Profiling Female Exclusion from Labour Markets: A Latent Class Approach

Delhi School of Economics Winter School 2021

Sankalp Sharma[1] Sowmya Dhanaraj[2] Vidya Mahambare[3]

[1] Development Data Lab / Madras School of Economics

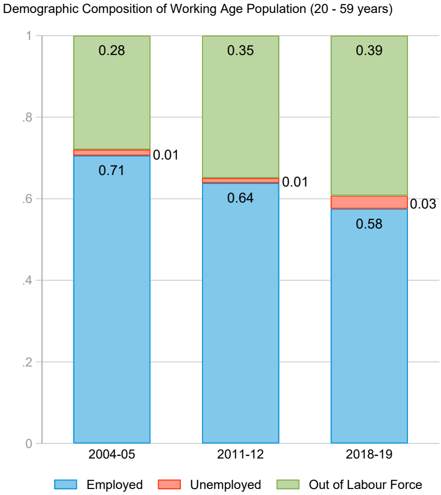
[2] Madras School of Economics

[3] Great Lakes Institute of Management

December 17, 2021

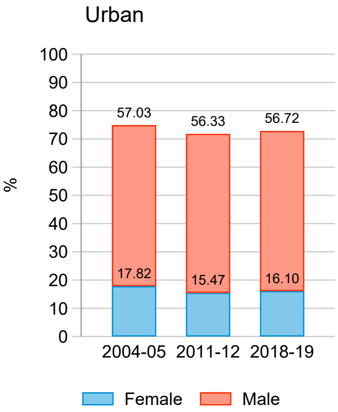
Context: Indian Workforce in the 21st Century

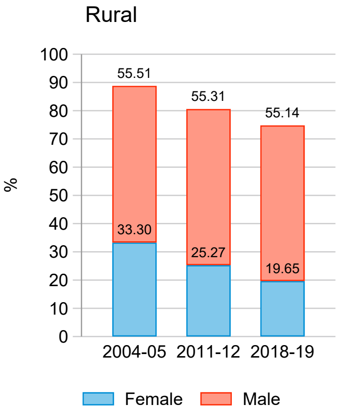
- Consistent decline in labour force participation rates despite having the second highest working age population in the world
  - 71% in 2004-05 $\rightarrow$ 58% in 2018-19
- Medium term economic prospects $\rightarrow$ trajectory of post-COVID job recoveries (Basole (2020), Deshpande (2020), Agarwal (2021))
- Critically important to identify vulnerable workers in the labour market



Declining Female Labor Force Participation

- Consistent decline in female labor force participation rates
- Decline in FLFP more severe in rural areas (Pande and Fletcher (2018))





Source: NSS61, NSS68, PLFS 2018-19

Motivation

- Vast dissimilarities in the gender composition of those not participating in labor markets (OLF)

Table: Duncan Dissimilarity Index for gender segregation among OLF

Year	2004	2011	2018
D-index	0.434	0.487	0.513

- Females face **excess burden** of being excluded from the labour market
- 57% women in India do not participate in the labour market (PLFS 2018-19)
- Vast heterogeneity in demographic characteristics of females excluded from the labour market, **making inclusionary policy targeting difficult and ineffective**

Policy Problem

- Labour market policies uninformed by typologies of targeted females $\rightarrow$ expected outcomes remain unachieved
- Awareness of heterogeneities in OLF females is important for atleast three immediate reasons:
  - Greater ease in policy targeting
  - Tailoring policies and programs for specific typologies of OLF women
  - Focused prioritization of policies that can reintegrate females into the labor market at the lowest public expenditure

Latent Class Approach

1. Multidimensional Profiling of Female Exclusion
2. Latent Class Model
3. Data and Covariates
4. Optimal Latent Classes

Multidimensional Profiling of Female Exclusion

- Observed characteristics of females are used to enable identification of latent classes that capture different dimensions of exclusion
- **Person-oriented approach**: Group sub-types of females that exhibit similar patterns of characteristics using a probabilistic framework
- Statistically, we compute the score for each female according to the likelihood of belonging to each of the computed latent classes (Bergman and Magnusson (1997))
  - Subsequently assigned to the latent class to which they have the highest posterior probability of belonging
- Frontier research on labour market segmentation is increasingly using variations of latent class approach (Lukac (2019), Yoon and Chung (2015))

Setting up the Latent Class Model

- Let $j = 1, \ldots, J$ be the observed indicator, and observed indicator has j has $r_j = 1, \ldots, R_j$ response categories.
- L is categorical latent variable with $c = 1, \ldots, C$ latent classes.
- $P[L = c] = \gamma_c$
- Item-response probability of indicator j is referred to as ρ_j

As latent classes are mutually exclusive and exhaustive, each female is a member of one and only one latent class

$$\sum_{c=1}^{C} \gamma_c = 1 \tag{1}$$

Item-response probability of a given indicator conditional on latent class is unity (since each individual provides a unique response to an indicator)

$$\sum_{r_j=1}^{R_j} \rho_{j,\gamma_j|c} = 1 \tag{2}$$

Estimation

- Let y_j represent element j of a response pattern $\mathbf{y}$.
- What is $P(Y = \mathbf{y}|L = c)$?

Let $I(Y_j = \gamma_j) = 1$ when $j = \gamma_j$ and 0 otherwise

Probability of observing a particular vector of response is a function of the probabilities of membership in each latent class c and the probability of observing each response conditional on latent class membership

$$P(Y = y|L = c) = \prod_{j=1}^{J} \prod_{\gamma_j=1}^{R_j} \rho_{j,\gamma_j|c}^{I(y_j=r_j)} \tag{3}$$

Parameters γ and ρ are estimated via **maximum likelihood**

Posterior Assignment

Using Bayes Theorem, we can calculate the posterior probability of belonging in latent class c conditional on the observed response pattern $\mathbf{y}$

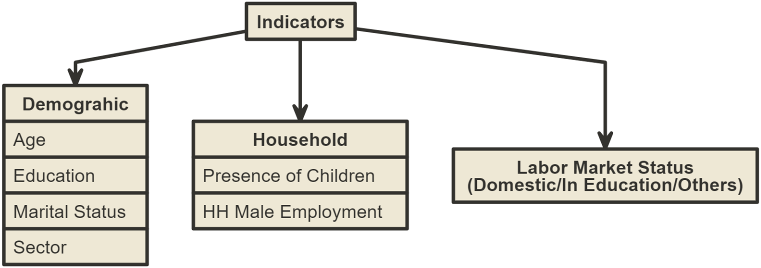
$$P(L = c \mid Y = \mathbf{y}) = \frac{P(Y = y \mid L = c)P(L = c)}{P(Y = y)} \tag{4}$$

Since we know that $P(L = c) = \gamma_c$,

$$P(L = c \mid Y = \mathbf{y}) = \frac{\left(\prod_{j=1}^{J} \prod_{\gamma_j=1}^{R_j} \rho_{j,\gamma_j|c}^{I(y_j=\gamma_j)}\right) \gamma_c}{\sum_{c=1}^{C} \gamma_c \prod_{j=1}^{J} \prod_{\gamma_j=1}^{R_j} \rho_{j,\gamma_j|c}^{I(y_j=\gamma_j)}} \tag{5}$$

Data and Key Covariates

- **Data**: NSS Employment Survey 2004-05 and Periodic Labour Force Survey 2018-19 ($N = 1,023,590$)
- **Sub-sample**: Females (20-59 years) out of labour force ($N = 164,214$)
- **Indicator variables**:



These indicators collectively constitute the response pattern vector **y** for each female in the sample

Determination of Optimal Latent Classes

- Models with upto 9 latent classes are estimated
- To compare model fit and parsimony → **Akaike information criterion (AIC)** and **Bayesian Information criterion (BIC)**
- Smaller values of AIC and BIC represent optimal balance of model fit and parsimony
- Using elbow test, **4 classes** are optimal for 2004-05 and **6 classes** for 2018-19

# Classes	2004-05				2018-19			
	AIC	BIC	$\Delta(AIC)$	$\Delta(BIC)$	AIC	BIC	$\Delta(AIC)$	$\Delta(BIC)$
2	1.31×10^9	1.31×10^9			2.38×10^9	2.38×10^9		
3	1.25×10^9	1.25×10^9	-4.1%	-4.1%	2.27×10^9	2.27×10^9	-4.7%	-4.7%
4	1.21×10^9	1.21×10^9	-3.2%	-3.2%	2.17×10^9	2.17×10^9	-4.3%	-4.3%
5	1.21×10^9	1.21×10^9	-0.6%	-0.6%	2.14×10^9	2.14×10^9	-1.4%	-1.4%
6	1.18×10^9	1.18×10^9	-1.7%	-1.7%	2.11×10^9	2.11×10^9	-1.4%	-1.4%
7	1.16×10^9	1.16×10^9	-2.2%	-2.2%	2.1×10^9	2.1×10^9	-0.3%	-0.3%
8	1.15×10^9	1.15×10^9	-0.5%	-0.5%	2.09×10^9	2.09×10^9	-0.7%	-0.7%
9	1.16×10^9	1.16×10^9	0.6%	0.6%	2.08×10^9	2.08×10^9	-0.4%	-0.4%

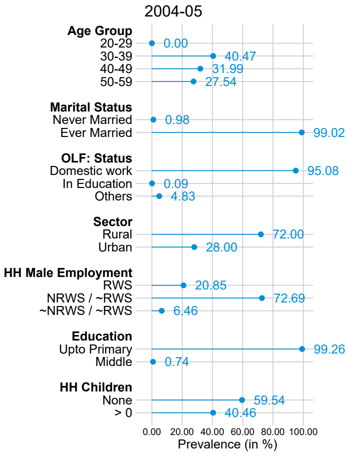
Latent Classes

2004-05	2018-19
Elder, Married and Unlettered (EMU)	Elder, Married and Unlettered (EMU)
Young, Married and Unlettered (YMU)	Young, Married and Unlettered (YMU)
Middle-aged, married, urban and financially secured (MFS)	Middle-aged, married, urban and financially secured (MFS)
Urban, young and educated (UY)	Middle-aged, Married and Unlettered (MMU)
	Young and married with HH dependents (YEH)
	Young women in education (YE)

Our focus: Similar latent classes across time (EMU, YMU, MFS)

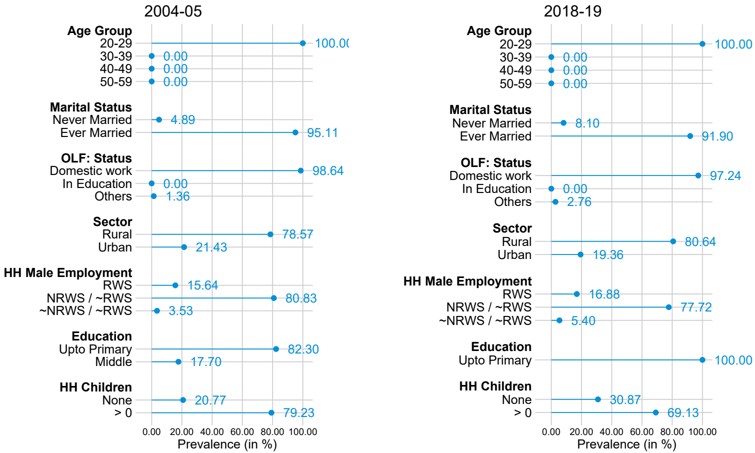
Comparison: Elderly, Married and Unlettered (EMU)

Prevalence in 2004-05: 43.6%; 2018-19: 13.2%



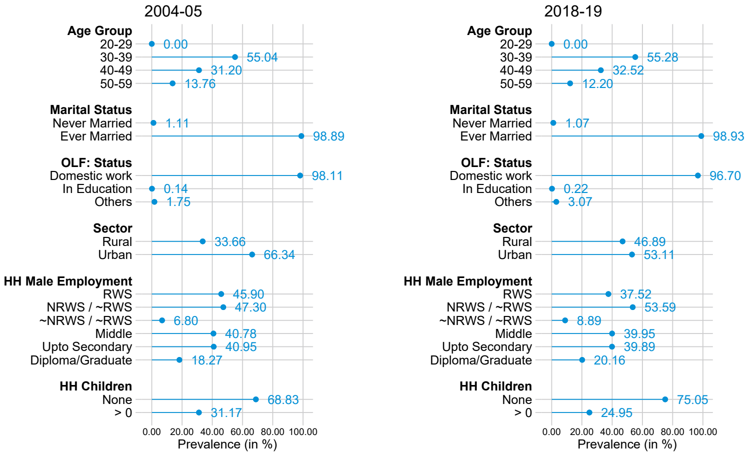
Comparison: Young, Married and Unlettered (YMU)

Prevalence in 2004-05: 24.8%; 2018-19: 9.8%



Comparison: (MFS)

Prevalence in 2004-05: 17.9%; 2018-19: 26.1%



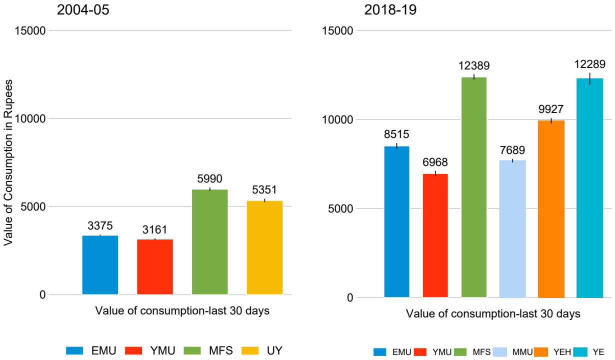
Quantifying Vulnerability: Axes of Exclusion

- 3 axes of exclusion:
  - Economic deprivation: lack of access to goods
  - Social marginalization
  - Location barrier
- Identifying vulnerability:
  - Compare average HH consumption across latent classes
  - Compare prevalence in social groups (ST/SC/OBC/Others)
  - Between and within-state variation

Household Consumption

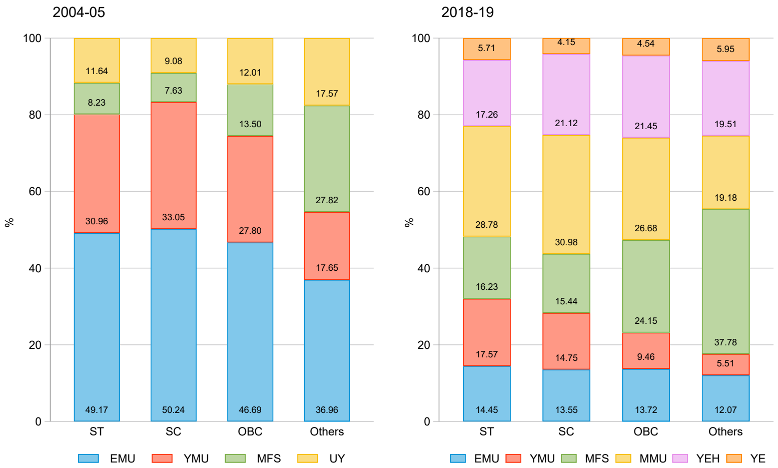
- Young, married and unlettered women (YMU) have lowest average HH consumption
- Average HH consumption increased for EMU but still low in 2018-19

Average HH Consumption by Latent Class



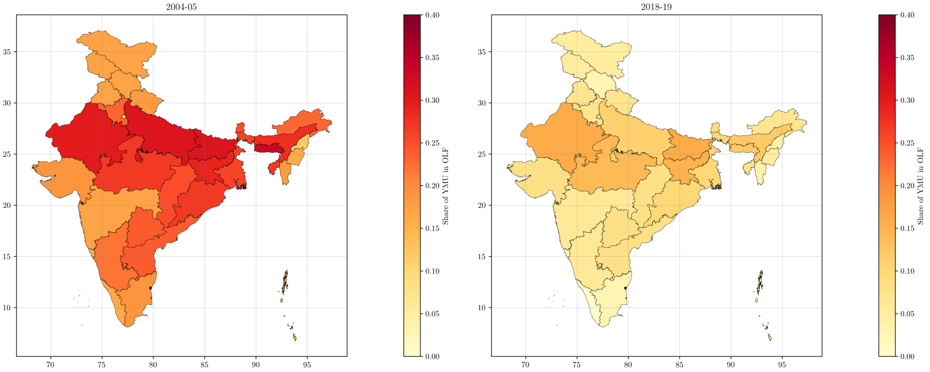
Social Group

Composition of Social Groups by Latent Classes

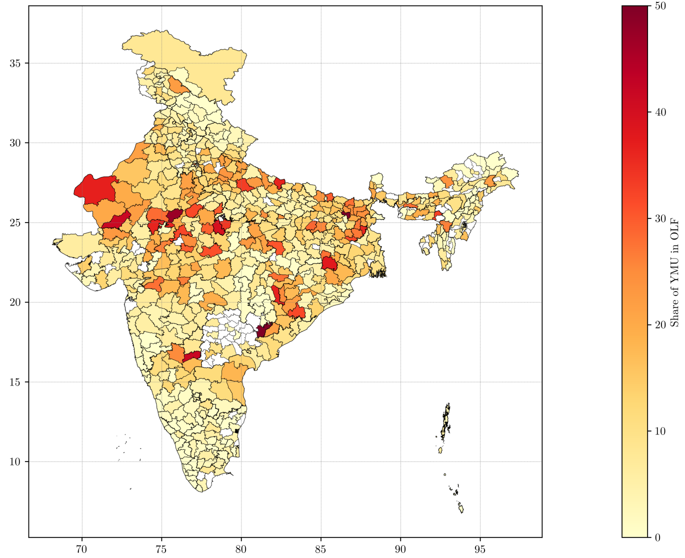


Spatial Distribution: Share of Vulnerable Latent Classes

- Aggregate decline in share of socio-economically vulnerable latent classes over time

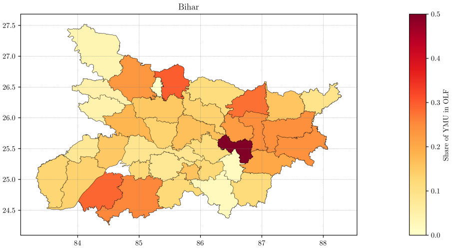


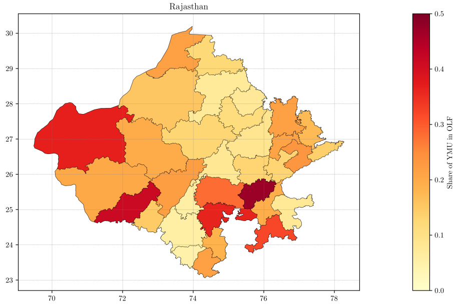
Spatial Distribution: District-wise Variation



Spatial Distribution: Within-State Variation

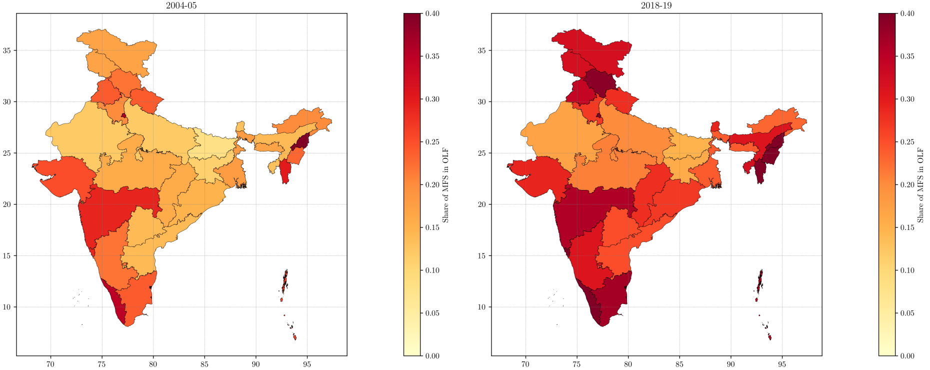
- Large within-state variation in prevalence of vulnerable latent classes





Spatial Distribution: MFS

- Rise in MFS: more difficult to reintegrate in the labour market due to high substitution effects



Key Takeaways

- Understanding composition of females not participating in the labor market is important for two reasons:
  - Demographically and socio-economically heterogenous
  - Prevalence of each sub-type varies between and within state
- Large part of rise in declining FLFP can be attributed to the rise in MFS women
  - High substitution effects within household
  - More difficult to integrate due to lesser incentives to work
- Latent class approach can give us actionable insight to understand policy-relevant compositional differences

Policy Recommendations

- EMU: Better social safety nets for women in regions with high prevalence of EMU women
- YMU: Increased public expenditure on public works programs in areas with high prevalence of YMU women
- Specific policy attention needed for reintegrating MFS women:
  - Targeted skill development
  - Improved design of market incentives
- Reduce barriers to female mobility

Thank you!