Comments Welcome

Gender, Generations, and Non-farm Participation

M. Shahe Emran[1]
Stanford University and
George Washington University

Misuzu Otsuka
Johns Hopkins University

Forhad Shilpi
World Bank

First Draft: April 2002
Revised: July 2002

Abstract

This paper presents some first empirical evidence on intergenerational linkages in non-farm participation in a developing country with a focus on gender effects. The econometric results, using household survey data from Nepal, show that while there are unambiguous evidence of strong linkages between mothers and daughters in non-farm participation, the link between fathers and sons is almost non-existent. The estimates imply that having a mother in the non-farm sector raises a daughter's probability of non-farm participation by almost four times. The effects are truly dramatic for skilled non-farm jobs, as the probability of participation increases by thirteen times when the mother is also in non-farm compared to the case where mother is employed in agriculture. Having a father in non-farm, on the other hand, does not have any significant (numerically or statistically) effect on a son's probability of non-farm participation when the endogeneity of education and assets are taken care of by Two Stage Conditional Maximum Likelihood approach. There are, however, evidence of a moderate intergenerational correlation between fathers and sons for skilled jobs. The analysis has important policy implications as the occupational mobility seems to have a distinct gender bias against women.

[1]Emails for correspondence: shahe_emran@hotmail.com, fshilpi@worldbank.org, motsuka@jhu.edu.

We would like to thank Gershon Feder, Hanan Jacoby, Peter Lanjouw, and David Ribar for useful discussions and/or comments on earlier drafts. The standard disclaimers apply.

Introduction

The evolution of income distribution, inequality and occupational structure across generations has attracted increasing attention in recent economic literature. This renewed interest reflects a widely shared consensus that the existence of intergenerational linkages in socioeconomic status has profound implications for mobility (or immobility) in a society. Starting from identical initial income distribution or occupational structure, the evolution of two societies might diverge dramatically if the strength of intergenerational linkages differs significantly.[1] A large body of empirical studies focusing mainly on developed countries find that intergenerational correlations in earnings are positive, statistically significant and numerically large, ranging between 0.2 to 0.5 (for a survey, see Solon (1999)). There is also a (relatively) small empirical literature that indicates significant positive correlations between parents and their offspring in occupational choices.[2] In this paper, we focus on intergenerational correlations in non-farm employment in a developing country, Nepal, with particular emphasis on the gender dimension of these correlations.[3] Although there is a substantial literature on the determinants of non-farm participation (see Lanjouw and Feder (2001) for an excellent survey), the issue of intergenerational linkages has so far not received any attention. A vast literature also explores gender effects in intra-household allocation of resources in developing countries (Kanbur and Haddad, 1994; Haddinott and

[1]At one extreme, the socioeconomic opportunities available to an individual in a society might be effectively determined by birth, as in a caste system. At the other extreme is a perfectly mobile society, in which the set of opportunities is same irrespective of the exogenous characteristics of an individual, like birth, gender, ethnicity. It is often argued that mobility is much more restricted in developing countries (see, for example, Lam and Schoeni, 1993).

[2]Dunn and Holtz-Eakin (2000) using U.S. data find that parent's self employment has large positive effect on son's probability of transition from wage employment to self employment. Sjogren (2000) finds that in Sweden people are more sensitive to economic incentives and more risk averse when considering occupations different from their parents, and poor family background (measured by father's education) increases this sensitivity to economic incentives.

[3]There are a couple of advantages in focusing on occupational correlation instead of income correlation. First, it is not fraught with the almost insurmountable measurement problems in permanent income, the variable of interest in income mobility studies. Second, as Goldberger (1989) pointed out early on, intergenerational linkages might be stronger for occupation choice (relative to income), and focusing on income correlations "could lead an economist to understate the influence of family background on inequality" (P.513).

Haddad, 1995, Thomas, 1997; Quisumbing and Maluccio, 1999). However, to our knowledge, the gender effects in intergenerational linkages in non-farm participation have not been analyzed before.

Although so far neglected in the literature, the analysis of intergenerational linkages and gender effects in non-farm participation has important implications that span a number of policy areas. The available evidence show that poverty in developing countries has a gender dimension; women are among the poorest and chronically deprived segment of the population (World Bank, 2001). At the early stages of development, access to non-farm employment can be a way out for the poor and landless people in general, and women in particular, as has been demonstrated by the micro finance programs like Grameen Bank in Bangladesh. A related but more important point is that the bargaining power of a woman is likely to be positively influenced by her participation in non-farm sector,[4] which in turn has desirable effects on the intra-household allocations.[5] Notwithstanding the benefits of non-farm participation by women, there is a strong gender bias against women in non-farm participation in many developing countries (For instance, in our Nepal data set, the participation rates in the non-farm sector are 41 percent for men and 10 percent for women.[6] It is important, especially from a policy perspective, to understand how much of this gender bias is due to occupational immobility caused by gender-specific intergenerational linkages. If there are strong intergenerational linkages in women's non-farm participation as the evidence discussed later suggest to be the case in Nepal, this implies that the long run benefits from women's participation in non-farm sector are likely to be much higher due to the intergenerational multiplier effect. Thus the standard cost-benefit analysis that ignores

[4]Since much of women's work in agriculture remain unpaid, access to non-farm employment bears special significance as a way to ensure control over income by women.

[5] A large number of empirical studies in the context of developing countries finds that greater resource control by women within the household leads to higher expenditure on family welfare (e.g. on food), especially higher expenditure on children (education, health etc.) and woman herself (for a recent survey, see World Bank, 2001).

[6]For similar evidence from India and Malaysia, see Feder and Lanjouw (2001)). It should be noted that while this pattern holds in many African countries as well, in Latin American countries, participation rates in the non-farm sector are significantly higher for women compared with men.

this multiplier effect is likely to underestimate the value of programs that simultaneously target poor women and encourage their participation in non-farm activities.[7]

We construct a simple theoretical model of education and occupation choice to identify the sources of intergenerational linkages in non-farm participation, which then forms the basis of our empirical analysis. Although there are many different channels through which intergenerational linkages can and do operate, our focus is on the role played by more 'intangible' factors like role model effects, learning externalities, preference and genetic endowment transmissions, and transfer of reputation capital. We explicitly control for education, assets, and network variables in the regressions, and the parental occupation variables thus capture the 'intergenerational linkages' over and above these more tangible (relatively easily measurable) factors. The econometric results, using household level survey data from Nepal, indicate strong intergenerational linkages in non-farm participation when we concentrate on the full children sample, including both sons and daughters.[8] The intergenerational correlations are pretty robust with respect to inclusion of an array of control variables, and with respect to different experiments regarding functional forms and sample definitions. When split into daughters and sons sub-samples, we find evidence of strong linkages that run along gender lines (mother-daughter, father-son) if, following the existing literature, simultaneity of education and occupation decision is ignored. However, the evidence unambiguously reject the null hypothesis of exogeneity, especially for education, in the non-farm participation decisions of sons. The estimated linkages become numerically small and statistically insignificant for sons when endogeneity of education and

[7]A different argument that underscores the importance of understanding the intergenerational linkages in non-farm participation relates to the role of non-farm entrepreneurship in the structural transformation of an economy. A dynamic non-farm sector can be the seedbed for experimentation and development of an entrepreneurial class that eventually graduates to industrial activities, as was the case in Japan's rise to a modern industrial state from late Tokugawa to Meiji period (See Smith, 1988). The existence of strong intergenerational linkages in non-farm participation means that the initial conditions assume paramount importance, and the emergence and development of an industrial entrepreneurial class might be severely constrained when an economy starts with a tiny non-farm sector and a large agricultural sector.

[8]In the full sample, the evidence against exogeneity of education and assets are weak, and the estimates from Two-Stage Conditional Maximum Likelihood are slightly smaller in magnitude, but are still numerically large.

assets are corrected by utilizing the Two-Stage Conditional Maximum Likelihood (henceforth TSCMLE) approach (Blundell and Smith, 1986; Voung and Rivers, 1988). When disaggregated according to skill levels, there are evidence of a positive correlation between father and son(s) only in the case of skilled jobs, thus indicating that the aggregate results are partly driven by the lack of intergenerational correlations in unskilled jobs. For daughters, there are convincing evidence that the endogeneity problems are not important, and the evidence show strong effects of parents', in particular mother's, non-farm participation regardless of levels of skill. Having a mother in the non-farm sector raises a daughter's probability of non-farm participation by four times when the sample consists of both skilled and unskilled groups. The effects are truly dramatic in the case of skilled jobs; daughter's probability of being in skilled job increases by 13 times if mother is employed in skilled job. The cross gender intergenerational links (mother-son, father-daughter) in non-farm participation appear to be unimportant.

The remainder of the paper is organized as follows. Section II provides a conceptual framework that underpins the empirical work presented in the subsequent sections. It presents a systematic analysis of possible sources of intergenerational linkages and gender effects in employment and occupational choice. The Section III describes the empirical specification, while the next section discusses the data and construction of variables. Section V, arranged in a number of sub-sections, presents the empirical results. The first sub-section focuses on the full children sample and discusses the preliminary evidence and thus provides a first impression of the importance of the intergenerational dependence in the non-farm participation. The next subsection presents the results of the empirical analysis that takes into account the potential endogeneity problems. The following two sub-sections are devoted to the analysis of gender effects and the role of skill differences in non-farm jobs. Section VI concludes the paper with a summary of the main findings and some remarks on future research.

II. The Conceptual Framework

In the following we present a simple model of non-farm participation that is based on the standard occupational choice model but is augmented to capture the essentials of the intergenerational linkages.

There are two sectors in the economy: agriculture (a) and non-farm sector (n). There are overlapping generations of people, each with a life span of two periods; in the first period (t_0) they live with the parents, and build up human capital (schooling and/or in-house learning externalities), and at the beginning of the second period (t_1) every person in the economy decides which sector to work for. Each individual is endowed with an innate ability $\theta_i \in [0, 1]$ that captures the genetic transmissions and idiosyncratic talents that are relevant for non-farm sector. So the higher is θ_i the better suited an individual is for non-farm employment. A fundamental source of intergenerational linkage arises from the fact that the genetic endowments of a child (θ_i) are likely to be correlated with those of parents. However, this correlation is less than perfect and varies even across the offspring of the same biological parents. Thus innate ability parameter θ_i is not known with certainty and every individual has to form an estimate utilizing all the available information contained in an appropriately defined information set.

In addition to the innate ability, the ability of a person also depends on the acquired skills $\mu_i(e_i, d_i^p)$, which is a function of education level e_i and the parental occupation $d_i^p \in \{a, n\}$(to reflect the learning externalities during the first period).[9] The higher is the level of education e_i the higher is the probability of getting a better paid non-farm job. The acquired skill is also a function of parental occupation d_i^p because children can gain valuable skills and experience by observing their parents at work, and by informal apprenticeship in parents' work place, especially when the nature of occupation is such that the workplace is in close proximity to home.[10] Such acquired specific human capital

[9]We use the superscript p to denote a parental variable. Analogously, superscripts m and f to a variable denote mother and father respectively.

[10]As noted by Lentz and Laband (1983), this proximity of work place to home is an important factor behind the observed strong intergenerational following in occupations like agriculture.

tends to give rise to increasing returns to following in ones parent's footstep, and thus may result in occupational lock-in or 'occupational following'. The total costs of acquiring skills μ_i is denoted as $K_i(e_i, d_i^p)$. A critical aspect of learning from the parents is that the acquisition of the occupation specific skill is almost costless to the children, as it is accumulated as part of their upbringing.[11] Thus, $K_i(0, d_i^p) \cong 0$. In contrast, acquiring skills through formal education which is valuable for other occupations is likely to involve substantial monetary and non-monetary costs, hence $K_i(e_i, d_i^p) > 0$.

Apart from the innate and acquired ability, parents can positively influence an individual's *access* to non-farm jobs. Information about and access to non-farm jobs are usually not widely shared, and the non-farm employers rely heavily on the informal referral process for new recruitment. As a result one's ability to secure a non-farm job (wage employment) depends on the thickness and cohesion of the social network one belongs to. The children usually have ready access to the parental network, which in turn depends on their occupation. The parents, when successful in self-employment, often transfer significant financial and reputation capital to their children. The transfer of financial capital may be critical in relaxing the credit constraint faced by the children when embarking on non-farm self employment (like trading, small shops etc.).[12] The reputation capital is associated with (brand or family) name, goodwill, and loyalty (of customers and suppliers) created and cultivated by parent's successful business, and thus is specific to the family and business. The children can reap the rent from such reputation capital only when they follow in their parent's footstep.[13] Let a vector σ_i denote measures of parental transfer of network, financial and reputation capital. An individual's earnings in the second period will be determined by innate and acquired ability $(\theta_i, \mu_i(e_i, d_i^p))$ and parental transfers σ_i.

[11] Marshall (1920) observed long ago: "As years pass on, the child of the working man learns a great deal from what he *sees* and *hears* going on around him....." (P. 172) (italics added).

[12] There is now a mature empirical literature that demonstrates the strong link between access to credit and propensity to become an entrepreneur, both in the context of developed and developing countries (for evidence on U.S.A. see, for example, Holtz-Eakin, Joulfaian and Rosen, 1994, and Blanchflower and Oswald, 1998).

[13] This is especially true in developing countries where the tradability of brand name is limited or non-existent because of underdeveloped capital markets.

The preference of an individual i is represented by a concave utility function, $U_i(.)$, that reflects, among other things, the risk preference. The preference of an individual depends on a vector of characteristics (like gender, marital status, ethnicity, parental characteristics). Like ability, the preferences of a child are likely to be correlated with those of her parents. In addition, parents can also induce changes in children's preferences by acting as their role models (Durlauf, 2000). The intergenerational correlation in preferences implies, for example, that, on an average, the children of the parents more inclined to taking risk will themselves be risk takers, and thus are more likely to become non-farm entrepreneurs.

Let C_i^0 and C_i^1 be the consumption in first and second periods respectively. The expected lifetime utility of an individual i can be expressed as:

$$V_i = U_i(C_i^0) + \delta_i E(U_i(C_i^1))$$

where $\delta_i \in (0,1)$ is the discount factor. At the beginning of first period, individual i receives a transfer of Y_i^{0p} from her parents. The consumption in the first period can be defined as:[14]

$$C_i^0 = Y_i^{0p} - K_i(e_i, d_i^p) \tag{1}$$

Credit constraint on borrowing can be introduced in the model using the restriction that $Y_i^{0p} - K_i(e_i, d_i^p) - C_i^0 \geqq 0$.[15] Another important source of intergenerational correlations is obvious from the budget constraint in period 1 (equation 1); a larger transfer from the parents can relieve credit constraint faced in acquiring education in the first period and hence enables children to aim for better paid non-farm jobs in second period. This induces an intergenerational correlation in occupations if non-farm jobs generate better income for parents and transfer to the children is a normal good. The second period consumption can

[14]Note that for simplicity, we ignore the savings decision which can be introduced in the model easily.

[15]This implies no access to credit markets. Partial access to credit markets can be permitted in the model by assuming that $Y_i^{0p} + B^* - K_i(e_i, d_i^p) - C_i^0 \geqq 0$ where B^* is the maximum that can be borrowed from outside.

be defined as:

$$C_i^1 = Y_i^1 \tag{2}$$

where Y_i^1 is the income earned during t_1. The optimization problem faced by an individual i of generation 0 at the beginning of first period (t_0) is to choose an optimal level of education given the information set Ω_i^0 summarizing all relevant information including the estimated innate ability ($\tilde{\theta}_i^0$), costs of education $K_i(.)$, and parental occupation (d_i^p). At the beginning of second period (t_1), he/she takes the education level e_i^* as given, and solves the optimal occupation choice problem utilizing the relevant information set Ω_i^1. Note that due to information revelation in the first period,[16] the information set at the beginning of the occupational choice is richer, implying $\Omega_i^0 \subset \Omega_i^1$.

The optimal Schooling Decision:

Let the possible education level $e_i \in [0, \bar{S}]$, $\forall i$ where $\bar{S}$ is maximum number of years of schooling possible. Each level of education induces a conditional distribution of income Y_i that incorporates the optimal choice of occupation in the following period given the information set Ω_i^0. Let $F(Y_i^1 \mid e_i; \Omega_i^0)$ be the conditional distribution of second period income when individual i chooses education level e_i given the information set Ω_i^0. The associated probability density function is denoted as $P(Y_i^1 \mid e_i; \Omega_i^0)$. The optimal schooling choice is as follows:

$$e_i^* = \arg\max_{e_i} \left\{ U_i(Y_i^{0p} - K_i(e_i, d_i^p)) + \delta_i \int U_i(Y_i^1) P\left(Y_i^1 \mid e_i; \Omega_i^0\right) dY_i^1 \right\} \tag{3}$$

The Optimal Occupational Choice

At the beginning of second period, individual i takes the accumulation of human capital and the consequent estimate of ability $(\tilde{\theta}_i^1, \mu_i)$ as given, and optimally chooses the occu-

[16]There are many types of information revelation that might occur during the first period, like sucess/failure in formal education might reveal information about ones innate ability θ_i, parental network might expand or contract etc.

pation $d_i \in \{a, n)$. Now given the information set Ω_i^1, a choice of occupation induces a probability distribution of income. Let $F(Y_i^1 \mid a; \Omega_i^1)$ denote the conditional distribution of income when individual chooses agriculture and the information set is Ω_i^1.

We define the expected utility from choosing agriculture as:

$$V_i(a, \Omega_i^1) \equiv \int U_i(Y_i^1) P(Y_i^1 \mid a; \Omega_i^1) dY_i^1 \tag{4}$$

Analogously the expected utility from choosing non-farm sector is:

$$V_i(n, \Omega_i^1) \equiv \int U_i(Y_i^1) P(Y_i^1 \mid n; \Omega_i^1) dY_i^1 \tag{5}$$

The individual chooses non-farm employment iff the following holds:[17]

$$V_i(n, \Omega_i^1) - V_i(a, \Omega_i^1) \geq 0 \tag{6}$$

The probability that an arbitrary individual drawn from the population will decide to work in the non-farm sector is $\Pr(V_i(n, \Omega_i^1) - V_i(a, \Omega_i^1) \geq 0)$. At the heart of the occupation selection process is the formation of expectation about pay-offs from different options using the information set Ω_i^1. A critical element of the information set is the occupational choices of the parents as they reveal two types of relevant information: (i) information about ones own genetic endowment (or innate ability), (ii) information about the characteristics of a certain occupation. For example, if parents (either or both) are successful (unsuccessful) non-farm entrepreneurs, the estimate of children's ability to be successful in similar occupation will be revised upward (downward). Another important channel is that revelation of information might reduce the uncertainty about the parental occupation, and thus induce risk-averse children to prefer the parental occupation to other alternatives. Thus, the information revealed by parental choices (and their outcomes) can influence children's occupation decision through their effects on the conditional distribution function of

[17] Assuming that the tie is broken in favor of non-farm sector.

income Y_i^1 giving rise to role model effects (Manski 1993; Streufert, 2000).[18] For example, consider a child's participation decision in non-farm sector ($d_i = n$). The parental role model effects imply that the conditional distribution of income when parents are in non-farm $F(Y_i^1 \mid n;\, n^p, \Omega_i^1)$ is stochastically dominant over the conditional distribution of income with neither of the parents is in non-farm $F(Y_i^1 \mid n;\, a^p, \Omega_i^1)$.[19]

The model presented above identifies a number of different sources of intergenerational linkages, but it leaves unexplored the sources of any gender effects in intergenerational linkages in employment and occupational choice. Why would one expect the correlations to be stronger along gender lines (mother-daughter and father-son)? Since the transfer of financial and reputation capital, and network effects can be reasonably argued to be largely gender neutral, we need to look for the answers in the other sources of intergenerational linkages discussed above. First, the genetic transmissions might have a gender dimension. For example, the preference of a daughter (son) is likely to be more aligned with that of her (his) mother (father) compared to that of her (his) father (mother). Second, and probably the most important factor behind gender effects in intergenerational linkages in occupational choices, is the gender dimension in role model effects. The information revealed by the choices (and consequent outcomes) of an older member of a society will be more informative for the choices of a given younger member the closer he/she is to the younger person in an appropriately defined socioeconomic space. The individuals can be grouped together by partitioning the socioeconomic space according to different exogenous (like ethnicity, gender) or endogenous (like schooling) characteristics. The finer the partitioning the more informative is the information revealed by the choices of a member of a given group for the other members of that same group. It immediately follows that, given the membership in a family, gender creates a finer partitioning, and the mother becomes the natural role

[18] The definition of role model adopted so far in economic literature is not uniform. For example, while Durlauf (2000) defines role model as the influence of "characteristics of older members" on the "preferences of younger members", Manski (1993) and Streufert (2000) define it as observations on older members whose choices reveal information relevant for the choice of younger members.

[19] Note that given a concave utility function both first and second order stochastic dominance are sufficient.

model for the daughter, and the father for the son. This has also implications for learning by doing and observing as the daughter (son) 'sees' and 'hears' primarily what her (his) mother (father) does and says. The existence of gender effects for a daughter means that the conditional distribution of income from non-farm employment when mother is in non-farm $F(Y_i^1 \mid n;\, n^m, \Omega_i^1)$ stochastically dominates the conditional distribution with father in non-farm $F(Y_i^1 \mid n;\, n^f,\, \Omega_i^1)$.

III. The Empirical Specification

For the econometric estimation, we employ a standard probit model taking inequality (6) as the basis for our empirical specification. Specifically, we consider the binary response model

$$y_i = 1\left\{y_i^* \equiv V_i(n, \Omega_i^1) - V_i(a, \Omega_i^1) \geq 0\right\}, \tag{7}$$

For estimation we impose linearity and assume that the latent variable y_i^* is generated from a model of the form

$$y_i^* = \breve{X}_i\beta + \varepsilon_i \tag{8}$$

Where $\breve{X}_i \subseteq \Omega_i^1$ is a vector of explanatory variables and ε_i is the idiosyncratic random disturbance term. For convenience, we partition $\breve{X}_i$ into four subsets: (i) X_i , the elements of which are individual specific characteristics (like education e_i^*, age, gender and marital status) that influence the productivity and preference, (ii) X_i^p, a vector of parental characteristics (mainly parent's occupation (d_i^p)), (iii) X_i^h, a vector of household characteristics that influence preference (household size and composition), and household's asset ownership that includes any transfers of financial capital from the parents, and (iv) X_i^g, a vector representing network variables like ethnicity, and measures of non-farm opportunities available in a village. When we explicitly control for education and assets of an individual in the regression, the parental occupation variable (s) then captures any intergenerational correlation resulting from similarities of preferences/tastes, transfer of intangible human and other types of capital (learning- by- watching, reputational capital etc.) and role model effects due to induced preference changes and information revelation.

There are two salient econometric issues that need to be dealt with in order to identify the effect of parent's occupation on children's occupational choices. First, and the most important concern is that the intergenerational correlation in occupation may result, *spuriously*, from the fact that parents and children may face similar labor market opportunities that induce them to choose similar occupations as well. For instance, if both parents and children lived in an area with better non-farm opportunities, then intergenerational correlation in non-farm participation may be an artifact of not adequately controlling for opportunities in the regression. Since non-farm opportunities are often clustered around urban areas, one can use distance to nearest urban center (or a non-linear function of it) as a control for unobserved heterogeneity in non-farm opportunities. Other candidate variables include the observed level of employment diversification and average income in a village. However, even with a wide range of village level controls, there could still be unobserved heterogeneity across villages in terms of non-farm opportunities. Thus, instead of including some observed village level controls, we allowed village level fixed effects in the estimation to control for non-farm opportunities. Second, according to our theoretical model, investment in education in first period (t_0) is dependent on the expected occupation in the second period (t_1), and thus is endogenous in the non-farm participation decision [20] Moreover, current assets of the household are also likely to be endogenous to the occupations of the household members. To deal with the endogeneity problem, we utilize the Two-Stage Conditional Maximum Likelihood (TSCMLE) approach.

IV. The Data

The data for our analysis come from the Nepal Living Standard Survey (NLSS) 1995/96. The NLSS consists of a sample of 274 primary sampling unit (PSUs) selected with prob-

[20] Observe that even though the education level is predetermined when the occupational choice is made, it can not be treated as exogeneous.

ability proportionate to population size, covering 73 of the 75 districts in Nepal. In each of the PSUs, 12 households were also selected randomly (16 households in the Mountain regions) providing a total sample size of 3373 households. With an average household size of about 5.6, the survey collected detail information for 18855 individuals. The NLSS is unique in the sense that it contained an entire section of questionnaire on parental information, including level of education, sector of employment and place of birth. The survey contains detail information on employment by sectors and by occupations at individual levels. However, for those parents who do not live in the household, or who are deceased, only four types of employment status were recorded; whether they were employed as wage labor in agriculture or non-agriculture, or self employed either in agriculture or non-agriculture.

Of the total individual level sample of 18347 for whom parents can be identified from the data, nearly 71 percent reported participation in the labor force, but about 20 percent did not report any occupation.[21] For the rest of the sample (9417 observations), 10 percent are either child labor (less than 14 years of age, 9 percent) or too old (more than 70 years of age) and thus are dropped.[22] But some of the parents of these individuals did not either participate in the labor force or report their labor force participation, further reducing the size of the sample.[23] Moreover, some of the rural PSUs showed no employment diversification (4 PSUs) which are dropped to avoid perfect fits in the regression analysis. As we focus on the rural areas, households residing in the urban areas are dropped leaving us with a final sample of 5820 observations.

The NLSS 1995/96 also contains a wide range of variables on household structure, education, income and asset ownership which can be used as controls in the regression. They include household size, composition of household (share of female adults, share of

[21] A large fraction of those not reporting occupation are in fact child labor with age 14 years or less.

[22] Note that the empirical results reported in the following sections remain unchanged even if we use any other cut-off age (e.g. dropping those below 20 years of age and above 65 years of age and so on).

[23] For 8394 individuals left in the sample, both parents reported participation in the labor force in the case of 6875 individuals, and for rest of the observations, employment status of either father (346 observations) or mother (1173) are missing. Note that if we use dummies to capture the missing parental information, the sample size can be increased but the qualitative results remain unaffected.

children, share of young and share of old in total household size) as well as individual's marital status and gender. The human capital of an individual is measured by her/his level of education and age. The education variable codes different levels of education (e.g. 0 if illiterate, 1 if literate but no schooling, and so on). We also define a set of dummies (15 to be exact) depicting the ethnicity of the individual. Finally, we use two alternative representation of the most critical variable in our analysis: parental occupation . First, we define a dummy indicating if at least one of the parents was employed in non-agriculture(n^e). Second, to allow for differential impacts, we disaggregate and define three different dummy variables: (*i*) whether both parents are employed in non-agriculture (n^p), (*ii*) if only father was in non-agriculture (n^f) and (*iii*) if only mother is employed in non-agriculture (n^m). The summary statistics of all the explanatory variables are presented in appendix Table A.1. The non-farm participation rate in our data is about 24 percent.

V. Empirical Results

5.1 Preliminary Results

Table 1 presents some preliminary results. Starting from a simple bivariate regression of children's occupation on parental occupation, we take a stepwise approach in presenting the results, introducing increasingly complex regressions in subsequent steps. This helps to demonstrate the robustness (or non-robustness) of intergenerational linkages in non-farm participation. The result from the the bivariate regression in column 1a shows that the parents' non-farm participation has significant positive influence on children's probability of participation in the same sector. Switching the indicator variable for either parent in non-farm employment from zero to one raises the probability of children's employment in non-farm sector from 0.21 to 0.36. This suggests a marginal effect[24] of about 0.15 which is quite large compared to the sample average probability of participation in non-agriculture of 0.24. The regression results, presented in column 1b of Table 1, suggest that having both parents in non-agriculture (n^p) has the strongest effect. While both father only (n^f)

[24]The marginal effect is estimated by holding all other explanatory variables at their sample mean values while switching the relevant indicator variable from zero to one.

and mother only (n^m) in non-agriculture have statistically significant (p-value=0.00) and numerically large marginal effects, mother's effect appears to be much larger in magnitude.

Regressions 1a and 1b attribute variations in children's employment choice entirely to parent's employment choice. Thus, the estimated positive correlation may not necessarily represent any genuine intergenerational externality but may simply pick up effects of omitted factors such as non-farm opportunities that influence both parents' and children's employment decisions in the same direction. The regression results using village level fixed effects to control for heterogeneity in non-farm opportunities are summarized in column 2a and 2b.[25] The inclusion of village level dummies leads to a significant reduction in the magnitude of the estimated coefficients and marginal effects of parents' occupational status on children's choice of occupation. However, the correlation remains positive and highly statistically significant (p-value=0.0), and the marginal effects are numerically large compared with the sample probability. Even in the case of the smallest marginal effect [0.06], switching the indicator variable n^f (father only in non-agriculture) from zero to one raises children's probability of non-farm participation by 30 percent, from 0.20 to 0.26. Also, controlling for common non-farm opportunities does not alter the finding that mother's employment status has a larger impact on the children compared with that of father.

The next set of results reported in columns 3a and 3b of Table 1 include network variables to represent access to non-farm jobs in addition to parental employment status and village dummies as explanatory variables. There are evidence that the web of networks often run along ethnic group/caste (see, for example, Dreze, Lanjouw and Sharma, 1998). Thus, we include a set of dummies depicting the ethnicity of the individual. We also include dummies showing if there is any short/long-term migrant in the household, as migration frequently occurs on the basis of personal networks. The results show a further but less dramatic decline in the magnitude of parental effects on children's non-farm

[25] If we use travel time to nearest urban center, village level median per capita expenditure, and share of non-farm employment in total village level employment as proxies for non-farm opportunities, the regressions results imply larger and more significant positive impact of parental variables compared with those obtained from regressions with village level fixed effects.

participation probability. Despite this decline, the overall results from regressions (1 & 2) regarding intergenerational correlation in employment choices still remain valid; parents exert significant positive influence on children's employment choice. Regression 4a and 4b add household size and composition variables to the set of explanatory variables. Although household size and composition have significant effect on an individual' employment choice, inclusion of these variables does not change the effect of parent's employment choice on children's non-farm participation probability. The marginal effects of parental non-farm employment remains nearly unchanged in comparison to what were found in regression 3a and 3b. Overall, the intergenerational positive correlations survive when a large number of control variables are added to the regression. However, there remain valid concerns as to whether the effects of parental variables are mainly driven by the omitted human and financial capital variables which, as discussed in the theoretical model, are important links in the intergenerational transmissions of socioeconomic status. It seems interesting to try to isolate the effects of other 'intangible' factors that include the genetic transmissions through endowment and preference and the role model effects by controlling for the effects of human and financial capital. This is the focus of the next sub-section.

5.2: Controlling for Human and Physical Capital

In addition to the level of education e_i^*, we include age of an individual as a human capital variable representing the work experience. We define two alternative measures of assets: (i) agricultural land, and (ii) non-land assets including house, non-farm assets and financial assets. Transfer income (remittances) received by the household and travel time to nearest commercial bank are also included as additional controls for access to capital. We include individual's marital status and gender (when the sample consists of all children, male and female) to account for taste and/or gender related differences. In Table 2, columns 1a and 1b present the results ignoring the potential endogeneity of human capital and assets.[26] While education and ownership of agricultural land have

[26] If parental occupation dummies were excluded from these regressions, 1a and 1b (Table 2) would

significant effects on children's probability of non-farm participation, the importance of parental variables remains nearly undiminished. For instance, having either parent in the non-farm sector raises children's probability of non-farm participation by 0.07 which is slightly lower than the marginal effect [0.08] reported in column 4a of Table 1. Compared with the results in 4a and 4b in Table 1, a marked decline in the statistical precision of estimates of coefficients of parental variables can be noticed, yet the estimates are still significant at 5 percent level or less. The inclusion of human capital and assets variables as additional control does not affect the strength of parental influence on children's non-farm participation probability appreciably. Thus parent's occupation choice variables did not act as a proxy for individual's human capital or access to financial capital.

The theoretical model and subsequent discussion on the empirical specification point out clearly that the level of education observed in the second period (t_1) can not be taken as exogenous in the non-farm participation regression due to simultaneity between optimal education and occupation choices. Likewise, current levels of land and non-land assets of a household are determined by the income and hence occupations of its members and can not be treated as exogenous. As mentioned before, to deal with the potential endogeneity problems, we utilize the Two Stage Conditional Maximum Likelihood Estimation (TSCMLE). At the first stage of the TSCMLE, the suspected endogenous variables are separately regressed on all exogenous variables in the regression model and a set of instruments. Given that we have three suspected endogenous variables (education, agricultural land, and other assets), we also identify three sets of instruments. They include a number of dummies depicting different levels of mother's education, another set of dummies for

correspond directly to the standard specification used in available literature (Ferrira and Lanjouw, 2001; Lanjouw, 2001, Lanjouw and Shariff, 2000, Kurosaki, 2001). Consistent with findings of these studies, we find that an individual's probability of non-farm participation varies significantly and positively with the level of education (marginal effect equal to 0.04 and p-value=0.0). Apart from the ethnicity dummies, having a migrant in the household also significantly improves one's probability of participation in the non-farm sector suggesting strong network effects in securing a non-farm job. On the other hand, participation in non-farm sector varies negatively with the ownership of agricultural land implying some occupational following in agriculture. Most of the other variables have expected signs and impacts on non-farm participation.

different levels of father's education, inherited land and its squared term. The positive impact of parent's, particularly mother's education on children's human capital is well documented in literature (Behrman, 1997) and can be taken as an identifying instrument in children's education equation. To check if parent's education have any independent impact on children's occupation choice apart from its effect through children's human capital, we run regressions including parent's education as additional explanatory variables. The regression results show that levels of parents education have no perceptible impact on children's occupation choice, as the estimated coefficients are statistically insignificant, often with wrong signs (negative) and rather small in magnitude.[27] Although NLSS 1995/96 provides us with detailed information on the household's *current* assets, it does not contain information needed to construct a comprehensive measure of parental asset transfers; only information on inherited agricultural land is available.[28] While in a predominantly agrarian country such as Nepal, much of the transfers from parents to children usually take the form of either agricultural land or houses, inherited land alone may not be adequate to capture the effects of all types of asset transfers. In the absence of a complete measure of asset transfers from parents to children, we use the *current* 'other assets' as a proxy for parental asset transfers other than agricultural land and instrument it with the level of parent's, particularly father's, education. For agricultural land, inherited land and its squared terms are used as the instruments. The first stage regressions explain a considerable amount of the variations in education, agricultural land and other assets while avoiding over-fitting. The adjusted R^2 is estimated to be 0.48 in education, 0.45 in agricultural land, and 0.31 in other assets regressions. The three sets of instruments are separately and jointly highly statistically significant in the relevant regressions.[29] The parental employment status variables are not statistically significant in education or other asset equations. Parent's non-farm participation appears to have negative influence on the amount of agricultural land owned

[27] These results are not presented here for the sake of brevity and can be obtained from the authors.

[28] The estimates based on the survey indicate that 85 percent of all agricultural land are inherited from the parents.

[29] The regression results are presented in the appendix Table A.2.

by the household.

For the second stage regression, the set of explanatory variables (including the suspected endogenous variables) is augmented with the residual (s) from the first stage and then probit regression of non-farm participation is run on this augmented set of explanatory variables. A nice feature of the TSCMLE is that the t-statistic on the residual is a valid test of the null hypothesis that the suspected endogenous variable is in fact exogenous. The estimated coefficients of the first stage residuals and their respective t-statistics are reported in the second panel of Table 2. In addition to the usual test based on t-statistics, we also carried out Wald tests of joint significance of the residuals from the first stage regression, results of which are reported in the last panel of Table 2. An important advantage of TSCMLE is that even when the null hypothesis of exogeneity is rejected, the estimated coefficients are consistent, although the standard errors and t-statistics have to be corrected for the fact that first stage residuals are added as regressors in the second stage. All t-statistics reported in Table 2 are corrected for two-stage nature of regression as well as for intra-cluster correlations due to clustered sampling. As before, we take a stepwise approach in presenting the results particularly to demonstrate the robustness of the regression results under alternative assumptions. In column 2a and 2b, education alone is assumed to be endogenous. The t-statistics (-1.01 and -1.13) in 2a and 2b suggest that the null hypothesis of exogeneity of education in non-farm participation regression can be rejected only at p-value of 0.26 or more. Moreover, the estimates of effects of parent's non-farm employment on children's probability of non-farm participation are nearly indistinguishable from those reported in 1a and 1b. Columns 3a and 3b report results from the second stage regression under the assumption that agricultural land alone is endogenous. Again there is weak evidence against the null hypothesis of exogeneity. We find similar evidence regarding the exogeneity of 'other assets'. Finally, columns 5a and 5b report the results when all three variables are treated as endogenous. The t-statistics on agricultural land and other assets are rather small (0.18 and 0.19 in 5a and 0.08 and 0.30 in 5b respectively). The absolute value of t-statistic is higher for education, yet the null hypothesis of exogeneity of education can be rejected only at 21 percent level of significance. The Wald test of joint significance of the residuals from education, land and other assets equations suggests that the null of exogeneity can be rejected again at 25 percent (29 percent in 5a) significance level only. The battery of exogeneity tests suggest that household's asset (land and other) ownership can be treated as exogenous. This is not entirely unexpected particularly if most of the assets owned by the household were inherited or transferred from the parents. Interestingly, the statistical evidence against exogeneity of education is also rather weak. None of the tests can reject the null hypothesis of exogeneity even at 20 percent level of significance suggesting that the problem of endogeneity is not serious in this data set.

The TSCMLE estimates in Table 2 suggest significant positive intergenerational linkages in non-farm participation in almost all cases. The coefficient of either parent employed in non-farm sector is positive and statistically significant in all different formulations (2a to 5a) with a p-value of less than 1 percent. The marginal effect, estimated by switching the indicator variable for 'either parent in non-farm' (denoted as n^e) from zero to one while holding everything else at their sample mean, ranges between 0.07 to 0.06. Even if we focus on the most conservative estimates (5a) where education and assets (land and non-land) are taken as endogenous, the marginal effect is quite substantial, as having a parent in the non-farm sector raises children's probability of non-farm participation by 37 percent from 015 to 0.21. Only education [0.12], being male[0.23] and having a migrant in the household [0.07] have marginal effects greater than that of having a parent in the non-farm sector. Turning to the formulation where parent's non-farm employment is decomposed into three different variables (1b to 5b), having both parents in the non-farm sector has the strongest impact according to each of the regressions. The estimated coefficient is positive and highly statistically significant (smallest t-statistic is 3.25). The estimates of marginal effects of having both parents in the non-farm sector range between 0.13-0.15. Again focusing on the most conservative estimates (5b), switching indicator variable n^p representing 'both parents in non-farm' from zero to one raises children's probability of non-farm participation by nearly 90 percent, from 0.15 to 0.28. This effect is significant not only statistically but also relative to the sample probability of participation in the non-farm sector of 0.24. In magnitude, the marginal effect of having both parent in the non-farm sector is only second to that of being male [0.23 in 5b]. Having only mother employed in the non-farm sector has also large positive impact on children's probability of participation in the non-farm sector. The most conservative estimate of the coefficient of 'only mother in the non-farm sector' (5b) has a marginal effect of 0.12 with a p-value of 0.08 which is comparable, in magnitude, to the marginal effect of having both parents in the non-farm sector (0.13). Compared with 'mother only in non-farm' (n^m), 'father only in non-farm' (n^f) has a much smaller impact on the children's probability of non-farm participation. The estimates of marginal effects range between 0.03 to 0.04. The estimated coefficient, though positive in all regressions, lacks statistical precision in 5b (t-statistics=1.21). The results in Table 2 indicate that parent's employment in the non-farm sector have significantly positive impact on children's probability of non-farm participation even after education and assets are treated endogenous. The effects are slightly larger and statistically more significant if education and assets are taken as exogenous as suggested by the exogeneity tests.

While the above results clearly show that the positive intergenerational linkages in non-farm participation is very robust with respect to the set of control variables, we run additional robustness checks that focus on alternative function forms and estimation techniques. The main results reported in the previous section hold true in all different specifications and across all different estimation techniques.[30] Also, note that the sample in the preceding analysis includes those who reported more than one sector of employment. Dropping these observations from both children and parents samples leaves the qualitative results on the strength and significance of parents non-farm employment as well as the differential effects of mother and father's non-farm participation virtually unchanged.[31]

[30] We do not report detail results for the sake of brevity though it can be obtained from the authors. To give some flavor of it, the results remain virtually unaffected if we introduce dummies for children's education level instead of log of levels, or if we use logarithms for some of the continuous variables (e.g. inherited land, unearned income) or some squared terms.

[31] In this smaller sample (size=5364), the marginal effect of switching either parent's employment indicator (n^e) from zero to one is estimated to be 0.07 with sample probability of non-farm employment 0.08.

5.3: Gender and intergenerational linkages

The theoretical analysis pointed out that the intergenerational linkages in non-farm participation are likely to be stronger along gender lines (mother-daughter, father-son) due to genetic transmissions and, more importantly, role model effects through information revelation and induced preference changes. The available evidence on developed countries confirm the existence of such gender effects in occupational choices (Altonji and Dunn, 2000; Holtz-Eakin and Dunn, 2000). In this section, we investigate the extent to which the intergenerational correlations identified in previous sections differ along gender lines. To this end, we split the sample into sons and daughters. Although daughters accounted for nearly 54 percent of our total children sample of 5820 observations, when we drop the villages lacking occupational diversity from the sample, we are left with 1696 observations in the daughter's sample. The son's sample consists of 2619 observations. The rate of non-farm participation for sons and daughters are 41 percent and 16.5 percent respectively.

As before, the results from the probit regressions are reported in Table 3. For each subgroup, results are shown for two alternative specifications; columns 1a, 1b, 3a and 3b present the estimates under the assumption that education and assets are exogenous in the non-farm participation decision, whereas in 2a, 2b, 4a and 4b assume these variables to be endogenous. In the lower two panels, we also report the estimated coefficients and respective t-statistics for the residuals (from the first stage regressions) in the second stage regressions as well as the results of Wald tests of their joint significance.[32] Several interesting differences emerge from the comparison of the results for the daughters with that of the full children sample. First, in contrast with the children's sample (Table 2,

Hence the effect remains not only large but also statistically significant (p value =0.0) similar to what we found in the case where all available observations are included. Switching of employment indicator n^b (both parents), n^f (father only) or n^m (mother only) from zero to one, on the other hand, raises probability of children's non-farm participation by 0.15, 0.04 and 0.13 respectively. These are smaller in magnitude compared with the estimates in Table 2, nevertheless significant and large relative to sample average of probability of participation in non-farm sector (0.08).

[32]The results from the first stage regressions for both son's and daughter's samples are very similar to that reported for the full children sample and thus not discussed here in order to save space. These results can be obtained from the authors.

columns1a &1b, and 5a & 5b), in the daughter's sample, the estimates of parameters and their respective t-statistics obtained under the assumption that education and assets are endogenous (column 2a & 2b in Table 3) differ only insignificantly from those obtained under the assumption that education and assets are exogenous (columns 1a & 1b in Table 3). Indeed, the hypothesis about the exogeneity of education and assets can not be rejected at 85 (2a) and 95 (2b) percent significance levels in the daughter's sample. The exogeneity of education to non-farm employment decision seems puzzling at first sight. However, it is not entirely unexpected in a traditional society such as Nepal where investments in daughter's education depend largely on marriage market considerations instead of expected labor market returns.[33] Second, parents' occupational choice appears to have stronger effect on the daughters compared with that on all children (1a and 2a in Table 3 compared with 1a and 5a in Table 2). Regardless of whether education and assets are considered as endogenous or exogenous (1a and 2a), having either parent in the non-farm sector raises daughter's probability of participation in the non-farm sector by 100 percent, from 0.09 to 0.18. This is indeed significantly higher than the largest estimate reported in Table 2 (48 percent in 1a). Third, in contrast with the children's sample, 'both parents in non-farm' (n^p) does not have the largest effect; mothers have far more important influence on daughter's non-farm participation decision. Switching the indicator variable representing 'only mother in non-agriculture' (n^m) from zero to one while evaluating all other variables at their sample mean value increases the probability of daughter's participation by 0.34 (column 2b table 3), from 0.12 to 0.46. This suggests nearly a 300 percent increase in the probability of participation in the non-farm sector for daughters if mother is also employed in the same sector compared with the case where mother is employed in agriculture. The impact of 'both parents in non-farm' (n^p), although smaller, is still substantial as it raises the daughter's probability of non-farm employment by 200 percent (column 2b). Both of these effects (n^m and n^p) are statistically significant at less than 1 percent. Compared with 'both parents in non-farm' and 'mother in non-farm', 'only father in non-agriculture' (n^f)

[33]For similar evidence from India, see Behrman, Foster, Rosenzweig and Vishishtha (1999).

appears to have no significant impact on daughter's non-farm participation rates. The estimated coefficients of n^f, though positive, are small (marginal effect=0.01) and statistically insignificant.

The results from son's sample stand in sharp contrast to those from both children's and daughter's samples. When education and assets are taken as exogenous, there is significant positive effect of having either parent in the non-farm sector; the estimated coefficient is significant at p-value=0.03 with a marginal effect of 0.07 (column 3a, Table 3). This marginal effect is comparable to that for the children's sample but much lower than that for daughters. When parent's employment choice is decomposed into three different indicator variables, we find positive intergenerational effects for 'both parents in non-farm' (n^p) and 'only father in non-farm'(n^f) variables. A Wald test on these two variables shows that they are jointly significant at 8 percent level. Though mother's impact on sons is positive, it is statistically insignificant with t-statistics equal to 0.08. [34] The moderate father-son link found in columns 4a and 4b in Table 3, however, turns out to be non-robust, once the endogeneity of education and assets to non-farm participation decision is recognized. Indeed, the tests presented in the lower two panels show that the exogeneity of education and assets can be rejected resoundingly at p-value equal to 0.02 or less. The t-statistics on the estimated coefficients of the residuals from the first stage regressions indicate that endogeneity of education is the primary source of the rejection of the exogeneity hypothesis. Thus, in contrast with daughters, for sons, expected occupation plays an important role in the decision to invest in education. Once endogeneity of education and assets is accounted for, the TSCMLE results show that the estimated coefficients have expected positive signs (4a and 4b), but their magnitudes decline drastically and they are not statistically significant, separately or jointly.[35] To summarize, when endogeneity problem is

[34] Thus the cross gender effects (mother-son, father-daughter) seem to be virtually non-existent in intergenerational links in non-farm participation.

[35] Endogeneity correction greatly enhances the impact of education on non-farm participation; the marginal effect of education on non-farm participation probability increases from 0.03 (in 3a & 3b ignoring endogeneity problem) to 0.18 (in 4a & 4b with proper correction for endogeneity).

properly accounted for, parent's non-farm participation does not appear to influence son's probability of participation in the non-farm sector significantly.[36]

Thus, consistent with the theoretical expectations, we find strong intergenerational linkages between mother and daughter, but the absence of any significant linkages for sons appears puzzling, especially in the light of contrary evidence in the extant literature on developed countries (e.g. Holtz-Eakin and Dunn, 2000; Altonji and Dunn, 2000; Sjogren, 2000). However, the studies mentioned above did not account for the endogeneity of education and part of the discrepancy can be attributed to this, as we also find relatively larger and more significant linkages for sons when exogeneity is maintained. Thus, at least part of the positive intergenerational correlations in occupational choice identified in earlier literature seems to have resulted from the simultaneity bias. Another possibility is that the non-farm sector, as broadly defined as in this paper, consists of a myriad of activities from unskilled to highly skilled occupations, and predominance of lower end and temporary jobs requiring little skill may have blurred the intergenerational linkages often reported for more skilled jobs in the context of developed countries. We explore the role of skill composition more thoroughly in the following section.

5.4: Skill and Intergenerational Linkages

We define an indicator variable depicting if an individual is employed in jobs requiring specialized skills or not.[37] Similar indicator variables are also defined depicting if parents were employed in skilled occupations. Since many of the parents did not report finer occupational details, the regressions were run on a much smaller sample of sons (735 observations) for whom we have complete information. The results are reported in Table 4, columns 1a-2b. The overall results represent a slight improvement over that reported in

[36] As an additional robustness check, we drop those individuals who reported multiple occupations (both in agriculture and non-agriculture), and run regressions similar to 4a and 4b (Table 3), we find results nearly indistinguishable from those reported in Table 3.

[37] An individual is assumed to be employed in a skilled job if she/he reported any of the following occupations; professional and technical workers, administrative and managerial workers, clerical workers and operators, skilled sales and services workers, skilled workers in agriculture, production workers requiring specialized skills (e.g. machine operator, metal processors etc.).

Table 3 (3a-4b). When endogeneity of education and assets are ignored, we find significant and positive impact of father's participation in skilled occupations on the probability of sons participation in similar jobs. Compared with the results for all non-farm jobs, the marginal effects are much larger in magnitude; having both parents or only father in skilled jobs improves son's probability by 0.19. Since the null hypothesis of exogeneity of education and assets can be rejected at 5 percent significance level, the appropriate TSCMLE estimates (2a and 2b) show that parental influence on sons' occupational choice is statistically weak. 'Both parents employed in skilled jobs' is significant at 11 percent and 'father only in skilled jobs' at 13 percent. Despite their statistical imprecision, the estimated coefficients bear correct signs and are large in magnitude. The implied marginal effects are also large; having a father in skilled job raises son's probability by 0.14, which represents nearly a 50 percent increase in probability over the case when father was employed in unskilled job. The results in Table 4 thus suggest that there is moderate positive correlations in occupations between sons and fathers in the case of skilled jobs.

Similar to sons' sample, we carried out the regression analysis for the choice of skilled versus unskilled jobs for daughters. The sample size in this case is even smaller comprising of 426 observations.[38] The results are reported in columns 3a-4b in Table 4. The qualitative results remain unchanged from those reported in Table 3. Parents, particularly mothers, exert great influence on daughters occupational choice. The marginal effects are smaller in magnitude compared with those reported in Table 3 (1a-2b). However, the sample

[38]Regressions also excluded village level dummies as only 17 villages have more than one women employed in skilled jobs (total observation: 49). Excluding these observations does not alter results in any significant way. Instead of the village level dummies, we include a number of variable in the regressions to capture the effect of common opportunities faced parents and daughters in the labor market. They include regional dummies, share of non-farm employment in total village level employment, log of median village level per capita expenditure. The comparison of results using these variables as explanatory variables with that of the case when village level dummies are included shows that the differences between these two sets of estimates are small in the daughter's sample. For instance, the implied marginal effects of either parents in non-agriculture is 0.08 compared with 0.09 in column 1a of Table 3. The implied marginal effects are slightly smaller (0.09) in the case of TSCMLE compared with 0.10 in 2a of Table 3. Results for other specifications (1b & 2b) are also similar. This suggests that the share of non-agriculture in total village level employment, village level median expenditure and region dummies are also good proxies for labor market opportunities.

probability of daughter's participation in skilled jobs is nearly half (0.08) of that in non-agriculture as a whole (0.16). When expressed in percentage terms, the impact of parent's participation in skilled jobs on that of daughters is spectacular. For instance, compared with the situation when none of the parents were employed on skilled jobs, either parent's participation in skilled jobs raises daughter's probability of having a skilled job by 4.6 times. Having only mother in skilled jobs increases daughter's probability by 13 times. Thus, consistent with our results for sons, the intergenerational positive correlations are much stronger in the case of skilled jobs for daughters.[39]

VI: Conclusions

Despite the recent surge in interests in the determinants of non-farm participation in developing countries, the issue of intergenerational linkages in non-farm participation has, to our knowledge, not been addressed in the literature. A burgeoning literature on socioeconomic mobility, on the other hand, emphasizes these intergenerational correlations. However, the empirical research in this literature has focused mostly on intergenerational income correlations in the context of developed countries.[40] In this paper, we present some first empirical evidence on intergenerational correlations in non-farm participation in a developing country, Nepal.

The empirical results show that there are strong intergenerational correlations, especially for daughters, in non-farm participation arising from 'intangible' factors like role model effects, learning externalities, genetic transmissions of preference and ability and transfer of reputation capital. The results of Two Stage Conditional Maximum Likelihood analysis that correct for potential enodgeneity of education and assets in the regressions,

[39] In addition to the choice of skilled vs. unskilled jobs, we also experimented with different sub-samples of daughters to check robustness of the linkages. For instance, when we consider the sample of only married women, the intergenerational correlations are found to be slightly stronger. Given the tradition of women leaving their natal family upon marriage and joining spouse's households, the sample of married women can correct for any unobserved household level factors common to both parents and children residing in the same household.

[40] Solon (1999) in his survey of studies on intergenerational mobility cited only one paper on income mobility in the context of developing countries; the paper cited being Lillard and Kilburn (1995) that uses data from Malaysia.

suggest dramatic impact of a mother's employment status on daughters' non-farm participation rate; probability of a daughter's non-farm participation increases by almost four times if mother is employed in the non-farm sector, even after controlling for a large number of relevant variables including education, assets, and non-farm opportunities. The linkage between mother and daughter is especially strong in case of skilled non-farm jobs, thus implying very restricted occupational mobility for women out of agriculture and low skilled non-farm activities. In contrast, the intergenerational links between father and sons are found to be nonexistent in case of sons, except for the case of skilled non-farm jobs. Thus the occupational mobility displays a strong gender bias against women. It will be interesting to see if the pattern of intergenerational linkages found in case of Nepal is valid for other developing countries, especially in South Asia and Africa where the initial non-farm participation rates are strongly skewed against women.

References

Altonji, J.G. and T. Dunn (2000), "An Intergenerational Model of Wages, Hours, and Earnings," *Journal of Human Resources*, Vol. 35, n.2, 221-58.

Blanchflower, D., and A. Oswald (1998), "What makes an Entrepreneur: Evidence on Inheritance and Capital Constraints," *Journal of Labor Economics*, vol.16, no.1, 26-60.

Behrman, J. R.(1997), "Women's Schooling and Child Education: A Survey," *mimeo*, University of Pennsylvania.

Behrman, J. R., A. Foster, M.R. Rosenzweig, and P. Vashishtha (1999): "Women's Schooling, Home Teaching, and Economic Growth," *Journal of Political Economy*, vol. 107, no.4, 682-714.

Blundell, R, and R. J. Smith (1986): "An Exogeneity test for a Simultaneous Tobit Model," *Econometrica*, vol. 54, 679-85.

Dreze, J., P. Lanjouw, and N. Sharma (1998), "Economic Development in Palanpur, 1957-93," in P. Lanjouw and N. Stern eds, *Economic Development in Palanpur over Five Decades*, New York: Oxford University Press.

Dunn, T. and D. Holtz-Eakin (2000), "Financial Capital, Human Capital, and the Transition to Self-Employment: Evidence from Intergenerational Links," *Journal of Labor Economics*, Vol. 18, no.2, 282-305.

Durlauf , S (2000), "The Membership Theory of Poverty: The Role of Group Affiliations in Determining Socioeconomic Outcomes," *mimeo*, University of Wisconsin.

Ferreira, F. H. G., and P. Lanjouw (2001), "Rural Nonfarm Activities and Poverty in the Brazilian Northeast," *World Development*, vol. 29, no. 3, 509-28.

Goldberger, A. (1989), "Economic and Mechanical Models of Intergenerational Transmission," *American Economic Review*, vol. 79, no.3, 504-513.

Hoddinott, J., and L. Haddad (1995), "Does Female Income Share Influence Household Expenditures? Evidence from Cote D'Ivoire," *Oxford Bulletin of Economics and Statistics*, vol. 57, 77-95.

Holtz-Eakin, D., D. Joulfaian, and H.S. Rosen (1994), "Entrepreneurial Decisions and Liquidity Constraints," *Rand Journal of Economics*, vol. 23, no. 2, 340-47.

Kanbur, R., and L. Haddad (1994), "Are Better off Households mor Equal or Less Equal?," *Oxford Economic Papers*, Vol. 46, 445-58.

Kurosaki, T. (2001), "Effects of Human Capital on Farm and Non-farm Productivity in Rural Pakistan," *mimeo*, Yale University and Hitotsubashi University.

Lam, D, and R. F. Schoeni (1993), "Effects of Family Background on Earnings and Returns to Schooling: Evidence from Brazil," *Journal of Political Economy*, 101, 710-740.

Lanjouw, P. (2001), "Non-farm Employment and Poverty in Rural El Salvador," *World Development*, vol. 29, no. 3, 529-47.

Lanjouw, P. and G. Feder (2001), "Rural Non-farm Activities and Rural Development," Rural Strategy Background Paper # 4, World Bank.

Lanjouw, P. and A. Shariff (2000), "Rural Non-farm Employment in India: Access, Incomes and Poverty Impact", *mimeo*, World Bank.

Lentz, B.S., and D. N. Laband (1983), "Like Father, Like Son: Toward an Economic Theory of Occupational Following," *Southern Economic Journal*, vol. 50, no.2, 474-93.

Lillard, L, and M. R. Kilburn (1995), "Intergenerational Earnings Links: Sons and Daughters," RAND, Labor and Population Program WP# 95-17.

Manski, C. F. (1993), "Dynamic Choice in Social Settings: Learning from the experience of others," *Journal of Econometrics*, 58, 121-136.

Marshall, A. (1920), *Principles of Economics*, 8th edition, London: MacMillan Press Ltd.

Quisumbing, A., and J. Maluccio (1999), "Intra- houseold Allocation and Gender Relations: New Empirircal Evidence," *mimeo*, IFPRI, Washington, D.C.

Rivers, D., and Q. H. Voung (1988), "Limited Information Estimators and Exogeneity Tests for Simultaneous Probit Models," *Journal of Econometrics*, vol. 39, no. 3, 347-66.

Sjogren, A. (2000), "Occupatoinal choice and Incentives: The Role of Family Background," IUI, The Research Institute of Industrial Economics, Sweden, Working Paper # 539.

Smith, T, C (1988), *Native Sources of Japanese Industrialization, 1750-1920*, University of California Press, Berkeley.

Solon, G (1999), "Intergenerational Mobility in the Labor Market," in O. Ashenfelter and D. Card eds *Handbook of Labor Economics*, Vol 3A, Elsevier Science, North-Holland, Amsterdam.

Streufert, P. (2000), "The Effect of Underclass Social Isolation on Schooling Choice," *Journal of Public Economic Theory*, vol. 2, no.4, 461-82.

Thomas, D. (1994), "Like Father, Like Son; Like Mother, Like Daughter: Parental Resources and Child Height," *Journal of Human Resources*, vol. 29, no. 4, 950-88.

Thomas, D. (1997), "Incomes, Expenditures, and Helath Outcomes: Evidence on Intrahousehold Resource Allocation," in L. Haddad, J. Hoddinott, and H. Alderman, eds., *Intrahousehold Resource Allocation in Developing Countries: Models, Methods, and Policy*, Baltimore, MD.: The Johns Hopkins University Press.

World Bank (2001), *Engendering Development*, Policy Research Report, New York: Oxford University Press for the World Bank.

Table 1: Intergenerational Correlation and Employment in the Rural Non-Farm Sector

	(1a)	(1b)	(2a)	(2b)	(3a)	(3b)	(4a)	(4b)
Either parent in non-agriculture (n^e)	0.44		0.32		0.26		0.28	
	(7.73)		(5.52)		(4.91)		(5.18)	
	[0.15]		[0.10]		[0.08]		[0.08]	
Both parents in non-agriculture (n^b)		0.81		0.6		0.49		0.49
		(8.36)		(5.15)		(4.40)		(4.58)
		[0.29]		[0.21]		[0.16]		[0.16]
Only father in non-agriculture (n^f)		0.29		0.21		0.16		0.19
		(4.63)		(3.29)		(2.78)		(3.09)
		[0.10]		[0.06]		[0.05]		[0.06]
Only mother in non-agriculture (n^m)		0.62		0.45		0.43		0.46
		(3.37)		(2.71)		(2.57)		(2.72)
		[0.22]		[0.15]		[0.14]		[0.15]
No of observation	5820	5820	5820	5820	5820	5820	5820	5820
Sample probability	0.24	0.24	0.24	0.24	0.24	0.24	0.24	0.24

Note.- Entries are probit coefficients. Standard errors are corrected for intra-cluster correlations due to clustered sampling
t-values are in parentheses and marginal effect of each variable (evaluated at sample means) are shown in brackets.
All regressions include an intercept term
Regression specifications differ in terms of explanatory variables: explanatory variables in 1: only parents employment status
Explanatory Var. in 2: parents employment status plus 210 village level dummies to control for labor market opportunities
Explanatory Var. in 3: parents employment status, 210 village dummies and network variables (ehtnicity & migrant in the household)
Explanatory Var. in 4: parents employment status, 210 village dummies, network variables and household size and composition

Table 2: Intergenerational Correlation, Human and Physical Capital, and Employment in the Rural Non-Farm Sector

	(1a)	(1b)	(2a)	(2b)	(3a)	(3b)	(4a)	(4b)	(5a)	(5b)
Either parent in non-agriculture (n^e)	0.27		0.26		0.25		0.26		0.22	
	(3.69)		(3.50)		(3.41)		(3.58)		(2.80)	
	[0.07]		[0.07]		[0.06]		[0.07]		[0.06]	
Both parents in non-agriculture (n^b)		0.51		0.50		0.49		0.50		0.47
		(3.68)		(3.59)		(3.56)		(3.65)		(3.25)
		[0.15]		[0.14]		[0.14]		[0.14]		[0.13]
Only father in non-agriculture (n^f)		0.15		0.14		0.14		0.15		0.10
		(2.15)		(1.91)		(1.91)		(2.04)		(1.21)
		[0.04]		[0.04]		[0.04]		[0.04]		[0.03]
Only mother in non-agriculture (n^m)		0.45		0.45		0.43		0.44		0.41
		(1.98)		(1.96)		(1.83)		(1.94)		(1.74)
		[0.13]		[0.13]		[0.12]		[0.13]		[0.12]
Estimates from second stage regression										
Residual (education)			-0.14	-0.15					-0.34	-0.39
			(-1.01)	(-1.13)					(-1.09)	(-1.25)
Residual (agricultural land)					0.04	0.03			0.03	0.01
					(1.35)	(1.36)			(0.18)	(0.08)
Residual (non-land assets)							0.06	0.06	0.09	0.14
							(0.99)	(0.97)	(0.19)	(0.30)
Test of Exogeneity										
X^2			1.02	1.28	1.81	1.85	0.97	0.94	1.25	1.37
p-value			0.31	0.26	0.18	0.18	0.33	0.33	0.29	0.25

Note.- Entries are probit coefficients for Reg 1(a&b), for the rest, entries are Two-stage conditional maximum likelihood coefficients. Standard errors are corrected for two stage regressions and for arbitrary heteroskedasticity due to clustered sampling. t-values are in parentheses and marginal effect of each variable (evaluated at sample means) are shown in brackets. All regressions include parental employment dummies, log of levels of education, age age squared, dummies for male and married, household size & composition, assets (land & other), distance to bank, un-earned income, dummy for migrant members, ethnicity dummies, 210 village level dummies, and an intercept term.

Assumed endogeneous variable (s) is education in Reg 2, land (Reg3), other asset (Reg4) and all three (Reg 5).

Table 3: Intergenerational Correlation, Gender and Employment in the Rural Non-Farm Sector

	Daughters				**Sons**			
	Probit		TSCMLE		Probit		TSCMLE	
	(1a)	(1b)	(2a)	(2b)	(3a)	(3b)	(4a)	(4b)
Either parent in non-agriculture (n^e)	0.45		0.49		0.19		0.09	
	(3.41)		(2.63)		(2.19)		(0.81)	
	[0.09]		[0.10]		[0.07]		[0.03]	
Both parents in non-agriculture (n^b)		0.93		0.90		0.27		0.19
		(3.71)		(2.88)		(1.61)		(1.04)
		[0.25]		[0.24]		[0.11]		[0.07]
Only father in non-agriculture (n^f)		0.06		0.04		0.16		0.05
		(0.36)		(0.20)		(1.57)		(0.40)
		[0.01]		[0.01]		[0.06]		[0.02]
Only mother in non-agriculture (n^m)		1.20		1.17		0.21		0.06
		(3.11)		(2.71)		(0.59)		(0.16)
		[0.37]		[0.34]		[0.08]		[0.02]
Estimates from second stage regression								
Residual (education)			0.26	0.22			-0.43	-0.44
			(0.57)	(0.47)			(-2.33)	(-2.43)
Residual (agricultural land)			-0.03	0.04			0.12	0.11
			(-0.13)	(0.19)			(1.31)	(1.26)
Residual (non-land assets)			0.02	-0.18			0.01	0.02
			(0.03)	(-0.27)			(0.01)	(0.09)
Test of Exogeneity								
X^2 (H_0: Education & Assets Exogenous)			0.33	0.12			3.55	3.68
p-value			0.80	0.95			0.02	0.01
Number of observations	1696	1696	1696	1696	2619	2619	2619	2619
Sample probability	0.16	0.16	0.16	0.16	0.41	0.41	0.41	0.41

Note.- Entries are probit coefficients for Reg 1 and 3 (a&b), for the rest, entries are Two-stage conditional maximum likelihood coefficients. Standard errors are corrected for two stage regressions and for arbitrary heteroskedasticity due to clustered sampling. t-values are in parentheses and marginal effect of each variable (evaluated at sample means) are shown in brackets. All regressions include parental employment dummies, log of levels of education, age age squared, dummy for married, household size & composition, assets (land & other), distance to bank, un-earned income, dummy for migrant members, 14 ethnicity dummies, a number of village level dummies, and an intercept term.

Table 4: Intergenerational Correlation in skilled jobs

	Sons				**Daughters**			
	Probit		TSCMLE		Probit		TSCMLE	
	(1a)	(1b)	(2a)	(2b)	(3a)	(3b)	(4a)	(4b)
Either parent in skilled jobs	0.57		0.54		0.78		0.89	
	(2.49)		(1.64)		(2.55)		(2.12)	
	[0.19]		[0.14]		[0.05]		[0.08]	
Both parents in skilled jobs		0.63		0.47		1.01		1.05
		(1.39)		(1.09)		(2.34)		(2.18)
		[0.22]		[0.12]		[0.10]		[0.12]
Only father in skilled job		0.56		0.54		0.22		0.29
		(2.28)		(1.55)		(0.52)		(0.62)
		[0.19]		[0.14]		[0.01]		[0.02]
Only mother in skilled job		0.60		0.64		1.6		1.67
		(0.97)		(0.67)		(2.64)		(2.03)
		[0.21]		[0.17]		[0.24]		[0.27]
Test of Exogeneity								
X^2 (H_0: Education & Assets Exogenous)			2.70	2.76			0.97	0.84
p-value			0.05	0.05			0.41	0.47
Number of observations	735	735	735	735	426	426	426	426
Sample probability	0.31	0.31	0.31	0.31	0.08	0.08	0.08	0.08

Note.- Entries are probit coefficients for Reg 1 and 3 (a&b), for the rest, entries are Two-stage conditional maximum likelihood coefficients. Standard errors are corrected for two stage regressions and for arbitrary heteroskedasticity due to clustered sampling. t-values are in parentheses and marginal effect of each variable (evaluated at sample means) are shown in brackets. All regressions include parental employment dummies, log of levels of education, age age squared, dummy for married, household size & composition, assets (land & other), distance to bank, un-earned income, dummy for migrant members, ethnicity dummies, a number of village level dummies (for sons' sample only), and an intercept term. Regressions for daughters include share of non-farm employment in total village level employment, log of median village level per capita expenditure, and 4 regional dummies.

Table A.1: Summary Statistics

	Mean	Standard Deviation
Participation rate in non-agriculture	0.240	0.011
Either parent in non-agriculture	0.180	0.010
Both parents in non-agriculture	0.040	0.005
Only father in non-agriculture	0.130	0.008
Only mother in non-agriculture	0.010	0.002
Inherited land (value in million Rs.)	0.210	0.040
Inherited land squared (value in million Rs.)	0.670	0.290
Father's Education		
Literate	0.120	0.008
Primary Education	0.070	0.005
Secondary education	0.040	0.004
Higher than Secondary Education	0.004	0.001
Mother's Education		
Literate	0.010	0.002
Primary Education	0.006	0.001
Higher than Primary Education	0.002	0.001
Children's Education		
Literate	0.050	0.004
Primary Education	0.126	0.006
Secondary education	0.150	0.008
Higher than Secondary Education	0.020	0.003
Age	34.22	0.200
Age squared	1365.3	15.38
Male	0.460	0.005
Married	0.786	0.007
log(household size)	0.178	0.014
Share of adult female	0.251	0.003
Share of children	0.155	0.003
Share of Young	0.341	0.005
Share of Old	0.025	0.002
Log(travel time to nearest market)	2.910	0.200
Un-earned income (million Rs)	0.005	0.001
Migrant in the household	0.360	0.016

Note: All figures have been weighted to population means

Table A.2: First Stage Regressions

Dependent variable		Education		Agricultural Land		Other Assets	
		β	t	β	t	β	t
Father's education	Literate	0.45	9.70	0.98	5.72	0.55	5.83
	Primary Education	0.50	8.72	0.79	3.67	0.43	3.44
	Secondary education	0.75	9.79	1.07	3.51	0.79	3.58
	Higher than Secondary Education	0.53	4.75	1.83	3.54	0.98	3.57
Mother's education	Literate	0.02	0.14	-1.54	-2.69	-0.16	-0.46
	Primary Education	0.40	2.59	2.07	3.87	1.00	3.07
	Higher than Primary Education	0.55	2.13	1.34	1.25	0.85	2.64
Either parent in non-agriculture		0.00	-0.11	-0.38	-2.09	-0.14	-1.09
Age		-0.04	-6.88	-0.01	-0.50	0.00	-0.22
Age squared		0.00	3.10	0.00	0.88	0.00	0.61
Male		0.79	26.21	-0.10	-1.81	-0.01	-0.13
Married		-0.08	-2.44	-0.06	-0.44	0.02	0.24
log(household size)		0.15	4.60	1.75	7.51	1.22	9.37
Share of adult female		0.45	3.67	0.44	0.61	0.32	0.65
Share of children		-0.13	-1.12	-2.85	-3.17	-2.22	-4.47
Share of Young		-0.01	-0.14	-1.80	-2.38	-1.11	-2.59
Share of Old		0.31	2.13	-0.36	-0.40	-0.24	-0.36
Inherited land (value in million Rs.)		0.13	1.99	3.39	9.76	1.28	6.57
Inherited land squared (value in million Rs.)		-0.01	-2.31	-0.20	-5.88	-0.08	-4.89
Log(travel time to commercial bank)		0.00	-0.76	0.01	0.63	-0.02	-1.29
Un-earned income (million Rs)		0.69	1.06	7.88	2.07	8.04	4.00
Migrant in the household		0.01	0.48	-0.17	-1.01	-0.08	-0.71
R^2		0.48		0.45			0.31
No. of observations		5750		5750			5750

Note: All regressions include 14 ethnicity dummies and 210 village dummies. Standard errors are corrected for intra-cluster correlations