Testing for separability in household models with heterogeneous behavior: A mixture model approach

by

Renos Vakis[1], Elisabeth Sadoulet[2], Alain de Janvry[2], and Carlo Cafiero[3]

August 2004

Abstract

Knowing whether a household behaves according to separability or non-separability is needed for the correct modeling of production decisions. We propose a superior test to those found in the literature on separability by using a mixture distribution approach to estimate the probability that a farm household behaves according to non-separability, and test that the determinants of consumption affect production decisions for households categorized as non-separable. With non-separability attributed to labor market constraints, the switcher equation shows that Peruvian farm households that are indigenous and young, with low levels of education, and lack of local employment opportunities are more likely to be constrained on the labor market.

Keywords: Labor, separability, mixture distributions, Peru.
JEL classification number: J43, C16, Q12

[1] The World Bank: rvakis@worldbank.org
[2] University of California at Berkeley: sadoulet@are.berkeley.edu, alain@are.berkeley.edu
[3] Università degli Studi di Napoli "Federico II": cafiero@unina.it

I. Introduction

Market failures affect household behavior. Indeed, an important aspect of the analysis of household behavior consists in observing and explaining the strategies they devise to mitigate the welfare costs market failures impose on them. With market failures pervasive for rural households in developing countries, there are many patterns of behavior that would be incomprehensible were it not for taking into account the effect of these failures. For example, Eswaran and Kotwal (1986) show how labor and credit market failures imply that households with different asset positions relative to labor endowments follow different strategies of labor deployment, and use this in turn to establish social classes. These same market failures imply existence of an inverse relation between land productivity and farm size that establishes the potential gains of redistributive land reforms.

Behavior under perfect markets implies separability between production and consumption decisions (Singh, Squire, and Strauss, 1986): the household can solve recursively first its production problem, and then allocate the full income obtained to consumption choices. By contrast, production and consumption decisions are non-separable when there are market failures: in this case, variables that affect consumption decisions (such as wealth, the household's total family labor endowment, consumer goods prices, and household characteristics affecting consumption) also affect production decisions.

Correct modeling of household production decisions thus requires knowledge of whether a specific household is likely to behave according to separability or non-separability decision rules. For this reason, numerous tests of the separability hypothesis have been made. These tests have been done following various approaches and assumptions, such as whether to use the reduced form or the structural form of a household model, to confinr the test to behavior on one specific market or not, to recognize or not heterogeneity across households, and to claim or not that sample separation into separable and non-separable behavior can be observed. We review these tests below, using these four categories in order to identify the distinguishing features of the approach proposed here. We show that a desirable approach combines the following four features: (i) use a reduced form approach, (ii) do not confine test to behavior on one specific market, (iii) recognize heterogeneity across households, and (iv) use unknown sample separation.

(1) Global tests that do not recognize heterogeneity across households

i) A reduced form approach can be used to test whether variables that affect consumption decisions also affect production decisions, without concern for heterogeneity in the sample of households observed. The test is not specific to failure on one particular market, as failure on any market will induce non-separable behavior. Using this approach, Benjamin (1992) and Bowlus and Sicular (2003) cannot reject separability for households in Java and China, respectively, while Lopez (1984) and Grimard (2000) reject it for households in Canada and Côte d'Ivoire, respectively.

ii) A structural form can be used, where estimation of a production or cost function allows to estimate the marginal productivity (i.e., the shadow price) of an input which can then be compared to the market price. Jacoby (1993) for Peru and Skoufias (1994) for India use this approach to reject separability in labor decisions across the sample of households, with no concern for heterogeneity. The test is not specific to a specific market failure since it does not presume that the constraint is necessarily in the market for the factor analyzed.

Nonetheless, global tests that do not take heterogeneity into account have limited usefulness since it is well known that market failures are largely idiosyncratic.

(2) Idiosyncratic tests that recognize heterogeneity across households

These tests have been pursued through four different approaches. The first two are confined to market failures on specific markets. The latter two tests diagnose failures on any market.

i) Whether a household is constrained or not on a specific market can be asked directly. This has been used by Feder, Lau, Lin, and Luo (1990) to partition Chinese rural households into constrained and un-constrained on the credit market, and to test at the reduced form level that production decisions are affected by liquidity available at the beginning of the season for constrained households, but not for the others. However, tests that account for heterogeneity on the basis of self-declared exposure to a constraint on a specific market may hide unrecognized constraints on other markets.

ii) Structural form tests can be done using a disequilibrium model to estimate the probability of being constrained on a specific market (Maddala, 1983). In this case, there is unknown sample separation into separable and non-separable behavior, and separation is derived from revealed determinants of behavior. Carter and Olinto (2003) thus specify a credit demand and supply function and estimate an idiosyncratic probability of being liquidity constrained. This allows assigning to each household a probability of making investment decisions according to separable or non-separable rules. Difficulties with the approach are: (a) specifying and estimating supply and demand functions for transactions on a specific factor market, and (b) confining the cause of non-separability to being constrained on that specific market, credit in this case. The probability of being constrained on that market is used as the probability of behaving in a non-separable way in all production decisions, which may be erroneous if households that are unconstrained on the credit market happen to be constrained on other markets.

iii) If observed non-participation in a particular market can be used to infer market failure, and this not only in that but possibly also in other markets, then market participation can be used to infer separability. Reduced form tests with regime specific partitions based on observed market participation have been done by Sadoulet, de Janvry, and Benjamin (1998) for Mexico, Carter and Yao (2002) for China, and Dutilly-Diane, Sadoulet, and de Janvry (2004) for Burkina Faso. They verify ex-post that the group of market-participating households indeed behaves according to decision rules in a separable model. Self-sufficient households are, by contrast, verified to behave in a non-separable fashion. Difficulty with the approach is that tests that account for heterogeneity on the basis of observed market participation may hide non-separable behavior due to constraints on that market.

iv) Finally, structural form tests have been used with unknown sample separation and without market failure being specific to one market. This has been done by Lambert and Magnac (1994) for Côte d'Ivoire and by Bhattacharyya and Kumbhakar (1997) for Bengal. They estimate a production function and, from this, an expected factor marginal productivity and its standard error for each household. This allows comparing an estimated idiosyncratic shadow price with an associated confidence interval to an observed effective market price for each household, classifying them ex-post into separable and non-separable groups, where separability is rejected if prices differ. Lambert and Magnac thus find that separability is rejected for 90% of men but only for 50% of women. Difficulties with the approach are in: (a) specifying and estimating a production function for agriculture, a difficult task both conceptually and empirically. The proposed test is highly sensitive to correct specification and estimation of this function; (b) specifying effective prices against which to compare the estimated marginal productivities.

(3) Tests that reveal idiosyncratic non-separability on the basis of observed behavior should consequently be done at the reduced form level. This is what we propose in this paper. Specifically, the test involves assigning to each household a probability of being exposed to market failure and of behaving according to non-separability rules, before testing for the contrasted behavior of the two groups. In the Peruvian case analyzed here, the test consists in analyzing on-farm labor decisions of farm households that participate in the off-farm labor market, as follows:

i) Unknown sample separation estimation is used to attach probabilities to households in the sample as to whether they make production decisions according to separable or non-separable rules. While this is revealed in labor allocation decisions in production, the cause of non-separability is not confined to existence of a constraint on the labor market.

ii) Verify that households that are ex-post classified in the separable category make production decisions that are not affected by variables that affect consumption decisions.

iii) The switcher equation helps reveal the causes of market segmentation between constrained and un-constrained labor market participants among farm households in Peru. Results show that differential skills, youth, ethnicity, as well as lack of regional opportunities are important impediments to unconstrained market participation.

The paper is organized as follows. Section II develops a household model with a labor market constraint and derives a testable hypothesis for separability among market participants. The econometric approach is presented in section III. The Peruvian data and findings are presented in section IV. Section V concludes.

II. Theory

Building on traditional farm household models (see Singh et al., 1986), we consider a farm household whose objective is to maximize utility. Utility is derived from income *y* and leisure time l^l.[1] The household is endowed with a total amount of time *E* to be allocated among on-farm work l^i, off-farm work l^o which will be paid a salary w^o, and leisure.[2] Finally, there exists an unknown upper limit $\overline{L}$ to the amount of labor that can be sold on the market.

The household's problem can be represented as follows:

$$\max_{l^i, l^o, l^l} U\left(y, l^l, z^h\right)$$

subject to:

$$y = pq\left(l^i, A, z^q\right) + w^o l^o;\ l^l = E - l^o - l^i;\ l^l, l^i \geq 0; \text{and } 0 \leq l^o \leq \overline{L},$$

where: *U* is the household's utility function, *y* is total household income, l^l is leisure, z^h is a vector of household characteristics relevant in consumption decisions, *p* is the output price, *q* is the quantity produced, z^q is a vector of farm characteristics relevant in production decisions, l^i and l^o are the amounts of family labor employed on and off-farm, respectively, w^o is the effective wage received by family labor outside the farm, *A* is exogenous farm size (for simplification, we assume no land market), *E* is total family labor endowment, and $\overline{L}$ is the maximum amount of family labor that can find work off-farm. The utility function $U(\cdot)$ is assumed to be increasing and quasi-concave; the production function $q(\cdot)$ is assumed to be increasing and concave.

Since the focus of this study is to explain labor allocation decisions and unobserved heterogeneity of small farmers that participate in the market as net sellers, we concentrate the rest of the analysis on them. Derivation of how regimes emerge is available in Sadoulet et al. (1998).

In absence of other constraints (like, for example, consumption cash constraints or food market constraints) for households that participate in the labor market as net sellers, whether there is separability between production and consumption decisions will depend on whether the maximum off-farm labor constraint is binding or not. The simple observation that a household is selling labor will not be sufficient to infer separability.

One can show that, for net sellers of labor for whom the off-farm labor constraint is not binding, production and consumption decisions are separable. The reduced form equation for l^i is expressed as a function of only "production side characteristics" and not of "consumption side characteristics":

$$l^i = f\left(p, w^o, A, z^q\right). \quad (1)$$

---

[1] For simplicity, we are assuming no imperfections on the commodity markets, which allows to include income directly in the utility function rather than other consumption goods.

[2] Hired labor could be also introduced in this setting. We nonetheless ignore it since our ultimate focus is on net sellers of labor.

In this case, the household will sell in the labor market all the excess labor and the decision price will be the market price w^o.

For net sellers of labor for whom the off-farm labor constraint is binding, i.e., $l^o = \overline{L}$, separability between production and consumption decisions breaks down. The household will sell $\overline{L}$ on the labor market and supply labor on-farm up to the point where the marginal product of labor equates the marginal utility from leisure. The decision price becomes a shadow price lower than w^o.

The reduced form equation for l^i is:

$$l^i = f\left(p, w^o, A, z^q, E, z^h, l^o = \overline{L}\right). \qquad (2)$$

As can be seen, the constrained optimal allocation for on-farm labor l^i also depends on consumption side variables E and z^h, and on the off-farm labor quota $\overline{L}$. Therefore, the separability hypothesis breaks down.

At this point, a brief discussion about the nature of the off-farm labor constraint is warranted. It can be verified that the off-farm labor allocation rule is given by:

$$l^o = \begin{cases} l^{o*}\left(p, w^o, A, z^q, E, z^h\right) \text{ if } l^{o*} < \overline{L} \\ \overline{L} \text{ if } l^{o*} \geq \overline{L} \end{cases},$$

where l^{o*} is the unconstrained level of off-farm labor. Denoting by λ the probability that a household is constrained, we have:

$$\lambda = \lambda\left[l^{o*}\left(p, w^o, A, z^q, E, z^h\right) - \overline{L} \geq 0\right]$$

which in reduced form becomes:

$$\lambda = \lambda\left(p, w^o, A, z^q, E, z^h, \overline{L}\right). \qquad (3)$$

To conclude, we show that under the assumption of no other market imperfections that might introduce non separability (for example, presence of credit constraints or food market imperfections), a labor selling farm-household will determine the amount of labor employed on farm l^i according to one of two alternative regimes, defined by equations (1) and (2), respectively. One empirical implication of this is that, if the researcher had information on households' classifications, a testable hypothesis on the separability assumption could be directly implemented. However, in most cases, this information is unobserved. The next section addresses this issue of ‘unknown sample separation’.

III. Econometrics

The preceding section established that it is conceivable for farm-households to participate in the market as net sellers, and yet to be constrained by unobservable quantity limitations or by transactions costs in their ability to respond to price changes. When the sample division is unknown, the farm labor supply response function of a group of market participating households can be econometrically represented by a switching regression model with unobserved sample separation (Maddala, 1983).

Formally, we can characterize the sample behavior in a three-equation model:

$$l^1 = x_1\beta + u_1$$

$$l^2 = x_2\gamma + u_2$$

$$\lambda^* = x_\lambda \xi + u_\lambda$$

where, using results from the household model,

$$x_1 = \left\{p, w^o, A, z^q\right\}$$

$$x_2 = \left\{p, w^o, A, z^q, E, z^h, \bar{L}\right\}$$

$$x_\lambda = \left\{p, w^o, A, z^q, E, z^h, \bar{L}\right\}$$

β, γ, and ξ are coefficients to be estimated, and

u_j's are normal iid disturbances with zero means and variances σ_j^2 (with $\sigma_\lambda = 1$ for identification purposes).

l^1, l^2, and λ^* are latent unobserved variables. Instead, for each observation, we observe variable l^i defined by:

$$l^i = \begin{cases} l^1 & \text{if } \lambda^* < 0 \\ l^2 & \text{if } \lambda^* \geq 0. \end{cases}$$

The problem is that of estimating the parameters $\left\{\beta, \gamma, \xi, \sigma_1, \sigma_2\right\}$ from the sample of *N* observations on $\left\{l_k^i, x_{k1}, x_{k2}, x_{k\lambda}\right\}$, $k = 1, ..., N$. Given that we cannot identify a priori the regime participation, a randomly selected observation l_k^i (household *k*'s on-farm labor supply) will have probability $1 - \lambda = \Phi(-x_{k\lambda}\xi)$ of belonging to the first regime, and probability λ of belonging to the second. The probability density function of observation l_k^i is thus:

$$f\left(l_k^i\right) = (1-\lambda)\varphi_1\left(l_k^i - x_{k1}\beta\right) + \lambda\varphi_2\left(l_k^i - x_{k2}\gamma\right) \tag{4}$$

where φ_1 and φ_2 are the probability density functions of u_1 and u_2, that is, the *mixture* of two distributions.

The likelihood function for a sample of *N* observations is then:

$$L(\beta, \gamma, \xi, \sigma_1, \sigma_2) = \prod_{k=1}^{N} f\left(l_k^i\right), \tag{5}$$

A natural way of estimating the parameters would be to maximize (5) with respect to $\{\beta,\gamma,\xi,\sigma_1,\sigma_2\}$.

Estimation of the maximum likelihood is obtained via the E-M method (see Hartley (1978) and Dempster, Laird, and Rubin (1977)), after having used other pre-estimation techniques to decide on the initial values of the parameters (Kiefer, 1978).

The steps of the E-M algorithm are as follows: using starting values for $\beta,\gamma,\sigma_1,\sigma_2$, we first obtain estimates of the classification vector λ (the E step). The starting values for $\beta,\gamma,\sigma_1,\sigma_2$ can be set equal to the estimated values for the pooled sample regression, the rational for which is that, if the observations were truly coming from the same population, those would be the values that maximize the likelihood function. Using the estimate for λ to weight the probability density functions of each observation in the two regimes, we obtain estimates for $\beta,\gamma,\sigma_1,\sigma_2$ (the M step). This iterative procedure is repeated until it converges.

IV. Data and results

Data

The data come from the 1997 Peruvian LSMS (World Bank, 2004). The survey was conducted on 4,500 households. From these, 1,131 allocate work to both on and off-farm activities and are used in the analysis. We postpone a discussion on descriptive statistics for now as it will be more relevant to present them after the estimation. We offer instead Figure 1, which plots the frequency distribution of individual off-farm hours worked and seems to suggest the existence of two distinct subpopulations, one in which individuals work less off-farm and another where they work more. What the econometric procedure attempts to do is to explore, among other things, this heterogeneity and characterize the two regimes. We superimpose a normal distribution to hint on the implications of assuming one homogeneous population.

Separability

According to a fully separable model, the decision about on-farm labor allocation should be purely a production decision, and thus household characteristics (such as $E,z^h,\overline{L}$) should not affect it. Our theoretical model postulates the possibility of two different regimes among farm-households in the sample. It also predicts that household characteristics should only affect the constrained regime. We, therefore, specify one of the two regimes by including the total family labor endowment (*E*, the number of adults), household characteristics in consumption z^h such as ethnicity and the number of boys and girls, and the constrained family off-farm employment l^o. We specify the second regime by only including production side characteristics z^q among household characteristics. We also include any variables $z^{q,h}$ that can be argued to affect both production and consumption decisions.

Apart from this restriction that comes directly from the theoretical model, we do not impose any other restrictions on the parameters of the two models in the belief that, if a dichotomy actually exists, it should be strong enough to let the econometric technique separate the two regimes. We apply the maximum likelihood procedure described in the previous section to determine the best way of dividing the observations into two groups. The dependent variable is the amount of hours allocated by the household to farm activities.[3]

Table 1 presents the results. The first column contains the results of an OLS estimation of the model on the pooled sample.[4] As expected, production characteristics such as human capital assets (farming experience) and land assets significantly affect on-farm labor allocation. In addition, these pooled estimates show a significant effect of household characteristics in consumption (number of adults and of boys and girls) on the on-farm labor use decision. This alone provides some indication that the issue of non-separability is important.

Columns 2-4 present the results from the mixture model. On average, a household has a probability of 0.51 of being constrained in its labor market participation. As expected, for those households in the constrained regime, we find that not only production characteristics affect their on-farm labor allocation but consumption ones as well, thus rejecting the separability hypothesis. In particular, we find that the numbers of children and of adults significantly increase on-farm labor allocation. Other variables that can be considered as household characteristics in both production and consumption also affect on-farm labor allocation. For example, education of the household head increases on-farm work. In addition, regional dummies that capture idiosyncratic effects and shocks of specific geographic areas affect on-farm labor decisions. Households that live in the Coast, Sierra, and Rainforest regions work more on-farm compared to those in Lima. This is expected as off-farm employment opportunities are much more abundant in Lima than in rural areas.[5]

For the second group of households allocated by the mixture model to the unconstrained regime, we find that production characteristics such as land and cattle ownership positively affect on-farm labor allocation. In addition, transaction cost in the form of time to get to the main market, negatively affects work on-farm.

Robustness

The analysis above would not be correct if any of the consumption side characteristics affect the on-farm labor decision for what we specify as the unconstrained regime. For this, we implement a counterfactual estimation including in the specification of the unconstrained regime all of the consumption characteristics used in the constrained one. The aim is to test whether any of these consumption variables have a significant impact on on-farm labor allocations for the unconstrained. The sample is divided into

[3] Remember that we do not consider hired labor in this analysis since we are only looking at households that are net sellers of labor.
[4] The coefficients of this OLS estimation were also used as starting values for the likelihood maximization routine.
[5] We also estimated these models using district dummies and obtained similar results.

constrained and unconstrained households. Following Hartley (1978), we assign observation $\left\{l_k^i, x_{k1}, x_{k2}, x_{k\lambda}\right\}$ to regime 1 if $E\left[\lambda_i^* \middle| l_k^i, x_{k1}, x_{k2}, x_{k\lambda}\right] < 0$, and to regime 2 otherwise. Note that this assignation uses all the information that is ex-post observed on the households, including their behavior on the labor market, rather than simply the predictors $x_{k\lambda}$ of the choice of regime. This conditional expectation is computed as:

$$
\begin{aligned}
E\left[\lambda^* \middle| l^i, x_1, x_2, x_\lambda\right] &= x_\lambda \xi + E\left[u_\lambda \middle| l^i, x_1, x_2, x_\lambda\right] \\
&= x_\lambda \xi + E\left[u_\lambda \middle| regime\,1\right] \Pr\left[regime\,1 \middle| l^i, x_1, x_2, x_\lambda\right] + E\left[u_\lambda \middle| regime\,2\right] \Pr\left[regime\,2 \middle| l^i, x_1, x_2, x_\lambda\right] \\
&= x_\lambda \xi + E\left[u_\lambda \middle| u_\lambda < -x_\lambda \xi\right] \omega_1\left(l^i\right) + E\left[u_\lambda \middle| u_\lambda \geq -x_\lambda \xi\right] \omega_2\left(l^i\right) \\
&= x_\lambda \xi - \omega_1\left(l^i\right) \frac{\varphi(-x_\lambda \xi)}{\Phi(-x_\lambda \xi)} + \omega_2\left(l^i\right) \frac{\varphi(-x_\lambda \xi)}{1 - \Phi(-x_\lambda \xi)}
\end{aligned}
$$

where $\omega_1\left(l^i\right)$ and $\omega_2\left(l^i\right)$ are the weights:

$$
\omega_1\left(l^i\right) = \frac{(1-\lambda)\varphi_1\left(l_k^i - x_{k1}\beta\right)}{f\left(l_k^i\right)} \quad \text{and} \quad \omega_2\left(l^i\right) = \frac{\lambda \varphi_2\left(l_k^i - x_{k2}\beta\right)}{f\left(l_k^i\right)}.
$$

This classifies 527 households, or almost half of the sample, as constrained.

In Table 2, we estimate the ex-post on-farm labor regressions for the two groups, assuming that ex-ante knowledge of the regimes was known. The results support the separability hypothesis for the unconstrained regime. Specifically, none of the consumption characteristics $\left(E, z^h, l^o\right)$ have a significant impact on the on-farm labor decisions of unconstrained households. By contrast, productive assets (land and cattle owned and years of farming experience) increase the level of on-farm work.

To summarize, applying the mixture model approach allowed us to separate the sample into two sub-populations of labor market participants. We find strong evidence that, in one of them, the separability hypothesis between production and consumption decisions is rejected, while for the other it is not. A counterfactual test strengthens our findings by showing that the households in the unconstrained regime behave in a separable fashion.

Understanding the labor constraint

The results above signal the presence of a group of farm households that, albeit participating in the labor market, are making their decisions on farm activity according to a non-separable model of behavior. This is, of course, not enough to conclude that the cause of non separability is the presence of a quantity constraint on the labor market of the kind we described in presenting the theoretical model. The last column of Table 1 contains the estimated coefficients for the equation that determines group separation. Given specification of the two regimes, we can interpret this switcher equation as representing the probability of being in the constrained regime. We use equation (3) from the theoretical model to specify the determinants of group separation.

Interesting patterns emerge regarding market integration and participation. A larger farm lowers the probability of being constrained as it increases labor demand for on-farm labor. In terms of household characteristics, larger numbers of adult members (*E*) and of girls in the household increase the probability of being constrained. Households with higher levels of education, with less years of farming experience, and with an older head are less likely to be constrained. Finally, indigenous households are more likely to be constrained, which could be reflecting greater difficulties for them in participating to labor markets.

At the regional level, households living in all the regions, as opposed to Lima, have a higher probability of being constrained as their off-farm opportunities are more limited. Greater availability of private jobs in the community is also effective in reducing the probability of being constrained in off-farm employment.

We can use the ex-post predicted groups of "constrained" and "unconstrained" households to get insights on the structure of the labor market constraint. From the theoretical framework, we know that the constrained households' off-farm labor supply l^o is also their binding off-farm labor allocation $\bar{L}$. This distribution is plotted in Figure 2. We would expect that, if there were a common market-based barrier on the amount of labor to be allocated, it would show up as a very narrow distribution for $l^o = \bar{L}$ for those households in the constrained regime. The fact that this is not true suggests that household idiosyncracies may be more important in terms of accessing the labor market. Figure 2 also plots the distribution of individual off-farm hours for the unconstrained. These individuals work more overall than constrained ones, as seen by the fact that the off-farm hours distribution is shifted to the right. This is consistent with our findings above.

Finally, a descriptive comparison between "constrained" and "unconstrained" households is presented in Table 3. Overall, constrained households are quite different from unconstrained households. They work almost three times as much on their own farm, although they have almost three times less land. They have lower levels of education, the key asset to gain access to off farm work. They are more often indigenous, they are poorer in terms of per capita consumption, income, and wealth (land owned), and have larger households to support with more adults and girls. They are located more in the Sierra, and definitely less in Lima. They are also more dedicated to farming, with more cattle owned and more years of farming experience. These observations suggest that being constrained on the labor market is closely related to lack of employment opportunities and to lack of control over the essential human and physical assets needed to successfully participate in the labor market.

V. Concluding remarks

We use mixture distribution techniques to develop a more accurate test of the separability hypothesis in household models than previously available. The approach has the advantage of using a reduced form estimate, detecting non-separability on all markets at once, and allowing for heterogeneity in separability behavior across households. It avoids reliance on labor market participation to conclude separability between production and consumption decisions of farm-households. Findings clearly show the existence of two distinct types of households among net-sellers of labor: those behaving as if unconstrained and hence in a separable way, and those that behave as if constrained and hence in a non-separable fashion.

In addition, the econometric technique helps understand the origins of market segmentation. In doing this, we look at the role of both demand and supply side effects on labor constraints. In the case of Peru, low educational attainments, youth, ethnicity, lower availability of job opportunities in the community, and handicaps associated with residence especially in the Sierra are important constraints on labor market participation.

References

Benjamin, Dwayne. 1992. "Household Composition, Labor Markets, and Labor Demand: Testing for Separation in Agricultural Household Models." *Econometrica,* 60(2): 287-322.

Bhattacharyya, Anjana, and Subal Kumbhakar. 1997. "Market Imperfections and Output Loss in the Presence of Expenditure Constraint: A Generalized Shadow Price Approach." *American Journal of Agricultural Economics,* 79(3): 860-71.

Bowlus, Audra, and Terry Sicular. 2003. "Moving Toward Markets? Labor Allocation in Rural China," *Journal of Development Economics*, 71(2): 561-583.

Carter, Michael, and Pedro Olinto. 2003. "Getting Institutions 'Right' for Whom? Credit Constraints and the Impact of Property Rights on the Quantity and Composition of Investment." *American Journal of Agricultural Economics* 85(1): 173-86.

Carter, Michael, and Yang Yao. 2002. "Local versus Global Separability in Agricultural Household Models: The Factor Price equalization Effect of Land Transfer Rights." *American Journal of Agricultural Economics* 84(3): 702-15.

Dempster, A., Nan Laird, and D. Rubin. 1977. "Maximum Likelihood from Incomplete Data Via the E-M Algorithm." *Journal of the Royal Statistical Society* 39: 1-38.

Dutilly-Diane, Céline, Sadoulet, Elisabeth, and Alain de Janvry. 2004. "Household Behavior under Market Failures: How Improved Natural Resource Management in Agriculture Promotes the Livestock Economy in the Sahel". Forthcoming in *Journal of African Economies.*

Eswaran, M., and Kotwal, A. 1986. "Access to Capital and Agrarian Production Organization." *Economic Journal,* 96: 482-498.

Feder, G., Lau, L. J., Lin, J. Y., and Luo, X. 1990. "The Relationship between Credit and Productivity in Chinese Agriculture: A Microeconomic Model of Disequilibrium." *American Journal of Agricultural Economics,* 72(4): 1151-1157.

Grimard, Franque. 2000. "Rural Labor Markets, Household Composition, and Rainfall in Côte d'Ivoire." *Review of Development Economics,* 4(1): 70-86.

Hartley, Michael. 1978. "Estimating Mixtures of Normal Distributions and Switching Regressions." *Journal of the American Statistical Association* 73(364): 747-48.

Jacoby, Hanan. 1993. "Shadow Wages and Peasant Family Labour Supply: An Econometric Application to the Peruvian Sierra." *Review of Economic Studies,* 60: 903-921.

Kiefer, Nicholas. 1978. "Discrete Parameter Variation: Efficient Estimation of a Switching Regression Model." *Econometrica* 46(2): 427-34.

Lambert, Sylvie, and Thierry Magnac. 1994. "Measurement of Implicit Prices of Family Labour in Agriculture: An Application to Côte d'Ivoire." In Caillavet, Gyomard, and Lifran (eds.), *Agricultural Household Modeling and Family Economics*. Amsterdam: Elsevier.

Lopez, Ramón. 1984. "Estimating Labour Supply and Production Decisions of Self-Employed Farm Producers." *European Economic Review,* 24: 61-82.

Maddala, G.S. 1983. *Limited-Dependent and Qualitative Variables in Econometrics*. Cambridge: Cambridge University Press.

Sadoulet, Elisabeth, Alain de Janvry, and Catherine Benjamin. 1998. "Household Behavior with Imperfect Labor Market." *Industrial Relations,* 37(1): 85-108.

Singh, I., Lynn Squire, and John Strauss, (eds.). 1986. *Agricultural Household Models*. Baltimore. MD: The Johns Hopkins University Press.

Skoufias, Emmanuel. 1994. "Using Shadow Wages to Estimate Labor Supply of Agricultural Households." *American Journal of Agricultural Economics,* 76(2): 215–227.

World Bank. 2004. Living Standards Measurement Study of the World Bank. http://www.worldbank.org/lsms/

Table 1. On-farm labor allocation: Pooled versus mixture model

		Pooled	Mixture		
			Constrained	Unconstrained	Switcher
Market wage (soles)	w^o	0.02	-0.11***	0.01	0.32**
Time to market (min.)	p	0.41	0.73	-1.14**	-9.1
Land owned (ha.)	A	0.14**	0.21	0.21***	-10.8***
Cattle owned (number)	z^q	0.49	0.29	0.65***	-2.23
Farming experience (years)	z^q	2.1***	1.36***	0.18**	28.5***
Coast (dummy)	z^q	17.8**	19.9*	2.87	214***
Sierra (dummy)	$z^{q,h}$	17.1**	21.9**	-3.91	216***
Rainforest (dummy)	$z^{q,h}$	18.4**	12.9	4.92	202***
Household head education (years)	$z^{q,h}$	0.54	2.4**	0.55*	-31.2***
Household head sex (male=1)	$z^{q,h}$	0.24	0.33	4.09	-10.9
Household head age (years)	$z^{q,h}$	1.42	5.16***	0.04	-41.6***
Household head age squared	$z^{q,h}$	-0.01	-0.05***	0.00	0.46***
Adults (number)	E	18.9***	16.1***		254***
Boys (number)	z^h	9.7***	21.8***		-46**
Girls (number)	z^h	23.2***	38.8***		53***
Ethnicity (indigenous=1)	z^h	6.8	-9.4		315***
Availability of private jobs in community (proportion of employed)	z^{lo}				-1428***
Hours worked off-farm	l^o	-0.13***	-0.19***		
Constant		25.6	-2.1	65***	1564***
Sample size: 1131					
Sample proportion	$\hat{\lambda}$	1.0	0.51	0.49	1.0
R^2 (pooled)	0.28				
Log likelihood (mixture)	-6440				

All coefficients of the switcher regression are multiplied by 1000.
Dependent variable: household's on-farm work (hours).
Switcher: probability of being constrained.
The missing dummy for regions is Lima.
Significance level: * 10%; ** 5%; *** 1%

Table 2. Counterfactual test with ex-ante separation

		OLS with sample separation "known" (predicted ex-post):	
		Constrained	Unconstrained
Market wage (soles)	w^o	-0.12***	0.004
Time to market (min.)	p	1.48	-1.14*
Land owned (ha.)	A	0.24	0.22***
Cattle owned (number)	z^q	0.33	0.60***
Farming experience (years)	z^q	1.03***	0.23**
Coast (dummy)	z^q	18.9	3.60
Sierra (dummy)	$z^{q,h}$	18.5	-1.98
Rainforest (dummy)	$z^{q,h}$	14.7	5.43*
Household head education (years)	$z^{q,h}$	2.27	0.41
Household head sex (male=1)	$z^{q,h}$	3.54	6.41**
Household head age (years)	$z^{q,h}$	6.02***	-0.18
Household head age squared	$z^{q,h}$	-0.07***	0.002
Adults (number)	E	9.84**	-0.17
Boys (number)	z^h	21.4***	-1.15
Girls (number)	z^h	38.3***	-1.30
Ethnicity (indigenous=1)	z^h	-15.8	-4.7
Hours worked off-farm	l^o	-0.16***	0.01
Constant		20.3	68.6***
Sample size		527	604
R^2		0.29	0.22

Dependent variable: household's on-farm work (hours).
The missing dummy for regions is Lima..
Significance levels:* 10%; ** 5%; *** 1%.

Table 3. Selected household descriptive statistics by predicted constrained regime

		Constrained	Unconstrained
Market wage (soles)	w^o	235.6	250.6*
Time to market (min.)	p	1.04	1.05
Land owned (ha.)	A	0.72	2.97
Cattle owned (number)	z^q	2.80	1.01*
Farming experience (years)	z^q	22.6	8.1*
Coast (%)	$z^{q,h}$	0.24	0.24
Sierra (%)	$z^{q,h}$	0.30	0.21*
Rainforest (%)	$z^{q,h}$	0.19	0.18
Lima (%)	$z^{q,h}$	0.28	0.37*
Household head education (years)	$z^{q,h}$	9.26	10.97*
Household head sex (male=1)	$z^{q,h}$	89.1	86.0
Household head age (years)	$z^{q,h}$	52.1	48.0*
Adults (number)	E	3.8	3.0*
Boys (number)	z^h	0.87	0.84
Girls (number)	z^h	0.87	0.71*
Ethnicity (indigenous %)	z^h	23.0	13.2*
Availability of private jobs in community (%)	z^{lo}	98.0	99.2*
Total hours worked on-farm (last week)	l^i	222	78*
Total hours worked off-farm (last week)	l^o	126.0	126.6
Income per capita (soles)	y	1401	1541*
Consumption per capita (soles)	C	2668	3478*
Sample size: 1131		527	604

Note: * means that there is a significant difference between the unconstrained and constrained groups at the 10% level or less.

Figure 1: Off-farm hours worked per individual

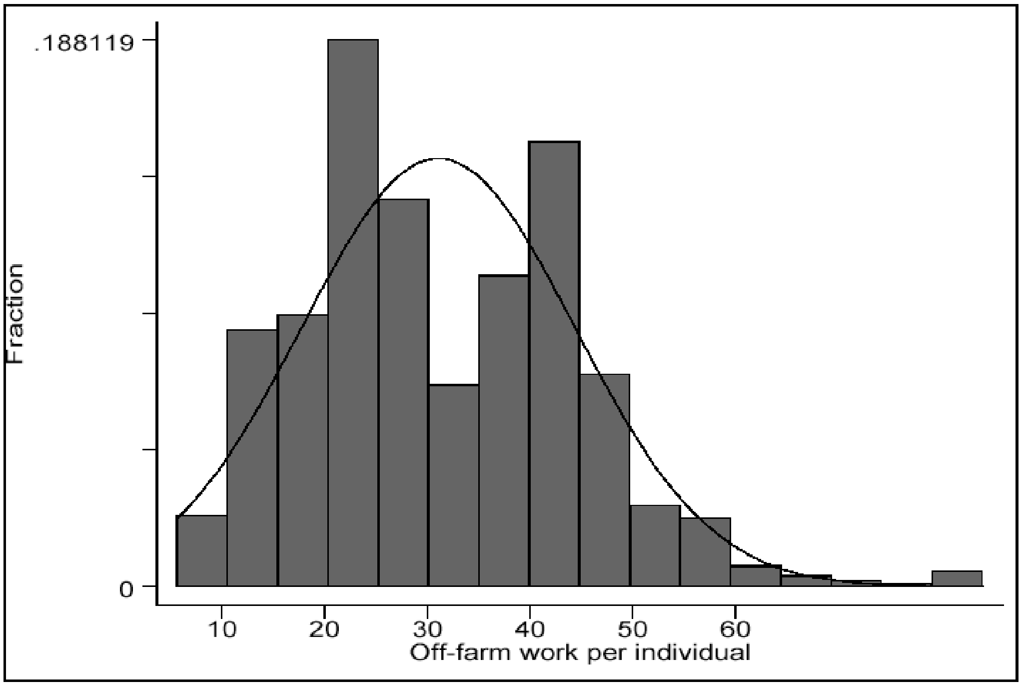

Figure 2: Kernel densities of off-farm hours worked per individual