Master's Thesis in Economics
Stockholm School of Economics

Determinants of Household Poverty in Rural Gujarat, India

Aziz Sunderji*

Abstract

Although India's economy is growing at a blistering pace, most Indians still live in rural areas, many of them in poor households. Based on data from a household survey conducted for an Indian NGO between March and August 2004, this study seeks to model household poverty in a sample village in the Western state of Gujarat. The study endeavours to answer questions such as, i) what are the economic and demographic characteristics of farming households, and ii) what characteristics are associated with increased household income. More specifically, the econometric model presented here aims to determine the impact of variability in economic and demographic characteristics on the household's level of per capita income. Two models are estimated; the first treats fertility as exogenous to household income, the second uses an instrumental variable in a two stage least squares regression, treating fertility as endogenous. Greater planted areas of cash crops and greater productivity of cash crops are found to significantly increase income in at least one of the models. Both models find that the farmed area planted with sesame and the productivity of the sesame crop are associated with income. The demographic variables associated with income are the age of the household head, the number of non-productive extended household members, and the number of productive extended household members. The model that treats fertility as endogenous finds that an increase in the household's total number of children decreases household per capita income.

Advisor:	Professor Mats Lundahl
Presentation:	Tuesday, February 13th, 2007, 15:15-17:00
Venue:	room 328
Discussant:	Ayse Abbasoglu

Acknowledgements:

I would like to thank Profs. Mats Lundahl and Per-Olov Edlund at the Stockholm School of Economics for their academic guidance. I would also like to express my appreciation to Niraj Joshi and his team at the Aga Khan Rural Support Programme (India) for their support during my internship. Finally, I am deeply grateful to my family for their support during my studies.

*80288@student.hhs.se

Table of Contents

Determinants of Household Poverty in Rural Gujarat, India

Aziz Sunderji

The rural household in the context of Indian economic development

While India suffered under what was mockingly described by economists as the "Hindu rate of growth" for much of the twentieth century, consistently high growth has been achieved since India's current Prime Minister, Manmohan Singh, implemented sweeping economic reforms and abolished the "licence raj" as finance minister in 1991. In 2007, most news stories about India that reach western audiences describe the rise of the service sector and the growing middle class in cities such as Bangalore.

However, India is still a poor country by any measure. It is ranked 127th in the UNDP's Human Development report. The literacy rate is about 60%. Thirty five percent of the population lives on less than $1 per day, 80% on less than $2 per day[1]. Seventy percent of India's inhabitants still live in rural areas[2], most of them employed in an agricultural sector which contributes to only 25% of the country's GDP[3].

Thus, while India's growth is being driven by the service sector in urban areas, its poverty is concentrated among farming households in rural areas. To a large extent then, the success of poverty eradication in India will depend on the economic development of these households. According to the World Bank, "the large number of poor agricultural households and their income vulnerability are major concerns among policymakers."

This paper belongs to a specific thread of research which aims to analyze the poverty of farming households—to answer questions such as, i) what are the economic and demographic characteristics of farming households, and ii) what characteristics are associated with increased household income. More specifically, the econometric model presented here aims to determine the impact of variability in economic and demographic characteristics on the household's level of per capita poverty.

Most of the literature in this field can be loosely described as "poverty profiling". Across the variety of poverty profiling papers, the goal is typically uniform: to establish a set of economic and demographic variables that have a statistically significant association with household poverty. However, each paper is unique in its a) geographic region studied, b) econometric methodology, and c) policy recommendation. Although a summary of all the work in this field is impossible here, a cross-section of the different types of studies should provide an adequate context for the methodology pursued in this study.

Using data from a survey of households in the north-west province of South Africa, Serumaga-Zake & Naude (2002) use a probit-specification model to determine the level of association of several variables with the probability of a household being below the poverty line. They draw conclusions at various levels of significance: i) at the 10% level, a higher proportion of elderly members increase the probability of poverty, ii) at the 5% level, migrant workers (captured using a dummy variable) decrease poverty, while children under 15 years and extended families (rather than nuclear families, using a binary variable) increase poverty, and iii) at the 1% level, education decreases poverty and a higher number of female adults in the household increases poverty. The authors argue that, in a probit regression, different methods of calculating poverty lines result in different categorizations of households into poor and non-poor, and can have significant effects on the results of such a regression.

---

[1] UNDP human development report 2004, poverty incidence based on 1990-2002, as quoted in Sunderji, 2005
[2] CIA world factbook, http://www.cia.gov/cia/publications/factbook/rankorder/2119rank.html
[3] World Bank (2004)

Sen & Yoon (2005) run an MLS regression, using an oaxaca decomposition to determine characteristic and coefficient effects of several variables on the income gap between SC/ST individuals (scheduled caste/scheduled tribe, the Indian terms for lower-caste groups) and non SC/ST individuals (a characteristic effect exists because lower caste groups often have less education, for example). They restrict their sample to households where the head is between 20 and 70 years old. They find that older households are generally less poor than younger households, with decreasing marginal benefits of older age. This latter effect was measured by including the household head's age squared as a regressor. Interestingly, the authors find that larger households are poorer within a limited range, and that households above a certain size are less poor. Total land cultivated also reduces poverty. The authors note that their results are robust to usage of equivalence scales (see footnote 5).

Bezemer & Lurman (2004) also use a probit-type regression on data from Armenia, and find that several regressors have a significant impact on poverty: land, irrigated land, cattle, household size, average age, working age women, borrowing, savings, cooperation between households, presence of wage employment, non-farm activities, and remittances.

Several studies grapple with estimating coefficients for variables that are endogenous to household poverty. Of these endogenous variables, household size or fertility (number of children) are perhaps the most important; one of the most consistent conclusions from poverty profiling is that larger families are poorer on a per capita basis. Although increased household size can arise under various scenarios—for example, grown children remaining at home or elderly parents joining the household—households typically become larger-than-average because of a greater-than-average number of children.

Since it is common for grown children to support elderly parents, fertility is endogenous to household income; lacking adequate savings facilities, poorer households will choose to have more children in order to gain a safety net in old age. Thus, greater fertility results in poorer households, but poorer households also choose greater fertility in the first place. Regressions which do not account for this reverse causality problem result in estimators which may be biased, inconsistent and correlated with the error term. In such cases, most authors use an exogenous instrumental variable (IV) to obtain predicted values of the desired regressor, followed by a regression of the dependent variable upon all exogenous variables and the predicted endogenous variable. The choice of IV is often one of the most interesting aspect of such studies.

Gupta & Dubey (2003) use the gender of the first two children born as an instrumental variable for fertility in India. In India, as in many LDCs, male offspring are more desirable than female children, as males are seen to be more valuable manual labourers, do not require a dowry, and are perceived as better "insurance and support to the family in old age". Thus, having two girls can proxy for an exogenous increase in fertility, as such families will be motivated to have additional children for reasons exogenous to their level of poverty. The authors find that, without accounting for the endogeneity problem, fertility has a significant effect on poverty, but the effect is roughly halved when fertility is instrumentalized. From the first-stage regression, the authors find that the several variables significantly "determine" fertility, including women's education levels, women's employment (binary), religion, household head's gender, mother's age, first two children being female (binary), and the mother's age at her first childbirth[4].

Chun & Oh (2002) also use the female gender of the first two born children to instrument for fertility, in this case to determine the effect of fertility on female workforce participation in Korea (the decision to have children is influenced by whether the would-be mother has a career). Interestingly, Angrist & Evans (1998), cited in Chun & Oh (2002), use *same-sex siblings* (both male or both female) to instrument for increased fertility in America, as American parents theoretically prefer both male and female children. It is clear that, in addition to econometric considerations, the cultural context of the sample studied is an important determinant of the validity of the instrument.

[4] Gupta & Dubey do not use mother's age as an instrument, but rather as an exogenous regressor in the main regression. Nevertheless, the mother's age at her first childbirth enters the first stage regression (as a predictor of fertility) along with the instrumental variable (binary for first two children being female) and all the second stage regressors, as is the typically methodology in a TSLS regression.

Methodological considerations

Choice of dependent variable

If the goal of a study is to determine the association of economic and demographic variables with poverty, then the definition of the metric of poverty is key. While some studies define welfare according to household utility functions or nutritional status, most studies use either a consumption-derived measure or an income-derived measure.

Deaton (1997) points out that the permanent income hypothesis is a frequently-cited theoretical motivation for using consumption rather than income. Assuming access to savings and credit, an individual's consumption will reflect his potential welfare over his lifetime, whereas income is sampled at one point in time, and may therefore be a poor proxy for actual welfare. However, Deaton suggests that empirical support for the permanent income hypothesis is "mixed at best", and the choice of using consumption may have more to do with practical data considerations. Coudouel et al (1999) outline some of these practical considerations as: a) interviewees may be unable to recall erratic cash inflows, and may find it difficult to assess input costs for agrarian activities, and b) it may be difficult or impossible to assign monetary values to farm produce which is consumed by the household.

This study makes use of both consumption and income data for the purpose of poverty profiling—that is, means of consumption and income variables are calculated across household income quartiles. However, the poverty regression uses only income per capita as the dependent variable. This choice was made because using income as a metric for welfare, in this study, does not suffer from most of the drawbacks outlined by Deaton and Coudouel et al, while offering several distinct advantages.

The permanent income effect is weak for very poor, agricultural households in Gujarat. Many households do not have access to savings and credit, and are therefore constrained in their ability to smooth consumption across years (indeed, this is especially problematic given the incidence of drought in the sample area). While households anticipate facing steep reversals in net cash flow throughout the year (positive during harvest season and negative for the remainder of the year), income in this study was gathered for the entire year in order to capture this variability.

The interview subjects were able to recall approximate quanta of inflows from their various sources of income without very much difficulty. Moreover, interview subjects were not asked for net income, but rather, were asked a series of questions for which the answers were relatively easy to recall. The responses to these limited-scope questions were then aggregated to arrive at a net income number . For example, rather than asking subjects for their net income from groundnut production, groundnut net income was derived from the amount of produce sold in kgs, the imputed value of the amount consumed by the household, approximate prices (constant for entire sample), and seed, fertilizer, pesticide, and labour input costs. The method of derivation of the income figures is shown in the appendix.

Perhaps most importantly, the use of consumption as a proxy for welfare assumes that households with higher income will spend more on consumption. When sampling very poor households, this may certainly be true. However, when studying households whose income exceeds their minimum requirements, consumption fails to keep up with welfare—above a certain level of income, consumption will level off and savings will increase. In such cases, the level of consumption plus any savings may be a good proxy for welfare. However, most respondents sampled here were unable to accurately estimate their savings. Thus, imputed income is a more accurate measurement of household welfare than consumption, or consumption plus savings.

Specification of dependent variable

Whether per capita income or consumption is taken as the proxy for household welfare, the chosen variable can either be specified in absolute terms (untransformed, or in natural log form) or alternatively, with reference to a poverty line, as shown by Ravallion (1996): y_i/z_i where y_i is the indicator of household welfare for household i, and z_i is the level of the indicator at which the household escapes poverty. In practice, there may be a single uniform z_i if all the households face similar prices and are of similar demographics (size, composition, etc.). For illustrative purposes, an aggregated measure of the poverty of the population can then be calculated. A common measure is the headcount index, which is the proportion of households for which $y_i/z_i < 1$. The poverty regression would then be:

$$y_i / z_i = \beta x_i + \varepsilon_i$$

Where x_i is a vector of observed household characteristics, and e_i is the error term.

A significant drawback of this specification is that the regressor coefficients may not be robust to varying definitions of z_i. As discussed above, Serumaga-Zake & Naude (2002) state that the results of their regression are not robust to varying equivalence scales[5], while Sen & Yoon (2005)'s results are robust. Using an absolute measure of welfare, rather than a measure relative to a poverty line, avoids this problem.

Type of regression

Having defined the dependent variable, the regression can take one of two forms: either a logistic or a standard OLS regression. In the probit/logit regression, a binary variable h is defined as h=1 when y/z <1 and h=0 when y/z>1, and the model specification is then:

$$\Pr ob[h = 1 | X] = \phi(\beta X + \varepsilon)$$

Where Φ is the cumulative standard normal distribution function and X is a vector of household characteristics. The type of regression therefore measures the association of the regressor variables with the probability of a household being poor.

By contrast, the standard linear OLS regression is specified:

$$y / z = \beta X_i + \varepsilon_i$$

Neither method dominates the other, as the choice of regression depends on the use to which the results will be put. The logistic regression typically benefits from better fit. Logistic regression can be used to distinguish between chronic poverty and transient poverty when using panel data, and "to assess the predictive power of various variables used for means testing" (Coudouel et al, 1999).

However, logistic regressions are more vulnerable to specification errors. Perhaps most importantly, the logistic regression treats the dependent variable as binary when it is in fact continuous. The transformation of the continuous income variable into the binary poor/non-poor variable truncates the available data and may result in less powerful conclusions. Coudouel et al (1999) argue that, "since there is no difficulty in predicting poverty from a linear regression, this type of regression should be used instead of probits/logits."

Data set[6]

The data modeled here is taken from a household survey which was originally produced by the author between March and August 2004 for the Aga Khan Rural Support Programme, India (AKRSP,I), an NGO-member of the Aga Khan Development Network (AKDN). AKRSP works to improve the livelihoods of the rural poor in a sustainable manner.

[5] Equivalence scales are used to adjust household income for demographic composition, eg: children require less food than adults. In practice, this method results in adjusted values of y_i. However, as discussed below, the result is the same as varying z_i. Thus, the use of equivalence scales and varying poverty lines should give the same result.

[6] This section is based on a paper previously written by the author. See references.

The data was gathered in the field by the author with the aid of two translators. A study entitled, "Rural Household Cash flows: the Mota Sakhpar Case" was written by the author in August 2004 based on the collected data and subsequent analysis (see references). As the study aimed at gathering cash flow information for rural households in order to allow microfinance practitioners to better understand the needs of their clientele, the survey was designed to this end. However, because of the broad range of information gathered, the data serves equally well as the basis for a poverty profile.

The survey was carried out on 50 households, 6 of which were eventually excluded as outliers.

Sample village characteristics

Choice of setting: Mota Sakhpar

Mota Sakhpar village is located within the Surendranagar district of Gujarat. Certain characteristics of the village make it a favourable location for the basis for a microfinance or a poverty profile study.

Mota Sakphar had not been studied extensively in the recent past. As a result, the residents of the village had not become averse to speaking at length with researchers, and were expected to be willing to share their experiences with the research team. Residents of villages that have been intensively studied have been known to be less forthcoming with detailed information ("survey fatigue")—a natural reaction to having to participate in a large number of studies.

Finally, it was felt that Mota Sakphar was a 'representative village', in that it has the typical social, economic and physical characteristics of the programme area. In other words, the information gathered from this village was expected to approximate the values that would be obtained from a sample of all the villages in the Surendranagar programme area.

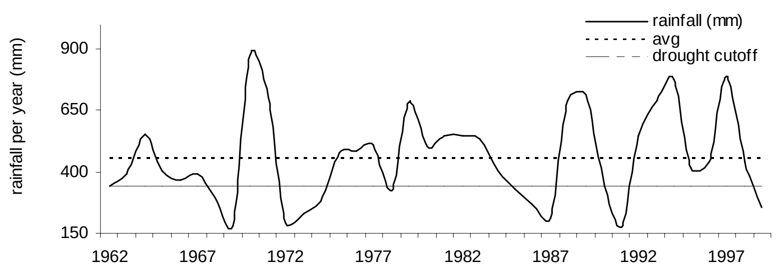
Physical attributes of sample village and province

The Surendranagar district is a landlocked portion of central Gujarat, bordered in the north by the Rann (desert) of Kutch and Mehsana districts, in the south by the Ahmedabad and Bhavnagar districts and in the west by the Rajkot district.

Within Surendranagar, Mota Sakphar is located approximately 30 kms away from the district (taluka) headquarters at Sayla. Mota Sakphar is reached from the cluster office by a well-paved, 12 km road that branches off from the Ahmedabad-Rajkot highway (CBA water, 2000).

The village is located in a severely drought-prone area. Drought[7] occurs approximately once every three years. Rainfall data is shown below.

Figure 1: rainfall in Sayla (district HQ) between 1962 and 1999



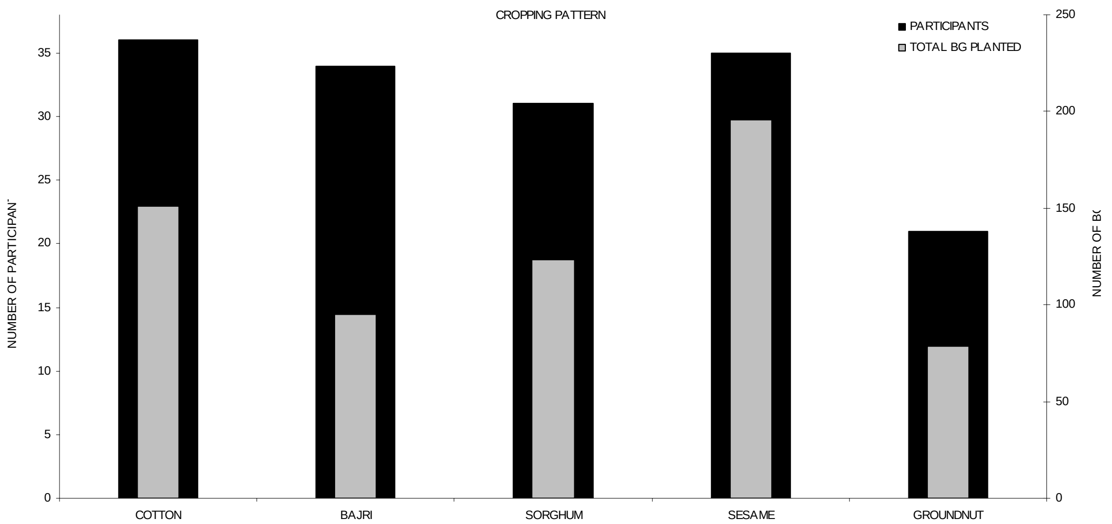
[7] Drought is defined as rainfall less than 75% of average rainfall

The soil in this area consists primarily of alluvial, light sandy soil, rocky medium black soil, and red soil. Only a small proportion of the land in Mota Sakphar has a slope of more than 3%. (CBA water, 2000). The village is located along a river that typically flows from July through December. Many wells are located along the river bed and river bank, providing water for household needs. Most households, due to an AKRSP(I) intervention, have individual water tanks for water storage. In drought years, government trucks replenish the water reserves in the holding tanks.

Economic attributes

Agriculture or agricultural labour is the primary occupation of 92% of the sampled households—landless households make up the other 8%. Crops under cultivation include cotton, bajri[8], sesame, sorghum (jowar) and groundnut. Sesame and groundnut produce are both consumed by the household and sold. Sorghum is used exclusively as animal fodder for farm and dairy animals (apart from a small fraction of those sampled who also sell sorghum). The entire cotton yields are sold.

Figure 2: cropping pattern in sample village



Most households keep animals for a variety of purposes, including milk and other dairy product production and farm use. 62% of sampled households produce enough milk from their animal husbandry operations to provide for household dairy consumption, and of these, several sell some portion of their dairy yield. Some households also produce ghee (clarified butter), primarily for sale.

Previous studies such as Chauhan (2003) had indicated that in households in villages nearby to Mota Sakphar, all but the wealthiest households suffer from persistent income-expenditure gaps, and credit needs are fulfilled by moneylenders, village institutions and friends and relatives.

[8] bajri is a type of wheat used to make flour

Comparison of data set attributes with village attributes

Figure 3: sample vs. village attributes

	Sample	Mota Sakphar
Sample size (frequency):		
Total households	44	133
Individuals	275	n/a
Wealth ranking (percentage, frequency in brackets)*:		
Category A	38% (17)	40% (53)
Category B	27% (12)	40% (53)
Category C	29% (13)	20% (27)
Ranking n/a	7% (2)	--
Landholding pattern:		
Landless	8% (4)	10%**

**wealth ranking categories according to AKRSP's methodology*
***proportion of landless across Surendranagar programme area*

As the sample comprised one-third of the village, the distribution across wealth categories and other attributes of the sample were expected to mirror those of the village. Figure 3 shows that the sample is representative of the larger village.

External validity of data set: comparison with Indian National Surveys

Selected variables from the sample data are compared with statistic gathered by the Indian National Statistics Society (NSS) below. The purpose of this comparison is two-fold.

Firstly, although this study is undertaken with a view to informing the policies of the NGO working in the sample village, the conclusion drawn herein can be more broadly applied if the households are deemed to be typical of the agricultural household within the state of Gujarat, and perhaps across the country. In other words, the presentation below aims to establish the external validity of the data set.

Secondly, the credibility of the data set is bolstered by establishing that its values lie within a reasonable margin of those gathered/calculated by the NSS. The sample need not necessarily mirror NSS data, as the households sampled are not identical to those sampled by the NSS. The sample should deviate even from NSS data on rural Gujarati households, as the households within the sample village benefit from NGO interventions, while having unique physical and other conditions (for example, being extremely drought-prone). These characteristics are perhaps most closely matched by the NSS data on rural Gujarati households, but are nevertheless not identical between the two samples.

Consumption data

Figure 4: monthly per capita consumption expenditures: sample vs. selected NSS-surveyed groups

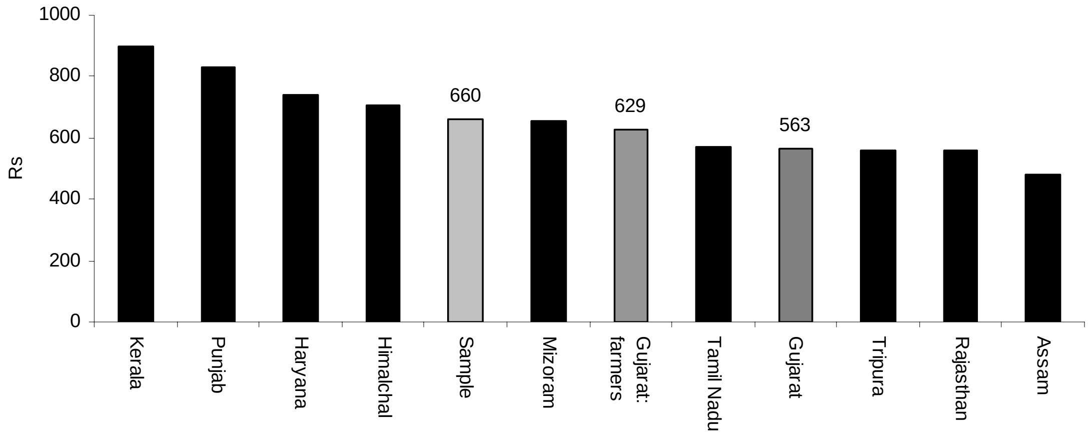


Figure 4 shows that monthly per capita consumption within the sample closely matches the relevant NSS subgroups. Monthly per capita consumption within the sample is 660 rs, only slightly lower for NSS-sampled Gujarati farming households (629 rs[9]), followed by all NSS-sampled Gujarati households (563 rs). The close match seems to validate the consumption figure, which in this study, was aggregated from a wide variety of individual variables.

The slightly higher consumption figure may be due to sampling error or a genuine higher level of wealth amongst the sampled households versus typical NSS-sampled households. This may be due to any number of factors: the sampled village farmers have access to improved seeds and fertilizer, pesticides, cooperative society loans, and for some households, self-help group (SHG) microfinance funds, as well as consistent, professional, long-running NGO interventions.

Figure 5: average monthly expenditure on selected consumer items (per person per month, rs)

Selected expenses	sample	Gujarat*	all India*
Bajra (millet flour)	8	15	2
Milk	52	65	46
edible oil	77	40	23
fruits & vegetables	51	49	37
Sugar	16	15	10
Salt	1	1	1
Spices	55	13	10
beverages/water	23	12	8
intoxicants (betelnut, etc)	24	17	14
fuel & light	31	52	47
clothing & footwear	77	44	43
Education	35	13	17
medical care	67	37	34
personal care/hygiene	16	14	15
Transportation	46	20	13

**Statewise and itemwise average value of consumption per person of farmer households per 30 days*

[9] all currency figures refer to rupees (rs) unless otherwise stated. At the time of the household survey, $1 USD = 33 rs

Figure 5 disaggregates consumption into selected items. Spending on most items is similar between all groups. Reassuringly, when the sample results deviate from NSS results, they are generally closer to the Gujarati household values than the all-India values. Spending on some items seems significantly different for sample households vs NSS households.

The sampled households spend more on edible oil and spices than NSS households. Indeed, Sunderji (2004) found that spending on cooking oil was a major expenditure for the sampled households. Rural Gujarati cooking is indeed very oily and very spicy, yet this does not explain the deviation from NSS data on rural Gujarati households. Expenditure on spice may be partly explained by the fact that the sampled households generally do not grow their own spices, whereas farming households elsewhere within the state and across India may grow their own spices.

Higher expenditure on items such as clothing and footwear, education, and medical care may reflect the higher incomes of the sampled households, as well as their access to such facilities (households without access to schools will not spend very much on education). Previous studies on nearby villages, such as Chauhan (2004), have concluded that expenditure on health care is a major cause of indebtedness among these households. However, Sunderji (2004) did not find that this was the case.

Higher spending within the sample on transportation may reflect access to relatively good quality roads for such a remote rural area, as well as the established autorickshaw-taxi business operating on these roads.

Income/revenue data

Figure 6: sources of monthly income/revenue: sample vs. selected rural NSS-surveyed households (rs)

income source	Sample	Gujarat	India
farming net income**	2472	1164	969
husbandry revenue*	2100	1141	593
wage work revenue	559	925	819
other business participation rate	29%	14%	17%
other business net income	637	140	236
landholding (ha)**	3.40	2.89	2.13

**imputed from output and prices for sample*

***NSS data for rural households self-employed in agriculture*

Figure 6 shows that the sampled households either over-estimated their income or do genuinely earn more than typical NSS-surveyed households. It should be noted that these figures do not adjust for households size, i.e.: they are not on a per capita basis. However, assuming similar demographics across samples, especially within-state, there is no reason to suspect that the results would not be robust to measuring them on a per capita basis. Moreover, higher incomes for the sampled households is consistent with their generally higher consumption spending versus NSS samples.

The sampled households earn significantly more than NSS-sampled households from farming and husbandry work, but earn significantly less from wage work. This is consistent with the fact that the sampled households are generally larger landowners, and should therefore earn more from farming. However, the sampled households earn much more from other business income. Indeed, 29% of the households participate in non-farm, non-agricultural labour (wage work) activities. In the sampled area, such activities include chakra (rural taxi) driving, diamond polishing, and selling vegetables from mobile trolleys. Sunderji (2004) found that non-farm business income was a very significant source of income for those households that participated in such activities. AKRSP is engaged in promoting non-farm work, especially given the drought-prone conditions in the sampled area and the consequent unreliability of farming as a sole income source. Throughout March-August 2004, the NGO was negotiating real-estate and a supply-chain for a milk-solids factory.

Poverty profile of sample village

While the poverty regression is a statistically more sophisticated method of data analysis, simply calculating each variable's mean across income quartiles can also be a useful method of gaining a further understanding of the poverty in the sample area. Deaton (1998) writes, "rather than use the theory to summarize the data through a set of structural parameters, it is sometimes more useful to present features of the data, often through simple descriptive statistics, or through graphical presentations of densities or regression functions, and then to think about whether these features tell us anything useful about the process whereby they were generated."

In the following graphs and charts, various economic and demographic attributes are calculated for each income per capita quartile. Number 1 denotes the poorest quartile, number 4 is the richest.

Selected demographic variables

Figure 7: number of children and dependency ratio across income quartiles

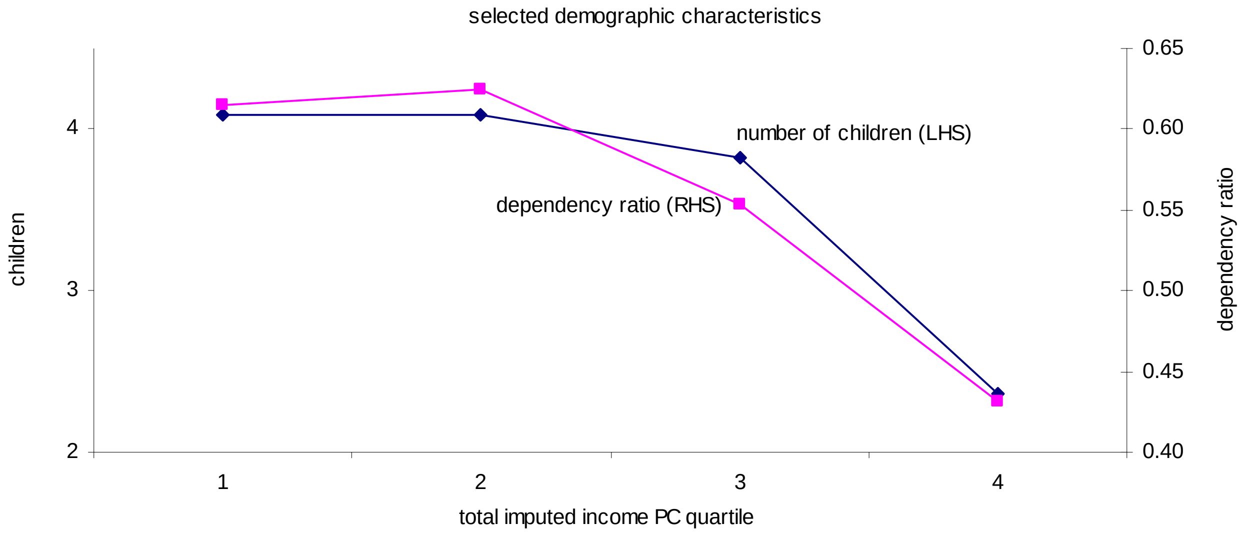


Figure 7 shows that wealthier households have fewer children, and that the dependency ratio (defined as the number of non-productive members of the household as a percentage of the total household) decreases for wealthier households. The endogeneity of the fertility variable maybe part of the cause of this relationship, but larger households also have more mouths to feed and hence often lower income on a per capita basis. Notably, the number of children and dependency ratio drops more sharply with each increase in wealth quartile. While the poorest 22 household do not have vastly different demographic profiles, the third quartile is composed of markedly smaller households, and the effect is even stronger for the wealthiest 11 households.

Figure 8: additional demographic characteristics

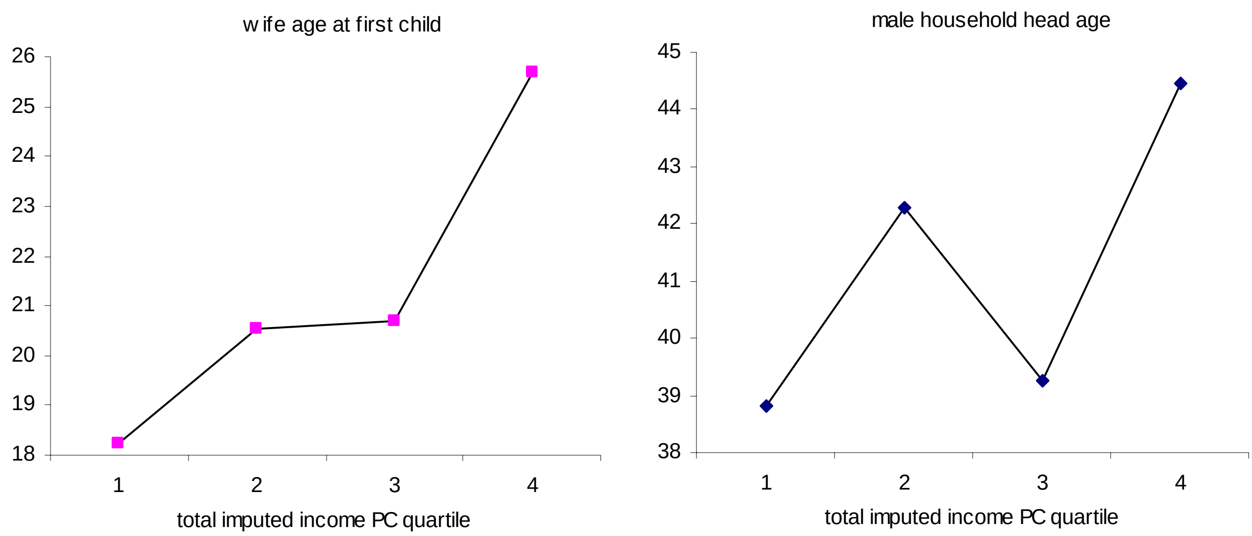


Figure 8a shows that women of wealthier households tends to wait longer before having children. Several effects may be at work here; wealthier household tend to be better educated and are consequently more aware of family planning; also, poorer households desire more children as a safety net in old age and therefore choose to begin having children early. The earlier the mother begins having children, the greater the number of children will she have—thus, younger mothers have larger families, and larger families are typically poorer on a per capita basis. This effect may also be causing the relationship shown in 8a.

The male head of the household's age demonstrates a less clear pattern, but is generally conforms to findings from other studies (see Sen & Yoon, 2005): older households are generally wealthier. This may be because older heads of households command greater authority and access to resources within the village, or have accumulated more land or productive assets. In the sample village, marriage of daughters is a major expense, and older households may have borne this financial burden in earlier times and since recovered (paid off debts, etc). Older household may also be smaller, as the grandparents generation may have passed on.

Figure 9: lower-caste content across income quartiles

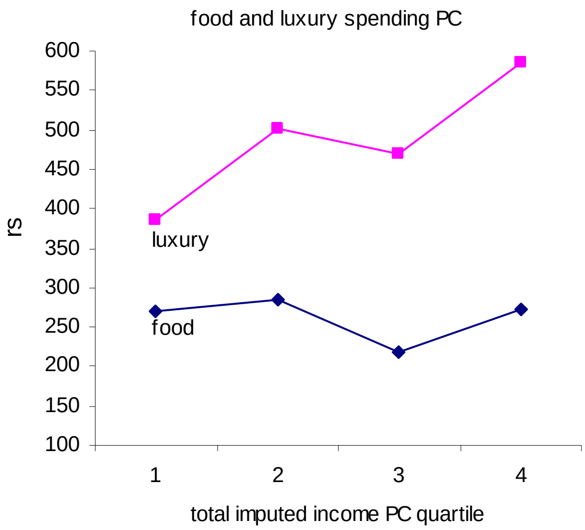
	income quartile			
	1	2	3	4
lower-caste members (%)*	0.27	0.27	0.09	0.00

**lower caste is Harijan, one of four caste groups in the sample village*

Sen & Yoon (2005) found that lower caste individuals are poorer than non-lower caste individuals. The Harijan (sometimes formerly called "untouchables") households seem to be clustered in the lowest two quartiles, while the wealthiest quartile contained no Harijan households. Systematic discrimination was not evident in the sample village and did not seem to be an active area of intervention for AKRSP. Land holding patterns may work against lower caste households, who may historically have been allowed only small or poor-quality plots of land.

Selected consumption variables

Figure 10: monthly food and luxury spending across income quartiles



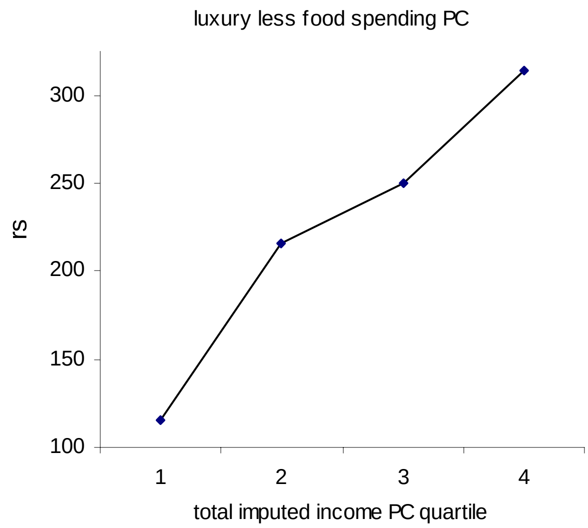


Figure 10a shows that food expenditure across all income quartiles is remarkably stable, while luxury spending generally increases for wealthier households. The relationship is even more clear in figure 10b—the difference between consumption expenditure on luxuries versus food increases linearly across income quartiles. Wealthier households spend the same amount of money on food as poor households, but spend more on luxuries.

Selected income/revenue/capital variables

Figure 11: income-related variables across quartiles

	income quartile			
	1	2	3	4
Average imputed total net income (rs)	24997	42558	59672	88724
Average imputed total net income PC (rs)	3431	6046	8768	17005
Average total land holding (bh)*	10.8	14.5	14.1	25.9
Average total land planted (bh)*	9.5	12.7	12.0	21.6
Average of total land holding PC (bh)*	1.5	2.0	2.1	5.1
Average no. animals owned	0.9	1.3	1.0	1.9
Average no. animals owned PC	0.12	0.19	0.14	0.39

**land holding measured in Gujarati bigha[10] (bh)*

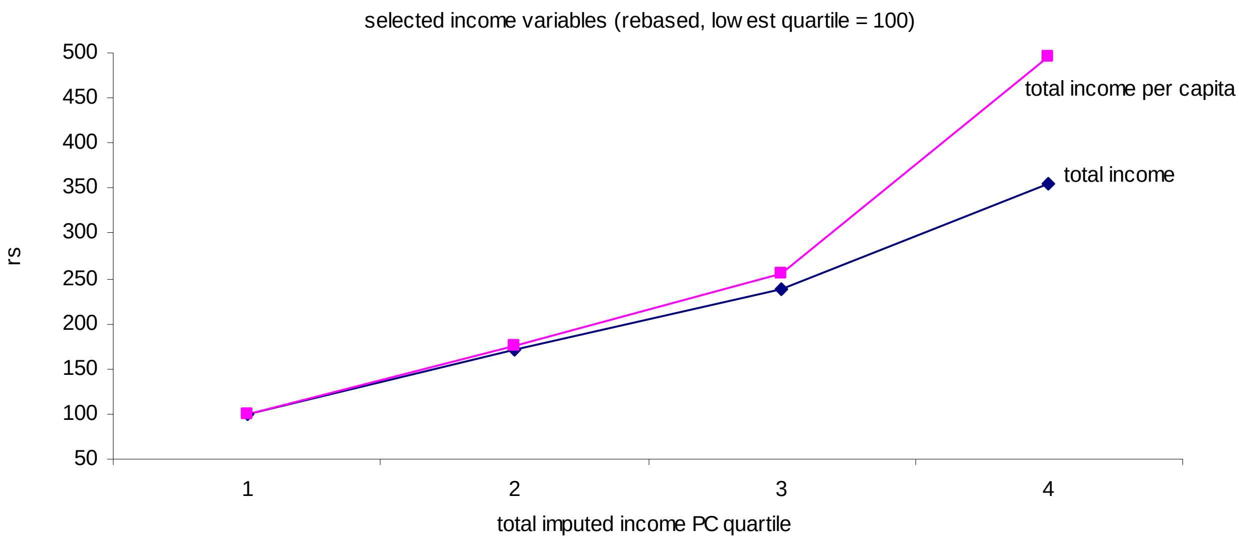
Figures 11 and 12 show that the households which are wealthier on a per capita basis are also wealthier on a household basis. This is an important conclusion: it is not simply the number of children which determines wealth on a per capita basis. Wealthier households are wealthier on an absolute basis, but the fact that they are also smaller exaggerates this effect. Especially in the case of households in the top quartile, wealthier households have higher net incomes *and* fewer mouths to feed. As Gupta & Dubey (2003) find, failing to account for household size results in exaggerated estimators for economic and demographic characteristics—i.e.: they seem to contribute to poverty more than they actually do. Once

[10] bigha in this study refers to the Gujarati measurement of land. 1 gujarati bigha = 0.24 acres

household size is taken into account, it becomes clear that both the economic and demographic attributes of households and their size jointly determine poverty on a per capita basis.

Figure 11 also shows wealthier households have more land, more planted land, and more land on a per capita basis. They also have more animals on an absolute and per capita basis. As farming and animal husbandry are the primary occupations in the sample village, these patterns should not be surprising.

Figure 12: income and income per capital across income quartiles



Model Specifications

Two models are presented here. The first treats the number of children in the household as an exogenous variable. In other words, the household's choice of number of children is unaffected by their economic status. The second model treats the fertility decision as endogenous to wealth. From a theoretical standpoint, this is consistent with the literature which suggests that poorer households choose to have more children.

Model 1: fertility as exogenous variable

The OLS model specification is:

$$\ln(incomePC)_j = \beta_{1,cj}x_{cj} + \beta_{2,cj}y_{cj} + \beta_{3,fj}z_{fj}$$

Where the subscript j in all cases denotes the household, x is a vector of landholding variables and y is a vector of agricultural productivity variables for each crop *c*, *z* is a vector of demographic variables (including fertility).

Model 2: fertility as endogenous variable

Fertility is modeled as an endogenous variable through a two stage least squares (TSLS) regression. The mother's age at her first child is the instrumental variable. In the first stage, fertility is regressed against the instrumental variable and all other exogenous variables from the structural model. In the second stage, the natural log of household per capita income is regressed against predicted fertility from the first stage regression, and all the exogenous variables.

The first stage regression is:

$$\hat{k}_j = \beta_{1,cj} x_{cj} + \beta_{2,cj} y_{cj} + \beta_{3,di} z_{dj} + \alpha(wife\_age_j)$$

And the second stage regression is:

$$\ln(incomePC)_j = \beta_{4,cj} x_{cj} + \beta_{5,cj} y_{cj} + \beta_{6,dj} z_{dj} + \delta_j \hat{k}_j$$

Where the subscript j in all cases denotes the household, *wife_age* is the age of the mother at her first child (the instrumental variable for children), x is a vector of landholding variables and y is a vector of agricultural productivity variables for each crop c, z is a vector of demographic variables for each variable d, and $\hat{k}$ is the predicted number of children based on the first stage regression.

Hypotheses

The variables examined and the signs (+/-) of their respective model coefficients are below.

variable	expected coefficient sign
agricultural ownings	
cotton area planted	+
sesame area planted	+
bajri area planted	+
sorghum area planted	+
groundnut area planted	+
no. animals owned	+
agricultural productivity	
cotton productivity	strongly +
sesame area planted	strongly +
bajri productivity	weakly +
sorghum productivity	weakly +
groundnut productivity	strongly +
agricultural inputs	
seeds	+
pesticide	+
fertilizer	+
labour	+
demographic attributes	
extended dependents	strongly -
extended productive members	weakly -
total children	-
male household age	+
male household age squared	-
caste dummy (1=lower caste)	-

The amount of land planted should increase income. Greater land planted should lead to greater output and higher farming income. Animals should also increase income. Although buffalo, for example, require food, shelter, and veterinary care, they should compensate for such expenses by providing a stream of income as dairy producers, and by increasing farming productivity as agricultural capital inputs (buffalo-drawn tractors, for example).

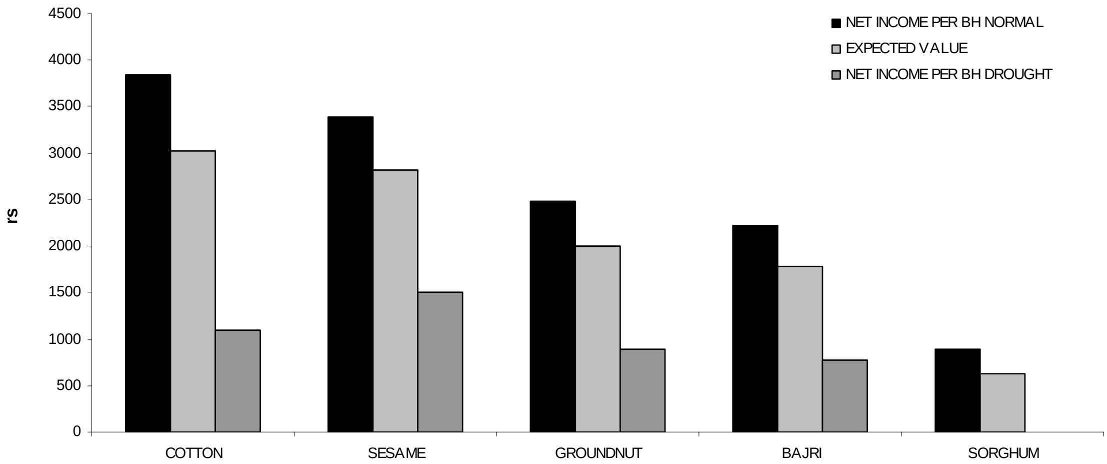
Greater agricultural productivity, defined as kgs of output per hectare, should increase income if that productivity is achieved without an offsetting increase in input costs. While the sign of these variables should be positive, their relative contributions to income (i.e.: their coefficients) are of interest.

Sunderji (2004) examines the productivity of the primary crops grown in the sample village, however, he defines productivity as the average net income per bh of crop grown (rather than as it is defined here, as output in kgs per BH). He finds that cotton is the most productive crop, followed by sesame, groundnut, bajri and sorghum. Sunderji also calculates expected values of crops, as:

$$EV_i = net\_income_{i,normal} \times p(normal) + net\_income_{i,drought} \times p(drought)$$

Where i denotes the crop, p(normal) is the probability of a normal (non-drought) rainfall, and p(drought) is the probability of a drought (defined as less than 75% of mean seasonal rainfall). Sunderji finds that although cotton is the most lucrative (highest NI per BH) crop during a normal year, because of the cotton crop's sensitivity to drought conditions, its expected value is only marginally higher than that of sesame.

Figure 13: expected values of net income of selected crops based on probability of drought



The productivity variable measured in this study is related to the net income per BH variable that Sunderji (2004) examines. Net income per BH for any crop depends on output per BH, prices per Kg, and input costs per BH. Productivity as measured in the current study is analogous to the output component contained within the net income variable measured by Sunderji. Thus, although cotton may have the highest net income per BH, this does not tell us about the increase in income for a unit increase in output productivity. If greater output is achieved only by investing disproportionately large amount of inputs in the farmed area, then net income may actually be reduced by greater productivity. In assigning a hypothesized positive value to crop productivities (for all crops), it is assumed that input costs rise more slowly than output increases.

Similarly, prices obtained for crops will enhance or reduce the effect of increases in productivity on net income. For crops with fetch higher prices in the market, increases in productivity will be more meaningful than crops which fetch poor prices. Since cotton, sesame, and groundnut fetch higher prices than sorghum or bajri (large portions of which are consumed by the household), ceteris paribus, the effect of increases in productivity for those crops on income should be larger than for the latter two.

Greater agricultural inputs such as seeds, pesticide, and fertilizer should increase income by increasing crop output. Theoretically, if such inputs had negative returns they would not be used in the first place. Extended family members should decrease income per capita—a greater number of mouths to feed means there is less food per mouth. While both extended family dependents and productive extended family dependents should both decrease income, productive dependents should decrease income by less than those members who do not contribute to income. The implicit assumption here is that extended family members consume more than their contribution to income, and are thus net losses for the household on a per capita basis.

Children are in some sense no different from other non-productive household members—they should decrease per capita income (the economic value of children accrues when they are grown and support their elderly parents). Previous research (Sen & Yoon, 2005) has found that older households (proxied by the male household head's age) are wealthier, but there are declining marginal benefits to older households (proxied by the square of the male household head's age). The theoretical reasons for this relationship are explained above. The sample village should not be idiosyncratic in this respect; older households are expected to be wealthier here too.

Lower caste household tend to be poorer in India for a variety of reasons—systematic discrimination, historically poverty, limited access to resources, poor allocation of land within village, limited social mobility, restrictions on occupations, etc. The lower caste households within the sample village should have lower incomes.

Results

Model 1

The model presented here is the reduced form OLS model regressing household per capita income on a vector of economic and demographic household characteristics. Several variable were dropped, as they were insignificant and interfered with the significance of other variables. These variables included all the inputs (seeds, pesticide, fertilizer, labour), number of animals owned, and the caste dummy.

Model Summary(b)

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	0.909	0.827	0.734	0.33651

a. Predictors: (Constant), total children, bajri_prod, sesame area, sorghum area, extended_deps, sorghum_prod, cotton area, sesame_prod, groundnut area, groundnut_prod, productive extended members, cotton_prod, husbandage2, bajri area, husband age

b. Dependent Variable: ln_imputedtotalincomePC

ANOVA(b)

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	15.166	15	1.011	8.928	0.000
	Residual	3.171	28	0.113		
	Total	18.336	43			

a. Predictors: (Constant), total children, bajri_prod, sesame area, sorghum area, extended_deps, sorghum_prod, cotton area, sesame_prod, groundnut area, groundnut_prod, productive extended members, cotton_prod, husbandage2, bajri area, husband age

b. Dependent Variable: ln_imputedtotalincomePC

Coefficients(a)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	6.540	0.821		7.968	0.000
	cotton area	-0.013	0.026	-0.050	-0.488	0.630
	bajri area	0.038	0.051	0.110	0.751	0.459
	sorghum area	0.010	0.021	0.047	0.455	0.652
	sesame area	0.058	0.022	0.361	2.687	0.012
	groundnut area	-0.007	0.027	-0.027	-0.255	0.801
	cotton_prod	0.002	0.000	0.428	3.479	0.002
	sesame_prod	0.003	0.001	0.307	3.111	0.004
	groundnut_prod	0.001	0.000	0.285	2.758	0.010
	bajri_prod	0.000	0.000	-0.061	-0.525	0.603
	sorghum_prod	0.000	0.000	0.049	0.534	0.597
	productive extended members	-0.177	0.056	-0.372	-3.156	0.004
	extended_deps	-0.269	0.110	-0.233	-2.455	0.021
	husband age	0.079	0.035	1.477	2.253	0.032
	husbandage2	-0.078	0.035	-1.432	-2.274	0.031
	total children	-0.021	0.041	-0.053	-0.512	0.613

a. Dependent Variable: ln_imputedtotalincomePC

Model 2

The model presented here is the reduced form two stage least squares (TSLS) model regressing household per capita income on a vector of economic and demographic household characteristics. Several variable were dropped, as they were insignificant and interfered with the significance of other variables. These variables included all the inputs (seeds, pesticide, fertilizer, labour), number of animals owned, husband age, husband age squared, and the caste dummy.

```
MODEL:  MOD_2.

Equation number:    1
```

Dependent variable.. ln_imputedtotalincomePC

```
Listwise Deletion of Missing Data

Multiple R           .87851
R Square             .77178
Adjusted R Square    .64817
Standard Error       .37320
```

Analysis of Variance:

```
               DF    Sum of Squares      Mean Square

Regression     13         11.304276         .86955971
Residuals      24          3.342677         .13927821

F =       6.24333       Signif F =  .0001
```

----------------- Variables in the Equation ------------------

Variable	B	SE B	Beta	T	Sig T
totalchildren	-.176542	.100001	-.436319	-1.765	.0902
cotton_prod	.000639	.000738	.186088	.866	.3950
sesame_prod	.002497	.001094	.313260	2.282	.0316
groundnut_prod	.000972	.000747	.190160	1.301	.2055
bajri_prod	.000150	.000405	.060194	.370	.7145
sorghum_prod	.000405	.000316	.153705	1.282	.2121
cotton area	-.013805	.030579	-.057519	-.451	.6557
bajri area	.014407	.059957	.044957	.240	.8122
sorghum area	.049309	.031767	.202765	1.552	.1337
sesame area	.052990	.027524	.340178	1.925	.0661
groundnut area	.026890	.041765	.118860	.644	.5258
extended deps	-.424180	.158445	-.365554	-2.677	.0132
productive ext mems	-.171509	.083505	-.280916	-2.054	.0510
(Constant)	8.937646	.460739		19.398	.0000

Analysis

Economic variables

Area farmed

One of the most robust results from both regression models presented here, as well as other model specifications not shown here, is that sesame has a strong positive effect on income along both the land planted and the productivity dimensions.

A 1 bigha increase in the amount of sesame area planted is found to increase net income per capita between 5%-6% (5.3% in model 1, significant at 5% level, 5.8% in model 2, significant at 10% level). The area farmed with other crops was not found to significantly affect income. Notably, increasing the area farmed with cotton (model 1 and 2) and groundnut (model 1) was found to have a negative effect on income, although the effect was not statistically significant.

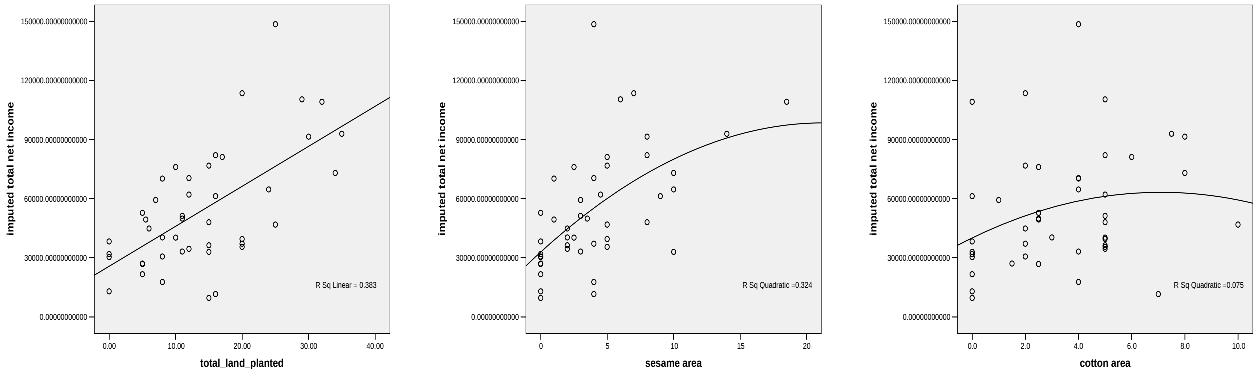
Sorghum and bajri, non-cash crops which are primarily consumed by the household, were found to be to be insignificant determinants of household income. It is not surprising that these low-margin crops which are primarily consumed by the household are not significant determinants of income—a kg of sorghum sells for only 7 rs/kg, while a kg of bajri fetches only 7.25 rs/kg. When the net income from sorghum is imputed, based on reported market prices, kgs of output, and input costs, 25% of households which grow sorghum make a loss. Sunderji (2004) examines this problem, concluding that households persist in growing sorghum for several reasons: a) the average household spends 5 rs to grow 1 kg of sorghum, while the cost in the market is 7 rs per kg, b) sorghum is often inter-cropped (planted around other crops) or on poor, infertile soil, and c) there is some non-monetary benefit of having an assured source of sorghum, as it is fed to animals which produce milk for household consumption, and milk products for sale in the market. Therefore, sorghum has a negligible impact on household income because it is a low-margin, frequently loss-making crop that serves primarily as an input for dairy production.

Similarly, bajri is a low-margin crop that is consumed by the household. Bajri is used to make flour, and it is the primary source of starch for households in Gujarat. In this sense, there is some intrinsic value (as an assured source of food) to growing bajri which is not captured in the imputation of its profitability. Moreover, input costs relative to market sale prices are high for both bajri and sorghum, resulting in widely varying profitability and an unclear relationship between area grown and income.

Cotton, sesame, and groundnut, on the other hand, sell for 22 rs/kg , 33 rs/kg, and 20 rs/kg, respectively. Perhaps the most interesting puzzle emerging from both regression models is that cotton and groundnut area are not significant determinants of income, while sesame area is.

Input costs may partly explain these results. Figure 14 shows that total land planted and income are highly linearly correlated, sesame area planted and income are less clearly but still significantly correlated, and cotton area planted and income are poorly correlated.

Figure 14: land planted vs. income, sesame area planted vs. income, cotton area planted vs. income



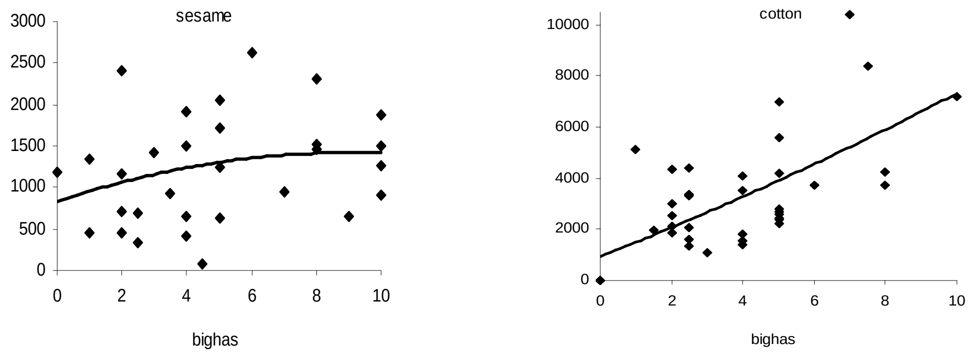
The sensitivity of the cotton strain to growing conditions is evidenced by its relatively low expected value; when the water supply is attenuated during drought seasons, cotton output falls sharply. Furthermore, Sunderji (2004) found an inverse relationship between farm size and farm productivity (where productivity is defined as kgs of output per bigha), suggesting that crops generally suffer from decreasing marginal returns. These two considerations—cotton's sensitivity to farming conditions and declining marginal returns for all crops—imply that farming a greater area would not necessarily contribute to more income. However, both of these effects would manifest themselves via productivity, which is explicitly controlled for in the regression through the crop productivity variable.

Input costs may be preventing greater area cropped from translating into greater net income. This may be a systematic relationship (input costs become increasingly high for larger farms), or input costs may vary across farms idiosyncratically, obscuring any relationship between larger farms and higher income. Figure 15 shows that total input costs (seeds+pesticide+fertilizer) are relatively low (0 to 3000 rs) and widely dispersed for sesame, with no clear relationship between area planted and costs. On the other hand, input costs for the cotton crop are relatively high (0-10,000 rs) and increase linearly with area planted (R^2=0.43).

These results may be partly explained by one of the primary differences between the two crops: sesame is a robust, drought-resistant crop which requires good conditions to become established but can continue to grow in sub-optimal conditions, while cotton is drought-sensitive, even after having become established[11]. Thus, input costs seem to explain at least some part of the regression results—with large fixed costs and small variable costs for growing sesame, greater areas planted lead to greater income. With cotton, small fixed costs and large variable costs make planting a wider area less income-augmenting.

[11] http://www.hort.purdue.edu/newcrop/afcm/sesame.html

Figure 15: land planted vs. total input costs (seeds/pesticide/fertilizer) for sesame, cotton



Crop productivity

Model 1 finds that cotton, sesame, and groundnut productivity significantly increase income. A 1 kg per bigha increase in sesame productivity raises net income per capita by 0.3%. Equivalent numbers are 0.2% for cotton, and 0.1% for groundnut. These results are all significant at the 1% level. Model 2 finds that only sesame productivity significantly increases income; a 1 kg per bigha increase in productivity increases income by 0.25% (significant at 5% level).

The results from model 1 are aligned with the hypothesis: greater agricultural productivity for cash crops, defined as kgs of output per hectare, should increase income if that productivity is achieved without an offsetting increase in input costs. Model 1 suggests that increased productivity in the primary cash crops can be achieved without matching escalations in input costs, while model 2 suggests that this is only the case for sesame.

Moreover, as stated in the hypothesis, the relative effects of productivity increases for different crops is of interest: increases in sesame productivity translate into greater income most strongly, followed by cotton, followed by groundnut. This root cause of this may again be found in the input requirements and profitability of the sesame crop (input requirements are discussed above). However, no clear relationship or pattern was found between crop productivity and input costs for any of the cash crops.

It is interesting to note that sesame productivity is the only productivity variable that is robust to both model specifications. However, there is no clear intuitive explanation for the effect of endogenizing fertility on the coefficients on crop productivity.

Non significant economic variables

All inputs (seeds, pesticide, fertilizer, and human labour) were not found to significantly affect household income. They were removed from both models as they were not variables of primary interest and were interfering with the significance of other variables.

Demographic variables

Model 1 finds that all demographic variables aside from total children were significant determinants of income, while model 2 finds that all were significant except husband age and husband age squared. The direction (+/-) of the coefficients are, in all cases, as hypothesized and are consistent with the theory found in the literature.

Model 1 finds that a one-year increase in the husbands age leads to a 7.9% increase in household income per capita, while the negative coefficient on husbands age squared signifies declining marginal returns to old age (these two variable were not included in model 2 as they were not found to be significant). There are numerous explanations for this relationship. Older households are theoretically able to command more resources, such as privileged access to shared resources such as water, or membership in group microfinance. Older households will also have become more established in their farming operations, with long-standing access to seeds, pesticide, fertilizer, and human labour inputs, and an established network for distributing produce or transporting it to the village market. Such household may have greater assets, such as animals required for husbandry (buffalo, cows, and goats) or capital goods for farm use (tractors, animals, utensils, water pumps, drainage systems, etc.) Older households may also receive remittances from children.

The declining marginal benefits of old age suggest that the benefits outlined above increase rapidly during a certain phase of aging, but decline thereafter. In this context, it is notable that the demographic factors that might influence the coefficient of the age variables have been controlled. For example, the variables for total children, productive extended members, and extended dependents hold the various components of household size constant, such that the observed effect of husband age is a measurement of non-demographic changes. In the absence of these controls, one may well have argued that in older households, the grandparents generation would have died, leaving the household with fewer dependents. In any case, one important variable has not been controlled for, and may be partly causing the age effect: older children (which were still counted in the survey as children if they are under the age of 18) may contribute to income, either by working on the farm of taking a job, such as selling tea or vegetables from a trolley.

The number of extended family members who do not contribute to income (extended dependents) was the single most important variable explaining low income—increasing the number of such dependents in the household by one decreases income by almost 27% according to model 1, and by over 42% in model 2. Extended family members who do contribute to household income (such as grown children) also significantly reduce income—an additional such member within a household decreases income by almost 18% (17% in model 2). A point to note here is that generally increasing members within a household will decrease income per capita. The negative coefficient on the productive extended members variable should not be interpreted to mean that such members are less hard-working than members of the nuclear family (no variable for non-child members of the nuclear family was gathered in the survey), but rather should be compared with the coefficient for extended dependents. As hypothesized, productive members contribute less to household poverty than extended dependents.

Model 1 finds that the total number of children within the household is not a significant determinant of poverty, while model 2 finds that total children is significant at the 10% level—an additional child reduces income per capita by over 17%. Thus, fertility when treated as exogenous to household income has no impact on income, but when instrumented by the age of the mother at her first child, is a significant cause of decreased income.

The difference in results between model 1 and 2 on fertility may suggest that the fertility decision truly is endogenous. Model 1 may find a significantly insignificant coefficient on fertility because of the reverse causality problem described in the introduction of this study.

Conclusions

This study aims to determine: i) what are the economic and demographic characteristics of farming households, and ii) what characteristics distinguish poor households from non-poor households. More specifically, the econometric model presented here aims to determine the impact of variability in economic and demographic characteristics on the household's level of per capita poverty.

The two models present differing conclusion on poverty for the sample village. One of the most robust findings emerging from both models is that the area planted with sesame and the productivity of the sesame crop (defined as output in kg per bigha) are significant determinants of household per capita income. Only one of the models finds that groundnut and cotton area planted and productivity is associated with greater income. Neither of the models finds that the non-cash crops (land planted or productivity) is associated with higher or lower income.

Demographic characteristics also significantly determine income. The number of extended household members who do not contribute to household income are significantly associated with decreased per capita incomes. Extended household members who do contribute to household income also decrease income per capita, but less than do non-productive extended members. Model 1, which treats fertility (number of children) as exogenous finds that households with greater numbers of children are not associated with either higher or lower income. However, model 2, which treats fertility as endogenous to household income, finds that households with more children tend to be significantly poorer.

Only model 1 finds that the age of the household head is associated with income; older age is associated with greater income, but there are decreasing marginal returns to old age.

Appendix

Variable definitions

Economic variables:

$$income_{crop} = (output\_consumed_{crop} + output\_sold_{crop}) \times price_{crop} - (agri\_inputs_{crop} + hired\_labour_{crop})$$

where

$$agri\_inputs_{crop} = seeds_{crop} + pesticide_{crop} + fertilizer_{crop}$$

and

$$hired\_labour_{crop} = workers\_hired_{crop} \times days\_worked_{crop} \times reported\_daily\_wage$$

$$income_{per\_capita} = \frac{(\sum_{n=1}^{x} income_{crop\_n} + income_{husbandry} + income_{wage\_work} + income_{other})}{household\_members}$$

where n=1…x denotes the various crops cultivated

$$productivity_{crop} = {output_{kg,crop}} \big/ {area_{crop}}$$

Definition of the household:

Gujarati households are typically comprised of parents, children, and in most cases, extended family members. Extended family members may be grandparents (the primary earner's parents) as well as adult children (the primary earner's children), either unmarried or married. In the case of married adult children, both the married child and their spouse are part of the household, as well as any children they may have. Thus, at its largest, a Gujarati household may include four generations. However, the typical household sampled here was not so large, including an average of only one extended family member per household. Married children usually move out of their parental home to start their own household.

The dependency ratio presented in the poverty profile above is calculated as:

$$dependency\_ratio = \frac{(children_{male} + children_{female}) + (dependents_{total} - dependents_{productive})}{(children_{male} + children_{female}) + dependents_{total} + parents}$$

where parents = 2 in all cases within the sample
and $dependents_{total}$ is the total number of extended members within the household
and $dependents_{productive}$ is the number of extended members who contribute to household income

References

Chauhan, Darmistha, A Study of Savings and Credit Situation in Surendranagar Programme Area, AKRSP (India), Ahmedabad, Gujarat, India, 2003

AKRSP(I), Baseline Studies (2000-2004)

Serumaga-Zake, Philip & Willem Naude, The determinants of rural and urban household poverty in the North West province of South Africa, Development Southern Africa, Vol 19, No.4, October 2002

Dupta, Nabanita Datta & Amaresh Dubey, Poverty and fertility – An instrumental variable analysis on Indian micro data, working paper 03-11, Department of Economics, Aarhus School of Business, unknown date

Bruck, Tilman, Determinants of rural poverty in post-war Mozambique: Evidence from a household survey and implications for government and donor policy, QEH working paper series: QEHWPS67, March 2001

Gang, Ira N. & Kunal Sen, Myeong-Su Yun, Caste, ethnicity and poverty in rural India, Dept. of Economics, Rutgers University, Dept. of Economics, University of East Anglia, Dept. of Economics, Tulane University, March 6, 2005

Deaton, Angus & Margaret Grosh, Chapter 17: Consumption, in Designing household survey questionnaires for developing countries: lessons from ten years of LSMS experience, July 1998

Coudouel, Aline & Jesko S. Hentschel, Quentin T. Wodon, Chapter 1 : Poverty measurement and analysis, from unknown volume, World Bank, date unknown

Bezemer, Dirk & Zvi Lerman, Rural livelihoods in Armenia, Post-communist economies, Vol 16, No. 3, September 2004

Ravallion, Martin, Issues in measuring and modelling poverty, The Economic Journal, Vol. 106, No. 438, pp. 1328-1343, Sep 1996

Lanjouw, Peter & Martin Ravallion, Poverty and household size, The Economic Journal, Vol. 105, No. 433, pp. 1415-1434, Nov 1995

Kofitse, Alexander K.A., The number and spacing of births and household poverty: an Economic analysis, A thesis in agricultural economics and demography, The graduate school, Pennsylvania State University, December 1999

Datt, Gaurav & Martin Ravallion, Is India's economic growth leaving the poor behind?, The journal of economic perspectives, Vol. 16, No.3, pp. 89-108, Summer 2002

Chun, Hyunbae & Jeungil Oh, An instrumental variable estimate of the effect of fertility on the labour force participation of married women, Applied economics letters, vol 9, pp. 631-634, 2002

Sunderji, Aziz, Rural household cash flows: the Mota Sakhpar case, submitted to AKRSP(I) as an independent research paper, August 6, 2004

Sunderi, Aziz, Overview of civil society: India, The Asian Development Bank, forthcoming publication

Deaton, Angus, The analysis of household surveys: A microeconomic approach to development policy, the international bank for reconstruction and development/the world bank, 1997

Assorted National Sample Survey Organization publications, Ministry of Statistics & Programme implementation, Government of India:

Consumer expenditure of farmer households, NSS 59th round, Oct 2005
Household asstet holdings, indebtedness, current borrowings and repayments of social groups in India, NSS 59th round, January-December 2003
Household consumer expenditure in India, NSS 60th round, January-June 2004
Household consumer expenditure and employment-unemployment situation in India, NSS 59th round
Household indebtedness in India as on 30.06.2002, NSS 59th round
Income, expenditure and productive assets of farmer household, NSS 59th round
Level and pattern of consumer expenditure, 2004-2005, NSS 61st round, July 2004-June 2005
Ownership of livestock, cultivation of selected crops and consumption levels, NSS 50th round, July 1993-June 1994
Some aspects of farming, NSS 59th round
Sources of household income in India 1999-2000, NSS 55th round, July 1999-June 2000