INSTITUTE OF DEVELOPING ECONOMIES

IDE Discussion Papers are preliminary materials circulated to stimulate discussions and critical comments

The Informal Labour Market in India: Transitory or Permanent Employment for Migrants?

Abu S. SHONCHOY* and P.N. (Raja) JUNANKAR

March 2014

Abstract

The informal economy is a very important sector of the Indian economy. The National Council of Applied Economic Research estimates that the informal sector - "unorganised sector" - generates about 62% of GDP and provides for about 55% of total employment (ILO 2002, p. 14). This paper studies the characteristics of the workers in the informal economy and whether internal migrants treat this sector as a temporary location before moving on to the organised or formal sector to improve their lifetime income and living conditions. We limit our study to the Indian urban (non-agricultural) sector and study the characteristics of the household heads that belong to the informal sector (self-employed and informal wage workers) and the formal sector. We find that household heads that are less educated, come from poorer households, and/or are in lower social groups (castes and religions) are more likely to be in the informal sector. In addition, our results show strong evidence that the longer a rural migrant household head has been working in the urban sector, ceteris paribus, the more likely that individual has moved out of the informal wage sector. These results support the hypothesis that, for internal migrants, the informal wage labour market is a stepping stone to a better and more certain life in the formal sector.

*Research Fellow, Micro-economic Studies Group, Development Studies Center, IDE-JETRO, Chiba, Japan (parves.shonchoy@gmail.com)

The Informal Labour Market in India: Transitory or Permanent Employment for Migrants? [1] [2]

P.N. (Raja) Junankar[#] and Abu Shonchoy[†]

Abstract

The informal economy is a very important sector of the Indian economy. The National Council of Applied Economic Research estimates that the informal sector - "unorganised sector" - generates about 62% of GDP and provides for about 55% of total employment (ILO 2002, p. 14). This paper studies the characteristics of the workers in the informal economy and whether internal migrants treat this sector as a temporary location before moving on to the organised or formal sector to improve their lifetime income and living conditions. We limit our study to the Indian urban (non-agricultural) sector and study the characteristics of the household heads that belong to the informal sector (self-employed and informal wage workers) and the formal sector. We find that household heads that are less educated, come from poorer households, and/or are in lower social groups (castes and religions) are more likely to be in the informal sector. In addition, our results show strong evidence that the longer a rural migrant household head has been working in the urban sector, ceteris paribus, the more likely that individual has moved out of the informal wage sector. These results support the hypothesis that, for internal migrants, the informal wage labour market is a stepping stone to a better and more certain life in the formal sector.

*JEL Codes***:** 017, J15, J61, J42
*Keywords***:** informal labour markets, migrant, caste, religion

---

[#] University of New South Wales, University of Western Sydney, and IZA, Bonn, raja.junankar@uws.edu.au
[†] Institute of Developing Economies (IDE), JETRO and University of Tokyo, parves.shonchoy@gmail.com

1. Introduction

In most developing countries there is a large sector of the economy that is called the informal sector or the unorganised sector. Employment in the informal labour market plays an important role in most developing economies. Very broadly, the informal labour market consists of workers in the informal sector plus casual workers in the formal sector. The informal labour market is a very large part of the agricultural sector, but is also a significant part of the urban sector. There is a difference between employment in the formal sector and the informal sector in terms of the conditions of work, whether workers are subject to government taxes, have access to social security or insurance, are casual or contract workers, and whether or not they receive minimum wages.

The informal economy is a very important sector of the Indian economy. The National Council of Applied Economic Research estimates that the informal sector - "unorganised sector" - generates about 62% of GDP, 50% of national savings and 40% of national exports (ILO 2002, p. 30). In terms of employment, the informal economy provides for about 55% of total employment (ILO 2002, p. 14). Urban areas (especially large cities) attract numerous migrants from both the rural areas and from smaller urban towns and cities in the hope of a better life.

The Indian labour market can be conceived of as a segmented market. One segment is the formal sector composed of workers who have salaried work, good working conditions, and are employed in organised business. The other segment is the informal economy consisting of small self-employed traders and business people, and casual workers in the informal or formal sectors. Some individuals are born into wealthy families who own large businesses and hence are in the formal sector by right of birth. Others who are born with parents from the professional classes would almost certainly have education from good schools and universities, and have a network of contacts that would ensure their joining the ranks of the employed in the formal sector. Some individuals may have built up sufficient assets over time to set up small businesses and hence enter the formal sector. However, most workers in the formal sector enter the sector through their educational achievements, or by birth (children of rich people) and through social networks. For someone who comes from a poor background (either in terms of income or belonging to a socially disadvantaged caste or

religion) the only way to enter the formal sector is via education in "good" schools[3] or universities. Even with a good education, entry into the formal sector is often based on family connections. The Indian government has for some time had a policy of positive discrimination for the Dalits, and as a result they may have a higher probability of finding a job in the formal (government) sector. Migrants (especially from rural areas) who come into urban areas would likely have to work in the informal sector for some time before they build good networks enabling them to move into the formal sector.

The literature on the role of the informal sector in developing countries has oscillated between treating the informal sector as a backward sector that is holding back economic development to a dynamic sector that is helping to develop the economy rapidly without straining foreign currency balances and with relatively low demands for (real) capital goods (see Mazumdar (1976), Weeks (1975), Bromley (1978), Gerxhani (2004)). The informal sector is considered to be a pre-capitalist form of production compared to the formal sector which is a profit maximising capitalist sector. There is a large volume of literature on rural-urban migration (see, Harris and Todaro (1970)) that examines migrants arriving in the city and initially finding work in the informal sector, then moving on to better paid work in the formal sector. Fields (1975) developed an early model of the informal sector as a "way station" along the path toward a formal job in urban areas (De Mel et al. (2010)) which has been followed by others. This view of the informal sector as a temporary abode for migrants has been disputed (among others) by Mazumdar (1976). The debate has also ranged over whether informal sector workers are living in poor conditions with low incomes, or whether some of the informal sector workers are there out of choice and have a comfortable life (see Meng (2001)). Some individuals may have employment in the formal sector and work in the informal sector as well.

Given the setup of the urban labour market in India, some of the important issues to investigate are (1) whether individuals working in the informal sector are migrants and whether they move out of the informal sector into the formal sector after a few years; (2) whether they are from disadvantaged social and ethnic groups who do not have social networks to enter the formal sector; and finally (3) whether those with low levels of education and skills are unable to enter formal sector employment and have to find low paid work in the informal sector.

---

[3] A "good" school would almost certainly be an established private school.

This paper is interested in studying the characteristics of the workers in the informal economy and whether migrants treat this sector as a permanent base or only as a temporary location before moving on to the organised or formal sector to improve their lifetime income and lifestyle. We limit our study to the Indian urban (non-agricultural) sector and study the characteristics of the household heads that belong to the informal sector (self-employed and informal wage workers) and the formal sector. We find that members who come from the lower social groups (castes and religions) are more likely to be in the informal sector. We distinguish between migrants in urban location who came from rural areas and those who came from other urban areas. We find that the longer duration of a rural migrant in the urban area, the lower the probability that the household head would be in the informal sector.

In Section 2 below, we clarify the definition of informal labour markets and briefly review the literature; Section 3 provides a detailed discussion of the properties of the urban informal sector in India; Section 4 discusses the lexicographic preferences that people have over formal sector, self-employment, and informal wage labour; Section 5 sets up an econometric model for estimating the probability of working in the informal sector and provides some results, while Section 6 provides results using a multivariate logit model; Section 7 concludes with a summary of the results. In general, we find that the longer the duration of a migrant in the urban sector the more likely s/he will have moved out of the informal sector.

2. The Informal Labour Market: Definitions and a review of some earlier studies

In the developing country context, the informal sector is sometimes defined in terms of the activities of the enterprises (ILO, 1972) and sometimes in terms of the kind of work done by individuals as employees or as self-employed people (Hart, 1973).

In 1972 the ILO characterised the informal sector as:

(a) Ease of entry

(b) Reliance on indigenous resources

(c) Family ownership of enterprise

(d) Small scale of operation, often defined in terms of hired workers less than (say) ten

(e) Labour-intensive methods of production and adapted technology

(f) Skills acquired outside the formal school system

(g) Unregulated and competitive markets

Whereas the formal sector was characterised by:

  (a) Difficult entry
  (b) Frequent reliance on overseas resources
  (c) Corporate ownership
  (d) Large scale of operation
  (e) Capital-intensive and often imported technology
  (f) Formally acquired skills, often expatriate
  (g) Protected markets (through tariffs, quotas, and licences)

Hart (1973) discussed the informal sector in terms of the working conditions of the individuals and whether they worked for wages with good conditions or informally as self-employed workers. Informal activities included:

  (a) Farming, market gardening, self employed artisans, shoe makers, tailors, etc.
  (b) Working in construction, housing, road building
  (c) Small scale distribution, e.g. petty traders, street hawkers, caterers in food and drink, etc.
  (d) Other services, e.g. barbers, shoe-shiners etc.
  (e) Beggars
  (f) Illegal activities like drug pushing

Formal sector income earning activities included:

  (a) Public sector wage earners
  (b) Private sector wage earners (on permanent contracts, not casual workers)

Sengupta (2009, p. 3) defined the informal economy thus:

> *Informal Sector*: The unorganised sector consists of all unincorporated private enterprises owned by individuals or households engaged in the sale and production of goods and services operated on a proprietary or partnership basis and with less than ten total workers.

Informal worker/employment: Unorganised workers consist of those working in the unorganised sector or households, excluding regular workers with social security benefits provided by employers and the workers in the formal sector without any employment and social security benefits provided by employers.

Informal economy: The informal sector and its workers plus the informal workers in the formal sector constitute the informal economy.

3. The Indian Informal Labour Market: Some Background Information

A recent report of the National Commission for Enterprises in the Unorganised Sector by the Government of India (Sengupta 2009) found that 86% of the total employment in 2004-2005 was in the informal sector. Further, the agricultural sector consisted almost entirely of informal workers. The non-agricultural workers in the informal sector were 36.5% of the total, most of whom were self-employed. From 1999-2000 to 2004-2005 most of the increase in employment in the formal sector consisted of informal workers (Sengupta 2009, p.14). The NSSO (2012, p ii) document found that in 2009-2010 in the non-agriculture sector, nearly 71% of the workers in rural areas and 67% in urban areas worked in the informal sector. It found that the informal sector activities are concentrated mainly in the manufacturing, construction, wholesale and retail trades, and transport, storage and communication industries.

Our study used data from the India Human Development Survey (IHDS) 2005, conducted by the Inter-university Consortium for Political and Social Research, Ann Arbor, Michigan, USA. The survey is a nationally representative, multi-topic survey of 41,554 households in 1,503 villages and 971 urban neighbourhoods across India. The data set has detailed information on household employment by industry and occupation, and detailed information about household characteristics including age, education, ethnicity, religion, and migration status. In this study we have limited our analysis to workers in the informal labour markets in the urban sector who are not engaged in any agricultural activities.

Our data set consisted of 12,056 heads of households for whom we had data on their age, education, marital status, gender, religion, caste, income source, assets, migration status and years since migration into the urban sector, and whether they were slum dwelling.

We define the urban informal sector as artisans, petty traders, small business people (who do not hire any labour), and non-agricultural casual workers in the informal or formal sectors. The informal sector consists of the self-employed and informal wage labour. We define self-employment as petty traders who do not hire any workers and those in the organised trade/business category who do not hire any workers. Note that this is a stricter definition than that suggested by, for example, Sengupta (2009). The informal wage labour category covers those who are in the informal sector but are not self-employed, i.e., the artisans, and non-agricultural labourers who are casually employed. The formal sector consists of salaried employees, professionals, and organised trade/business people who hire workers. In our study we limited our analysis to only the heads of households.

[Figure 1 is about here]

It is interesting to notice the industry and occupational distribution of the formal and informal sectors of the economy in our sample data. Most of the informal wage labour is in manufacturing, construction, wholesale and retail trades, restaurants, and hotels, and in community, social and personal services. Self-employment is concentrated (not surprisingly) in the wholesale and retail trades, restaurants, and hotels. Informal wage labour is concentrated in occupations: production and related workers, transport equipment operators and labourers (presumably unskilled workers).

[Figure 2 is about here]

If we look at the distribution of migrants moving into these sectors, we find that 38.88% of the migrants work in the formal sector; almost 21.58% are self-employed entrepreneurs and 17.30% are informal wage workers.

[Figure 3 is about here]

A high proportion of migrants (28%) are working primarily in the community, personal and social services, 21% in wholesale and retail trades, restaurants and hotels, and 17% in manufacturing.

A high proportion of migrants (34%) are employed as production and related workers, transport equipment operators and labourers while almost 20% are sales and service workers. It is interesting to note that the main income source for migrants (52%) is salaried employment; another 18% are in non-agricultural labour.

It is interesting to see the caste and religion breakdown for the formal and informal sectors (self-employed and informal wage labour). As we would suspect, Brahmins and high caste people are more likely to be in the formal sector, compared to the lower social castes and Muslims. If we look at the distribution of people by caste and religion for the principal source of household income, we see that Brahmins and high caste people are more likely to be salaried workers or professionals, whilst Dalits and Muslims are more likely to be non-agricultural labourers or artisans (see Table 1).

When we look at the distribution of occupations by caste and religion, we note that Brahmins and high caste people are more likely to be in the higher level occupations, while Dalits and Muslims are more likely to be in the lower level occupations. When we look at the distribution of industries that the different castes and religions are located in, we see that manufacturing, transport, and finance are important for most groups.

4. **The Informal Economy: Some Analytical Features**

We assume that individuals would prefer to work in the formal sector, either as employees or as owners/managers. This is based on the idea that the formal sector provides a better life not only in terms of present and future income, but also in terms of better working conditions (e.g., security of tenure, social security benefits, access to unions, safer working conditions). If they are unable to enter the formal sector, we assume that they would prefer to be self-employed (as long as their expected incomes are not below that in the informal wage sector). Employees in the informal wage sector would prefer to become self-employed if they

have access to credit to set up small businesses. Many of them may simply be "waiting" for a job in the formal sector. In the Harris-Todaro model, rural migrants come to the urban areas as long as their expected wages (urban wage multiplied by the probability of finding a job) are greater than their rural subsistence wage. Migrants who do not find work in the urban formal sector then enter the urban informal sector which is meant to be a form of "wait unemployment". Essentially, we are arguing that individuals have lexicographic preferences over these choices. However, what we observe is a reduced form depending on the household head's choice and success in the formal labour market, and the constraints in the credit market that determine whether s/he can become self-employed. Informal wage labour then is a residual category.

In fact if we look at the actual income (based on our sample), we find that the incomes of these three groups overlap to some extent, with the lowest income being in informal wage labour, followed by self-employment, followed by formal sector income. Figure 8 presents the kernel densities of the logs of informal wage labour, informal self-employment, and formal income respectively. As can be seen, informal wage-labour income is distributed to the left, informal self-employment income is in the middle, and formal income is to the right of the other distributions. There is some overlap at the lower tails of the distributions, but self-employment and formal incomes have tails spread out at the higher income levels.

[Figure 8 is about here]

[Table 3 is about here]

A Kolmogorov-Smirnov test reveals that there are significant differences in these kernel densities. (All pair-wise Kolmogorov-Smirnov tests are statistically significant with a p-value of 0.000). Table 3 provides some summary statistics to illustrate the differences in the distribution of incomes. As discussed above, the mean (log) income in the formal sector is greater than that for the self-employed and that is greater than that for informal wage labour. The only curious result seems to be that the minimum of the formal sector is lower than that of the other two groups.

To be in the formal sector, domestic capitalists need to have significant amounts of capital and access to credit. Inheritance plays a large part in providing either the original capital or access to credit. Multinationals come in with large amounts of capital with technology that is labour saving (embodied technological change). Employment in the formal sector is then limited by the use of imported technology and limited amounts of capital. Note that only a limited amount of labour-capital substitution is possible because of embodied technology.

Wages in the formal sector are fixed by the government (minimum wages), by unions or employers using efficiency wage ideas, or by multinational firms that constrain wages. Employers in the formal sector ration employment by using education/experience as an index of productivity, and using religion/caste as a signal for productivity (statistical discrimination). Given two people with the same education/skill levels, they would prefer a high caste Hindu to a low caste Hindu or a Muslim. Note: being in the formal economy is not a guarantee against poverty (see ILO 2002, p.31).

Self-employment (in the informal sector) is constrained by limited amounts of credit and access to capital. The higher the social class and the higher the level of education, the easier it is to access credit. Note: ILO (2002, p. 31) provides evidence that many in the informal economy, especially the self-employed, earn more than unskilled or low-skilled workers in the formal economy.

Informal sector employment is a residual; the lower the employment in the formal sector, the greater the number who look for work in the informal sector, and hence, the lower the wages (income) for this sector.

Figure 9 below shows that 43% of the self-employed have taken out loans for business purposes, compared to only 14% in the formal sector and 16% in the informal wage labour group. It is clear that the self-employed have to take out loans for setting up and running a small enterprise. Presumably many of the informal wage workers would be interested in setting up a small business but are unable to access credit.

[Figure 9 is about here]

To summarise this section, we argue that households have a lexicographic preference ordering over the different outcomes, formal, self-employment, or informal wage labour.

10

Migrants, especially rural migrants, have little access to credit or to the formal labour market, at least until they have spent some years in the urban sector.

5. Probability of Working in the Informal Sector

In this section we estimate the probability of a household head working in the informal wage employment sector to be self-employed or in the formal sector. As discussed earlier, we make three hypotheses. One is that those households that come from the lower social classes/groups are more likely to be working in the informal sector. Some of these households may have the entrepreneurial skills or have access to small amounts of capital to set up as self-employed workers. We also hypothesise that households coming from higher social classes/groups, and/or that have higher levels of education are more likely to be working in the formal sector. Further we hypothesise that migrants who come into the urban areas would initially find employment in the informal sector and after some time, when they have accumulated sufficient funds or developed social networks or skills, are more likely to move into the formal sector. In our analysis below we regard migrants as those whose origin is in a rural area. Individuals who have come from other urban areas are regarded as "urban natives". We hypothesise that the duration of migration from a rural origin influences the sector of employment.

5.1 Econometrics and Identification Strategy

The fundamental challenge of estimating the causal impact of migration duration on the probability of working in the informal sector is the possibility of unobserved individual characteristics that might influence the migration decision, survival in a migration destination, and the duration as well as the likelihood of working in the informal sector. For example, it might be possible that individuals with high unobserved ability or entrepreneurial skills might opt to move out of the rural area early in their life and remain in the urban area, and such unobserved skills and ability will also influence their choice of sector in the migration destination. Without controlling for this, estimation may be biased and inconsistent.

If we had panel data, we could have used methods to control for individual heterogeneity. Another ideal method that could be used to disentangle such unobserved

influences on migration duration and job status would be by using some natural experimental framework or by randomly inducing people to migrate out of rural areas to estimate the causal impact of migration on job choice. Lacking the availability of such methods, we need to opt for an instrumental variable approach (IV) where we can instrument migration duration with a set of variables which do not have a direct influence on job placement or current job status.

One recent method to instrument for migration is using the historic migration rate as an instrument for current migration status (for example see, Woodruff and Zenteno (2007), Hanson and Woodruff (2003); McKenzie and Rapoport (2007, 2011); López-Córdoba (2005); and Hildebrandt and McKenzie (2005)). Following the methodology of these influential works, we have used the historic state-level migration rates as an instrument for current migration duration. In particular, we use the Indian migration rates from data collected in the 1991 census at the state level and use this variable as an instrument in which the household is currently located.

It can be argued that these historic migration rates are the result of the massive development of railroads and other transportation systems in India coupled with the rapid economic expansion of large cities which created extended job demand. These historic migration rates can also be considered as signals of migration friendliness, of strong migration networks that can effectively lower the cost of migration for future potential migrants and increase the chances of their survival. These signals and networks become self-perpetuating, and as a result, continue to influence the migration decisions of households today.

Our identifying assumption is that historic state-level migration rates do not affect the current job placement of the individuals, apart from their influence through current migration. Instrumental variables estimation relies on this exogeneity assumption, and so it is important to consider and counteract potential threats to its validity.

One potential threat is that the historic level of inequality and lower economic class (lower caste and religious groups) could induce the historic migration rate, and that they are also influencing the current rate due to intergenerational transition. To tackle these potential pitfalls, we also used interaction terms of historic migration rate with the caste dummies as

additional instruments.[4] We have also controlled for city- and district-level fixed effects to control for spatial differences and location preferences, and report our results based on standard errors clustered at the state level to correct for arbitrary correlation in the error structure of individuals within a state (McKenzie *et al.* 2012).

As our main outcome of interest is whether migrants use the informal sector as their temporary base (like a stepping stone), we studied the impact of migration duration of individuals on their placement in the informal sector. The reduced form IV approach consists of estimating a two-stage model of the following form, where I_j is the outcome variable of interest (individual *j's* current employment sector), M_{jk} is individual *j's* migration duration, and who is currently staying in State *k* (years of migration since originally migrating), and Z_k is the set of instrumental variables. Hence the reduced-form first stage equation for migration M_{jk}, following Amemiya (1978), would be:

$$M_{jk}^* = \beta_0 + \beta_1 Z_k + \beta_2 X_{jk} + \gamma_k^m + \epsilon_{jk}^m, \qquad (1)$$

$$M_{jk} = \begin{cases} M_{jk}, & if\ M_{jk}^* > M_0 \\ 0, & if\ M_{jk}^* \le M_0 \end{cases},$$

and the equation for employment in the informal sector I_{jk} is

$$I_{ik}^* = \alpha_0 + \alpha_1 M_{jk} + \beta_2 X_{jk} + \gamma_k^i + \epsilon_{jk}^i, \qquad (2)$$

$$I_{jk} = \begin{cases} 1, & if\ I_{jk}^* < I_0 \\ 0, & if\ I_{jk}^* \ge I_0 \end{cases}.$$

Here M_{jk}^* is the latent variable for migration decision, and M_{jk} is the observed years of migration duration to the current state *k* from the time individual *j* decided to migrate to state *k* by comparing the costs and benefits using a net benefit function or latent index expressed in equation (1). Similarly, I_{ik}^* is the latent job placement and I_{jk} is the dummy of job placement in the formal and informal sectors for the same individual *j* living in state k which can be seen arising by comparing the job qualifications and job related network information (such as informal or formal referral system) required for the job placement expressed in equation (2). In this setup the first dependent variable, M_{jk}, appears in the second equation as an endogenous variable. Here, X_{jk} includes the following set of controls: personal and household characteristics, family background information, family composition information, religion, and

---

[4] As a robustness check, we ran regressions without land holding variables, and our regression remained consistent.

a dummy variable indicating whether or not the person is an urban native (the dummy is equal to one if the individual i who currently resides in state k is born in an urban area and zero if the person is a rural-to-urban migrant). Personal characteristics include age, age^2, sex, education and marital information, whereas household characteristics include wealth status of the household which has been constructed using the principal component analysis of the household non-durable assets.[5] Family background information contains variables on father's education and occupation history. γ_k^M and γ_k^I are unmeasured determinants of M_{ik} (e.g., migrant's own community network) and I_{ik} which is fixed at the state level (e.g., state's specialization in a particular occupational sector). M_0 and I_0 are unknown thresholds. Finally, ϵ_{ik}^M and ϵ_{ik}^I are non-systematic errors which follow $E(\epsilon_{ik}^M | X_{ik}, Z_k, \gamma_k^M) = 0$ and $(\epsilon_{ik}^I | X_{ik}, \gamma_k^I) = 0$.

Given the setup of binary outcomes with a continuous endogenous variable, we use maximum-likelihood to estimate a multivariate probit model, which we will refer to by following the common practice of mentioning it as a IV-Probit model.[6]

5.2 Estimation

As discussed above. we estimated a limited-information maximum-likelihood model for the probability of an individual being in the informal sector as a function of the duration of migration (for rural-to-urban migrants), demographic characteristics, household characteristics, religion and family background information, shown in Table 4. In addition we included district- and city-level fixed effects to capture unobserved geographical and regional impacts on an individual's job placement in the informal sector. Clustered standard errors at the state level are employed for all regressions to derive statistical inference. For robustness check of our estimations, we used full sample (column 1) as well as different sub-samples, such as males between 15 and 65 years of age (column 3) and males only (column 5). In all regressions, using different sub-samples, our results are largely consistent and none of the

---

[5] This variable ranks the value of non-durable assets from 1 to 6 Rank 1, the lowest, is households with a total value of non-durables of less than 500 rupees. Rank 6, the highest, is households with asset valued at more than 20,000 rupees. (On 12th March 2013 the exchange rate was: 100 INR=1.84 USD.

[6] Estimations were carried out by using the IVProbit command with MLE option in STATA version 11.2

variables changed sign. We have also reported the marginal effects of all estimations in the respective sub-sample estimations in columns 2, 4 and 6 respectively. To show consistency in our estimation, we have also estimated a simple probit model without treating the duration of migration as endogenous, shown in column 7. The probit result shows a small and negative but statistically weak significance of migration duration on probability of someone being in the informal sector. Once we instrument for migration duration in columns 1 to 6, however, this effect becomes larger and statistically more significant.

[Table 4 is about here]

We would expect the higher the education of an individual, the lower the probability of belonging to the informal sector. The evidence, see Table 4, shows clearly that the higher the level of education of the household head the lower the probability of being in the informal sector, and the coefficients get smaller (bigger in absolute value) respectively. The results for father's education are very similar to the household head's education level. Further, we would expect that if the father of the individual was of a higher social class (in terms of occupation), the probability of being in the informal sector would be lower. Again the evidence supports the view that the parent's occupation clearly influences an individual's employment placement: if the father's occupation is formal in nature like executive or clerk, the probability of being in the informal sector is lower, while if the father's occupation is sales (which is mostly informal in nature in the Indian context), then there is a higher probability of being in the informal sector. As discussed earlier, we would expect a person from a socially disadvantaged caste or religion more likely to be in the informal sector. Our findings show that OBC (Other Backward Classes) and Muslims are more likely to be in the informal sector. We did not find any statistically significance for Dalits (the lower social castes) in the informal sector compared with Brahmins, which may be attributable to the government's positive discrimination in employment in the government sector (reservation system) for Dalits. As we would expect, the wealthier the household head, the less likely s/he would be in the informal sector. Our results suggest that urban natives are more likely to be in the formal sector as they have more access to better schools, social networks and job information and referrals compared with the rural-to-urban migrants. Our main variable of interest is rural migration duration. In all cases it is negative and significant at the 1% level. In other words, the longer a rural migrant has been in an urban area the more likely that

15

individual would have moved to the formal sector. (Note the rural migration duration variable has been instrumented).

The validity of IV estimations depends on the power of instruments in explaining the predicted values at the first stage. As reported, all the first stage regressions have very high F-statistics (for example, for our preferred specification of column 1, the first stage F-statistic is 31.77). We have also tested for the joint significance of our IVs, and the results overwhelmingly reject the null of no joint significance. The Wald statistic of the exogeneity test has rejected the null hypothesis of no endogeneity. We have also tested the instruments using the typical 2SLS models to test for the over-identification test (Anderson canonical correlations test) and under-identification test (Sargan-Hansen test) which have duly supported our instruments.

6.0 Multinomial Estimation (Formal, Self-Employed, and Informal Wage)

In this section we have separated the informal sector into those who are self-employed and those who work in the informal or formal sectors as wage labourers to check whether or not highly qualified individuals are employed in the formal sector and also to check if migrants use the informal sector as their temporary base. We did this check by employing a multinomial logit job attainment model following the work of Xin Meng (2001).

6.1 Econometric modelling

Standard neo-classical economic rationality for an individual's job placement (labour supply) is a function of individual endowments and human resources (e.g., level of education and experience). However, other related factors that could also have an impact on an individual's labour supply, especially in the context of India, could be the family size (Brown *at el.* 1980), family background, caste and religious affiliation (see for example Banerjee and Knight (1985) or Ito (2009)). Another less frequently studied factor that might be critical is the job related network, for example, job-opening information, formal and informal channels of job search and referral (see for example Holzer 1987 or Calvó-Armengol, A., & Zenou, Y. (2005)). Since urban natives usually have a better endowment of job-related networks and referrals, we could hypothesise that migrants will acquire access to such networks as their migration duration increases and hence are less likely to be in informal wage labour.

A multinomial logit model is specified below to capture how these variables will influence an individual $j's$ probability of working in sector s. Formally the model is:

$$P_{js} = prob(y_j = sector_s) = \frac{e^{x_j'\beta_s}}{\sum_{s=1}^{S} e^{x_j'\beta_l}} \quad j = 1, \ldots\ldots, N; s = 1, \ldots\ldots, S. \quad (3)$$

where N is the size of the sample, S is the number of sectors and x_j is a vector of variables affecting the labour placement outcome y_j. The dependent variable y_j for equation (3) is the nature of an individual's current job in any of the three sectors: formal sector, self-employment or informal wage labour sector. Our main variable of interest, years of migration duration, is endogenous in nature. Hence to allow for the endogeneity in estimating equation (3), we first used the fitted value of migration duration using all the instruments (estimation done through OLS). In the second step we used the fitted value of the migration duration in equation 3. The standard errors of the estimates in the second step have been estimated through a bootstrapping process with 100 replications.

6.2 Estimations

The results of marginal effects of endogenous multinomial probit estimations are reported in Table 5 with full sample and in Table 6 with male only sub-sample. The dependent variable has been categorized into three groups with formal sector employment being used as a base category.

Tables 5 and 6 are broadly similar to those reported in section 5 of the IV-probit model (Table 4). For both the self-employed and informal wage labour sectors, education, father's characteristics, and the caste and religion coefficients have essentially the same signs and significance. The main point of difference is that a rural migrant's duration of migration does not significantly influence the probability of being in the self-employed sector, but is negative and significant for the informal wage labour sector. In other words, we find that the longer a rural migrant has been in the urban sector the less likely he would be in informal wage employment. For instance, an individual with one additional year's of migration duration from the rural area reduces his/her probability to be in informal wage employment by 2.52%. However, this variable has no statistically discernible effect on him/her being in the self-employed or formal sectors. Results on education are consistent with the other findings that with more years of education, individuals will less likely be in the informal sector. In the case of self-employment, education up to the tertiary level does not have any

17

statistically significant impact; however, for tertiary level education and higher, the probability of someone being self-employed is significantly reduced.

[Tables 5 and 6 about here]

The validity of multinomial regression lies on the strong assumption of the Independence of Irrelevant Alternatives (IIA), which means that adding or deleting alternative outcome categories does not affect the odds among the remaining outcomes. To check whether this assumption holds in our case, we have performed the test for the IIA assumption, and we find no evidence violating the assumption (using the full-sample specification of Table 5).

7. Conclusions

In our paper we have defined migrants as those individuals who have migrated from rural to urban areas. Those who were born in urban areas and migrated to another urban area are not considered as migrants.[7] Also note that in our multinomial logit regressions, for the sake of simplicity of estimation, we used only the rural-to-urban migration duration as endogenous and properly took care of such endogenous regression by using instruments to predict the fitted value of the variable and plugged in the fitted value in the final multinomial regression. One could argue, however, that urban-to-urban migration could also be endogenous. We have also used urban-to-urban migration as endogenous in separate regression estimations in the multinomial logit framework (not reported) and in a linear probability model, and in both cases the variable was insignificant and did not appear to be influential in explaining the likelihood of person's placement in the informal labour market.

In this paper we have argued that there are segmented labour markets in India's urban sector: people who are from the lower social classes (castes or religions) are more likely to work in the informal sector. We found that getting more education is one way of getting a job in the formal sector, but perhaps more important are family networks in providing an entry into the formal labour market. We argued that when rural migrants move to the urban sector, they initially find themselves working in the informal sector where they have lower incomes and work in industries like construction, manufacturing, wholesale and retail trades,

---

[7] Those who were born in other countries are not part of the sample in our estimations.

restaurants and hotels, transport, and social and personal services. Their occupations are mainly in the lower social grades, such as production and related work, transport, sales and service work, and working as labourers. We noted that caste and religion is important. The principal source of income for Dalits and Muslims is non-agricultural labour or as artisans. Brahmins and high caste people are more likely to be in higher level occupations.

We argued that there is a hierarchy of preferences, that people would prefer to work in the formal sector, with the self-employed sector being the second choice, and lastly working in the informal wage labour market. However, entry into the formal sector is constrained by education, social class, and family ties. Self-employment is constrained by access to the credit market.

We estimated a model of the probability of working in the informal sector as a function of demographic characteristics, education, father's education and occupation, caste and religion, and duration of a migrant in his/her present occupation. We distinguished between migrants who had come from rural areas from those who had moved from other urban areas. We treated the duration of the migrant as an endogenous variable and estimated a two-stage least-squares model. We found that most of the explanatory variables were significant and of the expected signs. In particular, we found that education and father's education and occupational status were important. Muslims and Other Backward Classes were more likely to be working in the informal sector.

The most interesting finding of our research is that the longer a rural migrant has been working in the urban sector, the more likely s/he will have moved out of the informal wage sector. The results support the view that for migrants, the informal wage labour market may be a stepping stone to a better life in the formal sector.

However, using a cross-sectional data set to analyse migration and urban employment is a challenging task. Migrants have a higher attrition probability due to the mobility of the population. Hence, when a researcher is confronted with a migrant population, it is difficult to define the population at hand, as there are constant inflows and outflows of individuals with different traits. Moreover, duration raises the possibility of right censoring which could not be addressed with the data at hand. These results need to be researched further using panel data, which unfortunately are not available as yet.

References

Amemiya, T. 1978. The estimation of a simultaneous equation generalized probit model. Econometrica 46: 1193–1205.

Banerjee, B., & Knight, J. B. (1985). Caste discrimination in the Indian urban labour market. *Journal of Development Economics*, *17*(3), 277-307.

Bennett, John and Matthew D. Rablen (2012) "Self-Employment, Wage Employment and Informality in a Developing Economy", IZA Discussion Paper No. 6406.

Bromley, Ray (1978) "Introduction-The Urban Informal Sector: Why Is It Worth Discussing?" *World Development*, 6 (9/10), 1033-1039.

Brown, R. S., M. Moon and B. S. Zoloth, (1980), Incorporating occupational attainment in studies of male/female earnings differentials. *Journal of Human Resources*, **15**, pp. 3–28.

Calvó-Armengol, A., & Zenou, Y. (2005). Job matching, social network and word-of-mouth communication. *Journal of Urban Economics*, *57*(3), 500-522.

De Mel, S., McKenzie, D., & Woodruff, C. (2010). Who are the Microenterprise Owners? Evidence from Sri Lanka on Tokman versus De Soto. In International Differences in Entrepreneurship (pp. 63-87). University of Chicago Press.

Fields, G. S. (2009). "Segmented Labor Market Models in Developing Countries." The Oxford handbook of philosophy of economics: 476-510.

Funkhouser, Edward (1996) "The Urban Informal Sector in Central America: Household Survey Evidence", *World Development*, 24 (11), 1737-1751.

Gerxhani, Klarita (2004) "The informal sector in developed and less developed countries: A literature survey", *Public Choice*, 120, 267-300.

Hanson GH, Woodruff C (2003) Emigration and educational attainment in Mexico. Mimeo, University of California at San Diego

Hart, Keith (1973) "Informal Income Opportunities and Urban Employment in Ghana', *Journal of Modern African Studies*, 11 (1), 61-89.

Hildebrandt N, McKenzie DJ (2005) The effects of migration on child health in Mexico. Economia 6(1):257–289

Hnatkovska, Viktoria, Amartya Lahiri and Sourabh Paul (2012) "Castes and Mobility", *American Economic Journal: Applied Economics*, 4 (2), 274-307.

Holzer, H. J. (1987). Informal job search and black youth unemployment. *The American Economic Review*, 446-452.

International Labour Organisation (1972) *Employment, Incomes and Equality: A Strategy for Increasing Productive Employment in Kenya*, ILO, Geneva.

International Labour Organisation (2002) *Decent Work and the Informal Economy*, International Labour Conference, Report VI, ILO, Geneva.

Inter-university Consortium for Political and Social Research (2005) *India Human Development Survey* (IHDS) 2005, Ann Arbor, Michigan, USA.

Ito, T. (2009). Caste discrimination and transaction costs in the labor market: Evidence from rural North India. *Journal of development Economics*, *88*(2), 292-300.

López-Córdoba E (2005) Globalization, migration, and development: the role of Mexican migrant remittances. Economia 6(1):217–256

Madheswaran, S. and Paul Attwell (2007) "Caste Discrimination in the Indian Urban Labour Market: Evidence from the National Sample Survey", *Economic and Political Weekly*, 42 (41), Oct. 13-19, 4146-4153.

Mazumdar, Dipak (1976) "The Urban Informal Sector", *World Development*, 4 (8), 655-679.

McKenzie DJ, Rapoport H (2007) Network effects and the dynamics of migration and inequality: theory and evidence from Mexico. *Journal of Development Economics*, 84(1):1–24

McKenzie, D., & Rapoport, H. (2011). Can migration reduce educational attainment? Evidence from Mexico. Journal of Population Economics, 24(4), 1331-1358.

Meng, Xin (2001) "The Informal Sector and Rural-Urban Migration – A Chinese Case Study", *Asian Economic Journal*, 15 (1), 71-89.

Mitra, Arup (2004) "Informal Sector, Networks and Intra-City Variations in Activities: Findings from Delhi Slums", The Applied Regional Science Conference, RURDS, 16 (2), 154-169.

Mitra, Arup (2008) The Indian Labour Market: An Overview, ILO Asia-Pacific Working Paper Series, New Delhi.

National Sample Survey Office (2012) *Informal Sector and Conditions of Employment in India*, NSS 66th Round, July 2009-June 2010, National Statistical Organisation, Ministry of Statistics and Programme Implementation, Government of India.

Rauch, James E. (1991) "Modelling the informal sector formally", *Journal of Development Economics*, 35, 33-47.

Sengupta, Arjun K. (2009) *The Challenge of Employment in India: An Informal Economy Perspective*, Report of the National Commission for Enterprises in the Unorganised Sector, Government of India, Volumes 1 and 2, Academic Foundation, New Delhi.

Thomas, J.J. (1992) *Informal Economic Activity*, LSE Handbooks in Economics, Harvester Wheatsheaf, Hemel Hempstead, Herts., UK.

Weeks, John (1975) "Policies for Expanding Employment in the Informal Urban Sector of Developing Economies", *International Labour Review*, 111, 1-13.

Woodruff C, Zenteno RM (2007) Remittances and micro-enterprises in Mexico. *Journal of Development Economics*, 82(2):509–528

Appendix:

Figure 1: Distribution of Employment across Industries
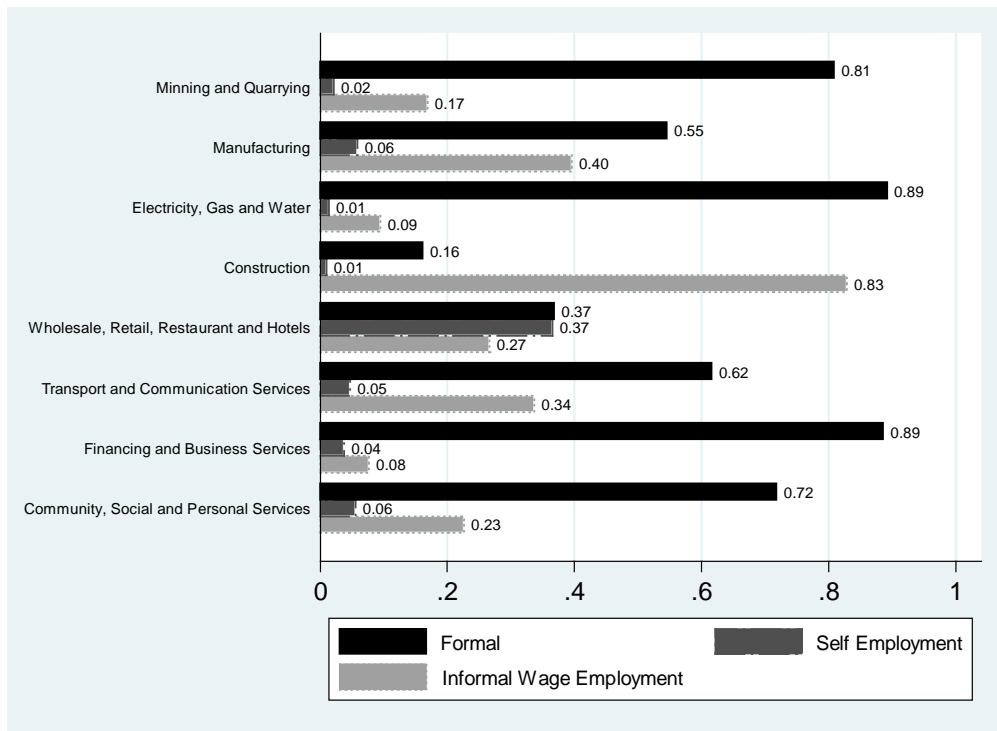
Figure 2: Distribution of Households across Occupations
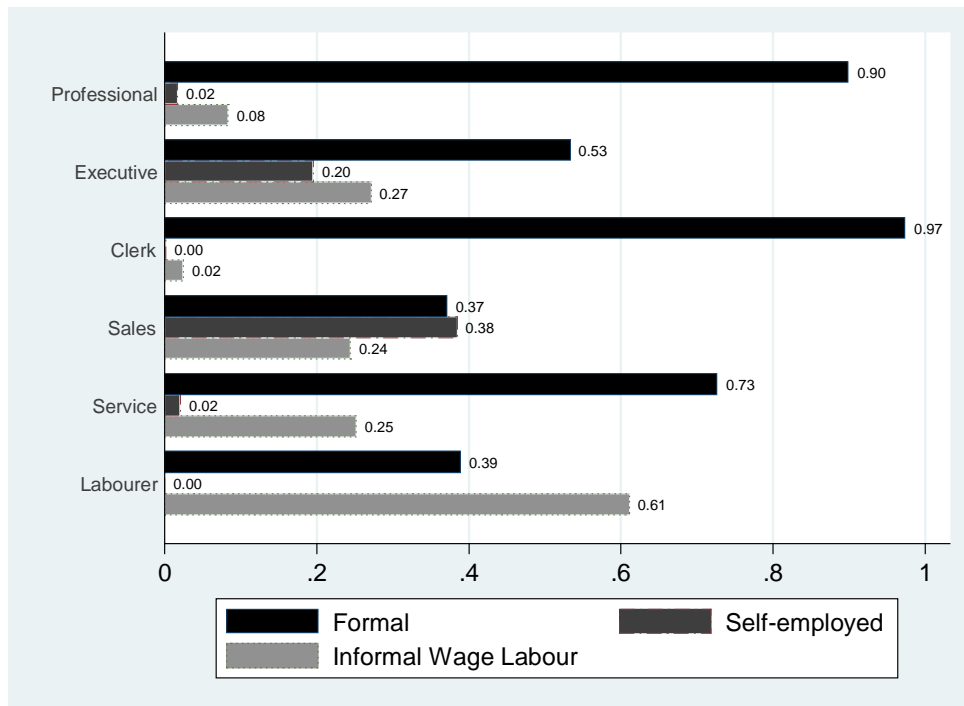
Figure 3: Employment Category Based on Migration Status
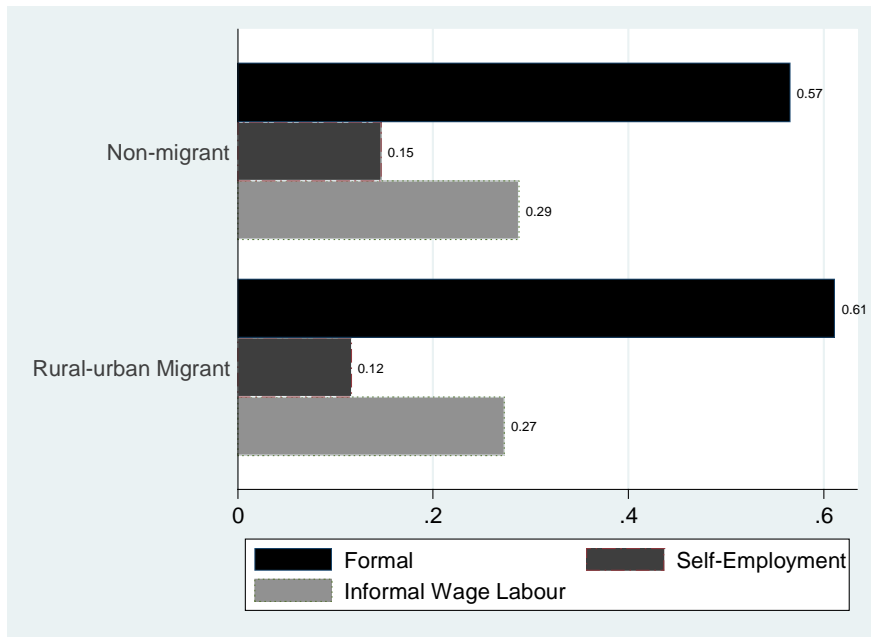
Figure 4: Migrants by Industry
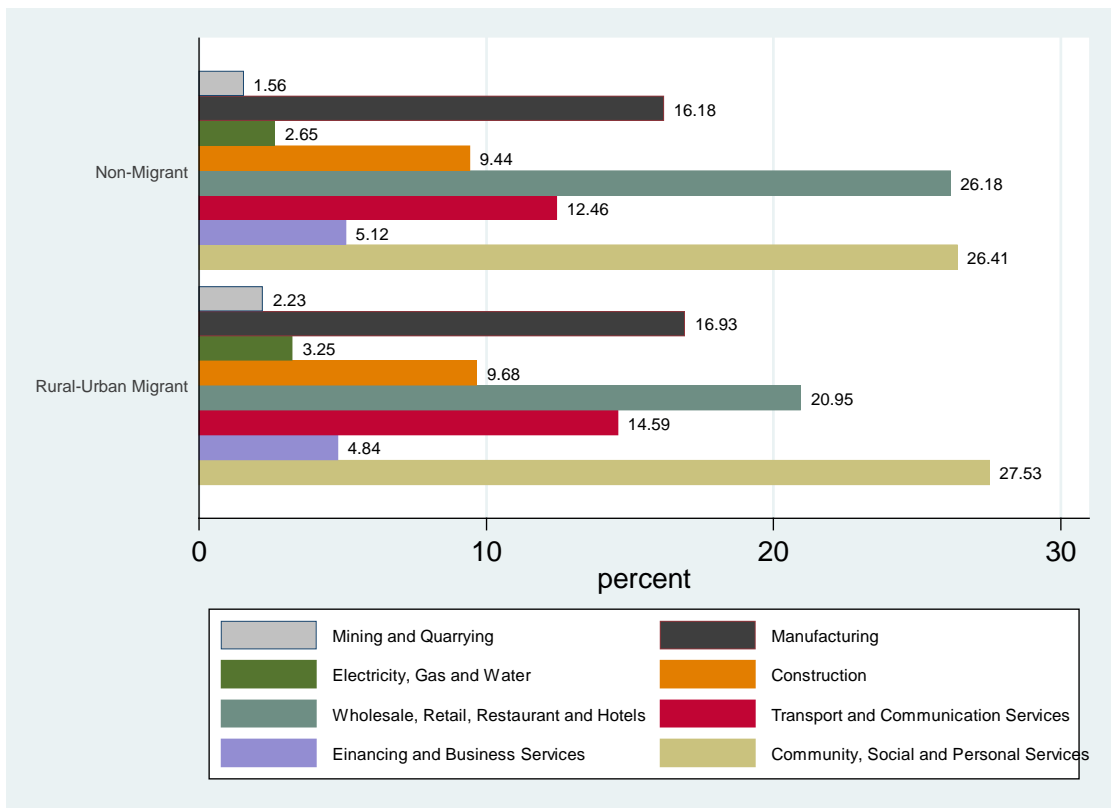
Figure 5: Migrants by Occupation
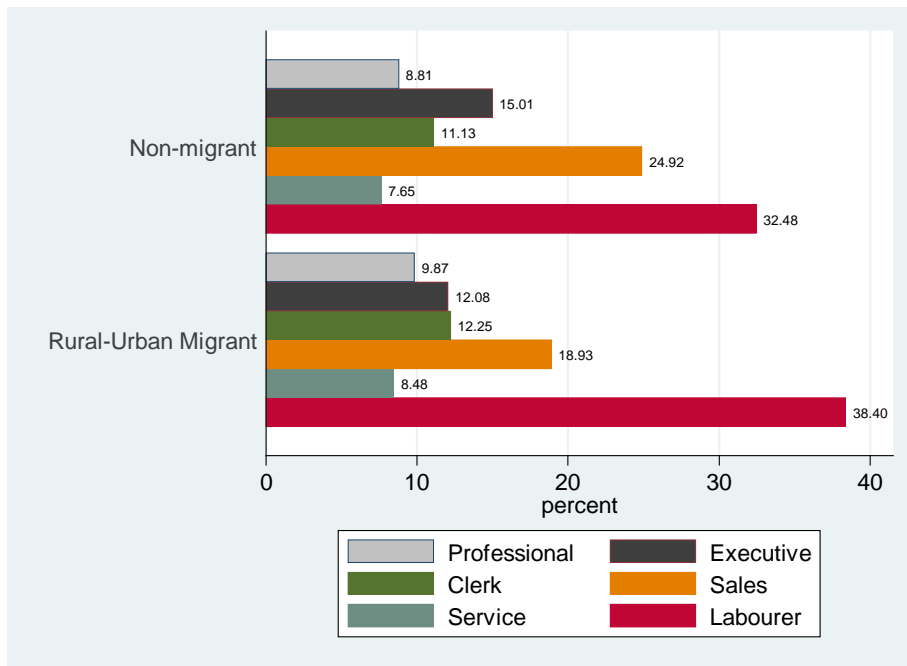
Figure 6: Migrants and Income Source
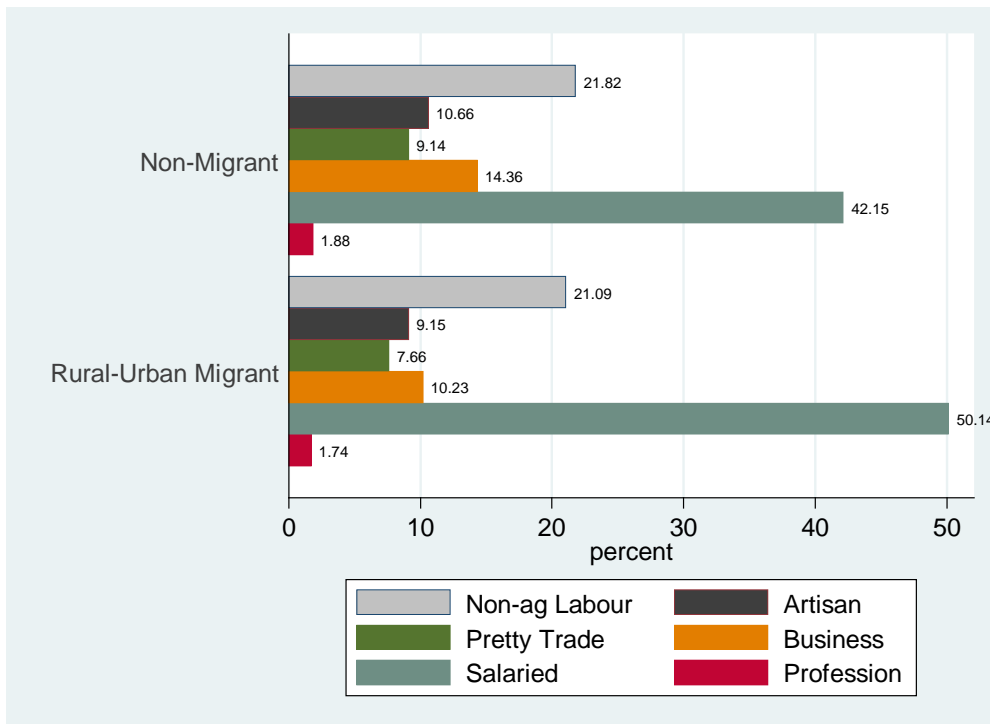
Figure 7: Caste and Religion by Sector
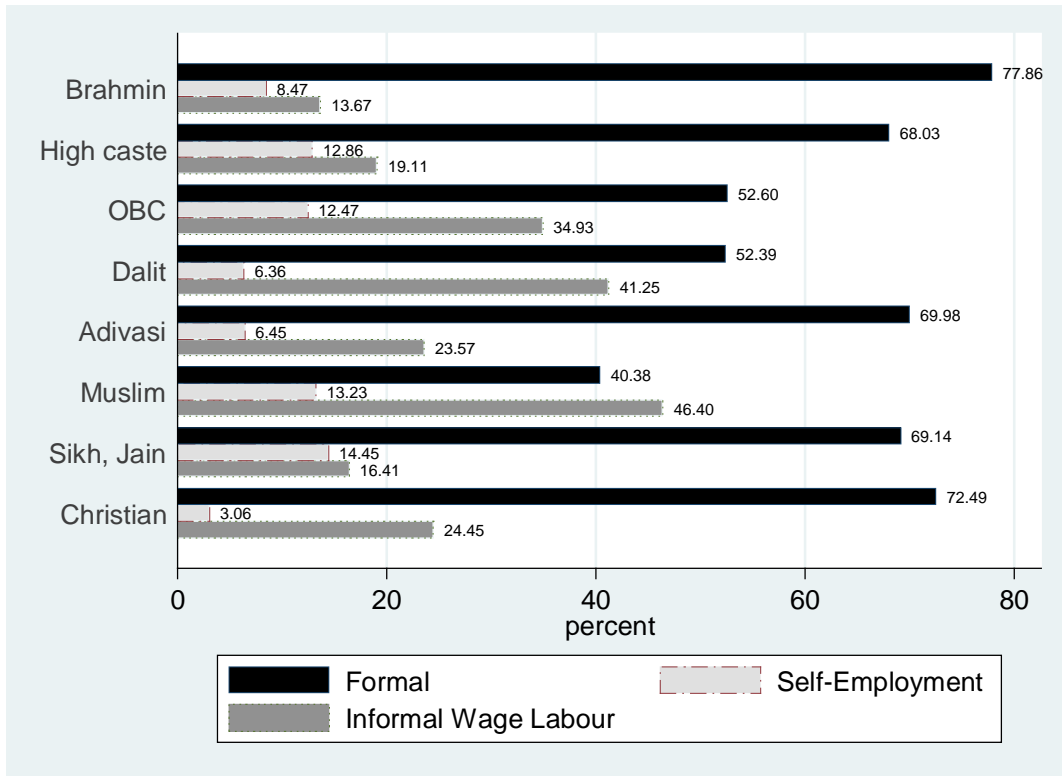
Figure 8: Kernel Densities of Log Income by Employment
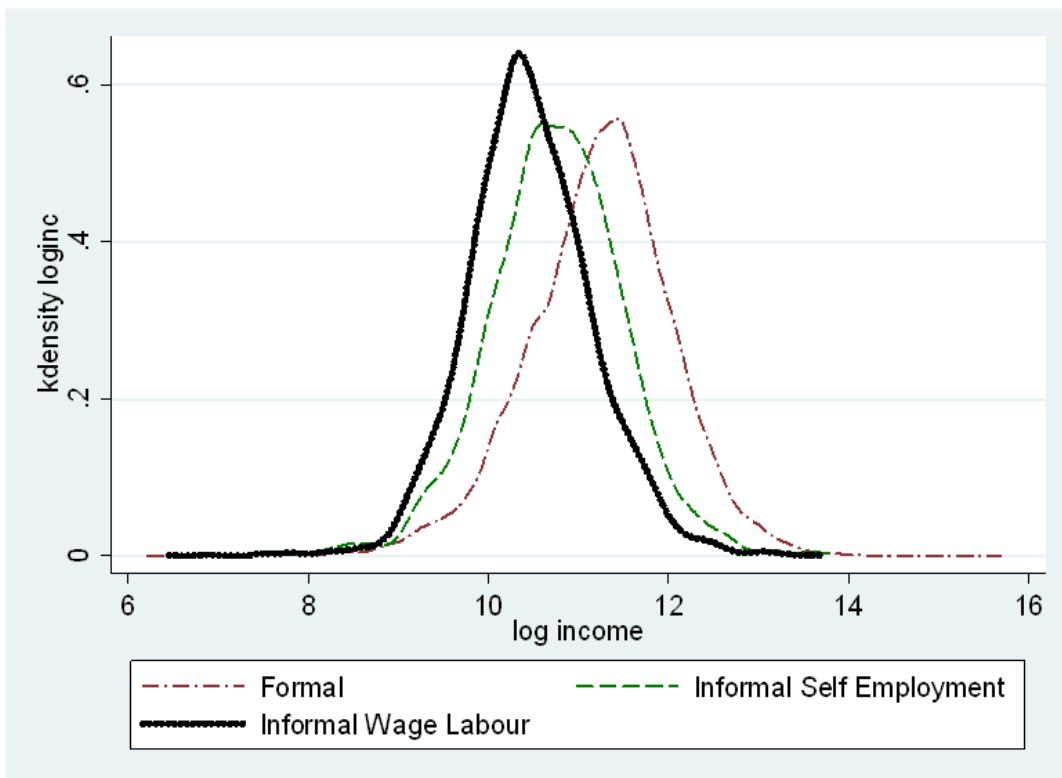
Figure 9: Purpose of Loan by Sector
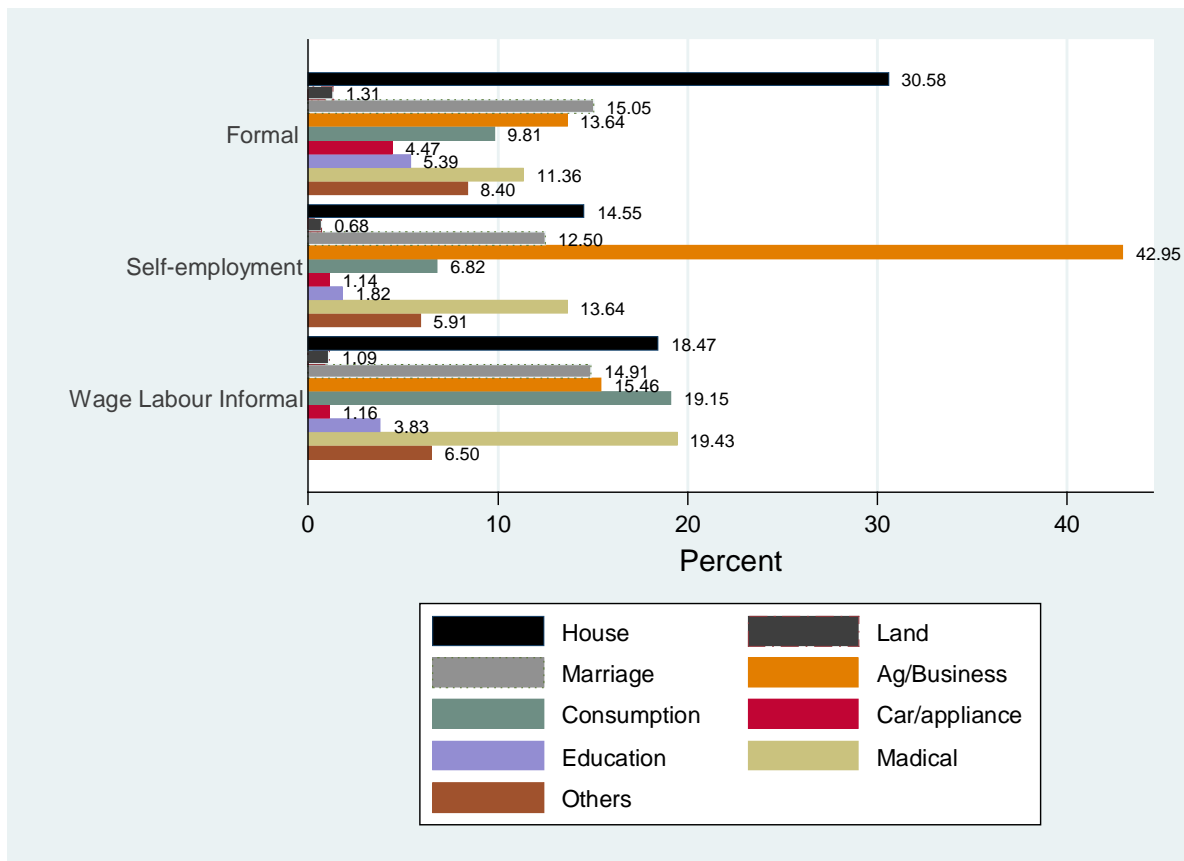
Table 1: Caste and Religion by Source of Income

	Non-Ag labour	Artisan	Petty traders	Business	Salaried	Professionals	Total
Brahmin	56	67	68	136	705	43	1,075
High Caste	254	182	277	536	1,429	59	2,737
OBC	875	437	341	446	1,438	56	3,593
Dalit	664	205	105	108	803	18	1,903
Adivasi	97	11	16	35	238	6	403
Muslim	598	295	211	256	471	29	1,860
Sikh, Jain	9	20	32	61	129	5	256
Christian	54	19	4	20	126	6	229
Total	**2,607**	**1,236**	**1,054**	**1,598**	**5,339**	**222**	**12,056**

Source: India Human Development Survey

Table 2: Caste and Religion by Occupation

	Brahmin	High caste	OBC	Dalit	Adivasi	Muslim	Sikh, Jain	Christian	Total
Professions, Technical and Related Workers	195	280	245	105	49	72	27	22	**995**
Administrative, Executive and Managerial Workers	135	357	427	147	28	243	35	27	**1,399**
Clerical and Related Workers	188	329	361	191	50	84	23	24	**1,250**
Sale Workers	190	746	765	235	52	445	97	21	**2,551**
Service Workers	71	172	210	248	44	87	11	18	**861**
Production, Transport and Labourers	159	551	1,236	799	141	732	41	67	**3,726**
Missing	137	302	349	178	39	197	22	50	**1,274**
Total	**1,075**	**2,737**	**3,593**	**1,903**	**403**	**1,860**	**256**	**229**	**12,056**

Source: India Human Development Survey

Table 3: Distribution of Log Incomes by Sector

Variable: Log of Income	Obs.	Mean	Std. Dev.	Min	Max
Formal	6916	11.2313	0.81603	6.21461	15.6904
Self-Employment	1324	10.7466	0.76783	6.8024	13.7695
Informal Wage Labour	3744	10.4617	0.70924	6.44883	13.731

Source: India Human Development Survey

Table 4: IV-Probit Estimates of the Probability for Informal Sector Employment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
			Male, Aged between 15 to 65				Full Sample Probit
Dependent Varible:	Full Sample				Male only		
Informal Sector Employment	Coefficient	M.E.	Coefficient	M.E.	Coefficient	M.E.	Coefficient
Urban Native	-0.535***	-0.195***	-0.551***	-0.202***	-0.541***	-0.199***	-0.152***
	(0.10)	(0.04)	(0.11)	(0.04)	(0.11)	(0.04)	(0.05)
Rural to urban migration duration	-0.076***	-0.029***	-0.079***	-0.031***	-0.076***	-0.029***	-0.004*
	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.00)
Age	-0.009	-0.003	-0.023	-0.009	-0.007	-0.003	-0.037***
	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)
Age Square	0.0000	0.0000	0.000**	0.000**	0.0000	0.0000	0.000***
	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Male	0.423***	0.153***					0.518***
	(0.09)	(0.03)					(0.09)
No. of Households	0.026***	0.010***	0.029***	0.011***	0.027***	0.010***	0.036***
	(0.01)	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.01)
Married	0.065	0.025	0.044	0.017	0.038	0.015	0.015
	(0.10)	(0.04)	(0.11)	(0.04)	(0.11)	(0.04)	(0.13)
Primary Education	-0.128*	-0.049*	-0.079	-0.03	-0.072	-0.028	-0.172**
	(0.07)	(0.03)	(0.07)	(0.03)	(0.07)	(0.03)	(0.09)
Secondary Education	-0.300***	-0.113***	-0.274***	-0.104***	-0.260***	-0.099***	-0.364***
	(0.08)	(0.03)	(0.08)	(0.03)	(0.08)	(0.03)	(0.08)
Matric Completed	-0.539***	-0.194***	-0.530***	-0.193***	-0.523***	-0.191***	-0.652***
	(0.10)	(0.03)	(0.09)	(0.03)	(0.10)	(0.03)	(0.06)
Tertiary Education	-0.693***	-0.240***	-0.682***	-0.239***	-0.678***	-0.237***	-0.865***
	(0.13)	(0.03)	(0.12)	(0.03)	(0.12)	(0.03)	(0.09)
Graduate	-0.928***	-0.327***	-0.946***	-0.334***	-0.921***	-0.328***	-1.150***
	(0.15)	(0.04)	(0.14)	(0.04)	(0.14)	(0.04)	(0.08)
High caste	0.112***	0.044**	0.062	0.024	0.105**	0.041**	0.119**
	(0.04)	(0.02)	(0.04)	(0.02)	(0.05)	(0.02)	(0.05)
OBC	0.140***	0.055***	0.132**	0.051**	0.149***	0.058***	0.220***
	(0.05)	(0.02)	(0.05)	(0.02)	(0.05)	(0.02)	(0.07)
Dalit	0.041	0.016	-0.013	-0.005	0.018	0.007	0.061
	(0.06)	(0.02)	(0.05)	(0.02)	(0.05)	(0.02)	(0.07)
Adivasi	-0.207**	-0.078**	-0.244***	-0.091***	-0.209**	-0.079**	-0.263***
	(0.09)	(0.03)	(0.09)	(0.03)	(0.10)	(0.04)	(0.09)
Muslim	0.166**	0.065**	0.158**	0.062**	0.186**	0.073**	0.318***
	(0.08)	(0.03)	(0.07)	(0.03)	(0.08)	(0.03)	(0.07)
Sikh, Jain	0.078	0.031	-0.032	-0.012	0.049	0.019	0.183*
	(0.09)	(0.03)	(0.08)	(0.03)	(0.08)	(0.03)	(0.09)
Christian	-0.003	-0.001	0.041	0.016	-0.002	-0.001	-0.002
	(0.13)	(.)	(0.15)	(0.06)	(0.15)	(0.06)	(0.12)
Father's Occupation: Professional	-0.103	-0.039	-0.084	-0.032	-0.076	-0.029	-0.209***
	(0.07)	(0.03)	(0.07)	(0.03)	(0.07)	(0.03)	(0.06)
Father's Occupation: Executive	-0.367***	-0.133***	-0.344***	-0.126***	-0.356***	-0.130***	-0.468***
	(0.10)	(0.03)	(0.12)	(0.04)	(0.11)	(0.04)	(0.11)

Father's Occupation: Clerk	-0.335***	-0.123***	-0.367***	-0.135***	-0.336***	-0.124***	-0.494***
	(0.10)	(0.03)	(0.10)	(0.03)	(0.10)	(0.03)	(0.08)
Father's Occupation: Sales	0.130***	0.051***	0.145***	0.057***	0.146***	0.057***	0.171***
	(0.05)	(0.02)	(0.05)	(0.02)	(0.05)	(0.02)	(0.05)
Father's Occupation: Service	-0.285***	-0.106***	-0.286***	-0.107***	-0.266***	-0.100***	-0.343***
	(0.07)	(0.02)	(0.08)	(0.03)	(0.07)	(0.03)	(0.06)
Father's Occupation: Agro	0.036	0.014	0.018	0.007	0.042	0.016	-0.241***
	(0.11)	(0.04)	(0.11)	(0.04)	(0.11)	(0.04)	(0.04)
Father's Education: Primary	-0.117***	-0.045***	-0.120***	-0.046***	-0.119***	-0.046***	-0.131***
	(0.03)	(0.01)	(0.04)	(0.02)	(0.04)	(0.01)	(0.04)
Father's Education: Secondary	-0.170***	-0.065***	-0.172***	-0.066***	-0.168***	-0.065***	-0.185***
	(0.05)	(0.02)	(0.04)	(0.02)	(0.04)	(0.02)	(0.05)
Father's Education: Tertiary	-0.257***	-0.096***	-0.293***	-0.109***	-0.259***	-0.097***	-0.330***
	(0.09)	(0.03)	(0.09)	(0.03)	(0.10)	(0.03)	(0.07)
Father's Education: Graduation	-0.311***	-0.115***	-0.314***	-0.116***	-0.319***	-0.118***	-0.334***
	(0.10)	(0.03)	(0.10)	(0.03)	(0.09)	(0.03)	(0.10)
Asset Status (1 to 6)	-0.120***	-0.046***	-0.132***	-0.051***	-0.123***	-0.048***	-0.148***
	(0.03)	(0.01)	(0.03)	(0.01)	(0.03)	(0.01)	(0.02)
City Dummies	Yes		Yes		Yes		Yes
District Dummies	Yes		Yes		Yes		Yes
Observations	10,521		9,685		9,668		10,521
Log pseudolikelihood	-42761		-38409		-39067		-5754
chi2	29420		672226		57984		
Wald test of exogeneity	6.579***		5.506***		6.38***		
F-Statistics at First Stage	31.77***		29.38***		30.00		
H0: Coefficient of IVs are zero	66.88***		58.92***		49.44***		

Source: Indian Human Development Survey 2005: Authors own Calculations.

Notes: Standard errors in parentheses, adjusted for clustering at the state level. Significance code: * $p<0.1$, ** $p<0.05$, *** $p<0.01$. M.E. stands for Marginal Effects which have been calculated at the mean. In all these specifications, we consider as migrants only those who have migrated from rural to urban areas for jobs. Those who were born in an urban setup and migrated to another urban area for jobs are not considered as migrants. The best specification is column (1). Estimations used in columns 1-6 are based on Maximum Likelihood Estimator (MLE). The instruments used in the first stage of the regressions are historic state-level migration rates and the interaction of the variable with caste dummies.

Table 5: Marginal Effects of Multinomial Logit Regression (full sample)

VARIABLES	(1) Formal	(2) Self-employment	(3) Informal Wage Labour
Urban Native	0.0466***	-0.00768	-0.0389***
	(0.01)	(0.01)	(0.01)
Rural to urban migration duration	0.0169	0.00723	-0.0241**
	(0.01)	(0.01)	(0.01)
Age	0.00770**	-0.00319	-0.00451
	(0.00)	(0.00)	(0.00)
Age Square	-8.11e-05***	3.34E-05	4.77e-05*
	(0.00)	(0.00)	(0.00)
Male	-0.171***	0.0614***	0.110***
	(0.02)	(0.01)	(0.01)
No. of Households	-0.0152***	0.0018	0.0134***
	(0.00)	(0.00)	(0.00)
Married	0.00122	-0.0383	0.037
	(0.05)	(0.04)	(0.03)
Primary Education	0.0319	0.0308*	-0.0627***
	(0.03)	(0.02)	(0.02)
Secondary Education	0.0935***	0.0252*	-0.119***
	(0.02)	(0.01)	(0.01)
Matric Completed	0.179***	0.00299	-0.182***
	(0.02)	(0.01)	(0.01)
Tertiary Education	0.218***	-0.0103	-0.208***
	(0.02)	(0.01)	(0.01)
Graduate	0.339***	-0.0396***	-0.300***
	(0.02)	(0.01)	(0.01)
High caste	-0.0474*	0.0238	0.0235
	(0.02)	(0.02)	(0.02)
OBC	-0.0743***	0.0322**	0.0421*
	(0.02)	(0.01)	(0.02)
Dalit	-0.0177	-0.0372***	0.0549**
	(0.03)	(0.01)	(0.03)
Adivasi	0.0571*	-0.0288	-0.0283
	(0.03)	(0.02)	(0.03)
Muslim	-0.115***	0.0415*	0.0737***
	(0.03)	(0.02)	(0.03)
Sikh, Jain	-0.0473	0.0457	0.00166
	(0.04)	(0.03)	(0.04)
Christian	0.0358	-0.0655	0.0297
	(0.07)	(0.07)	(0.04)
Father's Occupation: Professional	0.0505**	0.00458	-0.0551***
	(0.02)	(0.02)	(0.02)
Father's Occupation: Executive	0.144***	-0.0153	-0.129***
	(0.03)	(0.03)	(0.02)

Father's Occupation: Clerk	0.144***	-0.0331***	-0.111***
	(0.02)	(0.01)	(0.02)
Father's Occupation: Sales	-0.0764***	0.148***	-0.0716***
	(0.02)	(0.02)	(0.01)
Father's Occupation: Service	0.106***	-0.0123	-0.0940***
	(0.02)	(0.01)	(0.01)
Father's Occupation: Agro	0.0346	-0.0228	-0.0119
	(0.04)	(0.03)	(0.03)
Father's Education: Primary	0.0450***	-0.00915	-0.0358***
	(0.01)	(0.01)	(0.01)
Father's Education: Secondary	0.0706***	-0.0122	-0.0584***
	(0.02)	(0.01)	(0.02)
Father's Education: Tertiary	0.121***	-0.0234	-0.0973***
	(0.03)	(0.02)	(0.02)
Father's Education: Graduation	0.137***	-0.0203	-0.117***
	(0.04)	(0.02)	(0.02)
Asset Status (1 to 6)	0.0615***	-0.00687**	-0.0547***
	(0.00)	(0.00)	(0.00)
City Dummies	Yes	Yes	Yes
State Dummies	Yes	Yes	Yes
District Dummies	Yes	Yes	Yes
Observations	10,521	10,521	10,521

Source: Indian Human Development Survey 2005: Authors Own Calculation.

Note: The base outcome is formal employment. The variable "Rural to urban migration duration" is considered endogenous; hence the fitted value of migration duration has been estimated by OLS using all variables and instruments at the first stage. The instruments used are historic state-level migration rates and the interaction of the variable with caste dummies. Standard errors are in parentheses which have been computed using a bootstrapped method with 100 repetitions. Significance code: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 6: Marginal Effects of Multinomial Logit Regression (male only sample)

VARIABLES	(1) Formal	(2) Self-employment	(3) Informal Wage Labour
Urban Native	0.0502***	-0.00966	-0.0405***
	(0.01)	(0.01)	(0.01)
Rural to urban migration duration	0.0156	0.00831	-0.0239**
	(0.01)	(0.01)	(0.01)
Age	0.00781	-0.004	-0.00381
	(0.01)	(0.00)	(0.00)
Age Square	-7.71e-05**	0.0000397	0.0000375
	(0.00)	(0.00)	(0.00)
No. of Households	-0.0166***	0.00176	0.0149***
	(0.00)	(0.00)	(0.00)
Married	0.00762	-0.0359	0.0283
	(0.05)	(0.03)	(0.04)
Primary Education	0.0057	0.0434*	-0.0491***
	(0.03)	(0.02)	(0.02)
Secondary Education	0.0813***	0.0286	-0.110***
	(0.03)	(0.02)	(0.02)
Matric Completed	0.175***	0.00523	-0.180***
	(0.02)	(0.02)	(0.01)
Tertiary Education	0.215***	-0.00755	-0.207***
	(0.02)	(0.02)	(0.01)
Graduate	0.339***	-0.0363**	-0.302***
	(0.02)	(0.02)	(0.02)
High caste	-0.0412*	0.0239	0.0173
	(0.02)	(0.02)	(0.02)
OBC	-0.0727***	0.0332**	0.0395**
	(0.02)	(0.01)	(0.02)
Dalit	-0.00579	-0.0402***	0.0460**
	(0.03)	(0.01)	(0.02)
Adivasi	0.0588*	-0.0238	-0.035
	(0.03)	(0.02)	(0.03)
Muslim	-0.122***	0.0437**	0.0783***
	(0.03)	(0.02)	(0.03)
Sikh, Jain	-0.0336	0.0502	-0.0166
	(0.05)	(0.04)	(0.04)
Christian	0.0382	-0.0689	0.0307
	(0.07)	(0.06)	(0.04)
Father's Occupation: Professional	0.0392	0.00944	-0.0486**
	(0.03)	(0.02)	(0.02)
Father's Occupation: Executive	0.138***	-0.0172	-0.121***
	(0.04)	(0.03)	(0.03)
Father's Occupation: Clerk	0.144***	-0.0251*	-0.119***
	(0.02)	(0.01)	(0.02)

Father's Occupation: Sales	-0.0816***	0.158***	-0.0760***
	(0.02)	(0.02)	(0.01)
Father's Occupation: Service	0.100***	-0.0051	-0.0951***
	(0.02)	(0.01)	(0.02)
Father's Occupation: Agro	0.0377	-0.0239	-0.0138
	(0.04)	(0.03)	(0.04)
Father's Education: Primary	0.0455***	-0.0101	-0.0354***
	(0.01)	(0.01)	(0.01)
Father's Education: Secondary	0.0700***	-0.0131	-0.0570***
	(0.02)	(0.01)	(0.01)
Father's Education: Tertiary	0.123***	-0.025	-0.0983***
	(0.03)	(0.02)	(0.02)
Father's Education: Graduation	0.141***	-0.0205	-0.121***
	(0.03)	(0.02)	(0.03)
Asset Status (1 to 6)	0.0632***	-0.00656**	-0.0567***
	(0.01)	(0.00)	(0.00)
City Dummies	Yes	Yes	Yes
State Dummies	Yes	Yes	Yes
District Dummies	Yes	Yes	Yes
Observations	9668	9668	9668

Source: Indian Human Development Survey 2005: Authors Own Calculation.

Note: The base outcome is formal employment. The variable "Rural to urban migration duration" is considered endogenous; hence the fitted value of migration duration has been estimated by OLS using all variables and instruments at the first stage. The instruments used are historic state-level migration rate and the interaction of the variable with caste dummies. Standard errors are in parentheses which have been computed using a bootstrapped method with 100 repetitions. Significance code: *** p<0.01, ** p<0.05, * p<0.1

Table 7: Summary Statistics

	Formal	Informal	Total
Male	0.902	0.935	0.916
	(0.297)	(0.246)	(0.277)
Rural to urban migration	0.283	0.246	0.267
	(0.45)	(0.431)	(0.443)
Rural to urban migration duration	4.337	3.595	4.023
	(9.077)	(8.355)	(8.787)
Income (in Rupees)	103171.280	49812.840	80625.834
	(124176.363)	(51689.736)	(103573.581)
Age	45.968	43.181	44.791
	(11.936)	(12.299)	(12.169)
Size of the Household	4.842	5.066	4.937
	(2.089)	(2.089)	(2.092)
Married	0.984	0.983	0.984
	(0.124)	(0.128)	(0.126)
Primary Education	0.053	0.148	0.093
	(0.223)	(0.355)	(0.290)
Secondary Education	0.168	0.301	0.224
	(0.374)	(0.459)	(0.417)
Matriculation Complete	0.177	0.170	0.174
	(0.382)	(0.376)	(0.379)
Tertiary Education	0.163	0.099	0.136
	(0.370)	(0.299)	(0.343)
Graduate	0.399	0.123	0.283
	(0.490)	(0.328)	(0.450)
Adivasi	0.041	0.024	0.033
	(0.197)	(0.152)	(0.180)
Dalit	0.143	0.178	0.158
	(0.350)	(0.382)	(0.365)
Muslim	0.108	0.218	0.154
	(0.310)	(0.413)	(0.361)
Father's Occupation: Professional	0.111	0.056	0.089
	(0.314)	(0.230)	(0.285)
Father's Occupation: Executive	0.030	0.010	0.022
	(0.169)	(0.098)	(0.145)
Father's Occupation: Clerk	0.108	0.036	0.079
	(0.311)	(0.186)	(0.270)
Father's Occupation: Sales	0.132	0.189	0.155
	(0.338)	(0.392)	(0.362)
Father's Occupation: Service	0.113	0.085	0.102
	(0.316)	(0.279)	(0.302)
Father's Occupation: Agro	0.352	0.341	0.347
	(0.478)	(0.474)	(0.476)
Father's Occupation: Labourer	0.155	0.284	0.207
	(0.362)	(0.451)	(0.405)
Father's Education: Primary	0.222	0.226	0.224
	(0.416)	(0.418)	(0.417)
Father's Education: Secondary	0.236	0.140	0.196
	(0.425)	(0.347)	(0.397)
Father's Education: Tertiary	0.050	0.014	0.035
	(0.218)	(0.119)	(0.184)
Father's Education: Graduation	0.063	0.013	0.042
	(0.243)	(0.115)	(0.201)
N	6962	5094	12056