Institute for International Economic Policy Working Papers Series
Elliott School of International Affairs
George Washington University

The Extent of the Market and Stages of Agricultural Specialization

IIEP-WP#11

M. Shahe Emran
George Washington University

Forhad Shilpi
DECRG, World Bank

Institute for International Economic Policy
1957 E St. NW, Suite 501
Voice: (202) 994-5320
Fax: (202) 994-5477
Email: **iiep@gwu.edu**
Web: **www.gwu.edu/~iiep**

The Extent of the Market and Stages of Agricultural Specialization

M. Shahe Emran[1]

Department of Economics and ESIA

George Washington University

Forhad Shilpi

The World Bank

ABSTRACT

This paper provides empirical evidence of an U-shaped relationship between the extent of the market (size of the relevant urban market) and the pattern of crop specialization in a village economy. We use the recent two-stage estimator developed by Lewbel (2007) and exploit heteroskedasticity for identification of the causal effects of market size. The results suggest that the portfolio of crops in a village economy becomes more diversified initially as the extent of the market increases. However, after the market size reaches a threshold, the production structure starts to specialize again. This evidence on the stages of agricultural diversification is consistent with the stages of diversification identified in the recent literature for the economy as a whole and also for the manufacturing sector.

JEL Codes: O12, R11, R12

Key Words: Structural Change, Agriculture, Crop Specialization, The Extent of the Market, Market versus Home Production, Commercialization

[1]Emails for correspondence: shahe.emran@gmail.com (M. Shahe Emran), fshilpi@worldbank.org (Forhad Shilpi). We would like to thank Bryan Boulier, Jean Imbs, Tony Yezer and Stephen Smith for helpful comments and/or discussions. The standard disclaimers apply.

1 Introduction

Understanding the process of structural change has been one of the central focus of development economics from Lewis (1954), Kuznets (1973), Chenery et. al. (1986) to Lucas (1988, 2004), among others. Most of the theoretical and empirical literature on structural change and long run evolution of an economy focuses on the transition from a predominantly agrarian and rural economy to an industrialized and urban one (Chenery et. al., 1986; Locay, 1990 ; Laitner, 2000 ; Lucas, 1988; Buera and Kaboski, 2006; Matsuyama,2005). For a large number of developing countries where agriculture still predominates the economic landscape, the issue of structural transformation within agriculture– from a traditional subsistence based agriculture to more specialized and market oriented one– is, however, equally important. This is because structural transformation of an economy into more diversified non-agricultural (non-farm and industrial) activities is frequently triggered by productivity growth and increasing commercialization and specialization in agriculture (Johnson, 2000; Gollin, Parente and Rogerson, 2002, 2006). Moreover, the structural change from subsistence agriculture to a specialized and market oriented one leads to higher income and poverty alleviation in the village economies. This paper presents an empirical analysis of structural change within agriculture with a focus on the causal role played by the extent of the market.

The idea that the extent of the market is a principal driving force behind specialization dates back at least to Adam Smith (Smith, 1776).[2] In this classic Smithian account, a larger market allows greater division of labor and specialization by ensuring adequate demand for specialized skills and products. Although an integral part of economic thought over last few centuries, it is surprising that there is almost no formal econometric analysis of the role of the extent of market in determining pattern of specialization using household level data. The recent theoretical literature has underscored the importance of a large market in the adoption of increasing returns technologies that facilitates greater specialization in intermediate inputs and leads to higher economic growth (Murphy, Shleifer, and Vishny, 1989; Rodriguez-Clare,

[2] "As it is the power of exchanging that gives occasion to the division of labor, so the extent of this division must always be limited by the extent of that power, or, in other words, by the extent of the market" (Smith, A, 1776, Book I, Chapter III).

1996; Ciccone and Matsuyama, 1996).[3] An implication of this literature is that there is a *monotonic* relationship between the extent of the market and the degree of specialization. The more recent literature on the 'stages of diversification', however, uncovers non-linearity in the process of specialization. According to this literature, the production structure initially becomes more diversified as per capita income grows; and only after a threshold level of income is reached, the production structure becomes more specialized (Imbs and Wacziarg, 2003, Kalemli-Ozcan et. al. 2003). This inverted U pattern in stages of diversification holds both in the aggregate economy and within the manufacturing sector. This literature, however, does not address the pattern of structural change within the agricultural sector. Also, most of the literature tries to uncover robust correlations in the data, and thus does not address the issue of causality. An important exception is Kalemli-Ozcan et. al. (2003) where instrumental variables are used to explore causality. This paper contributes to the literature both in terms of identifying an U-shaped pattern of specialization in village economies and also by providing evidence on the causal role played by the extent of the market (i.e., the size of the relevant urban market). In the absence of credible exclusion restrictions, we use the recent two-stage estimator developed by Lewbel (2007) where identification relies on heteroskedasticity. This paper utilizes parametric and nonparametric techniques to uncover the nonlinearity in the effects of the extent of market on structural change in agriculture as measured by crop specialization and degree of commercialization.

We use data from the Nepal Living Standard Survey (NLSS) of 1995/96 to uncover the role played by the extent of the market in agricultural specialization and commercialization at the village level. Crop agriculture in Nepal, like many other developing economies, is characterized by low degree of commercialization and specialization on average. There are, however, striking differences among villages in terms of the level of agricultural development covering the entire range from completely specialized production of non-staple crops to nearly complete subsistence agriculture. The stark geographical differences in Nepal also resulted

[3] The positive influence of market size on specialization implies a positive correlation between initial income and subsequent growth of a country. Ades and Glaeser (1999), using cross country growth regressions, finds strong positive correlation between initial income and subsequent economic growth particularly for relatively closed economies.

in large variation in sizes of the urban centers (from population of 10,000 in smaller towns to 421,000 in the capital city Kathmandu in 1991). These large variations in the level of agricultural development along with that in access to and size of urban markets enable us to empirically characterize the relationship between agricultural specialization and the extent of the market. The results based on the Nepalese data are, however, of more general interest as they are likely to be relevant to many other developing countries which are characterized by relative isolation of rural areas due to poor infrastructure as well as low level of agricultural development.[4]

We analyze two dimensions of structural change in agriculture: the pattern of product diversification (crop specialization), and the degree of market production as opposed to home production (i.e., commercialization of agriculture). The pattern of crop specialization is measured by Herfindahl index of concentration of cropland use, and by the share of land devoted to non-cereal crops. Sales of non-rice crops and all crops as percentage of production are taken as measures of commercialization. The results from the empirical analysis based on Lewbel (2007) estimator show that the size of the relevant urban market exerts significant positive effects on both the pattern of crop specialization and commercialization of agriculture in a village economy. The evidence from non-linear parametric specification and nonparametric regressions show that the relationship between crop specialization (measured by Herfindahl index of cropland use) and the market size is U-shaped.[5] In contrast, the evidence of non-linearity in the effects of market size on measures of commercialization is not strong. The relationship between sales of non-cereal (non-staple) crops and market size is monotonically increasing with no significant evidence of nonlinearity. There is weak evidence of a concave relationship between the sales of all crops as percentage of production and the size of the market.

The U-shaped relation between the Herfindahl index of crop specialization and the extent of the market implies that when the farmers have access only to small urban markets, the

[4] These would include regions like Northern India and Pakistan, and many other developing countries in Africa and Asia.

[5] To formally test the quadratic effects of market size in a parametric specification we follow Lind and Mehlum (2007). We implement Sasabuchi t-test and use Fieller method to estimate the implied extremum point and its 95 percent confidence interval.

production structure in a village economy tends to be specialized in subsistence agriculture with most of the land devoted to a single subsistence crop (rice in case of Nepal) and only a limited degree of commercialization. As the extent (size) of the market increases, the portfolio of crops in a village economy becomes more diversified initially. However, after the market size reaches a threshold, the production structure starts to specialize again. This evidence on the stages of agricultural diversification is thus similar to the stages of diversification identified in the recent literature for the economy as a whole and also for the manufacturing sector (Imbs and Wacziarg, 2003). To the best of our knowledge, this paper provides the first evidence on the stages of agricultural specialization as it relates to the extent of the (urban) market with an emphasis on the causal effects[6].

The rest of the paper is organized as follows. Section 2 describes the simple conceptual framework underpinning the empirical analysis. Section 3 provides details about the data base used in the empirical analysis. Section 4, organized in a couple of subsections, presents the main empirical analysis and results. The paper is concluded in Section IV.

2 Conceptual Framework

To explore the relationship between market size and agricultural specialization, we start with the simple 'gravity model' which can be expressed as:

$$y_i = f(D_i, M_i, X_i) + u_{iz} \tag{1}$$

Where D_i is the distance to the urban market center relevant for village i, M_i is the size of that market, X_i is a vector of other relevant control variables and u_i is the error term. The vector y^z is a vector of dependent variables which in our case includes different measures of specialization and commercialization, to be defined precisely in the following paragraphs.

[6] Although there is a large literature on the role of urban market in agricultural specialization, the focus of that literature is on the *access* to urban markets (see, for example, Jacoby, 2000). For more general discussion on the role of urban markets in agricultural diversification and specialization, see World Development Report, 2008; Binswanger and Deininger, 1997.

The estimation of equation (1) requires identification of the most relevant market (s) for the households residing in village i. The standard practice in the current literature is to use the city/urban center closest to the village as the relevant market. Although convenient and useful as a first approximation, this can lead to misleading conclusions. To see the importance of including more than one city/town as the relevant market, consider the case shown in Figure 1; city A is located nearest to village i, city B is only a short distance farther but of a much larger size. It is quite likely that compared with city A, urban demand in city B exerts stronger influence on farm households' crop choice and market participation decisions. Moreover, the households in village i may trade different goods in different cities in which case distance to only the nearest city or market may not be the relevant measure of the extent of the market. The effective market for a village, given the transport infrastructure, consists of all the city/town market centers around a village where the villagers participate in. In order to allow for the possibility of trading by households of a village at multiple urban locations, we extend the traditional definition of gravity measure and define the effective market size for village i as:

$$M_i^e = \sum_{j=1}^{K} \omega_{ij} m_{ij} \tag{2}$$

where m_{ij} is the market size of urban center j and $1, 2, .., K$ are the urban centers where villagers trade, and ω_{ij} is the weight of urban center j.The weight for each market j is defined as:

$$\omega_{ij} = \frac{\delta(d_{ij})}{K}$$

where d_{ij} is the distance from village i to urban center j. We assume that $\frac{\partial \delta(.)}{\partial d_{ij}} < 0$ and $\frac{\partial^2 \delta(.)}{\partial d_{ij}^2} > 0$. These assumptions about the functional form of $\delta(.)$ imply that farther market centers are given smaller weights compared with the closer market centers. We assume that $\delta_{ij} = 1/d_{ij}$. Substituting $\delta(.)$ in equation (2), we have:

$$M_i^e = \frac{1}{K} \sum_{j=1}^{K} \frac{m_{ij}}{d_{ij}} = \frac{1}{K} \sum_{j=1}^{K} g_{ij} \tag{3}$$

where g_{ij} is the gravity measure widely utilized in studying the effect of market size on international trade flows. The effective market size for village i, as defined in equation (3), is thus an average of gravity measures of urban centers (1,2,...,K) where residents of village i trade. We also check robustness of our empirical results using alternative specification of the weight function. The empirical specification of M_i^e requires prior knowledge about K, the number of urban centers relevant for a village. As villagers may go to different markets for trading different products, empirical estimation is done for different values of K in order to establish robustness of our empirical results.

The estimating equation can now be specified simply as:

$$y_i^z = \beta_{z3} + \beta_{z4} M_i^e + X_i' \gamma_z + u_{zi} \tag{4}$$

The above specification allows for limited nonlinearity in the effects of market size by imposing diminishing marginal effects of distances. However linear specification in equation (4) may be inadequate to study the effect of the extent of market on agricultural specialization. The non-linearity in the relationship between agricultural specialization and market size may arise from a number of factors. First, in the case of crop agriculture, the yield and price risks are among the most important determinants of a farm household's land allocation across crops (Roumasset, 1976 , Newbery and Stiglitz, 1981, Islam and Thomas, 1996). When the relevant urban market is small, the price risk is likely to be higher due to imperfect matching in a thin market. This is especially important for non-staple (non-cereal) produce like fruits, vegetables and spices for which the extent of market is much more limited in a typical developing country because of the Engel's Law. This implies that a farmer facing a small urban market might not specialize in the production of high risk and potentially high return non-subsistence crops like fruits and vegetables although she might be willing to devote some land to such crops at the margin. When the extent of market reaches a threshold, the price risk is reduced significantly because of better matching in a thick market. A larger market also ensures adequate demand for large scale production and higher profit for non-staple crops. The higher profit may be due to more favorable prices for both inputs and output,

and adoption of increasing returns technology and agglomeration effects.[7] A large urban market allows scale economies in marketing of non-cereal crops like vegetables (fixed costs in transportation and storage by the wholesale traders) which in turn translates into better prices for farmers if entry into marketing of agricultural goods is not restricted. Another important point is that access to large urban areas means that the rural households have access to a rich set of markets including credit (banks) and insurance markets. As is well known, farmers are likely to behave in a risk averse fashion when markets are incomplete, especially when credit and insurance markets are missing (Newbery and Stiglitz, 1981). A more complete set of markets allows the farmers to take more production risk and devote more land to non-subsistence and cash crops. The interplay of subsistence and risk considerations, urban demand pattern and scale economies is likely to result in a non-linear relationship between agricultural specialization and the extent of the market.

However, the estimation of equation (4) can be implemented only when we have a measure of market size of each urban centers. Market size in a city is often represented by its population density (Ciccone and Alcala, 2003). This, however, does not take into account of the fact that the pattern of consumer demand in a city depends critically on the level of income of its population. According to the Engel curve relationship, poor people tend to spend a higher proportion of their income on staples relative to the non-poor. Moreover, with an increase in income, demand for non-staples rises more sharply than the demand for staples. Thus, it is the level of urban income which is likely to exert discriminating influence on demand for different agricultural crops and consequently on agricultural specialization. We thus use the total urban income as a measure of the extent of the market.

As to the vector of dependent variables y_i^z in equation (4), three measures of agricultural specialization and commercialization are analyzed in this paper. First, we define a Herfindahl index of concentration of crop land use as:[8]

[7] Note that a larger market is likely to induce farmers to adopt new technology in the production of the subsistence crop like rice also. This helps in specialization and commercialization as the farmers can allocate more land to non-subsistence crops without worrying about its own subsistence requirements.

[8] This index is similar to the specialization index in Imbs and Wacziarg (2003) and occupational specialization index used in Ades and Glaeser(1999).

$$S_j = \sum_{h=1}^{H} (\frac{l_{jh}}{lj})^2 \tag{5}$$

where l_{jh} is the amount of land devoted to crop h in village j, l_j is the total amount of land farmed in village j and H is total number of crops grown. Notice that if all land in a village is devoted to one crop, then specialization index S_j is equal to unity. The more the number of crops grown in a village, the lower is the value of S_j. A complete specialization ($S_j = 1$) could result from all land being devoted to non-cereal crops due to commercialization of agriculture. It could also be associated with subsistence nature of agriculture where virtually all land is allocated, for instance, to a cereal crop like rice in case of Nepal. This implies that the nature of specialization (whether due to commercialization or to subsistence nature of agricultural production) can be discerned only if the Herfindahl index of specialization is compared with an indicator of non-subsistence specialization. The share of land devoted to non-cereal crops at the village level is used as a measure of non-subsistence specialization. Our analysis focuses mainly on specialization and diversification during the dry season. This is due to the fact that the cropping pattern during the wet season is completely dominated by rice, a crop which benefits greatly from the monsoon conditions. Because of heavy monsoon and inundation associated with it, farmers have few options other than cultivating rice during the wet season. Sales of all crops and non-rice crops as a percentage of their respective production in the village are taken as measures of agricultural commercialization. It should be noted that the measures of specialization and commercialization are defined at the village level.

3 Data

The data for the empirical analysis of this paper come from the Nepal Living Standard Survey (NLSS) of 1995/96. The NLSS is a nationally representative survey which collected information from 3373 households spread among 274 primary sampling units (locally known as 'wards') covering 73 of Nepal's 75 districts. In addition to the comprehensive information on household and its members' characteristics, and household's expenditure levels, the

survey collected detailed information on agricultural activities including cropping pattern, crop production and sales. Of all households for which we have complete information (about 3344), 75% are engaged in crop production. Nearly all of the farmers engaged in crop production are located in rural areas (93%) and the rest located in and around rural towns. The farm households (about 2531 households) are distributed in 257 wards/villages. We drop wards/villages with less than three farm households from our sample. The empirical analysis of this paper is thus based on the sample of 237 villages/wards where at least three households are engaged in agricultural production. Both dependent and explanatory variables for the empirical analysis are defined at the village level using the information on farmers residing in respective ward/village.

Panel A of Table 1 provides the summary statistics for different measures of agricultural specialization and commercialization. The Herfindahl index (HI), defined in equation (5), is constructed from the cropping pattern observed at the village level. As already noted, the index takes a value of unity if all land in a village is devoted to a single crop and declines in value with an increase in the number of crops grown in a village. The median of the Herfindahl index is about 0.27 (mean=0.31). There are, however, considerable variations in the level of specialization across villages covering the entire spectrum from complete specialization (HI=1) to highly diversified cropping patterns. Our second measure of specialization (non-subsistence specialization) is the share of land devoted to production of relatively high value non-cereal crops (fruits, vegetables, oilseeds, spices, cash and other crops). According to Table 1, cereal crops dominate the cropping pattern. On average, less than a third of total land during dry season is used to produce non-cereal crops, the median is smaller about 0.25. However, there is considerable variation across villages, ranging from no land to all land allocated to non-cereal crops.[9]

Our final measures relate to commercialization of agriculture which captures the structural change away from home to market production emphasized in the recent literature (see, or example, Gollin et, al. 2002, 2006). Specifically, the shares of production of all crops and

[9] Both the Herfindahl Index and share of land devoted to non-cereal crops have a maximum of 1. There is, however, only one village with H=1. There are only 5 villages where all of the land is allocated to non-cereal crops. This should ameliorate any concerns for censoring in our data.

non-rice crops sold in the market are taken as measures of commercialization.[10] According to Table 1, on average, about 14.7% for rice output, and 21% of non-rice output are sold in the market. Overall, 19% of total crop output (median=15%) is sold for cash, rest being consumed at home or handed out as in kind payments. Despite the relatively low average degree of commercialization, in a number of villages, more than 50% of output are sold even in the case of rice, the main staple crop in Nepal.

Panel B of Table 1 reports the summary statistics for access to and size of the market located nearest to the surveyed villages. Despite its comprehensive data coverage, the NLSS 1995/96 lacks information on access to urban centers as well as market size in each urban centers. We complement the NLSS data by constructing measures of both access to urban centers and urban income. The Population Census of 1991 identifies 34 towns and cities in Nepal where a town is defined as a settlement of more than 10,000 inhabitants. We first compute the distance between each surveyed ward/village and each of these towns. Distances are normally taken along existing roads, except when roads do not exist, in which case we calculate the shortest arc distance to the nearest road, and then the distance to various cities along the road. Distances are then converted into travel time using available information about trucking and walking speeds along various types of roads in Nepal.[11] Off the road travel is assumed to take place by foot – a reasonable assumption for Nepal given the nature of the terrain. The median distance from surveyed wards to nearest town is about 2 hours and 21 minutes. The mean distance is however much higher; about 4 hours and 26 minutes, because a number of villages are located far off from nearest towns. Indeed, about 14% of villages are located at least 10 hours or more from the nearest town, the farthest one being about 29.5 hours away. Such wide variations in access to urban markets are outcomes of striking geographical disparities in Nepal, a country which extends from the relative plains

[10] The regression analysis focuses on all crops and non-rice crops sales because there is no rice sales in a large number of villages (about a third of the villages). This causes the problem of censoring in the case of regression for rice sales. The difference between the rice and non-rice sales can be highlighted by focusing on all vs. non-rice sales as well while avoiding the censoring problem.

[11] Travel speeds are calculated for various terrains and types of road. The travel times on different terrains and road types were obtained through discussion with various transportation experts and South Asia operations staff at the World Bank. Travel on highways and provincial roads is assumed to take place by truck; travel on secondary roads is assumed to be by cart.

of Terai to high Himalayas.

Data on city/town population are available from the Population Census of 1991. The population density of towns/cities in Nepal displays wide variations; the smallest town had barely 10 thousand people residing in it. The largest city, Kathmandu, on the other hand had a population of 421 thousand in 1991. We estimated per capita consumption expenditure for urban residents from the NLSS data. Total urban income is derived by multiplying per capita income by urban population. In the cases of smaller towns, per capita expenditure data are not always available. In those cases, average per capita consumption expenditure in the district where town is located is used to compute total urban income. The median income in the nearest town/city is about Rs. 302 million, and mean about Rs. 1.2 billion. Urban income displays high degree of variations. Using the formula in equation (3) and assuming $K = 34$ (total number of towns/cities in Nepal in 1991), we define the effective market size M_i^e.[12] The extent of market also displays considerable variation (SD= 0.137 and mean=0.126). These variations help us explore to the relationships between different measures of agricultural specialization and commercialization and the extent of the market.

4 Empirical Strategy

Establishing the causal effect of the extent of market on agricultural specialization and commercialization requires an empirical strategy to remedy the possible endogeneity of market size. The endogeneity problem arises because unobserved locational attributes may influence both the pattern of agricultural specialization, size of the urban agglomeration and placement of transport infrastructure. The standard approach to address such endogeneity problem is to look for instrumental variables. The candidates for possible instruments include geographic and topographic features which arguably can be treated as exogeneous. These instruments are, however, likely to violate the exclusion restrictions as they may also influence the range

[12] As we discuss later, our results are not sensitive to alternative definitions of market size corresponding to the different values of K chosen.

and intensity of agricultural production.[13] In the absence of credible exclusion restrictions, we rely on the recent approach to identification based on heteroskedasticity developed by Rigobon (2003), Vella and Klein (2003) and Lewbel (2007). In particular, we implement the two-stage approach developed in Lewbel (2007). We provide a brief intuitive discussion of the approach in what follows (for details, please see Lewbel (2007)).

In the first stage, the endogeneous variable (i.e.„ market size) is regressed on X, and the residuals $\hat{\xi}$ are retrieved. More specifically, we run the following first stage regressions:

$$M_i^e = \phi_o + X^{'}\gamma_k + \xi_k \tag{6}$$

Let $Z \subseteq X$ be a vector of exogeneous variables. The estimated residuals are then used to create instruments as follows:

$$\left(Z - \overline{Z}\right)^{'} \hat{\xi}_k \tag{7}$$

where $\overline{Z}$ is the mean of Z. As shown by Lewbel (2007), identification requires that the error terms in the first stage regression above (equation 6) are heteroskedastic [$cov(Z, \xi^2) \neq 0$]. Following Lewbel (2007), we use Breusch-Pagan test of heteroskedasticity to ensure that this identification condition holds in our data. In the second stage we use the above set of instruments to estimate equation (4).[14] As noted by Lewbel (2007), the Z vector can be the complete set of control variables or a subset of it (excluding the endogeneous explanatory variable, i.e., the extent of the market in our case). Following Lewbel (2007), we thus use the complete set of control variables in the regressions as the Z vector. It is important to emphasize that this does not require any more stringent assumptions than the standard instrumental variables approach when there are credible instruments satisfying both

[13] In an earlier version of the paper, we present evidence from an instrumental variables approach based on the geographic and topographic features (for details, please see Emran and Shilpi, 2008). Although the central conclusions reached in this paper are consistent with the conlcusions reached on the basis of geographic instruments earlier, we chose to omit these results precisely for the uncertianty regarding the exclusion restrictions.

[14] It should be noted that identification of casual effect of M_i^e is less problematic when the relationship between y_i^z and M_i^e is non-linear as as we find in the case of Herfindahl index of cropland use. Even in the absence of valid instruments, the identification can come from the functional form.

exogeneity and relevance criteria.[15]

The set of controls in the regressions include important determinants of agricultural specialization at the village level. The explanatory variables include average household size and composition (share of adult female members, share of children, share of old etc.) in the village.[16] In the sales regressions, the household size and composition variables control for possible subsistence considerations whereas in the case of land allocated to non-cereal crops and of Herfindahl index of concentration of cropland use, they control for labor supply and gender specialization. The average education level of adult male and adult female in the village are also introduced as possible controls for average human capital in the village. The dependent variables in the sales regressions are already normalized by production levels. In addition, we include a number of farm characteristics that can influence farm productivity and hence sales. The average characteristics of owned land at the village level are used as explanatory variables instead of that of operated land. These characteristics include size of owned landholding, a number of characteristics of owned land including land quality (share of khet land which is especially suitable for rice production, share of irrigated land, and share of land of different soil quality (such as awal, dwaim or sim). We also include a dummy for Hill region as well as log of land area of the district as regressors. In order to control for access to credit, we included average remittance income of the households residing in the village.

4.1 Empirical Results

We start by estimating the first stage regression specified equation (6) and perform the Breusch-Pagan test of heteroskedasticity of the residuals. The test results are reported in panel B of Table 2. The results show that the null hypothesis of homoskedasticity can be rejected resoundingly with a $\chi^2 = 335.47$ and P-value =0.00. Using the residual from this first stage regression, we generate the instruments utilizing the formula in equation (7).We

[15] It is standard in the literature on instrumental variables estimation to treat the set of control variaibles as exogeneous. In other words, we are relying on the same assumption regarding the set of control varaibles as is standard in an IV approach if we had had ideal instruments for endogeneous market size that satisfy the excluison restrictions beyond doubt.

[16] The omitted category is share of adult male.

then use two-step feasible efficient GMM estimator to estimate equation (4). The regression diagnostics are reported in Panel B of Table 2. The full regression results are reported in Appendix Table A.1, and we reported the coefficient of the extent of the market in Panel A of Table 2.

The regression diagnostics reported in Panel B (Table 2) clearly show that instruments generated using equation (7) satisfies all of the relevance and exogeneity criteria. The overidentification tests confirm the validity of the instruments: the largest value of Hansen's J-statistics is 19.06 with a P-value of 0.16 in the regression for Herfindahl Index of cropland use. As indicated by the Shea's Partial R^2 (= 0.74) and F-statistic for the test of joint significance of instruments in first stage regression (= 70), the instruments explain considerable variations in the extent of the market variable. The weak instrument test based on the Kleibergen-Paap rank statistic shows that the F-statistic (=70) is larger than the Stock-Yogo critical values for 5% maximal relative bias (=21.23) and 10% maximal IV size (=50.4) rejecting the null hypothesis of weak instruments comfortably.

The uppermost panel (A) of Table 2 reports the results regarding the effect of the extent of the market. The results show that the extent of the market has a statistically significant effect in all regressions except for that of Herfindahl index of cropland use. The coefficient of the extent of market is positive and statistically significant (P-value=0.01 or less) in the regression for percentage of land allocated to non-cereal crops and for sales of crops (non-rice and all). In terms of magnitude, market size has much larger impact on non-rice crop sales (coefficient=0.33) compared with all crops sales including rice (coefficient=0.15). The effect of market size in the case of concentration of crop land use measured by Herfindahl index is, however, small and statistically insignificant (t-statistic=0.44).[17]

The results from the parametric regressions suggest statistically significant and positive

[17] We check robustness of our results using an alternative definition of the weight function. We re-estimate equation (4) using a weight function where effect of distance declines at an exponential rate. The results are similar to those reported in Table 2. We also repeated the analysis defining the effective market size including only 5 nearest cities ($K = 5$) instead of all of Nepalese cities. The overall results are nearly unchanged. Additional robustness checks indicate that our results are not sensitive to the specifications of the weight function and to alternative values of K as long as K is not too small. We omit these results for the sake of brevity.

effect of the extent of market on commercialization as measured by sales of non-rice crops and all crops. The results on specialization (i.e., product diversification), on the other hand, are mixed at best. The market size has a statistically significant and numerically large effect on the share of land allocated to non-subsistence crops. The evidence, however, suggests that the extent of market does not exert any significant effect on the level of specialization as measured by the herfindahl index of cropland use. This seems to be counter-intuitive and contrary to the Smithian conjecture about the role of markets in fostering greater division of labor. As noted earlier, because of the critical influence of price risks, the relationship between the extent of market and agricultural specialization is likely to be non-linear. In the following section, we explore this possibility using nonparametric and parametric techniques.

4.2 Non-Linearity and Stages of Specialization

Following Imbs and Wacziarg (2003), we utilize nonparametric technique to explore the non-linearity in the relationship between the extent of market and different measures of agricultural specialization. We estimate the following regression using the locally weighted regressions smoother (LOWESS):

$$y_i^z = g(M_i^e)$$

In particular, we use the semi-parametric technique proposed by Robinson (1988) and Yatchew(1998).[18] The estimated relationships are shown in Figures 2a-2b and 3a-3b. According to Figure 2a, the relationship between market size and Herfindahl index of concentration of cropland use appears to be U-shaped. The share of land allocated to non-cereal crops increases with an increase in the effective market size, and there are indications of mild concavity up to a point before it becomes flat at very high values of the extent of markets

[18] The estimator involves stepwise procedure to estimate $g_z(.)$. At the first step, y_i^z and all explanatory variables in vector X'_{zi} are purged off the effect of M_i^e using standard non-parametric kernel regressions. Next, the residuals generated from the kernel regressions are then used to estimate the coefficient vector $\hat{\gamma}_z$. The effects of explanatory variables in vector X'_{zi} are then taken out of y_i^z using estimated coefficient vector $\hat{\gamma}_z$. Finally, a standard kernel regression is run with residual of y_i^z from the preceding step as the dependent variable and M_i^e as the explanatory variable. This final kernel regression provides the estimate of $g_z(.)$.

(see figure 2b). In the case of sales (Figures 3a-3b), both sales of all crops and of non-rice crops increases with an increase in the urban market size at the initial values of the extent of market with some indications of concavity in the relationships. The relationship has a steeper slope in the case of non-rice sales compared with all crop sales. A comparison of the two curves shows that for the entire range of urban income, the curve for non-rice sales lies above that for all crops including rice. This is consistent with the Engel curve prediction that non-rice crops face a higher income elasticity relative to rice, a subsistence crop. Figure 3b also points to flattening of the curves at high levels of urban market size.

Figures 2a-2b and 3a-3b indicate the presence of some non-linearity in the relationships between the extent of market and different measures of agricultural specialization (product diversification) and commercialization. However, a visual inspection of the figures make it clear that the number of observations at the higher values of the extent of market are sparse. It is thus not clear whether these nonlinearities are statistically significant. Since the potential nonlinearity identified from the semiparametric exercise is quadratic, we can use simple quadratic formulation of parametric regressions to check the statistical significance of the non-linear relationships. This also allows us to focus on the casual effect of the urban market size using the Lewbel (2007) approach. To correct for the possible endogeneity of the squared term in the quadratic formulation, we follow the same procedure elaborated in the preceding section to generate instruments that rely on heteroskedasticity for identification. While the instruments satisfy the overidentifying restrictions and relevance criteria, the two-step feasible efficient GMM estimation results indicate possible weak instrument problem. Instead of feasible efficient GMM estimator, we thus utilize the Continuously updated GMM estimator (GMM-CUE) which is found to perform better in the presence of weak instruments (Hahn et. al. (2004)).

The GMM-CUE estimation results along with regression diagnostics are presented in Table 3. The Bruesh-Pagan test of heteroskedasticity indicates clearly that the null hypothesis of homoskedasticity of the residuals of the first stage regression of square of the extent of market on the full set of explanatory variables can be rejected resoundingly. The instruments generated by interacting the first stage regression residuals with de-meaned explanatory vari-

ables appear to explain considerable variations in the extent of market and its squared term. The Hansen's overidentifying restrictions tests also confirm validity of instruments. The GMM-CUE estimation results do not display any weak instrument problems. Overall, the regressions comfortably pass all the relevant diagnostics tests.

The most interesting result in Table 3 is that the level and square of the extent of the market variable have now become highly statistically significant at 1 percent level in the regression for Herfindahl index of cropland use. The signs of these two terms imply a U-shaped relationship between the extent of the market and agricultural specialization. In the case of percentage of land devoted to non-cereal crops and sales of all crops, the squared term has a negative sign and is statistically significant at 5 percent level or less. The squared term in the regression for non-rice crop sales, however, lacks statistical significance. The estimated coefficients imply an extremum point in this case which lies outside the data range for the extent of market. Thus relationship between non-rice sales and the extent of the market can be suitably described as linear. The estimated coefficients imply a U-shaped relationship in the case of Herfindahl index, and a concave relationship in the case of land allocated to non-cereal crops and sales of all crops.

While the level and squared terms are statistically significant in all of the regressions with the exception of non-rice sales, a recent study by Lind and Mehlum (2007) shows that this widely used criterion for determining concavity or convexity in an empirical relationship is too weak. This is because the confidence interval for the implied extremum point may lie too close to either the lower or the upper bounds of the data range, and the curvature may not be enough to distinguish it from a monotonic relationship. A more appropriate test for non-linear relationship is a joint test of significance where the null hypotheses for a U-shaped relationship are that the slope is increasing or flat at a suitably chosen lower bound and is decreasing or flat at a upper bound of the market size variable. We implement the Sasabuchi t-test to detect the significance of the non-linear relationships. We also estimate the implied extremum point and its 95% confidence interval using the Fieller method.[19] The results are summarized in Table 4. In the case of Herfindahl index of cropland use, the

[19] Both of these are done by using the u-test program in Stata written by Lind and Mehlum (2007).

Sasabuchi test resoundingly reject the null hypothesis that the slope is negative or flat at the upper bound of the market size variable with a P-value=0.00 (t=5.31). The slope at the lower bound is significantly negative and the 95% confidence interval for the extremum point falls well within the data range. This provides us with robust evidence of a U-shaped causal relationship between the Herfindahl index of agricultural specialization and market size. In the case of share of cropland under non-cereal crops (a measure of non-subsistence specialization), the slope at the upper bound is negative but is not statistically significant from zero or positive (t=0.09). The extremum point (1.112) lies very close to the upper bound of market size (1.134) and the confidence interval extends well beyond the data range. These tests results thus contradict the more informal evidence in favor of a concave relationship discussed earlier. The evidence is more consistent with a linear and positive relationship between non-subsistence specialization and the extent of market. In case of all crop sales, although the null hypothesis of a positive or flat slope at the upper bound can not be rejected at 1 percent level, there is some evidence of non-linearity (P-value-0.013). The extremum point is within the data range but the 95% confidence interval is quite large. Thus the evidence in favor of a concave relationship between all crop sales and the extent of market is relatively weak.

The central result from the empirical analysis that there is a convex causal relationship between the Herfindahl index of cropland use and market size may appear puzzling to some readers. However, this can be understood in terms of the difference in the type of specialization at the two extremes of the U-shaped curve: the initial phase is characterized by subsistence specialization, while the specialization is in commercial crops at the higher end. This can be seen cleanly if the U-shaped curve is compared with that for the measure of non-subsistence specialization (i.e., share of land area planted with non-cereal crops). The higher values of Herfindahl index at lower levels of urban market size is due to the fact that the households in villages with access to smaller markets are basically self-sufficient and thus specialize in cereal crops. As market size faced by a village increases, it starts producing wider range of crops resulting in lower value of the Herfindahl index. However, as market size crosses a threshold, an increase in market size encourages more specialization with villages

specializing more and more in non-cereal crops like fruits and vegetables. When the village has access to larger markets, diversified consumer demand in the urban areas induce farmers to allocate more land to high value non-cereal crops. As market size increases further, it ensures more stable trading opportunity and a reduction in the price uncertainty for the non-subsistence crops thereby allowing farmers to completely specialize in non-subsistence crops.

5 Conclusions

The process of structural change that transforms a traditional subsistence based self-sufficient village economy into a more market oriented and specialized one is an important part of the long run evolution of an economy (Locay, 1990; Gollin et. al., 2002). This transformation process is of great importance to a majority of the developing countries where agriculture- still the mainstay of economic activity- is characterized by low levels of commercialization and specialization. The objective of this paper is to analyze structural transformation within agricultural sector with a focus on understanding the causal role played by the extent of market (i.e.„ the size of the relevant urban market).

Using village level data from Nepal, we analyze two dimensions of structural change in agriculture: the pattern of product diversification (crop specialization), and the degree of market participation (i.e., commercialization). The pattern of crop specialization is measured by Herfindahl index of concentration of cropland use, and by the share of land devoted to non-cereal (non-rice) crops (a measure of non-subsistence specialization). Sales of non-rice crops and all crops as percentage of production are taken as measures of commercialization. As opposed to the standard practice of defining market as the nearest urban center, we introduce a broader measure of the extent of market which incorporates the possibility that villagers may trade at multiple urban locations. In contrast to most of the current literature on structural change and stages of diversification, the empirical analysis also addresses the possible endogeneity of the extent of the market using heteroskedasticity for identification (Lewbel (2007)).

The linear parametric regression results indicate statistically significant and positive effect of the extent of market on the share of land devoted to non-cereal crops and sales of non-rice crops and all crops, but no effect on the Herfindahl index of crop specialization. The nonparametric and non-linear parametric regression analyses, however, uncover strong evidence in favor of a U-shaped relationship between the extent of market and Herfindahl index of specialization. There is no robust evidence of such non-linearity in the case of non-rice sales and share of land devoted to non-cereal crops. The results imply that when the farmers do not have access to large urban markets, crop production is dominated by subsistence considerations with villages specializing in the production of subsistence cereal crops. With an increase in the extent of market, crop production first becomes diversified with farmers producing both cereal and non-cereal crops. As the extent of market crosses a critical threshold, villages begin to specialize again–this time in the production of non-cereal crops. The evidence on the stages of agricultural diversification is thus similar to the stages identified earlier in the literature for the economy as a whole and for the manufacturing sector (Imbs and Wacziarg, 2003).

References

1. Ades, Alberto and Edward Glaeser, 1999, "Evidence on Growth, Increasing Returns and the Extent of the Market," *Quarterly Journal of Economics*, Vol.., p. 1025-1045.

2. Alcala, Francisco, and Antonio Ciccone, 2003, "Trade, Extent of the Market, and Economic Growth 1960-1996," *Working paper*.

3. Buera, F and J. kaboski, 2006, The Rise of the Service Economy, Working Paper, Northwestern University and Ohio State University.

4. Binswanger, H, and Klaus Deininger, 1997, "Explaining Agricultural and Agrarian Policies in Developing Countries," *Journal of Economic Literature*, Vol. 35(4), p.1958-2005.

5. Chenery, H, Robinson, M. Surquin, 1986, Industrialization and Growth: A Comparative Study, Oxford University Press.

6. Ciccone, A and K. Matsuyama,1996, Start-up Costs and Pecuniary Externalities as barriers to Economic Development, Journal of Development Economics, April, 1996.

7. Emran, M. Shahe and F. Shilpi (2008): "The Extent of the Market and Stages of Agricultural Specialization", Policy Research Working Paper No. 4534, World Bank.

8. Gollin, Douglas, Stephen Parente and Richard Rogerson, 2002, "The Role of Agriculture in Development," *American Economic Review Papers amd Proceedings*, Vol. 92(2), p.160-164.

9. Gollin, Douglas, Stephen Parente and Richard Rogerson, 2006, "The Food Problem and the Evolution of International Income Levels," forthcoming in *Journal of Monetary Ecnomics*.

10. Hahn, J., J. Hausman and G. Kuersteiner, 2004, "Estimation with weak instruments: Accuracy of higher-order bias and MSE approximations," *Econometrics Journal*, 7(1), p. 217-306.

11. Imbs, Jean and Romain Wacziarg, 2003, "Stages of Diversification," *American Economic Review*, Vol. 93(1), p.63-86.

12. Islam, N and S. Thomas, 1996, Foodgrain Price Stabilization in Developing Countries, *Food Policy Review* 3, IFPRI, Washington DC.

13. Jacoby, Hanan G., 2000, "Access to Markets and the Benefits of Rural Roads," *Economic Journal*, Vol.110, p.713-737.

14. Kalemli-Ozcan, sebnem, Brent E. Sorensen, and Oved Yosha, 2003, "Risk Sharing and Industrial Specialization: Regional and International Evidence," *American Economic Review*, Vol. 93(3), p.903-918.

15. Klein, R. and F. Vella, 2003, "Identification and Estimation of the Triangular Simultaneous Equations model in the absence of Exclusion Restriction through the presence of Heteroskedasticity," unpublished manuscript.

16. Kuznets, S, 1973, Modern Economic Growth: Findings and Reflections (Nobel Lecture), *American Economic Review*, 63, 247-258.

17. Lewbel, A, 2007, "Using Heteroskedasticity to identify and estimate mismeasured and endogenous rergessor models," *mimeo*, Boston College.

18. Laitner, J, 2000: "Structural Change and Economic Growth", *Review of Economic Studies*, July, 2000.

19. Lewis, A, 1954, "Economic Development with Unilimited Supplies of Labor", *Manchester School*, 22 (2), 139-191.

20. Lind, J. T., and H. Mehlum, 2007, "With or without U? The Appropriate test for a U shaped relationship," *mimeo*.

21. Locay, L, 1990, "Economic Development and the Division of Production Between Households and Markets", *Journal of Political Economy*, October, 1990.

22. Lucas, R, 1988, "The Mecahnics of Economic Development", *Journal of Monetary Economics*.

23. Lucas, R E. B., 2004, "Life Earnings and Rural-urban migration", *Journal of Policitical Economy*, February, 2004.

24. Matsuyama, K, 2005, Structural Change, Working Paper, NorthWestern University.

25. Murphy, Kevin, Andrew Shleifer, and Robert Vishny, 1989, "Industrialization and the Big Push," *Journal of Political Economy,* Vol. 97(5), p.1003-1026.

26. Newbery, D and J. Stiglitz, 1981, *The Theory of Commodity Price Stabilization*, Oxford University Press.

27. Rigobon, 2003, "Identification through Heteroskedasticity," *Review of Economics and Statistics*, 85.

28. Robinson, P. M., 1988, "Root-n-Consistent Semipraramtric Regression," *Econometrica*, Vol. 56(4), p.931-954.

29. Rodriguez-Clare, A, 1996, "The Divison of Labor and Economic Development", *Journal of Development Economics*, June 1996.

30. Roumasset, J, 1976, *Rice and Risk*, North Holland.

31. Smith, Adam, 1976, *An Inquiry into the Nature and Causes of the Wealth of Nations*, Chicago: The University of Chicago Press.

32. World Bank, 2008, *World Development Report 2008: Agriculture for Development.*

Table 1: Agricultural Specialization, Commercialization and Characteristics of Nearest Urban Market

	Median	Mean	SD	Min	Max	No. of observation
A. Crop Specialization						
Herfindahl index of concentration of cropland use	0.272	0.308	0.129	0.136	1	235
Share of total cultivated land devoted to non-cereal crops	0.251	0.317	0.251	0	1	235
B. Commercialization						
Percentage of production sold						
Rice	0.104	0.146	0.16	0	0.7	231
Non-Rice	0.149	0.206	0.185	0	0.882	237
All crops	0.151	0.19	0.155	0	0.724	237
C. Market Size and Access						
Distance to nearest town/city (hours)	2.35	4.43	4.98	0.047	29.55	237
Total income in nearest town/city (Rs. Million)	301.8	1209.91	2592.8	73	10748	237
Effective market size (000)	0.091	0.126	0.137	0.007	1.13	237

Table 2: Market Size and Agricultural Specialization and Commercialization

	Crop Specialization		Commercialization	
	Herfindahl index of crop land use	% of land used in Non-cereal crop	Sales as % of production Non-Rice Crops	All Crops
	Panel A			
Effective market size	0.0212	0.4304	0.3362	0.1536
	(0.44)	(3.37)***	(4.94)***	(2.63)***
R-squared	0.13	0.24	0.35	0.4
	Panel B			
Test of Heterocedascity (Breusch-Pagan/ Cook-Weisberg Test)				
χ^2[Mkt size Equation]	335.47	335.47	335.47	335.47
P-value	0	0	0	0
Over Identification Test				
Hansen's J Statistics	19.06	16.36	16.16	12.9
P-value	0.16	0.29	0.3	0.53
Validity of Instruments				
Shea's Partial R^2	0.74	0.74	0.74	0.74
F-statistics	69.8	69.8	70.06	70.06
P-value	0	0	0	0
Weak Identification Tests				
Kleibergen-Paap rk Wald F-Statistic	69.8	69.8	70.06	70.06
Stock-Yogo Weak Id Critical Values				
5% Maximal IV relative bias	21.23	21.23	21.23	21.23
10% Maximal IV size	50.39	50.39	50.39	50.39

Robust z statistics in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 3: Non-Linear Impact of Market Size on Agricultural Specialization and Commercialization

	Crop Specialization		Commercialization	
	Herfindahl index of crop land use	% of land used in Non-cereal crop	Sales as % of production Non-Rice Crops	All Crops
Effective market size	-1.6866	1.3831	0.4434	0.5442
	(5.52)***	(3.77)***	(2.37)**	(3.73)***
Effective market size^2	1.2662	-0.6216	-0.0724	-0.4168
	(5.50)***	(2.14)**	(0.44)	(3.01)***
Test of Joint Significance				
χ^2	30.50	57.74	42.00	17.52
P-value	0.00	0.00	0.00	0.00
Test of Heterocedascity (Breusch-Pagan/ Cook-Weisberg Test)				
χ^2[Mkt size^2 Equation]	1082.4	1082.4	1082.4	1082.4
P-value	0	0	0	0
Over Identification Test				
Hansen's J Statistics	20.62	31.90	33.12	22.64
P-value	0.84	0.28	0.23	0.75
Validity of Instruments				
Shea's Partial R^2				
Mkt Size Equation	0.55	0.55	0.55	0.55
Mkt Size^2 Equation	0.65	0.65	0.66	0.66
F-statistics				
Mkt Size Equation	111.2	111.2	109	109
Mkt Size^2 Equation	385.5	385.5	402.7	402.7
Weak Identification Tests				
Kleibergen-Paap rk Wald F-Statistic	5.87	5.87	5.96	5.96
Stock-Yogo Weak Id Critical Values				
10% Maximal LIML size	4.12	4.12	4.12	4.12

Robust z statistics in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 4: Tests of Non-linearity

	Herfindahl index of crop land use	% of land used in Non-cereal crop	Sales as % of production All Crops
Slope at lower bound of Mkt size	-1.668	1.374	0.538
	(5.52)***	(3.79)***	(3.73)***
Slope at upper bound of Mkt size	1.187	-0.0275	-0.402
	(5.31)***	(0.087)	(2.24)**
Sasabuchi test for U/inverse U shape	5.31	0.09	2.24
P-value	0.00	0.465	0.013
Extremum point	0.666	1.112	0.653
95% Confidence interval (Fieller method)	[0.634, 0.701]	[0.853, 7.105]	[0.548, 0.978]

Table A.1: Agricultural Specialization and Commercialization and the Extent of Market: Full regressions

	Herfindahl index of crop land use	% of land used in Non-cereal crop	Sales as % of production Non-Rice Crops	All Crops
Effective market size[1]	0.0212	0.4304	0.3362	0.1536
	(0.44)	(3.37)***	(4.94)***	(2.63)***
Household Size (log)	0.0133	-0.3350	0.0163	-0.1011
	(0.24)	(3.15)***	(0.24)	(1.72)*
Share of adult female	0.2112	-1.1310	0.2336	0.1894
	(1.29)	(3.70)***	(1.27)	(1.17)
Share of children	-0.1600	-0.0763	-0.0264	0.0249
	(1.00)	(0.27)	(0.15)	(0.17)
Share of Old	0.4095	0.0533	-0.4332	-0.1230
	(2.26)**	(0.14)	(1.95)*	(0.72)
Average female education (log)	0.0591	0.4119	-0.1228	-0.0602
	(0.95)	(3.49)***	(1.76)*	(0.98)
Average male education (log)	-0.0552	-0.0227	0.0329	0.0906
	(1.30)	(0.28)	(0.66)	(2.21)**
Average land area owned (log)	-0.0183	0.1242	0.0563	0.1382
	(0.51)	(1.63)	(1.27)	(3.32)***
Share of irrigated land (owned)	-0.0171	-0.0245	0.0237	0.0737
	(0.73)	(0.45)	(0.46)	(1.84)*
Share of Khet owned	-0.0628	-0.0058	-0.0150	-0.0523
	(1.68)*	(0.06)	(0.27)	(1.16)
Share of owned land of quality Awal	-0.0167	0.0255	0.1624	0.1984
	(0.38)	(0.26)	(2.51)**	(4.11)***
Share of owned land of quality Dwaim	-0.0285	0.0042	0.2043	0.1650
	(0.70)	(0.05)	(4.14)***	(4.20)***
Share of owned land of quality Sim	-0.0429	0.1344	-0.0313	-0.0364
	(0.93)	(1.44)	(0.83)	(1.02)
Remittance income (log)	0.0025	-0.0093	0.0037	0.0022
	(1.47)	(2.28)**	(1.62)	(1.24)
Area of District (log)	-0.0057	0.0265	-0.0571	-0.0441
	(0.35)	(0.75)	(2.61)***	(3.00)***
Ecological belt dummy (Hill=1)	0.0149	0.0688	-0.0558	-0.0557
	(0.73)	(1.71)*	(2.50)**	(3.08)***
Intercept	0.3226	0.8924	0.4040	0.4728
	(1.82)*	(2.34)**	(1.82)*	(2.95)***
R-squared	0.13	0.24	0.35	0.40
Observations	235	235	237	237

Note: Results from instrumental variable estimation.
1/:Definition: Effective market size= S_i (urban income$_i$/distance$_i^{\hat{c}}$) where i=1…34 cities and towns in Nepal
Robust t statistics in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Figure 1: Location of Markets

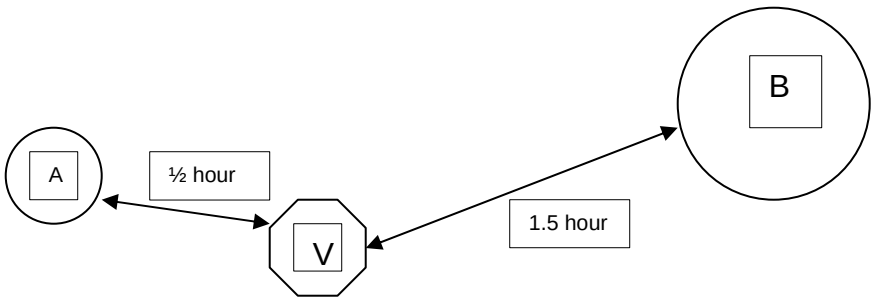

Figure 2a. Agricultural Specialization and the extent of the market: Herfindahl Index of Cropland use

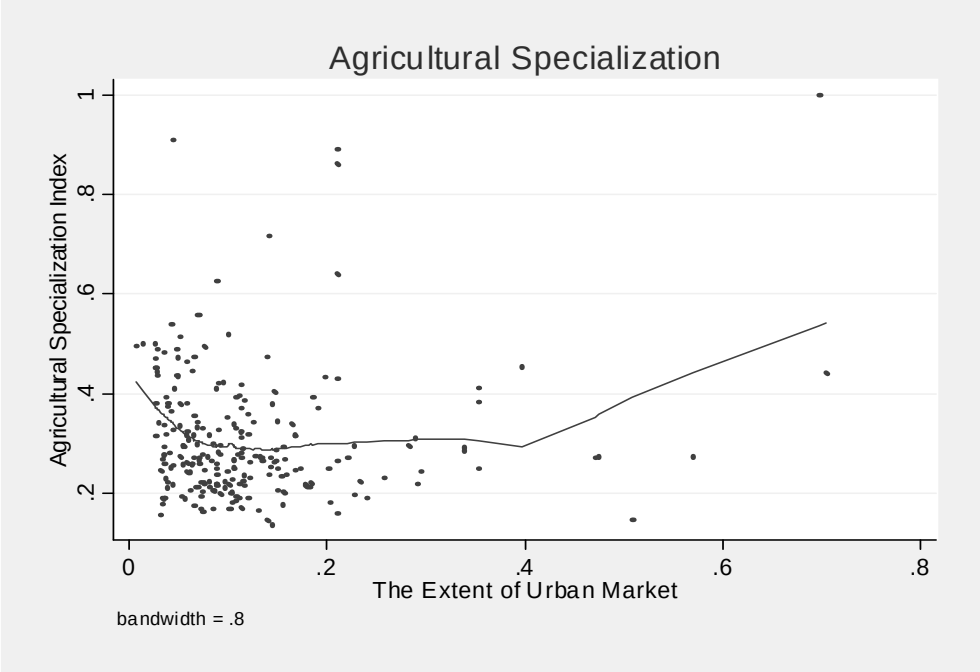

Figure 2b. Agricultural Specialization and the extent of the market: Land under Non-Cereal Crops

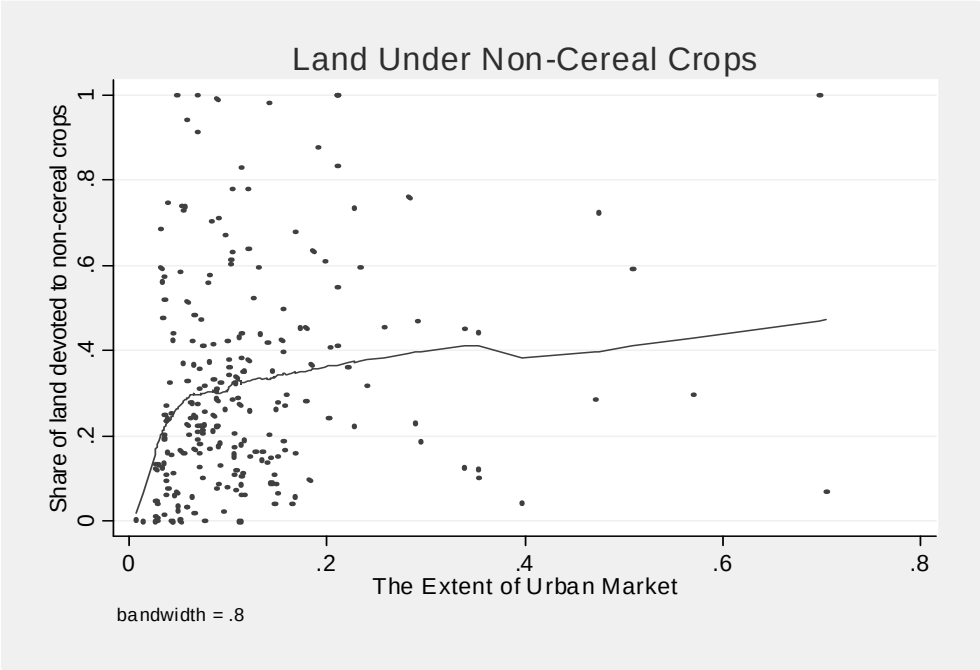

Figure 3a. Agricultural commercialization and the extent of the market : Non-Rice Sales

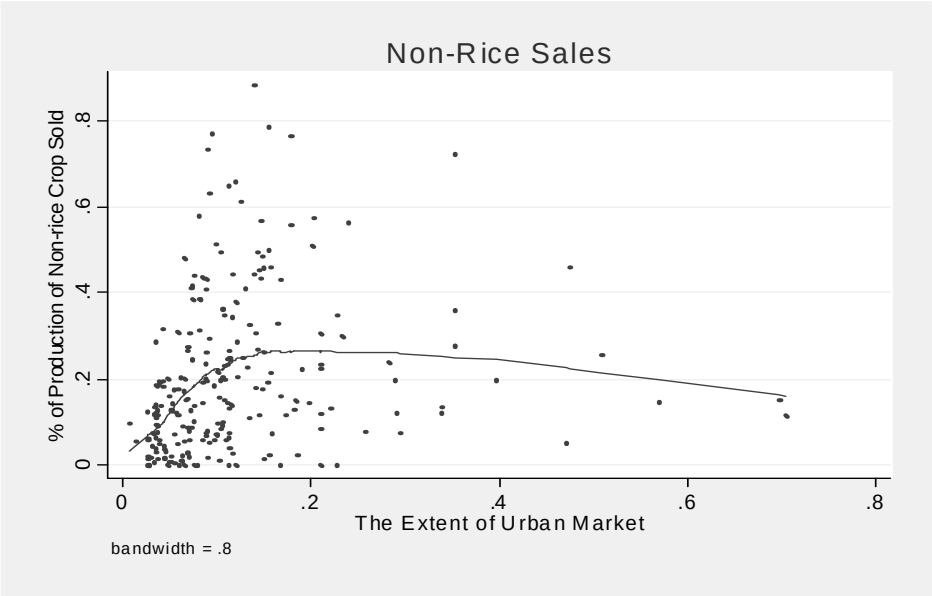

Figure 3b. Agricultural commercialization and the extent of the market: All Crop Sales