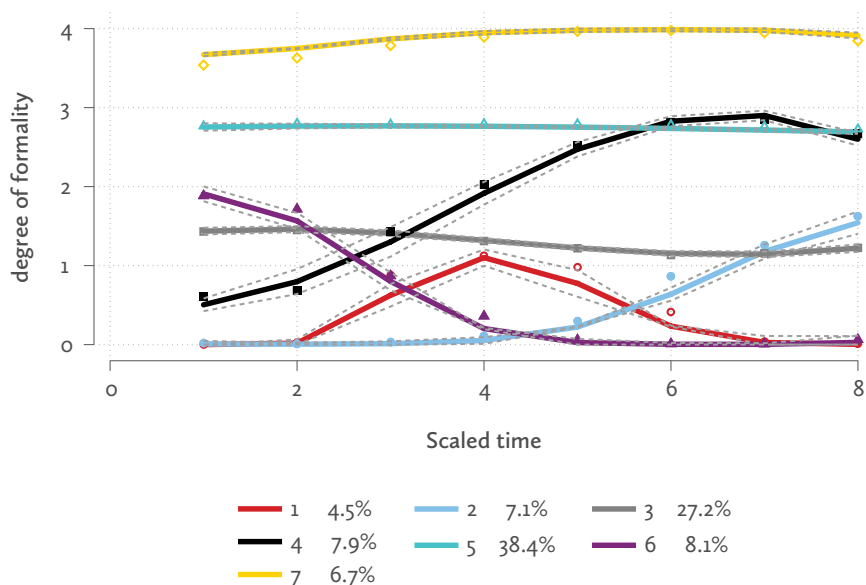
Sources and notes: CMIE Consumer Pyramids Household Survey

From out of wage work to - casual wage work - to out of wage work (4.5%)

Moving from out out of workforce towards informal wage work (7.1%)

Always informal wage work (27.2 %)

Moving into informal wage work / self-employment (7.9 %)

Always self-employed (38.4%)

Moving out of informal wage work and exiting the workforce (8.1%)

Always permanent salaried (6.7%)

A few notable things appear: First, there is a very small proportion of the workforce, about 6.7 percent, that makes up the consistently permanent salaried (likely the formal employment) trajectory group. In fact, there is almost no movement in the labour market from the various informal employment arrangements into the permanent salaried arrangement. Second, there is a huge proportion of the workforce concentrated across various informal arrangements, and is marked by a high degree of transition across these various employment

arrangements. Notably, the most dominant trajectory group among the workforce, i.e., comprising 38.4 percent of workers, is stable self-employment. The next major trajectory group, comprising 27 percent of the total workforce, are informal wage workers. Next, at about 8 percent each, are the trajectory groups that denote movement from out of the workforce toward different informal employment arrangements, including daily wage, temporary salaried, and self-employment, and those denoting movement from informal wage work towards out of the workforce.

The CMIE-CPHS data categorises households into Scheduled Caste, Scheduled Tribes, General, Intermediate castes and Other Backward Classes. Based on overall demographic similarities, we club the Intermediate and General castes together, while keeping SC and ST together. We use a simple representation index (RI) to understand what is the relative share of caste groups in each trajectory. The representation index is the share of a particular group in a

Figure 3: Caste-based representation index for the dominant trajectory groups in the Indian labour market
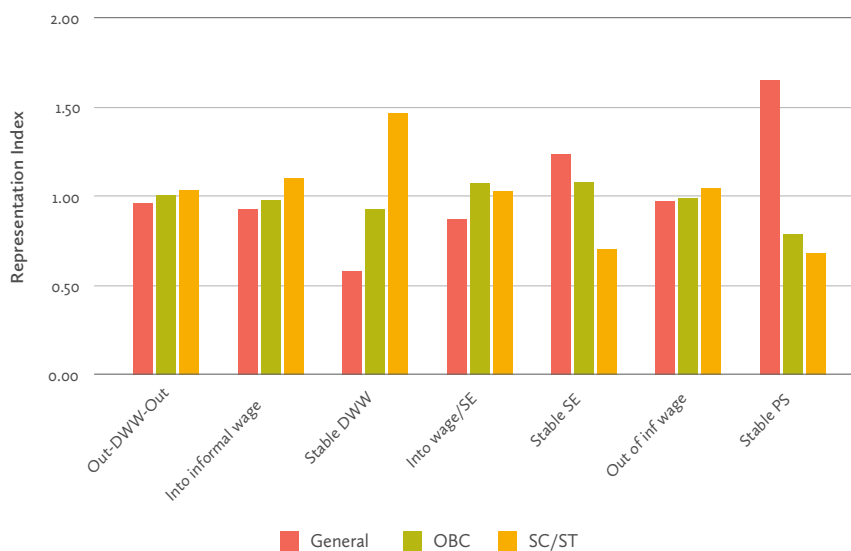
Sources and notes: CMIE Consumer Pyramids Household Survey

trajectory divided by their share in the overall workforce. A value of one indicates that a group is proportionately represented in a particular trajectory group, above one indicates over-represented while less than one indicates that their share in the trajectory is less than their share in the overall workforce.

Beginning with the most secure and stable of the trajectory group, which is also associated with the highest earning - always permanent salaried - we find that general caste workers are over-represented (with RI of 1.65), while the OBC and SC/ST are under-represented, with RI of 0.79 and 0.58, respectively. In sharp contrast, on the other hand, in the stable informal wage work trajectory, which comprises about 27 percent of the population and captures those who are consistently in precarious informal wage work, the SC/ST group is over-represented, with RI of 1.46, whereas the general category groups have least representation, with RI of 0.58. In the case of the stable self-employment trajectory, which is the most prominent trajectory in the

country, the General caste workers tend to be over-represented (with RI of 1.23) and SC/ST are under-represented (with RI of 0.7). While self-employed also tends to be mostly informal in nature, the earnings and consumption level for those deriving livelihood from this employment arrangement is better than those involved in casual wage work.

In the Indian labour market, marked by a high volume of churn as workers transitions across employment arrangements, a very small proportion of the workforce exhibits the stable salaried trajectory and a high proportion seems to be oscillating between different informal employment arrangement, with the likelihood of workers to climb the job ladder from informal arrangements onto salaried work being insignificant. These dynamics remain differentiated across caste lines.

5.2 Intergenerational mobility

In Chapter Three we took a macroeconomic perspective on the effectiveness of growth in delivering structural change for various social identities. If the structural transformation process is inclusive and successful in bringing previously excluded groups into dynamic and remunerative parts of the economy, then we should be able to see evidence for that in the form of intergenerational mobility at the household level as well. That is, on average sons and daughters should have better economic opportunities than fathers and mothers. This can be in the form of better quality jobs, better paying occupations and so on. Of course, educational mobility (sons and daughters

being more educated than fathers and mothers) is equally important and lies behind occupational and employment mobility (Jalan and Rinku Murgai 2015). In this section we examine mobility within caste and religion groups in terms of the kind of employment that a son has compared to his father. We also use cohort-level analysis to get some insights on intergenerational mobility for women. We discuss educational convergence by castes and jatis in Section 5.3.

5.2.1 Intergenerational mobility between fathers and sons

The lack of long-term panel data in developing countries like India, presents a significant challenge for analysis. Nevertheless, given the prevalence of households in India where multiple generations

of a family co-reside, examining intergenerational mobility through cross-sectional data can yield some insights into the phenomenon. Several papers have examined intergenerational mobility using NSSO EUS data (Reddy 2015; Hnatkovska, Lahiri, and Paul 2013; Majumder 2010). Reddy(2015) analysed mobility patterns from 1983 to 2012 whereas Majumder(2010) carried out the study for the time period 1993-2004. Both find that occupational mobility was comparatively less for Scheduled Castes and Tribes compared to others.[3] Hnatkovska et al.(2013) conducted a study on intergenerational occupational mobility between 1983 and 2004-05. The study found that there was an increase in the probability of intergenerational occupational switches among non-SC/ST males (includes both OBC and General castes), from 33 percent to 42 percent and among Scheduled Castes and Scheduled Tribes males, from 30 percent to 39 percent. The authors concluded that there was no significant convergence in intergenerational mobility between Scheduled Caste and Scheduled Tribe males and others during the period of study.

Here we examine the period after 2004-05. While previous studies have largely focused on occupational mobility, we take a different approach by classifying individuals based on their employment arrangements and sector of work. We are interested in examining mobility across regular salaried work, casual wage work and self-employment and away from agriculture. This is particularly pertinent given the structural transformation of the Indian economy in the past four decades with a steady increase in the share of non-agricultural salaried workers (see Chapter Three). Further, given the increasing informalisation of salaried work (State of Working India (2018) and Section 5.1 above), we differentiate between different degrees of informality within salaried work. Thus, we examine intergenerational mobility within caste and religious groups along two dimensions that capture the extent of the Lewis and Kuznets processes.

Data comes from the 61st (2004–05) and 68th

rounds (2011–12) of NSS EUS and the 2018-19 Periodic Labour Force Survey (PLFS). The 2019-20 and subsequent rounds of PLFS include Covid-19 pandemic effects, hence we do not include them. Our sample is restricted to employed fathers and sons who co-reside in the same household. Using information in the household roster (esp. relationship to the head of the household) for each individual in the sample, we map sons aged between 24 and 45 to their fathers.[4]

The most widely used method of measuring intergenerational mobility involves constructing a mobility matrix. This intergenerational mobility (IGM) matrix displays the percentage of sons in each category, given that their father is in a particular category. The next few tables present these mobility matrices (Tables 5.2 to 5.4). The diagonal cells in the mobility table show the percentage of sons in the same type of work as their father, that is immobility or intergenerational persistence. Values above the diagonal indicate upward mobility and those below it represent downward mobility. Warmer colours represent higher values. Thus, if the growth process is enabling sons to avail of better opportunities than their fathers, we would like to see warmer colours representing higher values above the diagonal.

Table 5.2 shows the IGM for three points in time, 2004-05, 2011-12 and 2018-19. At any point in time, the employment type with highest intergenerational persistence or "stickiness" is casual wage work, closely followed by self-employment. Formal regular wage work is far less sticky. Of course, it is possible that this is in part a life-cycle effect. That is, as sons age, their chances of securing better quality formal jobs may go up. But even with that caveat, in general it appears intergenerational persistence increases as the type of employment grows worse.

Encouragingly, overall mobility rose over the period of analysis. Note that, in 2004, the top half of the IGM matrix was dominated by cool colours and the bottom half by warm colours. But by 2018, we see warm colours appearing in the top half. Thus

Table 5.2: Intergenerational mobility matrices show rising upward mobility over time

Year 2018

Son's Employment	Father's employment				
	Formal RW	Semi-formal RW	Informal RW	Self-employed	Casual
Formal RW	33.8	4.9	8.6	4.4	3.1
Semi-formal RW	13.3	47.3	9.7	6.1	4.8
Informal RW	19.5	19.9	58.6	13.2	15.4
Self-employed	26.5	23.6	15.4	65.1	8.0
Casual	6.9	4.3	7.8	11.2	68.8
Col sum	100	100	100	100	100

Year 2011

Son's Employment	Father's employment				
	Formal RW	Semi-formal RW	Informal RW	Self-employed	Casual
Formal RW	25.8	4.2	0.7	1.9	0.4
Semi-formal RW	12.6	40.0	4.4	3.0	1.5
Informal RW	20.9	27.8	61.1	8.4	7.9
Self-employed	34.3	20.6	20.4	76.4	9.5
Casual	6.5	7.4	13.4	10.3	80.7
Col sum	100	100	100	100	100

Year 2004

Son's Employment	Father's employment				
	Formal RW	Semi-formal RW	Informal RW	Self-employed	Casual
Formal RW	14.6	1.5	2.8	1.9	0.3
Semi-formal RW	8.1	13.8	1.5	1.7	0.9
Informal RW	15.6	18.6	61.4	6.4	6.2
Self-employed	51.8	57.8	19.2	80.2	8.1
Casual	10.0	8.4	15.1	9.9	84.6
Col sum	100	100	100	100	100

Sources and notes: NSSO EUS and PLFS various rounds. Numbers are percentages adding to 100 along the columns. Warmer colour indicate higher values. RW - regular wage.

the intergenerational persistence of casual wage work declined between 2011 and 2018. In 2011, 80 percent of sons of casual wage working fathers were themselves casual workers. By 2018 this had fallen to less than 70 percent and the percentage of sons in better types of work rose between the two years. Correspondingly, the stickiness of regular wage work rose in the same period.

Table 5.3 shows a set of IGMs disaggregated by caste. The movement out of casual wage work

carries particular significance for Scheduled Caste and Scheduled Tribe workers due to their historical disadvantages in terms of access to property as well as education. In 2011, 87 percent of sons of casual workers belonging to the SC/ST community were themselves casual workers and this had not changed substantially since 2004. But in 2018, this number dropped to 75 percent.

But a clear difference emerges across caste in the pace of this decline. As of 2004 over 80 percent

Table 5.3: Intergenerational mobility matrices over time disaggregated by caste

SC/ST

Year 2018

Son's Employment	Father's employment				
	Formal RW	Semi-formal RW	Informal RW	Self-employed	Casual
Formal RW	38.6	0.2	2.2	3.0	2.0
Semi-formal RW	10.5	43.3	6.1	3.8	2.7
Informal RW	20.9	22.2	64.1	11.0	14.1
Self-employed	16.7	26.5	11.8	62.9	5.6
Casual	13.3	7.9	16.0	19.2	75.6
Col Sum	100	100	100	100	100

Year 2011

Son's Employment	Father's employment				
	Formal RW	Semi-formal RW	Informal RW	Self-employed	Casual
Formal RW	21.6	0.2	1.1	1.3	0.0
Semi-formal RW	8.7	47.1	9.5	1.9	1.3
Informal RW	28.0	30.0	43.0	8.9	5.2
Self-employed	25.6	14.7	13.6	71.0	6.6
Casual	16.1	8.0	32.7	17.0	86.9
Col Sum	100	100	100	100	100

Year 2004

Son's Employment	Father's employment				
	Formal RW	Semi-formal RW	Informal RW	Self-employed	Casual
Formal RW	14.8	2.7	0.2	1.6	0.2
Semi-formal RW	7.0	14.4	0.8	1.4	1.0
Informal RW	23.4	18.8	70.4	5.0	5.6
Self-employed	37.8	45.1	19.7	74.7	6.7
Casual	17.0	18.7	8.8	17.4	86.5
Col Sum	100	100	100	100	100

Sources and notes: NSSO EUS and PLFS various rounds. Numbers are percentages adding to 100 along the columns. Warmer colour indicate higher values. RW - regular wage.

of sons of casual wage working fathers were themselves casual wage workers, for all three caste groups. By 2018 this had declined the fastest for Other castes and slowest for SC/ST households, with OBCs in the middle. Among the General category, in 2018, only 53 percent remained in casual wage work like their fathers, and about 23 percent moved into informal regular wage work. In contrast, fewer proportion of SC/ST households saw such mobility (about 14 percent).

Interestingly, the stickiness in formal regular wage work is also the highest for SC/ST workers in 2018. Movement out of formal regular wage work

In 2004 over 80% of sons of casual workers were themselves casual workers, for all three caste groups. By 2018 this had declined the fastest for Other castes and slowest for SC/ST households, with OBCs in the middle.

Table 5.3: (cont'd) Intergenerational mobility matrices over time disaggregated by caste

OBC

Year 2018

Son's Employment	Father's employment				
	Formal RW	Semi-formal RW	Informal RW	Self-employed	Casual
Formal RW	29.2	11.9	4.9	3.8	4.0
Semi-formal RW	11.0	47.3	9.1	5.5	6.6
Informal RW	22.4	15.8	62.7	13.7	14.5
Self-employed	32.6	19.4	17.5	67.1	9.8
Casual	4.9	5.6	5.9	9.8	65.1
Col Sum	100	100	100	100	100

Year 2011

Son's Employment	Father's employment				
	Formal RW	Semi-formal RW	Informal RW	Self-employed	Casual
Formal RW	21.1	3.3	0.2	1.7	0.7
Semi-formal RW	16.9	37.6	0.0	3.0	1.7
Informal RW	18.3	28.6	63.7	7.9	8.7
Self-employed	40.4	24.1	24.2	77.1	10.3
Casual	3.3	6.5	11.9	10.4	78.7
Col Sum	100	100	100	100	100

Year 2004

Son's Employment	Father's employment				
	Formal RW	Semi-formal RW	Informal RW	Self-employed	Casual
Formal RW	10.3	0.9	2.9	1.3	0.4
Semi-formal RW	6.7	10.7	3.0	1.4	0.6
Informal RW	15.8	18.0	60.2	6.8	8.0
Self-employed	61.6	67.4	18.3	82.1	8.1
Casual	5.7	3.0	15.5	8.4	82.9
Col Sum	100	100	100	100	100

Sources and notes: NSSO EUS and PLFS various rounds. Numbers are percentages adding to 100 along the columns. Warmer colour indicate higher values. RW - regular wage.

into less formal types or into self-employment is higher for OBC workers as compared to SC/ST workers. Higher stickiness in both casual work and formal salaried work for SC/ST workers suggests a bimodal distribution. Of course, it should be kept in mind that the proportion of the SC/ST workforce in regular wage employment is itself quite small compared to the proportion for other castes. Even

for the Other castes, we see that only 34 percent of sons reported formal regular wage jobs conditional on their father holding such a job.

Between Hindu and Muslim households, the two salient distinctions that emerge in Table 5.4 are as follows. Movement out of casual salaried jobs is mainly into regular salaried jobs for Hindus, whereas

Table 5.3: (cont'd) Intergenerational mobility matrices over time disaggregated by caste

Others

Year 2018

Son's Employment	Father's employment				
	Formal RW	Semi-formal RW	Informal RW	Self-employed	Casual
Formal RW	34.4	2.5	16.8	6.3	4.7
Semi-formal RW	17.6	50.0	12.5	9.1	7.6
Informal RW	15.7	21.7	50.4	14.3	22.6
Self-employed	28.5	25.0	15.1	64.0	12.1
Casual	3.9	0.9	5.2	6.5	53.0
Col Sum	100	100	100	100	100

Year 2011

Son's Employment	Father's employment				
	Formal RW	Semi-formal RW	Informal RW	Self-employed	Casual
Formal RW	30.7	7.8	1.0	2.6	0.2
Semi-formal RW	12.2	37.1	5.9	3.8	1.4
Informal RW	18.6	25.5	68.1	8.8	12.2
Self-employed	35.3	21.9	20.3	79.0	14.3
Casual	3.2	7.7	4.8	5.8	71.9
Col Sum	100	100	100	100	100

Year 2004

Son's Employment	Father's employment				
	Formal RW	Semi-formal RW	Informal RW	Self-employed	Casual
Formal RW	17.3	1.3	4.2	2.7	0.0
Semi-formal RW	9.7	15.5	0.2	2.2	1.2
Informal RW	11.2	19.0	57.6	6.8	3.5
Self-employed	52.9	58.2	19.8	80.9	12.1
Casual	9.0	6.2	18.2	7.5	83.2
Col Sum	100	100	100	100	100

Sources and notes: NSSO EUS and PLFS various rounds. Numbers are percentages adding to 100 along the columns. Warmer colour indicate higher values. RW - regular wage.

it is primarily into self-employment for Muslims. And there is higher stickiness in formal regular wage jobs among Hindus compared to Muslims.

5.2.2 Intergenerational mobility for women[5]

Standard intergenerational mobility comparisons, such as the one we have just seen, compare fathers with their adult sons in the same household. This approach does not work for mother-daughter comparisons since women are typically not a part

Table 5.4: Intergenerational mobility matrices over time disaggregated by religion

Hindus

Year 2018

Son's Employment	Father's employment				
	Formal RW	Semi-formal RW	Informal RW	Self-employed	Casual
Formal RW	34.7	4.6	9.4	4.3	2.6
Semi-formal RW	13.2	50.7	10.3	6.5	4.9
Informal RW	19.3	16.5	57.3	13.3	15.7
Self-employed	26.0	24.1	15.5	65.4	5.8
Casual	6.8	4.1	7.5	10.5	71.0
Col Sum	100	100	100	100	100

Muslims

Year 2018

Son's Employment	Father's employment				
	Formal RW	Semi-formal RW	Informal RW	Self-employed	Casual
Formal RW	23.0	7.9	1.6	4.0	5.4
Semi-formal RW	16.4	23.9	6.7	3.4	3.5
Informal RW	18.4	37.8	69.5	13.5	15.0
Self-employed	36.0	25.4	11.9	62.9	15.9
Casual	6.3	5.0	10.4	16.2	60.2
Col Sum	100	100	100	100	100

Sources and notes: NSSO EUS and PLFS various rounds. Numbers are percentages adding to 100 along the columns. Warmer colour indicate higher values. RW - regular wage.

of their natal household owing to patrilocal norms around marriage. Mothers and daughters are thus not observed together in a household survey. A cohort level analysis allows us to track workers over their lifetime as well as across multiple generations.

In the Indian context, cohort-based analyses have been used to understand inter-caste social mobility. Hnatskova et al.(2012) explore the evolution of educational attainment, occupation and wages across different generations. They find a convergence in educational attainment and occupational distribution between SC/STs and non SC/STs, and a decline in the non SC/ST wage premium between 1983 and 2004-05. Deshpande and Ramachandran (2019) use birth cohorts and compare SC/ST, OBC and General category workers. They find that while absolute and relative gaps in

primary education have narrowed across castes, in secondary education, there has been decline in relative gaps, but not absolute gaps. Similarly, in the lower earnings category too, there has been a convergence, but not so in the above-median wages.

We use seven rounds of NSSO employment surveys to track cohorts over time. Beginning with the first of these, we identify the first or earliest cohort as individuals aged between 18-24 years in 1983.[6] The first cohort, by the time of the next EUS round in 1987-88, would belong to the 22-28 age group. In this manner, we track this cohort over subsequent EUS rounds, till 2018, till when they are between 52 to 58 years old. Likewise, we construct the second cohort by identifying individuals who would have freshly entered the labour market in that year, i.e.

Table 5.5: Construction of age-cohorts for women from NSSO-EUS surveys

NSS Year	1983	1987	1993	1999	2004	2011	2018	Number of years tracked	Birth year
Cohort1	18-24	22-28	27-33	33-39	38-44	45-51	52-58	34	1959-1965
Cohort2		18-24	23-29	29-35	34-40	41-47	48-54	30	1964-1970
Cohort3			18-24	24-30	29-35	36-42	43-49	25	1969-1975
Cohort4				18-24	23-29	30-36	37-43	19	1975-1981

18-24 years in 1988. Table 5.5 provides the details of each cohort. In total, we identify four cohorts, with the first being the generation born between 1959 and 1965, and potentially entering the labour market in 1983, and the final cohort being individuals born between 1975-1981 and entering the labour market in 1999. Note that this approach does not follow particular individuals over time (the data does not permit us to do that). Rather it follows an age group over time.

We have seen that the Indian economy has seen a steady increase in the share of salaried work with a reduction in casual wage work. We now examine how the increase in salaried employment manifests across cohorts of women over their lifetime. Typically, as individuals gain experience at work and build networks and social capital, there is an expected increase in their earnings as well as the kind of work they are engaged in. Individuals are expected to shift to more secure work arrangements like salaried work with job security.

For the earliest cohort, there is a steady increase in access to salaried work as women age (Figure 5.7). In addition, similar to men (not shown), there is a level shift in the cohorts, indicating that more recent cohorts have a greater likelihood of being salaried workers reflecting the general expansion of salaried work in the economy over the years. However, for the most recent cohort, there is a slowing down in the share in salaried work with age, and beyond the age of 30, the share in salaried work among the youngest cohort actually begins to fall. This suggests that the increase in salaried work

seen in a cross-sectional analysis is coming from women of older generations. More recent entrants to the labour market are less likely to have access to salaried work, compared to their counterparts of the same age from an earlier generation.

5.3 Educational Convergence across Social Groups

Caste based inequality in access to education is well documented and is regarded as one of the primary reasons for continued socio-economic disadvantage of the Scheduled Castes. A significant volume of research on caste and education shows that while affirmative action helped in increasing the educational status of marginalised social groups, inequality continues to persist at all levels of education, and specifically in higher education (Thorat and Khan 2023; Asher, Novosad, and Rafkin 2021; Cassan 2019; Deshpande 2013; S. Deshpande 2006). Empirical evidence on inequality at the level of jatis is more limited. It can be argued that in theory, all jatis within the SC category hold the right to reservation policy, and thus, the experience at the jati level may not deviate much from the average educational outcomes of SCs as a group. On the other hand, one can argue that some jatis may have experienced more educational progress than others due to differences in political assertion and other historical factors.

Figure 5.7: Proportion of regular wage workers among women in various age cohorts
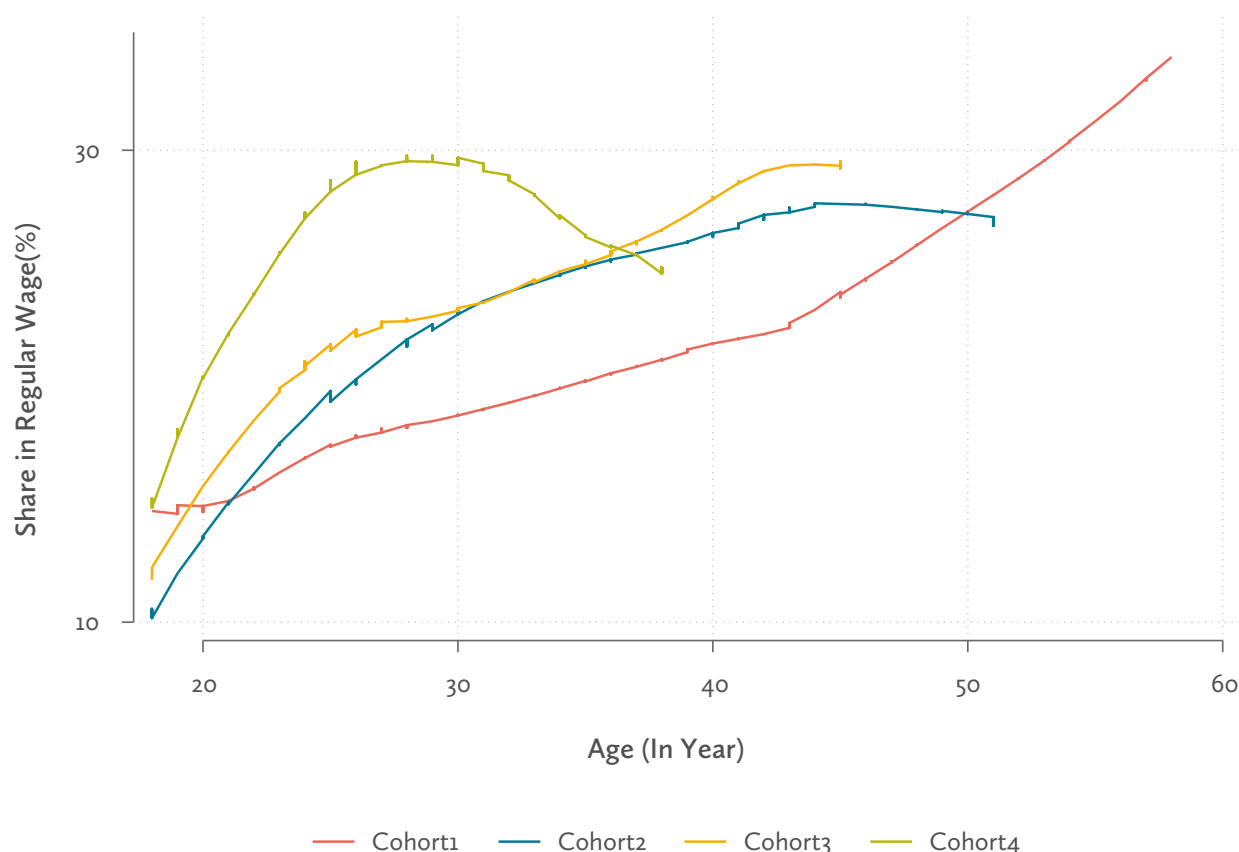
Sources and notes: NSS EUS and PLFS various years. Age-specific cohorts are constructed for women workers from various survey years. See Table 5.5 and text for cohort details.

As discussed in Chapter Three, one source of jati level data, albeit dated and only for Scheduled Castes, is the 2011 Population Census. Though jati information is collected for other caste categories in household surveys such as the India Human Development Survey (IHDS) and the CMIE Consumer Pyramids Survey, this information is not easy to use, as we discuss in Chapter 7. We discuss this issue further in Chapter Seven. Here we use data from the 2011 Population Census to compare the differences within jatis with the differences between constitutional categories (SC versus non-SC). Similar to Chapter Three, we restrict our analysis to jatis that have a share of 0.5 percent or higher within the SC population.

5.3.1 Convergence below secondary, divergence post secondary

Before proceeding to the jati level analysis, we first examine the extent of convergence for various educational levels at the level of broad caste categories. Each panel in Figure 5.8a plots the share of SCs and non-SC/STs for different levels of education in 1991 versus the change in the same between 1991 and 2011. Twelve states are represented -Andhra Pradesh, Bihar, Gujarat, Karnataka, Kerala, Madhya Pradesh, Maharashtra, Punjab, Rajasthan, Tamil Nadu, Uttar Pradesh and West Bengal. By construction, a negative slope of the fitted line indicates that the states with lower educational attainment in 1991 experienced higher increases between 1991 and 2011, implying convergence across states.

We observe that this holds true for both the groups, at all levels of education up to the secondary. Beyond the secondary level, there is no evidence of convergence and if anything we observe some

divergence (i.e. states with high initial levels saw faster increases in higher education attainment). There are also some interesting differences across the social groups. While the speed of convergence across states was similar for SCs and non-SCs below the secondary school level, at secondary level convergence was much faster for non-SC castes.

Convergence in educational attainment is observed at all levels of education up to the secondary. Beyond the secondary school level, there is no evidence of convergence.

Note also that the educational gap in the initial year (1991) increases as the level of education goes up and is the largest in the case of higher education. Also, for all groups, the scale of the X-axis for secondary education is the widest, and is comparatively narrower for both primary education and higher secondary. This implies that the overall share of people in higher education, irrespective of social groups, remained quite low throughout the period. Taken together, Figure 5.8a shows that convergence across states stopped at the level of secondary education for both SC and non-SC groups.

The Census data only takes the story till 2011. What has happened in the ten years after that? Figure 5.8b draws on the NSSO EUS and PLFS to show that the broad story of convergence till the secondary level and divergence after that, continues to hold. But on the positive side it also shows that the speed of convergence in secondary schooling increased for the Scheduled Castes post 2011.[7]

The Census data enables us to go beyond the broad SC category and examine educational experiences at the jati level. We now turn to jati-level analysis of educational convergence. As before (Chapter Three) we do not aggregate jatis across states, since the same jati may experience very different trajectories in different states. For example, the experience of Mahars in Maharashtra could be significantly different from that in Karnataka. DIfferentiating

between the same jati in different states, allows us to retain and compare the potentially different trajectories of a jati across states.

Figure 5.9 replicates the analysis shown in Figure 5.8 but this time for all the jatis that comprise more than 0.5 percent of the scheduled caste population in the 12 selected states. Here each point represents a state-jati combination. Again a negative relationship means that jatis with lower educational outcomes in 1991 became educated more rapidly than those with higher outcomes, thereby bridging the gap between relatively better educated and relatively less well educated jatis. It should be kept in mind that this discussion is now entirely within the Scheduled Caste category.

We observe that the overall experience of the larger constitutional categories, translates to the jatis as well. There is convergence among jatis upto the level of secondary education. much like the case of SCs and others across states. For educational levels higher than secondary, we find divergence. That is, whichever jati had a relatively higher share of graduates and/or higher secondary education, has also experienced higher increases between 1991 and 2011.

There is, however, one significant observation in case of jatis that differs from the constitutional categories. For education levels up to secondary, the fitted straight line is not the best approximation of the trend. Rather it seems to resemble an inverted U relationship. That is, jatis that had either lower or higher share at these educational levels (up to secondary), have not experienced much change whereas those at the middle, have. In the Results Appendix we provide figures with non-linear fits to show that this is indeed the case.

This can potentially be attributed to the differences across states. That is, in some states the underlying trend across jatis could be divergent, while in some other states, the jatis could be converging to each other. Since each of the points in these scatter

Figure 5.8: Convergence in educational attainment across states for SCs and Others

a. 1991 to 2011
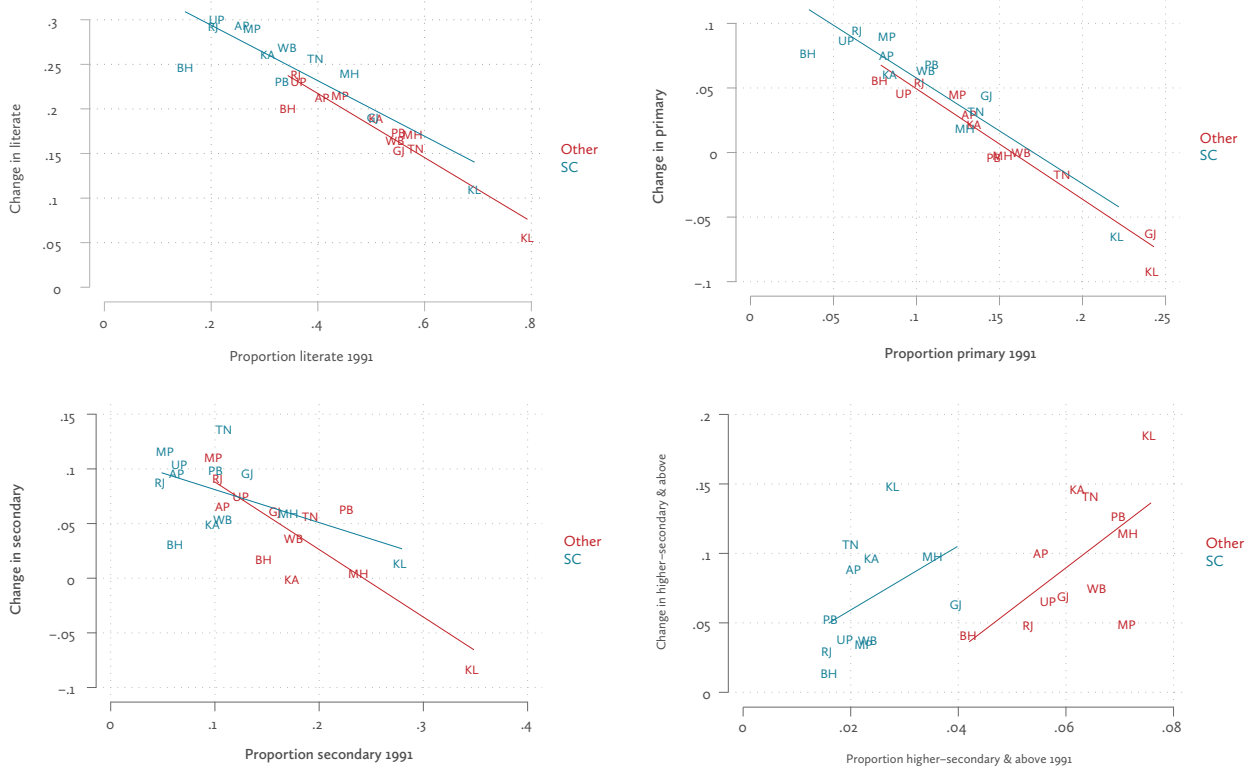
Sources and notes: Population Census 1991 and 2011. Others : Non SC /ST

b. 2011 to 2021
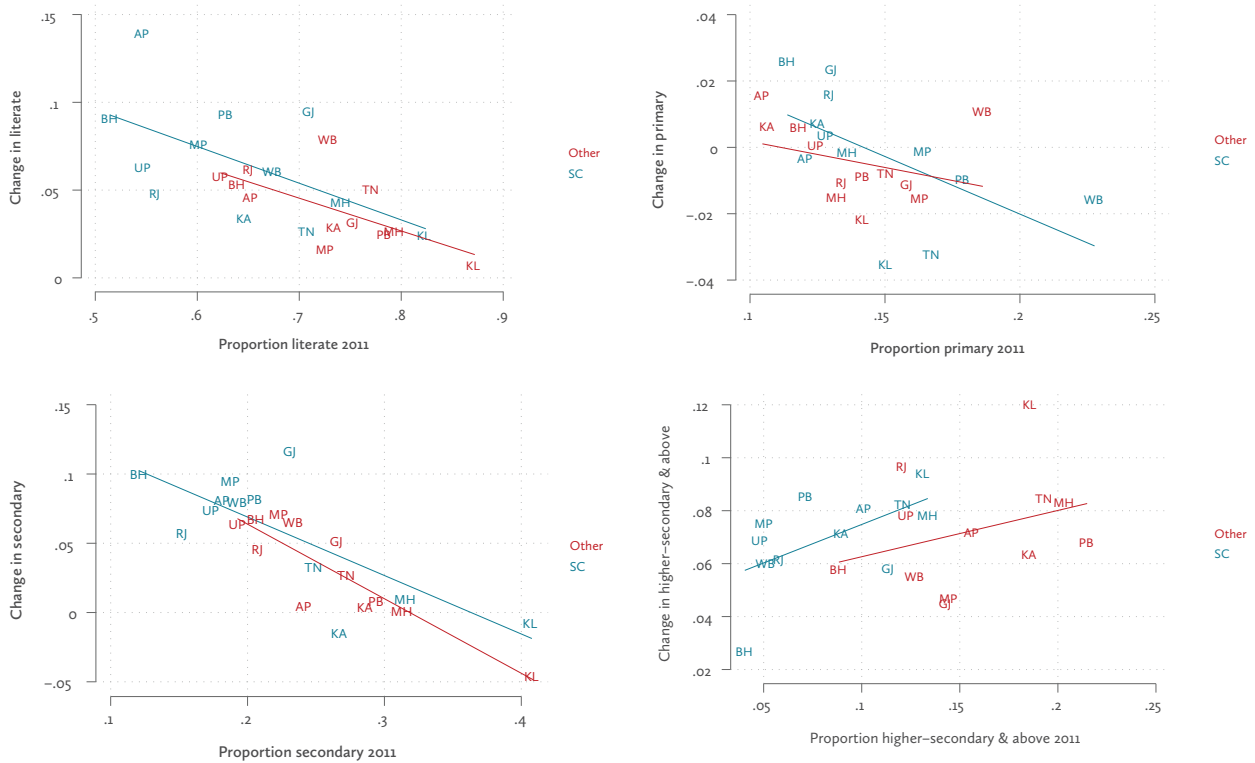
Sources and notes: NSS EUS 2011 and PLFS 2021. Others : Non SC /ST

Figure 5.9: Convergence in educational attainment across states for SC jatis
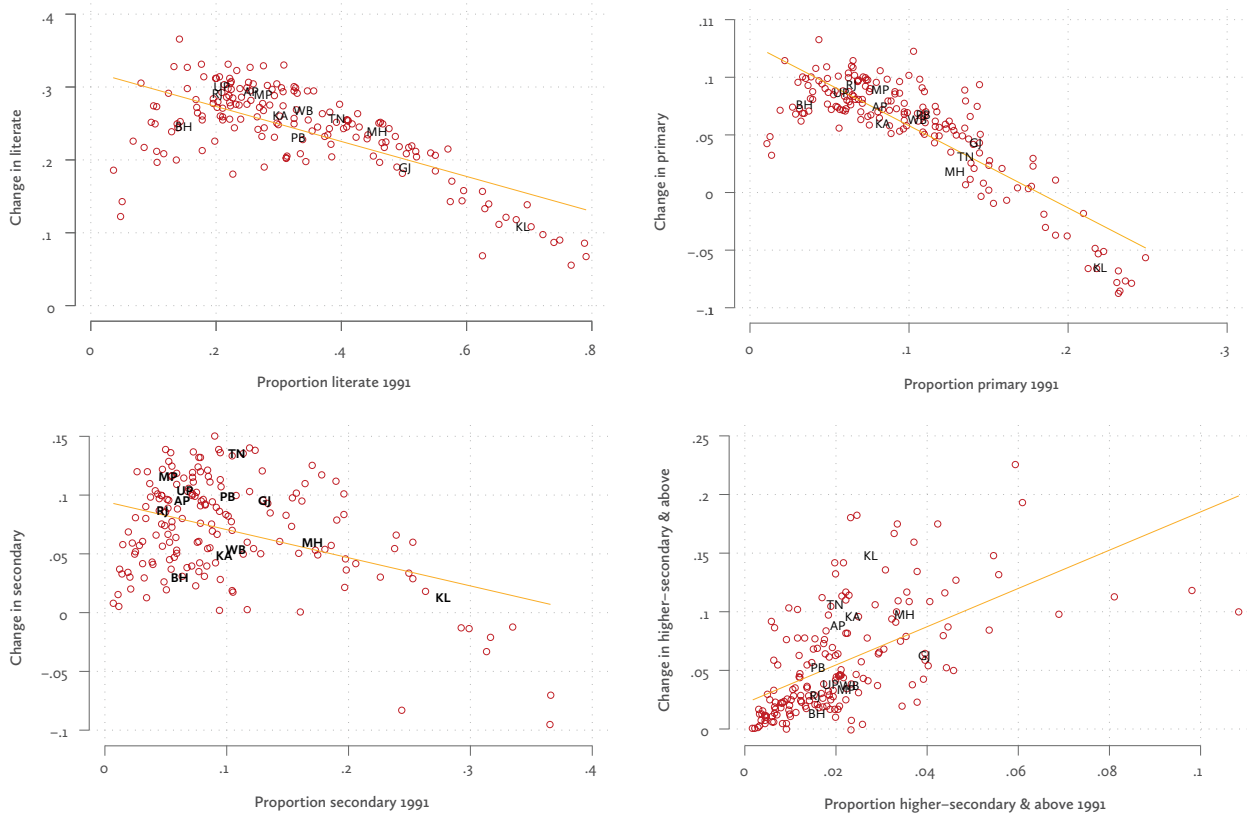
Sources and notes: Population Census 1991 and 2011

plots represent a specific jati in a specific state, the conflicting trends across states may confound each other to cause the inverted-U relationship. This motivates us to aggregate the individual states at the level of regions and take a closer look at the jati level dynamics at a regional level. Thus, we replicate the above exercise for northern (Punjab, Rajasthan, Uttar Pradesh, Madhya Pradesh and Bihar) and southern (Tamil Nadu, Kerala, Karnataka and Andhra Pradesh) states separately.

5.3.2 Convergence in the South, Divergence in the North

Indeed, we find a stark difference in the jati level experiences between the northern and southern states. The earlier noted trend in convergence up to the secondary level of education, is largely driven by the southern, and not the northern states. Figure 5.10 shows the by-now familiar educational

convergence graphs for jatis in the northern (top panels) and southern states (bottom panels). We observe that even for lower levels of education such as the share of literates and primary educated - the southern states show a strong and clear convergence, while the jatis in northern states do not.

We find a stark difference in the jati level experiences between the northern and southern states. Convergence up to the secondary level of education is largely driven by the southern states.

We find further interesting patterns in terms of secondary education. Jatis in the northern states start diverging at the secondary level itself, contrary to the case of south, where we observe convergence across jatis. This is in contrast to the earlier observation for the 12 selected states altogether,

that up to secondary level, jatis converged to each other. Clearly, that convergence therefore was driven by the southern states only. Additionally, one may note that the southern states also had a higher share of literates, as well as the primary and secondary-educated in 1991, implying a historically progressive state of education in these states.

For higher education beyond the secondary level, however, we find there is no north-south divide. There is divergence all around.

To summarise, we find that in terms of education, the broader constitutional categories as well as the jatis within the SCs, reflect similar experiences. Across states, these groups converge to each other up to secondary education, and for subsequent levels, they diverge. However, a look at the regional dynamics reveals that, in our sample of 12 states, most of this convergence is driven by the southern states. This leads us to ask, what are the potential drivers of such divergence across the southern and northern states?

One can think of multiple hypotheses as to why SC jatis in southern states may experience very different trajectories as opposed to the northern states. While a detailed analysis of the same is beyond the scope of this report, some observations from our ongoing work suggest the following.

First, we observe that the state level income indicators do not explain such variation. One may argue that except for Punjab, the northern states we have chosen are all relatively lower income states as compared to the southern states. While this is true, if we take a closer look at the state level stories, specifically in high income states such as Gujarat, Haryana or Punjab, the trends from the other northern states hold true for these relatively richer states as well. Therefore, there is no reason to believe that educational disparities at jati level across states have much to do with the income levels of the states.

Second possible channel could be the differences in educational infrastructure across the states. A look at the U-DISE report (2011) , specifically the four components of the educational index (for elementary education) reveals some more interesting insights. These components are, access to education, educational infrastructure, teacher quality and outcomes. A look at the report reveals that the states chosen for our analysis, whether in the south or north, do not differ significantly in terms of educational access. However, in terms of infrastructure, teacher quality and outcome, on an average the southern states rank higher than the northern states. Further, the southern states rank quite high in terms of the outcome based index, which includes a variable on caste based inclusion in elementary education. Thus, the overall state of education in terms of these three components (infrastructure, teacher quality and outcome) may offer a better explanation of the south-north divide.

While the above provide at least partial explanation of the regional differences, one observation remains unexplained. Even when we split the states in northern and southern regions, we observe an inverted U relationship for northern states. This essentially implies that jatis with medium share at different levels of education in 1991, at least among literates and primary educated, have experienced higher increases in the respective shares between 1991 and 2011. This calls for more detailed explorations than is possible with such secondary data and may require ethnographic work.

Figure 5.10: Southern states display convergence in educational attainment for SC jatis but northern states do not
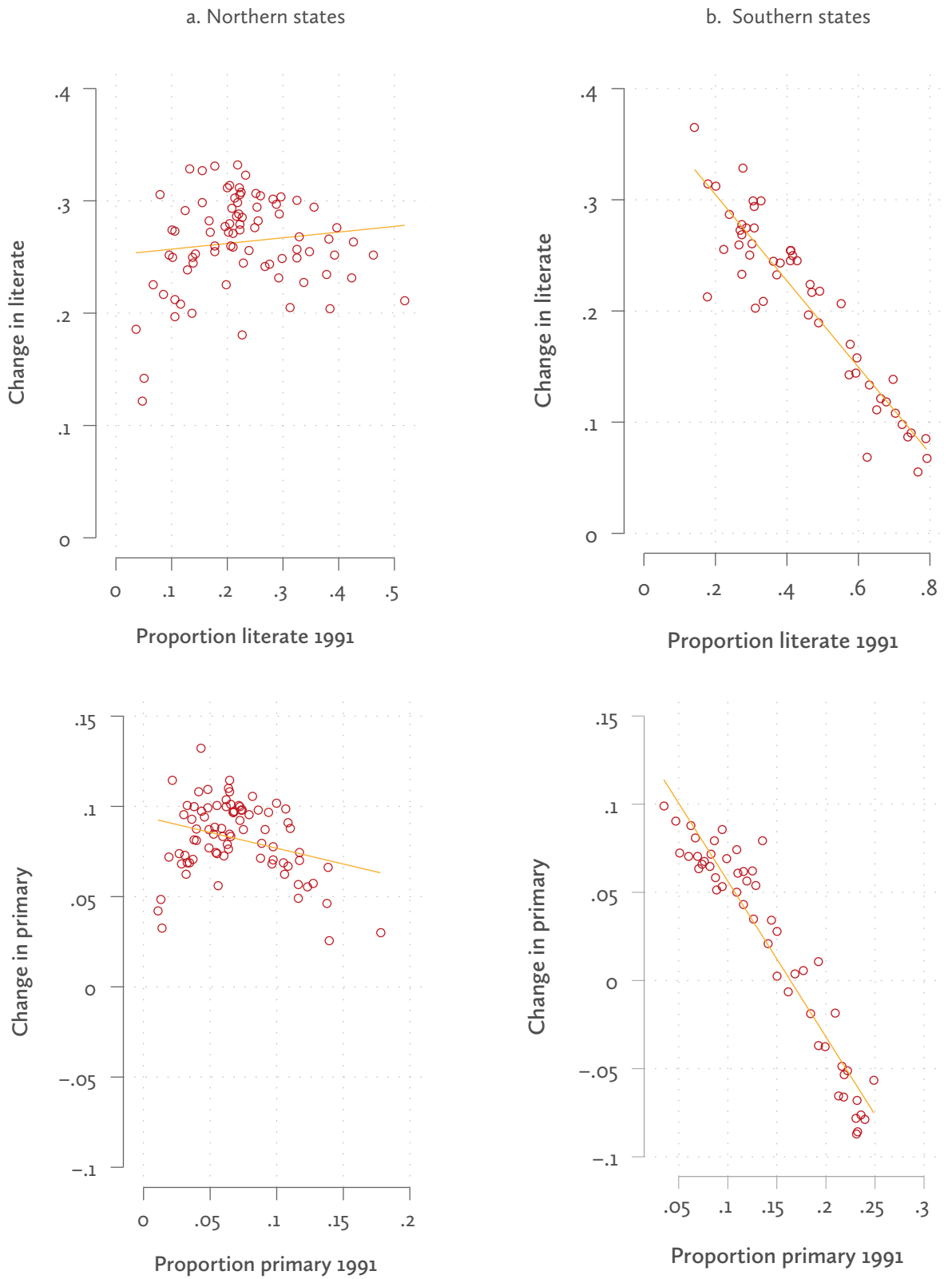
a. Northern states

b. Southern states

Figure 5.10: (Cont'd) Southern states display convergence in educational attainment for SC jatis but northern states do not

a. Northern states

b. Southern states
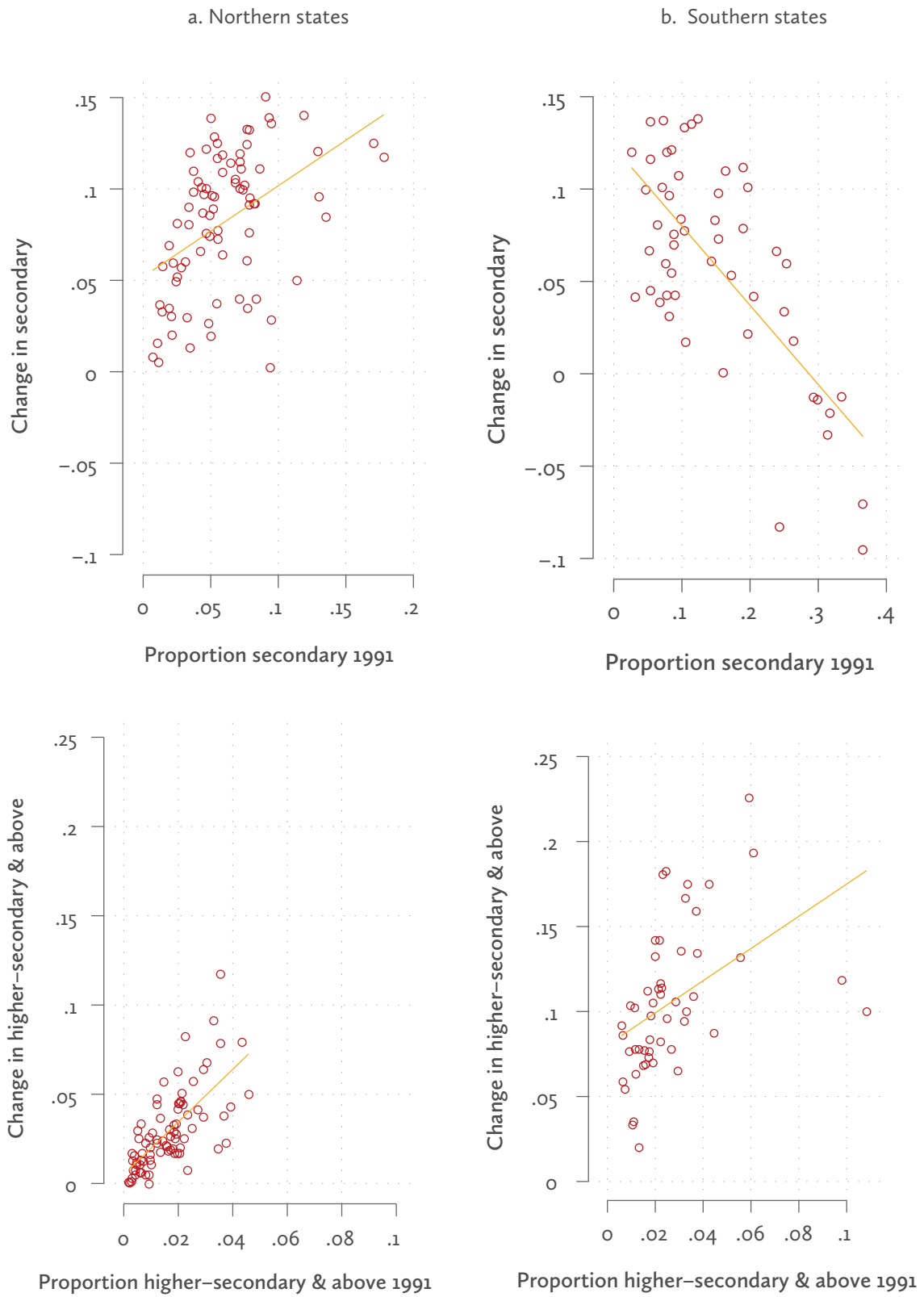
Sources and notes: Population Census 1991 and 2011

5.4 Conclusion

In this chapter, our continuing exploration of the labour market experiences of various social identities led us to look at the determinants of the type of employment, intergenerational mobility and educational convergence. While there are some long-standing inequities we find in India's labour market along the dimensions of gender, caste and religion, the recent two decades have also seen relatively big improvements, for e.g. women benefiting from the increase in regular wage work, an increase in intergenerational mobility for Scheduled Castes, and educational convergence across state for dominant as well as marginalised castes, albeit only till the secondary school level. As the Indian labour market recovers from the tremendous shock delivered by the Covid-19 pandemic, the challenge is to build on these positive developments so that the long-run persistence of identity-based divisions in employment can finally be put behind us.

Endnotes

1   This leaves out employers who run enterprises and employ hired workers. These constitute a very small fraction of the total workforce (less than 2 percent).

2   See Shrinking public employment and its costs by Pabitra Chowdhury and Rosa Abraham: https://www.hindustantimes.com/india-news/shrinking-public-sector-employment-and-its-costs-101661966849523.html

3   Due to a lack of disaggregated information on OBCs in older NSS rounds they were grouped together under the umbrella of Others.

4   Note that intergenerational mobility estimates can vary significantly depending on the age at which sons' occupational status is observed, this is mainly because of lifecycle effects. For instance, a 60 year and 18 year old father-son pair would not be entirely comparable with a 60 year and a 35 year old father-son pair.

5   The section draws on Abraham (2023). See the paper for more details.

6   If we were to choose narrower age bands, we might have more cohorts to track. However this comes at the cost of smaller sample sizes of each cohort. In the interest of smoother trends and representative estimates we use 6 year cohorts.

7   Prior to deciding that NSSO surveys could be used to extend the Census story beyond 2011, we verified that the story between 1991 and 2011 as seen in the NSSO survey is qualitatively the same as that seen in the Censuses.

6

Earnings disparities, identity-based segregation and entrepreneurship

6

Earnings disparities, identity-based segregation and entrepreneurship

In the preceding chapters we have seen how the social identities of caste, gender and religion continue to be salient in determining labour market outcomes, from participation in employment to the type of employment, to intergenerational mobility and pre-labour market outcomes such as educational attainment. We now turn to three more crucial aspects of the labour market where too identities play an important role. These are earnings, industrial and occupational segregation and patterns of entrepreneurship; that is, who earns how much, who is engaged in what kind of occupation, and who gets to be an employer and who does not.

6.1 Earnings disparities across social groups

The previous chapter focused on the type of employment in terms of self, casual or regular wage work. We saw that social identity plays an important role in determining the type of employment a person is likely to be engaged in. We now analyse disparities in earnings across employment types and social identities. On average, regular wage or salaried work tends to be the most remunerative, followed by self-employment and then casual wage work. Further, within each type, there exists significant heterogeneity in occupation, industry and other job characteristics that play an important role in determining labour earnings. A large number of studies have addressed earnings disparities by social identity in the Indian context. We do not review this literature here except to bring up specific studies

in the context of our findings.[1] Our aim is to bring estimates up to date with the most recent available data and to present a consistent series of estimates from the 1980s onwards.

In this section, we start by presenting data on the long-run evolution of earnings disparities across gender, caste and religion. Via an OLS regression approach as well as with the help of Oaxaca-Blinder decompositions, we analyse the gender, caste and religion earnings gaps in terms of the explained and unexplained components. We find that gender gaps tend to be the largest followed by caste and then religion. We also find that industrial and occupational segregation play an important role in creating earnings disparities across social identities. In the following section we present data on the extent of segregation prevalent in the economy and how it has changed over the past four decades.

6.1.1 Raw earnings gap across caste, gender and religion

We begin by examining earnings gaps for the most recent PLFS round, 2021-22. By "earnings gap" we mean the ratio of earnings for two different groups, e.g. female/male, SC-ST/Others or Muslim/Hindu. Table 6.1 shows the average earnings by gender, caste and religion for the three kinds of employment. Gender gaps are the largest among the three and gaps along religious lines are the narrowest. The starkest difference is seen in self-employment with women reporting earnings only 40 percent that of men.

Table 6.1: Labour earning disparities across social identities

Identity	Regular wage (₹ per month)	Self-employed (₹ per month)	Casual wage (₹ per day)
Gender			
Women	13,666	4,809	230
Men	17,910	12,099	358
Ratio	0.76	0.40	0.64
Caste			
SC/ST	13,735	8,271	309
Others	18005	11,539	344
Ratio	0.76	0.72	0.90
Religion			
Muslims	13,550	10,395	361
Hindus	17,197	10,663	323
Ratio	0.79	0.97	1.12

Sources and notes -PLFS 2021-22.

Several questions arise here. How far are these disparities explained by differences in characteristics of workers? For example, as we saw in the last chapter, educational attainment varies greatly by caste. Could this be the entire explanation for the caste earnings gap? How have the gaps changed over time as the economy has grown and new opportunities have been created? We now turn to these issues.

The starkest difference is seen in self-employment with women reporting earnings only 40 percent that of men.

Figure 6.1 breaks down the gender earnings gap for the three kinds of employment by decile of the earnings distribution. This figure answers the question, are women at the bottom of the earnings distribution (the lowest earners) equally disadvantaged as compared to women at the top

(the highest earners). A few points are worth noting. First, for salaried or regular wage workers the gap reduces as we move up the distribution. Duraisamy and Duraisamy (2016) as well as Deshpande, Goel and Khanna (2018) report this "sticky floor" phenomenon for gender gaps. Smaller gaps at the upper end may reflect the better bargaining position that higher educated women have or lesser incidence of discrimination in more regulated and formal jobs. Conversely, larger gaps at the bottom may point to the existence of more severe gender discrimination and weak bargaining position of women for relatively more informal regular wage jobs. It may also be due to a more segregated division of labour in lower-end jobs, with clearly demarcated "men's work" and "women's work" as is often found in factories.

In the case of self-employment too, a similar fall in disparity is observed across the earnings distribution, though of course, as we saw in Table 6.1, in general

Figure 6.1: The gender earnings gap decreases with higher earnings
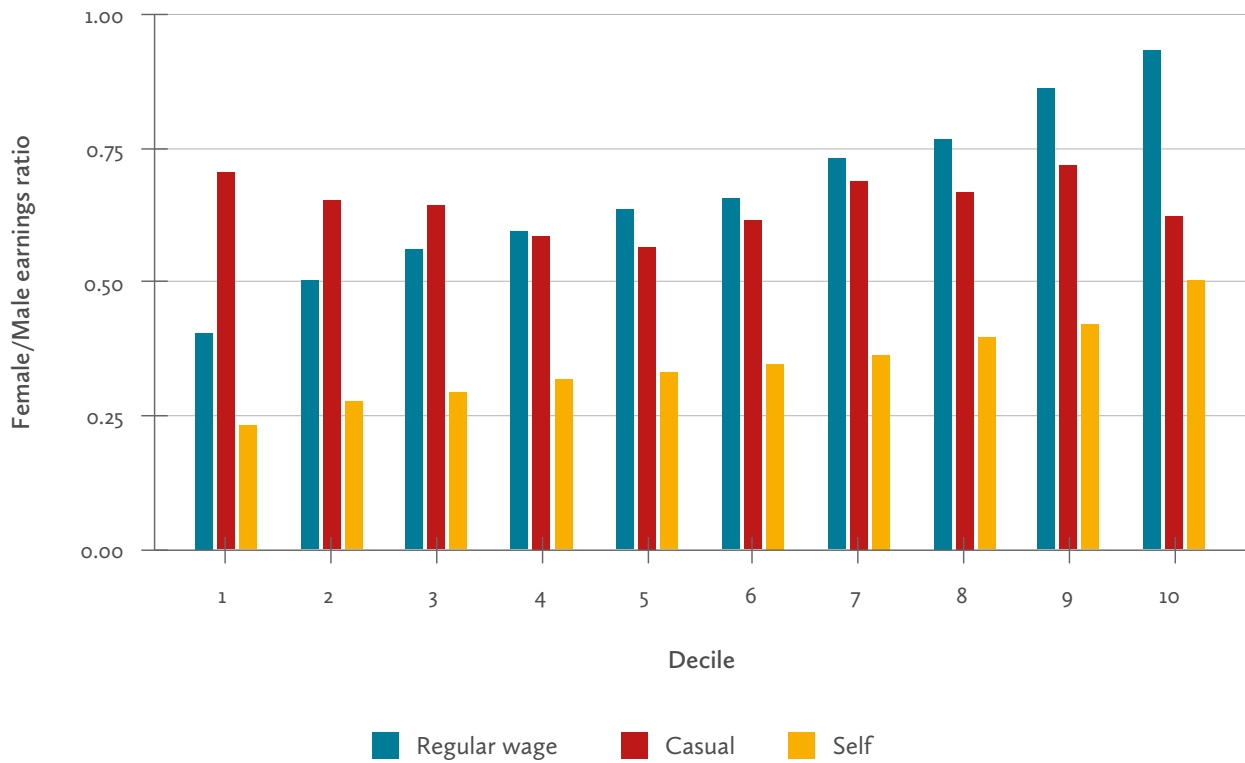
Sources and notes: PLFS 2021-22

the disparity is much higher in self-employment than in regular wage work. Even in the top earnings decile, self-employed women earn only half as much as self-employed men. We analyse the possible reasons for higher self-employment disparities later in the chapter. For now, we note that, as with regular wage work, it is possible that self-employed women near the top of the earnings distribution have better education as well as higher social and/or physical capital. It is also possible that there is more segregation at the bottom of the distribution forcing women into lower paid work. A classic example of the latter is the large overrepresentation of women in bidi rolling and other home-based work, an extremely low-earnings form of self-employment.

Next, focusing just on regular wage work, we examine decile-wise earnings gaps by caste -SC/ST versus Others (including OBCs), gender (women by men) and religion (Muslim by Hindu). Figure 6.2 shows that religion and gender gaps move in

opposite directions. As we saw earlier, women at the upper end of the earnings distribution are less disadvantaged than men but Muslim workers at the upper end of the distribution are more disadvantaged vis-a-vis Hindu workers.

At the top decile Muslim workers earn 75 percent of what Hindu workers earn. For the 3rd and 4th decile, this number is much higher at around 94 percent indicating that Muslim salaried workers earn almost as much as Hindu salaried workers at the lower end of the earnings distribution.

Women at the upper end of the distribution are less disadvantaged than men but Muslim workers at the upper end are more disadvantaged vis-a-vis Hindu workers.

Conversely, women at the bottom of the distribution only earn 40 percent of what men earn but at the

Figure 6.2: The gender gap falls while the religion gap widens over the earnings distribution
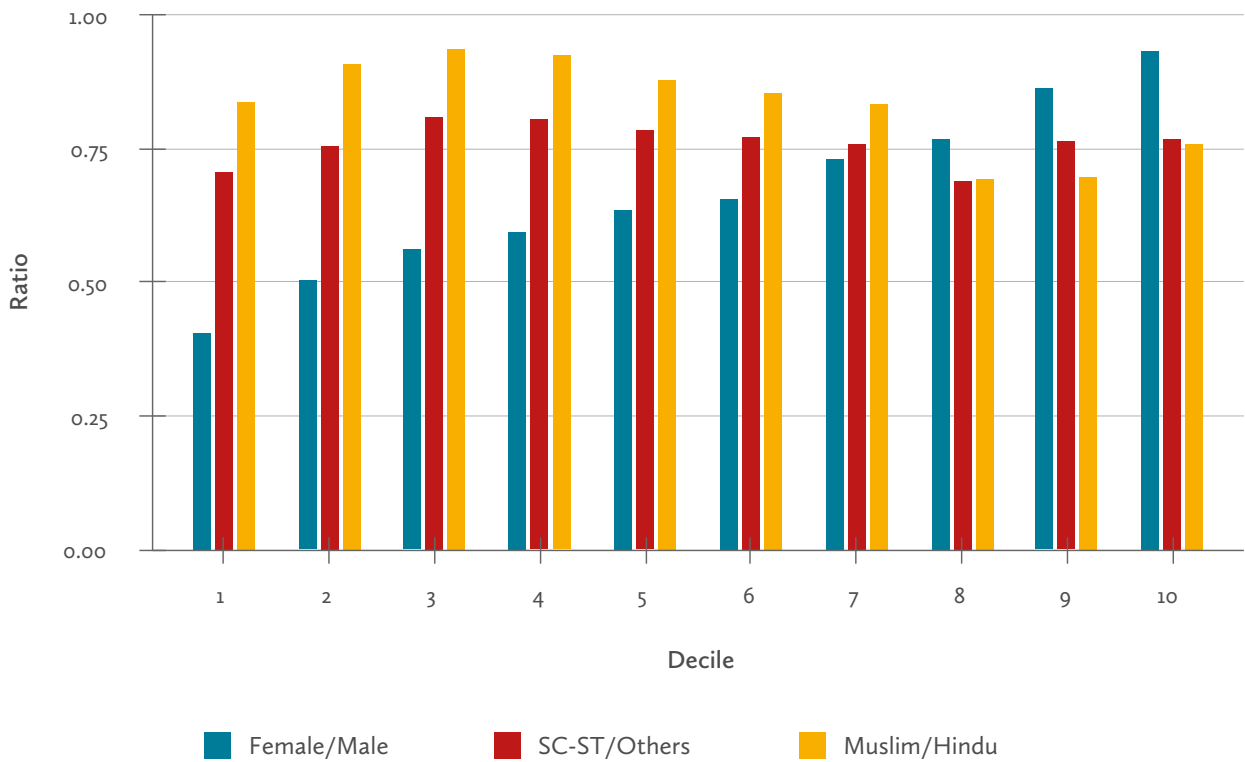
Sources and notes: PLFS 2021-22

Figure 6.3: Comparing caste, gender and religion–based earnings gaps for regular wage workers
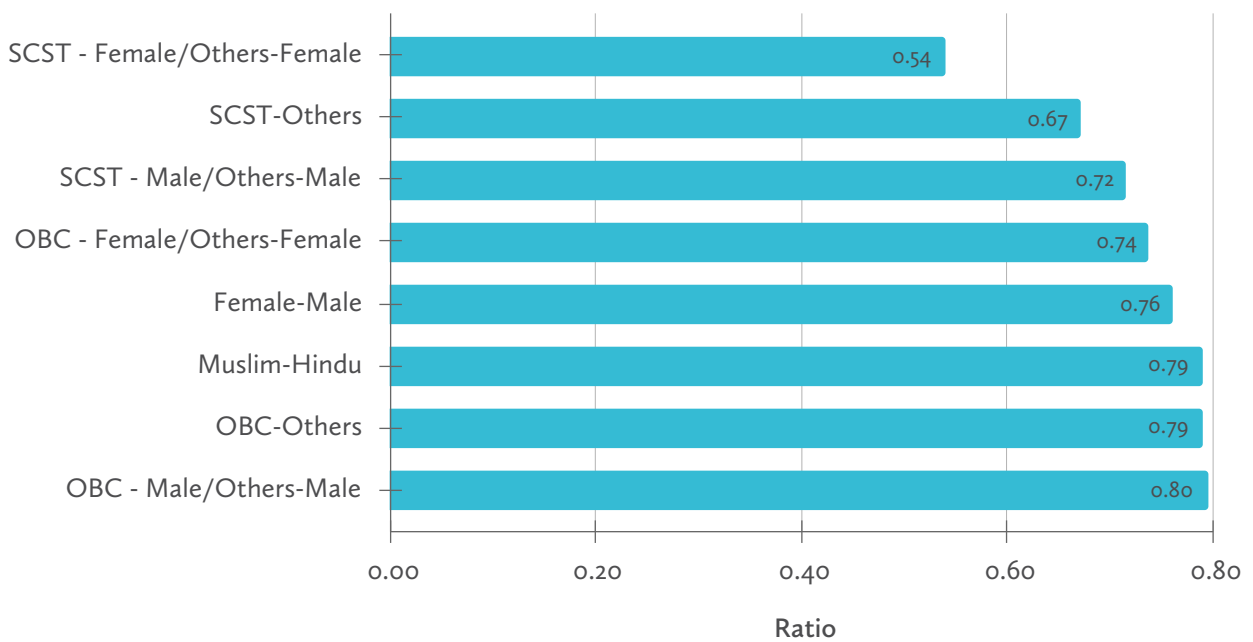
Sources and notes: PLFS 2021-22

top decile this number is 93 percent. An increasing Hindu-Muslim gap for higher paid (and possibly more formal jobs) is a matter of concern and needs further investigation. Finally, the gap in earnings by caste does not vary a great deal across the distribution (70 percent and 80 percent).

The decile-wise analysis shown above does not allow us to separate OBCs from Others for sample size reasons. We separate OBCs from other non-SC/ST castes and calculate the overall regular wage gap (Figure 6.3). The earnings ratio is 0.67 for SC/ST to other workers (excluding OBCs). The gap is narrower between OBC workers and Others (0.79). Thus, as expected, SC and ST regular wage workers are more disadvantaged compared to OBCs. Incidentally, the religion gap (Muslim/Hindu) at 0.79 is identical to the OBC-Others gap. Finally, Figure 6.3 also shows some important intersectional effects. When we calculate the caste earnings gaps within each gender, we see that the largest disadvantage is

observed for women SC/ST workers who earn only 54 percent of what women from general category households earn in regular wage work.

The analysis thus far was limited to the most recent PLFS data (2021-22). We now take a longer view and present data on the gender earnings gap for regular wage workers from the 1980s. Figure 6.4 shows the data for each quartile. Quartiles or 25 percent groups are shown instead of deciles due to sample size limitations in earlier rounds.

The largest disadvantage is observed for women SC/ST workers who earn only 54 percent of what women from general category households earn in regular wage work.

Two things are clear. First, as we have already seen, women at the upper parts of the earnings distribution suffer a smaller penalty as compared

Figure 6.4: The gender earnings gap has decreased over time among regular wage workers
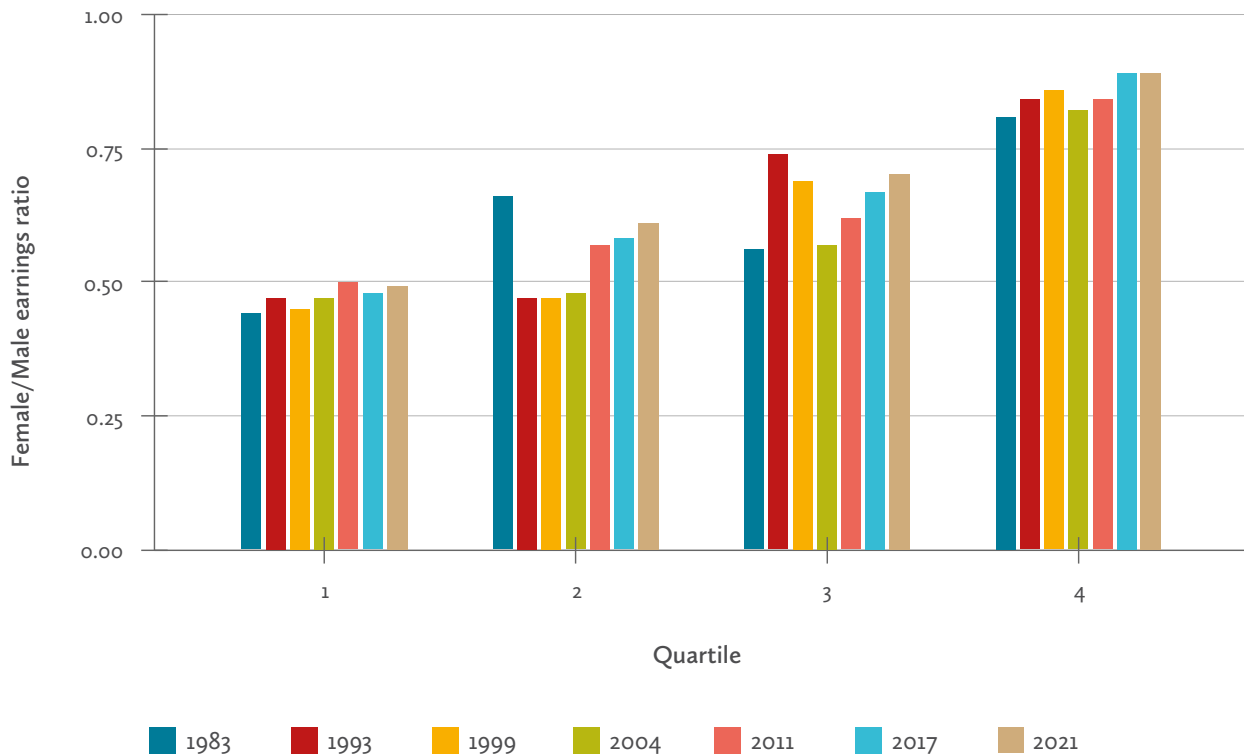
Sources and notes: PLFS 2021-22

to women at the lower end. In the bottom quartile, over the entire period, women only earned half as much as men in regular wage work. In the top quartile this ratio was 80 percent.

Second, starting in 2004, the gap narrowed in each quartile. This is particularly clear for the 2nd, 3rd, and 4th quartiles. This could be either because the extent of gender discrimination narrowed over time or because increased educational attainment as well as an increase in regular wage opportunities allowed women access to better paid work. The analysis thus far cannot answer this question because we do not control for any differences in either individual or household characteristics or the nature of the industry or occupation in which individuals work. It is well known that at least part of these gaps are explained by such observable differences. But part of the gap remains unexplained. In the labour economics literature the unexplained portion of the gap has generally been understood as a measure of identity-based discrimination. In the next section we present estimates of gender and caste unexplained gaps over the past four decades.

6.1.2 Measuring unexplained earnings gaps – some technical issues

Before we proceed to discuss how much of the raw earnings gap is a result of differences in observed worker or job characteristics versus unobserved factors, we flag two issues of importance for labour economists. First as we saw in Chapter Four, there is a well understood set of factors that determines whether a woman is employed or not. In India, the process of structural transformation has resulted in a large change in the composition of the female workforce. Older, less educated women have exited while younger more educated women have entered. Since earnings are only observed for women who are currently employed, it follows that the evolution of the earnings gap over time will be a result of any changes in discrimination faced by women at work

as well as the changing characteristics of women who happen to be employed. Thus the observed gender earnings gap could fall either because women face less discrimination over time, or because of the entry of relatively more educated and therefore, likely higher wage earning women into the labour market (Blau, Koebe, and Meyerhofer 2021).

Economists have tried to deal with this problem by employing a two-step process. In the first step selection into the workforce is modelled using a probit or linear probability regression model. Then a selection-corrected earnings equation is estimated. But this method relies on finding a variable that can affect only employment without having an impact on wages. Variables such as number of children a woman has, are sometimes used for this purpose. But many authors have noted that it is hard to justify that such variables have no impact on earnings. Hence several recent studies present un-corrected results and acknowledge that selection bias will play a role in the estimated gap. The analysis here does the same.[2]

The second general point we make before presenting the results pertains to the method of measuring unexplained differences. The standard technique of decomposing the earnings gap for any two groups (say men and women or black and white) into its explained and unexplained components is the Blinder-Oaxaca decomposition. These are also called "endowments" and "returns" components (see Methods Appendix for more details). This is an accounting technique (with no causal interpretation) that tells us what proportion of the earnings gap is explained by differences in levels of education, marital status or other observable characteristics. The second part (returns) is usually interpreted in the labour economics literature as measuring discrimination but it can also include the effect of omitted observable characteristics (which the researcher has not been able to include due to lack of data).

However, there exists a simpler approach to estimate the unexplained portion of the earnings gap. This involves estimating a pooled OLS wage regression with a dummy variable for the group identity (say women) along with other observable characteristics. The coefficient on the dummy variable can be interpreted as the unexplained component of the earnings gap. The Blinder-Oaxaca approach is more popular because it also allows us to estimate the contribution of each observable characteristic to the explained and unexplained parts of the gap. However, the decomposition is sensitive to our assumption of what constitutes a non-discriminatory earnings structure. In other words, do we believe that men are paid the market wage while women face negative discrimination or that men face positive discrimination (favouritism) while women are paid the market rate, or are there perhaps both kinds of discrimination prevalent? Different counterfactuals will yield different results. To get around this problem authors often present results from different counterfactuals (e.g. see Deshpande and Sharma 2016). The OLS approach does not allow us to get separate estimates for how each variable contributes to the explained and unexplained parts of the gap but its advantage is that we do not have to make any assumptions such as the ones mentioned above. Elder et al.(2010) show that the OLS approach is as good as the decomposition approach.[3]

6.1.3 Decomposing the earnings gaps

We start by measuring the unexplained part of the gender, caste and religion gaps for regular wage, casual wage and self-employed workers. We have done this using both methods discussed above - a simple OLS with a variable that captures the social identity as well as a standard Oaxaca-Blinder decomposition. We first show the OLS results. The OLS model is log earnings regressed on a set of individual and household characteristics such as age, education, marital status, gender, caste, religion, education level of the household head as well as industry and occupation. The full model is provided in the Methods Appendix and the complete regression results are in the Results Appendix.

We start by presenting results for the latest PLFS round, 2021-22 and then move to analysing trends over time. Table 6.2 shows the coefficients observed on the gender, caste and religion variables for the three types of employment. Note that, for ease of interpretation, each identity is captured by a binary variable (male-female, SC/ST - non SC/ST and Hindu-Muslim). Multiplying each number by 100 tells us how much percent less (due to the negative sign) women earn than men or SC/ST workers earn compared to the rest and so on. Also shown are results with and without the inclusion of a set of

Table 6.2: The unexplained component of earnings gap is largest for women workers

2021-22	Regular wage		Casual		Self-employed	
	Without industry or occupation controls	With industry and occupation controls	Without industry or occupation controls	With industry and occupation controls	Without industry or occupation controls	With industry and occupation controls
Female	-0.455[***]	-0.346[***]	-0.433[***]	-0.294[***]	-0.948[***]	-0.789[***]
SC/ST	-0.0503[***]	-0.0432[***]	-0.007	-0.001	-0.171[***]	-0.169[***]
Muslim	-0.0435[**]	-0.0567[***]	0.0582[**]	0.0420[**]	-0.013	-0.017

Sources and notes: PLFS 2021-22. Shown are the regression coefficients from an OLS regression with indicator variables for each social identity. See Methods Appendix for details. The full regression results are provided in the Results Appendix. *** p<0.01, ** p<0.05, * p<0.1

variables to capture the industry of work (NIC 2 digit) and the occupation (NCO 2 digit).

Several points are worth noting. First, the unexplained portion of the earnings gap is much larger for women than for SC/ST workers or for Muslim workers in any type of work. In fact, for caste in casual wage employment and religion in self-employment the coefficient is insignificant. This means that the raw earnings gaps we saw in the previous section are fully explained by differences in education and other observable characteristics. For women, on the other hand, a large gap (34.6 percent) remains unexplained. Earlier studies have also pointed to this important fact - unexplained earnings gaps tend to be much larger for gender than for caste (Chakraborty and Bohara 2021; Agrawal 2014).

Unexplained earnings gaps are much larger for gender than for caste.

Second, as in the case of the raw gaps, for women and for SC/ST workers, the size of the penalty is much larger for self-employment than for wage work (whether regular or casual). One way to understand this is that wage employment exposes workers to possible discrimination only in the labour market while self-employed workers can face discrimination in factor markets (such as land, credit, or raw materials) as well as product markets and additional burdens due to harassment by officials etc, thereby increasing the possible size of the total penalty.

The last point to note is that the inclusion of variables that capture industry and occupation makes a large difference for the gender gap but not the caste or religion gap. This suggests that a significant part of the gender penalty comes from segregation of women into worse paying or less productive work. We take up the issue of industrial and occupational segregation in the next section.

Here one may ask if it is acceptable to add information on industry and occupation to such a model, or should it be limited to individual and household characteristics only? Banerjee and Knight (1985) point out that the earnings function is a reduced form equation that captures the outcome of an interaction of the forces of demand and supply. Thus, no theoretical justification exists for restricting the independent variables to the conventional supply-side ones such as the worker's experience and education. Rather variables measuring the nature of labour demand can be included. Conceptually what this means is that labour market discrimination can operate via channelling women into certain kinds of work, which limits their ability to earn as much as men. The unexplained gap then captures the remaining earnings penalty within each industry-occupation combination.

Have the earnings penalties suffered by women, marginalised caste workers or Muslims, after accounting for differences in their demographic and employment characteristics, reduced over time? We address this question using data from the NSSO-EUS rounds (1983 to 2011) and the more recent PLFS rounds. Both approaches discussed earlier, the OLS model as well as the Blinder-Oaxaca decomposition were used. Table 6.3 shows the coefficient on the gender dummy in the OLS regression. As before, this number multiplied by 100 gives us the percent difference in male and female earnings that is not explained by various individual and household characteristics or by industrial and occupational segregation. Self-employed earnings are only available starting with the PLFS rounds (2017-18).

Over time the gender penalty for regular wage workers has risen from 24 percent to 34.6 percent.

We see that over time the gender penalty for regular wage workers has risen from 24 percent to 34.6 percent while it shows no clear trend for casual

Table 6.3: The unexplained gender earnings gap increased over time for regular wage workers

Year	Regular wage	Casual wage	Self-employed
1983	-0.24	-0.33	
1993	-0.28	-0.31	
1999	-0.31	-0.30	
2004	-0.33	-0.32	
2011	-0.33	-0.29	
2017	-0.37	-0.34	-0.74
2021	-0.35	-0.29	-0.79

Sources and notes: NSSO EUS and PLFS various rounds. Shown are the regression coefficients from an OLS regression with indicator variables for each social identity. See Methods Appendix for details. The full regression results are provided in the Results Appendix. All coefficients are significant at the p<0.01 level

wage workers. Note that these results are counter to what we saw earlier in Figure 6.4 where there was a convergence in male-female earnings over time.

How do we reconcile these two findings? Recall that the raw gender earnings gap reported earlier did not take into account the changing female workforce over time. Therefore, the convergence in earnings could be an outcome of a higher educated female workforce working in better paid jobs. The gender penalty estimated from the regression accounts for individual characteristics and changes in the

Figure 6.5: The unexplained gender gap for regular wage workers rose in rural areas but not in urban areas
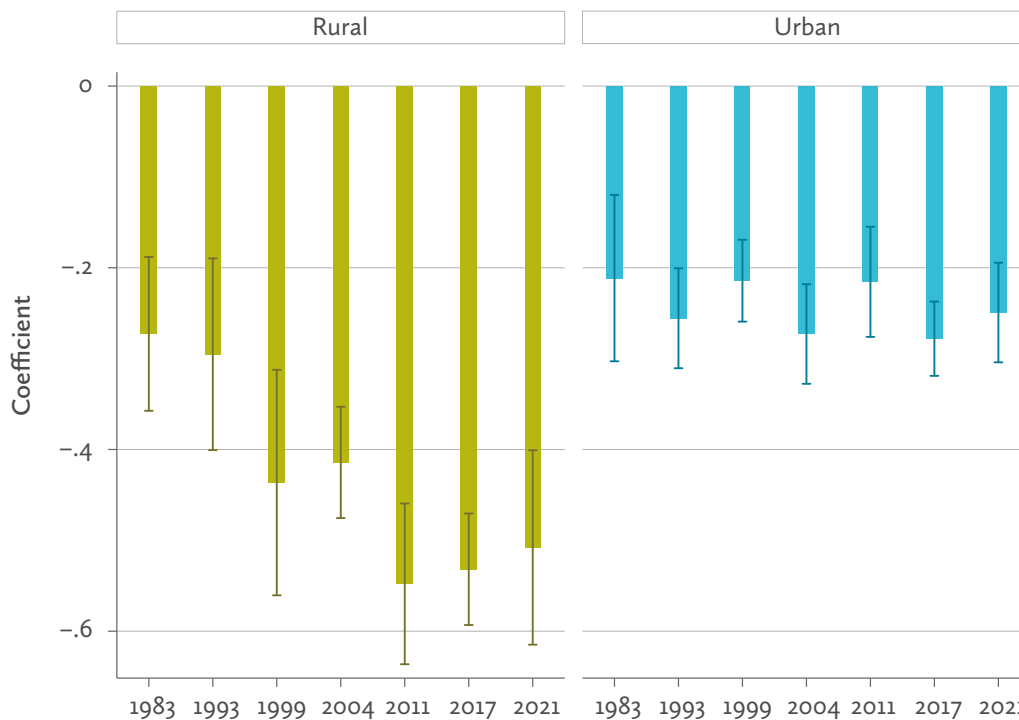
Sources and notes: NSSO EUS and PLFS various rounds. Shown are the regression coefficients from an OLS regression with indicator variables for each social identity. See Methods Appendix for details. The full regression results are provided in the Results Appendix.

Table 6.4: Results of the Blinder-Oaxaca decomposition for regular salaried workers

a. 80% of the gender earnings gap remains unexplained while only 15% is unexplained for SC/ST workers after controlling for observable characteristics

By SC/ST

Year	Actual			Percent	
	Difference	Explained	Unexplained	Explained	Unexplained
1983	0.356	0.248	0.108	69.6	30.4
1993	0.305	0.270	0.034	88.7	11.3
1999	0.225	0.217	0.008	96.6	3.4
2004	0.297	0.253	0.044	85.3	14.7
2011	0.259	0.214	0.045	82.6	17.4
2017	0.243	0.200	0.043	82.3	17.7
2021	0.272	0.232	0.039	85.5	14.5

By gender

Year	Actual			Percent	
	Difference	Explained	Unexplained	Explained	Unexplained
1983	0.384	0.138	0.245	36.0	64.0
1993	0.495	0.205	0.290	41.4	58.6
1999	0.539	0.225	0.314	41.8	58.2
2004	0.568	0.241	0.328	42.3	57.7
2011	0.493	0.155	0.338	31.5	68.5
2017	0.467	0.077	0.390	16.5	83.5
2021	0.440	0.083	0.357	18.9	81.1

b. In rural areas, the unexplained component of the gender earnings gap rose steadily since 2004

Rural

Year	Actual			Percent	
	Difference	Explained	Unexplained	Explained	Unexplained
1983	0.410	0.140	0.270	33.1	66.9
1993	0.650	0.330	0.320	51.0	49.1
1999	0.710	0.270	0.440	38.3	61.7
2004	0.690	0.270	0.420	38.8	61.2
2011	0.630	0.090	0.540	14.2	85.8
2017	0.590	0.020	0.560	4.1	95.9
2021	0.560	0.030	0.530	6.2	93.8

Urban

Year	Actual			Percent	
	Difference	Explained	Unexplained	Explained	Unexplained
1983	0.370	0.150	0.220	40.7	59.3
1993	0.420	0.150	0.260	36.2	63.8
1999	0.430	0.210	0.220	49.2	50.8
2004	0.500	0.230	0.270	46.0	54.0
2011	0.420	0.200	0.220	47.2	52.8
2017	0.420	0.120	0.290	29.9	70.1
2021	0.390	0.140	0.250	36.2	63.8

Sources and notes: NSSO EUS and PLFS various rounds. See Methods Appendix for details.

nature of the female workforce in terms of their age, education and family incomes, as well as the industry and occupation of employment. If we see that the gender penalty has persisted and actually increased, this shows that even with comparable education levels, women continue to earn far less than men with the disparity increasing over time. Intriguingly, upon disaggregating by rural-urban regions, we find that the increase in the gender penalty in regular wage work comes entirely from rural areas. In urban areas, no trend is seen (Figure 6.5). Finally, note that, as seen earlier in Table 6.2, the unexplained gender

Figure 6.6: The unexplained earnings gap for SC/ST workers fell over time in rural and urban areas
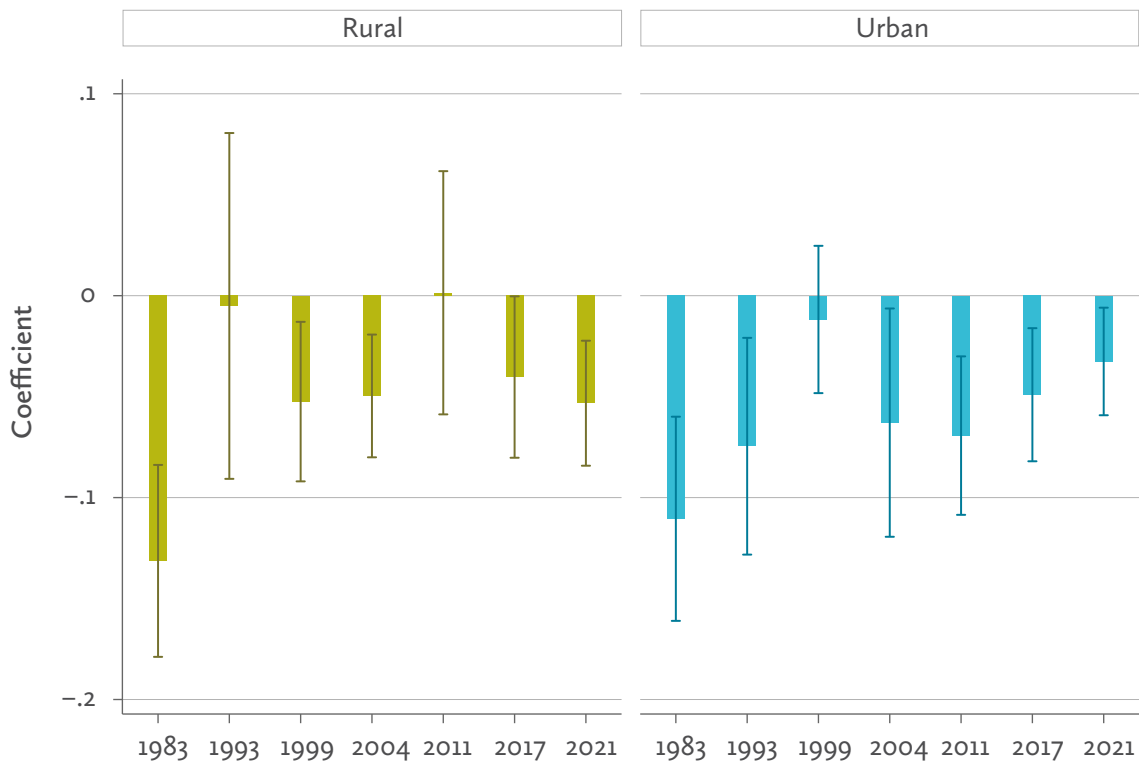
Sources and notes: NSSO EUS and PLFS various rounds. Shown are the regression coefficients from an OLS regression with indicator variables for each social identity. See Methods Appendix for details. The full regression results are provided in the Results Appendix.

Figure 6.7: A significant earnings gap for Muslim workers is observed since 2017 and only in urban areas
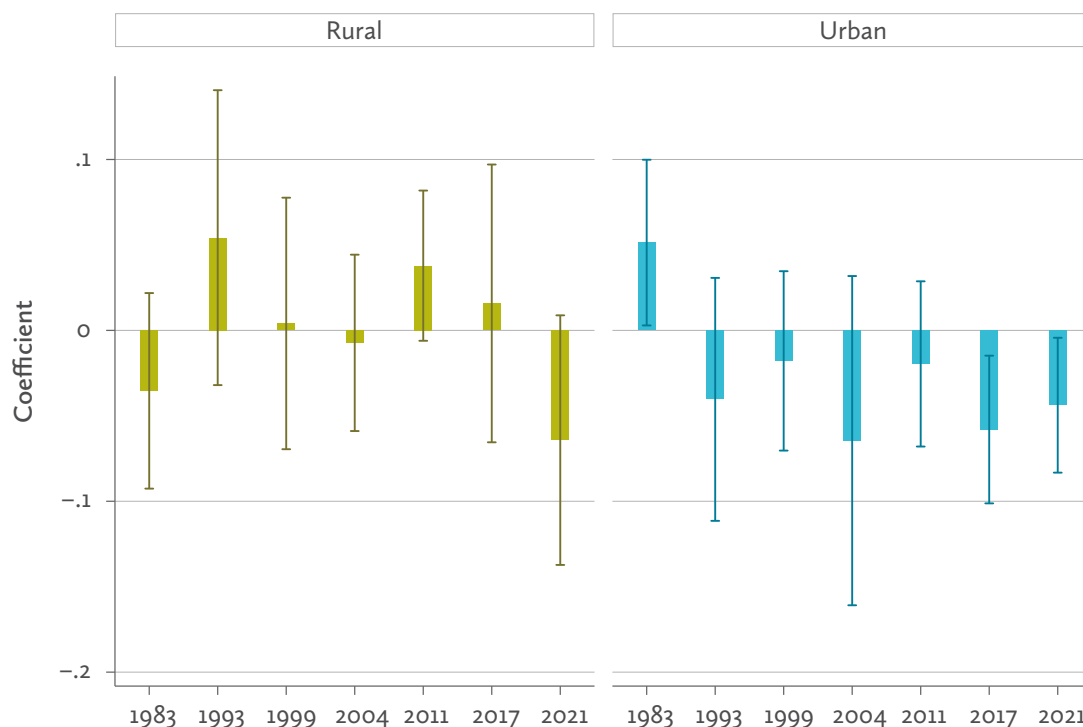
Sources and notes: NSSO EUS and PLFS various rounds. Shown are the regression coefficients from an OLS regression with indicator variables for each social identity. See Methods Appendix for details. The full regression results are provided in the Results Appendix.

gap for self-employed workers is far larger.

Coming to the evolution of the caste and religion penalties in regular wage work, Figures 6.6 and 6.7 show, first, that the size of these penalties has never been as large as the gender penalty (note the difference in scale of the Y axis with respect to Figure 6.5). Second, the caste penalty has declined over time. And third, there is a penalty for Muslim regular wage workers in urban areas in the last two time points (2017 and 2021). For the other years the coefficient is not significantly different from zero.

To summarise the findings so far, we see that raw gender gaps are far larger than raw caste or religion gaps. Furthermore, even after controlling for a range of factors, gender penalties remain larger than the penalties faced by other disadvantaged groups and have increased over time. In the case of religion and caste, most of the raw gap is explained by differences in education, occupation and industry.

Before ending this section, we reinforce the main conclusions by showing results from a Blinder-Oaxaca decomposition on earnings for regular wage workers (Table 6.4). This approach allows us to show the division between explained and unexplained parts of the earnings gap directly. The gaps as well as the proportion explained and unexplained are given for each year, for caste as well as gender (a) and only for gender by rural and urban separately (b).[4] We can see that, with the exception of 1983, the unexplained part of the caste gap only goes as high as 17.7 percent, while that for gender never falls below 57 percent.

Further, both the overall gender gap as well as its unexplained part have risen steadily since 2004, with the latter reaching more than 80 percent in recent years. Table 6.4b shows that almost all this increase comes from the rural areas, as we mentioned earlier. More than 90 percent of the gender wage gap in rural areas in 2021-22 was not explained by the

workers' characteristics. This conclusion is worth emphasising. The same period that saw a rapid rise in regular wage work among women (2004 onwards) also saw an increase in the wage gap as well as the unexplained part of the gender wage gap.

We should note that while interpreting the unexplained portion of the earnings gap as resulting from discrimination a few caveats should be kept in mind. First, recall that in the case of gender, the selection bias can play an important role. This is less relevant in the case of caste or other kinds of identities where selection into employment does not vary much by identity.

The same period that saw a rapid rise in regular wage work among women also saw an increase in the unexplained part of the gender wage gap.

Second, the explained component also embodies the effects of past discrimination (such as educational opportunities). Based on the second point, Deshpande and Sharma (2016) caution that estimates of the unexplained component are not precise measurements of "true" discrimination but only give a sense of the magnitude and trend.

6.2 Industrial and occupational segregation

We saw in the previous section that accounting for industry of work as well as occupation of the worker can make a significant difference in explaining the disparities in earnings across gender, caste and religion. Our earlier report, State of Working India 2018 (Chapter Five) presented some data on this theme. In this section we expand on our earlier work and examine the changing nature of industrial and occupational segregation in India over the long-run. This analysis is based on data from the NSSO EUS

and PLFS for the years 1983, 1993, 2004, 2011, 2017 and 2021.

Several studies have shown that the disadvantaged groups in India are concentrated in low-skilled low-wage industries or occupations (Agrawal 2016; Kaufman 2010; Blau, Brinton, and Grusky 2009; Anker 1998). There is also evidence that in Indian manufacturing, gender pay gaps increased within blue collar workers between 2004-05 and 2011-12 (Mondal et al. 2018). There is also rigidity in intergenerational occupational mobility, particularly among marginalised castes and in rural areas (Motiram and Singh 2012).

It is important to note that segregation in India operates both vertically (dominant and marginalised groups clustered in two ends of the occupational ladder) and horizontally (dominant and marginalised groups have different tasks and differential wages within an occupation category). For instance, in 2011-12, SC women earned 25.2 percent less than the per capita average wage of the economy due to occupational segregation and an additional 19.5 percent less owing to wage differentials within occupations. Similar trends were also observed for Muslim women (Rammohan, Goli, and Reddy 2017).

Industrial Segregation
In this section, we first calculate the widely used Duncan Index as an overall measure of industrial segregation. The index takes values between 0 and 1 where 0 represents no segregation and 1 represents complete segregation. The value of the index indicates the proportion of the workforce that needs to be reallocated between industries to achieve a state of no segregation (e.g. 0.3 indicates that 30 percent of the workforce needs to be reallocated). We also calculate a representation index for gender, caste and religious identities. This is a ratio obtained by dividing the share of a particular group found in a particular industry by the share of that group in the workforce. We do this for every broad industry groups concorded using 2-digit level of aggregation

Table 6.5: Industrial segragation is highest along gender lines, followed by caste and religion

Duncan Index

Year	Male-Female	SCST-Others	SC-Others	ST-Others	OBC-Others	Hindu-Muslim
1983	0.34	0.23				0.21
1993	0.35	0.23				0.22
2004	0.40	0.25	0.29	0.35	0.19	0.21
2011	0.39	0.25	0.32	0.35	0.17	0.20
2017	0.42	0.23	0.28	0.31	0.14	0.18
2018	0.41	0.24	0.29	0.33	0.15	0.18
2021	0.41	0.23	0.28	0.32	0.15	0.17

Sources and notes: NSSO EUS and PLFS various years. Numbers are the Duncan Segregation Index. See Methods Appendix for details. For column 2 (SC/ST-Others), "Others" include OBCs

of the National Industrial Classification (NIC) in the non-agriculture sector (see Methods Appendix for details). A representation index of 1 indicates representation in proportion to workforce share, less than 1 indicates under-representation and more than 1 indicates over-representation.

Among the different social identities examined in this analysis, the highest level of industrial segregation in the Indian workforce was observed between men and women. The segregation index between men and women in 1983 was 0.34 which increased to reach a value of 0.41 in 2021 (Table 6.5). By contrast, caste-based industrial segregation between SC/ST and Others (including OBCs) remained almost the same, declining slightly between 2004 and 2021. Industrial segregation between Hindus and Muslim was even lower at around 0.17 to 0.21. This means that 40 percent of the workforce would need to be redistributed along gender lines across industries to achieve gender parity.

The number is 23 percent for industrial segregation based on SC/ST-Others and 17 percent for Hindu-Muslim industrial segregation. Since 1999, the data also allows us to separate OBC workers from Others. We see that segregation is lowest for OBC-Others, followed by SC-Others and ST-Others.

40% of the workforce needs to be redistributed along gender lines across industries to achieve gender parity. The number is 30% for caste and 20% for Hindu-Muslim industrial segregation.

Most of the increase in gender-based segregation across industries occurred in the early post-reform period between 1993 (0.35) and 2004 (0.4). But there are several industries in which women are consistently over-represented over the entire four decade period. For example, in tobacco products manufacturing, domestic work (maids etc), health and social work, education, and manufacture of textiles not only are women consistently over-represented, but their over-representation has also increased over time (Figure 6.8a). In 2017 the share of women in the tobacco industry reached a high of more than 5 times their share in the workforce, before declining slightly by 2021. However, as a positive change, women's over-representation in the waste management and sewerage industry has reduced from an index value of 2.4 in 1983 to 1.2 in 2021. As we will see later, this may be a result of a

Figure 6.8: Industries that show consistent over–representation for disadvantaged identities

a. Women

b. Scheduled Castes

c. Scheduled Tribes

d. Muslims

Legend: 1983 · 1993 · 2004 · 2011 · 2017 · 2021
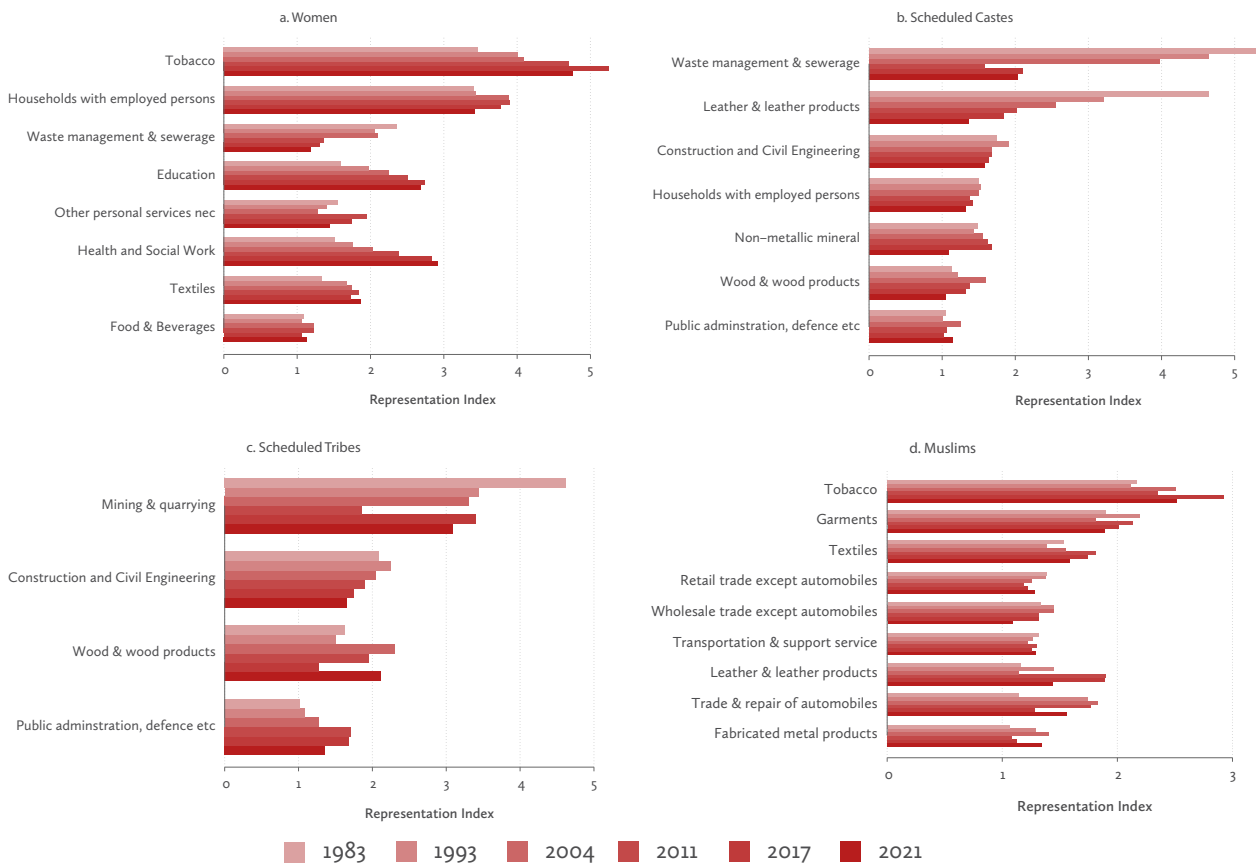
Sources and notes: NSSO EUS and PLFS various rounds. Shown are Representation Indices. See text and Methods Appendix for details.

desegregation along caste lines.

Caste-based industrial segregation between SC/ST and Others (including OBCs) remained constant at a Duncan index value of 0.23 in 1983 and 2021, with marginal changes in the years in between. Later EUS rounds allow us to separate out OBCs from the Others category (Table 6.5). As expected, the level of segregation between Schedule Castes and Tribes versus Others is much higher when we remove OBCs (Duncan of 0.28 in 2021). And correspondingly, the segregation index between OBCs and general category workers was much lower to start with (0.19 in 2004) and declined further (0.15) over the years.

Scheduled Castes have remained consistently over-represented in waste-management and sewerage and manufacture of leather and leather products - the two industries that have strong historical ties with their caste identities (Figure 6.8b). But

encouragingly, in both these industries, the extent of over-representation has declined substantially. In the case of the leather industry, the representation index declined from 4.6 in 1983 to 1.4 in 2021. In case of waste management and sewerage, over-representation of SCs decreased from over 5 times their workforce share in 1983 to 1.6 times in 2011, with the largest drop occurring between 2004 and 2011 which was also the high growth period. However, in 2017 it increased to 2.1 and has remained above twice their workforce share in 2021.

For STs, mining and quarrying has been a consistently over-represented industry, but here too the extent of segregation declined until 2011 with the largest decline recorded between 2004 and 2011 (Figure 6.8c). But between 2011 and 2018 this trend was somewhat reversed and in 2021, the value of the representation index was above 3.

Finally, industrial segregation between Muslims and Hindus is relatively less-pronounced than gender- and caste-based segregations and it has also declined over time. The segregation index declined from a peak of 0.22 in 1993 to 0.17 in 2021. Among industries in which Muslims are consistently over-represented since 1983, their over-representation is highest in manufacture of tobacco, garments, and textiles in 2021 (Figure 6.8d).

The flip-side of the over-representation of disadvantaged social groups (women, marginalised castes, adivasis, Muslims) in labour intensive industries, traditional services and traditional caste-based occupations is obviously that they are under-represented in capital intensive manufacturing as well as modern services, both of which are more desirable jobs from an earnings as well as social respect perspective. The one exception to this is the over-representation of women in health and education.

The foregoing analysis demonstrates the following. Industrial segregation in India remains high. Overall it has increased for gender and declined slightly for caste and religion. In case of some traditionally caste-segregated industries, the level of segregation has declined substantially.

Occupational Segregation

To gain a complete picture of how social identities correlate with the type of work people do, we need to study occupational segregation as well. Within a given industry, we routinely find occupational hierarchies, such as production workers, supervisors and managers. In this section we take advantage of the fact that we have data on earnings for various occupations and examine both, the extent of occupational segregation and its implications for earnings. The analysis is carried out only for the year 2021-22 using the most recent PLFS round. So we do not discuss time trends.

In the case of gender and religion, industrial segregation index is marginally higher than the occupational segregation index but for caste, occupational segregation is more pronounced than industrial segregation.

The overall segregation index for the different identities shows that industrial and occupational segregation indices in 2021 are quite similar to each other for any given identity pair (Table 6.6). However, it is a telling fact that in the case of gender and religion, the industrial segregation index is marginally higher than the occupational segregation

Table 6.6: Occupational segregation is the highest for gender, followed by caste and then religion

Identity	Duncan Index
Male-Female	0.37
ST-Others	0.34
SC-Others	0.32
OBC-Others	0.16
Hindu-Muslim	0.15

Sources and notes: PLFS 2021-22. Shown are Duncan Segregation Index values. See Methods Appendix for details.

index but for caste, occupational segregation is more pronounced than industrial segregation (compare Table 6.5 and 6.6).

Bringing in the monthly earnings data, for women we see a clear negative relationship between their occupational representation index and average monthly earnings from that occupation (Figure 6.9). The occupation titles for the codes shown in the figure are as per NCO 2015 2-digit sub-divisions. Although there is a lot of scatter in the data at the lower end of the earnings distribution, note that in one of the lowest paid occupations of personal care workers, with monthly average real earnings of less than ₹10,000, there are four times as many women than there are in the workforce as a whole (RI = 4). Similarly, other low-earning occupation categories such as cleaners and helpers, woodworking, garment and other craft and related trades workers, and food preparation assistants

have an over-representation of women amounting to at least twice their workforce share. On the other hand, women are under-represented in some of the highest earning occupations which include information and communications technicians; science and engineering professionals and associate professionals; production and specialised service managers; and business and administration associate professionals.

A similar negative relationship between representation index and average earnings also exists for Muslims (Figure 6.10). Street and related sales and services workers; and food processing, woodworking, garment and other craft and related trades workers are two low-earning occupations where Muslims are most over-represented. Information and communications technicians and professionals, science and engineering professionals, production and specialised service managers are

Figure 6.9: Women are over–represented in low paying occupations while men are over–represented in high paying ones
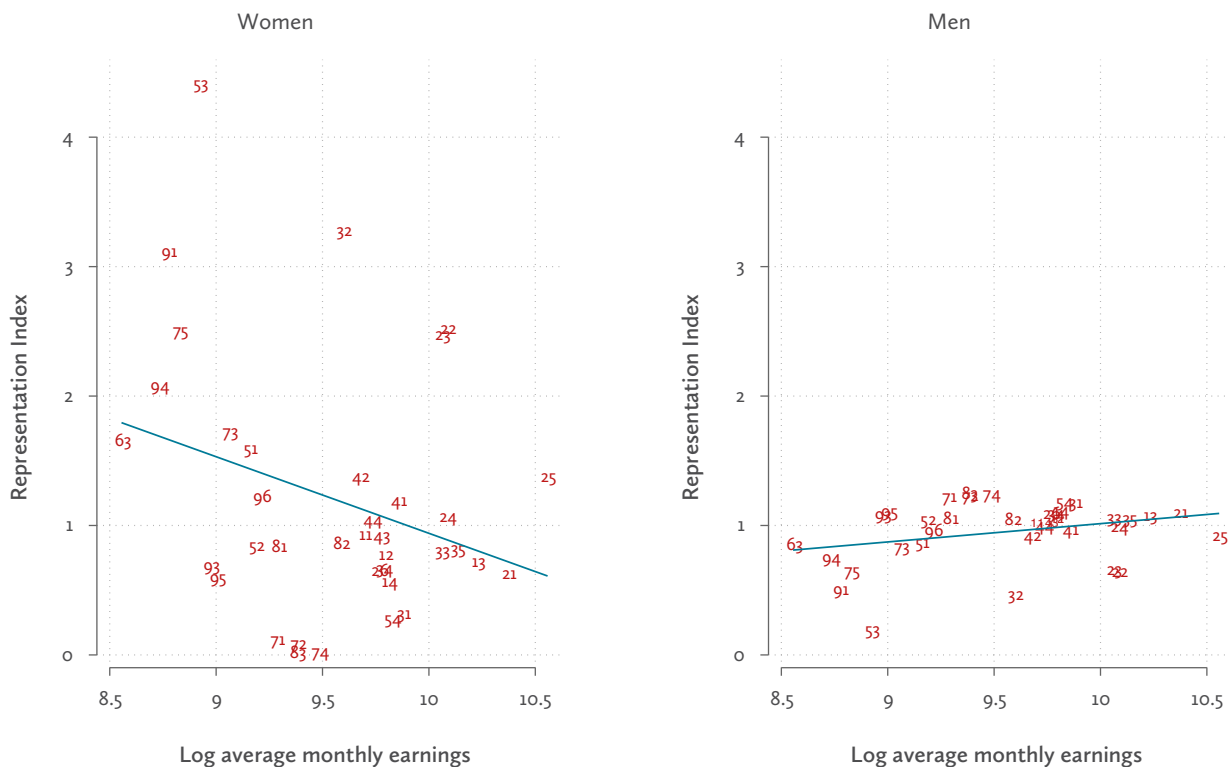
Sources and notes: PLFS 2021-22. See Results Appendix for occupation names corresponding to the codes. The X-axis shows log earnings. Actual earnings range between ₹5000 and ₹36000 per month.

Box 6.1: Occupational segregation at the level of jatis

Evidence from the limited body of empirical research on persistence of traditional occupations at the jati level suggests that caste-based ties continue to determine occupational outcomes especially for those from marginalised groups. Cassan et al. (2021), based on IHDS data from 2011, note how workers from different jatis are over-represented in their traditional occupations. More recently, Oh (2023) reports the findings from a field experiment where labourers who were offered various job opportunities were less willing to accept jobs linked to caste identities other than their own, especially if they belonged to historically marginalised jatis.

Findings from both these papers suggest that the link between caste and traditional occupation continues to persist. However, as noted in Chapter Seven of this report, we cannot comment on the strength of such persistence without data over time at the jati level. For example, Census 2011 data shows that Chamars and Mochis are over-represented in occupations that involve working with leather. But has this connection declined over time? If so, how fast? Cross-sectional data cannot answer this question.

Despite this caveat, Census 2011 data reveals three important features. First, the relationship between higher education and participation in occupations with higher income (and social prestige) is not a straightforward one. Across states, dhobis, for example, are 9 to 30 times overrepresented in their traditional occupation, 'Hand Launderers and Pressers' (NCO code 9133). A look at the educational outcome of Dhobis in Bihar, reveals that they have experienced the highest increase among all SC jatis in terms of their share of graduates and yet, are 11 times overrepresented in their traditional occupation. Similar to the experience of Dhobis,

independent of the educational experience, Valmikis are massively overrepresented among sweepers and cleaners across the states. Such evidence indicates that persistence of traditional occupations may not decline as fast as desired even with significant change in educational levels.

Second important feature we note, speaks to the differences across states. When contrasted with the experience of Mahars in MP, Mahars in Maharashtra experienced a relatively higher increase in higher education. And yet, the latter are over-represented in construction work, and slightly underrepresented among occupations such as corporate managers, while in MP they do better in occupational terms. This implies, similar levels of rise in education, even for the same jati, may not imply similar occupational trajectories across states. This further emphasises how the historical, political and social standing of jatis, which varies across regions, come together to shape the occupational outcome of any jati.

Third, we combine the NCOs at one digit level, based on whether the occupations are non-elementary in nature (that is, NCO codes one to eight) or elementary (NCO code nine) or unclassifiable (code X). We find that a majority of unique state-jati pairs (61 percent) have a higher share in either elementary occupations, or have reported occupations that are not classifiable under the existing NCO framework. Further, even non-elementary occupations (NCO code 1-8) are not free from traditional ties. For example, "shoemaking and other leather related crafts", which can traditionally be linked to Chamars and Mochis, appears under NCO code 7. Indeed, digging deeper into the non-elementary occupations, we note that as the NCO codes move from eight to one, that is, from lower to higher ranked occupations, the share of almost all

SC jatis starts declining, with very few jatis even participating in the highest ranked occupations.

These three observations, taken together, show the persistence of jati-level segregation and suggest that much more empirical work remains to be done on jati level occupational outcomes.

again the higher earnings occupations where Muslims are most under-represented.

It should be noted here that for men (in the case of Figure 6.9) and Hindus (in the case of Figure 6.10) the denominator for the RI is much higher than women and Muslims respectively (i.e. the former are a much greater share of the workforce). Hence the range over which the RI varies is smaller for the majority groups.

Among the caste identities, the relationship between occupational representation and earnings

is clearly negative for SCs and STs and positive for Others whereas for OBCs, the fitted line is almost flat indicating no clear relationship between the two (Figure 6.11). For SCs, occupation categories such as refuse workers and other elementary workers; cleaners and helpers; and street vendors (excluding food) are among the low earnings occupations where they are most over-represented. On the other hand, for STs, these are occupations like labourers in mining, construction, manufacturing and transport; food preparation assistants; handicraft and printing workers; and personal care workers. Others are clearly over-represented in all occupations with

Figure 6.10: Muslims are over–represented in low paying occupations while Hindus are over–represented in high paying ones
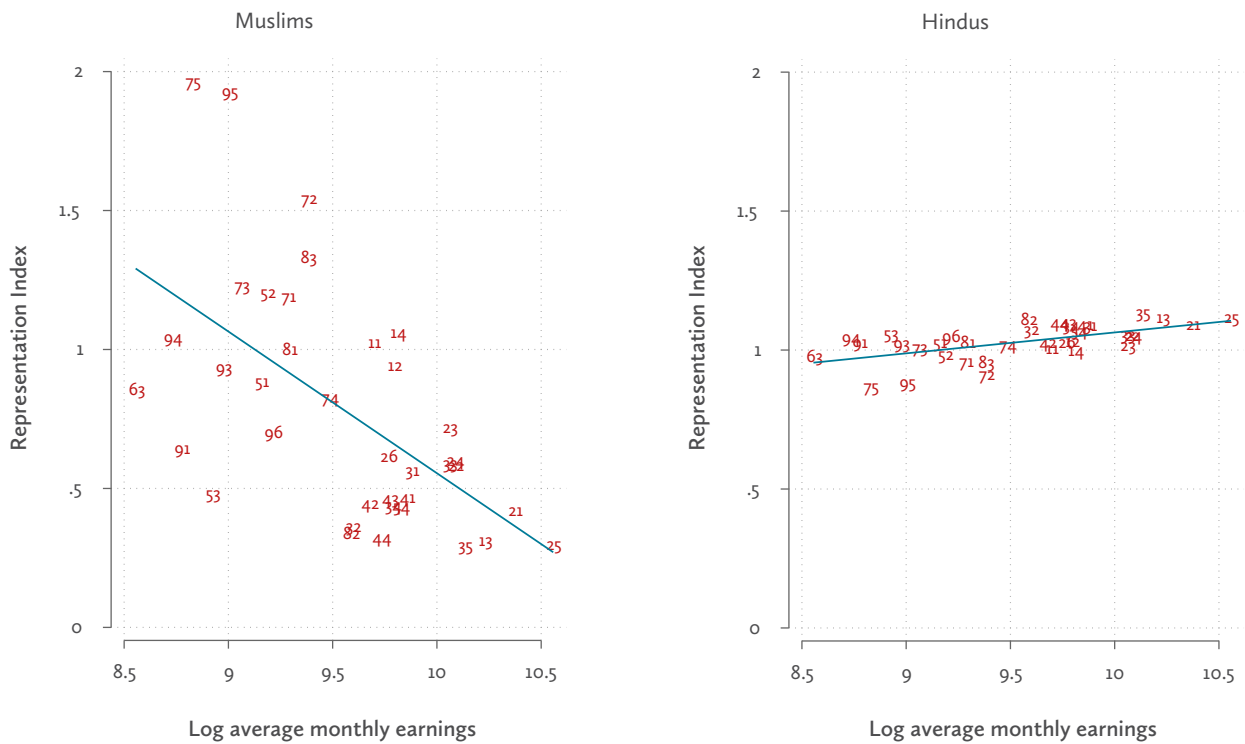
Sources and notes: PLFS 2021-22. See Results Appendix for occupation names corresponding to the codes. The X-axis shows log earnings. Actual earnings range between ₹5000 and ₹36000 per month.

Figure 6.11: Marginalised caste groups are over-represented in low paying occupations while dominant caste groups are over-represented in high paying ones
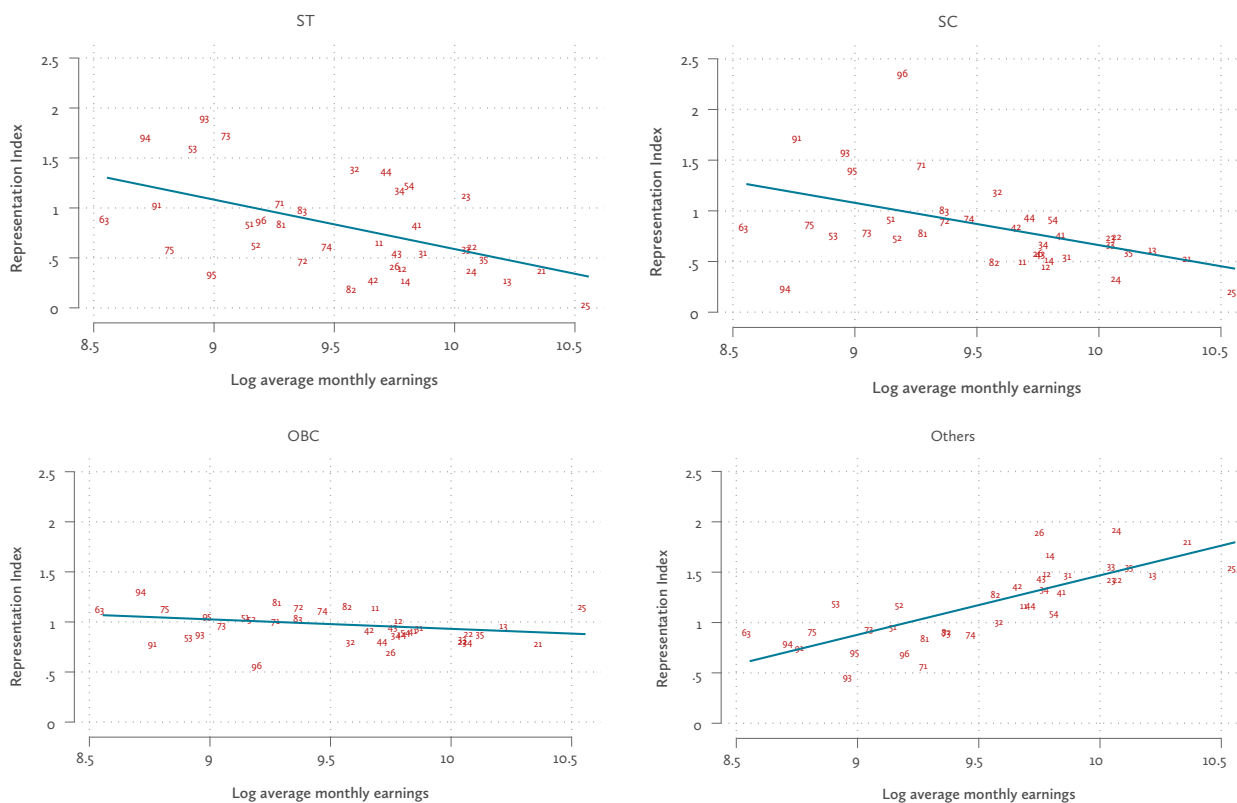
Sources and notes: PLFS 2021-22. See Results Appendix for occupation names corresponding to the codes. The X-axis shows log earnings. Actual earnings range between ₹5000 and ₹36000 per month.

earnings greater than ₹14,000 including the highest earning professional occupations in science and engineering, business and administration, and information and communication technology.

The scatter plots just discussed clearly demonstrate the continued penalty that disadvantaged groups pay due to being segregated into lower paying occupations. This explains why controlling for occupation accounts for a large part of the earnings gap. Thus discrimination operates at multiple levels. There is pre-market discrimination that affects educational attainment as well as quality of education, followed by labour market discrimination that channels different types of individuals into different types of work as per non-market social norms and finally discrimination that results in workers with the same characteristics and work

being paid different wages.

6.3 Caste penalties in entrepreneurship

Most of the literature on discrimination in labour economics focuses on wage work. In an economy like India, this leaves out roughly half of the workforce which is self-employed and where discrimination operates via other markets (credit, land, inputs and outputs). Such discrimination can make it harder to secure loans, hire labour, negotiate terms on input and output and so on. It is not easy to estimate the extent of discrimination in these markets but the combined effect of it can be to crowd entrepreneurs from disadvantaged backgrounds out of ownership of larger enterprises

and force them to operate mainly in the microenterprise sector.

Under-representation of disadvantaged social groups in large firm ownership can occur through processes of discrimination that manifest both at the level of individuals and institutions. The former can occur when there is active opposition from traditionally dominant business communities to the entry of persons from marginalised social groups in the sphere of business (Jodhka 2010). On the other hand, an important route of institutional discrimination is found to operate via exclusion or discrimination in the credit market, lowering the probability and amount of loan that can be obtained by entrepreneurs from socially disadvantaged social groups (Raj and Sasidharan 2018).

The Bhopal Declaration of 2002 adopted at the conference of Dalit intellectuals and activists under the sponsorship of the Madhya Pradesh government sought redistribution of land and democratisation of capital, to enable the dalit community to claim its fair share in the country's resources and assets (Nigam 2002). Soon after, the Dalit Indian Chamber of Commerce and Industry (DICCI) was founded in April 2005. Subsequently, a few studies have measured the extent of under-representation using Economic Census data.

Studies on caste and ownership of enterprises using the 2005 Economic Census have found that the share of Scheduled Castes and Scheduled Tribes in ownership of enterprises was much less than their shares in the population in both rural and urban areas (Sadana and Thorat 2009). Further, there had been little change in such trends between 1990 and 2005. Moreover, the average size of enterprises owned by SCs and STs was smaller than those owned by non-SC/ST owners (Varshney, Iyer, and Khanna 2013). A substantial earnings gap between SC/ST and non-SC/ST households engaged in business has also been reported (Deshpande and Sharma 2016).

Other studies using the Medium and Small Manufacturing Enterprises (MSME) Census data for 2001-02 and 2006-07 have found that similar caste disparities in ownership exist in the registered segment of the MSMEs and such disparities marginally increased between 2001-02 and 2006-07 (Deshpande and Sharma 2013). There is also evidence of homophily since major proportions of the SC/ST workforce are employed in SC/ST owned enterprises and much fewer proportions are in non-SC/ST owned enterprises.

In this section we examine the extent of under-representation of marginalised groups in firm ownership in the Indian economy, as of the most recent Economic Census data (6th round, 2013). Despite the dated nature of this data, it offers important insights into the persistence of caste-base disparities during India's high growth period. It also points to some important changes and improvements in the situation. We also quantify the extent of the penalty borne by marginalised groups in rupee terms, due to their exclusion from firm ownership, particularly their under-representation among owners of larger sized enterprises.

Data used in our analysis comes from different sources. Data on ownership of enterprises comes from three latest rounds of the Economic Census for the years 1998, 2005 and 2013. Data on the share of different social groups in the workforce comes from three rounds of the Employment and Unemployment Survey (EUS) for the years 1999-00, 2005-06 and 2011-12. Data on Gross Value Added (GVA) for enterprises in the formal sector is estimated from the Annual Survey of Industries (ASI) for 2005-06 and 2010-11. For enterprises in the informal sector, GVA data is estimated from the NSS Unorganised Manufacturing Enterprises Survey for 2005-06 and from the Unincorporated Non-Agricultural Enterprises Excluding Construction (NSS) for 2010-11. The EUS, ASI and NSS rounds are chosen such that they are the closest years corresponding to the Economic Census rounds.

Table 6.7: Share of different social groups in workforce and ownership of enterprises

Social Groups	Share in the workforce			Share in enterprise ownership		
	1998	2005	2013	1998	2005	2013
ST	10.7	10	10.1	4.4	4.1	5.4
SC	20.6	20.1	19.3	8.4	9.8	11.4
OBC	36.6	41.4	43.5	36.3	42.4	41
Others	32.2	28.5	27.1	50.9	43.7	42.1

Sources and notes: NSSO EUS various rounds and Economic Census various rounds.

The analysis here is carried out in two steps. The first involves constructing a representation index of enterprise ownership for different social groups and observing the patterns of under- and over-representation of certain castes in ownership of enterprises over time. The representation index for each caste group is constructed by computing the ratio of the share of that caste group in ownership of enterprises to its sharein the workforce. This representation index is then traced across enterprise size to understand if under- and over-representation patterns intensify as the size of the enterprise increases.

In the second step, the cost or benefit resulting from such under- and over- representation in the manufacturing sector is computed in the form of the value of output produced. This is referred to as the caste-penalty or caste-premium associated with skewed representation of social groups in enterprise ownership. For the manufacturing sector, a caste penalty/premium is assigned to under- or over-represented social groups by calculating the difference in total GVA *actually controlled* by a caste group and the total GVA that *would be controlled* by the caste group if their share in ownership of enterprises was same as their share in the workforce. This is computed separately for each enterprise size and then aggregated to arrive at the total caste penalty/premium.

Total GVA *actually controlled* by a caste group for any enterprise size is obtained by multiplying the average GVA of that enterprise size with the total number of enterprises of that size owned by the caste group. To obtain total GVA that *would be controlled* by a caste group if their share in ownership was same as their share in the workforce, the workforce share of the caste group is first multiplied by the total number of enterprises of a given size and this is then multiplied by the corresponding average GVA. Total number of enterprises is always obtained from the Economic Census data. These steps are explained in detail in the Methods Appendix.

6.3.1 Representation of Social Groups in enterprise ownership

The share of the four broad administrative categories of social groups in India – SC, ST, OBC and Others in the workforce and in the ownership of enterprises is shown in Table 6.7. It shows that as of 2013 around 30 percent of the Indian workforce belonged to the SC/ST categories. Approximately another 30 percent were Others and the rest were OBCs. Between 1998 and 2013 there were some changes to the exact percentage shares of these categories, particularly for OBCs and Others. Despite these changes, the rankings of the four groups have remained unaltered such that OBCs constitute the largest social group component of the workforce followed by Others, SCs and STs.

Figure 6.12: Marginalised caste owners tend to own smaller enterprise as compared to privileged castes
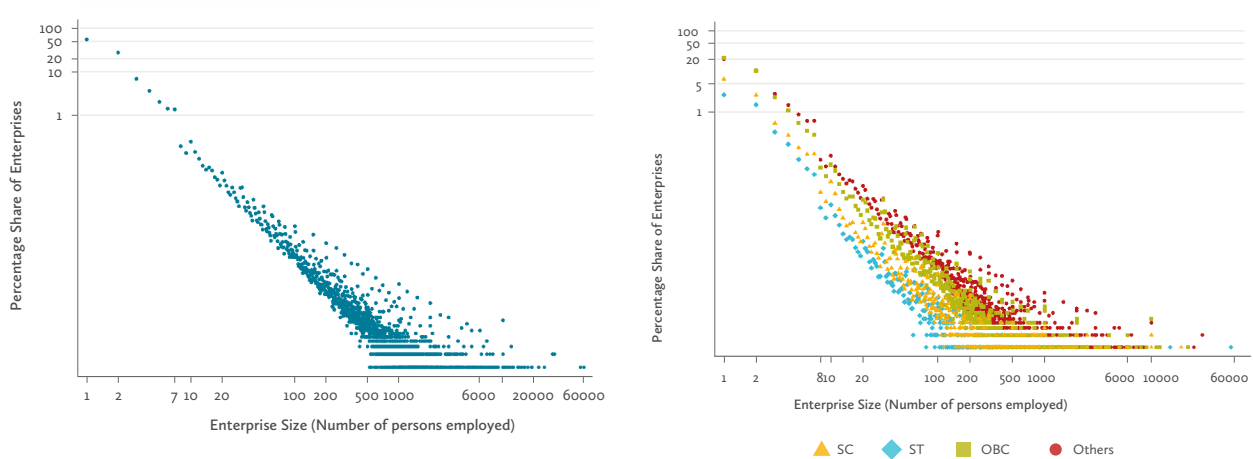
Sources and notes: 6th Economic Census, 2013. Both X and Y axis are shown in log terms for clarity

However, the picture is quite different when it comes to ownership of enterprises. The largest share belongs to Others and in all three years their share in ownership of enterprises has been significantly higher than their share in the workforce. In 1998, nearly 51 per cent of all enterprises were owned by them which declined over the years with a fall in their share in the workforce. However, despite this decline, they owned 42.1 per cent of all enterprises in 2013 which was still the highest share among the four social groups. The value of their representation index remained far above one, never falling below 1.5.

For SCs and STs, on the other hand, share in ownership of enterprises is significantly lower than their share in the workforce and it has remained so despite some increase in their ownership shares in 2013. In 1998 the representation index was only 0.4 for both SCs and STs. By 2013, there was some improvement in representation for both groups but even with this improvement, their representation index remains far below 1 (0.5 for STs and 0.6 for SCs). OBC is the only caste group whose share in workforce and enterprise ownership are similar in magnitude. As a result their representation index was 1 in 1998 and 2005 which in 2013 has declined slightly to 0.9.

This aggregate representation index is useful in understanding the broad trends in over-representation of privileged castes and under-representation of marginalised castes and their persistence. However, an aggregate measure like this treats all enterprises alike and conceals differences that might exist in enterprises owned by different social groups. One critical difference is the size of the enterprise. Owning a one-person (own-account) enterprise is very different from owning an enterprise that employs 20 persons both in terms of monetary value as well as economic power. We now examine the role of enterprise size in representation of different social groups in ownership of enterprises.

The size distribution of enterprises in the Indian economy is heavily skewed towards small and tiny enterprises. Majority of enterprises are very small and often operate in the informal sector. Panel (a) of Figure 6.12 shows the proportion of enterprises of each size in the Indian economy in 2013, where size is measured by the total number of persons employed. This figure shows that 60 percent of all enterprises in India were tiny one person enterprises as of 2013. The firm size distribution for private proprietary enterprises is even more skewed towards small and tiny enterprises since most of the very

Figure 6.13: Privileged castes are over-represented to a greater degree in larger enterprises
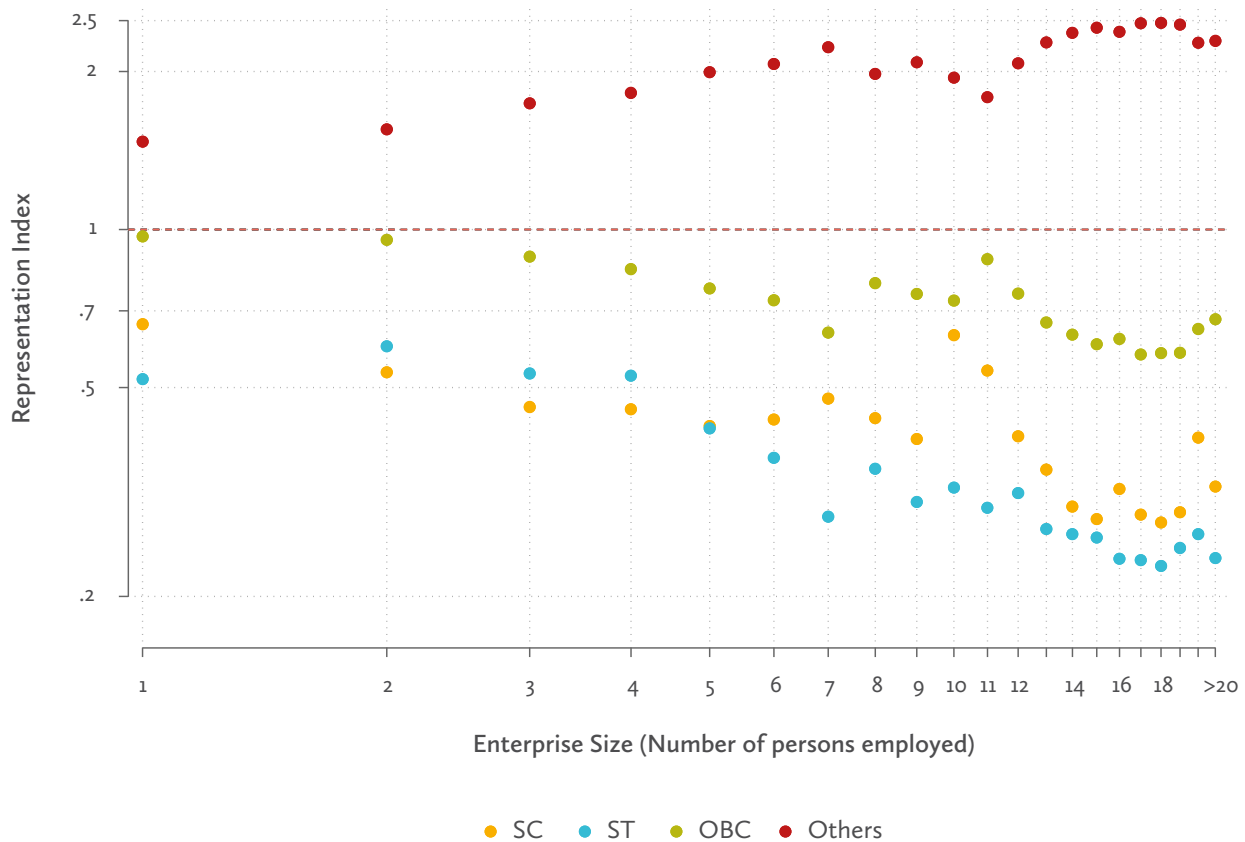
Sources and notes: 6th Economic Census, 2013. Both X and Y axis are shown in log terms for clarity.

large enterprises are non-proprietary and owned in the form of partnerships or limited companies. The mean size of proprietary enterprises in India in 2013 was 1.96 whereas the mean size of all enterprises in the economy was 2.2 workers.

Information on the caste of the owner is available in the Economic Census only for private proprietary enterprises. In the rest of the analysis we focus on these enterprises. In 2013, 89.5 per cent of all enterprises in the Indian economy were owned by private proprietors.

Panel (b) of Figure 6.12 shows the size distribution of private proprietary enterprises by the caste of the owner in 2013. The shape of this distribution is similar to that of the aggregate size distribution of enterprises. However, what stands out is that for the most marginalised social groups - SCs and

STs the distribution is clearly left-shifted compared to that of Others. That is, the proportion of larger enterprises owned by SCs and STs are substantially less.

It is worth examining the caste representation index across the firm size distribution. We construct the representation index for each caste group for every enterprise size up to 20 employee enterprises. Since the proportion of proprietary enterprises with more than 20 employees is only 0.3 per cent, all enterprises with more than 20 persons are clubbed together into one bin representing large private proprietary enterprises. Figure 6.13 plots this representation index for all four social groups against enterprise size. First note that not only are general category owners over-represented in all enterprise sizes but their over-representation increases with increase in size of the enterprise. On the contrary,

Figure 6.14: Representation index for different social groups by enterprise size in 1998, 2005 and 2013
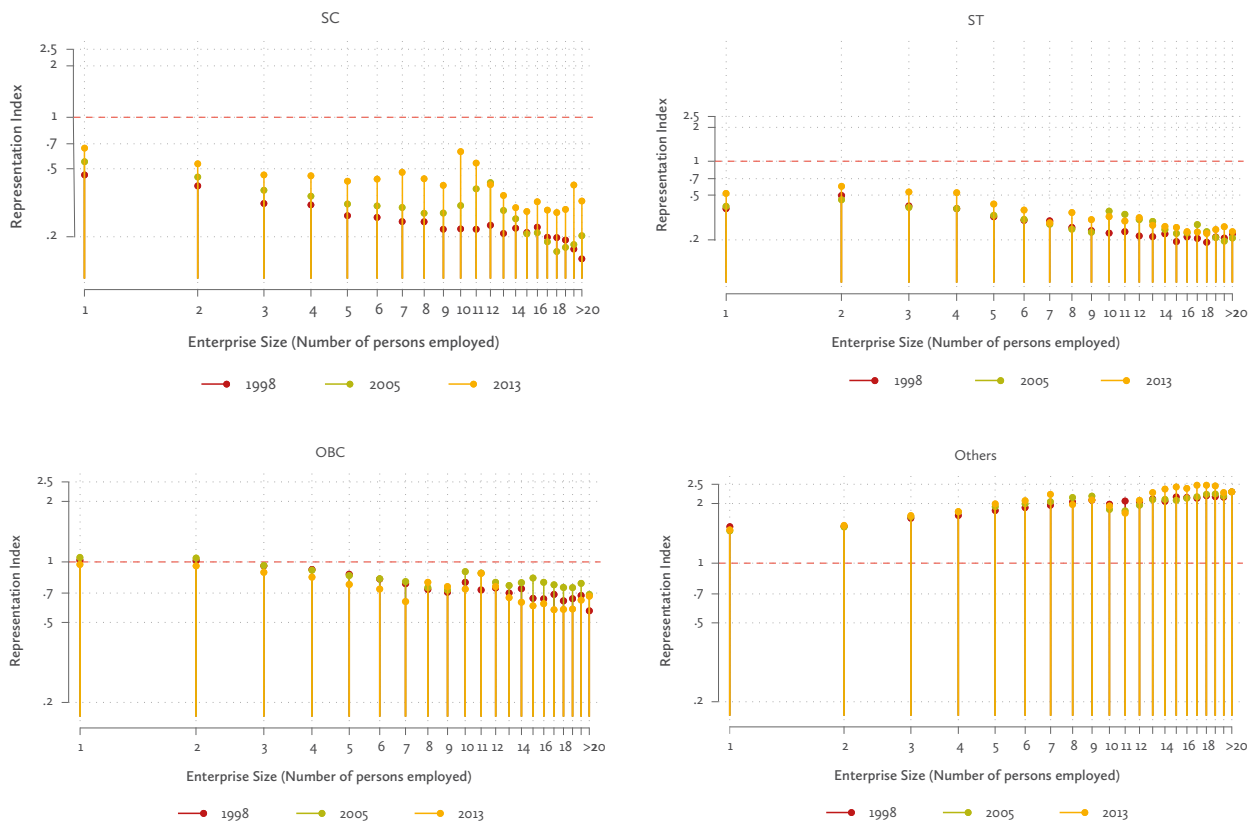
Sources and notes: Economic Census, various years. Both X and Y axis are shown in log terms for clarity.

SCs and STs are under-represented in all enterprise sizes including the smallest ones and under-representation intensifies as enterprises become larger.

In most of the larger enterprises, the representation index falls below 0.3 for these social groups. For OBCs, whose representation index is close to 1 (0.9) at the aggregate, this is only on account of their high representation in very small enterprises employing 1 to 3 persons. As enterprise size increases beyond 3, OBC representation in ownership falls and continues to remain far below one.

General category owners over-represented in all enterprise sizes and their over-representation increases with increase in size of the enterprise.

Interestingly, although over- and under-representation of privileged and marginalised groups intensifies with increase in enterprise size, this trend does not appear to be continuous. Visual inspection reveals a break at enterprise size of 10-11 for Others, OBCs and SCs. For STs the break is less clear. Since the point of break coincides with one threshold for formal sector entry, it appears that the size effect operates separately within the informal and formal sectors.

Constructing representation indices with data from two previous rounds of the Economic Census (1998 and 2005), makes it possible to track changes for different social groups over one and a half decades. Figure 6.14 shows that despite persistent under-representation in all enterprise sizes, there was some improvement in representation of SCs and STs in this time period, particularly between

Figure 6.15: Distribution of caste penalty/premium across different enterprise size in 2013
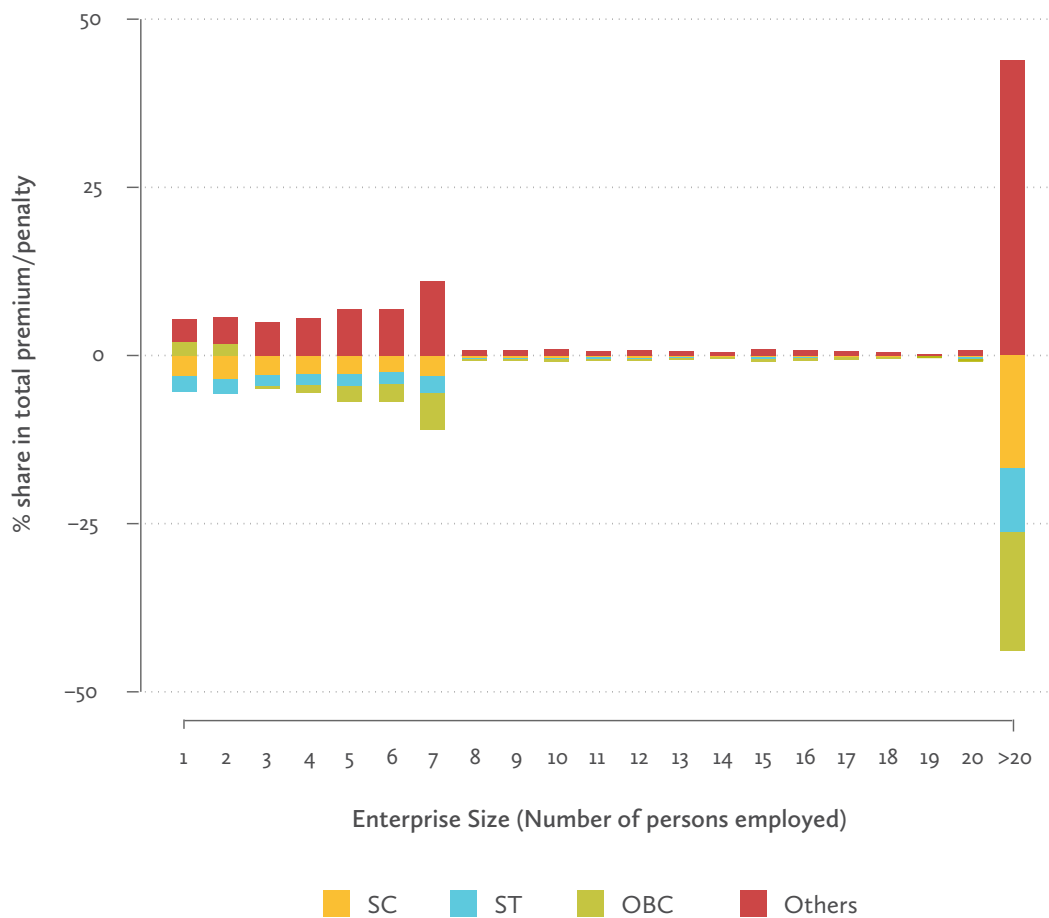
Sources and notes: Economic Census, EUS, ASI and NSS data. See text and Methods Appendix for details on calculation of the premium/penalty.

Figure 6.16: Caste premium/penalty as a share of proprietary manufacturing GVA
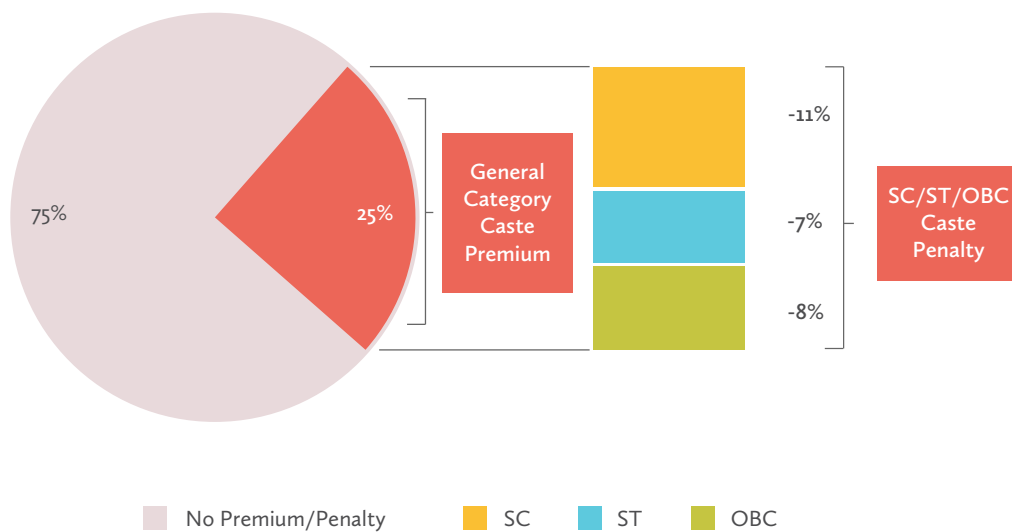
Sources and notes: Economic Census, EUS, ASI and NSS data. See text and Methods Appendix for details on calculation of the premium/penalty.

2005 and 2013. The improvement was faster for Scheduled Castes, particularly in medium and large enterprises. However, what is interesting is that this improvement was not accompanied by a decrease in over-representation of Others. On the contrary, their over-representation, particularly in larger enterprises, increased between 2005 and 2013. As a result, in 2013 the share of Others in ownership of enterprises employing 15 or more workers ranged between 2.3 to 2.5 times their share in the workforce.

The improvement in SC/ST representation since 2005 came at the cost of an increase in under-representation of the OBCs. Except for enterprises employing 8 and 9 persons, the representation index for OBCs declined in all other enterprise sizes between 2005 and 2013. This reversed whatever improvement in OBC representation in enterprise ownership had occurred between 1998 and 2005 and the representation index for OBCs in 2013 was lower than their values in 1998 in most parts of the enterprise size distribution. Thus, when the improvement in SC and ST representation is seen in comparison with changes in representation of OBCs and Others, it is clear that the improvement is the result of some flux at the lower end of the historical caste hierarchies. The domination of Others in ownership of enterprises and thereby in consolidation of economic power consistently persisted since 1998 and further strengthened since 2005.

6.3.2 Estimating the caste penalty/ premium

From the above discussion on representation of different social groups in ownership of enterprises, it is evident that the marginalised social groups in India are consistently disadvantaged. One way to quantify the penalty borne by marginalised groups due to lack of representation in firm ownership is to express it in terms of lost value of output. We calculate the caste-penalty for marginalised castes (which in magnitude is the same as a caste premium

for General castes) for each enterprise size and then aggregate it to arrive at the total value of the penalty/premium. Figure 6.15 shows the contribution of enterprises of different sizes to this aggregate penalty/premium in 2013. More than 45 percent of the total penalty came from under-representation of social groups in the largest enterprises employing 20 or more persons. This is because these are relatively more productive enterprises, and the extent of under-representation is also very high in these enterprises.

Another 44 percent of the total value of penalty/ premium came from the smaller enterprises employing up to 7 persons. The contribution of the smaller enterprises is sizable despite these being relatively less productive and under-representation being less severe. This is because the total number of such small enterprises is huge and therefore the aggregate proprietary GVA produced by all these enterprises together is also substantially large. The remaining 10 percent of the penalty/premium comes from the middle-sized enterprises employing 8 to 20 persons. The lower contribution of these enterprises in total penalty/premium is owing to the fact that they are both relatively low in value and fewer in number.

We estimate the total value of General-caste premium in the ownership of private proprietary enterprises in 2005 to be ₹16,400 crores in 2004-05 prices. This increased to ₹28,400 crore in 2013 (in 2004-05 prices). Of this, the caste-penalty value was ₹12,200 crores for SCs, 8,600 crores for OBCs and 7,700 crores for STs. Adjusting to 2013 prices, the total penalty comes to ₹42,000 crores.

Although the real value of the penalty/premium in absolute terms increased between 2005 and 2013, total manufacturing GVA also increased in this time period. As a fraction of the total private proprietary GVA in Indian manufacturing, the caste penalty/ premium remained constant at around 25 percent. Therefore, consistently around one-fourth of all

private proprietary GVA that is being controlled by the General caste owners would be under the control of marginalised social groups if there was no bias in caste representation in ownership of enterprises (Figure 6.16).

One-fourth of all private proprietary GVA that is being controlled by the General category owners would be under the control of marginalised social groups if there was no bias in caste representation in ownership of enterprises.

What has changed between 2005 and 2013 is the distribution of the penalty. While SCs continue to remain the biggest losers even in 2013, a part of the penalty accruing to them in 2005 has been transferred to OBCs in 2013.

6.4 Conclusion

This chapter brings us to a conclusion of the in-depth analysis begun in Chapter Two of the report. We have examined both short-run and long-run changes in labour market outcomes of key social identities in India. In this chapter, we saw that even as raw earnings gaps have narrowed, disadvantaged groups still pay significant penalties with respect to earnings, segregation into less desirable industries and occupations as well as exclusion from entrepreneurship. In some cases, penalties have even increased over time, such as the gender penalty for women engaged in regular wage work in rural areas. Gender-based segregation has also increased over time, even as caste-based segregation has fallen. Going beyond administrative caste categories, whatever limited evidence exists from secondary data also points to significant persistence of jati-based occupational segregation.

All this is by way of saying that, if we wish growth to be truly inclusive, policies that target the reduction of identity-based disparities are essential. We discuss a few of these briefly in the last chapter. Before that,

however, in the next chapter we discuss how surveys can be improved to provide better information on social identities and labour market outcomes.

Endnotes

1   Interested readers are pointed to the Sachar Committee Report (2006), Madheswaran and Attewell (2007) Majumder (2007), Sengupta and Das (2014), Duraisamy and Durasamy(2016), Deshpande and Sharma (2016), Thorat and Madheswaran (2018), Deshpande, Goel and Khanna (2018) .

2   For examples of selection-corrected measurements for earnings gaps see Das and Dutta (2007). But these authors also note the concerns with selection correction methods and present uncorrected estimates as well. Deshpande, Goel and Khanna (2018) present only selection-uncorrected estimates for the same reason.

3   More specifically they show that the pooled Oaxaca-Blinder can overstate the role of observables in explaining mean outcome as compared to OLS with a group indicator. They recommend that pooled O-B decomposition should not be used to distinguish between explained and unexplained gaps, although this method may be useful to assess how much of an unexplained gap represents discrimination if specific assumptions are met. We have verified that the results obtained from a Oaxaca-Blinder exercise are not qualitatively different from the OLS results and in some cases are quantitatively almost the same as well.

4   Interested readers may note that, as expected from the discussion in Section 6.1.2, the unexplained gender gap in the year 2021 as measured by the Blinder-Oaxaca decomposition (0.357) is very close to the coefficient on the gender dummy for that year (0.346, see Table 6.3).

7

Improving survey-based measurement of labour market outcomes

7

Improving survey-based measurement of labour market outcomes

We live in a data-saturated world. Not only are large amounts of data being generated every day during the course of normal economic and administrative activities, but also governments everywhere are keen to use data for design and implementation of policy. In these circumstances, it becomes ever more imperative to understand the strengths and limitations of different kinds of data. Such an understanding enables us to build a data architecture that does justice to our complex needs and helps us steer clear of unproductive debates; such as is administrative data better than survey data or the other way around? Or, in this new world of "big data", do we still need to bother with household surveys? In this chapter, we begin with a few brief remarks on the need for multiple data sources. Then we proceed to a detailed discussion on how to improve the ability of surveys to capture labour market outcomes for disadvantaged social identities. We end with some remarks on phone surveys, which have become increasingly popular since the Covid-19 pandemic.

7.1 Strengths and weaknesses of various data sources

Data generated in the process of policy administration is called administrative data. Examples relevant to employment include the Management Information System (MIS) for MGNREGA or the Employee Provident Fund

Organisation (EPFO) database. This data is generally expected to be a "census" of the relevant population. That is, every worker who has worked under MGNREGA is part of the MIS and every employee who contributes to PF is part of EPFO. These databases are usually updated in the course of policy administration and running of programmes. Thus their main strengths are that they provide a complete picture of the relevant population in real time.

The quality of administrative data is as good as the processes that generate it. It is also subject to political pressures. This means that such datasets may record information that is inaccurate because it suits a political or administrative purpose. But a deeper issue with administrative data is that it cannot (indeed it is not designed to) give a comprehensive picture of the entire workforce or the population. That is, it only has a selected sample. For example, for EPFO data, this is the sample of regular wage workers who contribute to the PF system. This is an important limitation when it comes to drawing conclusions about the entire workforce only from this selected sample.

Databases generated in the course of normal economic and social activity, such as Google mobility or Indian railways data, online transactions on digital platforms are voluminous and useful for fast frequency analysis. In this respect, they are similar to administrative data. Unlike administrative data, they are less likely to suffer from reporting issues since actual transactions are involved. But they suffer from

selection problems like administrative data, in that the information only pertains to that part of the population that engages in the relevant activity (e.g. owns a smartphone or travels by train or uses digital platforms). The fraction of the population in this net may be increasing, but this does not alter the fact that this is also a highly selected sample from which general conclusions (say on earnings or spending etc) cannot be drawn.

This brings us to survey data. There are three crucial ways in which survey data is valuable. First, a well-designed and executed sample survey can be far more representative of an entire population than data generated through either administrative or commercial activities which suffer from selection problems. Second, surveys can be designed to gather information on specific questions or demographic groups while administrative or commercial data is a by-product of activities undertaken with different purposes in mind. Third, by asking the right questions, surveys can not only provide us with outcomes but also insights into the mechanisms driving these outcomes.

Survey data can suffer from both sampling and non-sampling errors. An example of the former kind is incomplete or out-of-date sampling frames from which the sample is drawn or inability to reach a part of the sample (say missing out affluent sections due to denial of entry). It is important to emphasise that sample surveys are in principle representative of the entire population, though in practice eliminating all sampling error is an ongoing process. In contrast, administrative data, especially in India, where informal activities predominate, is not representative of the population by design.

Some examples of non-sampling error are incorrect self-reported earnings, incorrect information provided by the informant about other members of the household or deliberate undercounting of employees by employers. Such errors can result in noisy estimates as we show later in the chapter.

However, the existence of such problems does not mean survey data is not useful. It only means that our estimates will have error margins around them and we must be careful not to draw large conclusions from small differences. It also means that generally trends are more reliable than levels. Until the Indian economy reaches the point where the vast majority of workers are in the tax database or the vast majority of firms are registered, sample surveys will continue to be relevant.

In recent years, the growth of administrative data has raised questions on the continued relevance of sample surveys. The draft National Policy on Official Statistics (NPOS) issued in August 2023 by the Ministry of Statistics and Programme Implementation (MOSPI) states "Reduction in respondent burden through minimization of surveys and optimizing use of Administrative Statistics" as an objective, adding that "as the volume, content and coverage of administrative sources grows, so does their attractiveness as an alternative to statistical surveys" (MOSPI 2023, p. 18).

However, when it comes to employment, consumption and several other questions of living standards, surveys are indispensable in societies such as India where the majority of workers and firms are outside the formal tax or administrative net. Administrative, commercial and survey data are complements and not substitutes. Careful work using all three sources can give us a much better picture of the economy than any one in isolation.

Surveys are indispensable in societies such as India where the majority of workers and firms are outside the formal tax or administrative net. Administrative, commercial and survey data are complements and not substitutes.

In this chapter we reflect on ways in which surveys can be strengthened in the measurement of labour market outcomes, particularly for socially

disadvantaged groups. We draw on results from the India Working Survey and supplement these with other research. Employment and social identities - the central themes of this report, are both vulnerable to different kinds of issues with respect to their measurement. These are conceptual - how well concepts are defined or how clearly categories are delineated, as well as operational - how these concepts are used in survey instruments, who is being asked what and how it is being asked. The conceptualisation of measures and categories in which social identities are captured, as well as how surveys are implemented, are all social processes, embedded within political contexts. Recognising this helps us understand the limitations of the data generating process and forces us to be careful in our interpretation of labour statistics.

7.2 Improving and refining our concepts and categories

Limiting ourselves to the theme of this year's report, we reflect on issues of measurement surrounding employment and identities.

7.2.1 Concepts of employment, work and skills

Debates and definitions on what constitutes employment and work have a long history. Recent international attempts at formalisation have come through the ILO's International Conferences on Labour Statistics (ICLS). In 1982, the 13th ICLS had defined the employed, viz the "economically active population" to include "all persons who furnish the supply of labour for the production of economic goods and services as defined by the United Nations systems of national accounts and balances" (ILO, 1982,par.5).

In 2013, the 19th ICLS updated this approach, defining employment as "work performed for others in exchange for pay or profit" (ILO 2013). The updated definition diverged from the earlier understanding of employment in excluding production of goods for own use from employment. Instead, the 19th ICLS introduced "forms of work" of which employment was one, with activities under employment aligning with the System of National Accounts (SNA) production boundary (Figure 7.1). With this change, for example, individuals engaged in farming where the produce was consumed within the household

Figure 7.1: Conceptualisation of forms of work in 19th ICLS

Intended destination of production	for own final use		for use by others					
Forms of work	Own-use production work		Employment (work for pay or profit)	Unpaid trainee work	Other work activities	Volunteer work		
	of services	of goods				in market and non-market units	in households producing	
							goods	services
Relation to 2008 SNA		Activities within the SNA production boundary						
	Activities inside the SNA General production boundary							

Sources and notes: ILO (2013)

would no longer be counted in the labour force, although their activity was still recognised as work.

The definitional changes have been criticised both for exclusion of own use production, as well as the break in statistics. For example, in Rwanda, the shift to the revised definitions resulted in a drop in employment rate especially for women from 84 percent in 2014 to 52 percent in 2019 (Beegle 2023). Gaddis et al (2023) also observed similar implications for sub-Saharan African countries. Although India has yet to adopt the 19th ICLS recommendations into its labour force surveys, the outcomes are likely to be similar with a fall in the recorded number of women in employment.

From a measurement point of view what is important is that all non-leisure activities should be clearly identified and counted, whether they result in income (cash or kind) or in production for consumption of goods and services. Once the information has been gathered, stricter or looser definitions can be constructed depending on the needs of the analysis. The same applies to the intensive margin of the number of hours a person needs to be engaged in a particular activity in order to qualify being employed. The lowest threshold used in practice is one hour per week. Obviously, this is a very low bar and a person who only engages in remunerative work for an hour of a week is not employed in any meaningful sense. However, the bar is kept purposefully low so as to capture any activity however small. Subsequent collection of information on actual hours spent per week in the activity then allow us to construct any definition of employment we desire. In the India Working Survey. we took this approach of gathering information of all non-leisure activities at the extensive (yes/no) and intensive (hours spent) margins so that different definitions could be constructed as required.

Finally, it has proved consistently difficult to adequately capture the level and types of skills that are prevalent in the informal economy. Labour force surveys usually produce very low estimates of the

proportion of the workforce that is skilled. This is not only due to the lack of formal certification systems, curricula, diplomas etc among informal workers, but also due to deeper conceptual problems that result in a conflation of skills with formal education. Targeted surveys as well as ethnographic work reveal that workers in the informal economy routinely go through apprenticeships and other forms of on-the-job training, as well as acquire significant skills during the course of normal work experience. This includes hard skills such as welding, carpentry, repairs, or weaving, as well as soft skills such as speaking English (Sengupta, Gaurav, and Evans 2021; Basole 2018; 2014). There is, thus, a clear need to refine our concept of skills as well as train enumerators better to capture the skill landscape in labour force surveys.

Labour force surveys usually produce very low estimates of the proportion of the workforce that is skilled. This is not only due to the lack of formal certification systems, curricula, diplomas etc among informal workers, but also due to deeper conceptual problems that result in a conflation of skills with formal education.

7.2.2 Defining identity–based categories

Large-scale surveys usually depend on mutually exclusive and exhaustive categories. When it comes to social identities, the level at which information is gathered in surveys and the level at which identities actually matter for outcomes can be very different. A classic case of this, in the India context, is caste. Most surveys collect data on caste using administrative categories of Scheduled Castes (SC), Scheduled Tribes (ST), Other Backward Classes, and Others (residual category). While these categories are obviously relevant for public policy, for most Indians, their lived experience is not that of their administrative caste category, but rather of their jati. As we have seen in the preceding chapters,

there is large variation within the administrative category at the jati level for Scheduled Castes when it comes to occupational segregation or educational outcomes. There is no reason to doubt that this will be true for other administrative caste categories as well, in particular for OBCs, a large and highly heterogeneous group. Thus, if we are interested in understanding to what extent social norms influence labour market outcomes, we need to identify individuals by their jatis and not only their caste category.

But jati-level data is seldom collected. This is because there are a large number of jatis and the list changes across regions of India. Not only this, their hierarchical ordering also varies across regions, as does the fact of inclusion or exclusion from a particular administrative category. The principal source of labour statistics in India, the NSSO EUS and PLFS surveys do not collect jati information. The last available population census (2011) for India collected information on jatis only for Scheduled Caste and Scheduled Tribe categories (see Box 7.1). The IHDS and the CMIE-CPHS do collect jati information for all households. But the data is difficult to use because it comes in the form of a large set of text strings where the same jati may be spelt differently in different places, the same jati called by different names in different regions, or different jatis in different regions being each other's cognates. Recent advances in data cleaning technologies do allow us to use such data and some interesting research has started coming out, but it is early days (Cassan, Keniston, (also see Box 7.2) and Kleineberg 2021; Asher, Novosad, and Rafkin 2021).

Continuing on the theme of caste, the second issue is the limitation of currently available data on the intersection of caste and religion. In recent years, there has been some progress in understanding the intersections between caste and gender, as well as jati and gender in determining economic and social outcomes (Joshi, Kochhar, and Rao 2018; Deshpande 2002). However, in the case of religion and caste,

there are very few studies that look at the two together. The common understanding of caste as a Hindu system of social organisation has led to a blind spot when it comes to differential outcomes for individuals from different castes among other religions. But caste-based stratification exists among Muslims, Sikhs and Christians too (Azam 2023; Ansari 2018; Teltumbde 2010; Government of India 2006). Legally, Sikh and Buddhist Dalits have been included in the Scheduled Castes list while the demand from Muslim and Christian Dalits to be recognised officially continues (Ansari 2016).[1] That said, NSSO surveys such as the PLFS do allow for collecting administrative caste groups for all religions. We have used this data in this report to estimate some within religion caste-based differences for Hindus and Muslims. But the quality of this data needs improvement since they are self-reported and not verified against the official schedules published by states (Paliath 2021).

7.3 Improving survey conduct and operations

Large-scale representative surveys are powerful tools that have the potential to generate unbiased estimates of important economic outcomes. However, unbiasedness is not guaranteed. Advances in survey methodologies as well as in social sciences have highlighted errors (sampling and non sampling) in the process of data collection that can produce bias (Perez 2021; Agrawal and Kumar 2020; Biemer and Lyberg 2003). Consistent efforts are needed to ensure that sampling is random, errors of exclusion do not occur, and questions are asked to people who can provide accurate answers, in ways that get the information needed. This requires careful sampling and questionnaire design, but also sensitisation and training of survey enumerators. In this section we address these questions.

Box 7.1:   Census data on jatis: uses and limitations

Over the past decade, most of the media coverage and scholarship around the Socio-Economic Caste Census (SECC) has pointed out that jati level data has not been officially collected at the national level since the census of 1931. Such a statement, however, is only partially correct. Every Census conducted since 1951, has collected varying amounts of jati level information, but only for SC and ST groups. The obvious implication is that jati level analysis across SC/ST and other groups is not possible. While this restricts researchers from commenting on the socio-economic trajectories of those at different gradations of the caste hierarchy, at least it is possible to track education and/or employment outcomes of different jatis within the SC and ST categories over the decades. Specifically from 1961 onwards, the population statistics for each such jati, as well as the employment statistics by industrial categories are available either in PDF or XLS format from the census website.

There are a few challenges worth noting though. First, for each census, the coverage of jatis has changed based on their status of SC or ST membership. Second, the levels of aggregation for industrial categories of employment (as well as education) varies across different census reports. For example, data on the industrial category of workers by jati identity is restricted to only four industrial categories in 2001 as opposed to nine or more categories in 1991 and 2011. These two aspects, together, make it difficult to create a balanced jati panel at the state level.

The third difficulty arises from the fact that the same jatis may have different names across states, and even within states. The only way to harmonise the jati names across regions is to first minimise the spelling variation (for example, Chamar and Chambhar), and subsequently, systematically categorise jati names using one of the volumes from the People of India project (Anthropological Survey of India, 1985) that list the history, equivalence, and nomenclature of the jatis across regions. A recent example of such a matching exercise is the work of Cassan et al. (2021), where the authors match jati names from IHDS with the People of India project. But as the authors themselves noted, it is a laborious process and may still miss some of the jatis based on the coverage of the data itself.

Fourth, some of the most important information from the perspective of jati level analysis is available only for one round of census (2011). For example, the four-digit classification of occupation (NCO) at jati level is potentially helpful in capturing the persistence of traditional occupations within the SCs, but without a comparison over time, such an analysis is incomplete. That is, in the absence of data from other rounds, one cannot comment on the degree to which the caste based ties are weakening (or not) over time.

The jati level analysis presented in this report, is thus conditional on the extent to which we could overcome the challenges listed above. To summarise, without a systematic collection of jati level data either through surveys or through exercises such as the Socio-Economic Caste Census, it remains extremely difficult to compare the socio-economic dynamics across jatis. The limited pool of data available on the jati level outcomes for SCs and STs clearly indicates significant differences at the jati level underlining the need for such data.

Box 7.2:   Collecting and using jati level data – the IWS experience

As part of the India Working Survey, we collected self-reported jatis for all respondents. The enumerators were provided a drop-down menu of jati names based on the state's official list, but they could also type in a different jati that was not on the list. Enumerators would first note the administrative category and then would be directed to a drop down list of state-specific jatis corresponding to that category. While we had expected a few instances of out-of-list names, we ended up having around 20 percent of all respondents giving names outside the selected list. Some of these were because of alternative spellings, some because they were genuinely not in the list and some because the administrative caste category reported by the respondent did not have the jati they belonged to. As a result we ended up recording more than 1,800 unique jati names in just a few districts of Rajasthan where the entire state's official list had a total of 378 jatis. In Karnataka we recorded more than 1,500 jati names against the official list of 1,124.

An important source of duplication, as mentioned earlier, are alternative spellings of the same jati's name. For example, Lingayat can be spelled variously as Lingayata, Lingayath, Lingayatha etc. We used an algorithm called Soundex that assigns a phonetic code to a word, which represents how it sounds. For example, all the four variants of the spelling of Lingayat (including ones that have typographical errors like Lengayeetha) are given the same Soundex code L523, whereas a different sounding jati name, like Kuruba (and variants of this spelling), will have a different code K610. This allowed us to reduce the number of unique jatis in Rajasthan to 182 and in Karnataka to 668.

Some problems still persist. There are similar jati names that are spelled differently enough for Soundex to assign different codes to them, and perhaps more importantly, similar sounding jati names can represent different communities in different parts of the same state (eg. Bovi and Bhovi jatis in Karnataka), which Soundex would attach the same code to.

In sum, a combination of using a drop-down menu based on official jati lists, training enumerators on possible spelling variations and using post-data collection statistical analysis can generate some usable jati level data.

7.3.1   Who is counted? Exclusion from sampling frames

In a sample survey, exclusion may occur in two ways -a unit may not appear in the sampling frame despite being in the population or may not be interviewed despite being selected. Surveys often suffer from "sedentary bias" wherein the sampling frame excludes itinerant populations or migrant workers who may not have fixed residences (Carr-Hill 2013). Geographical remoteness, conflict and political instability also play a role in making regions inaccessible to surveys (Agrawal and Kumar 2020; Carr-Hill 2013). Further micro-segregation of communities along identity lines means that the distribution across space is not random. If a listing (frame-creation) exercise misses a region, once again, the result is a biased sampling frame. Finally, if the sample is not drawn from a list, but rather via a random walk rule of some kind, the starting point or other aspect of local geography may interact with micro-segregation to create the possibility of bias (Somanchi 2021).

In urban areas, another problem that arises is that of lack of access to households in gated communities and apartment complexes. In dense localities an apartment complex may represent an entire urban primary sampling unit (PSU). A refusal at the point of entry can lead to the entire PSU being excluded from the survey. Under-representation of the rich in sample surveys has been a persistent problem , one that we faced in the case of the India Working Survey too.

Finally, certain groups may refuse to participate in surveys out of fear or suspicion regarding the object of the exercise. This was the case for the India Working Survey fieldwork which coincided with the ongoing protests around the Citizenship Amendment Act and the National Register of Citizens in 2019-20. We faced high rates of refusal for participation from Muslim households.

7.3.2 Who is asked? An analysis of "self-proxy" data from the India Working Survey[2]

Once the household has been selected for survey, the next question is, who, in the household is administered the questionnaire. Respondents within a household are often non-randomly chosen for accessibility reasons (available to answer the survey). The choice of informant has implications for measurement when questions are asked about other household members on whose activities they may not have information on or may have preconceived notions about.

The role of proxy respondents has been studied in the high income country context (Bound, Brown, and Mathiowetz 2001; Blair, Menon, and Bickart 1991), and more recently in middle and low income countries, largely with respect to the reporting of employment, child labour and asset ownership (Kapur, Vaishnav, and Verley 2021; Galdo, Dammert, and Abebaw 2021; Ambler et al. 2021; Koolwal 2021; Muller and Sousa 2020; Bardasi et al. 2011). In the

domain of autonomy, decision making and asset ownership, Ambler et al.(2021) find that spousal disagreements on asset ownership are substantial and systematic especially with regard to women's ownership. Disagreements are higher in the case of assets that are easier to hide. For employment, Bardasi et al (2011) find that male employment is sensitive to respondent selection with women under-reporting men's work, while the reporting of women's work by a male proxy had no significant deviation from self reports. In a study of five urban clusters in India, Kapur et al. (2021) find no correlation between gender of respondent and estimation of women's employment.

In this context, one of the research objectives of the India Working Survey (IWS) was to examine how informant identity (specifically gender) affected labour estimates. One adult male and one adult female were randomly chosen from each household after completion of the household roster. Both were interviewed separately. If this male-female pair happened to be spouses they were asked to report about their own work and employment activities as well as that of their partner. A detailed weekly module with six questions to ascertain labour market participation was administered. Table 7.1 details the questions asked to the respondents about themselves and their spouses.3

Table 7.2 reveals the differences in self-proxy reports for men and women. Typically, convergence between self and proxy reports is taken as an indicator of reporting accuracy (Blair, Menon, and Bickart 1991). In the absence of a third source of validation data to triangulate the different reports, we consider the self-reported estimates to be the benchmark and under-reporting and over-reporting are estimated using self-reports.

Before proceeding further, it is worth noting that the employment rates for women obtained in this survey, irrespective of whether self or proxy reported, are well above what is found in other

Table 7.1: Detailed weekly questions on employment asked to two randomly chosen individuals from a household

Employment type	Asked to respondent about themselves (self-reported)	Asked to respondent about their spouse (proxy-reported)
Self employment	Last week, did you do any kind of business, farming or other self-employed activity to generate income, even if only for one hour?	Last week, did your spouse do any kind of business, farming or other self-employed activity to generate income, even if only for one hour?
Contributing family worker	Last week, did you assist without pay in a business/farm/livestock of a household or family member even if only for one hour?	Last week, did your spouse assist without pay in a business/farm/livestock of a household or family member even if only for one hour?
Wage work (casual or salaried)	In the last week, did you work for a wage, salary, commission or any payment in kind, including doing paid domestic work, even if only for one hour?	In the last week, did your spouse work for a wage, salary, commission or any payment in kind, including doing paid domestic work, even if only for one hour?
Apprentice/intern	In the last week, did you work for pay as an apprentice, intern or trainee even if only for one hour?	In the last week, did your spouse work for pay as an apprentice, intern or trainee even if only for one hour?
Small-scale production of goods/services for sale	Last week, did you engage in small scale production of goods or services at home that were exchanged for cash or kind even if only for one hour?	Last week, did your spouse engage in small scale production of goods or services at home that were exchanged for cash or kind even if only for one hour?
Unpaid volunteer	Last week, did you work as an unpaid volunteer or do any kind of unpaid social work even if only for one hour?	Last week, did your spouse work as an unpaid volunteer or do any kind of unpaid social work even if only for one hour?
Recovery question	Did you miss out reporting any work activities that led to you earning an income, or helping household members with an activity that generates an income even if only for one hour?	Did you miss out reporting any work activities that led to your spouse earning an income, or helping household members with an activity that generates an income even if only for one hour?

Sources and notes - India Working Survey questionnaire

Table 7.2: Difference in self and proxy reported labor market outcomes for spousal pairs

	Women			Men		
	Self	Proxy	Difference between proxy and self	Self	Proxy	Difference between proxy and self
Laborforce participation rate	69.5	63.8	-5.7***	79.7	81.5	1.8
Workforce participation rate	63.2	57.9	-5.4***	76.9	78.7	1.8
Unemployment rate	9	6	-3	3.5	2.8	-0.7

*Sources and notes - India Working Survey 2020, *** - p<0.01. See Abraham et al (2023) for details.*

surveys.[4] Possible reasons include probing questions (discussed in Section 7.3.3), training and sensitisation of enumerators, and matching enumerator gender with respondent gender. Coming to the self-proxy results, while 70 percent of women report themselves as being part of the labour force, only 64 percent of women were reported to be a part of the labour force by their husbands, indicating a significant discrepancy (Table 7.2). In contrast, no statistically significant differences exist between self-reported and proxy-reported labour market outcomes for men.

Since the IWS questionnaire asks for self and proxy reports on each kind of employment, we are able to estimate if, and to what extent, specific kinds of employment are susceptible to over/under-reporting by proxies. An individual may be self-employed, or contributing to the family farm or business or in casual or salaried wage work.[5] The data presented in Table 7.3 shows the distribution of activities for husbands and wives based on self and proxy reports. We find some significant differences in both men's and women's employment type when reported by themselves versus their proxies. This is particularly the case for agricultural employment.

Coming to women first, according to self-reports, approximately 14 percent of women identify as own-account workers in agriculture. However, when their husbands report on their behalf, this proportion increases to 24 percent. Conversely, women are more likely to report themselves as contributing family workers. While 52 percent of women self-report as contributing workers in agriculture, only 46 percent are classified as such by their husbands. This suggests that whereas women report themselves as contributing workers, their husbands are likely to report them as self-employed. Self-employment involves earning direct income, whereas contributing work entails contributing to the household farm

or enterprise without receiving direct payment. It is possible that women acknowledge the lack of payment or reimbursement associated with their work and classify themselves as contributing workers, whereas men may not perceive it the same way and categorise women as self-employed. Finally, husbands also under-report casual wage work in agriculture performed by women.

It is possible that women acknowledge the lack of payment associated with their work and classify themselves as contributing workers, whereas men categorise women as self-employed.

Although Table 7.2 showed that there were no significant differences in self and proxy reports of whether men are employed or not, we see that there are significant variations in the types of activities identified by self and proxy reports (Table 7.3). In general women under-report men's self-employment and over-report wage work. About 38 percent of men reported themselves as own-account workers in agriculture, whereas only 31 percent of wives report their husbands as so. The differences in reporting are most pronounced in the case of wage work, particularly in casual employment.

Intriguingly, even the PLFS data shows disagreements of this kind if we match husband-wife pairs using the roster and examine the differences in the employment type distribution for men and women depending on who is reporting the information (i.e. whether the husband or the wife is the informant). That is, the share of men involved in, say salaried work, may vary by as much as 5 percentage points depending on whether the informant is male or female (data not shown). In Section 7.4 we discuss the implications of these findings for conducting surveys as well as using survey data.

Box 7.3: Comparing husband's and wife's perceptions women's decision making role

Kasturi Kandalam

Besides self and proxy reporting of women's employment, the IWS survey instrument also sought to understand to what extent reporting of women's participation in various intra-household decision making processes differed between husbands and spouses. Three dimensions of decision making were examined - participation of women in the labour force, purchase of expensive household items and mobility/travel outside the home. These were asked to women to gauge their perception regarding their participation in such decisions. If the respondent pair were

Table 1 – Self and proxy decision-making questions from the IWS

Women	Men
Who in your family are involved in the following decisions? (Select all that apply)	
Participation of women in the labor force	
Whether you should work outside the home or not	Whether women in the family should work outside the household
Purchasing	
Whether to buy any expensive household item	Whether to buy any expensive household item
Whether to buy immoveable property (land or house)	Whether to buy immoveable property (land or house)
Options Provided: You, Your Husband, Father/ Father- in-law, Mother/ Mother- in-law, Brother/ Brother- in-law, Son, Other	Options Provided: You, Wife, Father/Father-in- law, Mother/Mother-in-law, Son, Other
Mobility	
Please tell us whether you have to ASK PERMISSION of your husband (if married) or a senior family member to go –	Please tell us whether your wife has to ASK YOUR PERMISSION or ask a senior family member to go -
To the local health centre/doctor	To the local health centre/doctor
To the home of friends or relatives in the neighbourhood	To the home of friends or relatives in the neighbourhood
To the Kirana shop	To the Kirana shop
To a short distance by train or bus	To a short distance by train or bus
Options Provided: Yes –must ask for permission: Your Husband, Father/ Father-in-law, Mother/Mother-in-law, Brother/ Brother-in-law, Son, Other	Options Provided: Yes –must ask for permission: You, Father/Father- in-law, Mother/Mother- in- law, Brother/Brother- in- law, Son, Other
No – need not ask for permission	No – need not ask for permission

Sources and notes -Decision-making module of the India Working Survey (2020)

Box Figure 1: Comparing husband-wife responses on perceptions of decision making - Employment and purchase

N=727 couples in Employment Dimension & N=1180 in Purchasing Dimension

- Wife Reporting - Participating in labor force
- Husband Reporting - Participating in labor force
- Wife Reporting - Purchasing household items
- Husband Reporting - Purchasing household items
- Wife Reporting - Purchasing immoveable property
- Husband Reporting - Purchasing immoveable property

Sources and notes: India Working Survey (2020)

spouses, then husbands were also asked similar questions to understand how husbands perceived their wife's involvement in decision making. Table 1 details the questions asked and potential

Box Figure 2: Comparing husband-wife responses on perceptions of decision making - Mobility

N=1132 couples in Mobility Dimension

- Wife Reporting - Visiting local Health Centre
- Husband Reporting - Visiting local Health Centre
- Wife Reporting - Visiting Neighbourhood
- Husband Reporting - Visiting Neighbourhood
- Wife Reporting - Going to a Kirana shop
- Husband Reporting - Going to a Kirana shop
- Wife Reporting - Travelling short distance
- Husband Reporting - Travelling short distance
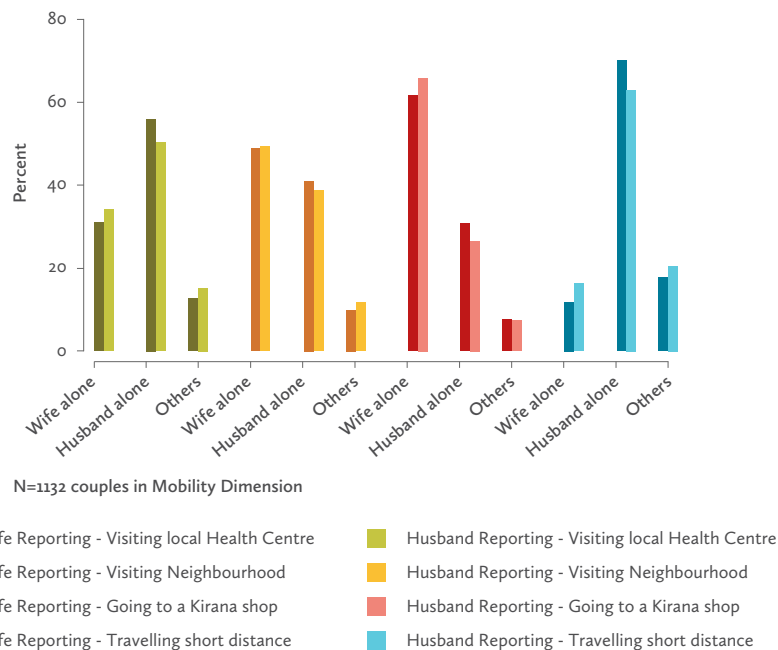
Sources and notes: India Working Survey (2020)

response options. These are modified from questions posed in the National Family Health Survey (NFHS).

For participation in the labour force, 12 percent of women reported that they alone made the decision. In contrast, only 3 percent of men reported their wife having a sole decision-making role. Rather, husbands were more likely to report making the decision jointly with their spouses. Notably, 65 percent of men reported that their wife was not involved in any capacity (i.e. either they or someone else made the decision) compared to 50 percent of women reporting the same (Figure 1).

 In the purchasing dimension, unlike participation in the labour force, women were less likely to report themselves as sole decision-makers. With regard to purchasing expensive household items, there was high convergence between spouses

(Figure 2). In purchasing immovable property, husbands and wives disagree as well. 33 percent of women reported that their husband alone made the decision, while 27 percent of men reported themselves as sole decision makers.

The extent and nature of divergence varies by the dimension of decision making being asked about. In most surveys, questions on decision making are fielded to women themselves. Therefore, the problem of proxy misreporting does not arise. However, what these results indicate is that even when women are reporting participation in decisions, their perceptions may not often match with that perceived by their spouses. These divergences in self-proxy reporting particularly in contentious aspects - employment and travel outside the home, imply that self-reported perceptions of autonomy may differ from lived experiences as well as perceptions of other members of the household.

7.3.3 Designing questions to get better estimates of women's employment

The survey instrument including the framing of questions and their sequencing can affect the quality of data collected (Moya 2021; Biemer and Lyberg 2003). The framing of questions can have significant impacts on the measurement of employment particularly in economies dominated by agricultural employment and where gendered social norms around what men or women ought to be doing prevail (Benes and Walsh 2018; Comblon and Robilliard 2015). For example, Muller and Sousa (2020) show that rural female labour force participation in Honduras is under-reported due to women primarily recognising themselves as "housewives" although they are engaged in

economic activities besides household work. They also find that women are more stringent in applying the definition of employment, characterising it as work done in exchange of money, performed outside home etc., while men did not identify such constraints in defining work.

Bardasi and Beegle (2011) estimate differences in labour statistics comparing between a short module questionnaire versus a longer module in the context of Tanzania. The short module consists of one question on work - "Did you do any type of work in the last 7 days?". The detailed module consists of screening questions, specifying three main groups of economic activity. Comparing employment rates across these two instruments they find that female employment is about five percentage points lower in the short module compared to the detailed

Table 7.3: Differences in self and proxy reports of employment activity distribution

	Women			Men		
	Self	Proxy	Difference between proxy and self	Self	Proxy	Difference between proxy and self
Own account worker agriculture	13.7	23.8	10.1[***]	37.6	31.3	-6.3[*]
Own account worker non-agriculture	6.1	7.8	1.7	12.4	11.1	-1.3
Employer agriculture	0.7	1.2	0.5	5.2	6.6	1.4[***]
Employer non-agriculture	0	0.5	0.5	1.6	1.7	0.1
Contributing worker agriculture	52.3	46	-6.3[***]	17.2	11.8	-5.4[*]
Contributing worker non-agriculture	2.4	2.6	0.2	0.3	1.1	0.8
Casual agriculture	15.8	9.8	-6.0[***]	7	9.2	2.2[**]
Casual non-agriculture	4.3	3.7	-0.7	9.5	14.7	5.2[*]
Salaried	4.7	4.7	0	9.2	12.6	3.4
N	717	656		872	892	

Sources and notes - India Working Survey 2020, [***] - p<0.01, see Abraham et al (2023) for details.

module. Benes and Walsh (2018) also emphasise the importance of asking "recovery questions" in capturing the work of unpaid family helpers who are otherwise reported as unemployed.

In the Indian context, Anker (1983) shows that calling out a "list of activities" as against the keyword questions (questions that contain a typically recognisable keyword about overall employment such as "main activity", "secondary activity", "pay or profit", rather than question(s) about specific activities) in employment surveys leads to higher estimates of women's work. The "list of activity" approach has been found to provide higher estimates of women's employment by capturing women's part-time, home-based work in studies in other low and middle income contexts too (Langsten and Salem 2008; Anker 1983).

Another more recent study by Deshmukh et al. (2019) in the Indian context reiterates these findings. Based on questionnaires administered to respondents in selected districts of Delhi National Capital Region, the study finds that asking about primary and secondary activities (analogous to the keyword questions identified in Anker (1983)) results in a larger share of women being listed as home-makers. On the other hand, follow-up questions on the major sources of household income and who contributed to this income resulted in much higher estimates of women's employment participation with a large share of this increase coming from the reporting of women's work in caring for livestock.

In the main IWS survey, as highlighted in Table 7.1, questions on employment activities in the last week were framed in such a way that three broad kinds of employment - wage, self-employment or

contributing family work, were explicitly called out. Besides the main survey, the IWS also had two experiment sub-samples. In each of these sub-samples, two different modules were administered to enquire about employment status - a short weekly module and a short daily module.

The short weekly and short daily modules resemble those used in the NSSO surveys. In the short weekly module, respondents were asked a single question: "In the last week, what were the activities you were doing, even if only for an hour?" They were allowed to report multiple activities, but enumerators did not provide a specific list of potential activities. In the short daily module, the framing question was the same, but the reference period was the previous day. The same question was asked for each

of the seven preceding days, resulting in a total of seven questions. The weekly labour force status is calculated from these questions asked about each day. If they are employed even a single hour or any one day in the previous week, they are deemed as employed according to weekly status.

The detailed weekly module differs from the shorter modules in two ways. First, there is a specific question for each potential employment activity asking respondents if they engaged in any of them in the past week. These employment activities encompass self-employment activities, assistance in family farms or businesses (unpaid), wage or salaried work, paid apprenticeships or internships, and small-scale production of goods or services for sale. Five questions, pertaining to each one of the activities,

Table 7.4: Detailed weekly, short weekly and short daily questions

Detailed weekly questions (main survey)	Employment Questions
Self-employment	Last week, did you do any kind of business, farming, or other self-employed activity to generate income even if only for one hour?
Contributing family worker	Last week, did you assist in a business/farm/livestock of a household or family member without pay even only for one hour?
Wage work (casual/salaried)	In the last week, did you work for a wage, salary, commission, or any payment in kind even if only for one hour?
Apprentice/intern	In the last week, did you work for pay as an apprentice, intern or trainee even if only for one hour?
Small-scale production	Last week, did you engage in the small-scale production of goods or services at home that were exchanged for cash or kind even if only for one hour?
Unpaid volunteer	Last week, did you work as an unpaid volunteer or do any kind of unpaid social work even if only for one hour?
Short weekly question (experiment)	In the last week, what were the activities you were doing for even if only for an hour?
Short daily question (experiment)	Think about the <n-th> day prior to today. What were the activities you were doing on this day even if this is for an hour or less? (asked about each 7 days of the week)

Sources and notes -India Working Survey 2020.

Figure 7.2: Employment estimates from main and experiment surveys
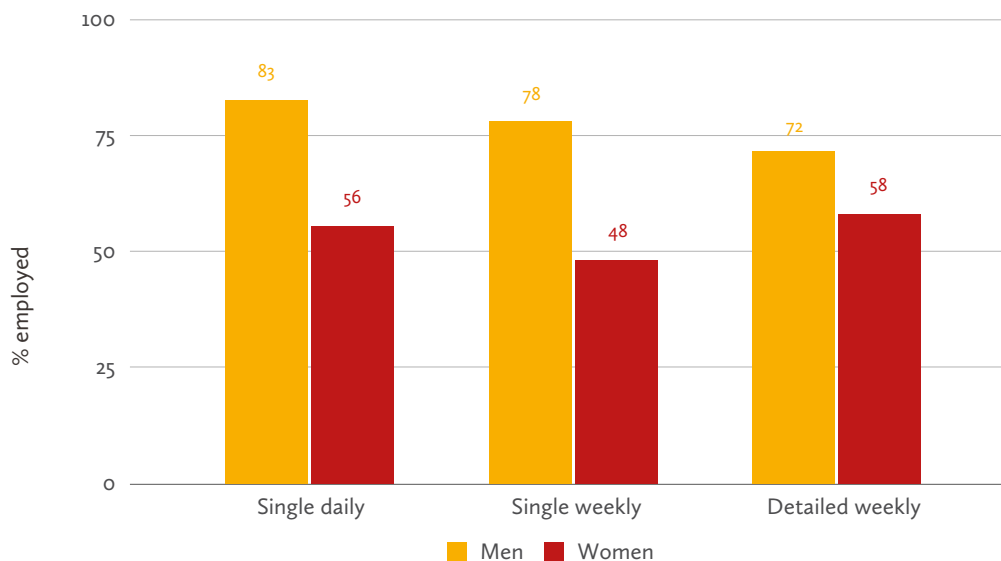
Sources and notes: India Working Survey (2020).

were asked, with respondents indicating "yes" or "no" for each activity. If the response was "yes", further details regarding hours worked, industry/sector of work, and income were collected.

This module also included a recovery question that enquired if the respondent missed out on reporting any other income-generating activity that they were involved in. The questions in the detailed module adhere to the recommendations of the International Labor Organization (ILO) for measuring key labour market indicators developed based on extensive pilot studies conducted in ten countries (Benes and Walsh 2018). Moreover, these questions are similar to those used in the World Bank's Living Standards Measurement Study (LSMS) surveys. For a full list of questions asked in the different modules of the survey, refer to Table 7.4.

The short modules (weekly and daily) were administered only in Karnataka and to randomly selected sets of households based on a household listing exercise. Every household was randomly assigned either the single weekly module or the single daily module. Like in the main survey, an

adult male and an adult female were selected as respondents in each household and both administered the same module. 260 and 252 rural households were administered the short weekly and short daily module, respectively.[6]

These two modules in accompaniment with the detailed weekly module in the main survey tells us to what extent the framing of the question (detailed versus single) and the reference period (daily or weekly) affects estimates of employment. Note that the appropriate comparisons are (i) between estimates from the short weekly and the detailed weekly - which reveals how keeping reference period unchanged, moving from a single question to multiple questions affects estimates, and (ii) between estimates from the short weekly and short daily - which reveals how changing the reference period from weekly to daily while keeping the question framing intact affects estimates.

Comparing between the single weekly and detailed weekly reveals that asking detailed questions about different kinds of employment rather than a single question increased estimates of women's

employment by 10 percentage points while it *decreased* men's employment rate by 6 percentage points. Both differences are statistically significant. Moving from asking about the week to asking about each day of the week (single weekly versus single daily) increases women's employment estimates by about seven percentage points and men's by four percentage points but these differences are not statistically significant.[7]

Asking detailed questions about different kinds of employment rather than a single question increased estimates of women's employment by 10 percentage points.

Finally, we investigate what kinds of activities are better captured when we move to a more detailed framing. We find that higher employment estimates from the detailed module come from higher estimates of women in contributing family work which increases overall employment estimates. This segment of work tends to be invisible, particularly in agricultural households and is often not cognised by women themselves (Koolwal 2021; Muller and Sousa 2020; Deshpande and Kabeer 2019). This is similar to results from Deshmukh et al (2019) who find that livestock care (largely done by women as contributing family workers) is what drives up higher estimates of employment when question framing is changed.

7.4 Social norms, measurement of employment and implications for survey data

We saw in Section 7.3.2 that proxy-reporting can give significantly different results from self-reports. And Section 7.3.3 showed that even when respondents are reporting about themselves, how questions are asked can impact the estimates derived. We now discuss the possible reasons (based on Ambler et al 2021) and broader implications of these findings.

Two types of error are relevant here. Asymmetric measurement error and asymmetric information error. Asymmetric measurement errors may stem from two sources - differences in the understanding of what constitutes employment and gender norms-based differences in identifying employment. First, the perceptions of employment for men and women may differ, even though there is complete information on activities performed by either within the household. For instance, men may consider employment to involve only paid labour, while women may also include work on the family farm which is often not explicitly remunerated. Second, women themselves may not identify as "employed" but rather conform to their socially prescribed role of a "home-maker". Indeed, Basole (2016) found among home-based embroidery workers in Banaras that women thought of their paid work as being conducted in "spare time" even though time-use surveys revealed that they were engaged in such work for nearly 6 hours of the day. In such cases, further probing that breaks down these socially prescribed roles allows for better convergence between the actual employment activities and reported employment.

It is important to note that, even if men and women do not conform to socially prescribed roles themselves, reporting may still occur in line with social norms. For example, the reporting of women's employment by men may be driven by perceived social censure associated with a certain activity (Jayachandran 2020). Bernhardt et al. (2018) report how second-order beliefs - beliefs of husbands about what they perceive are the social costs of their wives working, are a major constraint to women's participation in the workforce. Therefore, even if they do recognise women as working, it is possible that men may not report them as such.

However, despite the difficulties in measuring women's employment, it should be noted that above factors are unlikely to explain the long run trends in declining female workforce participation. While conceptual and operational issues can affect the levels of the estimates, they cannot be used to explain the trend, given that neither concepts or ways of asking employment questions have changed.

Asymmetric information error occurs due to differences in the information available to men and women about each other's activities. This may arise as spouses strategically hide information about their activities from each other to keep income private or avoid censure as they are going against social norms. The information asymmetry could also be unintentional due to issues of observations (spouse could be employed in different locations - farm, house etc) or gendered domains of responsibility. In cases where the work is carried out intermittently, and alongside care or other domestic work, it is likely that the employment activity is invisibilised or not recognised. This results in proxy reports differing from self reports.

In the IWS, we collected information relating to the respondents' demographic profile as well as information on their household structure. Many of these variables can proxy for the possibility of these two different errors. To capture asymmetric measurement error and the role of norms, we include (i) a variable that captures whether the husband reported his mother as employed, (ii) the difference in the number of hours spent by the husband and the wife in household work including cooking, cleaning, taking care of children and elderly, and fetching water. For asymmetric information error we include variables that capture the number of hours engaged in that activity in a week and the major activity in the year for the person being reported about. We find that self employment and contributing family work are susceptible to asymmetric information and asymmetric measurement error, while the

differences in reporting of wage work comes largely as a result of asymmetric information error.

The foregoing analysis reinforces our understanding that women's employment is difficult to measure accurately. Both, what questions are asked and to whom they are asked, matter a great deal. What can be done?

For one, surveys should not rely on one single question.[8] The survey instrument used in IWS which calls out employment types (and are in-built probing questions) addresses this concern without having to rely on the enumerators' discretion of asking probing questions.

Second, as we discussed earlier, the self-proxy problem is not limited only to women's employment. We get different answers for men's work too depending on who is asked. Hence, ideally, information related to employment, earnings and other aspects of the labour market should be gathered from the informant themselves rather than relying on proxies. This is not possible in the way surveys such as the PLFS are currently conducted. These are household surveys where information on everyone in the household is collected from the survey informant. Thus, one gets a large sample of individuals while visiting only a fraction of households. It is more expensive and time consuming to conduct individual surveys. This is not only because more households will need to be visited to get the same sample of individuals, but also because the randomisation now has to be carried out at the individual level and repeat visits may be required to a household to interview the selected person.

The self-proxy results also provide us useful information on how best to use household survey data. Trends in employment overall as well as employment type can safely be interpreted, at least over a few years, since there is no reason to expect the errors giving rise to self-proxy divergence to

change quickly over time. Self or proxy reporting also does not change the rank ordering of employment types in terms of their shares. For example, for men, own-account work remains the most frequently observed type of work, while for women it is contributing family work. Salaried or regular wage work remains the smallest for men with casual wage work in between while for women both are at par. However, one should be cautious in drawing large conclusions based on small differences in levels. For example, if we see that two states differ in terms of their regular wage employment share by 5 percentage points, then we know that we have to research the source of these differences and rule out explanations that are related to survey design.

Methodological innovations in sampling, survey design, and questionnaire development can be piloted on a regular basis to ensure data collection processes are continuously improving. It is important for us to go this route, rather than dismiss survey data as unreliable because sample surveys remain our best source of unbiased information.

Improvements in survey design and enumerator training can go a long way in addressing the foregoing issues. Methodological innovations in sampling, survey design, and questionnaire development can be piloted on a regular basis to ensure data collection processes are continuously improving. It is important for us to go this route, rather than dismiss survey data as unreliable because sample surveys remain our best source of unbiased information. Administrative and commercial data, as discussed in the beginning of this chapter, suffer from strong selection issues which prevent us from generalising on their basis.

7.5 Improving the quality of phone surveys[9]

The last few years have seen an upheaval in practices of data collection and survey methods. Even before the pandemic, several data collection endeavours had begun the transition to digital, computer-assisted and tablet-based surveys. NSO surveys have also moved away from traditional paper-based to computer-assisted PI techniques (CAPI). During the Covid-19 pandemic ongoing surveys had to be prematurely terminated or put on hold in the interest of the safety of enumerators and interviewers. However, as the pandemic progressed, data on the impacts of the pandemic at the economic, social and health levels became crucial for researchers, policymakers and the general public.

As a result a number of phone and web-based surveys were conducted by independent agencies and national research organisations to address the urgent need for recent data on the impact of the economic lockdown across various dimensions.[10] Since the pandemic, the urban revisits for the PLFS are also being conducted via phone.[11]

7.5.1 Challenges, new and old

The transition to a phone survey raises several issues including access to respondents, ensuring proper administering of consent, monitoring of enumerators and data quality, ensuring the privacy of interviews and respondent safety when administering sensitive questions. These concerns may be broadly classified into those related to limitations in the reach and coverage of phone surveys and those related to the quality of data emerging from phone-based surveys. The former, we refer to as issues related to sampling and the latter issues related to reporting.

The most immediate sampling concern in the context of a phone survey is the coverage and reach of such surveys. The available sample is automatically restricted to those individuals who have access to a phone. This in itself introduces an implicit bias into the sample. However, even with access to a phone, there may be further constraints in participation. For many, an active connection

may not be available despite owning a phone. Participating in a phone call with a near stranger for an extended period may not be an option for some owing to a number of restrictions including social norms that frown on conversations with outsiders (particularly for women), the burden of household responsibilities and other time constraints. This is particularly the case when the survey has a particular intended respondent (say, female respondent) rather than any individual who answers the phone.

The nature of phone surveys can also raise certain reporting issues since enumerators are not able to make actual observations (as in field surveys) or engage in a more detailed dialogue to verify the certain response. Phone surveys are also typically shorter and do not often allow for free flow of conversation which may further restrict enumerators from asking more probing questions, including questions to aid recall or recovery type questions. Finally, privacy, something which is often difficult even in field-based surveys particularly in developing countries may not be possible and may not even be verifiable when doing a phone survey.

Both sampling and reporting errors are likely to disproportionately affect certain populations. Women, older people and those from marginalised communities are more likely to not have access to a functional phone connection. The digital gender divide has been well documented in India and elsewhere (GSMA 2019). Alvi *et al.* (2020) find that women were less likely to own a phone and even if they did, were less likely than men to have a working recharged connection. Although the gender gap in women's access to mobiles has narrowed globally in recent years, in India, there is still substantial variation in the ownership of mobiles between men and women, with only 59 percent women owning mobile phones compared to 80 percent of men. Besides the issue of comparatively limited access to women over phone, there are also issues of

privacy particularly when asking sensitive questions, being able to provide own responses to questions (rather than being prompted by someone) as well as availability of time for women to participate in such surveys (Alvi et al. 2020; Gupta and Gupta 2020; Mathur 2020).

A detailed protocol can go a long way in reducing sampling and reporting errors. A protocol that envisages the possible scenarios that can arise when households are contacted via phone and account for these in enumerator responses minimises confusion and streamlines processes. It can also increase response rates. Another tool for monitoring and assessing survey quality is to incorporate post-survey interviewer observations. These can provide some insight into the nature of the interview and whether there are any systematic differences across certain groups.

7.5.2 The IWS phone survey

With the India Working Survey field operations being prematurely terminated in March 2020 due to the Covid-19 pandemic, in August 2020 we transitioned to a phone survey. The aim was to understand the impact of the pandemic on already interviewed respondents. The design of the IWS phone survey protocol aimed to minimise sampling and reporting issues discussed in the previous section. We now discuss our experience of conducting this survey and the lessons learned.

The IWS field survey interviewed approximately 5,951 individuals and 3,646 households over the course of two months. To the 5,171 individuals who agreed to share their number, we also asked their permission to call back on the number provided in case of any follow up questions or clarifications. Almost all respondents consented to a call back, giving us a final sample of 5,117 individuals. Some of the numbers provided were incomplete or non

functional leaving us with a final sample of 4515 individuals.

The workflow (see Abraham and Mohan 2023 for a schematic) begins with the supervisor assigning respondents to enumerators for the day on the previous night. Following this, an automatic SMS is sent to these respondents informing them about the survey. The enumerator then calls the respondent and fills out the questionnaire (hereafter referred to as "forms") depending on the response and in the end, marks the form as "closed".

To begin with, we paid particular attention to the allocation of respondents to be surveyed. Rather than giving the enumerators a complete list of respondents they need to call at once, we assign the supervisors to allocate a fixed number of respondents they should call each day. During the allocation, we also ensured that only female respondents are assigned to female enumerators and male respondents to male enumerators.

In contrast to face-to-face surveys, telephone surveys have an additional hurdle of getting the phone call connected. Sometimes respondents are missed because they are temporarily unreachable because of network issues. To minimise such attrition, we also included such connection related scenarios in the protocol. A way in which it is possible to maximise reach is through reschedules. To ensure that we do not lose a potential respondent because their phone was unreachable, the survey is automatically rescheduled to four hours after the first call. If this automatically calculated time-slot does not work for the enumerator, they are given an option to reschedule the survey to a date and time of their convenience.

We also allowed the enumerators or respondents themselves to reschedule a survey up to 6 times. To ensure that the enumerators remembered to call the

respondent, we sent an automatic SMS reminder to the enumerators one hour before the rescheduled time. This made it easier for the enumerator to keep track of the rescheduled calls easily. A separate SMS reminder is also sent to the respondent to ensure their availability.

To help survey monitoring, there was a provision to add a supervisor to the call during a survey, and time-stamps and questions for the enumerators at the end of the survey. In addition, specific information from each call that helps monitoring was published into a spreadsheet automatically and in real-time and shared with both the enumerators and supervisors.

Of the 4,515 individual phone numbers we had, we were able to reach nearly 75 percent. We did not see any difference between men and women in this aspect. In terms of response rate (where response rate is defined as the total successfully completed interviews as a share of all numbers available), within the sample of all numbers available, we were able to complete interviews with about 66 percent of the respondents. There is a marginal difference between men and women, with response rates higher among men, and a higher share of women refusing,(9.7 percent compared to 8.2 percent).

We found that even with having a phone number to reach a woman at, there are still difficulties in having direct access. Indeed this is apparent when we look at how many times the respondent themselves answered the phone when our enumerators called.

For male respondents, 71% of the times we called, our intended respondent, the man, themselves answered the phone. For women, there was only a 40% chance that it was the woman themselves who answered the phone.

For male respondents, 71 percent of the times we called, our intended respondent, the man, themselves answered the phone. However, when we contacted women on the phone number they had provided, there was only a 40 percent chance that it was the woman themselves who answered the phone. The phone survey protocol anticipated this and provided the enumerators with the appropriate responses in this scenario as described in the earlier section. In the event that the main respondent themselves did not answer the phone, we requested to speak with the main respondent, or if they were unavailable, rescheduled the interview. Interestingly, even though women did not answer the phone call themselves, when requested to come to the phone, a relatively high share of women were available to speak. Therefore, although women were less likely to answer the phone, ultimately, there were no major differences in reaching women over the phone as compared to men.

As detailed in the protocol, up to six attempts would be made to a given phone number. Any call made by the enumerator to the assigned number is counted as a call attempt, whether or not the call is connected. The only reason a call would be closed with less than six attempts being made is if the respondent or the person who answered the phone refused to participate or if the intended respondent was no longer available.

A successful call is one where the respondent agrees to participate and the interview is completed. Table 7.5 shows the distribution of completed and successful calls by the call attempt at which they were completed. Among men, nearly half of calls (44 percent) were likely to be completed in the first attempt. For women, on the other hand, the first attempt yields comparatively lower response rates, with only 30 percent of calls being completed. Half of the women's completed calls occur only by the second attempt. Notably, for women, at every successive call attempt, there is a relatively higher share of completed calls compared to men. By the fifth attempt, up to 93 percent of the completed calls had occurred. The similar number for men occurred earlier at the fourth attempt. So for women, the initial low response coming at the first attempt is made up for by subsequent call attempts to the same number. In fact, a large share of completed interviews occur at the sixth and final attempt. This highlights the need for a strict protocol and multiple re-attempts being made to the respective number so as to secure a response. This can ensure higher response rates particularly among women.

Finally, we examine the response rate by time of call. While this was not used in our protocol, the information from this paradata can be useful for planning future surveys and protocol. For men and

Table 7.5: Higher number of call attempts can help reach women respondents better

Call attempt at which complete interview made	Men (%)	Women (%)	Overall (%)
1	44.1	29.5	36.05
2	26.9	28.89	27.99
3	13.3	17.38	15.55
4	7.74	11.08	9.58
5	4.58	6.79	5.8
6	3.38	6.36	5.02
Total	100	100	100

Sources and notes - IWS Phone Survey 2020.

women response rates are highest in the latter half of the day. For men, there is also a time in the early morning where response rates are high, as well as some time mid-day when there is small increase in response rates. For women, response rates are highest after 6 PM, and there is no other intermediate time during which their response rates are high. This likely points towards the lesser availability of women owing to work within and outside the home at all times of the day.

At the end of a completed survey, we asked our enumerators to make a series of observations on the nature of participation of the respondent. Specifically, we asked them the following five questions:

(i) Were the responses given by someone else on behalf of the respondent ? (proxy)

(ii) Did the respondent put the call on speaker phone while the interview was going on? (speaker)

(iii) Did you hear any other individuals helping the respondent in understanding the questions? (helped)

(iv) Did you hear any other individuals prompting the respondent in their responses? (prompting)

(v) Did he/she seem in a hurry to finish the interview? (hurry)

(vi) Did they seem interested and willing to provide answers? (interested)

We find that in the case of female respondents, responses were more likely to be given by proxy, 13 percent compared to 11 percent of men. However, in general, there are only a few instances where proxies are reporting on behalf of the main respondent. Nearly a quarter of female interviews had the phone placed on speaker, much smaller than the 18 percent observed among male interviews. In the case of help in understanding questions as well as enumerators observing a second person prompting on responses, there is substantial difference between male and female interviews. Women respondents were about three times more likely to receive help or be prompted in responses. Women respondents

seemed more hurried in responding to questions perhaps as a result of the multiple demands on their time at home/work. There is not much difference between men and women in their perceived interest in participating in the survey.

The IWS survey did not ask any obviously sensitive questions to the female respondents (domestic violence, autonomy at home etc). However, even then, we see that female respondents are likely to not have privacy when participating in the phone survey. Therefore, phone surveys that seek to ask more sensitive questions that may put women at risk must heed extreme caution.

Meta and para data from surveys - data about the process of surveying - are useful inputs to judge survey quality, monitor progress and provide feedback to enumerators. Given that this data is easily available in CAPI / CATI based surveying, it is mostly under utlised in most surveys. Goel et al (2022) demonstrate, using IWS data, how paradata can be used to monitor survey quality (moving beyond the standard parameters of interview time or number of interviews conducted per day) and provide tailored enumerator feedback that can improve enumerator performance in subsequent surveys.

7.6 Conclusion

Reliable, timely, and representative estimates of key labour market indicators such as the employment and unemployment rates, sectoral shares, earnings and hours worked are indispensable for effective policy-making. Moreover, these estimates are needed at various levels of disaggregation such as regions (e.g. state and district) as well as by key identities. Under current conditions, sample surveys are the most effective tool for doing so. Combining them with administrative and other data sources,

a complete picture of the labour market can be constructed. This chapter has discussed several ways in which survey estimates can be improved. This is necessarily a partial discussion. Several key issues such as resource constraints, level of granularity (e.g. district-level estimates) etc have not been covered. Nevertheless, we hope that this discussion and the results reported therein can inform the ongoing debate on how India's statistical architecture can be strengthened.

Endnotes

1 See https://theprint.in/opinion/indias-scheduled-caste-list-must-be-religion-neutral-muslims-christians-are-also-dalit/1734311/ for a recent article on this issue.

2 The discussion in this section draws from Abraham, Anjum, Lahoti, Swaminathan (2023).

3 Among the 5,951 respondents in the survey, 3,750 (63 percent) were spousal pairs. However, due to non-participation by one or both respondents and data recording issues, the final sample for spousal pairs consists of 2,674 observations, including 1,337 husbands and 1,337 wives. Since the IWS was prematurely terminated owing to the pandemic, the sample is predominantly rural and not representative of either state. Unless otherwise stated, all estimates pertain to rural areas.

4 However, due to the interruption of the IWS during Covid-19, the samples are not representative at the state level. Hence the rates cannot be directly compared to PLFS estimates for the two states.

5 The questions allowed for an individual to report in more than one activity in the past week. To determine the main activity of each individual, we employ an hour-based criterion, considering the activity in which the highest amount of time is spent during the week as the individual's primary activity.

6 The survey instrument for these experiments, like the main survey, collected details on household demographics, asset ownership. However, sections on experiences of discrimination, life history calendar, time spent on household production were not included. In total, 299 and 300 individuals responded to the single weekly and single daily modules respectively.

7 Given the difference in size and composition of each sample, we match the sample size between the main and experiment surveys by randomly drawing a smaller subset of individuals from the main survey equivalent in size to the experiment. This is pooled with the experiment data and we estimate bootstrapped coefficients of difference in estimates between the main and experiment surveys. We find that women's employment estimate increases significantly when we move from single weekly to detailed weekly while for men there is no significant change. If we control for differences in the composition of the samples, women's employment estimates from detailed weekly continue to be significantly higher, while for men, it is significantly lower. Therefore, asking detailed questions to elude employment information results in a lower share of men being reported as employed.

8 Notably, the Population Census 2011 recognises this problem and in its manual states that 'Women and children may often be classified as Non-workers because of non-reporting of their work. It also happens that women and children who work for six months or more are sometimes reported as working for less than six months. You should therefore make special efforts for listing women's and children's work by asking probing questions'.

9 This section draws on Abraham and Mohan (2023)

10 In India, multiple phone surveys emerged collecting crucial information on various aspects including the impact of the pandemic on employment, health, awareness about Covid-19, access to government schemes, the welfare of migrant workers etc. (NCAER 2021; RCRC 2021; Action Aid 2020; Dalberg 2020).

11 https://pib.gov.in/PressReleaseIframePage.aspx?PRID=1928124

8

Conclusion

8 Conclusion

In the foregoing chapters we took a detailed look at the relationships between economic growth, structural change, and labour market outcomes for the key social identities of caste, gender and religion. The analysis provides reasons for hope and satisfaction, but also raises concerns over persistent, and in some cases, growing, disparities. In this concluding chapter, we return to the key points revealed by the analysis and also discuss limitations of the study.

In Chapter Two we saw that the recent rise in female workforce participation rate after years of decline is a result of a shift from domestic duties to self-employment. Coming on the heels of the growth slowdown and the pandemic, this increase in self-employment appears to be distress-led. As a result, earnings from self-employment have declined in real terms since 2019. This points to the danger of taking an increase in the overall LFPR or WPR as an indicator of improvement in employment conditions. The increased supply of female labour has to be matched with an increased demand for it in modern, productive activities. Otherwise, it only entails crowding into an already crowded self-employment sector.

Creating labour demand (i.e. job opportunities) for women, is thus a key policy priority. As Mehrotra (2019) points out, a standard set of macroeconomic policies that aim to promote structural change are part of the solution when it comes to addressing missing women from the workforce. While taking the concerns around "premature deindustrialization" seriously, it is necessary to imagine a National Employment Policy framework that directly and indirectly addresses gender imbalances. We presented one such framework in the State of Working India 2021. But note that generating job opportunities for female labour need not only mean actual job creation (though it does mean that too). Rather, policies that enable women to travel safely over longer distances to avail of job opportunities and policies that make it possible for them to be relieved of care-work responsibilities in order to leave the home are equally important (Mehrotra and Sinha 2019)

Chapter Three showed that India has performed more poorly than the average developing country when it comes to linking growth to jobs. But the period since 2004 is interesting for a few reasons. First, growth resulted in an unprecedented withdrawal of workers from agriculture in this period. Second, the vast majority of non-farm employment was generated in construction and not in manufacturing. Third, the period saw a large shift in the composition of the female workforce. Older, less educated, agricultural women workers left the workforce, in part due to rising incomes and in part due to mechanisation and loss of work. Younger, more educated women joined the workforce and as a result the share in regular wage work increased sharply for women.

But despite these changes, Chapter Four shows the continuing importance of gender norms when it comes to women's employment. But equally so, if norms are important on the labour supply-side by influencing the availability of women for paid work, labour demand is important too, as our analysis clearly shows.

The analysis in Chapter Five illustrated the ways in which gender, caste and religion all play a role in labour market outcomes. Even after controlling for a range of individual and household factors, most importantly education level and socioeconomic

status, we see that SC/ST workers are far more likely to be in casual wage work, and Muslim workers more likely to be in self-employment. Decent jobs (regular wage work with some benefits and contracts) still remain out of reach for the majority. Of course, a limitation of this kind of analysis is that we cannot account for the quality of education, the local geography (at a district level or even lower) and other factors that are likely to be very important in determining such outcomes.

The analysis of earnings disparities in Chapter Six builds on a very rich literature that has long studied such gaps in India and elsewhere. We see that gender penalties are by far the highest among the three identities and further that the unexplained part of the gender gap has been rising over time for salaried workers in rural areas. On the other hand, for caste and religion, the raw gaps are smaller and most of the gap is explained by differences in observed factors including industrial and occupational segregation. On segregation, we see significant decline in the extent of overrepresentation of Scheduled Castes in waste and leather-based work, though they still remain overrepresented.

We end with a few remarks of the limitations of the study. The trends and estimates presented in the foregoing pages are the result of a complex interaction of prevailing social norms with public policy. Norms shape policy implementation and outcomes, and policies in turn, over time, change norms. Policies can affect norms directly by incentivising certain behaviour that brings about changes in beliefs or indirectly by stimulating growth and structural change, which change beliefs as well. Needless to say, the aim of the report was not to tease apart these connections. Rather, it was to present a set of evidence that can be used in policy-making and in formulating future research questions.

The report has not engaged substantially with public policies that aim to reduce or remove social disparities. The history of such policies is as old as the republic itself and in some cases older. Their scope too is vast, encompassing long-standing affirmative action (reservation) policies in public institutions for marginalised groups, targeted welfare policies that aim to improve nutritional, health, educational or employment outcomes for such groups, subsidies for public transit and many more. A separate report-length study is required to engage with these questions and this is left for future work.

We have relied largely on official secondary datasets (with the exception of the India Working Survey). This enables us to make claims at the national and state level but it also comes with its own limitations. As discussed in detail in Chapter Seven, this limits us to categories used in such surveys. For example, data is available only for administrative caste categories and not at the jati level. Second, caste data is not always available for non-Hindu households. Though we have attempted to use EUS and PLFS data to estimate caste-based differences in outcomes within Muslims, the data does not allow us to make strong claims.

Further, many insights on social mechanisms, why people act in certain ways, how they negotiate the structure of norms to find agency for themselves, are simply not available from such datasets. For this, specially designed surveys, both quantitative and qualitative, are needed. For example, Chapter Four discusses the correlates of women's work based on PLFS and NFHS data. While this information is useful, it cannot yield the kind of insights that the 20 Voices project of the India Working Survey did with its detailed interviews or that Bhattacharya (2021) provides on women's agency and work.

The next few decades in India's journey are crucial ones if we are to realise the potential of our demographic dividend while making progress on the elimination of social disparities of the kind documented in this report. The progress reported here on several fronts gives us hope that the goal can be achieved through concerted efforts on part of

governments, civil society and the private sector. We hope that the analysis presented can inform such efforts.

Methods appendix

This Appendix provides details on the data and methods used in the analysis presented in the main report. Some of the details on data and methods are found in the following background papers available on the CSE website and are not repeated here.

- Abraham, Rosa. 2023. "Tracking workers across generations- a cohort-based analysis," Centre for Sustainable Employment Working Paper #57. Azim Premji University, Bangalore.https://cse.azimpremjiuniversity.edu.in/publications/#working-papers

- Abraham, Rosa, and Mridula Mohan. 2023. "Telephone Surveys for Data Collection - Some Reflections." Centre for Sustainable Employment Working Paper #53. Azim Premji University, Bangalore.https://cse.azimpremjiuniversity.edu.in/publications/#working-papers

- Abraham, Rosa, Nishat Anjum, Rahul Lahoti, and Hema Swaminathan. 2023. "What Did They Say? Respondent Identity, Question Framing and the Measurement of Employment." Centre for Sustainable Employment Working Paper #55. Azim Premji University, Bangalore.https://cse.azimpremjiuniversity.edu.in/publications/#working-papers

- Dasgupta, Zico, and Amit Basole. 2023. "Jobless Growth and Structural Transformation: Some Theoretical Considerations and Emperical Evidence from India." Centre for Sustainable Employment Working Paper #54. Azim Premji University, Bangalore.https://cse.azimpremjiuniversity.edu.in/publications/#working-papers

Before we provide details on specific methods used, a few general points are in order. All rupees values are adjusted for inflation with base as noted, unless otherwise specified. NIC codes used in the construction of broad sectoral categories introduced in Chapter Three and used throughout the report are reported in the online Results Appendix. This appendix only provides Method details. All descriptive statistics and regression results are provided in the online Results Appendix.

Chapter Two

Analysis of Periodic Labour Force Survey data

The Periodic Labour Force Survey (PLFS) is a nationally representative annual survey conducted by the National Statistical Office (NSO) since 2017-18. Before PLFS, the earlier National Sample Survey Organisation used to carry out quinquennial Employment and Unemployment Surveys (EUS). The last round of EUS was released in 2011-12. There are five rounds of PLFS unit level data available (2017-18 to 2021-22). The survey collects information on demographics for all the individuals in the surveyed household, details on

their employment status, individual labour incomes and monthly household consumption expenditure.

PLFS adopts a stratified multi-stage sampling design, where the first-stage units (FSUs) are urban frame survey blocks for the urban areas and the 2011 population census villages for the rural areas. The ultimate stage units are households. For urban households, the PLFS follows a rotational panel sampling design. In this sampling design, out of all sample households selected from the urban frame, only 25 percent are visited in each quarter. During this visit, a 'first visit' schedule is administered. In the second quarter, another 25 percent of households are surveyed for the first time, while the previous set of households is interviewed again with a revisit schedule.[1] This process continues until by the fourth quarter, all households in the original sample have been interviewed at least once. From the second year onwards 25 percent of existing households for whom four visits have been completed are replaced by new households and the remaining 75 per cent of the sample is approached again for collecting information. For our analysis, we use only the first visit data for urban areas. In the case of rural households, there is no revisit therefore each household is surveyed once. In each quarter of the survey period, 25 percent of FSUs are covered. The PLFS design allows the generation of quarterly estimates for both rural and urban areas.

For all of our analysis regarding employment outcomes and income calculation, we have used the current weekly activity status (CWS) of an individual. CWS is more acceptable due to the shorter reference period and is expected to provide a more accurate employment situation than the other reference periods. All the calculations for the labour force participation rate, workforce participation rate, unemployment rate, employment and industry type are done using CWS. Usual Principal Status is used only in cases where the relevant information is not available in the weekly schedule (such as information on contracts, benefits etc). To generate absolute numbers (workforce, labour force etc) in any given year, the ratios calculated from the sample survey are multiplied by the population projections for that year (see reports issued by the National Commission of Population, Ministry of Health & Family Welfare). The quarterly and annual earnings numbers from PLFS are adjusted for inflation using CPI-R and CPI-U, base April-June 2022 quarter.

Chapter Three

The first year for which detailed unit-level employment data is available from the employment surveys corresponds to the NSS 38th round: Employment & Unemployment Survey (EUS), 1983. The last EUS round came out in 2011, between 1983 and 2011 there were four more rounds in 1987, 1993, 1999 & 2004. EUS was replaced by the annual Periodic Labour Force Survey in 2017. Both surveys are similar in terms of their sampling methodology and can be used together for long term analysis. Since 2017, there have been four more rounds, the latest one for the year 2021-22.

The EUS has a gap of six years on average though PLFS data is available at an annual frequency since 2017-18, we do not include all available PLFS rounds since this would weigh the sample excessively towards the latter period. Therefore, we have employment data available for eight time points; 1983-84, 1987-88, 1993-94, 1999-00, 2004-05, 2011-12 & 2018-19. The last year in this corresponds to 2018, pre covid year because the labour markets all over the world witnessed severe shock due to the COVID pandemic.

We restrict our analysis to eighteen states; Andhra Pradesh, Assam, Bihar, Delhi, Gujarat, Goa, Himachal Pradesh, Haryana, Karnataka, Kerala, Madhya Pradesh, Maharashtra, Odisha, Punjab, Rajasthan, Tamil Nadu, Uttar Pradesh and West Bengal. We have excluded north-eastern states (other than Assam), Jammu and Kashmir, and union territories (except Delhi) due to sample size issues. We also took precautions to see that identity-based slices were possible and were not resulting in very small sample sizes. While generating estimates for an identity at the state level, we replaced the estimates with missing values if the estimates were not significantly different from zero. For example, while calculating the share of the SC population in the agriculture sector in the state of Goa, the confidence interval includes zero and hence this slice was not used in the analysis.

For per capita State Domestic Product (SDP) data, the RBI's Database on Indian Economy is used. The numbers are in terms of 2011 constant prices. For the long run analysis states created since 2000 were considered as part of earlier existing states. To calculate per capita SDP of these states, a weighted average is used where weights are the proportion of the population in earlier existing state and newly formed states. For example, to calculate per capita income for undivided Bihar in 2004, the per capita SDP of Jharkhand and Bihar is used with respective population proportions (population projections are taken from the Ministry of Health & Family Welfare population projection report).

In the state panel regression, the independent variables are log per capita SDP, state fixed effects and interaction of log per capita SDP and state fixed effects. The regression model captures the effect of growth on particular employment outcomes, for instance, the proportion of women employed in the agriculture sector for a given state i at time point t. The cross-state regression framework allows us to estimate the semi-elasticities for structural change for each state and for different social identities. The standard errors in this model are clustered at the state level. The regression model was run separately for each employment outcome and for different identities to assess how the correlates changed over time.

$$\text{prop\_employment}_{it} = \alpha + \beta_1 \text{log\_percapita\_sdp}_{it} + \beta_2 \text{state}_i + \beta_3 \text{log\_percapita\_sdp}_{it} * \text{state}_i + \epsilon_i$$

Here $\beta_1 + \beta_3$ gives the semi-elasticity. Multiplying this by 100 gives us an easy interpretation- how doubling of growth (100 percent change in per capita SDP) affects a particular share in percentage point terms.

Jati Panel

Census of India collects information on jati identity for individuals belonging to Scheduled Castes (SC) and Tribes (ST). This includes information on education, work, industry of work and more recently occupation of work for these individuals. The jati level data is aggregated at the state level, therefore we have data about education and work for all SC jatis in a state. To compare jatis over time, firstly we matched jatis within the state. This includes matching across spelling variations. For example, in 1991, a jati called 'Adi Dravida' was recorded in Andhra Pradesh but in 2011 it was recorded as 'Adi Dravid'. For most of our analysis, we focus on 12 states: Andhra Pradesh, Bihar, Gujarat, Karnataka, Kerala, Madhya Pradesh, Maharashtra, Punjab, Rajasthan, Tamil Nadu, Uttar Pradesh and West Bengal.

In the 1991 census, there were 10 industrial categories recorded for a main worker (person who has worked for more than 6 months in a year), which included cultivators, agriculture labourers, plantation workers, mining, household manufacturing, non-household manufacturing, construction, trade and commerce, transportation storage & communication and other services. The number of industries recorded in 2011 increased to 19, for our analysis we match the 2011 industrial categories to the 1991 categories. A further restriction is that we limit the analysis to jatis that constitute at least 0.5 percent of the SC population in a state.

We utilise information on six industrial categories of main workers. These are: cultivation, agricultural labour, household manufacturing (HHI), non-household manufacturing (non-HHI), construction, and services. Additionally, for each state, we compute the distribution of main workers across these six categories for SCs as a whole as well as the Others (that is, for non-SC/STs). The boxplot shown in Figure 3.24 is created by computing the total number of main workers in a particular industrial category divided by the total workers in a jati. Therefore boxplots highlight the distribution of jatis across six industrial groups for different states. With the help of a boxplot, we can show how the variation across jatis in a particular industry has changed over time. For instance, on average the spread of the boxplot has increased for the construction sector, which shows that the proportion of people employed in the construction sector is very different across jatis and ranges from as low as 1 percent to as high as 30 percent. Since the boxplot shows for an individual jati, what proportion of people are employed in a given industry, the sum of boxplots for all jatis should add up to 100 percent.

Chapter Four

Women Labour Force Participation Analysis

We used a linear probability model to estimate determinants of employment participation for women using PLFS 2021-22 data. The dependent variable in all the regressions is a binary variable taking the value 1 if the individual is employed, and 0 otherwise. All regressions are estimated for the working age population. Instead of separate categories for religion and caste we introduce a religion-caste variable that combines the two. The categories include Hindu General, Hindu OBC/SC/ST, Muslim General, Muslim OBC/SC/ST, Others General and Others OBC/SC/ST with Hindu General category as base. We could not add the religion-caste variable in the long run analysis, since OBC data was only included in the EUS from 1999 onwards.

$$Employed_i = \alpha + \beta_1 age_i + \beta_2 age_i^2 + \beta_3 married_i + \beta_4 rel\_caste + \beta_5 number\ of\ children_i + \beta_6 household\ size_i + \beta_7 education\ group_i + \beta_9 head\ of\ HH\ education_i + \beta_{10} state + \varepsilon_i$$

We also estimated a linear probability model of employment for the sample of married women alone. This sub-sample allows us to introduce specific variables including the presence of mother-in-law and her employment status, and husband's earnings. For the former, we construct a variable that identifies when the mother-in-law is not present in the same household, mother-in-law is present and not employed, and mother-in-law is present and employed. The base used is mother-in-law not present.

$$Employed_i = \alpha + \beta_1 age_i + \beta_2 age_i^2 + \beta_3 married_i + \beta_4 rel\_caste_i + \beta_5 number\ of\ children_i + \beta_6 household\ size_i + \beta_7 head_i + \beta_8 mother\ in\ law_i + \beta_9 husband\ earnings_i + \beta_{10} husband\ earnings_i^2 + \beta_{11} education\ group_i + \beta_{12} head\ of\ HH\ education_i + \beta_{13} state + \varepsilon_i$$

We construct a variable that captures the husband's earnings. For this, we first needed to identify spousal pairs in the sample. We do this by using the data on individuals and their relationship to the head of the household within the sample, encompassing four distinct cases. The first case pertains to households where the head is male and the wife is the spouse of the head. In the second case, the wife is the head of the household, and the husband is the spouse of the head. It is commonly observed that married sons can co-reside with their parents in Indian households. So, in the third case, husband and wife pairs are matched based on the husband being the married child of the head of the household, and the wife being the spouse of the married son. Finally, if the wife is the married child of the head of the household, and the husband is the spouse of the married child, we are also able to identify husband-wife pairs. In order to allow a perfect match and not mix up between siblings residing in the same household, we have imposed a restriction on households to include only one married child and one daughter. This is due to the absence of an identifier in the PLFS data that would enable us to precisely identify husband-wife pairs when multiple married sons co-reside with the head of the household. Husband's earnings is the monthly income from wage, salaried or self-employed work. We include a squared term to account for a non-linear relation. We also included a control for whether the married woman is the head of the household.

IWS regression on gender norms

This is an individual-level regression on the India Working Survey sample. The dependent variable is either Employment, which is a dummy variable taking the value 1 if a person is employed for more than 20 hrs in a week, or it is the actual number of hours worked per week. The main explanatory variable is High autonomy, which is a binary variable that takes the value 1 if the woman has reported more than the median level of the autonomy index, which is calculated on a number of questions on autonomy as explained in the main text. The control variables are: ST, SC, OBC, age, age squared, education level of self, father's education level, mother's education level, marital status, wealth quintile, and household size. The "restrictive norms" dummy takes the value one if women out of the labour force reported that their family would not like them to work.

Analysis of NFHS data for impact on gender norms on employment

The National Family Health Survey (NFHS) collects data from eligible women aged 15-49 under various heads (reproduction, marriage and cohabitation, contraception, pregnancy, postnatal care, child nutrition, family planning, maternal and child health, sexual activity, fertility preferences, HIV, woman's work, household relations, domestic violence).[2] The five rounds were carried out in 1992-93, 1998-99, 2005-06, 2015-16, and 2019-20. To construct the social norms, data collected in the sections on women's work, household relations, and domestic violence is used.

Construction of norms based indices

The existing literature on empowerment led to the formation of different indices of social norms. They are body autonomy, beliefs on body autonomy, mobility, justification of domestic violence, ownership of large assets, ownership of small assets, decision making, husband's control over wife, and domestic violence. A thorough examination of the women's recode questionnaire of the five rounds of the survey led to short listing of questions under different indices. In this process, some measures were taken to arrive at the final indices with selected questions. They are as follows.

Separation of perception and actual experiences - there were questions on body autonomy and partner violence about perceptions and experiences. Since they are conceptually different, we separated them. So, separate indices were constructed and named as body autonomy and body autonomy belief, justification of domestic violence and experience of domestic violence.

Autonomy v/s infrastructure - while narrowing down the questions, an effort was made not to select questions that are connected to the availability of services and infrastructure. The focus was retained on household dynamics, autonomy and norms. For example, there is a question for women: In getting medical treatment, which of the following things is a problem? The sub questions are getting permission to go, the distance to the health facility, having to take a transport and other questions. We chose responses to getting permission over transport/distance because the former constitutes the dynamics within the household, while the latter depends on how connected the place is and the access to various modes of transport.

Similar questions in a single index - we checked whether the selected questions in an index go together by carrying out the pairwise correlations of all the variables in each index at the unit level and the collapsed averages of the variables at the state level. This exercise helped to eliminate certain variables that did not go with other variables in a category. For example, in the decision-making category, one question was, "are you allowed to have money that you can use as you wish?" This question showed a negative correlation with other variables in the category. Other questions were about "who makes the decisions on obtaining your health care? household purchases? your visits to family/relatives? use of your earnings? husband's earnings? contraception?" In this case, we eliminated the question on being allowed to have money as it did not go with other questions in this category. Similarly, for the domestic violence category, there was one question, "at any time in your life, as a child or adult, has anyone ever forced to have sexual intercourse or perform any other sexual acts?". This question was not correlated with other questions because the remaining were related to partner violence, while this question was not. Hence, this question was removed from this index.

Simple averages - We constructed indices of social norms with simple averages of individual questions.[3] In the regression analysis, we average together the district averages of all the individual components in each index to construct district norms indices.

Binary coding - A woman's response to each individual norms question was coded 1 or 0. The coding of a regressive response was 0, the code of a progressive response was 1, and responses of don't know were coded as missing values. For example, for the question "is it justified for a husband to beat his wife if she disrespects her in laws?" The responses are yes, no, and don't know. Yes was coded as 0, no as 1, and don't know as missing. A simple average of a set of 1/0 responses gives us an index that ranges from 0 to 1, where 0 is the regressive beginning of an index, and 1 marks the progressive end.

Strict v/s lenient definition - in the process of doing binary coding for questions with multiple responses, we experimented with strict and lenient approaches to coding. The rationale for doing this was to understand if the results were sensitive to the coding. For the questions where this approach is adopted, the strict coding mainly refers to the woman alone doing a particular thing. Whereas, lenient coding refers to the woman carrying out something along with her partner. For e.g. for the question on decision making of a woman's health care, a woman alone deciding it is coded as 1 in the strict definition. Her husband, woman and her partner jointly, someone else is 0. In the lenient definition, a woman alone and a woman with her partner deciding is coded as 1, and someone else is 0. Refer to Results Appendix for the list of strict and lenient definitions.

Sample - we categorised the selected questions as asked to all women or married women. NFHS rounds 1 (1992-93) and 2 (1998-99) were asked to ever married women aged 15-49. But NFHS round 3 (2005-06) till round 5 (2019-20) collected data from all women aged 15-49. In our categorisation, the questions on decision making were asked only to married women and those on husband's control over wife, experience of domestic violence were asked to selected married women in the domestic violence module.[4] The questions on body autonomy, body autonomy belief, mobility, justification of domestic violence, and ownership of assets were asked to all women aged 15-49.

Division of indices based on time periods - among the final set of selected questions, different questions were introduced in different rounds. The time frame of the different indices was divided to suit their inclusion over rounds. For example, the three selected questions of body autonomy were introduced in the first round itself and were asked till the last round. This index was constructed from 1992-93 to 2019-20. Decision-making, mobility, and justification over domestic violence were from 1998-99 to 2019-20. Husband's control over wife and partner violence was from 2005-06 to 2019-20. Ownership of large and small assets was introduced only in 2015-16 and was also asked in 2019-20. There were new questions of a particular index in subsequent rounds, for example, questions on husband's earnings and decisions on contraception in 2005-06. In such a case, two versions of the decision-making index, a longer time period of 1998-99 to 2019-20 and a shorter one from 2005-06 to 2019-20 were done. See Results Appendix for indices, selected questions, variable name, and year.

Privacy settings - some of the indices had sensitive questions about partner violence, refusal to have sex etc. We ensured the privacy levels of the interviews. Two questions were checked for this - one about the presence of others for the household relations module (where domestic violence questions are asked) and another about interruptions in interviews due to someone listening, coming inside the room, or being interrupted in any other way. In all the rounds, 96-97 percent of the overall sample's privacy was retained.

Robustness checks - some of the selected questions were based on the beliefs and perceptions of the respondents. To strengthen its validity, we checked the overall trends of some other variables and used them as controls in the regression analysis. These variables were chosen based on the literature. They are age at marriage, difference between husband's and wife's ages, son preference (sex of the last born child and difference between the ideal number of sons and daughters). The assumption is that the higher the age of marriage and the lesser the difference between the husband and wife's ages, the woman may be in a progressive position in the discussed indices. If the last born is male and the ideal number of sons is higher than daughters, this is taken as a proxy for relatively more conservative social norms.

Regression analysis

A linear probability model is fit to the data to understand the relationship between social norms and women's labour market outcomes. The regression is carried out on NFHS-4 (2015-16) data (since we wanted to avoid Covid-related effects which may be present in the most recent round). It should be noted that district-level estimates can only be produced from NFHS-4 onwards. The model specification is as follows:

$$\text{Women's employment}_i = \alpha + \beta_1 \text{distnorm} + \beta_2 \text{distlabourdemand} + \beta_3 \text{age}_i + \beta_4 \text{agesq}_i + \beta_5 \text{agediff}_i + \beta_6 \text{caste}_i + \beta_7 \text{hhsize}_i + \beta_8 \text{religion}_i + \beta_9 \text{education}_i + \beta_{10} \text{children}_i + \beta_{11} \text{wealthindex}_i + \beta_{12} \text{state} + \varepsilon_i$$

Here, women's employment for the last seven days is the dependent variable. It is a binary variable and takes the value 1, if the woman did any other work apart from the housework for the last seven days. *distnorm* is the average of any given variable or the average index (when we combine variables into one) in a district the woman resides. For variables with strict and lenient definitions, only the strict definitions of the norms are taken as explanatory variables. *distlabourdemand* is a proxy for labour demand in the district generated from the type of husband's occupation. It is coded as a binary variable where not working and work in agriculture is coded as 0, and work in non-agriculture is 1. The norm and the labour demand variables were collapsed at the district level and used as explanatory variables to reduce the problem of endogeneity that results when norms are asked of the same woman whose employment outcomes are being modelled.

One point is important to make here with regard to district level estimates. The NFHS runs two different samples, a large sample with a smaller questionnaire and a smaller sample that has a larger number of questions. Many of the questions on gender norms that we are interested in (e.g. mobility restrictions) are only asked in the smaller sample. For these, the NFHS only releases state-level estimates. However, the sampling method is such that unbiased district level averages can be produced, however, due to a small sample size, the confidence intervals around these averages will be larger. We include district level averages for the State sample as well. Our reasons for doing so is that if the point estimates at the district level are too imprecise it will not be possible to detect any statistically significant effects. However, as shown in the main report, we do find significant results.

The other controls at the individual or household level are as follows: *age* of the woman, *agesq*, *agediff* - age difference of the woman and her husband, *caste*, *hhsize* - household size, *religion*, *education* - years of education of the woman, *children* - number of children of the woman, *wealth index* of the household - a quintile score given to a household based on the number of consumer goods owned and housing characteristics using principal component analysis by NFHS, and *state* are the individual and household characteristics included as explanatory variables. *age, agesq, agediff, education, hhsize, children* are continuous explanatory variables. *caste, religion, wealth index, and state* are categorical explanatory variables.

The Role of Public Infrastructure and Local Demand

An ordinary least squares (OLS) regression is used to estimate the role of public transport and demand on the proportion of women working outside the home. We run two regressions for rural and urban areas separately.

$$\text{prop\_female\_out}_i = \alpha + \beta_1 \text{prop\_female\_graduate}_i + \beta_2\, \text{prop\_sc\_st}_i + \beta_3\, \text{median\_male\_} \text{distance}_i + \beta_4\, \text{prop\_male\_bus}_i + \beta_5\, \text{log\_manufacturing}_i + \beta_6\, \text{log\_} \text{construction}_i + \beta_7\, \text{log\_traditional\_svcs}_i + \beta_8\, \text{log\_modern\_svcs}_i + \beta_9 \text{log\_working\_age\_population}_i + \beta_{10}\, \text{state}_i + \varepsilon_i$$

The dependent variable is the proportion of female workers employed in other than agriculture work and household industries travelling more than 1.5 km from home to workplace in a district. In the regression model, we control for the proportion of female graduates in a district (as an indicator of both labour supply and locally operative gender norms), proportion of SC/ST population in the district (as an indicator of identity-based privilege/vulnerability), median distance travelled by men (as an indicator of labour demand), proportion of men travelling by bus (as an indicator of the availability of public transport), log of manufacturing firms in a district hiring at least 1 worker, log of construction firms in a district hiring at least 1 worker, log of traditional services firms in a district hiring at least 1 worker, log of modern services firms in a district hiring at least 1 worker (the modern and traditional services are defined above), and finally, the working age population in the district (indicating the overall labour supply of the region). For this regression analysis, standard errors are clustered at the district level.

Chapter Five

Employment type analysis

Data from EUS 1983, 1993, 2004, 2011, and PLFS 2017 and 2021 are used for this analysis. We use a multinomial logistic model to estimate the determinants of different kinds of employment (own account work, unpaid work, salaried work and casual work). These regressions were run separately for each year to understand how the correlates changed over time.

$$\text{Pr(Employment type)} = \alpha + \beta_1 age_i + \beta_2 age_i^2 + \beta_3 female_i + \beta_4 married_i + \beta_5 number\ of\ children_i + \beta_6 head_i + \beta_8 urban_i + \beta_9 religion_i + \beta_{10} sc/st_i + \beta_{11} education\ group_i + \beta_{11} head\ of\ HH\ education_i + \beta_{12} state + \varepsilon_i$$

The dependent variable in all the regressions is a categorical variable with four categories - Own Account Worker, Unpaid Worker, Salaried Worker and Casual Worker, with Casual worker as the base. All regressions are estimated for the working age population. We control for individual level characteristics, namely gender of the individual, their age (and age-squared term to capture non-linear relation), marital status (1 if married, 0 otherwise) and education level (illiterate, literate, primary & middle, secondary & higher secondary, diploma & above with illiterate as base). We include controls for household attributes including religion (Hindu, Muslim, Others with Hindu as the base), caste (SC/ST or not), number of children aged below 5 years in the household and education of the head of the household. We control for region (urban or rural) and include state dummies so as to control for state-level differences.

The regression coefficients from such a logit regression model are not interpretable straightforwardly. The numbers reported in the chapter are marginal effects obtained using the Stata margins command. A marginal effect is the change in the probability of a certain outcome compared to the base outcome with a change in the independent variable at a certain specified point in the logistic function.

We did a similar kind of multinomial logistic regression (4 categories as above) with the margins analysis for only women workers. PLFS 2021 data was used for this analysis. Instead of the religion and caste variable, we have introduced a religion-caste variable that combines the two. The categories include Hindu General, HIndu OBC/SC/ST, Muslim General, Muslim OBC/SC/ST, Others General and Others OBC/SC/ST with Hindu General category as base. The education variable is similar with the exception that secondary and higher secondary were retained as separate categories here. This is because the separation between secondary and higher secondary education was included in EUS from 1993 onwards.

$$\begin{aligned} Pr(Employment\ Type) \quad = \quad & \alpha + \beta_1 age_i + \beta_2 age_i^2 + \beta_3 married_i + \beta_4 number\ of\ children_i + \\ & \beta_5 head + \beta_6 household\ size + \beta_7 rel\_caste_i + \beta_8 education\ group_i \\ & + \beta_9 head\ of\ HH\ education_i + \varepsilon_i \end{aligned}$$

Intergenerational mobility

The unit level data used in this analysis comes from the 61st (2004–05) and 68th rounds (2011–12) of NSS EUS and the 2018-19 Periodic Labour Force Survey (PLFS) undertaken by the National Statistical Office (NSO). Although previous rounds of NSS EUS are available, since other studies have used this data to look at similar questions (Reddy 2015, Hnatskova et al.2013), we do not consider them here. Similarly, while an earlier round of PLFS data is available (2017-18), we do not consider this since we are interested in long term mobility (rather than year on year). Finally, the 2019-20 and subsequent rounds of PLFS would include pandemic periods and we do not include these.

Using the data on an individual and their relationship to the head of the household for each individual in the sample, we map sons aged between 24 and 45 and their fathers. Intergenerational mobility estimates can vary significantly depending on the age at which sons' occupational status is observed, mainly because of lifecycle effects. For instance, a 60 year and 18 year old father-son pair would not be entirely comparable with a 60 year and a 35 year old father-son pair. Maintaining a lower age limit of 16 for sons could introduce potential bias in intergenerational estimates due to the high likelihood of continued education and relative job instability during the early stage of individual's careers, particularly in more recent cohorts as tertiary education enrolment has increased. To address these issues, we set a lower bound of 24 years of age for the sons and an upper bound of 45 years of age for fathers. Our first step in drawing our sample is to restrict it to employed fathers and sons who co-reside in the same household. While estimating intergenerational mobility from cross-sectional data is not ideal, the lack of appropriate nationally representative long run longitudinal household data in India leaves us with no other option.

We studied intergenerational mobility in the labour market mainly through three different categorisations. First, we use the skill-based categorisation of Hnatskova et al (2013) and others and classify individuals into four different occupational groups using the information on their main occupation(three digit level occupational description of the National Classification of Occupations NCO-2004) available from each EUS and PLFS. These three digit occupation code were aggregated to one digit code.The individuals are grouped into white collar workers, skilled/semiskilled workers, farmers and unskilled workers.[5] We extrapolate Reddy (2015)'s framework to understand the long run trend, adding the 2018-19 time period. Second, we use the employment arrangement categories such as regular salaried employees with both contract and benefits,

those with either contract or benefits, and those with neither contract nor benefits, casual workers and the self-employed (including unpaid workers). The above classification helps us to understand another form of employment mobility in the Indian labour market in the long run. Third, we also examine intergenerational mobility across industries to answer the question of whether the movement out of agriculture has percolated across all social groups, and into what industries. We use broad industry categories of agriculture,manufacturing, construction, traditional services and modern services. In the report we only focus on the second dimension noted above.

The main axis of our analysis is caste and religious identities. The most widely used method of measuring intergenerational mobility involves constructing a mobility matrix. This intergenerational mobility (IGM) matrix is created using categorical variables, such as occupation, employment, or industry, and displays the percentage of sons in each category, given that their father is in a particular category. Each cell in the matrix, denoted as Pij, represents the share of sons who belong to the jth category, given that their father belongs to the ith category. The diagonal cells in the mobility table reflect immobility or intergenerational persistence.

Education Convergence

As mentioned above, the Census data collects information about the education of individuals belonging to different jatis. It allows us to compare how a particular jati has performed in terms of educational attainment over the years. Using the same jati panel, we measure the change in education levels between jatis for 12 major states. The analysis is restricted to jatis that constitute at least 0.5 percent of the SC population in a state. The four education levels that we check for convergence or divergence are: literate, primary, secondary, higher secondary and above.

The proportion of people in any particular education category is calculated as the total number of people with that level of education divided by the total population in that group. For instance, the proportion of literate in 'Adi Dravid' jati of Andhra Pradesh in 1991 is calculated by dividing the literate population of 'Adi Dravid' by the total jati population. Similarly, the proportion for various education categories can be calculated for 2011. The difference between the two proportions is the change in literate proportion for a jati over the time period of 20 years.

Chapter Six

Analysis of earnings gaps

Data from EUS 1983, 1993, 1999, 2004, 2011 and PLFS 2017 and 2021 were used for the analysis. We used an OLS model to estimate a Mincerian wage equation. The regression model was run separately for each year to assess how the correlates changed over time. We ran separate models for regular wage or salaried earnings, casual wage earnings and earnings from self-employment.

$$\text{Log Salary}_i = \begin{aligned}&\alpha + \beta_1 age_i + \beta_2 age_i^2 + \beta_3 femalei + \beta_4 married_i + \beta_5 head\ of\ the\ HH_i + \beta_6 religion_i + \\ &\beta_7 scst_i + \beta_8 urban_i + \beta_9 number\ of\ children_i + \beta_{10} education\ group_i + \beta_{11} head\ of\ HH \\ &education_i + \beta_{12} state + \beta_{13} Industry\&occupation + \varepsilon_i\end{aligned}$$

The dependent variable here is the log of real salaried earnings, which is a continuous variable. Earnings are adjusted for inflation using CPI-AL and CPI-IW series since 1983 (base 2020). For the long run period five regression models were used. One for all salaried workers, one each for male salaried workers and female salaried workers. In addition to this, separate regressions were carried out by sector wise (urban & rural) for salaried workers. We control for individual level characteristics, namely their age (and age-squared term to capture non-linear relation), gender (female dummy), marital status (1 if married, 0 otherwise), whether the person is the head of the household and education level (illiterate, literate, primary & middle, secondary & higher secondary, diploma & above with illiterate as base). We include controls for household attributes including religion (Hindu, Muslim, Others with Hindu as the base), caste (SC/ST or not), number of children aged below 5 years in the household and education of the head of the household. We include state dummies to control for state-level differences and an industry-occupation concatenated dummy variable to control for industry and occupation level differences.

We carried out concatenation of the industry-occupation variable by combining 2 digit NIC (industry)with 2 digit NCO (occupation) codes. The categories with frequency less than 20 were combined into one category depending on their first digit.

$$\text{Log Salary}_i = \begin{aligned}&\alpha + \beta_1 age_i + \beta_2 age_i^2 + \beta_3 female_i + \beta_4 married_i + \beta_5 head\ of\ the\ HH_i + \beta_6 religion\_ \\ &caste + \beta_7 urban_i + \beta_8 number\ of\ children_i + \beta_9 education\ group_i + \beta_{10} head\ of\ HH \\ &education_i + \beta_{11} state + \beta_{12} Industry\&occupation + \varepsilon_i\end{aligned}$$

Similarly, 5 regression models were used for the short run analysis. The analysis was based on the EUS 1999, 2004 and 2011 data and PLFS 2017 and 2021 data. The only difference from the long run analysis is that in the short run, instead of using separate religion and caste dummy, we use a religion-wise caste variable. The categories include Hindu General, HIndu OBC/SC/ST, Muslim General, Muslim OBC/SC/ST, Others General and Others OBC/SC/ST with Hindu General category as base. We could not add the religion-caste variable in the long run analysis, since OBC data was only included in the EUS from 1999 onwards. Also a small change in the education variable, the secondary and higher secondary categories are retained as separate categories here. This is because the separation between secondary and higher secondary education was included only from EUS 1993 onwards.

$$\text{Log weekly wage} = \begin{aligned}&\alpha + \beta_1 age_i + \beta_2 age_i^2 + \beta_3 female_i + \beta_4 married_i + \beta_5 head\ of\ the\ HHi + \beta_6 number \\ &of\ days\ worked\ in\ a\ week_i + \beta_7 religion_i + \beta_8 scst_i + \beta_9 urban_i + \beta_{10} number \\ &of\ children_i + \beta_{11} education\ group_i + \beta_{12} head\ of\ HH\ education_i + \beta_{13} state + \\ &\beta_{14} Industry\&occupation + \varepsilon_i\end{aligned}$$

The dependent variable in all the regressions is the log of real weekly wages. We are using weekly wages and not monthly wages because for casual employment the wage data is available only for the last week of work. Three regressions models were used for this analysis. One for all casual workers, and one each for casual workers in urban and rural sectors. We control for individual level characteristics, namely their age (and age-squared term to capture non-linear relation), gender (female dummy), marital status (1 if married, 0 otherwise), whether the person is the head of the household, number of days the person worked in a week and education level (illiterate, literate, primary & middle, secondary & higher secondary, diploma & above with illiterate as base). We include controls for household attributes including religion (Hindu, Muslim, Others with Hindu as the base), caste (SC/ST or not), number of children aged below 5 years in the household and education of the head of the household. We include state dummies to control for state-level differences and an industry-occupation concatenated variable to control for industry and occupation level differences.

For an analysis of earnings from self-employment data from PLFS 2017-18 and 2021-22 was used because earlier EUS rounds did not collect this data.

$$
\begin{aligned}
\text{Log monthly earnings} = {} & \alpha + \beta_1 age_i + \beta_2 age_i^2 + \beta_3 female_i + \beta_4 married_i + \beta_5 head\ of\ the\ HH_i + \\
& \beta_6 religion_i + \beta_7 scst_i + \beta_8 urban_i + \beta_9 number\ of\ children_i + \beta_{10} education\ group_i \\
& + \beta_{11} head\ of\ HH\ education_i + \beta_{12} state + \beta_{13} Industry\&occupation + \varepsilon_i
\end{aligned}
$$

The dependent variable is the log of real monthly earnings of self employed workers. We control for individual level characteristics, namely their age (and age-squared term to capture non-linear relation), gender (female dummy), marital status (1 if married, 0 otherwise), whether the person is the head of the household and education level (illiterate, literate, primary & middle, secondary & higher secondary, diploma & above with illiterate as base). We include controls for household attributes including religion (Hindu, Muslim, Others with Hindu as the base), caste (SC/ST or not), number of children aged below 5 years in the household and education of the head of the household. We include state dummies to control for state-level differences and an industry-occupation concatenated variable to control for industry and occupation level differences.

Blinder–Oaxaca decomposition

This technique and its variations have been widely used to decompose gender and racial wage gaps. The decomposition allows us to quantify the extent to which the outcome gap between two groups can be accounted for by differences in observed characteristics or differences in returns to those characteristics. This is a two-way decomposition. It is also possible to calculate a three-way decomposition with an interaction term. But in the labour economics literature, two-way analysis is more common with the endowments component being thought of as the "explained" and the returns component as the "unexplained."

$$
R = [E(X_A) - E(X_B)]\beta^* + E(X_A)'(\beta_A - \beta^*) + E(X_B)'(\beta^* - \beta_B)
$$

In the above equation (Jann 2008), the first component is the explained -that part of the overall outcome gap that is accounted for by differences in observed characteristics. The second and third terms are together the unexplained part resulting from the difference between a group's returns (the betas) and β^* a nondiscriminatory coefficients vector. β_A and β_B come from separate regressions run on the two groups (e.g. a wage equation for men and a wage equation for women). The non-discriminatory vector can be either taken as the dominant group's coefficients or the marginalised group's or those coming from a pooled regression (men and women together) depending on the theoretical position adopted. See Chapter Six for more details.

Industrial and occupational segregation

This analysis is based on data from EUS 1983-84, 1993-94, 2004-05, 2011-12, and PLFS 2017-18 and 2021-22. There are two measures that have been used to estimate the extent of segregation for different identity groups across industries and occupations. These are - an overall industrial segregation index and a representation index within each industry and occupation.

The occupational segregation analysis is done using cross-sectional data from PLFS 2021. For this, occupations at 2-digit NCO 2015 and earnings for the current weekly activity status have been used. Only non-farm occupations have been used for this analysis.

For the industrial segregation analysis, industries for the principal activity status have been used for all years listed above. For consistency, industry codes from NIC 1970, 1987, 1998, 2004 and 2008 are first concorded to harmonise all industries according to NIC 1998 classifications. Similar industries at 2-digit NIC-1998 are then clubbed together to form 46 broad (non-farm) industry groups (see online Results Appendix for the Industry codes and descriptions).
The two measures of segregation used in this analysis are explained below:

Index of Segregation

The Duncan Segregation Index is constructed within castes, gender, and religions across Indian industries. For any identity, the segregation Index is calculated between two groups using the formula for below:

$$\frac{1}{2}\sum_{i=1}^{N} \ \left|\frac{xi}{X} - \frac{yi}{Y}\right|$$

where,

x_i = workers of group 1 in the ith industry
X = total group 1 workers in the workforce
y_i = workers of group 2 in the ith industry
Y = total group 2 workers in the workforce

For each identity, the two groups are as follows:

Gender: Male and Female

Religion: Hindus and Muslims

Caste (2004 – 2021):   Others and Scheduled Caste

                          Others and Scheduled Tribes

                          Others and Other Backward Classes

Caste (1983 – 2021):   (Others + Other Backward Classes) and Scheduled Caste

                          (Others + Other Backward Classes) and Scheduled Tribe

The segregation index takes values between 0 and 1 where 0 represents no segregation and 1 represents complete segregation. A value of 0.5 indicates that 50 percent of the workforce needs to be reallocated between industries to achieve a state of no segregation.

Representation Index

The representation index for each group within any identity category (caste, gender, religion) is constructed by computing the ratio of the share of that group in any industry or occupation and in the workforce. For any identity group i and industry or occupation j, the representation index is given by:

$$RI_{ij} = \frac{(I_{ij}/\Sigma_i I_{ij}) *100}{(WF_i/\Sigma_i WF_i) *100}$$

where, I_{ij} is the total number of workers of identity group i employed in industry j and WF_i is total number of persons in the workforce belonging to group i.

The representation index is always a non-negative value with a lower bound of zero and no upper bound. If the representation index is equal to 1, it indicates representation in proportion to workforce share. Value of the representation index less than 1 indicates under-representation and a value greater than 1 indicates over-representation.

Caste and entrepreneurship

This analysis uses data on ownership of enterprises comes from three latest rounds of the Economic Census for the years 1998, 2005 and 2013. Data on the share of different social groups in the workforce comes from three rounds of the Employment and Unemployment Survey (EUS) for the years 1999-00, 2005-06 and 2011-12. Gross Value Added data for enterprises in the organised sector is estimated from the Annual Survey of Industries (ASI) for 2005-06 and 2010-11. For enterprises in the unorganised sector, GVA data is estimated from the NSS Survey on Unincorporated Non-Agricultural Enterprises Excluding Construction (NSS) for 2010-11 and 2015-16. The EUS, ASI and NSS rounds are chosen such that they are the closest corresponding years to the Economic Census rounds.

The analysis in this paper is carried out in two steps as mentioned in Chapter Six (Section 6.3). These steps are explained in detail in the following.

Representation Index

The representation index for each caste group is constructed by computing the ratio of the share of that caste group in ownership of enterprises and in the workforce. Workforce share is taken as the reference to have a conservative estimate of the representation index, so that the representation in ownership is conditioned on the share of the population that are of working age and engaged in some economic activity. Moreover, since data on the caste of the owner is available only for enterprises with private proprietary ownership, all analysis in this paper is restricted to enterprises that are owned by private proprietors. The representation index is first constructed at the aggregate for all enterprises and then separately for different sizes of the enterprise where size is measured by total number of persons employed. Therefore, for a caste group i and enterprise size j, the representation index is given by:

$$RI_{ij} = \frac{(O_{ij}/\Sigma_i O_{ij}) *100}{(WF_i/\Sigma_i WF_i) *100}$$

Where, O_{ij} is total number of enterprises of size j owned by caste group i and WF_i is total number of persons in the workforce belonging to caste group i.

Assuming that each caste group has non-zero representation in the workforce, the representation index is always a non-negative value with a lower bound of zero and no upper bound. If the representation index for any caste group is equal to 1, it means that the caste group is represented in ownership of enterprises in proportion to its workforce share. For any caste group, if the value of the representation index is less than 1, then the caste group is under-represented in enterprise ownership and if it is greater than 1 then it is over-represented in ownership of enterprises.

Caste-Penalty/Premium

For the manufacturing sector, a caste penalty/premium is assigned to under-/over-represented social groups by calculating the difference in total GVA *actually controlled* by a caste group and the total GVA that *would be controlled* by the caste group if their share in ownership of enterprises was same as their share in the workforce. This is computed separately for each enterprise size and then aggregated to arrive at the total caste penalty/premium.

GVA *actually controlled* by a caste group for any enterprise size is obtained by multiplying the average GVA of that enterprise size with the total number of enterprises of that size owned by the caste group. To obtain total GVA that *would be controlled* by a caste group if their share in ownership was same as their share in the workforce, the workforce share of the caste group is first multiplied by the total number of enterprises of a given size and this is then multiplied by the corresponding average GVA. Total number of enterprises is always obtained from the Census data.

The average GVA for a particular enterprise size is obtained from the ASI and NSS data for the organised and unorganised sector enterprises respectively. Therefore, for enterprises employing up to 10 persons that should be in the unorganised sector, the average GVA value comes from NSS data. For enterprises employing more than 20 persons which should be in the organised sector, the average GVA value is obtained from the ASI data. If the enterprise employs between 10 and 20 persons, then it could be in the organised sector

if it uses electricity and in the unorganised sector if it does not. Hence, for these enterprises the average GVA for each enterprise size is calculated by taking ASI and NSS data together. For estimating GVA values from either ASI or NSS or both, only private proprietary enterprises are considered to keep parity with the representation index.

The penalty/premium for caste group i in enterprise size j is given by:

$$p_{ij} = (\text{Mean GVA}_j * O_{ij}) - (\text{Mean GVA}_j * (\underline{\Sigma_i O}_{ij} * (WF_i/\Sigma_i WF_i)))$$

And aggregate caste penalty/premium for any caste group i is given by:

$$\mathbf{P}_i = \Sigma_j (p_{ij})$$

If the value of p_{ij} or $\mathbf{P}_i$ is positive, it denotes a caste-premium and if the value is negative, it denotes a penalty. If there is no penalty or premium for a social group, then the value of p_{ij} or $\mathbf{P}_i$ should be equal to zero.

Endnotes

1   The revisit schedule remains largely the same as the first visit schedule. However, there are some points of departure. First, except for age, demographic information on erstwhile household members is not collected. Second, it collects information only on the activities of each member in the last seven days and their weekly status, and usual principal activity (i.e. what they were involved in for the majority of the year is not asked about.)

2   NFHS has different questionnaires asked to men and women. This study is based on the women's questionnaire only.

3   We tried various methods of calculating the social norms index based on the literature, such as Principal Component Analysis (PCA), the method of Women's Empowerment in Agriculture Index (WEAI) (see https://www.ifpri. org/publication/women%E2%80%99s-empowerment-agriculture-index) and the Patriarchy Index (see https:// link.springer.com/article/10.1007/s11205-021-02752-1). We found the simple averages method to be the most straightforward in the absence of good reasons to weight different questions differently.

4   In the state module sub-sample, one eligible woman per household was randomly selected and asked questions in the domestic violence section to adhere to ethical requirements.

5   White collar workers in our classification consist of people with occupations which require the highest levels of skills. They were formed by aggregating three one-digit occupations: legislators, senior officials and managers; professionals; and associate professionals. We have classified all individuals who are engaged in skilled occupation— such as clerks, service workers and shop and market sales workers, craft and related trades workers, and plant and machine operators and assemblers as semi-skilled or skilled workers (Reddy, 2015).

Bibliography

Aaronson, Daniel, Rajeev Dehejia, Andrew Jordan, Cristian Pop-Eleches, Cyrus Samii, and Karl Schulze. 2017. "The Effect of Fertility on Mothers' Labor Supply over the Last Two Centuries." Working Paper Series. National Bureau of Economic Research. https://doi.org/10.3386/w23717.

Abraham, Rosa. 2023. "Tracking Workers across Generations –a Cohort-Based Analysis." Centre for Sustainable Employment Working Paper #57. Azim Premji University, Bangalore. https://cse. azimpremjiuniversity.edu.in/publications/#working-papers

Abraham, Rosa, and Anand Shrivastava. 2022. "How Comparable Are India's Labour Market Surveys?" *The Indian Journal of Labour Economics* 65 (2): 321–46.

Abraham, Rosa, and Mridula Mohan. 2023. "Telephone Surveys for Data Collection - Some Reflections." Centre for Sustainable Employment Working Paper #53. Azim Premji University, Bangalore. https://cse.azimpremjiuniversity.edu.in/publications/#working-papers

Abraham, Rosa, Nishat Anjum, Rahul Lahoti, and Hema Swaminathan. 2023. "What Did They Say? Respondent Identity, Question Framing and the Measurement of Employment." Centre for Sustainable Employment Working Paper #55. Azim Premji University, Bangalore.https://cse.azimpremjiuniversity.edu.in/publications/#working-papers

Abraham, Vinoj. 2013. "Missing Labour or Consistent 'De-Feminisation'?" 48 (31). https://www.epw.in/journal/2013/31/special-articles/missing-labour-or-consistent-de-feminisation.html.

Action Aid. 2020. "Workers in the Time of Covid-19: Round II of the National Study on Informal Workers." Action Aid.

Afridi, Farzana, Monisankar Bishnu, and Kanika Mahajan. 2023. "Gender and Mechanization: Evidence from Indian Agriculture." *American Journal of Agricultural Economics* 105 (1): 52–75. https://doi.org/10.1111/ajae.12315.

Agrawal, Ankush, and Vikas Kumar. 2020. *Numbers in India's Periphery: The Political Economy of Government Statistics*. Cambridge, United Kingdom ; New York, NY: Cambridge University Press.

Agrawal, Tushar. 2014. "Gender and Caste-Based Wage Discrimination in India: Some Recent Evidence." Journal for Labour Market Research 47 (4): 12. https://doi.org/10.1007/s12651-013-0152-z.

———. 2016. "Occupational Segregation in the Indian Labour Market." *The European Journal of Development Research* 28 (2): 330–51. https://doi.org/10.1057/ejdr.2015.10.

Agüero, Jorge M, Mindy Marks, and Neha Raykar. 2020. "Economic Development and the Motherhood Wage Penalty." Working Paper 2020-06. University of Connecticut, Department of Economics. https://www. econstor.eu/bitstream/10419/248342/1/wp2021-128.pdf.

Alvi, Muzna, Shweta Gupta, Ruth Meinzin-Dick, and Claudia Ringler. 2020. "*Phone Surveys to Understand Gendered Impacts of COVID-19: A Cautionary Note*." CGIAR. Phone Surveys to Understand Gendered Impacts of Covid-19: A Cautionary Note. (blog). 2020. https://pim.cgiar.org/2020/07/14/phone-surveys-to-understand-gendered-impacts-of-covid-19-a-cautionary-note/.

Ambler, Kate, Cheryl Doss, Caitlin Kieran, and Simone Passarelli. 2021. "He Says, She Says: Spousal Disagreement in Survey Measures of Bargaining Power." *Economic Development and Cultural Change* 69 (2): 765–88. https://doi.org/10.1086/703082.

Angelov, Nikolay, Per Johansson, and Erica Lindahl. 2016. "Parenthood and the Gender Gap in Pay." *Journal of Labor Economics* 34 (3): 545–79. https://doi.org/10.1086/684851.

Anker, Richard. 1983. "Female Labour Force Participation in Developing Countries : A Critique of Current Definitions and Data Collection Methods." *International Labour Review.*

——— 1998. Gender and Jobs: Sex Segregation of Occupations in the World. Geneva: International Labour Office.

Ansari, Khalid Anis. 2016. "The 'Muslim' Affirmative Action Debate: Post-Sachar Reflections." In *What Ails Indian Muslims?*, 147–66. Aakar. http://www.aakarbooks.com/details.php?bid=56.

———. 2018. "Pasmanda Politics and Muslim (Minority) Discourse in India." In *Gender, Caste and the Imagination of Equality*, 310–23. Women Unlimited. https://www.amazon.in/gender-caste-imagination-equality-anupama/dp/938560614X.

Anukriti, S, Catalina Herrera-Almanza, Praveen K. Pathak, and Mahesh Karra. 2020. "Curse of the Mummy-Ji: The Influence of Mothers-in-Law on Women in India†." *American Journal of Agricultural Economics* 102 (5): 1328–51. https://doi.org/10.1111/ajae.12114.

Asher, Sam, Paul Novosad, and Charlie Rafkin. 2021. "Intergenerational Mobility in India: New Methods and Estimates across Time, Space and Communities." Working Paper. GLM|LIC WORKING PAPER NO. 66. Bonn, Germany: Institute of Labor Economics(IZA). https://g2lm-lic.iza.org/publications/wp/intergenerational-mobility-in-india-new-methods-and-estimates-across-time-space-and-communities/.

Azam, Shireen. 2023. "Scheduled Caste Status for Dalit Muslims and Christians" 58 (27). https://www.epw.in/journal/2023/27/commentary/scheduled-caste-status-dalit-muslims-and.html.

Banerjee, Biswajit, and J. B. Knight. 1985. "Caste Discrimination in the Indian Urban Labour Market." *Journal of Development Economics* 17 (3): 277–307. https://doi.org/10.1016/0304-3878(85)90094-X.

Bardasi, Elena, Kathleen Beegle, Andrew Dillon, and Pieter Serneels. 2011. "Do Labor Statistics Depend on How and to Whom the Questions Are Asked? Results from a Survey Experiment in Tanzania." *The World Bank Economic Review* 25 (3): 418–47. https://doi.org/10.1093/wber/lhr022.

Basole, Amit. 2014. "The Informal Sector from a Knowledge Perspective." *Yojana*, no. 58: 8–13.

———. 2016. "Informality and Flexible Specialization: Apprenticeships and Knowledge Spillovers in an Indian Silk Weaving Cluster: Informality and Flexible Specialization in a Silk Weaving Cluster." *Development and Change* 47 (1): 157–87. https://doi.org/10.1111/dech.12216.

———. 2018. "The Skilled And The Schooled:India's Struggles over What Counts as Knowledge." *The Caravan: A Journal of Politics and Culture*, 2018. https://caravanmagazine.in/reviews-essays/india-struggle-knowledge.

———. 2022. "Structural Transformation and Employment Generation in India: Past Performance and the Way Forward." *The Indian Journal of Labour Economics* 65 (2): 295–320. https://doi.org/10.1007/s41027-022-00380-y.

Basole, Amit, and Paaritosh Nath. 2021. "Did Employment Rise or Fall in India between 2011 and 2017?" *Economic and Political Weekly* 56 (34). https://www.epw.in/journal/2021/34/special-articles/did-employment-rise-or-fall-india-between-2011-and.html?0=ip_login_no_cache%3Df4e0301295e3c51e9ff30f89628e3e8f.

Beegle, Kathleen. 2023. "Has Women's LFP Rate in Rwanda Fallen from 84% in 2014 to 52% in 2019?" *Development Impact World Bank Blog* (blog). April 19, 2023. https://blogs.worldbank.org/impactevaluations/has-womens-lfp-rate-rwanda-fallen-84-2014-52-2019.

Benes, Elisa, and Kieran Walsh. 2018. "Measuring Employment in Labour Force Surveys." ILO. Statistical Methodology Series. ILO. https://www.ilo.org/wcmsp5/groups/public/---dgreports/---stat/documents/publication/wcms_635732.pdf.

Bernhardt, Arielle, Erica Field, Rohini Pande, Natalia Rigol, Simone Schaner, and Charity Troyer-Moore. 2018. "Male Social Status and Women's Work." *AEA Papers and Proceedings* 108 (May): 363–67. https://doi.org/10.1257/pandp.20181086.

Bhattacharya, Shrayana. 2021. *Desperately Seeking Shah Rukh: India's Lonely Young Women and the Search for Intimacy and Independence*. Noida, Uttar Pradesh, India: HarperCollins Publishers India.

Biemer, Paul, and Lars E Lyberg. 2003. "Sampling Error." In *Introduction to Survey Quality,* 305–50. John Wiley & Sons, Ltd. https://doi.org/10.1002/0471458740.ch9.

Biswas, Arpita, and Anjani Thampi. 2021. "Women's Workforce Participation and Spousal Violence: Insights from India." UMass Amherst Economics Department Working Paper, 2021-17, https://doi.org/10.7275/RBE5-JN42.

Blair, Johnny, Geeta Menon, and Barbara Bickart. 1991. "Measurement Effects in Self vs Proxy Responses to Survey Questions: An Information Processing Perspective." In *Measurement Errors in Surveys.* John Wiley & Sons, Ltd.

Blau, Francine D., Mary C. Brinton, and David B. Grusky. 2009. *The Declining Significance of Gender*? New York: Russell Sage Foundation.

Blau, Francine D., Josefine Koebe, and Pamela A. Meyerhofer. 2021. "Who Are the Essential and Frontline Workers?" *Business Economics* 56 (3): 168–78. https://doi.org/10.1057/s11369-021-00230-7.

Bound, John, Charles Brown, and Nancy Mathiowetz. 2001. "Measurement Error in Survey Data." In Handbook of Econometrics. Vol. 5.

Bussolo, Maurizio, Jessy Amarachi Ezebuihe, Ana Maria Munoz Boudet, Stavros Poupakis, Tasmia Rahman, and Nayantara Sarma. 2022. "Social Norms and Gender Equality : A Descriptive Analysis for South Asia." World Bank Policy Research Working Paper Series 10142. https://ideas.repec.org/p/wbk/wbrwps/10142.html.

Carr-Hill, Roy. 2013. "Missing Millions and Measuring Development Progress." *World Development* 46 (June): 30–44. https://doi.org/10.1016/j.worlddev.2012.12.017.

Cassan, Guilhem. 2019. "Affirmative Action, Education and Gender: Evidence from India." *Journal of Development Economics* 136 (January): 51–70. https://doi.org/10.1016/j.jdeveco.2018.10.001.

Cassan, Guilhem, Daniel Keniston, and Tatjana Kleineberg. 2021. "A Division of Laborers: Identity and Efficiency in India." Working Paper 28462. National Bureau of Economic Research. https://doi.org/10.3386/w28462.

Chakraborty, Soumyajit, and Alok K. Bohara. 2021. "The Cost of Being 'Backward' in India: Socio-Religious Discrimination in the Labour Market." *Indian Journal of Human Development* 15 (2): 252–74. https://doi.org/10.1177/09737030211029634.

Chaudhary, Ruchika, and Sher Verick. 2014. "Female Labour Force Participation in India and Beyond." Delhi: ILO. https://www.ilo.org/wcmsp5/groups/public/---asia/---ro-bangkok/---sro-new_delhi/documents/publication/wcms_324621.pdf.

Comblon, Virginie, and Anne-Sophie Robilliard. 2015. "Are Female Employment Statistics More Sensitive than Male Ones to Questionnaire Design? Evidence from Cameroon, Mali and Senegal." *DIAL Working Papers*, Working Papers, , no. DT/2015/22, (December). https://ideas.repec.org//p/dia/wpaper/dt201522.html.

Dalberg. 2020. "Efficacy of Government Entitlements for Low-Income Families during Covid-19." https://impactsofcovid.in/assets/front_assets/download/20.07.27_Efficacy%20of%20government%20entitlements%20for%20low-income%20families%20during%20Covid-19.pdf

Das, Maitrey, and Puja Dutta. 2007. "Does Caste Matter for Wages in the Indian Labor Market." https://www.researchgate.net/publication/228809672_Does_Caste_Matter_for_Wages_in_the_Indian_Labor_Market.

Das, Maitreyi Bordia, and Ieva Zumbyte. 2017. *The Motherhood Penalty and Female Employment in Urban India*. Policy Research Working Paper No. 8004, World Bank, Washington, DC. https://doi.org/10.1596/1813-9450-8004.

Dasgupta, Zico, and Amit Basole. 2023. "Jobless Growth and Structural Transformation: Some Theoretical Considerations and Empirical Evidence from India." Centre for Sustainable Employment Working Paper #54. Azim Premji University, Bangalore.

De, Rahul. 2020. "Mechanisms of Surplus Appropriation in the Informal Sector: A Case Study of Tribal Migrants in Ahmedabad's Construction Industry." Centre for Sustainable Employment Working Paper #32. Azim Premji University, Bangalore. https://cse.azimpremjiuniversity.edu.in/publications/mechanisms-of-surplus-appropriation-in-the-informal-sector-a-case-study-of-tribal-migrants-in-ahmedabads-construction-industry/.

Deshmukh, Neerad, Sonalde Desai, Santanu Pramanik, and Dinesh Tiwari. 2019. "Improving the Measurement of Women's Work: Sensitivity of Work Participation Rates to Survey Design." NCAER National Data Innovation Centre Measurement Brief 2020–1. National Council Of Applied Economic Research.

Deshpande, Ashwini. 2002. "Assets versus Autonomy? The Changing Face of the Gender-Caste Overlap in India." *Feminist Economics* 8 (2): 19–35. https://doi.org/10.1080/13545700210160005.

——— 2013. "Social Justice through Affirmative Action in India: An Assessment." In *Capitalism on Trial*, edited by Jeannette Wicks-Lim and Robert Pollin. Edward Elgar Publishing. https://doi.org/10.4337/9781782540854.00029.

Deshpande, Ashwini, Deepti Goel, and Shantanu Khanna. 2018. "Bad Karma or Discrimination? Male–Female Wage Gaps Among Salaried Workers in India." In *World Development*, 102:49. https://doi.org/10.1016/j.worlddev.2017.07.012.

Deshpande, Ashwini, and Naila Kabeer. 2019. "(In)Visibility, Care and Cultural Barriers: The Size and Shape of Women's Work in India." Working Papers,10. Ashoka University,Department of Economics. https://dp.ashoka.edu.in/ash/wpaper/paper10_0.pdf.

———. 2021. "Norms That Matter: Exploring the Distribution of Women's Work between Income Generation, Expenditure-Saving, and Unpaid Domestic Responsibilities in India." WIDER Working Paper 2021. UNU-WIDER. https://doi.org/10.35188/UNU-WIDER/2021/070-2.

Deshpande, Ashwini, and Rajesh Ramachandran. 2019. "Traditional Hierarchies and Affirmative Action in a Globalizing Economy: Evidence from India." *World Development* 118 (June): 63–78. https://doi.org/10.1016/j.worlddev.2019.02.006.

Deshpande, Ashwini, and Smriti Sharma. 2013. "Entrepreneurship or Survival? Caste and Gender of Small Business in India." *Economic and Political Weekly* 48 (28): 38–49.

———. 2016. "Disadvantage and Discrimination in Self-Employment: Caste Gaps in Earnings in Indian Small Businesses." *Small Business Economics* 46 (2): 325–46. https://doi.org/10.1007/s11187-015-9687-4.

Deshpande, Ashwini, and Jitendra Singh. 2021. "Dropping Out, Being Pushed Out or Can't Get in? Decoding Declining Labour Force Participation of Indian Women." SSRN Scholarly Paper. Rochester, NY. https://doi.org/10.2139/ssrn.3905074.

Deshpande, Satish. 2006. "Employment, Exclusion and 'Merit' in the Indian IT Industry." *Economic and Political Weekly* 41 (24): 2438–44.

Dhamija, Dhruvika, and Akshi Chawla. 2023. "Decoding Women's Labour Force Participation in 2021-22: What the Periodic Labour Force Survey Shows." Centre for Economic Data and Analysis, CEDA, Ashoka Univeristy, March 20, 2023. https://ceda.ashoka.edu.in/decoding-womens-labour-force-participation-in-2021-22-what-the-periodic-labour-force-survey-shows/.

Dhanaraj, Sowmya, and Vidya Mahambare. 2019. "Family Structure, Education and Women's Employment in Rural India." *World Development* 115 (March): 17–29. https://doi.org/10.1016/j.worlddev.2018.11.004.

———. 2022. "Male Backlash and Female Guilt: Women's Employment and Intimate Partner Violence in Urban India." *Feminist Economics* 28 (1): 170–98. https://doi.org/10.1080/13545701.2021.1986226.

Dhar, Diva. 2021. "Indian Matchmaking: Are Working Women Penalized in the Marriage Market in India?" Working Paper.

Drèze, Jean, and Anmol Somanchi. 2022. "Weighty Evidence? Poverty Estimation with Missing Data." Ideas for India. https://www.ideasforindia.in/topics/poverty-inequality/weighty-evidence-poverty-estimation-with-missing-data.html.

Duraisamy, Malathy, and P. Duraisamy. 2016. "Gender Wage Gap across the Wage Distribution in Different Segments of the Indian Labour Market, 1983–2012: Exploring the Glass Ceiling or Sticky Floor Phenomenon." *Applied Economics* 48 (43): 4098–4111. https://doi.org/10.1080/00036846.2016.1150955.

Elder, Todd E., John H. Goddeeris, and Steven J. Haider. 2010. "Unexplained Gaps and Oaxaca–Blinder Decompositions." *Labour Economics* 17 (1): 284–90. https://doi.org/10.1016/j.labeco.2009.11.002.

Farré, Lídia, and Francis Vella. 2013. "The Intergenerational Transmission of Gender Role Attitudes and Its Implications for Female Labour Force Participation." *Economica* 80 (318): 219–47. https://doi.org/10.1111/ecca.12008.

Finnoff, Kade. 2012. "Intimate Partner Violence, Female Employment, and Male Backlash in Rwanda." *The Economics of Peace and Security Journal* 7 (2). https://doi.org/10.15355/epsj.7.2.14.

Fletcher, Erin, Rohini Pande, and Charity Maria Troyer Moore. 2017. "Women and Work in India: Descriptive Evidence and a Review of Potential Policies." HKS Working Paper RWP18-004. https://www.ssrn.com/abstract=3116310.

Gaddis, Isis, Gbemisola Oseni, Amparo Palacios-Lopez, and Janneke Pieters. 2023. "Who Is Employed? Evidence from Sub-Saharan Africa on Redefining Employment." *Journal of African Economies* 32 (2): 151–74. https://doi.org/10.1093/jae/ejac021.

Galdo, Jose, Ana C Dammert, and Degnet Abebaw. 2021. "Gender Bias in Agricultural Child Labor: Evidence from Survey Design Experiments." *World Bank Economic Review* 35 (4): 872–91. https://doi.org/10.1093/wber/lhaa021.

Gautham, Leila. 2022. "It Takes a Village: Childcare and Women's Paid Employment in India." *Population and Development Review* 48 (3): 795–828. https://doi.org/10.1111/padr.12504.

Ghose, Ajit K. 2016. "Globalization, Growth and Employment in India." *Indian Journal of Human Development* 10 (2): 127–56. https://doi.org/10.1177/0973703016663536.

Glasner, Tina, and Wander van der vaart. 2009. "Applications of Calendar Instruments in Social Surveys: A Review." Quality And Quantity 43: 333–49.

Goldin, Claudia. 1994. "The U-Shaped Female Labor Force Function in Economic Development and Economic History." w4707. Cambridge, MA: National Bureau of Economic Research. https://doi.org/10.3386/w4707.

Government of India. 2006. "Sachar Committee Report." Social, Economic and Educational Status of the Muslim Community of India. https://www.minorityaffairs.gov.in/WriteReadData/RTF1984/7830578798.pdf.

———. 2019. "Population Projections for India and States 2011-2036." https://nhm.gov.in/New_Updates_2018/Report_Population_Projection_2019.pdf

Government of India, 2013. 2011. "Elementary Education in India :Progress towards UEE." DISE 2011-12: Flash Statistics. New Delhi: National University of Educational Planning and Administration.

Gruber, Siegfried, and Mikołaj Szołtysek. 2016. "The Patriarchy Index: A Comparative Study of Power Relations across Historical Europe." *The History of the Family* 21 (2): 133–74. https://doi.org/10.1080/108160 2X.2014.1001769.

GSMA. 2019. "Connected Women: The Mobile Gender Gap Report." https://www.gsma.com/mobilefordevelopment/wp-content/uploads/2019/02/GSMA-The-Mobile-Gender-Gap-Report-2019.pdf.

Gupta, Kamla, and P. Princy Yesudian. 2006. "Evidence of Women's Empowerment in India: A Study of Socio-Spatial Disparities." *GeoJournal* 65 (4): 365–80. https://doi.org/10.1007/s10708-006-7556-z.

Gupta, Saini, and Shweta Gupta. 2020. "Just a Call Away? Learnings from Phone Surveys in India and Nepal." *SurveyCTO* (blog). 2020. https://www.surveycto.com/case-studies/ifpri-covid19-phone-surveys/

Heintz, James, Naila Kabeer, and Simeen Mahmud. 2018. "Cultural Norms, Economic Incentives and Women's Labour Market Behaviour: Empirical Insights from Bangladesh." *Oxford Development Studies* 46 (2): 266–89. https://doi.org/10.1080/13600818.2017.1382464.

Hnatkovska, Viktoria, Amartya Lahiri, and Sourabh Paul. 2012. "Castes and Labor Mobility." *American Economic Journal: Applied Economics* 4 (2): 274–307. https://doi.org/10.1257/app.4.2.274.

Hnatkovska, Viktoria, Amartya Lahiri, and Sourabh B. Paul. 2013. "Breaking the Caste Barrier: Intergenerational Mobility in India." *Journal of Human Resources* 48 (2): 435–73. https://doi.org/10.3368/jhr.48.2.435.

International Labour Organization(ILO). 1982. "Resolution Concerning Statistics of the Economically Active Population, Employment, Unemployment and Underemployment, Adopted by the Thirteenth International Conference of Labour Statisticians." In , Para 5. https://www.ilo.org/wcmsp5/groups/public/---dgreports/---stat/documents/normativeinstrument/wcms_087481.pdf.

———. 2018. "Statistics on Work Relationships, Report for Discussion at the Meeting of Experts on Labour Statistics in Preparation for the 20th International Conference of Labour Statisticians (Geneva, 5–9 February 2018)." MEPICLS/2018. https://www.ilo.org/wcmsp5/groups/public/---dgreports/---stat/documents/event/wcms_615360.pdf.

Jalan, Jyotsna and Rinku Murgai. 2015. "Intergenerational Mobility in Education in India." https://doi.org/10.13140/RG.2.1.4160.9124.

Jann, Ben. 2008. The Blinder-Oaxaca Decomposition for Linear Regression Models, *The Stata Journal*, 8(4): 453-479.

Jayachandran, Seema. 2020. "Social Norms as a Barrier to Women's Employment in Developing Countries." Working Paper. Working Paper Series. National Bureau of Economic Research. https://doi.org/10.3386/w27449.

Jayaraman, Vijayshree. 2023. "The Unhappy Marriage of Women's Work and Spousal Violence in India." Centre for Sustainable Employment Working Paper #52. Azim Premji University, Bangalore.

Jha, Mrinalini, and Amit Basole. 2022. "Labour Incomes in India: A Comparison of Two National Household Surveys,*The Indian Journal of Labour Economics*, 66(1), pp.181-201.

Jodhka, Surinder S. 2010. "Dalits in Business: Self-Employed Scheduled Castes in North-West India." *Economic and Political Weekly* 45 (11): 41–48.

Joshi, Shareen, Nishtha Kochhar, and Vijayendra Rao. 2018. "Are Caste Categories Misleading? The Relationship Between Gender and Jati in Three Indian States." In *Towards Gender Equity in Development*,

edited by Siwan Anderson, Lori Beaman, and Jean-Philippe Platteau, 0. Oxford University Press. https://doi.org/10.1093/oso/9780198829591.003.0016.

Kannan, KP, and G Raveendran. 2009. "Growth Sans Employment: A Quarter Century of Jobless Growth in India's Organised Manufacturing." *Economic and Political Weekly* 44 (10): 80–91.

Kapsos, Steven, Evangelia Bourmpoula, and Andrea Silberman. 2014. "Why Is Female Labour Force Participation Declining so Sharply in India?" ILO Research Paper No. 10. International Labour Organization. https://www.ilo.org/wcmsp5/groups/public/---dgreports/---inst/documents/publication/wcms_250977.pdf.

Kapur, Devesh, Milan Vaishnav, and Dawson Verley. 2021. "What Is Women's Work? Reflections from Four North Indian Urban Clusters." Urbanisation 6 (1_suppl): S20–39. https://doi.org/10.1177/24557471211038394.

Kaufman, Robert L. 2010. *Race, Gender, and the Labor Market: Inequalities at Work*. Boulder, Colo: Lynne Rienner Publishers.

Kesar, Surbhi. 2023. "Economic Transition, Dualism, and Informality in India: Nature and Patterns of Household-Level Transitions." *Review of Development Economics*, no. forthcoming.

Khanna, Madhulika, and Divya Pandey. 2021. "Reinforcing Gender Norms or Easing Housework Burdens? The Role of Mothers-in-Law in Determining Women's Labor Force Participation." Unpublished Manuscript.

Kishor, Sunita, and Kamla Gupta. 2004. "Women's Empowerment in India and Its States: Evidence from the NFHS." *Economic and Political Weekly* 39 (7): 694–712.

Klasen, Stephan, and Pieters Janneke. 2013. "What Explains the Stagnation of Female Labor Force Participation in Urban India?" *Courant Research Centre PEG*. https://ideas.repec.org/p/got/gotcrc/146.html.

Klasen, Stephan, and Janneke Pieters. 2015. "What Explains the Stagnation of Female Labor Force Participation in Urban India?" *World Bank Economic Review*, October. https://elibrary.worldbank.org/doi/10.1093/wber/lhv003.

Kleven, Henrik, Camille Landais, Johanna Posch, Andreas Steinhauer, and Josef Zweimüller. 2019. "Child Penalties across Countries: Evidence and Explanations." *AEA Papers and Proceedings* 109 (May): 122–26. https://doi.org/10.1257/pandp.20191078.

Kohli, Vaishali. 2023. "Association between Caste and Class in India: Evolution of Caste-Class Dynamics during Economic Growth." Centre for Sustainable Employment Working Paper #56. Azim Premji University, Bangalore.

Koolwal, Gayatri B. 2021. "Improving the Measurement of Rural Women's Employment: Global Momentum and Survey Priorities." *World Development* 147 (November): 105627. https://doi.org/10.1016/j.worlddev.2021.105627.

Lahoti, Rahul, and Hema Swaminathan. 2016. "Economic Development and Women's Labor Force Participation in India." *Feminist Economic*s 22 (2): 168–95. https://doi.org/10.1080/13545701.2015.1066022.

Lahoti, Rahul, Rosa Abraham, and Hema Swaminathan. 2021. "Childbirth and Women's Labour Market Transitions in India." WIDER Working Paper 2021. Vol. 2021. WIDER Working Paper. UNU-WIDER. https://doi.org/10.35188/UNU-WIDER/2021/068-9.

Langsten, Ray, and Rania Salem. 2008. "Two Approaches to Measuring Women's Work in Developing Countries: A Comparison of Survey Data from Egypt." *Population and Development Review* 34 (2): 283–305.

Lei, Lei, Sonalde Desai, and Reeve Vanneman. 2019. "The Impact of Transportation Infrastructure on Women's Employment in India." *Feminist Economic*s 25 (4): 94–125. https://doi.org/10.1080/13545701.2019.1655162.

Lundborg, Petter, Erik Plug, and Astrid Würtz Rasmussen. 2017. "Can Women Have Children and a Career? IV Evidence from IVF Treatments." *American Economic Review* 107 (6): 1611–37. https://doi.org/10.1257/aer.20141467.

Madheswaran, S, and Paul Attewell. 2007. "Caste Discrimination in the Indian Urban Labour Market: Evidence from the National Sample Survey." *Economic & Political Weekly* 42 (41): 4146–54.

Majumder, Rajarshi. 2007. "Earning Differentials across Social Groups: Evidences from India." SSRN Scholarly Paper 1291791. Rochester, NY. https://papers.ssrn.com/abstract=1291791.

———. 2010. "Intergenerational Mobility in Educational and Occupational Attainment: A Comparative Study of Social Classes in India." *Margin: The Journal of Applied Economic Research* 4 (4): 463–94. https://doi.org/10.1177/097380101000400404.

Mathur, Mitali Roy. 2020. "Reflections on Reaching Women over the Phone in Rural India." Medium. 2020. https://medium.com/idinsight-blog/reflections-on-reaching-women-over-the-phone-in-rural-india-29fffdbe296a.

Mehrotra, Santosh. 2019. "Inequalities in the Gendered Labour Market: What Can Be Done?" Centre for Sustainable Employment Working Paper #22. Azim Premji University, Bangalore. https://cse.azimpremjiuniversity.edu.in/publications/inequalities-in-the-gendered-labour-market-what-can-be-done/.

Mehrotra, Santosh, and Jajati K. Parida. 2017. "Why Is the Labour Force Participation of Women Declining in India?" *World Development* 98 (October): 360–80. https://doi.org/10.1016/j.worlddev.2017.05.003.

Mehrotra, Santosh, and Sharmistha Sinha. 2019. "Towards Higher Female Work Participation in India: What Can Be Done?" Centre for Sustainable Employment Working Paper#20. Azim Premji University, Bangalore. https://cse.azimpremjiuniversity.edu.in/publications/towards-higher-female-work-participation-in-india-what-can-be-done/.

Misra, Sangita, and Suresh. 2014. "Estimating Employment Elasticity of Growth for the Indian Economy." RBI Working Paper Series. https://rbi.org.in/Scripts/PublicationsView.aspx?id=15763.

Mondal, Bidisha, Jayati Ghosh, Shiney Chakraborty, and Mitra Mitra. 2018. "Women Workers in India: Labour Force Trends, Occupational Diversification and Wage Gaps." CSE Working Paper#3. Azim Premji University, Bangalore. https://publications.azimpremjiuniversity.edu.in/4334/1/Mondal_et_al_Women_Workers_In_India_March_2018.pdf.

Mondal, Dinabandhu, and Pintu Paul. 2021. "Associations of Power Relations, Wife-Beating Attitudes, and Controlling Behavior of Husband With Domestic Violence Against Women in India: Insights From the National Family Health Survey–4." Violence Against Women 27 (14): 2530–51. https://doi.org/10.1177/1077801220978794.

Morselli, Davide, Jean-Marie Le Goff, and Jacques-Antoine Gauthier. 2019. "Self-Administered Event History Calendars: A Possibility for Surveys?" Contemporary Social Science 14 (3–4): 423–46. https://doi.org/10.1080/21582041.2017.1418528.

Motiram, Sripad, and Ashish Singh. 2012. "How Close Does the Apple Fall to the Tree? Some Evidence from India on Intergenerational Occupational Mobility." Economic and Political Weekly 47 (40): 56–65.

Moya, Cristóbal. 2021. Inferential Statistics and Complex Surveys. https://bookdown.org/cristobalmoya/iscs_materials/.

Mukherjee, Anirban, and Sukanya Sarkhel. 2021. "Patriarchal Norms and Women's Labor Market Outcomes." SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3770404.

Muller, Miriam, and Liliana D. Sousa. 2020. "'She Helps Me All the Time': Underestimating Women's Economic Engagement in Rural Honduras," April. https://doi.org/10.1596/1813-9450-9217.

Nagin, Daniel S. 2005. Group-Based Modeling of Development: Cambridge, MA: Harvard University Press.

Nandwani, Bharti, and Punarjit Roychowdhury. 2023. "British Colonialism and Women Empowerment in India." GLO Discussion Paper Series, GLO Discussion Paper Series, . https://ideas.repec.org//p/zbw/glodps/1275.html.

Nayyar, Deepak. 2019a. "India's Path to Structural Transformation: An Exception and the Rule." In The Oxford Handbook of Structural Transformation, by Deepak Nayyar, edited by Célestin Monga and Justin Yifu Lin, 464–88. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198793847.013.20.

———. 2019b. Resurgent Asia: Diversity in Development. 1st ed. Oxford University PressOxford. https://doi.org/10.1093/oso/9780198849513.001.0001.

Neff, Daniel F., Kunal Sen, and Veronika Kling. 2012. "The Puzzling Decline in Rural Women's Labor Force Participation in India: A Reexamination." SSRN Electronic Journal. https://doi.org/10.2139/ssrn.2143122.

Nigam, Aditya. 2002. "In Search of a Bourgeoisie: Dalit Politics Enters a New Phase." *Economic and Political Weekly* 37 (13).

Oh, Suanna. 2023. "Does Identity Affect Labor Supply?" *American Economic Review* 113 (8): 2055–83. https://doi.org/10.1257/aer.20211826.

Paliath, Shreehari. 2021. "How Unreliable Data On Dalit Christians & Muslims Expose Them To Discrimination." *IndiaSpend*, March 15, 2021, sec. Data Gaps. https://www.indiaspend.com/data-gaps/how-unreliable-data-on-dalit-christians-muslims-expose-them-to-discrimination-734972.

Papola, TS, and Sahu. 2012. "Growth and Structure of Employment in India Long-Term and Post-Reform Performance and the Emerging Challenge." New Delhi: Institute for Studies in Industrial Development.

Paul, Sohini. 2016. "Women's Labour Force Participation and Domestic Violence: Evidence from India." *Journal of South Asian Development* 11 (2): 224–50. https://doi.org/10.1177/0973174116649148.

Perez, Caroline Criado. 2021. Invisible Women. New York, NY: Harry N. Abrams.

Raj S. N., Rajesh, and Subash Sasidharan. 2018. "Does the Caste of the Firm Owner Play a Role in Access to Finance for Small Enterprises? Evidence from India: CASTE OF FIRM OWNER AND ACCESS TO FINANCE." *The Developing Economies* 56 (4): 267–96. https://doi.org/10.1111/deve.12183.

Rammohan, Anu, Srinivas Goli, and Bheemeshwar Reddy. 2017. "Occupational Segregation by Caste and Gender in India." MPRA Paper No. 101969. https://mpra.ub.uni-muenchen.de/id/eprint/101969.

Rangarajan, C, Padma Iyer Kaul, and Seema. 2011. "Where Is the Missing Labour Force?" *Economic and Political Weekly* 46 (39). https://www.epw.in/journal/2011/39/special-articles/where-missing-labour-force.html.

RCRC. 2021. "RCRC Household Survey Round3." https://www.rcrc.in/wp-content/uploads/2021/03/RCRC_householdsurvey_round3.pdf.

Reddy, A. Bheemeshwar. 2015. "Changes in Intergenerational Occupational Mobility in India: Evidence from National Sample Surveys, 1983–2012." *World Development* 76 (December): 329–43. https://doi.org/10.1016/j.worlddev.2015.07.012.

Reskin, B.F. 2001. "Sex Segregation at Work." *In International Encyclopedia of the Social & Behavioral Sciences*, 13962–65. Elsevier. https://doi.org/10.1016/B0-08-043076-7/03994-2.

Sadana, Nidhi, and Sukhadeo Thorat. 2009. "Caste and Ownership of Private Enterprises." *Economic and Political Weekly* 44 (23): 13–16.

Schmitz, Sophia, and C Katharina Spiess. 2022. "The Intergenerational Transmission of Gender Norms—Why and How Adolescent Males with Working Mothers Matter for Female Labour Market Outcomes." Socio-Economic Review 20 (1): 281–322. https://doi.org/10.1093/ser/mwab023.

Sengupta, Anindita, and Panchanan Das. 2014. "Gender Wage Discrimination across Social and Religious Groups in India: Estimates with Unit Level Data." *Economic and Political Weekly* 49 (21): 71–76.

Sengupta, Nandana, Sarthak Gaurav, and James Evans. 2021. "The Skills Space in Informal Work: Insights from Bangalore Slums." *The Journal of Development Studies* 57 (10): 1662–89. https://doi.org/10.1080/0022038 8.2021.1898593.

Shah, Arpit, Deepak Malghan, Sneha Thapliyal, Anish Sugathan, and Vimal Mishra. 2023. "Social Demography and Climate Inequality in India." Paper Under Review.

Singh, Abhishek, Praveen Chokhandre, Ajeet Kumar Singh, Kathryn M. Barker, Kaushalendra Kumar, Lotus McDougal, K. S. James, and Anita Raj. 2022. "Development of the India Patriarchy Index: Validation and Testing of Temporal and Spatial Patterning." *Social Indicators Research* 159 (1): 351–77. https://doi. org/10.1007/s11205-021-02752-1.

Singh, S.K., Bhawana Sharma, Deepanjali Vishwakarma, Gudakesh Yadav, Swati Srivastava, and Barsharani Maharana. 2019. "Women's Empowerment and Use of Contraception in India: Macro and Micro Perspectives Emerging from NFHS-4 (2015–16)." *Sexual & Reproductive Healthcare* 19 (March): 15–23. https://doi. org/10.1016/j.srhc.2018.11.003.

Sinha, Babita, Smita Jha, and Nalin Singh Negi. 2012. "Migration and Empowerment: The Experience of Women in Households in India Where Migration of a Husband Has Occurred." *Journal of Gender Studies* 21 (1): 61–76. https://doi.org/10.1080/09589236.2012.639551.

Small, Sarah F., and Yana Van Der Meulen Rodgers. 2023. "The Gendered Effects of Investing in Physical and Social Infrastructure." *World Development* 171 (November): 106347. https://doi.org/10.1016/j. worlddev.2023.106347.

Somanchi, Anmol. 2021. "Missing the Poor, Big Time: A Critical Assessment of the Consumer Pyramids Household Survey." SocArXiv. https://doi.org/10.31235/osf.io/qmce9.

Srinivas, M. N. 1977. "The Changing Position of Indian Women." Man 12 (2): 221. https://doi. org/10.2307/2800796.

"State of Working India 2018." 2018. Centre for Sustainable Employment,Azim Premji University. https://cse. azimpremjiuniversity.edu.in/state-of-working-india/swi-2018/

"State of Working India 2019." 2019. Centre for Sustainable Employment,Azim Premji University. https://cse. azimpremjiuniversity.edu.in/state-of-working-india/swi-2019/.

"State of Working India 2021." 2021. Centre for Sustainable Employment,Azim Premji University. https://cse. azimpremjiuniversity.edu.in/state-of-working-india/swi-2021/

Tejani, Sheba. 2016. "Jobless Growth in India: An Investigation." *Cambridge Journal of Economics* 40 (3): 843–70. https://doi.org/10.1093/cje/bev025.

Teltumbde, Anand. 2010. The Persistence of Caste: *The Khairlanji Murders and India's Hidden Apartheid*. New Delhi: Navayana.

The Ministry of Statistics & Programme Implementation (MoSPI). 2023. "National Policy on Official Statistics (NPOS)." https://www.mospi.gov.in/sites/default/files/announcements/Draft_Revised_NPOS.pdf

Thorat, Sukhadeo, and Khalid Khan. 2023. "Why Inter-Caste Inequality in Educational Attainment Still Persists? Wealth and Caste Discrimination Matters." *The Indian Economic Journal* 71 (1): 256–75. https://doi.org/10.1177/00194662221146656.

Thorat, Sukhadeo, and S. Madheswaran. 2018. "Graded Caste Inequality and Poverty: Evidence on Role of Economic Discrimination." *Journal of Social Inclusion Studies* 4 (1): 3–29. https://doi.org/10.1177/2394481118775873.

Varshney, Ashutosh, Lakshmi Iyer, and Tarun Khanna. 2013. "Caste and Entrepreneurship in India." *Economic and Political Weekly* 48 (06): 52–60.

Virtanen, Pekka, Liudmila Lipiäinen, Anne Hammarström, Urban Janlert, Antti Saloniemi, and Tapio Nummi. 2011. "Tracks of Labour Market Attachment in Early Middle Age: A Trajectory Analysis over 12 Years." *Advances in Life Course Research* 16 (2): 55–64. https://doi.org/10.1016/j.alcr.2011.03.001.