Job Opportunities along the Rural-Urban Gradation And Female Labor Force Participation in India

Urmila Chatterjee, Rinku Murgai and Martin Rama

The World Bank *

Draft, December 16, 2014

Abstract

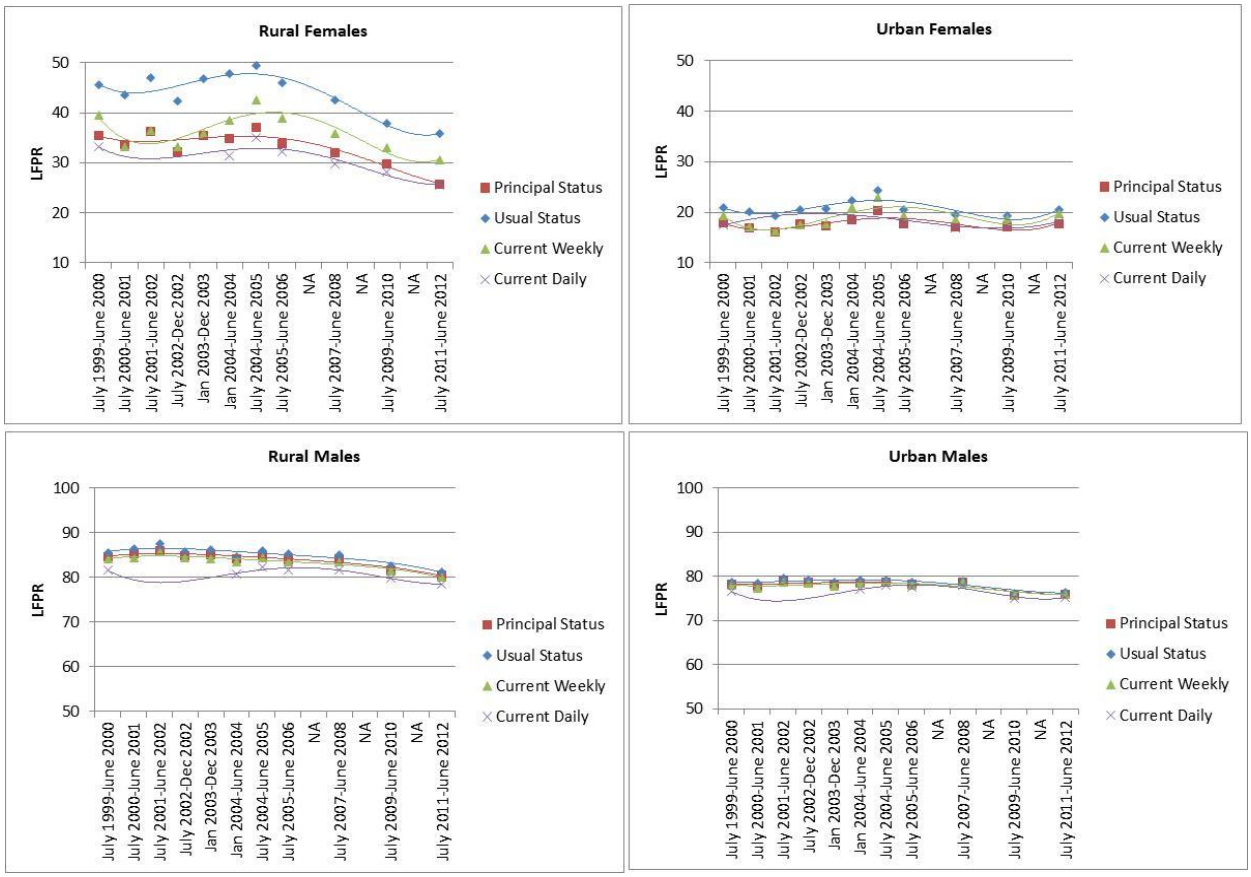
The recent decline in India's rural female labor force participation is generally attributed to higher household incomes. Together with the growing share of the urban population, where female participation rates are lower, this alleged income effect does not bode well for the empowerment of women as the country develops. This paper argues that such a narrow supply-side interpretation is wrong, because it ignores the transformation in the structure of employment at local levels. A salient trait of this period is the collapse in the number of farming jobs without a parallel emergence of non-farm regular jobs and other employment opportunities considered suitable for women. The paper develops a novel approach to capture the structure of employment at the village or town level, along seven ranks in the rural-urban gradation. It also takes into account the possible misclassification of urban areas as rural, as a result of survey instruments lagging behind India's rapid urbanization process. When using an empirical specification that includes both supply- and demand-side factors, the alleged income effect vanishes. For a similar employment structure, shifting to bigger villages and towns does not affect female labor force participation either. The results are robust to changes in the definition of employment and to the introduction of a range of additional control variables. Simulations suggest that for India to revert the decline in female labor force participation rates it needs to boost job creation, and especially the creation of non-farm regular employment.

* The corresponding author is Urmila Chatterjee (email: uchatterjee@worldbank.org). The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

1. Introduction

Female labor force participation in India is unusually low. According to the International Labor Organization's Global Employment Trends Report 2013, India ranks 120 among 131 countries. Even within South Asia, India ranks 6th among eight countries, just above Pakistan and Afghanistan (World Bank, 2012). Two other intriguing patterns stand out when analyzing the labor force participation rate (LFPR) of women aged 15 years and above in India (figure 1).[1] One is the significant gap between rural and urban areas; the other is the dramatic drop in female LFPR in rural areas.

Figure 1: Female labor force participation is low in urban areas and is declining in rural areas



[1] The sources of data are the consumption expenditure surveys and the employment and unemployment surveys of the National Sample Survey (NSS). The trends are presented for three definitions of LFPR used in the NSS; Principal Status implies at least 3 months of work over the past year, Usual Status at least 1 month of work over the past year, Current Weekly at least one hour of work over the past week. A detailed description of the definitions is provided in Appendix A of this paper.

The decline in rural female LFPR has received considerable attention in academic research and in the public debate in recent years. But it is intuitively clear that this decline ought to be connected to the gap in LFPR between urban areas. And the two patterns taken together are a matter for concern. As argued in the recent *Gender* and *Jobs* World Development Reports (World Bank, 2011 and 2012) gainful work by women, and especially paid employment, are correlated with agency and contribute to stronger development outcomes. In a context of rapid urbanization, these two patterns do not bode well for women's empowerment in India.

Some decline in LFPR can be expected with development, as young people spend more time in school and old people can afford to stop working at earlier ages. In India, the 3-4 percentage point decline in LFPR for rural men over the period 1999-00 to 2011-12 is consistent with such prediction. But the magnitude of the decline was much larger for rural women, with their LFPR falling by 8-10 percentage points regardless of the measure of employment used. The decline was particularly pronounced after 2004-05, when the female LFPR fell by 12-13 percentage points, in contrast with the much steadier and slower decline in the male LFPR. In urban areas, there was little change in the LFPR of both males and females in this period. But the rural-urban gap in LFPR for females was consistently higher than for males (15-25 percentage points versus 6-10 percentage points).[2]

The conventional wisdom is that these patterns are driven by "supply side" factors, with income gains from development gradually allowing more women to stay at home, a preferred choice in a culturally traditional environment. Authors like Olsen et al. (2006), Chowdhury (2011), and Neff et al. (2012) have argued that social and cultural barriers in a predominantly patriarchal society like India can explain women's work choices. A large body of literature is devoted to quantifying the role of education and income in explaining the decline in female LFPR in rural areas, including studies by Chowdhury (2011), Himanshu (2011), Rangarajan et al. (2011), Kannan et al. (2012), Neff et al. (2012), Abraham (2013), and Klassen and Peters (2013).

It has also been argued that the lack of crèches and institutional child support for working women contributes to the decline in female LFPR. As the number of multi-generational households shrinks, women with young children have no choice but to stay at home.

Supply-side explanations for the decline in female LFPR are not totally convincing, however. While more schooling and the improvement in living standards can account for some of the decline of LFPR in rural areas, the empirical evidence provided so far is not conclusive. Almost

[2] Some authors have argued the spike in LFPR in 2004-05 does not square with the general economic trends in the period preceding the survey. Among them are Sundaram and Tendulkar (2006), Unni and Raveendran (2007) and Chandrashekhar and Ghosh (2007). But the data from the thin year rounds of the NSS shows that LFPR of females was increasing at a steady rate in the years between 1999-00 and 2004-05 and dropped steadily thereafter, thus weakening the argument that the NSS 2004-05 survey results are an outlier.

all the studies mentioned above use the same data namely, the NSS employment and unemployment surveys, but they fail to arrive at a consensus on the magnitude of effects. While higher school enrollment and staying longer in school can explain the LFPR decline for younger women, they cannot explain the drop in participation among older aged cohorts. And the decline in female LFPR between 2004-05 and 2011-12 has been the largest for the age group 30-39 years, a group that is clearly out of school (figure 2).

Figure 2: Much of the decline in female labor force participation is for prime working ages

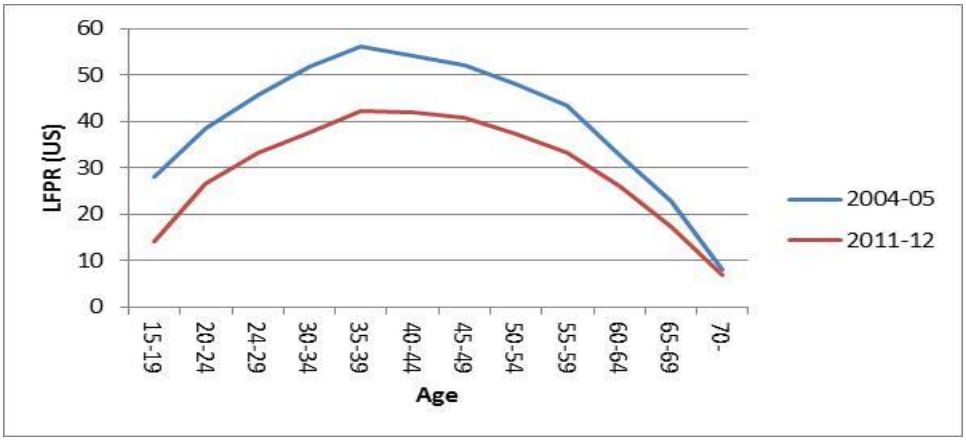
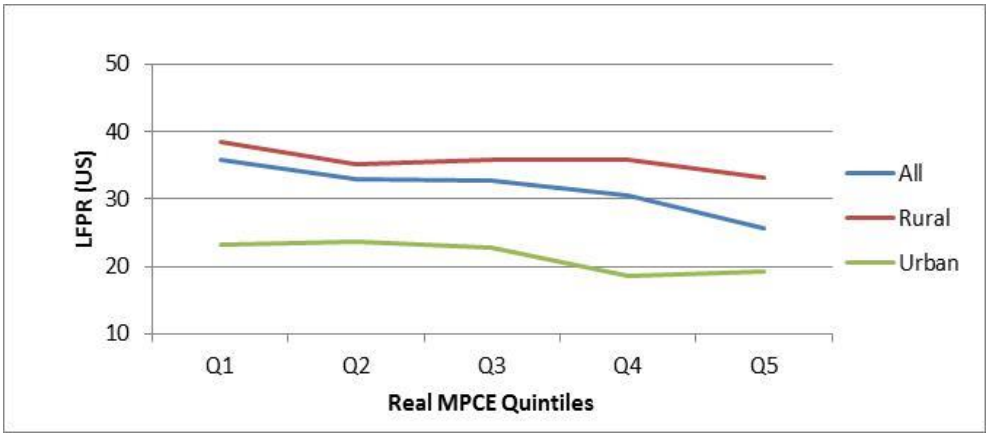


Figure 3: For the same household income, urban female labor force participation is lower



The income-effect hypothesis is further weakened by the fact that there is a large gap in female LFPR between rural and urban households with similar living standards. Monthly expenditures per capita provide a defensible proxy for household income. If all households are classified by

quintiles of expenditure per capita, independently of where they live, LFPR for is higher in rural areas in every quintile (figure 3)[3].

An additional reason to be skeptical about the income effect-hypothesis comes from short-term economic fluctuations. The drought of 2009-10 was the worst in three decades, and it must have resulted in a decline in household income in rural areas in spite of the support provided by transfers programs such as the Mahatma Gandhi Rural Employment Guarantee (MNREGA). Yet, female LFPR did not increase in that year but rather fell dramatically in rural areas (Himanshu, 2011).

A few studies have paid attention to "demand-side" factors to explain the observed decline in female LFPR. The implicit hypothesis is that the number and type of jobs available matters. According to Hirway (2012) a sizable part of female employment is related to home based and subsidiary work, which are not adequately captured by the NSS. Klassen et al. (2013) focus on female LFPR in urban areas and find that the decline is explained by a combination of rising household incomes and declining white collar jobs, especially for educated women. Rodgers (2012) examines rural Bihar and finds that the decline in female LFPR can be attributed to limited job opportunities for women outside agriculture. Kannan et al. (2012) and Chand et al. (2014) suggest that the decline in female LFPR could be due to poor agricultural performance and the diversification of jobs in rural areas.

Last but not least, in a context of rapid urbanization there is a potential misclassification of spatial units. Much of the growth in India's urban population in recent years has taken place in administratively rural areas. The multiplication of "census towns" (formally panchayats, but with urban characteristics) over the last decade is a testimony to this trend4. The census categorizes urban areas into Statutory Towns (ST) and Census Towns (CT). STs are places with a municipality, corporation, cantonment board or notified town area committee. CTs are places with a minimum population of 5000, a density of at least 400 per sq. km. and at least 75% of male main workers are engaged in non-farm work. Most countries use population size to distinguish between rural and urban areas. Based on this international comparison it can be argued that some of the criteria used by the census in India to identify census towns are too stringent. If only population size was used to identify urban areas in India, then the villages in the census with more than 5,000 inhabitants would be considered urban, and the share of the rural population would decrease by close to 15 percentage points. While the census may lag behind reality in capturing the reality of urbanization, the NSS lags behind the census in

[3] The Tendulkar poverty lines are used to correct for spatial price differences and to express real monthly per capita expenditure in All India Rural Rupees.

[4] There were 2,542 Census Towns and 242 Statutory Towns created between 2001 and 2011.The total urban population share went up by an unprecedented 31.8% between 2001 and 2011. 36.3% of this increase came from CTs and the rest from STs

reclassifying rural areas as urban. The latest survey of the NSS on employment and unemployment, conducted in 2011-12, used the census 2001 as the rural sampling frame. As mentioned before there was unprecedented growth in Census Towns in the period 2001 to 2011 in India. Thus, the share of urban population may be underestimated by a considerable proportion in the NSS. Part of the apparent decline in rural female LFPR could thus be a composition effect, reflecting urban outcomes in allegedly rural areas.

The goal of this paper is to take a fresh look at female LFPR in India factoring in not only the standard supply-side considerations, but also the demand-side and the urbanization perspectives. This broader, contrarian view is intuitively articulated in Section 2. Section 3 describes the model, data and methodology used in the analysis. The econometric results are presented in Section 4. Section 5 assesses the robustness of the analysis. Section 6 brings together the main findings by decomposing the decline in LFPR between income effects, household composition effects, employment effects, urbanization effects and measurement bias, along the entire rural-urban gradation. The results suggest that the changing nature of local employment opportunities as India urbanizes accounts for much of the observed decline in female LFPR.

2. A contrarian view

As in other parts of the world, the rapid urbanization process of India has been accompanied by a massive decline in the number of agricultural jobs. What is perhaps more unusual about India is that this structural transformation has not been associated with a substantial increase in manufacturing jobs, or in wage employment. Instead, employment in the services sector has surged, as also has the number of casual jobs. Most of this expansion of employment out of agriculture has been in informal activities.

The central hypothesis of this paper is that the rapid change in the composition of employment, and the way it materialized along the rural-urban gradation, are fundamental to understand why female LFPR is so low in India, and why it is declining. This central hypothesis can be seen as a contrarian view, because it attributes much less importance to income effects or the lack of crèches than the dominant view does.

Casual jobs and informality make the measurement of employment, and hence the assessment of employment opportunities at the local level, particularly challenging. The NSS defines participation and work in four different ways (see Appendix A for a detailed description). In this paper we focus on two definitions that come up often in the discussion on measurement of jobs and their quality, namely principal status and subsidiary status. Principal status (PS) applies to a person who participates in the labor market for at least 6 months and is employed

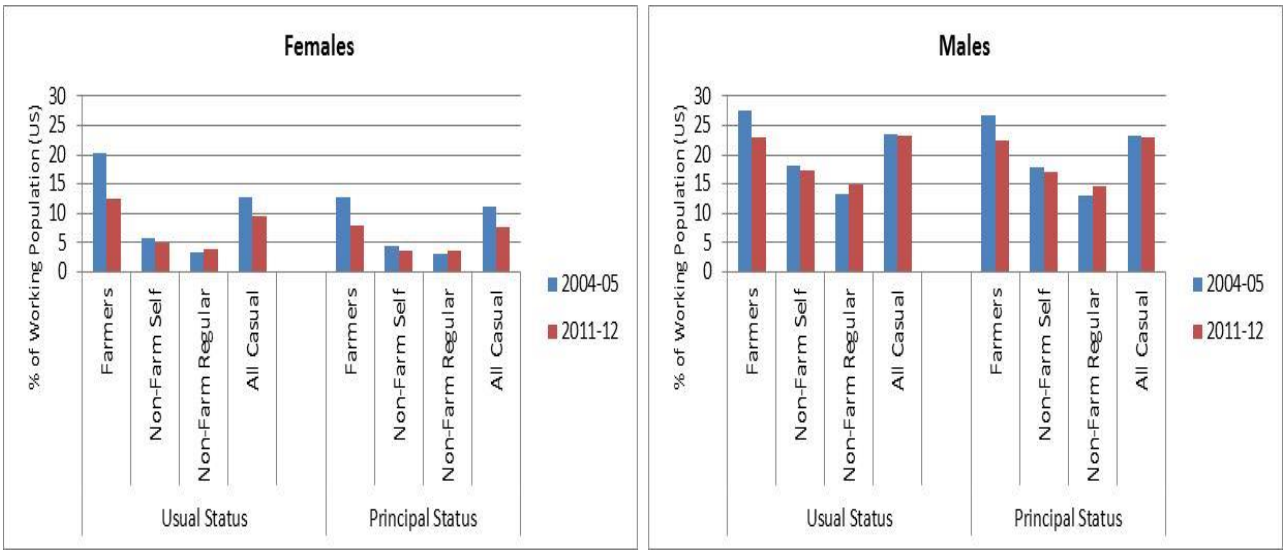
for at least 3 months over the past year. Subsidiary status (SS) refers to a person who is employed for at least 1 month. Usual Status (US) refers to a person who falls under either the PS or SS category. Thus, the US definition provides a broad measure of participation and employment, whereas PS can be seen as a proxy for steady and regular employment.

The definition of employment matters more for women than for men. To characterize the employment opportunities available at the local level, in this paper we combine information on activity status and sector of work. This allows us to define four types of jobs: farmers, non-farm self-employed, non-farm regular wage workers, and all casual wage workers. Employment shares can be computed for each of these categories as the ratio of working-age people holding each of the four types of jobs relative to the working-age population. For men, employment shares are almost identical over the period 2004-05 to 2011-12 when using the PS and US definitions (figure 4).[5] For women, however, the US definition results in higher levels of employment, and therefore in higher LFPR. This difference between the US and PS definitions reflects the greater proportion of women engaged in subsidiary work, especially in farm and casual jobs.

Employment shares also reveal a greater diversification of jobs for men than for women, and this under both the US and the PS definitions. In rural areas, women's employment is concentrated in farming and casual jobs. In urban areas fewer women work; but those who do are typically in non-farm self-employment or in regular employment (mostly in services). This greater concentration of employment among women matters, because from 2004-05 to 2011-12 female farming jobs declined much more rapidly than male farming jobs. Moreover, female farming jobs declined much more under the US definition. There was also a decline in the share of casual jobs for women in rural areas, accompanied by a slight increase in the share of non-farm regular jobs, for both men and women, under the PS definition. But the increase took place mainly in urban areas and was not large enough to offset the decline in farming jobs.

[5] The share of the unemployed is marginal. Thus, working-age people who are out of the labor force account for the bulk of the left-out group.

Figure 4: The decline in farming jobs was not offset by an increase in non-farm jobs



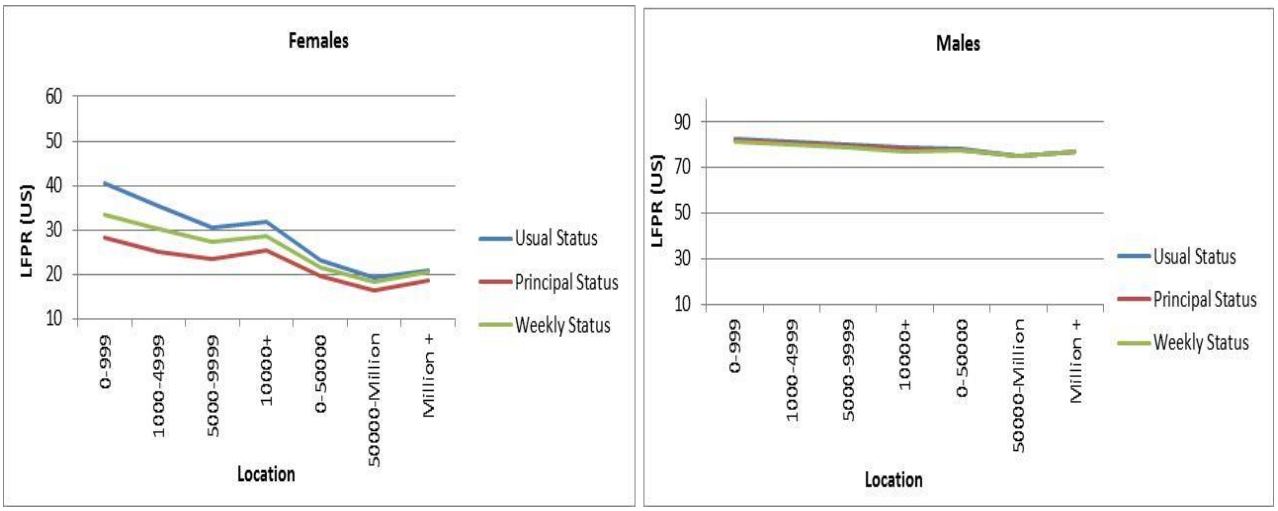
The change in the structure of employment at local levels was different along the rural-urban gradation. While most analyses (including the discussion in the introduction to this paper) consider a rural-urban divide, this dichotomous approach results in an important loss of granularity. Urbanization in developing countries has been driven to a large extent by the rapid development of peri-urban areas, and the densification of formally rural areas. This is even more so in India, where relatively weak city governance has resulted in a particularly "messy" urbanization process.

To better capture this granularity of the rural-urban gradation, this paper introduces a novel approach to classify NSS sampling units along a continuum. Rural areas fall under one of four ranks based on the population of the average village in the corresponding substratum (below district): 0-999 inhabitants, 1,000-4,999, 5,000-9,999 and 10,000 and above. The selection of these ranks are motivated by the 5000 population size used for the identification of Census Towns. Urban areas can be classified into three groups: below 50,000 inhabitants, between 50,000 and one million, and above one million. Details on the methodology employed to construct the ranks is in section 3 of the paper. The rural-urban cutoff point considered in most analyses is thus located between the fourth and the fifth ranks. But taking all seven of them into account supports a richer analysis of the impact of local employment opportunities on female LFPR.

There is a striking difference between the LFPR of men and women along this rural-urban gradation. While male LFPR is relatively stable across the seven ranks, female LFPR declines

steadily (figure 5)[6]. The exception to this trend is a "kink" around the fourth and fifth ranks, where the standard rural-urban divide is supposed to happen. Such kink is consistent with the hypothesis of misclassification of large villages with urban characteristics as rural, as in rank 4. Indeed, if many villages in rank three and especially in rank four are actually urban, the true female LFPR in rank four should be lower than the estimate suggests.

Figure 5: Female labor force participation declines along the rural-urban gradation



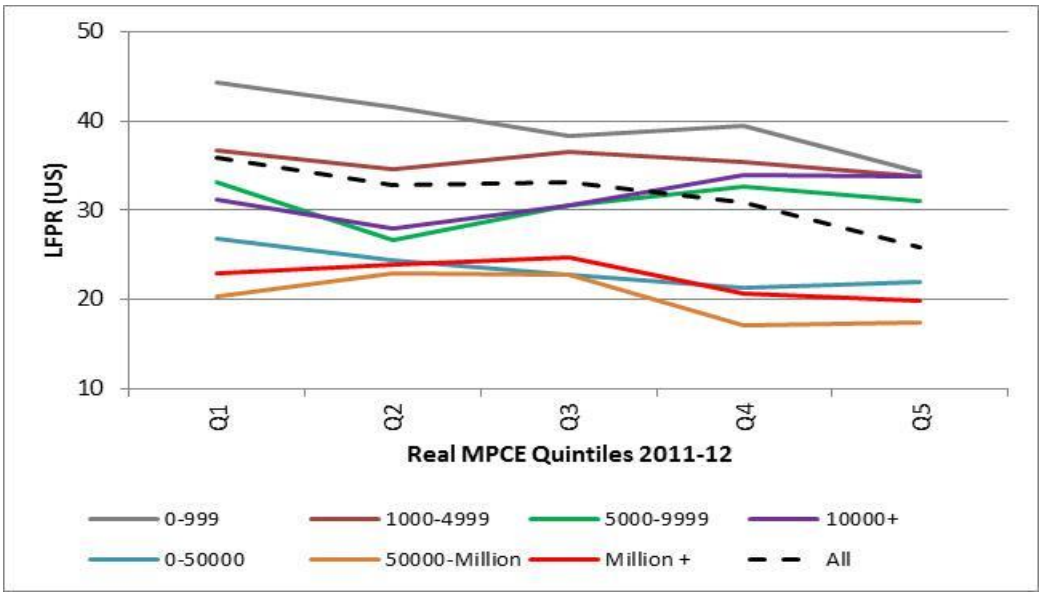
The definition of employment does not affect the male LFPR but it does result in a different pace of decline in female LFPR along the rural-urban gradation. In smaller villages, female LFPR is much higher based on US than it is based on PS. This is because women in rural areas are more frequently engaged in subsidiary work and casual jobs. The difference in female LFPR between these different definitions of employment shrinks along the rural-urban gradation, and almost completely disappears in large urban centers.

Supply-side interpretations emphasize the role of higher incomes as a key contributory factor to the decline in LFPR. Authors such as Goldin (1994) and Paxson et al. (2000) have argued that there is a "U-shaped" pattern in female LFPR associated with economic development. At low levels of income female LFPR tends to be high, as poor people need to work more to maintain a certain standard of living. Moreover, if there is a shift from agriculture to manufacturing, the nature of factory work could discourage women from participating. But as the economy develops and there is an associated expansion in the services sector, there could be an increase in LFPR, especially that of educated females. The nature of the jobs in the services sector may

[6] We observe similar patterns for 2004-05 as well.

be more attractive to females compared to factory work. In addition, women's relative wages may rise due to the comparative advantage they have over men in the service sector jobs. Thus, a combination of preferences and relative wages may counteract the income effect leading to a higher LFPR at higher levels of income for women.

Figure 6: Female participation rates do not decline with income in large villages and small towns

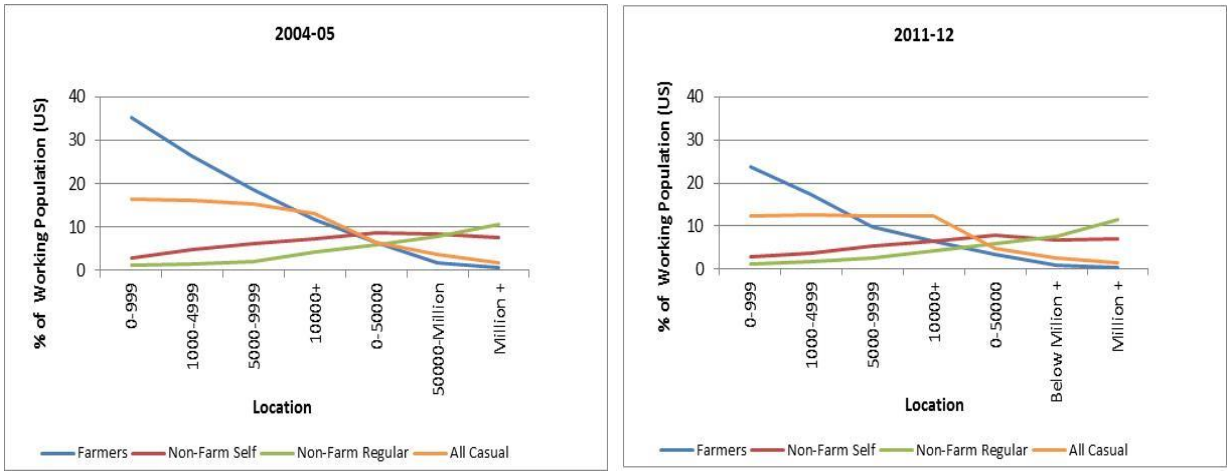


At an aggregate level, India seems to be in the declining portion of the U-shaped income effect. Households with higher monthly expenditure per capita had lower female LFPR in 2011-12 (figure 6), and the result was similar in 2004-5. However, the income-effect hypothesis also implies that LFPR for women from households with similar living standards should on average be the same regardless of where they live. This second implication is not supported by the data: for the same level of monthly expenditure per capita, female LFPR is generally higher in lower ranks of the rural-urban gradation, and declines with the average population size of the area of residence. The relationship between monthly per capita expenditures and female LFPR is actually flat in large villages and small towns, where the rural-urban divide is supposed to lie.

Areas of residence along the rural-urban gradation may differ in many important respects, from infrastructure to amenities. But one key difference between them is the availability of jobs for women. The structure of employment does not change in a linear way when moving from purely rural to increasingly urban settings. In both 2004-5 and 2011-12 there is a decline in the share of agricultural employment, relative to the working-age population (figure 7), along the rural-urban gradation. But the decline was steady and agricultural jobs were abundant in 2004-5. In 2011-12 however there was a sharper decline in agricultural jobs in smaller sized villages.

On the other hand, the share of non-farm regular employment is consistently low along the rural-urban gradation, and only picks up modestly in larger urban areas.

Figure 7: There was a sharp decline in the availability of jobs for females in villages and small towns



In a traditional society, where women's work is acceptable only if it takes place in environments perceived as safe, female LFPR can be expected to depend on the availability of farming jobs (which are mainly "at home") and of regular jobs. From this perspective, in India there is a "valley" of suitable job opportunities along the rural-urban gradation. And in recent years suitable job opportunities have declined precipitously in large villages and small towns.

3. Model and data

A simple analytical framework can be used to test the impact of demand-side factors, characterized by job opportunities and location, on female labor force participation in India. The framework takes the form of a series of nested specifications for the individual decision to participate in the labor force by working-age women. The standard supply-side specification, called Model A in what follows, emphasizes the characteristics of the woman making the decision and those of her household. It also allows for changes in household preferences regarding participation:

$$LFP = f_A(Individual, Household, Time)$$

where:

LFP: is the binary choice of participating in the labor market.

Individual are the characteristics of the woman making the decision, including her age, education and marital status.

Household are the characteristics of her household, including its composition by age and gender, its asset ownership, its social group, and its religious beliefs.

Time is the year of the NSS survey, when multiple cross-sections are pooled.

This basic specification can be enriched so as to take into account the characteristics of the area where the household lives, which yields Model B:

$$LFP = f_B(Individual, Household, Time, Location)$$

where:

Location is the rank of the area the household lives in along the rural-urban gradation (from one for less than 1,000 inhabitants to seven for more than one million)

A central hypothesis of this paper is that location matters not just in itself, but because the employment opportunities available for women along the rural-urban gradation are different. This hypothesis is captured in Model C:

$$LFP = f_C(Individual, Household, Time, Location, Employment)$$

where:

Employment is the local employment structure characterized by the share of farming, non-farm self-employment, non-farm regular wage employment and all casual employment relative to the working-age population of the area.

The paper also argues that in a context of rapid urbanization, the misclassification of actually urban areas as rural can result in important estimation biases. This gives us the broadest specification, called Model D:

$$LFP = f_D(Individual, Household, Time, Location, Employment, Gap)$$

where:

Gap is the difference between the actual share of the urban population in the district the household lives in and the urban share estimated based on the NSS. The actual urban share is computed on the basis of census data for 2001 and 2011, treating all villages with more than 5,000 inhabitants as urban.

The key hypotheses to be tested are:

1. When more jobs are available in the area where the household lives, other things equal the probability for a women to participate in the labor force is higher.

$$\frac{\partial f_D(.)}{\partial Employment} > 0$$

2. The misclassification of urban areas as rural biases the estimates. It makes the effect of urbanization look as a change in participation in rural areas.

$$\frac{\partial f_D(.)}{\partial Gap} \neq 0$$

3. Not taking into consideration the location where people live ("corrected" for the possible misclassification of urban areas) biases the estimates. The standard supply-side specification over-emphasizes the role of household income and composition.

$$\frac{\partial f_A(.)}{\partial Individual} \neq \frac{\partial f_D(.)}{\partial Individual}$$

$$\frac{\partial f_A(.)}{\partial Household} \neq \frac{\partial f_D(.)}{\partial Household}$$

4. When the employment structure is not taken into account the role of urbanization in accounting for the declining LFPR is over-estimated.

$$\left|\frac{\partial f_B(.)}{\partial Location}\right| > \left|\frac{\partial f_D(.)}{\partial Location}\right|$$

5. The standard supply-side specification attributes to a change in preferences a decline in LFPR due to a change in employment opportunities.

$$\frac{\partial f_A(.)}{\partial Time} < \frac{\partial f_D(.)}{\partial Time}$$

Models A, B, C and D are estimated in reduced form using Probit regressions. Data are individual records from the 61st (2004-05) and 68th (2011-12) rounds of the NSS employment and unemployment surveys. The descriptions of the variables used in the regressions and their descriptive statistics are in Table 1 and 2 respectively.

Both these rounds use the Census 2001 as the sampling frame for rural areas.[7] Linking the Census to the NSS at the substratum level allows us to reclassify rural areas in every district by average village size of the substratum. By doing so, we can construct employment variables below the district level (substratum) for rural areas and thereby control for the local labor market structure. We reclassify rural areas into four ranks based on the average population size of a village in a substratum: 0-999, 1,000-4,999, 5,000-9,999 and 10000 and above. We are able to match 571 districts in the NSS 61st round and 595 districts in the NSS 68th round with the Census 2001 districts.[8]

The same procedure cannot be replicated for the other three ranks considered in this paper, as the NSS uses the Urban Frame Survey (UFS) for urban areas, instead of the Census. Moreover, in urban areas the NSS data is not representative below the district level. For the urban sample we therefore use the population size of the location (below 50,000, between 50,000 and one million, and above one million) to attribute the rank, but construct the employment variables at the district level.

[7] The exception is Kerala, where the panchayat wards are used as the sampling frame.

[8] For the districts that were newly created after 2001, we merge the data of the new district with its parent to ensure comparability over time. Also, due to missing substrata information we drop 2 states, Delhi and Nagaland, and 2 Union Territories, Daman & Diu and Andaman & Nicobar Islands.

Table 1: Description of the variables used in the empirical analysis

Variables	Definition
Dependent Variable	
Labor Force Participation (Usual Status) for persons aged 15 years and above	
Independent Variables	
Individual Variables	
Age	Age
age_sq	Age Squared
schooling	Number of years of schooling (Lahiri et. al. 2013)
schooling_sq	Schooling Years Squared
marital_dummy	Whether currently married or not
Household Variables	
log_landown	Log of Land Owned by Household (in hectares)
log_hhsize	Log of Household Size
children_under_6	Share of Children below 6 years in the household
children_above_6	Share of Children 6 years or more in the household
female_adult	Share of female adults in the household aged 15-59
female_dependent	Share of female dependents (aged 60+) in the household
male_dependent	Share of male dependents (aged 60+) in the household
female_hh_dummy	Household head is a female or not
max_schooling	Maximum schooling of household
st_dummy	Household belong to Scheduled Tribes or not
sc_dummy	Household belong to Scheduled Castes or not
obc_dummy	Household belong to Other Backward Castes or not
hindu_dummy	Household belongs to Hindu religion
muslim_dummy	Household belongs to Muslim regilion
Time Variable	
survey	NSS61 =1, NSS68=2
Indicator Variable for Location	Average size of village in substratum for rural, Type of stratum for urban
rank	Population (0-999=1, 1000-4999=2, 5000-9999=3, 10000>=4) for rural Below million=5, Million plus=6 for urban
Employment Variables (Substratum /district level for rural , district level for urban)	**all=Usual Status, Working Age Population =Population 15 years and above**
all_farmers_share	Share of Agricultural Self Empoyed and Regular workers aged 15+ in working age population
all_non_farm_self_share	Share of Non Farm Self Empoyed workers aged 15+ in working age population
all_non_farm_regular_share	Share of Non Farm Regular Wage workers aged 15+ in working age population
all_casual_share	Share of All Casual workers aged 15+ in working age population
Misclassification of urban areas	
gap	Urban share in Census (based on 5000 population)-Urban share in NSS, district level

Table 2: Mean values of the variables used in the empirical analysis

	NSS 61	NSS 68	Percentage Change
age	30.86	31.53	2.2
schooling	7.33	7.77	6.1
marital_dummy	0.67	0.68	1.3
land (hectares)	0.09	0.08	-11.8
hhsize	5.26	4.86	-7.6
share children_under_6	0.07	0.06	-8.5
share children_above_6	0.09	0.09	2.4
share female_adult	0.21	0.22	6.8
share female_dependent	0.04	0.04	6.1
share male_dependent	0.04	0.04	6.9
female_headed_household	0.11	0.11	0.0
max_schooling_household	10.12	10.45	3.3
st_dummy	0.05	0.07	33.5
sc_dummy	0.14	0.16	9.2
obc_dummy	0.38	0.43	11.5
hindu_dummy	0.81	0.82	0.7
muslim_dummy	0.11	0.12	6.1
all_farmers_share	0.20	0.17	-15.0
all_non_farm_self_share	0.13	0.11	-14.5
all_non_farm_regular_share	0.10	0.09	-7.4
all_casual_share	0.17	0.17	-1.8
gap	0.20	0.21	5.0

4. Main results

The four models presented above are estimated for all women aged 15 years and above, with participation defined based on the US definition of employment in the NSS (Table 3). The estimated marginal effects are reported for each explanatory variable in all four models. Marginal effects indicate by how much the probability of labor force participation changes for a unit increase in the corresponding explanatory variable.

The overall explanatory power of Models C and D is much higher than that of the other two models, suggesting that the local structure of employment and the misclassification of urban areas as rural matter. The similarity in the fit of the regression for Models A and B suggests that taking into account location alone, while ignoring the employment structure or the potential spatial misclassification, improves the overall explanatory power of the model only marginally.

The estimated marginal effects under Model A are consistent with the conventional wisdom. The income effect cannot be inferred directly, as monthly expenditures per capita are endogenous and therefore cannot be used as an explanatory variable. But educational attainment is predetermined and provides a good proxy for income generation potential. The marginal effects for the education variable do reflect an inverted-U pattern, with participation

first decreasing with schooling, and then increasing. The turning point where higher educational attainment leads to higher LFPR is at the secondary level. However, it must be noted that the inverted-U implied in the estimates is almost flat in its downward-sloping part, and quite steep in its upward-sloping part.

Table 3: Marginal probability effects for working-age women based on Usual Status

Dependent Variable: Labor Force Participation				
	Model A	Model B	Model C	Model D
Age	0.049***	0.050***	0.056***	0.056***
age_sq	-0.001***	-0.001***	-0.001***	-0.001***
schooling	-0.032***	-0.032***	-0.033***	-0.034***
schooling_sq	0.002***	0.003***	0.003***	0.003***
marital_dummy	-0.093***	-0.098***	-0.118***	-0.120***
log_land	0.026***	0.020***	-0.004**	-0.002
log_land_sq	0.001***	0.002***	-0.002***	-0.002***
log_hhsize	-0.021**	-0.016*	0.019**	0.021**
children_under_6	0.008	-0.011	-0.027	-0.030
children_above_6	-0.028	-0.032	0.007	0.007
female_adult	0.115***	0.115***	0.126***	0.123***
female_dependent	0.108***	0.109***	0.125***	0.123***
male_dependent	0.071***	0.076***	0.103***	0.099***
female_hh_dummy	0.086***	0.086***	0.102***	0.099***
max_schooling	-0.021***	-0.018***	-0.018***	-0.017***
st_dummy	0.155***	0.136***	0.060***	0.065***
sc_dummy	0.089***	0.071***	0.059***	0.059***
obc_dummy	0.068***	0.060***	0.040***	0.038***
hindu_dummy	-0.065***	-0.062***	-0.049***	-0.046***
muslim_dummy	-0.142***	-0.135***	-0.090***	-0.091***
survey	-0.091***	-0.098***	-0.008**	-0.008*
rank==2		-0.015***	0.015**	0.007
rank==3		-0.046***	0.027***	0.003
rank==4		-0.019**	0.118***	0.049***
rank==5		-0.110***	0.080***	0.060***
rank==6		-0.148***	0.036***	0.018*
rank==7		-0.137***	0.009	0.007
all_farmers_share			1.559***	1.576***
all_non_farm_self_share			1.500***	1.524***
all_non_farm_regular_share			1.657***	1.702***
all_casual_share			1.455***	1.448***
gap				0.140***
Number of Observations	195457	195317	177782	175929
Wald Chi Sq	6072	7211	11984	11946
Pseudo R2	0.10	0.11	0.21	0.21
note: *** p<0.01, ** p<0.05, * p<0.1				

The marginal effect of land ownership is positive as well, but this should not be interpreted as a positive income effect. Farm land is only one among several household assets, such as housing ownership and financial savings. Holdings of these other assets are not captured by the NSS. Assume that across the population at large total household wealth is negatively correlated with agricultural land holdings. This would be so if the increase in the value of housing ownership and financial savings as households become more urban was larger than the decline in the value of farm land ownership. With a negative correlation between total wealth and farm land ownership, the positive marginal effect on the latter would amount to a negative income effect.

The estimates for Model A also confirm the conventional wisdom regarding the role of marriage and child-rearing in depressing the female LFPR. Married women are less likely to participate in the labor force. The marginal effects of having children are not statistically significant, but having more adults in the household does increase LFPR, consistent with the hypothesis that lack of support constraints women's ability to participate.

Importantly, all five hypotheses advanced in this paper are supported by the estimates. Consistent with the first hypothesis, the marginal effects of the employment shares are positive and highly significant in Models C and D, implying that the availability of jobs at the local level increases LFPR. While the effects are positive for all types of jobs, they are the largest for non-farm regular jobs, followed by farmers, non-farm wage jobs and then casual jobs. The hypothesis that the coefficients of all the employment variables are the same can be rejected at 1 percent significance level. The gap variable is also large and highly significant, implying that the misclassification of urban areas as rural matters. In districts where misclassification is substantial the LFPR is higher than the observed explanatory variables would suggest. This is because in such districts part of the apparent decline in rural LFPR is not really an "in situ" decline, but the result of shifting from a more rural to a more urban environment.

While most individual and household effects remain statistically significant across specifications, their magnitudes are attenuated when going from Model A to Models C and D. Consistent with the second hypothesis, income effects become less important (figure 8). The U-shaped curve for educational attainment changes only slightly, but it does so in the direction of a positive income effect. Land ownership becomes much less relevant too. The corresponding marginal effect is now close to zero, implying that for any value of the correlation between total household wealth and farm land holdings the overall income effect is much attenuated relative to Model A.

Other individual and household characteristics lose importance as well. In all specifications, women from Scheduled Castes and Scheduled Tribes are more likely to participate in the labor force. But the social group effects are attenuated when controlling for the availability of jobs. The same holds true for religious beliefs. The LFPR is lower among Muslim women, but the

marginal effect is smaller (in absolute terms) in Model C. Only the importance of having other adults around seems to be enhanced when moving from Model A to Models C and D.

Figure 8: The income effect is overstated when ignoring local circumstances

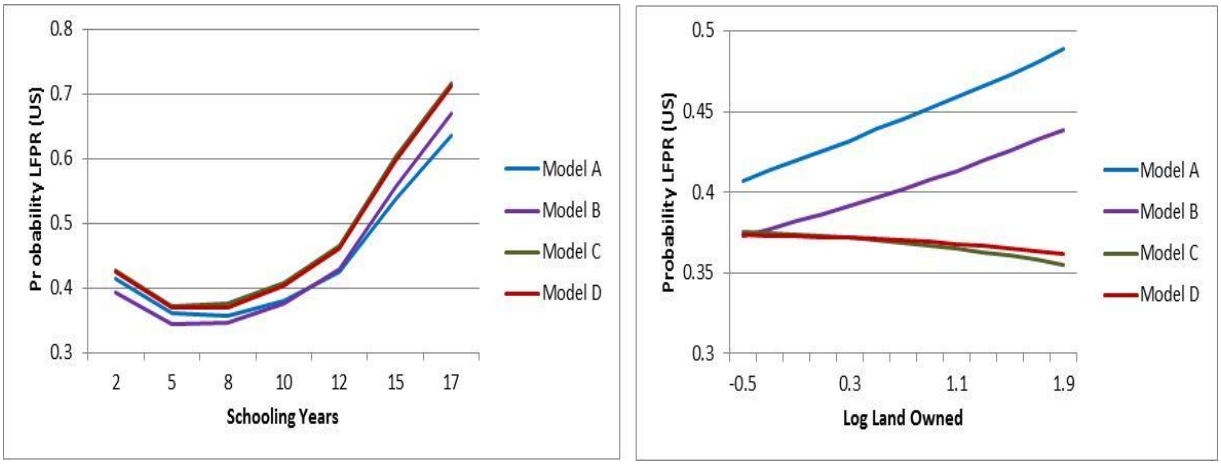
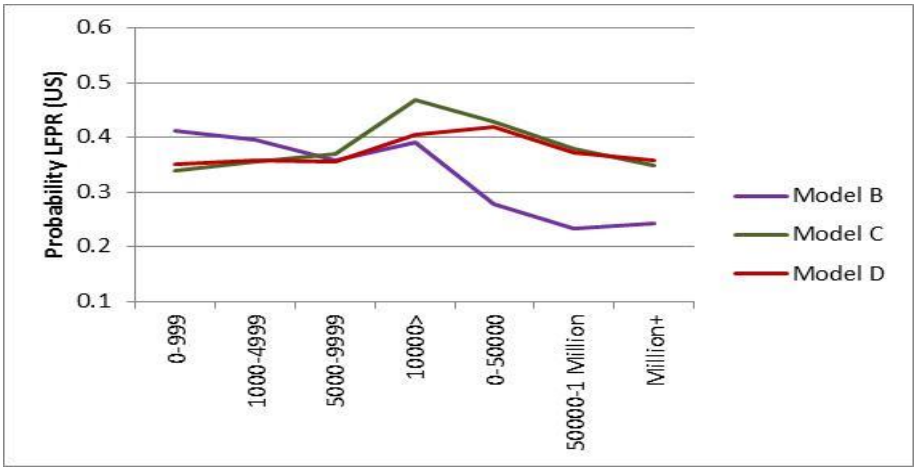


Figure 9: Urbanization by itself does not lead to lower female labor force participation



The fourth hypothesis in this paper is that the impact of urbanization on the LFPR is overstated in the standard analyses. This hypothesis is confirmed by the estimation results too (figure 9). Specifications not controlling for local employment opportunities and for the possible misclassification of urban areas as rural suggest that urbanization leads to a reduction in the female LFPR. But based on Models C and D, it is not urbanization by itself that causes the decline. For a similar local employment structure, the actual rank in the rural-urban gradation becomes mainly irrelevant. There is virtually no difference in the participation rates between smaller sized villages and larger sized towns. The kink that we observe in the middle, more

specifically at ranks 4 and 5, is considerably diminished with the inclusion of the gap variable in Model D. Thus, the jobs around where you live and not where you live per se matter for labor force participation of women.

Finally, the claim that household preferences in India are becoming more aligned with a patriarchal culture becomes much weaker when relying on the broader specification. Consistent with the fifth hypothesis in this paper, the marginal effect of the time variable decreases by two orders of magnitude when moving from Model A to Models C and D, and its statistical significance declines as well.

5. Robustness checks

The results presented above could be questioned if the population considered was not the most appropriate, or if important variables were omitted from the analysis, or if some of the variables included in it were not adequately measured.

The latter concern is particularly relevant when considering the various definitions of jobs in the NSS, and the fact that the female LFPR varies substantially with the definition used. A first robustness check thus consists of re-running the analysis with the PS definition of jobs, instead of the US definition used above (Table 4). The estimates are very similar to the ones reported in Table 3 using the US definition. Because the results do not change much, the other robustness checks are run reverting to the US definition of jobs.

It can also be argued that the characterization of the local context was too parsimonious. Infrastructure tends to improve along the rural-urban gradation, and amenities tend to increase, but these changes do not happen at the same pace along all locations. These and other characteristics of local areas could influence the participation decision by local households. Considering only the rank of the local area and its employment structure could thus bias the estimates. One way to address this concern is to introduce fixed effects at the district level (table 5). The main consequence of this change is to further reduce the significance of the marginal effects for the rank variables.

Last, the empirical analysis can be replaced for married women only. The goal of this robustness check is to account for the fact that individual preferences may differ depending on marital status, and household characteristics may play out differently for women in the two groups. Focusing on married women only allows removing the composition effect resulting from combining two heterogeneous sub-populations. But again, the results do not differ substantially (Table 6). But it is worth pointing out that the coefficients of variables measuring the share of children in the household, both under and above 6, are statistically significant for

this sub-group. But the coefficients of the employment variables remain larger. In fact, they are larger for married females than for the entire sample of women aged 15 and above.

Table 4: Marginal probability effects for working-age women based on Principal Status

Dependent Variable: Labor Force Participation				
	Model A	Model B	Model C	Model D
Age	0.042***	0.043***	0.045***	0.045***
age_sq	-0.001***	-0.001***	-0.001***	-0.001***
schooling	-0.031***	-0.032***	-0.030***	-0.031***
schooling_sq	0.002***	0.003***	0.002***	0.002***
marital_dummy	-0.112***	-0.116***	-0.135***	-0.136***
log_land	0.020***	0.017***	-0.008***	-0.006***
log_land_sq	0.002***	0.003***	-0.002***	-0.001***
log_hhsize	-0.031***	-0.028***	0.013*	0.016**
children_under_6	-0.023	-0.035*	-0.024	-0.026
children_above_6	-0.050***	-0.053***	-0.014	-0.013
female_adult	0.094***	0.092***	0.122***	0.119***
female_dependent	0.079***	0.077***	0.095***	0.092***
male_dependent	0.093***	0.095***	0.132***	0.128***
female_hh_dummy	0.086***	0.085***	0.102***	0.099***
max_schooling	-0.016***	-0.014***	-0.013***	-0.012***
st_dummy	0.178***	0.167***	0.065***	0.070***
sc_dummy	0.070***	0.057***	0.043***	0.043***
obc_dummy	0.070***	0.062***	0.035***	0.033***
hindu_dummy	-0.001	0.001	-0.016***	-0.012**
muslim_dummy	-0.078***	-0.074***	-0.054***	-0.054***
survey	-0.074***	-0.077***	-0.011***	-0.011***
rank==2		-0.002	0.019***	0.011**
rank==3		-0.016**	0.033***	0.009
rank==4		0.018**	0.132***	0.057***
rank==5		-0.059***	0.073***	0.051***
rank==6		-0.091***	0.036***	0.017*
rank==7		-0.087***	0.008	0.004
all_farmers_share			1.264***	1.281***
all_non_farm_self_share			1.157***	1.188***
all_non_farm_regular_share			1.305***	1.354***
all_casual_share			1.191***	1.185***
gap				0.133***
Wald chi2(24)	195457	195317	175820	174091
Prob > chi2	6072	5953	10443	10373
Pseudo R2	0.10	0.11	0.22	0.22
note: *** p<0.01, ** p<0.05, * p<0.1				

Table 5: Marginal probability effects with fixed effects at the district level

Dependent Variable: Labor Force Participation				
	Model A	Model B	Model C	Model D
Age	0.050***	0.051***	0.055***	0.055***
age_sq	-0.001***	-0.001***	-0.001***	-0.001***
schooling	-0.042***	-0.041***	-0.040***	-0.040***
schooling_sq	0.003***	0.003***	0.003***	0.003***
marital_dummy	-0.098***	-0.105***	-0.115**	-0.115***
log_land	0.022***	0.011***	0.002***	0.002
log_land_sq	-0.000**	-0.000*	-0.001**	-0.001**
log_hhsize	0.012	0.020	0.022	0.022**
children_under_6	0.028	0.010	-0.018	-0.018
children_above_6	0.036***	0.036***	0.041***	0.040*
female_adult	0.140***	0.143***	0.127***	0.127***
female_dependent	0.072***	0.071***	0.091***	0.091***
male_dependent	0.057***	0.067***	0.079***	0.079***
female_hh_dummy	0.082***	0.086***	0.087***	0.087***
max_schooling	-0.022***	-0.019***	-0.018***	-0.018***
st_dummy	0.149***	0.124***	0.102***	0.101***
sc_dummy	0.087***	0.070***	0.064***	0.065***
obc_dummy	0.058***	0.050***	0.043***	0.043***
hindu_dummy	-0.020***	-0.016***	-0.018***	-0.018**
muslim_dummy	-0.092***	-0.078***	-0.070**	-0.070***
survey	-0.087***	-0.096***	-0.014**	-0.014***
rank==2		-0.023***	0.010***	0.009
rank==3		-0.069**	0.011***	0.010
rank==4		-0.098***	0.007***	0.007
rank==5		-0.136***	0.059***	0.060***
rank==6		-0.169***	0.015	0.015
rank==7		-0.157***	0.007***	0.007
all_farmers_share			1.408***	1.409***
all_non_farm_self_share			1.400***	1.401***
all_non_farm_regular_share			1.296***	1.296***
all_casual_share			1.275***	1.275***
gap				0.088
Number of Observations	195420	195280	177760	175907.00
Wald Chi Sq	14672	15544	17587	17374
Pseudo R2	0.18	0.19	0.24	0.24
note: *** p<0.01, ** p<0.05, * p<0.1				

Table 6: Marginal probability effects among married women

Dependent Variable: Labor Force Participation				
	Model A	Model B	Model C	Model D
Age	0.033***	0.034***	0.043***	0.043***
age_sq	-0.000***	-0.000***	-0.001***	-0.001***
schooling	-0.022***	-0.021***	-0.022***	-0.023***
schooling_sq	0.002***	0.002***	0.002***	0.002***
log_land	0.038***	0.031***	0.001	0.004
log_land_sq	0.002***	0.003***	-0.002***	-0.001***
log_hhsize	-0.042***	-0.037***	0.008	0.010
children_under_6	-0.077***	-0.100***	-0.104***	-0.106***
children_above_6	0.105***	0.100***	0.164***	0.163***
female_adult	0.058	0.054	0.109*	0.105*
female_dependent	0.181***	0.175***	0.211***	0.208***
male_dependent	-0.085***	-0.077**	-0.048	-0.055*
female_hh_dummy	0.044***	0.043***	0.074***	0.069***
max_schooling	-0.021***	-0.018***	-0.017***	-0.017***
st_dummy	0.168***	0.147***	0.068***	0.074***
sc_dummy	0.100***	0.080***	0.070***	0.070***
obc_dummy	0.078***	0.069***	0.045***	0.043***
hindu_dummy	-0.083***	-0.079***	-0.062***	-0.058***
muslim_dummy	-0.184***	-0.176***	-0.127***	-0.128***
survey	-0.082***	-0.090***	0.009*	0.010*
rank==2		-0.020***	0.020**	0.010
rank==3		-0.049***	0.044***	0.016
rank==4		-0.017	0.154***	0.074***
rank==5		-0.137***	0.102***	0.078***
rank==6		-0.177***	0.058***	0.037***
rank==7		-0.164***	0.032	0.029
all_farmers_share			1.837***	1.857***
all_non_farm_self_share			1.614***	1.641***
all_non_farm_regular_share			1.876***	1.930***
all_casual_share			1.666***	1.657***
gap				0.161***
Wald chi2(24)	130410	130321	119161	118018
Prob > chi2	3501	4655	9027	8967
Pseudo R2	0.08	0.09	0.21	0.20
note: *** p<0.01, ** p<0.05, * p<0.1				

6. Interpreting the findings

The estimated results can be used to decompose the change in female LFP rate between 2004-5 and 2011-12 into changes due to different subsets of variables, including household income, household composition, local employment opportunities and measurement error. Let t=0 be 61st NSS round and t=T the 68th round. The predicted decision to participate in the labor force by a working-age woman is:

$$\widehat{LFP_t} = f_D(Individual_t, Household_t, Location_t, Employment_t, Gap_t)$$

Therefore the predicted change in the female LFPR can be approximated as:

$$\Delta LFP(total) = \widehat{LFP_T} - \widehat{LFP_0}$$

with both terms in the right hand side evaluated at the mean.

The same logic can be applied to predicting what the LFP would have been if only a subset of explanatory variables had changed values between 2004-5 and 2011-5. For instance, the effects of urbanization can be captured by replacing the mean value of the location variables in 2004-5 by their mean value in 2011-12, while leaving all the other variables unchanged. The contribution of urbanization to the change in the LFP is thus given by:

$$\Delta LFP(location) = f_D(Individual_0, Household_0, Location_T, Employment_0, Gap_0) - \widehat{LFP_0}$$

Similarly, it is possible to compute the effect of changes in individual characteristics, including educational attainment. This change can be interpreted as the impact of the income effect on LFP. Other changes refer to household characteristics, and in particular the presence of children and of other adults, as well as to job opportunities and to the extent of misclassification of urban areas as rural.

The results of this decomposition exercise show that almost all of the decline in female LFPR between 2004-05 and 2011-12 can be attributed to the changes in the structure of employment at the local level (table 7). Individual and household characteristics, including the educational

attainment and land ownership variables usually associated with the income effect, only play a marginal role.

Table 7. A simple decomposition of the decline in labor force participation

	2004-05	2011-12
Actual LFP	43.1	31.4
Predicted LFP, all variables at means	43.0	32.5
Predicted LFP in Model D 2004-05 by replacing a group of variables at a time using the 2011-12 means while keeping others at 2004-05 means		
Variable	**Predicted LFP**	**Change in LFP**
Individual	44.2	1.2
Household	42.3	-0.7
Location	42.9	-0.1
Employment	32.8	-10.2
Gap	42.8	-0.2

Appendix A: Definitions of Labor Force Participation in the NSS

The persons surveyed by the NSS are classified into various activity categories on the basis of the activities pursued by them during three reference periods namely, (i) one year, (ii) one week, and (iii) each day of the reference week. Based on these three periods, three different measures of activity status were arrived at, which are termed as 'usual status' (principal and subsidiary), 'current weekly status' and 'current daily status' respectively.

Usual Principal Status: The activity status on which a person spent relatively long time (i.e. major time criterion) during the 365 days preceding the date of survey was considered as the usual principal activity status of the person. To decide the usual principal activity of a person a two stage dichotomous classification was followed. At the first stage, persons were first categorized as those in the labour force and those out of the labour force depending on the major time spent during the 365 days preceding the date of survey. At the second stage, for persons belonging to the labour force, the broad activity status of either 'working' (employed) or 'not working but seeking and/or available for work' (unemployed) was ascertained based on once again the time criterion.

Usual Subsidiary Status: A person whose principal usual status is determined on the basis of the major time criterion could pursue some economic activity for a relatively shorter time during the reference period of 365 days preceding the date of survey or for a minor period, which is less than 30 days (not necessarily for a continuous period) during the reference year. The status, in which such economic activity is pursued, is the subsidiary economic activity status of that person.

Usual Status: The usual status, determined on the basis of the usual principal activity and the usual subsidiary activity of a person taken together, is considered as the usual activity status (PS+SS) of the person. According to the usual status, workers are those who perform some work activity either in the principal status or in the subsidiary status. Thus, a person who is not a worker in the usual principal status is considered as worker according to the usual status, if the person pursues some subsidiary economic activity for 30 days or more during 365 days preceding the date of survey.

Current Weekly Status: The broad current weekly activity status of a person is decided on the basis of a certain priority cum major time criterion. A person was considered working (or employed) if, while pursuing any economic activity, had worked for at least one hour on at least one day during the seven days preceding the date of survey. A person was considered 'seeking or available for work (or unemployed)' if, during the reference week, no economic activity was pursued but he/she made efforts to get work or had been available for work any time. A person who had neither worked nor was available for work any time during the reference week

was considered as out of the labour force. After the broad current weekly status of a person was determined using priority criterion, the detailed current weekly status was derived from the intensities[9] (time disposition during the week) assigned for the daily activities performed by a person during the seven days of the reference week.

Current Daily Status: The current daily activity status for a person was determined on the basis of his/her activity status on each day of the reference week using a priority-cum-major time criterion. A person is considered "working full day" if he/she had worked for four hours or more during the day, considered as "working half day" if he/she had worked for one to four hours in a day and "unemployed or not in the labor force " for the other half of the day. If a person had not worked for even one hour in a day but was seeking/available for work for four hours or more, he/she would be considered "unemployed" for full day and if he was seeking/available for work for less than 4 hours, he would be considered "unemployed" for half day and "out of labor force".

[9] Intensity is calculated for up to a maximum of two economic activities per day. Thus, on a day, a person may either have only one activity with 'full' intensity or two activities with 'half' intensity for each. If the activity is pursued with intensity 'half' on a particular day, the entry will be 0.5 against that activity and if that is pursued with intensity more than half, 1.0 will be recorded against that activity. Generally, an activity, which is pursued for more than 1 hour but less than 4 hours is considered to have been pursued with 'half' intensity. If it is pursued for 4 hours or more, the activity is considered to have been pursued with 'full' intensity.

References

Abraham, Vinoj (2013). "Missing Labor or Consistent De-feminization?" *Economic and Political Weekly*. August.

Bhalla, Surjit S. and RavinderKaur (2013). "Labor Force Participation of Women in India: Some facts, some queries". *Asia Research Center Working Paper*, 40. London School of Economics.

Chand, Ramesh and Srivastava, S.K. (2014). "Changes in the Rural Labour Market and Their Implications for Agriculture". *Economic and Political Weekly*. March.

Chandrashekhar, C P and Jayati Ghosh (2007). "Recent Employment Trends in India and China: An Unfortunate Convergence". Paper presented at JNU-IIAS conference "Making Growth Inclusive with Reference to Employment Generation", 28-29 June, Jawaharlal Nehru University, New Delhi. *Processed*.

Chowdhury, Subhanil (2011). "Employment in India: What Does the Latest Data Show?" *Economic and Political Weekly*. February.

Goldin, Claudia (1995). "The U-Shaped Female Labor Force Function in Economic Development and Economic History", in: T. P. Schultz (ed.): *Investment in Women's Human Capital and Economic Development*. University of Chicago Press: Chicago, 61-90.

Himanshu (2011). "Employment Trends in India: A Re-examination". *Economic and Political Weekly*. September.

Hirway, Indira (2012). "Missing Labor Force. An Explanation". *Economic and Political Weekly*. September.

ILO (2013). *Global Employment Trends 2013: Recovering from a Second Jobs Dip*. Geneva

Kannan, K P and G Reveendran (2012). "Counting and Profiling the Missing Labor Force". *Economic and Political Weekly.* February

Klasen, Stephan and Janneke, Pieters (2013). "What Explains the Stagnation of Female Labor Force Participation in Urban India?" *IZA DP,* 7597. Institute for the Study of Labor, August.

Mammen, Kristin and Christina Paxson (2000). "Women's Work and Economic Development". *Journal of Economic Perspectives*, 14(4), 141-164.

Neff, Daniel, Kunal Sen and Veronika Kling (2012). "The Puzzling Decline in Rural Women's Labor Force Participation in India: A Rexamination". *German Institute of Global and Area Studies, Working Paper*, 196, May.

Olsen, Wendy (2006). "A Pluralist Account of Labour Participation in India," Economics Series Working Papers, University of Oxford, Department of Economics, GPRG-WPS-042.

Rangarajan, C, Padma Iyer Kaul, Seema (2011). "Where Is the Missing Labor Force?" *Economic and Political Weekly*. September.

Rodgers, J. (2012). "Labour Force Participation in Rural Bihar: A Thirty-year Perspective Based on Village Surveys". Institute of Human Development, New Delhi. *Processed*.

Sundaram, K and S Tendulkar (2006). "Changing Structure of India Workforce, Quality of Employment and Real Earnings, 1983-2000", in *India: Meeting the Employment Challenge*, Conference on Labour and Employment Issues, Institute for Human Development, New Delhi.

Unni, Jeemol and G Raveendran (2007). "Growth of Employment (1999-00 to 2004-05): Illusion of Inclusiveness". *Economic and Political Weekly*. January.