The impact of market failures on household behavior: explaining labor market segmentation, technology adoption patterns and transaction costs in rural Peru

by

Renos Nicos Vakis

B.A. (University of California, Davis) 1994
M.S. (University of California, Berkeley) 1997

A dissertation submitted in partial satisfaction of
the requirements for the degree of
Doctor of Philosophy

in

Agricultural and Resource Economics

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, BERKELEY

Committee in charge:
Professor Alain de Janvry, Co-Chair
Professor Elisabeth Sadoulet, Co-Chair
Professor David Zilberman
Professor David Card

Fall, 2002

UMI Number: 3082440

UMI Microform 3082440

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

The impact of market failures on household behavior: explaining labor market segmentation, technology adoption patterns and transaction costs in rural Peru

by

Renos Nicos Vakis

Abstract

The impact of market failures on household behavior: explaining labor market segmentation, technology adoption patterns and transaction costs in rural Peru

by

Renos Nicos Vakis

Doctor of Philosophy in Agricultural and Resource Economics

University of California, Berkeley

Professor Alain de Janvry, Co-chair
Professor Elisabeth Sadoulet, Co-chair

The existing economic literature on transaction costs and market failures illustrates how they alter economic behavior by reducing market participation and affecting the response to policies of heterogeneous populations. Still, empirical work has often been inadequate and unable to capture the non-trivial intricacies and complications that arise by introducing them in statistical analyses.

This dissertation, divided into three chapters, addresses this gap. The first chapter proposes a theoretical model of farm households that participate in the labor market. The aim is to identify those households that are potentially constrained in their off-farm labor allocation and those that are not. Using mixture distribution estimation techniques we find two regimes among market-participating households depending on whether the off-farm labor allocation constraint is binding or not. The results also suggest that labor markets in Peru are segmented and that aspects like ethnicity, gender, as well as regional attributes such as the level of unemployment or low population density can prevent market integration.

The second chapter examines the relationship between credit market failures, technology adoption, and income portfolio diversification. Using a semi-structural approach, we show that income diversification can complement the adoption of new technologies by relaxing cash liquidity constraints. This finding is the reverse of the existing literature on adoption, which posits that changes in the agricultural product mix are a result of adoption of new technologies.

Finally, the third chapter attempts to measure the magnitude and role of different types of transaction costs on behavior. A market search model that incorporates both variable and fixed transaction costs is developed to understand how farmers choose where to sell their marketed surplus. The empirical findings show that, in addition to distances and access to good road infrastructure, a number of other transaction costs attributes such as information about markets and prices, relationships with buyers, as well as bargaining abilities also affect market choices. As such, these results suggest that policies aiming at reducing transaction costs should address not only road and infrastructure but also create mechanisms to enhance information flows and bargaining.

Alain de Janvry	8/4/2002
Professor Alain de Janvry, Co-chair	Date
E Sadoulet	8-4-2002
Professor Elisabeth Sadoulet, Co-chair	Date

Contents

List of Figures

List of Tables

Acknowledgments

Despite what they say, a dissertation is NOT the result of a person's brilliant efforts to develop a new idea. Instead, it is the non-linear sum of the interaction and collaboration between the author and advisors, teachers, colleagues, friends and family. Reaching this point where I am left to write these acknowledgments in order to complete this dissertation makes me realize how much I have grown and learned over the last few years here in Berkeley from all the people that have shared my life, each one of them providing me with their unique blend of inspiration, knowledge and support. While the list is endless, there are a few people that deserve eponymous acknowledgement.

Since my first year in the program, Alain de Janvry and Elisabeth Sadoulet (aka the "Jefes") have been role models for me and their incredible support and guidance has been instrumental for my development as an economist. David Zilberman inspired and helped me explore new ideas and think beyond the box. My other thesis and orals committee members, David Card, Jeff LaFrance and Michael Ward all gave me useful insights at various stages of my work. Among my colleagues, Carlo Cafiero really served as a catalyst for me to begin my journey in economics. He also taught me how to make gnocchi, great pizza and fix refrigerators (and the list goes on). Erin (McCormick) Godtland and Melanie Raymond are jewels as friends, and along with Craig McIntosh are all guilty for fixing all of my mistakes during our endless dissertation sessions in Odyssia. My dear friends Chris Costello and Hayley Chouinard always motivated me to work hard and cheered me up with their optimism and love (yes you are both invited!). Rinku Murgai allowed me to begin learning about econometrics and empirical work by borrowing my hands! While Gail Vawter did not help with my empirical work, she was partly responsible for me gaining weight through her generous provisions of chocolate (all other blame can go to Zachary's pizza). She also made sure that I kept up with all deadlines and requirements in the program. Javier Escobal and my friends at GRADE in Lima were really wonderful and through their hospitality and collaboration I managed to better understand Peru. Many other friends in the department and in Berkeley in general have offered me their love, friendship, advice and support. You know who you are. My parents, Nicos and Olga Vakis back in Cyprus have been a constant source of energy and love. I would not be here without their unconditional support. Finally, this experience would have never been as exciting and fulfilling without the help, support, love and wisdom of my wife Caridad Araujo. I promise not to talk about transaction costs with you again, even though I cannot say the same about practicing my magic!

Thank you all for everything. Now back to work...

Chapter 1

Searching for Failures in the Peruvian Labor Market via Mixture Models

1.1 Introduction

That institutional failures have important effects in economic behavior has been explored in both theoretical and empirical literature. The existing theoretical literature shows that market failures significantly alter economic behavior; they impede market participation and can affect the response to policies of heterogeneous populations. For example, Roemer [50] analyzes class formation based on differences in assets position. Building on Roemer's work, Eswaran and Kotwal [15] derive a model where differences in asset position and the presence of two market failures (unobserved effort and a non-uniform access to credit) determine the emergence of different market participation regimes.

The implications of these models are multiple. First, they illustrate how social classes or regimes emerge through rational choices. In addition, these choices can help explain efficiency, like the existence of the inverse relation between yields and farm size

(Sadoulet et al. [51]). In the context of household economics, the emergence of different regimes has also important policy consequences as they imply differential responses to policy interventions. For example, Sadoulet et al. [51] look at labor markets and show that the presence of a price band for wage, i.e. a positive difference between the wage to pay for hired labor and the effective wage received by family members when working outside, determines endogenous selection of households in labor market participation. Such a difference, induced by fixed transaction costs, determines the appearance of a finite set of self-sufficient farm-households. As a result, market participating households use the market wage as the decision price, while self-sufficient households use an endogenously determined shadow price that also depends on consumption and thus breaking the separability assumption between consumption and production decisions [1].

Empirical literature that looks at the separability hypothesis has focused on developing ways to incorporate and test it. Lopez [39], Jacoby [31], Lambert and Magnac [37] and recently Barrett et al. [2] use a direct approach where they estimate an aggregate production function that is then used to get implicit prices of family labor. These prices are then compared and found different with the market wage, thus rejecting separability. Skoufias [54] uses Jacoby's methodology to find similar results with an ICRISAT data from India. Using demographic variables to explain production decisions, Benjamin [4] finds that they do not affect the demand for pre-harvest labor in rural Java and thus does not rejects separability.

Another approach is to get information directly from the data and to then use

[1]The separability hyporthesis posits that consumption and production decisions are recursive within the household. First, the households maximizes profits and then allocates its full income for consumption (Singh et al.[53]).

it to test separability. Feder et al. [24] use survey questions on access to credit to divide their sample between credit constrained and unconstrained. They find that consumption decisions enter the production decision for constrained households only and use this result to infer non-separability. Finally, Sadoulet's et al. [51] find selective recursiveness for seller and employer regimes but not for self-sufficient households in rural Mexican data.

While these papers make advances in both the theoretical and empirical aspects of separability and market failures, some questions remain unanswered. This paper addresses some of these issues and offers a number of refinements of the previous literature. Using labor markets, we question whether the common conclusion that market failures divide people between market participants and those who choose not to participate is enough. We argue that even among market participants, unobserved heterogeneity (either at the individual or a regional level) can affect the separability assumption. Our attempt is that of explicitly allowing for the fact market participation *per se* does not imply belonging to a particular regime. In particular, the presence of a quantity constraint on the labor market makes it plausible for some farmers to participate in the market and yet be constrained; that is, among those who participate in the market, there are farms that refer to the endogenously determined *shadow wage* as the decision price. These farms are not expected to react to an infinitesimal change in the market price. The practical difficulties of implementing such a model comes from the fact that, usually, the constraints are not observable. Our econometric methodology allows us to identify between constrained and unconstrained households.

In addition, while past work has provided a number of ways to test separability, little has been done to address the source of the market imperfection. The effect of policies does not depend on the knowledge of the existence of market distortions but understanding their source. In this context, our econometric approach allows us to explore and assess market integration and market distortions in more detail. The findings suggest that labor markets in Peru are very segmented and that ethnic and gender discrimination impedes market integration. In addition, on the supply side, the labor market also plays an important role in determining households' equilibria. Regional infrastructure, off-farm opportunities and geographic location affect the probability of finding off-farm work. Policy makers can use such results to target and implement policies that reduce labor market failures and enable households to better anticipate and protect themselves. For example, while our results show that policy schemes that address discrimination and education are important in the Peruvian context, additional programs that enhance off-farm job opportunities are also necessary to complement the labor market development process and ameliorate market participation by marginal groups.

Section 1.2 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 1.3. The Peruvian data and findings are presented in section 1.4. Section 1.5 concludes.

1.2 Theory

Building on traditional farm-household models (see Singh et al.[53]), let us consider a farm–household whose objective is to maximize utility. Utility is derived from income,

y, and leisure time l^l [2]. The household is endowed with a total amount E of time to be allocated among in–farm work, l^i, off–farm work, l^o, which will be paid a salary w^o, and leisure. Farm labor can be supplied by household members or can be hired on the market, h, at a given wage rate w^h. Finally, there exist an unknown, upper limit to the amount of labor that can be sold on the market, $\bar{L}$.

The household's problem can be represented as follows:

$$\max_{h,l^i,l^o,l^l} U(y, l^l, \mathbf{z^h}) \tag{1.1}$$

subject to:

$$y = pq(l^i + h, A, \mathbf{z^q}) - w^h h + w^o l^o \tag{1.2}$$

$$l^l = E - l^o - l^i \tag{1.3}$$

$$l^l \geq 0 \tag{1.4}$$

$$h \geq 0 \tag{1.5}$$

$$l^i \geq 0 \tag{1.6}$$

$$l^o \geq 0 \tag{1.7}$$

$$\bar{L} \geq l^o \tag{1.8}$$

where:

U is the household's utility function,

y is total household income,

l^l is leisure,

[2]For simplicity, we are assuming no imperfections on the commodity markets, so that it makes sense to include directly income in the utility function, rather than other consumption goods.

$\mathbf{z^h}$ is a vector of *household* characteristics relevant in *consumption* decisions,

p is the output price,

q is the quantity produced,

h is the amount of hired labor,

$\mathbf{z^q}$ is a vector of *farm* characteristics relevant in *production* decisions,

w^h is the effective price to be paid for hired labor,

l^i and l^o are the amounts of family labor employed in and off–farm respectively,

w^o is the effective wage received by family labor outside the farm,

A is farm size,

E is the total family labor endowment, and

$\bar{L}$ is the maximum amount of family labor that can find work off-farm.

The utility function $U(\cdot)$ is assumed to be increasing and quasiconcave; the production function $q(\cdot)$ is assumed to be increasing and concave.

Substituting from (1.2) and (1.3) for y and l^l the objective function can be written as:

$$\max_{h,l^i,l^o} U(pq(l^i + h, A, \mathbf{z^q}) - w^h h + w^o l^o, E - l^i - l^o, \mathbf{z^h}) \tag{1.9}$$

subject to the constraints given by equations (1.5) through (1.8).

The first order Kuhn Tucker conditions that describe this problem are:

$$\frac{\partial U(\cdot)}{\partial h} : [U_1(\cdot)[pq_L(\cdot) - w^h + \mu^h]]h = 0,\ U_1(\cdot)[pq_L(\cdot) - w^h + \mu^h] \leqq 0,\ h,\ \mu^h \geq 0 \tag{1.10}$$

$$\frac{\partial U(\cdot)}{\partial l^i} : [U_1(\cdot)pq_L(\cdot) - U_2(\cdot) + \mu^i]l^i = 0,\ U_1(\cdot)pq_L(\cdot) - U_2(\cdot) + \mu^i \leqq 0,\ l^i,\ \mu^i \geq 0 \tag{1.11}$$

$$\frac{\partial U(\cdot)}{\partial l^o} : [U_1(\cdot)w^o - U_2(\cdot) + \mu^o - \mu^{\bar{L}}]l^o = 0,\ U_1(\cdot)w^o - U_2(\cdot) + \mu^o - \mu^{\bar{L}} \leqq 0 \tag{1.12}$$

with l^o, $\bar{L}$, μ^o, $\mu^{\bar{L}} \geq 0$, $\bar{L} - l^o \geq 0$.

In addition:

$U_1(\cdot) = U_1(p, w^o, w^h, l^i, h, A, \mathbf{z^q}, E, l^o, \mathbf{z^h})$ and

$U_2(\cdot) = U_2(p, w^o, w^h, l^i, h, A, \mathbf{z^q}, E, l^o, \mathbf{z^h})$

are the marginal utilities of income and leisure respectively,

$q_L(\cdot) = q_L(l^i + h, A, \mathbf{z^q})$ is the marginal productivity of labor and

μ^t, $t = h, i, o, \bar{L}$ are the multipliers associated with the non-negativity constraints.

In this simplified setting and assuming that $w^h > w^o$, that is the price to pay per unit of hired labor is higher than the price received per unit of family labor sold outside, the relative size between farm size A and the total family labor endowment E will determine in which of four possible alternative regimes the farm-household will optimally operate: workers, net sellers, net buyers and self-sufficient in labor. Since the focus of this study is to explain labor allocation decisions and unobserved heterogeneity of small farmer that participate in the market as net sellers, we concentrate the rest of the analysis on them. Appendix A summarizes how these regimes emerge. For a more complete description of this classification process, we refer the reader to earlier work by Roemer [50], Eswaran and Kotwal [15] and Sadoulet et al. [51].

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

Proposition 1: For net sellers of labor for whom the off-farm labor constraint is not binding, production and consumption decisions are separable.

To see this, we analyze the Kuhn Tucker first order conditions for net sellers. For these households $h = 0$, $l^i > 0$, $l^o > 0$, $\mu^h = \mu^i = \mu^o = 0$. If, in addition, the off-farm labor constraint (1.8) is not binding, we also have $l^o < \bar{L}$ and $\mu^{\bar{L}} = 0$ The Kuhn Tucker conditions reduce to:

$$\begin{cases} \frac{\partial U(\cdot)}{\partial l^i} : U_1(\cdot)pq_L(\cdot) - U_2(\cdot) & = 0 \\ \frac{\partial U(\cdot)}{\partial l^o} : U_1(\cdot)w^o - U_2(\cdot) & = 0 \end{cases} \quad (1.13)$$

where in this case:

$U_1(\cdot) = U_1(p, w^o, l^i, A, \mathbf{z^q}, E, l^o, \mathbf{z^h})$,

$U_2(\cdot) = U_2(p, w^o, l^i, A, \mathbf{z^q}, E, l^o, \mathbf{z^h})$ and

$q_L(\cdot) = q_L(l^i, A, \mathbf{z^q})$.

Combining the two first order conditions we obtain:

$$pq_L(l^i, A, \mathbf{z^q}) = w^o \quad (1.14)$$

that can be solved for l^i leading to a single, reduced form equation in which the variable l^i is expressed as a function of **only** "production side characteristics" and not of "consumption side characteristics":

$$l^i = f(p, w^o, A, \mathbf{z^q}) \quad (1.15)$$

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.

Proposition 2: For net sellers of labor for whom the off-farm labor constraint is binding, the separability between production and consumption decisions breaks.

If, instead, the quantitative constraint on the labor market is binding (and thus $l^o = \bar{L}$), the separability between production and consumption decisions no longer holds. Using the Kuhn Tucker conditions for this case we have:

$$\begin{cases} \frac{\partial U(\cdot)}{\partial l^i} : U_1(\cdot) p q_L(\cdot) - U_2(\cdot) = 0 \\ \frac{\partial U(\cdot)}{\partial l^o} : U_1(\cdot) w^o - U_2(\cdot) - \mu^{\bar{L}} = 0, \text{ and } \bar{L} = l^o \end{cases} \tag{1.16}$$

with:

$U_1(\cdot) = U_1(p, w^o, l^i, A, \mathbf{z^q}, E, l^o = \bar{L}, \mathbf{z^h})$,

$U_2(\cdot) = U_2(p, w^o, l^i, A, \mathbf{z^q}, E, l^o = \bar{L}, \mathbf{z^h})$ and

$q_L(\cdot) = q_L(l^i, A, \mathbf{z^q})$.

The household will sell $\bar{L}$ 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.

To find the optimal quantity of l^i, the equations above must be solved jointly for l^i and $\mu^{\bar{L}}$, recognizing that for this set of households, constraint (1.8) holds with equality and thus $l^o = \bar{L}$ (which implies that there are only two unknowns):

$$l^i = l^i(p, w^o, A, \mathbf{z^q}, E, \mathbf{z^h}, l^o = \bar{L}) \tag{1.17}$$

and

$$\mu^{\bar{L}} = l^i(p, w^o, A, \mathbf{z^q}, E, \mathbf{z^h}, l^o = \bar{L}) \tag{1.18}$$

As it can be seen, in this case, the constrained optimal allocation for on-farm labor l^i also depends on consumption side parameters, E, $\mathbf{z^h}$, and on the off-farm labor endowment $l^o = \bar{L}$. Therefore, the separability hypothesis breaks.

At this point, a brief discussion about the nature of the off-farm labor constraint is warranted. In particular, from equations (1.13) and (1.16) it can be easily verified that the off-farm labor allocation rule is given by:

$$l^o = \begin{cases} l^o(p, w^o, A, \mathbf{z^q}, E, \mathbf{z^h}) & \text{if } l^o < \bar{L} \\ \bar{L} & \text{if } l^o \geq \bar{L} \end{cases} \tag{1.19}$$

Denoting λ as the probability that a household is constrained we have:

$$\lambda = \lambda(l^o \geq \bar{L}) = \lambda(l^o(p, w^o, A, \mathbf{z^q}, E, \mathbf{z^h}) - \bar{L} \geq 0) \tag{1.20}$$

which in reduced form becomes:

$$\lambda = \lambda(p, w^o, A, \mathbf{z^q}, E, \mathbf{z^h}, \bar{L}) \tag{1.21}$$

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 equation (1.15) and equation (1.17) 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".

1.3 Econometrics

The preceding section establishes that it is conceivable for farm-households to participate in the market as net sellers, and yet be constrained by unobservable quantity limitations or by transaction costs in their ability to respond to price changes. Prior knowledge by the researcher of the sample division could be used to examine each labor regime separately. The problem, however, is that in our case such classification is unknown. This can be translated in econometric terms saying that the farm labor supply response function of a group of market participating households can be represented by a *switching regression model with unobserved sample separation* (Quandt [47], Maddala [40] p.302)[3].

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

$$l^1 = l_1(\mathbf{x_1}; \boldsymbol{\beta}) + u_1 \tag{1.22}$$

$$l^2 = l_2(\mathbf{x_2}; \gamma) + u_2 \tag{1.23}$$

$$\lambda = \lambda(\mathbf{x}_\lambda; \boldsymbol{\xi}) + u_\lambda \tag{1.24}$$

In our case, $\mathbf{x_1} = \{p, w^o, A, \mathbf{z^q}\}$, $\mathbf{x_2} = \{p, w^o, A, \mathbf{z^q}, E, \mathbf{z^h}, \bar{L}\}$ and $\mathbf{x}_\lambda = \{p, w^o, A, \mathbf{z^q}, E, \mathbf{z^h}, \bar{L}\}$. In addition, $\boldsymbol{\beta}, \gamma$ and $\boldsymbol{\xi}$ are coefficients to be estimated while u_j's are normal i.i.d. disturbances with zero means and variances σ_j^2.

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

$$l^i = \begin{cases} l_1(\cdot) & \text{if } \lambda \leqq 0 \\ l_2(\cdot) & \text{if } \lambda > 0 \end{cases} \tag{1.25}$$

The problem then becomes how to estimate the parameters $\{\boldsymbol{\beta}; \gamma; \boldsymbol{\xi}; \sigma_1, \sigma_2, \sigma_\lambda\}$ from the sample of N observations on $\{l^i, \mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_\lambda\}$, $i = 1, \ldots, N$. Given that a priori we cannot identify the regime composition, a randomly selected observation l^i (household i's

[3]Empirical work using a similar approach includes Lee and Porter [38] on cartels, Bash and Paredes-Molina [3] on dual markets in Chile, Murdoch and Stern [43] on sex bias, Pape and van Dijk [46] on growth rates convergence and Conway and Kimmel [10] on moonlighting.

on-farm labor supply) will have probability λ of belonging to the first regime, and probability $(1-\lambda)$ of belonging to the second one. As such, the conditional density of observation l^i given regime 1 can be written as:

$$f(l^i|regime1) = f(l^i - l_1(\mathbf{x}_1;\boldsymbol{\beta}))/\lambda$$

while the conditional probability of l^i given regime 2 is:

$$f(l^i|regime2) = f(l^i - l_2(\mathbf{x}_2;\gamma))/(1-\lambda)$$

where $f(\cdot)$ are probability density functions of u_1 and u_2.

The unconditional density of l^i is then:

$$\begin{aligned} f(l^i) &= \lambda f(l^i|regime1) + (1-\lambda)f(l^i|regime2) = \\ &= f(l^i - l_1(\mathbf{x}_1;\boldsymbol{\beta})) + f(l^i - l_2(\mathbf{x}_2;\gamma)) \end{aligned} \tag{1.26}$$

that is, the *mixture* of two distributions.

Focusing on the simpler case of mixture of two components, starting from the unconditional density function (1.26) and assuming that $f(\cdot)$ is normal, $\sigma_1 \propto \sigma_2$ and $\sigma_\lambda = 1$ (required for identification purposes) the likelihood function for a sample of N observations is then:

$$L(\lambda,\boldsymbol{\beta},\gamma,\sigma_1,\sigma_2) = \prod_N \left[\lambda\phi(l^i - l_1(\mathbf{x}_1;\boldsymbol{\beta}),\sigma_1) + (1-\lambda)\phi(l^i - l_2(\mathbf{x}_2;\gamma),\sigma_2)\right] \tag{1.27}$$

where $\phi(\cdot)$ denotes the normal density function. A natural way of estimating the parameters would be that of maximizing (1.27) with respect to $(\lambda,\boldsymbol{\beta},\gamma,\xi,\sigma_1,\sigma_2)$.

Properties of mixture distributions and the estimation of their parameters are extensively covered in Titterington [57] and Everitt [18]. Appendix B summarizes some econometric issues and provides insights on how to approach the problem of estimation in practice, as presented in Quandt [47], their extensive commentary (Hartley [27], Bryant [6], Clarke and Heathcote [7], Johnson [32], Hosmer [30], Kiefer [36], Binder [5] and Fowlkes [25]), and summarized in Maddala [40] pp. 302–305.

Agreement seems to have been reached on that estimation of the maximum likelihood, obtained via the so called E-M method (see Hartley [27] and Dempster et al. [13]) is a feasible approach, perhaps after having used other pre-estimation techniques to decide on the initial values of the parameters (Kiefer [35]). Moreover, the E-M approach can be used to estimate the model when λ is considered to be endogenously determined and specified as above. The only required assumption is that the variance of the error term of the classification equation (u_λ) is taken to be one.

The steps of the E-M algorithm are as follows: using starting values for $\beta, \gamma, \xi, \sigma_1, \sigma_2$, we first obtain estimates of the classification vector λ.(the E step). The starting values for the $\beta, \gamma, \sigma_1, \sigma_2$ 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 were the values would maximize the likelihood function. Using the estimate for λ to weight the probabilities of each observation to be in each regime and obtain estimates for $\beta, \gamma, \xi, \sigma_1, \sigma_2$ (the M step). This iterative procedure is repeated until the maximum likelihood function (in our case Equation (1.27)) converges.

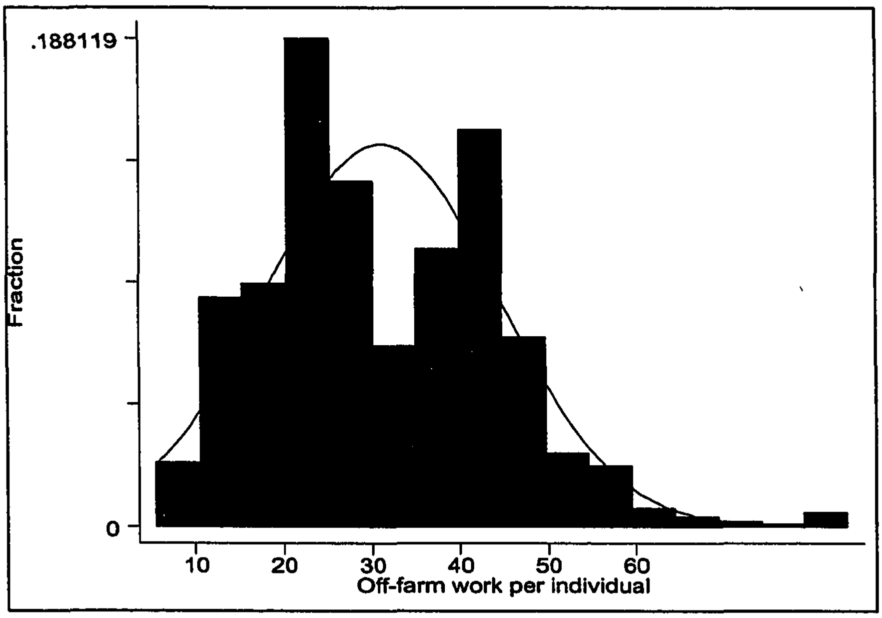
1.4 Data and Results

Data The data come from the 1997 Peruvian LSMS. The survey was contacted on 4500 households. From these, 1131 allocate work between 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 1.1, which plots the distribution of individual off-farm hours worked and seems to suggest the existence of two underlying processes. In particular, there seem to be two distinct subpopulations, one for which individuals work less off-farm and another that 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 on labor allocation on farm should be purely a production decision, and thus household characteristics (such as E, z^h, $\bar{L}$) should not affect it. Our theoretical model postulates the possibility of the presence of two different regimes among the 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 variables such as household composition z^h (like children, elder members or ethnicity). We specify the second regime by only including production side characteristics z^q.[4] Apart from this restriction that comes directly from the theoretical model, we do not impose any other restrictions on the parameters of the two

[4]Notice that in this specification of the unconstrained regime we exclude any variables that can be argued to affect both consumption and production decisions. On the other hand, our counterfactual estimation (below) includes such variables in both regimes.

Figure 1.1: Off-farm hours worked per individual



models in the belief that, if a dichotomy actually exists, it should be strong enough to let the econometric technique to separate between the two regimes. We apply the maximum likelihood procedure described in the previous section to determine the best way of dividing the observations in two groups. The dependent variable is the amount of hours allocated by the household for farm activities[5].

Table 1.1 presents the results. The first column contains the results of an OLS estimation of the model on the whole pooled sample[6]. As expected, production characteristics such as livestock assets, human capital and land assets significantly affect on-farm labor allocation. In addition, these pooled estimates seem to show a significant effect of

[5]Remember that we do not consider hired labor in this analysis since we are only looking at households that are net sellers of labor.

[6]The coefficients of this OLS estimation were also used as starting values for the likelihood maximization routine.

some household characteristics on the decision of labor use on-farm. This alone provides some indication that the issue of non-separability is important.

The second column in Table 1.1 presents the results from the maximum likelihood procedure. On average, a household has a probability of 0.48 to be constrained in its labor market participation. As we 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 small children and the elder significantly increase on-farm labor allocation. In addition, other variables that can be argued as both production and consumption ones also affect on-farm labor allocation. For example, education decreases the on-farm work, implying that human capital is an important asset for off-farm work. In addition regional dummies that capture idiosyncratic effects and shocks of specific geographic areas also affect on-farm labor decisions. Households that live in the mountains work more on-farm compared to Lima, an urban center. This is expected as off-farm opportunities are much more abundant in Lima than in the more isolate mountain regions[7].

For the second group of households estimated by the econometric procedure (and specified as the unconstrained regime), we find that production characteristics such as land and cattle ownership positively affect on-farm labor allocation. In addition, transaction costs in the form of the time to get to the main market, negatively affects work on-farm. One interesting finding is that women have no impact for on-farm labor. One hypothesis is that since these households are not constrained in labor, women may actually have off-farm jobs or do not need to work.

[7]We also estimated these models using district dummies and got similar results.

Table 1.1: On-farm labor allocation: pooled versus mixture model

		Pooled	Mixture		
			Constrained	Unconstrained	Switcher
Market wage (soles)	w^o	-0.03	-0.14***	-0.05**	0.03**
Land owned (ha.)	A	0.15**	0.92*	0.21***	-1.37***
Adult males (#)	E	25.66***	36.25***	2.54**	19.27***
Adult women (#)	E	25.90***	34.74***	1.24	17.05***
Time to market (min.)	p	0.96	6.20	-3.80***	6.98***
Time to market sq. (min. sq.)	p	0.02	-0.95	0.16***	-0.61***
Cattle owned (#)	z^q	1.30***	0.96***	1.39***	0.06
Coast (dummy)	$z^{q,h}$	13.14	12.02		17.04***
Sierra (dummy)	$z^{q,h}$	28.88***	27.32***		37.95***
Rainforest (dummy)	$z^{q,h}$	25.37***	7.03		32.96**
hh average education (years)	$z^{q,h}$	-10.41**	-12.18**		-7.62***
hh average ed. sq. (years sq.)	$z^{q,h}$	-0.39**	0.67***		-0.01
hh head sex (male=1)	$z^{q,h}$	11.72	14.66		11.08**
hh average ed.* indigenous hh	$z^{q,h}$	-0.71**	-0.42		-0.56***
head age (years)	$z^{q,h}$	0.26	1.35		-0.73
head age sq.(years sq.)	$z^{q,h}$	-0.003	-0.02*		0.002
Boys (#)	z^h	6.51**	18.48***		-9.57***
Girls (#)	z^h	17.18***	38.53***		-2.29
Elder (#)	z^h	12.66**	43.07***		-2.30***
Indigenous hh (yes=1)	z^h	37.97**	19.23		17.38***
Hours worked off-farm	l^o	-0.20***	-0.32***		-0.14***
Private jobs in community (yes=1)	z^q	-49.96	-18.76		-65.10***
Public jobs in community (yes=1)	z^q	8.29	-1.81		18.40***
Constant		130.98**	122.9**	73.98***	33.29
Sample proportion	$\widehat{\lambda}$	1.0	0.48	0.52	1.0
R^2 (pooled)		0.2			
Log likelihood (mixture)		-261			

Dependent variable: household's on-farm work (hours).

Switcher: probability of being constrained.

The coefficients of the switcher equation are all multiplied by 100 ($\widehat{\xi} * 100$).

The missing dummy for regions is Lima.

Significance levels:*: 90%, **: 95%, ***: 99%

Sample size: 1131

For both regimes, we also expected a significantly different value for the coefficient on the wage rate. For the unconstrained regime it should be significant and negative, in which the wage rate plays just the role the marginal product of labor. However, for the constrained regime, the relevant price for labor to be used on farm is, instead, a shadow price determined by the subjective equilibrium of the farm-household, and therefore the wage rate will not necessarily reflect the marginal product of labor. Instead the market wage will affect the production decisions only indirectly, through its effect on total family income.

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 for the unconstrained. Table 1.2 reports the results of this specification. The results do support the separability hypothesis for the second regime. Specifically, none of the consumption characteristics have a significant impact on the on-farm labor decisions of unconstrained households. Two of the regional dummies affect this decision, but as we discussed above, this could reflect production aspect differences in marketing opportunities among regions.

To recapitulate, applying the maximum likelihood procedure allowed us to separate the sample in two sub-populations of labor market participants (net sellers). We find strong evidence that in one of them, the separability hypothesis between production and consumption decisions is rejected, while for the second it is not. A counterfactual test

Table 1.2: Counterfactual estimation: On-farm labor allocation using mixture model

		Mixture		
		Constrained	Unconstrained	Switcher
Market wage (soles)	w^o	-0.12***	-0.06**	0.05***
Land owned (ha.)	A	0.89	0.22***	-1.03***
Adult males (#)	E	18.43***	-0.41	32.93***
Adult women (#)	E	26.82***	-0.44	23.14***
Time to market (min.)	p	8.42	-3.61***	2.86
Time to market sq. (min. sq.)	p	-0.59**	0.16***	-0.27***
Cattle owned (#)	z^q	0.66*	0.69***	0.90***
Coastal dummy	$z^{q,h}$	17.78	1.77	19.05***
Sierra dummy	$z^{q,h}$	36.97***	-4.85*	31.91***
Rainforest dummy	$z^{q,h}$	13.69	5.49*	32.65***
hh average education (years)	$z^{q,h}$	-16.45***	1.72	-2.41
hh average ed. sq. (years sq.)	$z^{q,h}$	0.72***	-0.06	-0.18
hh head sex (male=1)	$z^{q,h}$	24.13**	1.93	19.07***
hh average ed.* indigenous hh	$z^{q,h}$	-0.31	-0.18	-0.37*
head age (years)	$z^{q,h}$	4.52***	-0.17	-4.35***
head age sq.(years sq.)	$z^{q,h}$	-0.05***	0.001	0.05***
Boys (#)	z^h	9.01**	0.99	-4.92**
Girls (#)	z^h	27.55***	-0.89	4.51**
Elder (#)	z^h	29.12***	-1.23	-13.52***
Indigenous hh (yes=1)	z^h	11.09**	3.70	44.55***
Hours worked off-farm	l^o	-0.17***		-0.18***
Private jobs in community (yes=1)	z^q	-31.25	-36.26	-27.35
Public jobs in community (yes=1)	z^q	5.53	0.70	15.12***
Constant		102.91*	107.18***	16.71
Sample proportion	$\widehat{\lambda}$	0.49	0.51	1.0
Log likelihood (mixture)		-254		

Dependent variable: household's on-farm work (hours).

Switcher: probability of being constrained.

The coefficients of the switcher equation are all multiplied by 100 ($\widehat{\xi} * 100$).

The missing dummy for regions is Lima.

Significance levels:*: 90%, **: 95%, ***: 99%

Sample size: 1131

strengthens our findings by showing that the households in the unconstrained regime behave in a separable fashion.

Understanding the labor constraint The results above provide sensible indication of the presence of a group of farm household which, albeit participating in the labor market, are making their decision 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 1.1 contains the estimated coefficients for the equation that determines the group separation. Given the specification of the two regimes, we can interpret this switcher equation to represent the probability of being in the constrained regime. We use equation 1.21 from the theoretical part to specify the determinants of group separation.

Interesting patterns seem to emerge that relate to market integration and participation. The market wage itself increases the probability of being constrained. A higher wage may make it harder to find an off-farm job due to increased competition (by other workers) and decreased demand by employers. In addition, a larger farm (in terms of land size) lowers the probability to be constrained via the increase in labor demand for on-farm labor. In terms of other household characteristics, a larger number of both male and female members (E) in the household as well as high time costs of accessing markets (p) increase the probability to be constrained, while a larger number of children and elder members (z^h) decrease it.

Interactions between ethnic classification and education are also important to characterize market opportunities. Being indigenous increases the probability to be constrained, perhaps implying the presence of ethnic discrimination. On the other hand, households with higher levels of education are less likely to be constrained. Moreover, an interaction term between the indigenous dummy and education levels captures the importance of education: higher levels of education among indigenous households may mitigate their access to labor markets by reducing the probability to be constrained.

At the regional level, households living closer to mountains and the rainforest (as opposed to Lima) have a higher probability to be constrained as their off-farm opportunities there are more limited. The same is true for the coastal areas but the effect does not seem as strong. We also wanted to explore further the condition of the labor market beyond the regional dummies. While the data set is limited for this, we include two variables that capture market demand and availability for off-farm work at the village level. We differentiate between availability of private and public off-farm job opportunities. Interestingly, while the availability of private jobs decreases the probability to be constrained as expected, the availability of public jobs increase it.

Finally, working more off-farm is negatively related with being constrained. Off-farm work can be thought as representing the off-farm labor constraint ($\overline{L}$) derived in the theoretical model. Relaxing this constraint therefore makes it less likely to be constrained.

The probability to be constrained One of the benefits of our estimation methodology is that we can divide the sample between those that are more likely to be constrained and those that are not (using the predicted probability to be in a given regime $\widehat{\lambda}$). This leads to an ex-post predicted sample of "constrained" and "unconstrained" households. We use different ways to get insights on the structure of the labor market constraint from these sub-samples.

From our theoretical framework, we know that the constrained households' off-farm labor supply l^o is also their binding off-farm allocation labor $\bar{L}$. This distribution is plotted in Figure 1.2. We would expect that if there was a common market based barrier on the amount of labor to be allocated, it would show up as a very narrow distribution for $l^o = \overline{L}$ 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 1.2 also plots the distribution of individual off-farm hours for the unconstrained. These individuals overall work more 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.

Another way to take of advantage of our methodology is by looking at the predicted probabilities of being constrained. Figures 1.3 and 1.4 plot two such distributions by ethnic group and regions respectively. In Figure 3, we see that indigenous people have overall higher probabilities of being constrained, supporting our findings above and the hypothesis of possible ethnic discrimination. While the contrast is not as strong when we look at this in terms of regional differences, there is some evidence to suggest that households living in mountain regions are more likely to be constrained. This may reflect the fact the off-farm labor opportunities in the Peruvian sierra are usually limited and even non-existent.

We also look at the link between being constrained and poverty. We use both income and consumption quintiles to compare the probabilities of being constrained (Ta-

Figure 1.2: Kernel densities of off-farm hours worked per individual

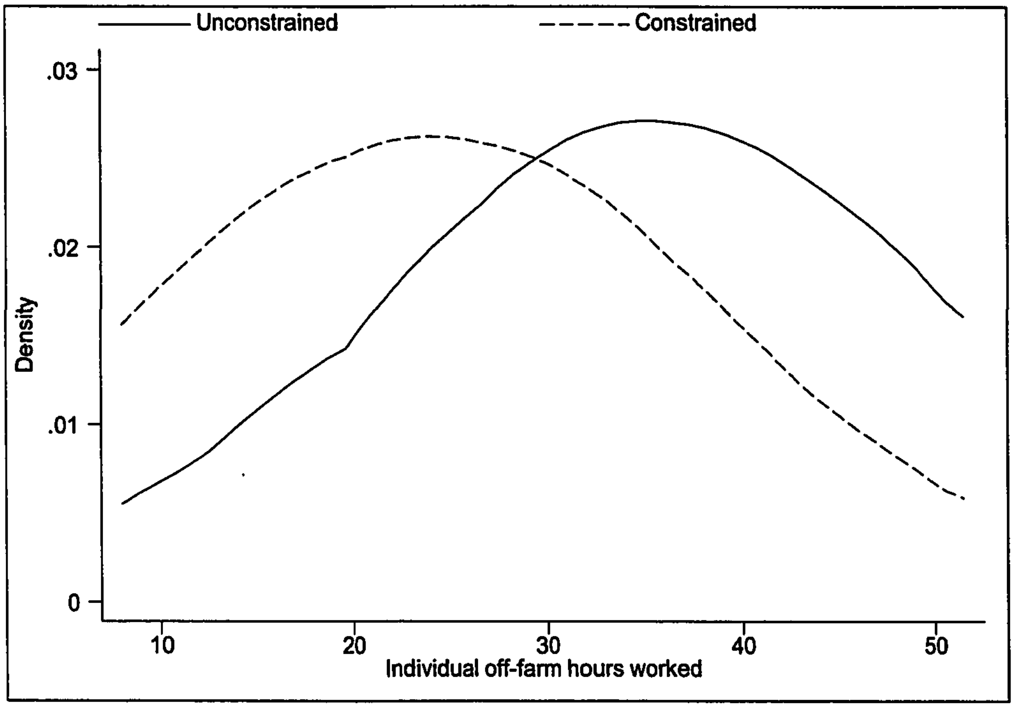


Figure 1.3: Kernel densities of the probability to be constrained by ethnicity

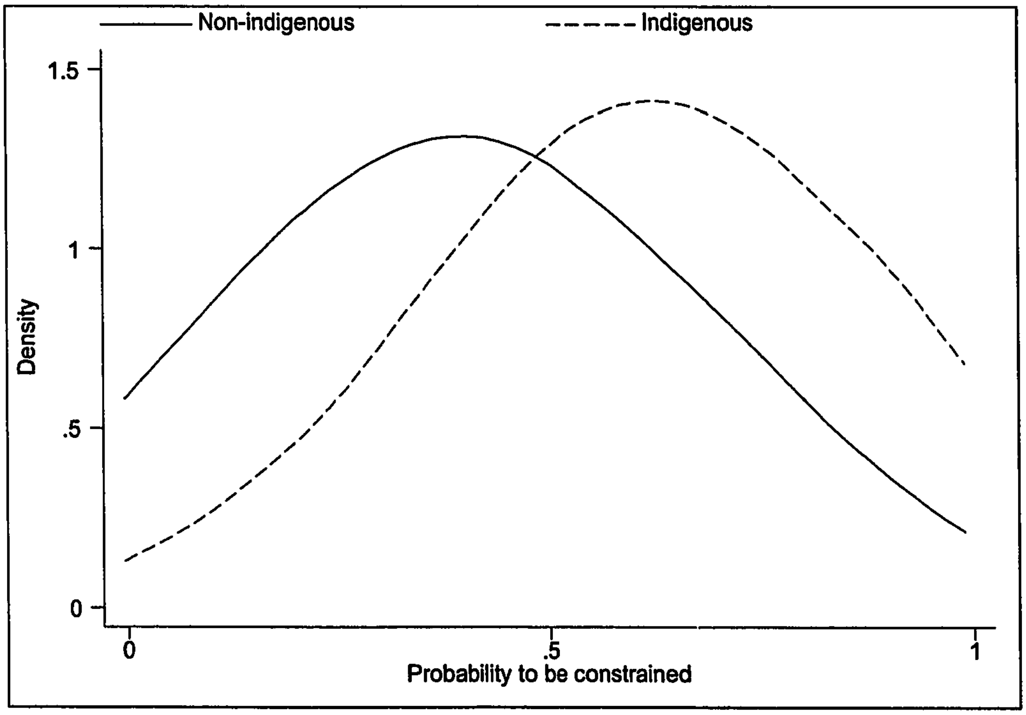


Figure 1.4: Kernel densities of the probability to be constrained by geographic region

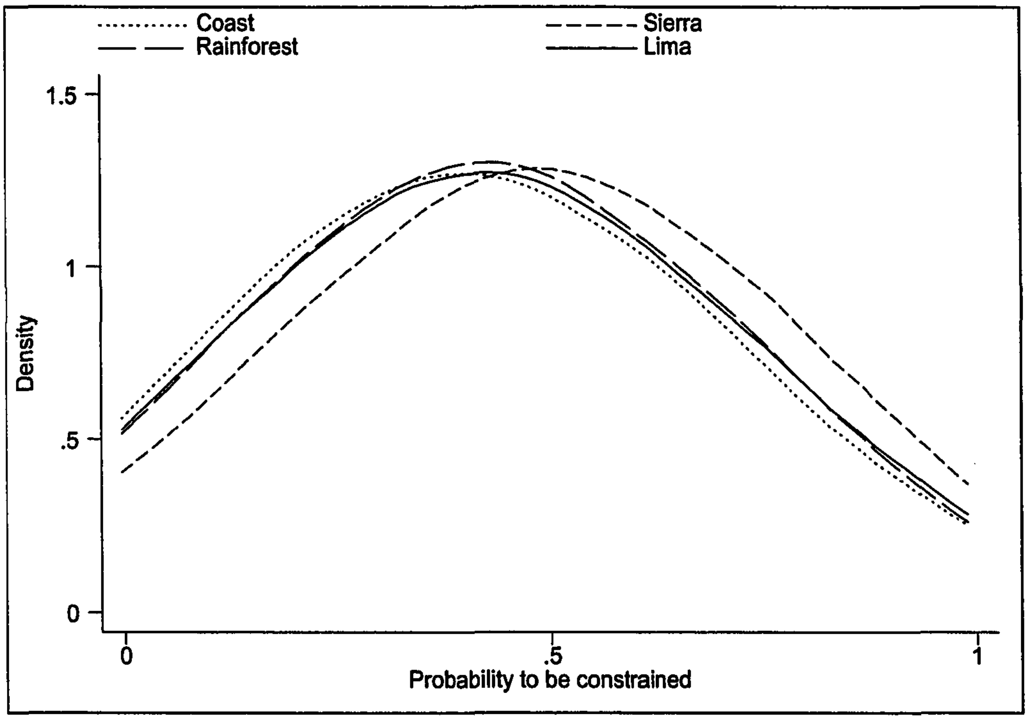


Table 1.3: Probabilities to be constrained by quintiles

	First (lowest)	Second	Third	Fourth	Fifth (highest)
Income	53	49	43	43	37
Consumption	48	50	45	42	36

ble 3). Indeed, there is a strong correlation between being constrained and poverty. For both indicators, being poorer implies having a higher chance of being constrained. This probability decreases at higher levels of consumption and income.

Finally, a more descriptive comparison between "constrained" and "unconstrained" households is presented in Table 1.4. Overall, constrained households have lower levels of key assets: they own less land, are less educated and are poorer in terms of both income and consumption. Constrained households are also larger and predominantly indigenous. These observations suggest that access to off-farm opportunities may be closely related to access and accumulation of both human and physical assets.

1.5 Concluding remarks

We use mixture distribution techniques to show that labor market participation is not enough to conclude separability between production and consumption decisions of farm-households. Our 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 some that behave as if constrained. These results provide an important refinement of separability studies by extending and expanding the concept of heterogeneity.

In addition, we also take advantage of the econometric technique to explore and understand the market segmentation and its source. In doing this, we try to look at the role of both demand and supply side effects on labor constraints. In the case of Peru, ethnic discrimination, differential education attainments as well as differences in regional opportunities seem to be some important deterrents for market participation.

Some interesting policy implications arise. First, if we are able to characterize well the source of the sample separation, it enables us to assess (at least qualitatively) the impact of different policies. For example, in our case, we find that almost half the sample is likely to be constrained in the labor market, even though they participate in it. Therefore, a wage policy that ignored this observation would not only be incorrect but quite ineffective as well. In particular, policies using pooled wage elasticities would underestimate the effects of a wage change since the constrained share of the pooled sample would not alter its labor market behavior based on a wage policy (also see Figure 1.5). The methodology proposed

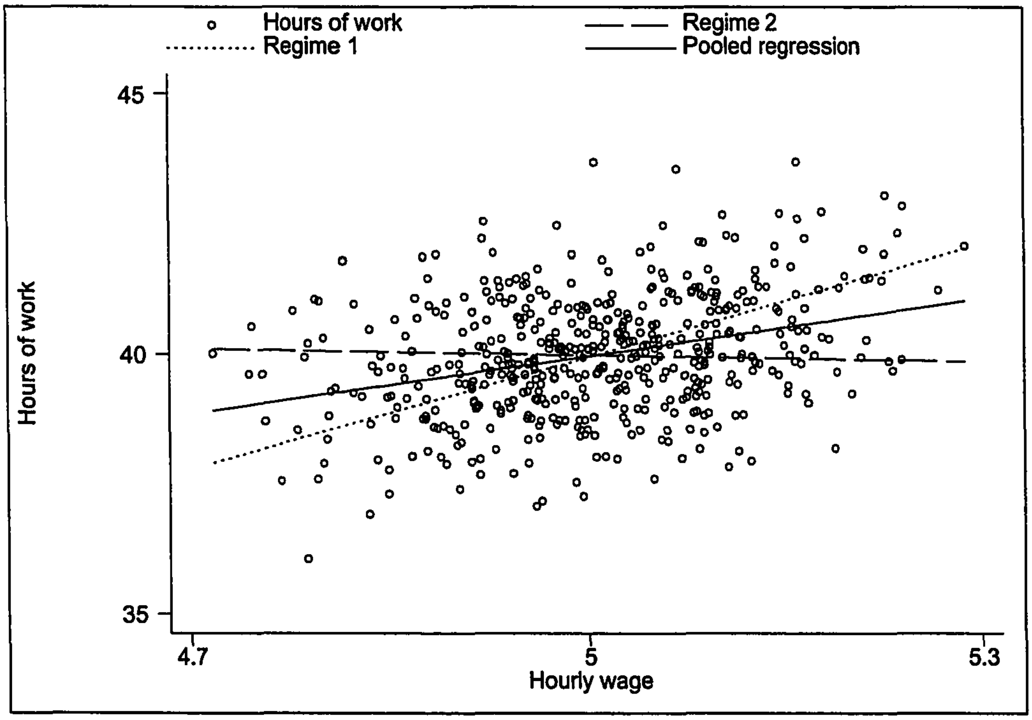
Table 1.4: Household descriptive statistics by predicted constrained regime

		Constrained	Unconstrained	Pooled
hh on-farm work (hours per week)	l^i	153	105*	124
hh off-farm work (hours per week)	l^o	112	108	110
Per hh member on-farm work (hours per week)	l^i	39	44*	42
Per hh member off-farm work (hours per week)	l^o	27	42*	36
Market wage (soles)	w^o	189	219*	207
Adult men in hh (#)	E	2.1	1.3*	1.6
Adult women in hh (#)	E	2.2	1.4*	1.7
Land owned (hectares)	A	0.7	2.8*	1.9
Cattle owned (#)	z^q	3.2	0.9*	1.8
Time to market (min.)	p	54	66*	60
Average hh education (years)	$z^{q,h}$	9.5	9.9*	9.6
hh head education (years)	$z^{q,h}$	8.6	11.2*	10.2
head age (years)	$z^{q,h}$	55	46*	50
hh head sex (male=1)	$z^{q,h}$	0.8	0.9*	0.9
Coastal (%)	$z^{q,h}$	20	26*	24
Sierra (%)	$z^{q,h}$	30	21*	25
Rainforest (%)	$z^{q,h}$	17	19	18
Lima (%)	$z^{q,h}$	31	34	33
Boys in hh (#)	z^h	1.1	1.2*	1.1
Girls in hh (#)	z^h	1.2	1.0*	1.0
Elder in hh (#)	z^h	0.4	0.4	0.4
hh size (#)	z^h	7.1	5.3	6.0
Dependency ratio $(\frac{\#\ children + elder}{\#\ adults})$	z^h	0.7	1.1*	0.9
hh indigenous (%)	z^h	32	7*	17
Per capita income (soles)	y	1111	1557*	1362
Per capita consumption (soles)	C	2467	3532*	3101
	obs	458	673	1131

Note: * means that there is a significant difference between the unconstrained and constrained groups at the 90% level or more here could be then used to correctly estimate more meaningful elasticities based on the unconstrained and thus relevant portion of the population.

The characterization of the source of the market constraint itself also provides us with the means to better understand and implement different policies that can be effective to ameliorate market integration and participation. An exhaustive use of possible individual, regional and market heterogeneity (for example unemployment rates, infrastructure, human capital differences as well as other sources for non separability such as credit and risk) could allow us to make inferences about the structure of the market failure. While the nature of our data and their lack of such a list of supply side variables limits our analysis somewhat, the methodological insights we get are important. Without our data constraints, this methodology can allow future analyses to carefully compare different policy options that may arise which can be used to guide policy makers to devise and target policies that maximize social welfare.

Figure 1.5: Pooling effects



Chapter 2

Overcoming Credit Market Failures: A Paradigm of Diversification for Technology Adoption

2.1 Introduction

Technology adoption in developing countries has received renewed attention in recent years. A lot of importance has been placed on the role of agricultural innovations and the tremendous opportunity to increase production and incomes through the use of these new technologies and improved crop varieties. Yet, the vast heterogeneity of the economic environments where new technologies are being applied, as well as the low observed rates of adoption of different technologies around the developing world open a debate as to how to increase both the speed and the extent of adoption.

Adoption literature, both theoretical and empirical, has focused on identifying the factors that affect adoption decisions and the different constraints that may exist when such decisions are contemplated (Feder et al. [22] and Feder and Umali [23] provide an extensive literature review). In general, these factors can be divided in two categories: farm attributes and technology attributes. Literature on farm attributes looks at the links between adoption and farm characteristics. In these studies, issues like small vs. large, subsistence vs. market oriented farming (Kaliba [33]), human capital (Strauss et al. [55]), risk and risk management strategies (Hiebert [28], Saha [52]), institutional support systems, asymmetric information (Feder [20]), production factors availability, factor endowments, level of off-farm income, other income sources and credit constraints (Dimara and Skuras [14]), and their links with adoption decisions are investigated.

The second body of literature addresses how technological attributes themselves can affect adoption. For example, one technology can be better in one aspect of production (e.g. the new seed is resistant to a pest) but another technology may have other advantages (e.g. lower risk). As a result, these attributes combined with the idiosyncratic characteristics of the farmer can affect adoption (Misra et al. [41]).

In spite of the vast literature that addresses adoption and the influence of the factors mentioned above on the adoption process, there is still not a consensus on why some farmers adopt new technologies and others do not. Many reasons have been debated as to the lack of such consensus. On one hand, it may not be sensible to talk about one adoption policy given the complexity and heterogeneity of both new technologies and the setting where they are applied. However, even for the same technology, research methodologies to assess the barriers to adoption vary considerably. For example, most adoption studies treat adoption as a discrete phenomenon. Thus, a farmer that adopts ten percent of a new input will be treated the same as a farmer who adopts one hundred percent. Therefore, the researcher may fail to capture particular intricacies that the continuous case offers (Feder et al. [22], Rauniyar and Goode [48]).

Another shortcoming of adoption studies has been the type of link they have assumed between adoption and other farm practices. Most adoption studies generally assume (implicitly or explicitly) that adoption of new technologies generates important ex-post changes in the agricultural product mix (Ellis [19]). It is argued that adoption increases farmers' incomes so that those who have adopted new technologies are more likely to engage in other new activities, thus creating a more diversified income portfolio. In addition, adoption of new technologies may interact with agroclimactic and other biological constraints and enable the introduction of new crops or varieties, altering the product mix. An example is the introduction of soybean cultivation and Zebu cattle due to new water conservation technologies into areas of Brazil considered otherwise too hot and dry to support them (Nerlove [44]).

Yet, the reverse may also be true. Farmers facing an adoption decision and confronted with a cash constraint in the growing season may be able to relax it by altering their product mix and shifting resources to cash generating activities. In particular, while new technology profit realizations will come at the end of the harvest season, as in the case of potato cultivation, by reallocating resources to activities that generate cash income in the interim, farmers may be able to increase their ability to invest more in the new technology. This cash may not only mitigate the seasonality problem of unstable stream of income but may also help reducing risk aversion in farm production decisions (Eswaran and Kotwal [16], [17]). Therefore, self-financing has important policy ramifications as it suggests an alternative mechanism to alleviate and overcome institutional failures such as those in credit markets.

In the case Peru, the adoption of high-yield potato varieties has been slow. While the technology exists, adoption of such varieties has not diffused at the level of small and poor farmers. In the highlands of Peru, the majority of the farmers are small-scale ones with usually less than five hectares of land (Crissman [11]). They generally farm with oxen or by hand. Moreover, even if access to markets is relatively easy, the demand for new varieties of potato seeds is low due to the high costs associated with it. Specifically, potato seed is an input whose quality is observed only after it has been purchased. In addition, unlike other crops, clonal potato seed is bulky and can easily transmit diseases. As a consequence, the costs of buying new seed varieties are high. It is estimated that seed costs represent up to fifty percent of the total production costs for potato cultivation in Peru (Monares [42]). In most instances, small-scale farmers cannot afford these high monetary costs to buy these inputs. The unavailability or failure of institutional support such as credit institutions further impedes or binds the extent of adoption. Thus, understanding the adoption process and the role of credit in this setting is a high priority.

This paper addresses the "diversification for adoption" idea with data of small-scale farm potato producers in rural Peru. Using a continuous measure to capture adoption intensity, we argue that while credit constraints dramatically decrease adoption rates, farmers can alleviate and overcome these market failures via alternative mechanisms. In particular, by diversifying the farm's income sources through inclusion of cash generating activities, credit constrained farmers can increase their liquidity and access the new technologies. Cash income, such as dairy, flows in the farm in many frequent time intervals, as opposed to the large lump-sum payment that the farmers get during the harvest of seasonal crops like potatoes.

In the area of Cajamarca, Peru, where our data comes from, dairy farming has been a recent phenomenon. Nestle and a local company have established dairy centers in the region where farmers can sell their milk. In this sense, at least in the short run, only farmers that a priori had the capability to engage in dairy farming would do so. Therefore, while there is a high positive correlation between wealth and dairy farming, as we discuss below, the exogenous introduction of dairy farming allows us to link dairy income to adoption of new potato varieties. Our empirical results show that such an inflow of cash relaxes the cash liquidity constraint and allows farmers to access expensive inputs (seeds of new improved potato varieties) and therefore increase their adoption rates.

The paper proceeds as follows: section 2.2 develops the theoretical framework. The empirical methodology is discussed in section 2.3, the data is introduced in section 2.4, while section 2.5 presents the results. Section 2.6 concludes.

2.2 An adoption model with two technologies and a cash activity

We use a static adoption model that links diversification with credit market failures. We begin by assuming a profit maximizing farm-household and a missing labor market. While the latter may seem a strong assumption, it corresponds to the labor market conditions (or lack of) faced by the Peruvian small farm-households that comprise the data.

The farmer allocates labor between crop production and a cash activity. In addition, the farmer chooses between two crop varieties: either adopting new variety (N) that requires additional inputs X_N, or the traditional variety (T) that only requires labor. Formally, the farmer's problem is:

$$\underset{l_N, l_T, l_C, x_N}{Max} \Pi = p_N q_N(l_N, x_N, z^q) + p_T q_T(l_T, z^q) + p_C q_C(l_C, z^c, z^q) - p_{x_N} x_N \tag{2.1}$$

where:

N, T, C are subscripts for the new crop variety, traditional variety and the cash activity respectively,

p_j is the exogenous price of product j,

q_j is a production function of j,

l_j is the labor allocated to the production of j,

x_N are variable inputs for the production of the new variety N,

z^q are production related characteristics for all activities,

z^c are characteristics specifically related to the cash activity.

The farmer faces a labor constraint:

$$\overline{L} - l_N - l_T - l_C \geq 0 \tag{2.2}$$

where $\overline{L}$ is the total labor endowment of the farm,

a cash liquidity constraint:

$$C - p_{x_N} x_N \geq 0 \tag{2.3}$$

where C is the total income from the cash activity:

$$C = p_C q_C(l_C, z^c, z^q) \tag{2.4}$$

and the non-negativity constraints:

$$l_N \geq 0 \tag{2.5}$$

$$x_N \geq 0 \tag{2.6}$$

$$l_T \geq 0 \tag{2.7}$$

$$l_C \geq 0 \tag{2.8}$$

We can solve the maximization problem in two parts. Below we present the main results. Appendix B has a complete derivation of the solution to this maximization problem. First, we assume that the labor constraint is binding (i.e. equation (2.2) holds with equality) and focus on the case where farmers produce both the traditional and the new variety. This implies that none of the corresponding non-negativity constraints are binding. Substituting equation (2.2) for l_T, the maximization problem becomes:

$$\underset{l_C}{Max}\{Max \;\; p_N q_N(l_N, x_N, z^q) + p_T q_T(\overline{L} - l_N - l_C, z^q) + p_C q_C(l_C, z^c, z^q) - p_{x_N} x_N\} \tag{2.9}$$

subject to constraints (2.3), (2.8) as well as (2.5), (2.6), (2.7) being strictly positive.

The first order Kuhn Tucker conditions that describe the interior solution are:

$$\frac{\partial \Pi(\cdot)}{\partial l_N} : p_N \frac{\partial q_N}{\partial l_N} - p_T \frac{\partial q_T}{\partial l_N} = 0 \tag{2.10}$$

and

$$\frac{\partial \Pi(\cdot)}{\partial x_N} : p_N \frac{\partial q_N}{\partial x_N} - p_{x_N}(1 + \mu) = 0, \;\; \mu \geq 0 \tag{2.11}$$

where μ is the multiplier associated with the cash liquidity constraint.

Two cases arise depending on whether the cash constraint is binding or not.

Case 1 *Unconstrained*

If the cash constraint is not binding ($\mu = 0$), the optimal quantities of x_N and l_N are given by:

$$x_N^u = x_N^u(p_N, p_{x_N}, p_T, z^q, \overline{L} \,|\, l_C^u) \tag{2.12}$$

and

$$l_N^u = l_N^u(p_N, p_{x_N}, p_T, z^q, \overline{L} \,|\, l_C^u) \tag{2.13}$$

where the superscript u refers to the unconstrained case.

In addition, since

$$l_T = \overline{L} - l_N - l_C$$

and given (l_C):

$$l_T^u = l_T^u(p_N, p_{x_N}, p_T, z^q, \overline{L} \,|\, l_C^u) \tag{2.14}$$

Notice that the optimal levels of these inputs depend on the labor demand for the cash activity (l_C). Substituting for l_N^u, x_N^u and l_T^u in equation (2.9), the reduced form for the cash activity's labor demand l_C can be solved as:

$$l_C^u = l_C^u(p_N, p_{x_N}, p_T, z^q, \overline{L}, p_C, z^c) \tag{2.15}$$

Case 2 *Constrained*

On the other hand, if the cash constraint is binding ($\mu > 0$) then:

$$x_N^c = x_N^c(p_N, p_{x_N}, p_T, z^q, \overline{L}, l_C^c, C) \tag{2.16}$$

$$l_N^c = l_N^c(p_N, p_{x_N}, p_T, z^q, \overline{L} \,|\, l_C^c, C) \tag{2.17}$$

and

$$l_T^c = l_T^c(p_N, p_{x_N}, p_T, z^q, \overline{L} | l_C^c, C) \tag{2.18}$$

where the superscript c refers to the constrained case.

In this case, these quantities do not only depend on the labor demand for the cash activity, but also the cash activity income itself. These are given by:

$$\begin{cases} l_C^c = l_C^c(p_N, p_{x_N}, p_T, z^q, \overline{L}, p_C, z^c) \\ C = C(p_N, p_{x_N}, p_T, z^q, \overline{L}, p_C, z^c) \end{cases} \tag{2.19}$$

The important implication of the model above is that while cash income (C) does not influence the adoption decision for liquidity unconstrained farmers, it does so for the constrained. By engaging in the cash activity, cash constrained farm-households generate income that allows them to substitute for their inability to access credit. This in turn allows them to purchase cash inputs (X_N) and increase their adoption rates of the new technology. It is in this context that diversification can be thought as an alternative mechanism to overcome credit market failures that impede adoption.

2.3 Econometric specification

The findings of section 2.2 show that cash income (C) affects the adoption decision only for those farm-households that are cash constrained. This has important consequences for statistical modeling. In particular, while in the unconstrained case crop production and cash income practices are unrelated in terms of cash liquidity, this is not true for those farmers that are cash constrained. The implication is that to correctly model the adoption process, we need to simultaneously model both decisions.

To formalize, we first define adoption (A) to be the share of the seeds planted of the new variety over the total quantity of seeds planted. Then, from equations 2.16 and 2.19, adoption for cash constrained farm-households can be specified in the following system:

$$A^c = A^c(x, C, l^c_C; \beta) + \varepsilon_1 \tag{2.20}$$

$$C = C(x, w; \gamma) + \varepsilon_2 \tag{2.21}$$

$$l^c_C = l^c_C(x, w; \delta) + \varepsilon_3 \tag{2.22}$$

where:

$x = \{p_N, p_{x_N}, p_T, z^q, \overline{L}\}$ is a vector of characteristics that affect both the adoption decision and the cash activity,

$w = \{p_C, z^c\}$ are instruments correlated with the cash activity but not with adoption,

$\varepsilon_1, \varepsilon_2, \varepsilon_3$ are uncorrelated, normally distributed, i.i.d. disturbances and

β, γ and δ are vectors of coefficients to be estimated.

Equations (2.20) through (2.22) constitute a recursive system. However, both the cash activity labor demand and cash activity income are endogenously determined with the adoption decision. We can consistently estimate this system using instrumental variable techniques.

On the other hand, our theory suggests that cash income does not affect the adoption decision for farmers that are unconstrained. For them, the correct specification using equations (2.12) and (2.15) is:

$$A^u = A^u(x, l^u_C; \nu) + \omega_1 \tag{2.23}$$

$$l_C^u = l_C^u(x, w; \xi) + \omega_2 \tag{2.24}$$

where

x and w are defined as above,

ω_1, ω_2 are uncorrelated, normally distributed, i.i.d. disturbances and

ν and ξ are vectors of coefficients to be estimated.

This latter case, as we discuss below, provides us with a counterfactual test between the relationship of cash income and adoption by testing whether cash income indeed does not affect the adoption process for these farmers. For both specifications we implement a two-stage least squares estimation.

2.4 Data

The data comes from the province of San Miguel, which is located in the state of Cajamarca, in northern Peruvian Andes. The economy in the region is dominated by small farms with potatoes being the main agricultural crop. In addition to potatoes, production of dairy (cash activity) has evolved in recent years as another important activity. The farmers sell their milk to Nestle, which operates refrigeration posts throughout the region. Farmers bring their milk daily to these posts. Recently, another dairy company has entered the market, offering competitive prices for milk. The World Bank collected this data in 1999 with the goal of obtaining baseline data to evaluate the impact of a pilot farmer field school program being administered by the International Potato Center and CARE International.

Following Feder et al. [21], we classify a farmer as credit constrained if she wanted credit and could not get it either because collateral was missing or credit was not avail-

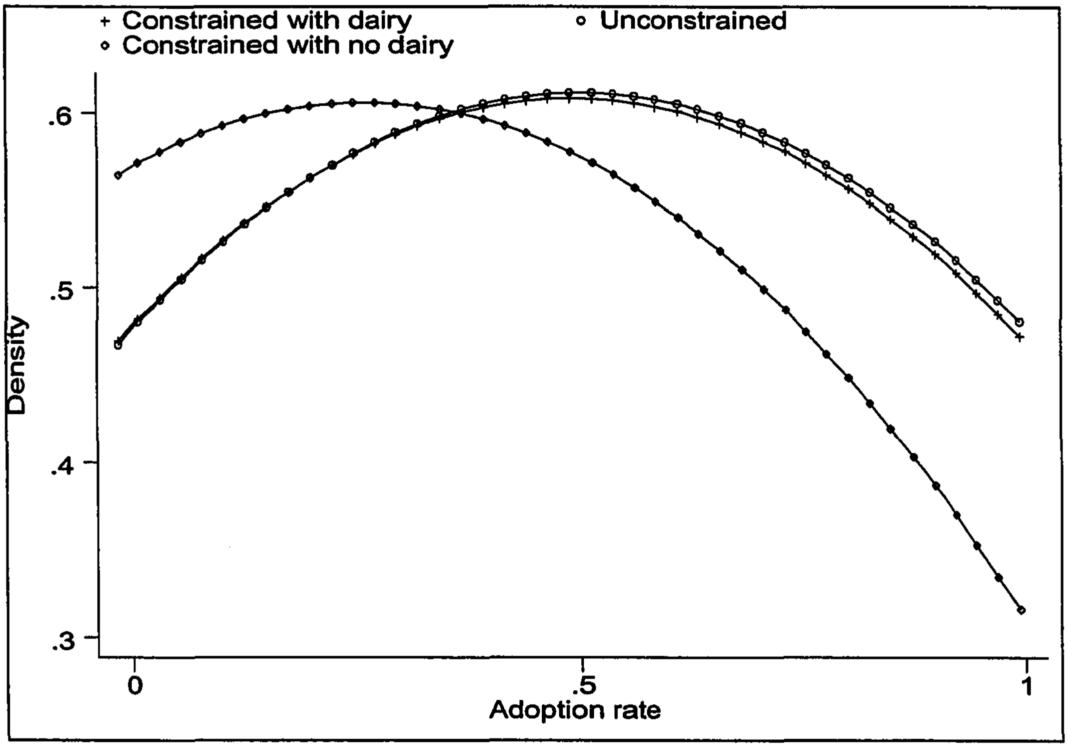
Table 2.1: Typology of farm-households

	Cash activity		
Credit constrained	Yes	No	Total
Yes	187	46	233
No	224	29	253
Total	411	75	486

able. A farmer who did not get credit because credit was too expensive is classified as credit unconstrained (Table 2.1). Since the level of our analysis is within the constrained subpopulation, we treat this classification as exogenous. That is, since we are interested in understanding behavior among constrained farmers only (by showing that dairy income allows those who engage in dairy farming to adopt more), we do not endogenize the credit constraint.

The distribution of adoption intensities of the new potato variety is presented in a kernel density in Figure 2.1. Intensity of adoption is defined as the ratio between the quantity of new variety seeds planted to the total seeds planted. The striking observation from the figure is to notice that credit unconstrained farm-households have a very similar adoption distribution to that of the credit constrained farm-households that produce dairy. In addition, credit constrained farmers that do not engage in the dairy activity have much lower rates of adoption. On average, farmers in dairy have adoption rates of around fifty percent (also see Tables 2.2 and 2.3) compared with only twenty five percent for those with no dairy income. These observations suggest the existence of complementarities between the dairy activity and adoption of the new varieties of potatoes. In fact, if the adoption rates for the unconstrained are thought to be the outcome of "optimal" rules, then dairy income seems to compensate those farm-households that are credit constrained so that they

Figure 2.1: Adoption intensities of new variety



can achieve similar adoption rates to those of the unconstrained.

Descriptive statistics for the *credit constrained* households are presented in Table 2.2. In terms of income, more than half of a typical farm-household's income is derived from potatoes[1]. Yet, for credit constrained households engaged in dairy, potato income only comprises half of total income compared to more than two thirds for those not engaged in dairy. In addition, farm-households who are engaged in the cash activity are wealthier than those who are not. They own more than twice the amount of land and have both more farm and household assets. They are also closer to the main market and have better road infrastructure. The average annual dairy income for those who engage in it is 2877 soles

[1]At the time of the survey, not all households had finished harvesting. Therefore, the results are likely to be underestimating potato income. Any comparisons are only suggestive and should be treated cautiously.

(around US $ 800), a considerable amount.

Finally, Table 2.3 compares credit constrained and unconstrained farm-households. Unconstrained farm-households are wealthier and they adopt more than constrained ones. Interestingly, while they adopt almost twice as much of the new variety seed in quantity, their adoption rate is not significantly different from that of constrained households that have dairy income (from Table 2.2).

To summarize, the descriptive statistics seem to indicate that while access to credit may hinder some households to adopt new potato varieties, credit constrained farmers engaging in dairy have significantly better adoption rates. This suggests that dairy income may be compensating, to some extent, for the lack of credit. Yet, even though dairy income may compensate for adoption **rates**, the **level** of adoption is lower for the credit constrained.

Section 2.5 presents the empirical findings and analyzes the impact of dairy on adoption.

2.5 Results and analysis

Dairy income and adoption

Estimation of equations (2.20), (2.21) and (2.22) is presented in Table 2.4. Indeed, dairy income positively and significantly affects adoption. This supports the hypothesis that diary (cash) income complements adoption of new varieties. However, the strength of these results depends to the extent that dairy income "relaxes the credit constraint"-and thus increases the cash liquidity for these farm-households- and is not the product of a "wealth effect." In other words, for our results to be consistent with our story of diversification for

Table 2.2: Household descriptive statistics for the credit constrained only by dairy activity

	No dairy activity	With dairy activity
Potato seeds planted		
New variety(kg.)	119	313*
Traditional (kg.)	277	251
Total (kg.)	396	564*
Adoption rate (%)	26	50*
Dairy (cash) activity		
Dairy income (soles)	0	2887*
Herd size (#)	1.8	7.4*
Dairy pickup stop (%)	85	79*
Total labor applied (days)	64	310*
Dairy refrigeration in the community (%)	33	55*
Inherited cows (#)	0.2	0.3
Wealth-assets		
Total land owned (hectares)	4.7	10.0*
Total farm assets (soles)	293	394*
Total household assets (soles)	459	709
Rooms (#)	2.0	2.2
Income sources		
Potatoes (%)	66	54*
Dairy (%)	0	36*
Off-farm (%)	31	9*
Other crops (%)	3	1*
Production characteristics		
Total arable land owned (hectares)	4.3	8.6*
Household labor force (#)	2.7	2.8
Household head age (years)	43	44
Household head education (years)	4.5	4.8
Transaction costs		
Distance to main market (km)	362	258*
Good road quality (%)	0	7*
CARE in the community (%)	72	83*
obs	46	187

* means that there is significant difference between the two groups at the 90% level or more

Table 2.3: Household descriptive statistics by credit constrained

	Constrained	Unconstrained
Potato seeds planted		
New variety(kg.)	275	470*
Traditional (kg.)	256	389*
Total (kg.)	531	859*
Adoption rate (%)	45	49
Dairy (cash) activity		
Dairy income (soles)	2317	3502*
Herd size (#)	6.3	9.3*
Total labor applied (days)	271	262
Dairy pickup stop (%)	80	58*
Dairy refrigeration in the community (%)	50	62*
Inherited cows (#)	0.3	0.5
% with cash activity	80	88*
Wealth-assets		
Total land owned (hectares)	9.0	11.6*
Total farm assets (soles)	374	541*
Total household assets (soles)	660	671
Rooms (#)	2.2	2.5*
Income sources		
Potatoes (%)	55	49
Dairy (%)	32	41
Off-farm (%)	12	10
Other crops (%)	1	1
Production characteristics		
Total arable land owned (hectares)	7.8	9.5*
Household labor force (#)	2.8	3.0*
Household head age (years)	43	45*
Household head education (years)	4.8	4.8
Transaction costs		
Distance to market (minutes)	278	241*
Good road quality (%)	6	11*
Member or CARE (yes/no)	74	64*
obs	233	253

* means that there is significant difference between the two groups at the 90% level or more

Table 2.4: Adoption of new variety for the credit constrained (2SLS)

	Adoption (A^c)	Dairy	
		Income (C)	Labor (l_C^c)
Dairy (cash) income ($\widehat{C}$)	**0.00008****		
Total labor applied to dairy activity ($\widehat{l_C^c}$)	-0.003		
Household labor force (#) ($\overline{L}$)	-0.0006	101	-0.6
Production characteristics (z^q)			
Total arable land owned (hectares)	-0.015	45***	1.1**
Household head age (years)	0.0009	58***	1.9***
Household head education (years)	-0.004	130*	6.1**
Transaction costs (p)			
Distance to market (minutes)	-0.003***	-5.7***	-0.5***
Good road quality (yes=1)	0.220***	-728	3.4
Dairy (cash) activity (z^c)			
Dairy refrigeration availability (yes=1)		890**	-11.0
Dairy pickup stop (yes=1)		1195**	54.4**
Inherited cows (#)		413***	4.3
Constant	1.73***	-1317	241***
obs	233	233	233
R^2	0.23	0.27	0.20

Adoption: $\frac{\textit{Quantity of new variety seed planted}}{\textit{Quantity of total seed planted}}$
Dairy income: income from dairy production
Dairy labor: labor used for dairy production
Significance levels:*:90%, **:95%, ***:99%
The Sargan test for overidentification cannot be rejected at the 95% confidence level

adoption, dairy income should have no impact on adoption decisions for credit unconstrained households. This would correspond with the predictions of our theoretical model and in particular with equation (2.12). We implement a counterfactual test by estimating equations (2.20), (2.21) and (2.22) for the credit unconstrained. Table 2.5 presents the results. As expected, for the credit unconstrained farm-households, dairy (cash) income does not affect adoption decisions.

Combining the above findings with the summary statistics from section 2.4, we

Table 2.5: Adoption of new variety for the credit unconstrained (counterfactual) (2SLS)

	Adoption (A^u)	Dairy	
		Income (C)	Labor (l_C^u)
Dairy (cash) income ($\widehat{C}$)	**-0.00006**		
Total labor applied to dairy activity($\widehat{l_C^u}$)	0.002		
Household labor force (#) ($\overline{L}$)	0.014	254*	2.0
Production characteristics (z^q)			
Total arable land owned (hectares)	0.0007	105***	1.6***
Household head age (years)	-0.001	96***	1.5***
Household head education (years)	0.003	384***	3.1
Transaction costs (p)			
Distance to market (minutes)	-0.002***	-4.8*	-0.3***
Good road quality (yes=1)	0.162	1951**	1.3
Dairy (cash) activity (z^c)			
Dairy refrigeration availability (yes=1)		144*	2.2
Dairy pickup stop (yes=1)		68**	37**
Inherited cows (#)		70*	3.3*
Constant	0.351	-3654***	224***
obs	253	253	253
R^2	0.21	0.36	0.18

Adoption: $\frac{\textit{Quantity of new variety seed planted}}{\textit{Quantity of total seed planted}}$

Dairy income: income from dairy production

Dairy labor: labor used for dairy production

Significance levels:*:90%, **:95%, ***:99%

The Sargan test for overidentification cannot be rejected at the 95% confidence level can conclude that while dairy enables credit constrained farmers to achieve desired rates of adoption by overcoming liquidity constraints, there are still significant differences in the level (quantity) of adoption between credit constrained and unconstrained. Therefore, dairy income has a wealth effect only to the extent that it allows farmers to produce a higher quantity for the new potato variety and it does not alter the production behavior of those who do not need it (credit unconstrained).

The cost of dairy

As we have shown, for the credit constrained, shifting resources away from potato production and into dairy production positively affects the adoption intensity of the new potato varieties. However, this shift of resources may come at a cost. Under our assumption that potatoes are a more profitable activity than dairy, it would imply that there is an excess resource use (and hence a cost) in producing dairy via the necessary reallocation of resources (and especially labor) to the dairy activity. That is, even though credit constrained households improve adoption rates by producing dairy, the fact that they shift labor to do that implies that they may be overproducing milk in order to ameliorate their ability to adopt.

One way to capture and test such an effect is to estimate the additional labor that is shifted towards the dairy activity by credit constrained households. To implement this, we use the estimated coefficients ($\widehat{\xi}$)and predicted error ($\widehat{\omega_2}$) from equation (2.24) to predict the dairy activity's labor demand for credit constrained households that also engage in the dairy activity. This demand can be interpreted as the quantity of labor that would have been used for dairy if these credit constrained households were not constrained, that is:

$$l_C^c|\ if\ unconstrained = l_C^c(x, w; \widehat{\xi}) + \widehat{\omega_2} \tag{2.25}$$

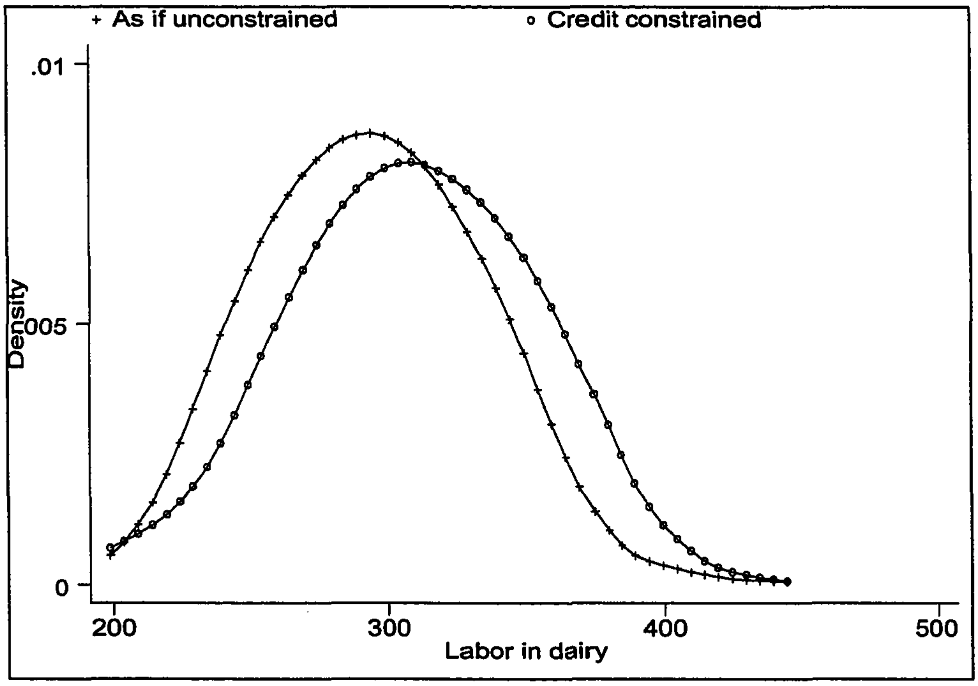
The hypothesis is that if we compare this labor demand to the predicted values of labor demand for dairy from equation (2.22), we should find that the latter predicted values are larger, implying that indeed credit constrained overproduce dairy (to reach higher adoption rates).

We compare the two distributions of the predicted labor demand functions in Figure 2.2. Indeed, on average, credit constrained households work more on the dairy activity as opposed to the amount they would use if they were not constrained (seen by comparing the peaks of the two curves). Perhaps a more interesting observation is the fact that the overall distribution of the simulated unconstrained labor demand is shifted to the left, indicating that indeed the credit constrained would use less labor in dairy. These simulations support our previous results to show that while resource reallocation is essential for adoption, it comes at a cost.

2.6 Conclusions

This paper explores the relationship between cash constraints, diversification and technology adoption. Building on a simple model of adoption, we show that income diversification can serve as an alternative source for cash liquidity that allows farmers with limited credit access to adopt new technologies. While this diversification comes at a cost, it provides an otherwise non-existent mechanism for these farmers to overcome market failures and adopt better technologies. In this context, diversification complements adoption.

Figure 2.2: Labor in dairy for credit constrained (observed and counterfactual)



The results show that dairy income, a new practice in the area of study, is not displacing potatoes. On the contrary, it complements the adoption of improved potato varieties for credit constrained farm-households. While programs to help farmers' adoption rates such as input-credit or input subsidies are important, facilitating farmers' ability to diversify income sources such as dairy income, at least in the short-run, seems to provide another channel from which farmers can achieve higher adoption rates. Further research to understand these complementarities between diversification and adoption is an important next step for implementing better policies and apprehending the adoption process and long-run diffusion of new technologies.

Chapter 3

Transactions Costs and the Role of Information: Evidence from Peru

3.1 Introduction

Unlike the zero transaction costs assumption in neoclassical economic theory, transaction costs economics (TCE) posits that agents making decisions on different types of actions do so in a costly way (Williamson [58]). For example, farmers deciding where to sell a particular crop will base their decision not only on the price they expect to receive in each market but also on additional costs related to transacting in these markets.

These transaction costs can be divided in two broader categories: variable and fixed. Variable transaction costs change according to how much a household sells (or buys). For example, the cost related to transporting a product to a market will depend on the quantity but also the time it takes to reach the market.

On the other hand, fixed transaction costs are independent of the quantities sold or bought. They can be further distinguished between information, bargaining and monitoring. Information costs occur before the exchange takes place and include aspects such as searching for attributes that could facilitate the transaction, finding prices, and searching for potential buyers. In addition, bargaining or negotiation costs take place during the exchange and can include the cost of time to reach an agreement, contractual and payment arrangements. The extent of which a person will be able to minimize these costs will be a factor of individual characteristics (education, skills, gender), product attributes like quality or the relationship between the agents participating in the transaction. Finally, monitoring can ensure that the conditions of an exchange are met (for example enforcing the payment schedule agreed or that the quality of the product is the correct one).

Based on the above, the main insight from TCE is that an agent engaging in a transaction will choose the strategy which will maximize overall benefits (Coase [8]). As such, this paper looks at the marketing decision patterns of net sellers of potatoes in the Peruvian Andes. Using transaction-specific data, the paper explains how differential transaction costs influence the selection of markets where the producers choose to sell. In particular, our findings show that besides price differentials and access to good road infrastructure, access to information about markets and prices, relationships with buyers and the farmers' bargaining abilities, also affect market choices. In addition, we find that bargaining explains a large percentage of the variation in prices received in the markets. These results offer an empirical example of the importance of transaction costs in marketing decisions and market integration. They suggest that policies aiming at reducing transaction costs should address not only road and infrastructure but also create mechanisms to enhance information flows. Furthermore, as the results indicate that different aspects of bargaining can affect transaction outcomes via both market participation and prices received, increasing bargaining power should also be a key issue for policy makers.

While the body of theoretical literature on transaction costs is extensive, the empirical literature on transaction costs has been lagging. As transaction costs are often unobserved, the empirical challenge has been to develop methodologies to measure them indirectly. Most data sets usually lack explicit information on variables relevant to these types of costs. With the exception of transaction costs attributes like distances to markets and transportation costs, aspects like market information or search and bargaining procedures are rarely included in most surveys. Therefore, much of the empirical work on transaction costs has focused on testing their existence but not on their actual measurement or their effect on behavior.

Nonetheless, there are a number of studies that try to empirically address these limitations. Cogan [9] estimates a model of labor supply for married women with fixed transaction costs associated with entry into the labor market. Estimates of these fixed transaction costs confirm his hypothesis that they are relevant. Earlier studies, on the contrary, had found large own-wage elasticities for married women. Cogan's results suggest that such large values may be due to ignoring the existence of fixed costs of labor market entry. Goetz [26] proposes a way of empirically estimating the supply response of coarse grains producers using data from rural household in Senegal. Recognizing that the decision of participation in the market is endogenous, Goetz first estimates a probit model to assess the probability of market participation and then uses the results to correct the estimation of the response of two distinct variables: quantity sold and quantity bought. He finds that the availability of better information about the market raises the probability of market participation. Hobbs [29] looks at the factors affecting the choice between different types of cattle marketing and concludes that various transaction costs in addition to farm characteristics affect the choice of marketing channel.

More recently, Key, Sadoulet and de Janvry [34] develop a theory of household behavior under the presence of both fixed and variable transaction costs. Applying their model to Mexican data of corn producers, they estimate a censored regression with unobserved threshold and confirm the importance of both variable and fixed transaction costs. Finally, Renkow and Hallstrom [49] develop a conceptual framework for quantifying fixed transaction cost of semisubsistence farm-households in Kenya and find that on average, these costs are equivalent to a 15 percent ad-valorem tax.

The paper proceeds as follows: section 3.2 discusses the decision process that farm-households face in allocating their marketed surplus to a market. Insights from the transactions specific survey are presented in section 3.3, while the results and subsequent analysis are in section 3.4. Section 3.5 concludes.

3.2 Decision timeline

In order to conceptualize the problem, consider an economy that produces an agricultural product. A farm-household's production decision process in this economy can be divided in three phases. Initially, during the planting season, farm-household i chooses the optimal allocation of resources to determine the total quantity to be produced. Following a typical farm-household setting, this decision will be based on the expected price of the product, available resources such as labor, land and other incomes.

The second decision phase, at the harvest season, entails household i's realization of the effective total quantity produced Q_i and an assessment of how to allocate it given available information (such as current product prices, market conditions). In particular, Q_i can be consumed or sold.[1] Denoting the quantity consumed c_i the marketed surplus for household i, q_i is given by:

$$q_i = Q_i - c_i \tag{3.1}$$

The final decision for the farm-household is where to sell the marketed surplus. Focusing on farm-households that are net-sellers (that is $q_i > 0$), and assuming that there exist J available markets where farm-household i can sell q_i, the farm-household's decision will be based on a number of factors. First, for farm-household i, selling in market j will be associated with variable transaction costs TC_{ij}^v. These costs are a function of the distance (d_{ij}) and time (m_{ij}) to reach market j, as well as other individual-specific characteristics that are associated with variable transaction costs (z_{ij}^v).[2] Variable transaction costs are thus given by:

$$TC_{ij}^v = TC^v(d_{ij}, m_{ij}, z_{ij}^v) \tag{3.2}$$

In addition, the farm-household will consider the expected price to be received at each candidate market j. In particular, the expected expected price $E[p_{ij}]$ to be received at market j is decomposed in:

$$E[p_{ij}] = E[\overline{p_j} + B(q_i, z_{ij}^b)] \tag{3.3}$$

[1]Of course, other options that could be incorporated in this decision are the possibility for storage or payment in kind. However, these do not add much to our analysis at this point so we omit them for simplicity.

[2]For example, owning a truck versus an animal affects both the quantity able to transport and the time to reach all the markets.

where $\overline{p_j}$ is a market specific exogenous expected price; and $B(q_i, z^b_{ij})$ is a potential price markup that farmer i expects to receive in market j. This markup depends on the quantity sold (q_i) as well as other bargaining related attributes such as bargaining ability, experience or product quality (z^b_{ij}). Notice that for the same farmer i these may differ across markets.

The expected price received and the variable transaction costs can be used to comprise the expected revenues R_{ij} associated with market j:

$$R_{ij} = q_i * (E[p_{ij}] - TC^v_{ij}) \tag{3.4}$$

Selling to market j is also associated with fixed costs $TC^f(z^f_{ij})$. Such costs are related to searching for the best market and buyer or obtaining information about prices. In addition, they are invariant to the specific quantity sold and may also be market specific. For example, knowledge of market specific attributes such as prices and relevant infrastructure (that could potentially affect the probability of finding a buyer) can influence the decision of a farm-household as to where to sell. Individual specific characteristics such as experience, education, gender and age could also enhance the ability to collect and analyze relevant information and thus can affect market choices. Finally, contractual agreements between a farmer and a specific buyer or information about types of contractual agreements available at different markets may also affect market choice.

Based on the above, farm-household i will sell q_i in the market that yields the highest net profits (Π_{ij}). Specifically, farm-household i will choose to sell in market j such that:

$$j : \arg\max\{\Pi_{ij} = R_{ij} - TC^f(z^f_{ij})\} \tag{3.5}$$

or in reduced form:

$$j : \arg\max\{\Pi_{ij} = q_i * (E[\overline{p_j} + B(q_i, z_{ij}^b)] - TC^v(d_{ij}, m_{ij}, z_{ij}^v)) - TC^f(z_{ij}^f)\} \tag{3.6}$$

As such, observing farm-household's i selling in market j implies that:

$$\Pi_{ij} > \Pi_{ik} \quad \forall j \neq k \tag{3.7}$$

Two questions can be tested empirically using this framework: (i) whether the farm-household's market choice is influenced by both variable (distance, time) and fixed (information, search) transaction costs (Eq. (3.6)); and (ii) how does bargaining affect the seller's prices received (Eq. (3.3)).

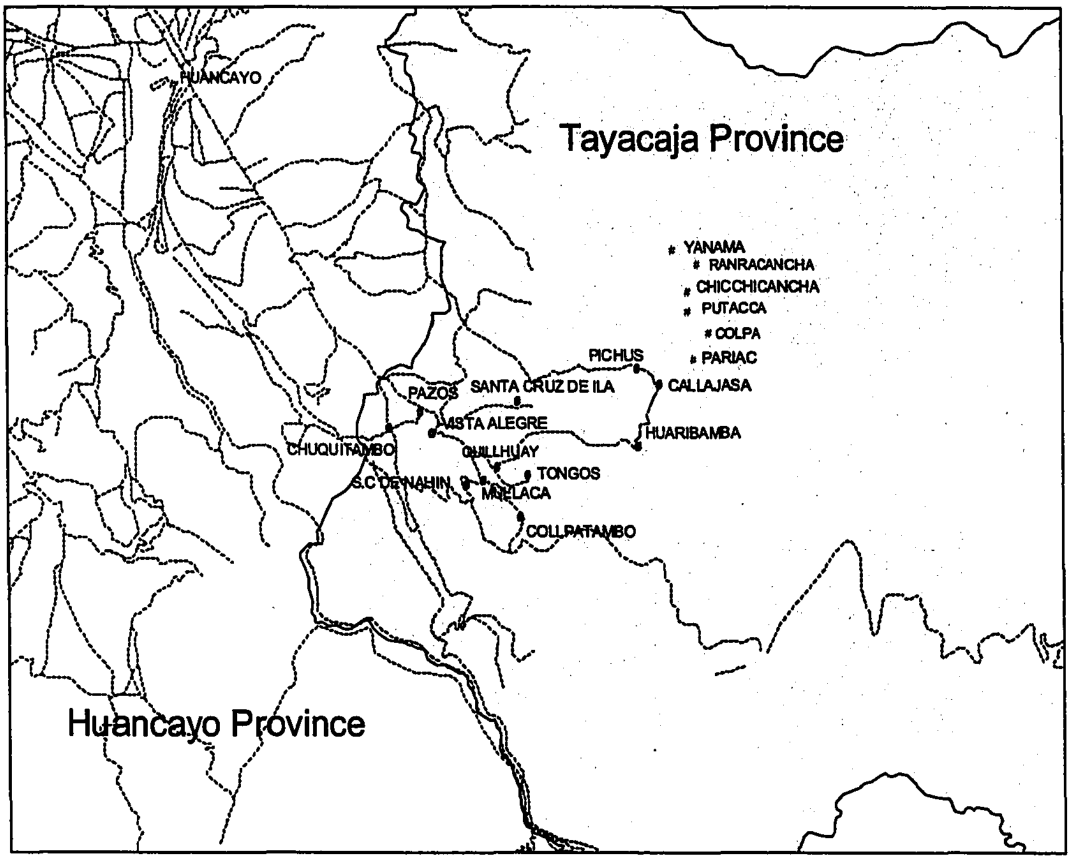
3.3 Transaction costs insights from rural Peru

Markets, transactions and location The data used in this paper, collected in early 2001, comes from a survey of 229 small-scale farm-households in the province of Tayacaja in Peru (Figure 3.1). The main objective of the survey, designed and implemented by Javier Escobal at the Grupo de Análisis para el Desarrollo (GRADE), was to study transaction costs. These farm-households are part of about 1500 households in the region dedicated mainly to potato production.[3] A number of different locations exist where they can sell potatoes: the farmgate, two local markets (Pazos and Pichus) and two distant ones (Huancayo and Lima).[4] The two local markets open twice a week while the two distant ones are open every day. Farmers can also arrange directly with buyers to come at the farmgate to

[3]For an extensive description of the region see Escobal [1].

[4]There are a number of other small markets where farmers could potentially transact, but are not included them in the analysis due to too few observations.

Figure 3.1: Survey map



- - - - - - Good roads

purchase their potatoes. Potential buyers also travel around these communities in search of potatoes, which allows informed farmers to sell at the farmgate. All markets are open markets with minimal infrastructure. However, the distant markets are larger and thus attract more sellers and buyers. The average quantities sold in these markets is also higher. As such, the fact that not everyone sells in those markets can be partly attributed to high transaction costs, as discussed next.

The transactions typically occur as follows: a seller is approached by a buyer (even though the opposite can also occur). The buyer proceeds to evaluate the condition and quality of the potatoes and then the two parties negotiate a unit price. If there is an agreement, they weight the quantity and conclude the transaction.[5]

Farm-households can also be distinguished based on their access to road infrastructure: about two thirds of the households have good road access, while the rest reside in a region with limited road network (North-East region in Figure 3.1). This latter group of farmers is also significantly farther away from the available markets (Table 3.1), and as seen below have fewer assets, produce less and are overall poorer than the farmers that have access to good roads.

Market destination Table 3.2 presents the average quantities sold per transaction in the different markets. In general, the average quantity sold in distant markets is significantly higher than the quantities sold locally or at the farmgate. In addition, the sales quantity per transaction is strongly correlated by the road access. Farm-households with good road access sell on average 3 times more per transaction than those with bad road access. As mentioned above, road access (and in effect the distance and time to reach a specific market) can play an important role in the market choice and as such, good road access enables farm-households to sell in distant markets more easily.

In fact, sales in distant markets represent more then two thirds of total sales for farm-households with good road access, as opposed to only 25 percent for farm-households

[5]In most cases, each party uses his own weighting scale out of concern that the other agent may tamper the machine.

Table 3.1: Sample structure, transactions and location

	Good road access	Bad road access	All
Number of farm-households	138	91	229
Sales transactions:			
Total number	671	533	1204
At the farmgate	291	43	334
In local markets	175	436	611
In distant markets	205	54	259
Average number per household:	5.5	6.7*	6.0
Average distance to (km):			
Pazos (local)	22	82*	47
Pichus (local)	45	48	27
Huancayo (distant)	79	139*	100
Lima (distant)	377	437*	426
Average time to (minutes):			
Pazos (local)	77	372*	101
Pichus (local)	187	152*	54
Huancayo (distant)	136	432*	148
Lima (distant)	436	732*	270

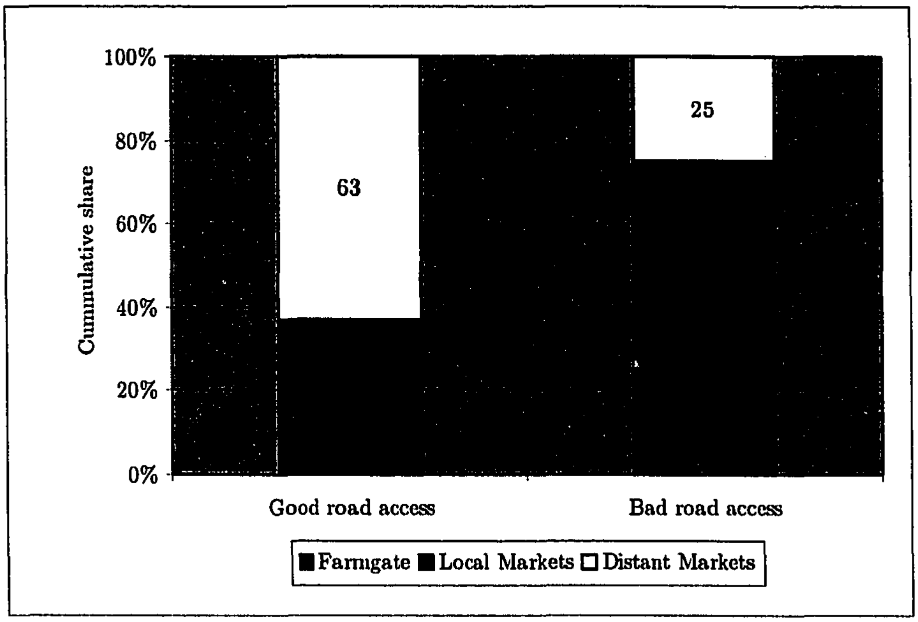
* significantly different from those with good road access at the 90% level or more.

Table 3.2: Quantities sold per transaction (by market and road access, in kilos)

	Good road access	Bad road access	All
Farmgate	5285	1787*	4555
Local markets:	3683	2375*	2750
Pazos	3683	3088*	3433
Pichus	n.a.	1914	1914
Distant markets:	10452	8128[1]	10236
Huancayo	8948	8128[1]	8821
Lima	15481	n.a.	15481
All	7108	2831*	5215

* significantly different from those with good road access at the 90% level or more.
[1] Small sample.

Figure 3.2: Distribution of total quantities sold, by market destination and road access



with bad road access (Figure 3.2). For farm-households with bad road access, almost 70 percent of the marketed surplus is sold in local markets, suggesting the importance of access to good roads in choosing markets.

The survey also included a question about why the farm-household chose the market where they sold. These perceptions signal not only what the farmers may value the most in terms of market choice, but indirectly the level of information they may have. The main reasons that farmers considered for making the market choice were: the expectation for higher prices, availability of more buyers and a higher trust level in the potential buyers in that market (Table 3.3). These results indicate that the expected price is not the only consideration for farmers in deciding where to sell. Instead, transaction related aspects like

Table 3.3: Perceptions by market and by road access

	Good Road access			Bad Road access		
	Farmgate	Local	Distant	Farmgate	Local	Distant
The farmer prefers to sell in this market because of (%):						
Higher prices	23	44	38	4	55	89
More buyers	30	26	33	0	30	0
More trust in buyers	27	20	25	4	5	11
Only option available	20	9	3	86	2	0
Other reasons	0	1	0	6	8	0
The price received was [...] compared to the expected price (%):						
Higher	5	8	3	0	0	0
Lower	30	26	25	53	19	3
Same	65	66	72	47	81	97

* significantly different from the group to its left at the 90% level or more.

** significantly different from the distant market at the 90% level or more.

the availability of buyers (that could signal lower transaction costs in terms of the higher probability to find a buyer), or the trust-worthiness for buyers in specific markets (reflecting information asymmetries) are also aspects that the farmers take into consideration to decide where to sell.

Transaction specific data Perhaps the most interesting aspect of the survey is the availability of transaction costs related information. In particular, the survey contains data on 1204 potato sales transactions in the five markets described above. As discussed earlier, one way to classify transaction costs is between transportation, information, bargaining and monitoring. First, transportation costs are expected to be higher the furthest a market is. Indeed, per kilo transportation costs are higher for transactions that take place in the more distant markets (Table 3.4). Specifically, the transportation costs for farm-households that sold in local markets was on average three cents per kilo, as opposed to eleven cents for those that sold in the two distant markets.

Table 3.4: Transportation costs by market and road access (in soles)

	Good road access		Bad road access		All	
	Total	Per kilo	Total	Per kilo	Total	Per kilo
Local markets:	204*	0.06*	58*	0.02*	100*	0.03*
Pazos	215	0.06	99	0.04	153	0.05
Pichus	n.a.	n.a.	27	0.02	27	0.02
Distant markets:	1121	0.12	309	0.04	1025	0.11
Huancayo	857	0.11	309[1]	0.04[1]	772	0.10
Lima	2003	0.15	n.a.	n.a.	2003	0.15
All	550	0.07	77	0.02	341	0.05

* significantly different from distant markets at the 90% level or more.
[1] Small sample.

Table 3.5: The role of information: knowing prices

	Among farmers who sold in:				
		Local		Distant	
	Farmgate	Pazos	Pichus	Huancayo	Lima
% of farmers who knew prices in:					
Farmgate	24	13	1	21	10
Pazos	60	60	8	64	32
Pichus	9	9	40	6	7
Huancayo	50	24	1	84	51
Lima	14	4	0	15	85
Obs	257	329	184	258	68

Prior to deciding where to sell, farm-households form expectations about the prices in each market. This will require time to collect information from different sources and knowledge about markets and seasonal idiosyncrasies that could affect prices. Interestingly, with the exception of the farmgate, the market that farmers chose to sell is the one that they are more likely to know the prices (Table 3.5). For example, among farmers that sold in Pazos, a local market, almost two thirds of the farmers knew the price in Pazos as opposed to only 24 percent for Huancayo. Similarly, 85 percent of those farmers that sold in Lima knew the price there, the highest among all other markets for these farmers. This strongly indicates the importance of information in deciding where to sell.

In addition, farmers that sell in distant markets seem to be more informed overall in terms of information about prices in different markets. (Table 3.5). By contrast, farmers that sell in local markets have information mainly about the prices of the market where they sell. One explanation is that, as discussed above (Table 3.1), most of the farmers that sell in local markets are farmers with bad road access to the distant markets. For them accessing these markets is too costly and therefore collecting information about them may not be as important. Consistent with the above is the fact that compared with households that sold in other markets, more farm-households that sold in distant markets reported that the price received was the one expected (Table 3.3). Still, very few farm households reported receiving a higher price than expected, while a third of the farm households reported receiving on average a lower price, suggesting the importance of information in forming expectations and thus influencing the market choice decision.

Another aspect of information costs is that related to search costs for finding market specific information or potential buyers. One would expect that this type of information to be more difficult or costly to collect for distant markets, as a person will generally know more about the place he resides. Interestingly, almost two thirds of the farm-households that sold in distant markets found the buyer prior to the transaction, compared with only a third of the farm-households that sold in local markets (Table 3.6). Still, finding a buyer is costly: half of the farmers that sold outside the farmgate needed on average more than two hours to sell their product, a third needed between one to two hours while the rest did not find a buyer in the same day, implying that the transaction costs for finding buyers are important.

Table 3.6: Transaction cost related variables by market destination

	Farmgate	Local	Distant
Information and search (ex ante)			
Found a buyer and fixed price before the sale (%)	37**	35	63*
Knew buyer (%)	64**	81*	88*
Buyer lives in same area (%)	35**	11*	0*
Time to sell in market (%)			
less than an hour	n.a.	12	16*
between 1-2 hours	n.a.	37	29*
more than 2 hours	n.a.	43	43
not in the same day	n.a.	8	11*
Bargaining and negotiation			
Number of negotiation rounds before agreeing on price	1.3**	1.5*	1.9*
Farmer bargained himself (%)	58**	79*	63*
Number of available buyers if sold at farmgate	2.7	n.a.	n.a.
Had problems agreeing on quality (%)	42**	29*	36*
Managed to agree on quality (%)	40**	18*	50*
Buyer paid with cash (%)	60**	65*	37*
Land owned (hectares)	4.8**	4.7	6.6*
Farm experience (years)	17**	19*	19
Improved variety (%)	65**	52*	82*
hh head age (years)	47**	50*	49
hh head education (years)	5.3**	5.3	5.7*
hh head gender (%)	95	91*	94*
hh head indigenous (%)	57	64*	57*
Monitoring and enforcement (ex post)			
Time to get paid (days)	3.0**	2.1*	4.5*
Number that farmer had to ask for payment (%)	1.5**	1.6*	2.0*
Confidence in buyer (1:worst, 10 best)	4.4**	4.6*	5.0*
Signed an agreement (%)	21**	28*	41*

* significantly different from the group to its left at the 90% level or more.

** significantly different from the distant market at the 90% level or more.

The ability to affect the price received will depend on a number of factors: the farmer's negotiating skills, the product's attributes (such as quality), or the relationship with the other party. The survey reveals a number of interesting insights. For example, farmers that sell in distant markets are wealthier, more educated and have more farming experience compared with those that sell in local markets (Table 3.6), implying that they may be better equipped to negotiate. In addition, most of the farmers negotiate the sales transaction themselves. Still, while the negotiation may be more effective in this way, it also comes at the cost of additional time spend negotiating.

The quantity sold can also can offer a bargaining advantage for a seller by lowering the search costs for the buyer. The quantity can also constrain a farmer from selling to a particular markets as it may not be profitable to sell there below a minimum quantity (in order to recover transportation or fixed costs). More than 80 percent of the farmers that sell in distant markets sell improved potato varieties, as opposed to 65 percent of those who sell at the farmgate and 50 percent of those who sell in local markets suggesting that potato quality may be more of a desirable attribute for buyers from distant markets rather than local ones (Table 3.6).[6]

Still, and reflecting higher fixed transaction costs, a third of all farmers had problems with agreeing on the quality of the product.[7] Even though some farmers reported that they managed to resolve these types of problems, the majority did not. As such, farmers may not only incur the time costs associated with the negotiation but also settle for a lower

[6] In fact, field observations revealed that native potato varieties have higher consumption incidences among local communities than the improved varieties.

[7] This entails the verification and subsequent agreement between the two parties that the potatoes are of a specific quality and variety (for example between good and bad condition, or betwen native or improved variety). While the variety is usually easier to verify, agreeing on the quality can be more challenging.

price if there is uncertainty about the availability of other buyers.

Finally, even after the transaction takes place, the farmer may still incur additional costs. For example, the farmer may enter an agreement to get paid in the future and as such, he can incur costs related to enforcing the sales agreement. Interestingly, for farmers that sold in distant markets, it took more than four days to get paid, compared to only two for those that sold in local markets (Table 3.6). This could be one explanation as to why 40 percent of the farmers that sold in distant markets signed a contractual agreement with the buyer.

3.4 Understanding market choices and the role of transaction costs

Market choice The empirical patterns in the previous section suggest that both fixed and variable transaction costs related variables such as transportation, information, bargaining and monitoring could be linked to farmers' market choice. Based on Eq. (3.6) from section 3.2, household i will choose market j to sell, as long as the net profits are the highest than any other market. The probability that individual i will choose market j is:

$$\text{Prob}_{ij} = \frac{\exp(\beta' x_{ij} + a_j' y_i)}{\sum_{k=i}^{J} \exp(\beta' x_{ik} + a_k' y_i)} \tag{3.8}$$

where x_{ij} is a vector of characteristics for market j as perceived by farmer i; y_i is a vector of individual characteristics for farmer i; α_j and β are coefficients to be estimated.[8]

[8]While β shows the effect of an attribute specific parameter (such as distance to a market), a_j captures the market specific impact of individual characteristics (such as the role of education) on the market choice.

Table 3.7 reports the results for a reduced form conditional logit of market selection (Eq. (3.8)).[9] The results indicate a number of interesting patterns. First, holding other variables constant, farmers are less likely to sell in any of the markets compared with selling at the farmgate. It is interesting to note that for the distant markets the effect is stronger. In other words, if these dummies can be interpreted as capturing the role of location (and therefore of variable transaction costs), they suggest that the further away a farmer is from a market, the more likely it is that he will choose to sell at the farmgate instead of that particular market.

In terms of variable costs, and as expected, the longer it takes to reach a specific market, the less likely it is to choose to sell in that market, reflecting the higher variable costs associated with reaching more distant markets. In addition, access to good roads makes it less likely for farmers to sell in local markets, compared to selling at the farmgate. This implies that buyers may be more willing to come at the farmgate if roads are better. Conversely, since transaction costs are lower for farmers with good road access they are more likely to sell in distant markets compared to selling at the farmgate.

A number of fixed transaction costs related variables also affect the market choice decision. For example, indigenous farmers are less likely to sell in distant markets and more likely to sell in local markets (as opposed to selling at the farmgate). This could be suggesting that fixed costs such as language barriers may be constraining the ability of indigenous farmers to integrate in some markets. It could also indicate that indigenous farmers face discrimination in some markets so that their overall participation is limited.

[9]We report relative risk ratios. Thus, for a market specific attribute (such as distance to market j), a coefficient greater (less) than 1 implies that a unit increase (decrease) for that attribute will increases (decrease) the likelihood of choosing the corresponding market.

Table 3.7: Market choice: conditional logit

		Market choice
Markets		
Pazos (Local market)		0.72
Pichus (Local market)		0.82
Huancayo (Distant market)		0.08***
Lima (Distant market)		0.01**
Quantity sold (predicted)		
Quantity * Pazos	q_i	1.0002***
Quantity * Pichus	q_i	1.0001
Quantity * Huancayo	q_i	1.0005***
Quantity * Lima	q_i	1.0007***
Quantity * Access to good road * Pazos	$q_i * z_{ij}^p$	0.9997***
Quantity * Access to good road * Pichus	$q_i * z_{ij}^p$	0.9999
Quantity * Access to good road * Huancayo	$q_i * z_{ij}^p$	0.9995***
Quantity * Access to good road * Lima	$q_i * z_{ij}^p$	0.9995**
Proportional Transaction Costs		
Distance to market (in km)	d_{ij}	1.007
Distance to market squared	d_{ij}	0.99997
Time to market (in minutes)	m_{ij}	0.99367***
Time to market squared	m_{ij}	1.000007
Access to good road * Pazos	z_{ij}^p	0.17***
Access to good road * Pichus	z_{ij}^p	0.01***
Access to good road * Huancayo	z_{ij}^p	2.43
Access to good road * Lima	z_{ij}^p	18.79
Bargaining and negotiation		
Indigenous * Pazos	z_{ij}^b	1.15
Indigenous * Pichus	z_{ij}^b	1.60**
Indigenous * Huancayo	z_{ij}^b	0.84
Indigenous * Lima	z_{ij}^b	0.53**
Improved variety * Pazos	z_{ij}^b	1.75**
Improved variety * Pichus	z_{ij}^b	0.61**
Improved variety * Huancayo	z_{ij}^b	1.63**
Improved variety * Lima	z_{ij}^b	2.58**
Experience * Pazos	z_{ij}^b	1.07***
Experience * Pichus	z_{ij}^b	1.07**
Experience * Huancayo	z_{ij}^b	1.05***
Experience * Lima	z_{ij}^b	1.08***
Information		
Farmers in village know the market price (%)	z_{ij}^f	3.09***
Date of sale		
Big harvest season * Pazos		1.42**
Big harvest season * Pichus		1.64**
Big harvest season * Huancayo		1.40**
Big harvest season * Lima		1.40
Observations		5480

Ommited market is farmgate.
Relative risk ratios reported
Significance of underlying parameters: * at 10%; ** at 5%; *** at 1%

Table 3.8: Quantity sold (OLS estimation)

Production characteristics	
Access to good road (yes=1)	4,744***
Farming experience (years)	-26
Indigenous (yes=1)	-622
Improved variety (yes=1)	1,500***
Date (at harvest=1)	939***
Number of adults	858***
Land owned (hectares)	1,267***
Household characteristics	
Number of children	351*
Number of elder	1,134***
Constant	-9,437***
Observations	1096
R-squared	0.42

Dependent variable: quantity sold (in kilos)
Significance of underlying parameters: * at 10%; ** at 5%; *** at 1%

Furthermore, experience (reflecting the ability to negotiate), makes farmers more likely to interact with all of the markets compared with selling at the farmgate. Finally, the higher the quantity the farmer has available to sell sold the higher the likelihood of selling in a particular market, especially in distant markets[10].

Knowing the prices in different markets can allow a farmer to make a more informed decision about where to sell. We thus expect a positive effect of this knowledge on the probability to sell in a specific market. Nonetheless, as collecting such information is likely to be endogenous to the market selection, for the estimation, we proxy this by the share of farmers in a farmer's village that knew prices in a specific market. Indeed, we find that a higher level of information about prices in a specific market increases the likelihood of selling in that market, corroborating the story that information is indeed crucial for market selection.

[10]Since the decision of how much to sell is likely to be simultaneously taken with the decision of where to sell, we instrument the quantity sold (q_i) using a number of production and household characteristics (Table 3.8). The identifying instruments were the number of kids and elder in the household.

In summary, while variable transaction costs are key elements for the market decision for potato farmers, a number of fixed transaction costs relating to bargaining and information are also important determinants for market choice. We now try to further explore these results in order to understand the relative importance of these effects.

Prices received and bargaining According to Eq. (3.3), the expected price received at a given market will depend on the actual market price observed and any additional markup that the farmer could obtain. Therefore, the ability to bargain with buyers will be a crucial factor in the determination of this market price. We thus expect prices received both within and across specific markets to vary significantly. Table 3.9 presents the prices received per kilo of potatoes by market and road access. Indeed, there is a more than thirty percent variation in the prices received, and the pattern remains even after decomposing the prices by market access.[11] One additional interesting point worth indicating is that the farmgate offers the lowest prices while the more distant market the highest. Still, these prices do not account for transportation costs so that comparisons across markets are not appropriate per se (the next section addresses this issue). These findings seem to corroborate the hypothesis that while there is a significant variation in prices received within a specific market, it is not entirely explained by market location or market access.[12]

To empirically explore the source of this price variation, we estimate a model of the price received at market j, as a function of farmer attributes that represent different aspects of bargaining ability. Since the choice of selling in a particular market j is not

[11] The prices reported in the survey are prices received at the market destination and at the time of the sale. Hence, to make them comparable within each market, these prices are deflated and expressed in constant prices of December 2000.

[12] We also compared the price variation based on quantities sold and whether the sale was near the harvest season or not and still found high variation in the prices.

Table 3.9: Prices received by market and road access (in soles/kilo)

	Good road access	Bad road access	All
Farmgate	0.20 (0.05)	0.17 (0.03)*	0.19 (0.05)
Local markets:	0.26 (0.07)	0.24 (0.08)*	0.25 (0.08)
Pazos	0.26 (0.07)	0.26 (0.02)	0.26 (0.05)
Pichus	n.a.	0.22 (0.02)	0.22 (0.02)
Distant markets:	0.34 (0.09)	0.35 (0.05)[1]	0.34 (0.08)
Huancayo	0.32 (0.08)	0.35 (0.05)[1]	0.32 (0.08)
Lima	0.41 (0.07)	n.a.	0.41 (0.07)
All	0.28 (0.09)	0.25 (0.08)*	0.26 (0.09)

* significantly different from the group to its left at the 90% level or more.
Standard deviation in parenthesis.
[1] Small sample.

random, as the farmer chooses to sell in the market j that leads to the highest expected net profits, we estimate a price equation correcting for market selection bias using the results from Eq. (3.8) above. Therefore, for each market j, we estimate an equation of the form:

$$p_{ij} = p(B(q_i, z_{ij}^b), \widehat{\lambda}_{ij}) \quad \forall \; j \tag{3.9}$$

where $\widehat{\lambda}_{ij}$ is a market selection correction term derived from equation (3.8).

The results show that a number of bargaining related attributes significantly affect the price received at the market (Table 3.10). For example, the quantity sold seems to positively affect the price received, reflecting the fact that the buyer could be willing to pay more if he can buy everything from the same seller (and thus minimizing his transaction costs of finding other sellers).

Given the above and the fact that we are more interested in assessing the relative importance, rather than the magnitude, of bargaining on the price received, we decompose the price variation and evaluate the relative importance of different bargaining aspects. In particular, we classify the explanatory variables between those that relate to the exogenously

Table 3.10: Explaining prices received (by market)

		Farmgate	Local		Distant	
			Pazos	Pichus	Huancayo	Lima
Quantity sold (in kg.)	q_i	-3.2	-0.2	12.1	60.9*	-39.3
Hh head farm exper. (years)	z_{ij}^b	0.1	-9.2	-12.5	-10.1	-13.2
Hh head farm exper. squared	z_{ij}^b	-0.2	0.2	0.5	0.4	0.2
Improved variety (yes=1)	z_{ij}^b	10.7	-21.2	-92.3**	-174.2	-51.1
Knew buyer (yes=1)	z_{ij}^b	39.3	8.7	0.4	-172.1	-97.1
Harvest season (yes=1)		-17.1	-112.7***	-274.4***	-83.7	6.9
Selectivity	$\hat{\lambda}_{ij}$	-175.7***	-70.9	95.4**	47.1	66.4
Constant		-1399***	-1133***	-1478***	-1367	-358.7
Observations		257	329	184	258	68
R-squared		0.12	0.08	0.58	0.18	0.11
F-test		4.94***	2.18**	25.84***	9.01***	2.24*

Dependent variable: log of price received
Coefficients reported are multiplied by 1000
Significance of underlying parameters: * at 10%; ** at 5%; *** at 1%

given market prices and those that reflect bargaining (such as ability, product attributes or relationship with buyers). We keep the variation observed due to the market selection term separate. Therefore, for each market j, we decompose the total explained variance of the estimated price received in each market (i.e.$Var(\widehat{p_{ij}}) = Var(\widehat{\gamma}w)$) in M components (4 in our case):

$$Var(\sum_{m=1}^{M}(\widehat{\gamma}w)_m) = \sum_{m=1}^{M} Var(\widehat{\gamma}w)_m + \sum_{m=1}^{M}\sum_{\substack{k=1 \\ k\neq m}}^{M} Cov((\widehat{\gamma}w)_m(\widehat{\gamma}w)_k) \tag{3.10}$$

where $\widehat{\gamma}$ and w are the parameter estimates and explanatory variables from Eq. (3.9). Each component m represents a subset of w that captures a particular aspect of the price determination process. For example $m = 1$ could reflect those variables that affect bargaining via individual skills while $m = 2$ the attributes of the product (such as quality).

The decomposition reveals that while for distant markets and the farmgate the role of bargaining is important, it is not true for local markets (Table 3.11). In particular, for the distant markets, more than 70 percent of the price variation can be attributed to variation

Table 3.11: Variation of price received by component

	Farmgate	Local		Distant	
		Pazos	Pichus	Huancayo	Lima
Market (time of sales)[1]	0.02	0.92	0.91	0.18	-0.01
Bargaining (ability)[2]	0.38	0.02	0.00	0.38	0.52
Bargain (product attributes)[3]	0.00	0.03	0.06	0.42	0.17
Selectivity	0.61	0.03	0.04	0.02	0.31

[1] Harvest season

[2] Hh head farn experience, knew buyer

[3] Improved variety, quantity sold

in aspects related to bargaining. For Huancayo, one of the distant markets, individual ability, the types of relationships with buyers and product-specific traits are all important in explaining the price variation. Interestingly, for Lima which is the biggest market but also the furthest away, more than half the price variation is due to bargaining ability. By contrast, the price variation in local markets is mainly explained by non-bargaining aspects like the time of the sale. One explanation is that local markets have overall fewer buyers and as such sellers may face monopsonistic pricing which may leave little room for bargaining. By contrast, in the bigger distant markets where there are more potential buyers, bargaining will be a crucial mechanism for receiving higher prices, especially due to the high level of asymmetric information about both prices and buyers.

Revisiting the market choice: a semi-structural approach The analysis so far has established that transaction costs are indeed present and significantly affect farmers' behavior. While the last exercise explored the role of bargaining, the final exercise presented below compares the relative importance between information and expected prices. In particular, we want to assess whether, in addition to expected prices, information (and to that

Table 3.12: Transportation costs (estimations by market)

		Local		Distant	
		Pazos	Pichus	Huancayo	Lima
Access to good road (yes=1)	z_{ij}^p	4.14	-20.68**	1.78	0.03
Distance to market (km)	d_{ij}	-0.024	0.033**	0.015	-0.02
Distance to market squared	d_{ij}	-0.0016**	0.0058***	0.0005**	-0.0005
Time to market(in minutes)	m_{ij}	0.06***	-0.007	0.004	0.007
Time to market squared	m_{ij}	-0.00014***	-0.0003***	-0.0001	-0.0006
Big harvest season (yes=1)		-1.12	1.44	2.13**	0.03
Selectivity	$\hat{\lambda}_{ij}$	-5.48**	8.55**	-0.25	-1.01**
Constant		-4.93	-23.34***	-6.91	3.73
Observations		329	184	258	68
R-squared		0.20	0.35	0.06	0.41

Dependent variable: log of transportation costs paid
Significance of underlying parameters: * at 10%; ** at 5%; *** at 1%

extent fixed transaction costs) is a key determinant for market choice. Unlike the reduced form analysis above, in this case we want to compare the relative importance between the two.

To implement the above, we use the regressions on prices received to predict (expected) prices for each farmer and market. In addition, as these prices are not comparable across markets, we discount them by substracting the unit costs related to transporting potatoes in the particular market. This entails estimating a transportation costs model, based on Eq. (3.2) and predicting the transportation costs for all markets associated with each farmer (Table 3.12).[13] For each farmer, we then calculate the price net of transportation costs ($\widehat{p}_{ij} - \widehat{TC}^v_{ij}$), which are now comparable across markets (Table 3.13). As expected, these net prices are more similar across the different markets than the prices received at the market (Table 3.9). Interestingly, even after discounting for transportation costs, the price variation in prices is still high (about 50 percent).

[13]We also correct for market selection.

Table 3.13: Prices received net of transportation costs by market and road access (in soles/kilo)

	Good road access	Bad road access	All
Farmgate	0.20 (0.05)	0.17 (0.03)*	0.19 (0.05)
Local markets:	0.20 (0.10)	0.22 (0.08)*◇◇	0.21 (0.09)◇◇
Pazos	0.19 (0.10)	0.23 (0.03)*	0.21 (0.07)
Pichus	n.a.	0.21 (0.03)	0.21 (0.03)
Distant markets:	0.22 (0.19)◇	0.31 (0.05)*◇	0.23 (0.13)◇
Huancayo	0.21 (0.15)	0.31 (0.05)*	0.23 (0.14)
Lima	0.26 (0.28)	n.a.	0.26 (0.28)
All	0.21 (0.14)	0.22 (0.08)*	0.22 (0.12)

* significantly different from the group to its left at the 90% level or more.
◇significantly different from local and farmgate at the 90% level or more.
◇◇significantly different from farmgate at the 90% level or more.
Standard deviation in parenthesis.

With the net prices calculated, we can re-estimate a semi-structural market selection model of Eq. (3.6):

$$\text{Prob}_{ij} = f(q_i, \widehat{p}_{ij} - \widehat{TC}^{v}_{ij}, z^{f}_{ij}) \tag{3.11}$$

The main result is that while as expected, higher (net) prices significantly increase the likelihood of selling in a market, information about prices (capturing the impact of fixed transaction costs) is also important (Table 3.14). In addition, comparing the relative risk ratios between expected prices and information, knowledge of market prices is as important as expected prices.

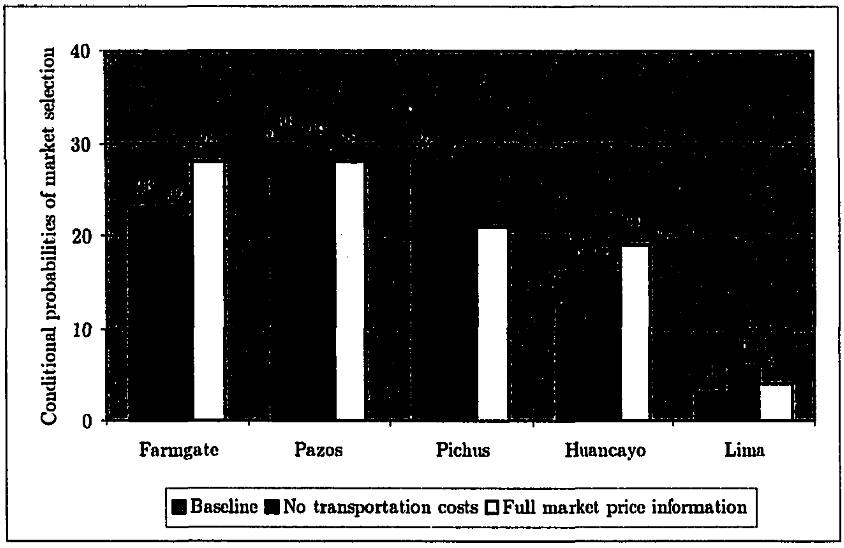
In order to further explore these results, we simulate the effect on the conditional probability for market selection from reducing transportation costs to zero or having complete information about prices (Figures 3.3 and 3.4). The two figures suggest that, on average, reduction in information gaps have a stronger impact on the conditional probabilities than changes in transportation costs. For example, among farmers with bad road access, eliminating all transportation costs would leave the probability to sell in Huancayo

Table 3.14: Revisiting market choice: semi-structural conditional logit

		Market choice
Markets		
Pazos (Local market)		0.995
Pichus (Local market)		1.208
Huancayo (Distant market)		0.302***
Lima (Distant market)		0.031***
Quantity sold (predicted)		
Quantity * Pazos	q_i	0.9998***
Quantity * Pichus	q_i	0.9997***
Quantity * Huancayo	q_i	1.0001***
Quantity * Lima	q_i	1.0002***
Effective price		
Predicted net price received (soles)	$\hat{E}[p_{ij}]-\hat{TC}_{ij}^{p}$	7.23***
Information		
Farmers in village know the market price (%)	z_{ij}^{f}	5.58***
Date of sale		
Harvest season * Pazos		1.87***
Harvest season * Pichus		2.41***
Harvest season * Huancayo		1.95***
Harvest season * Lima		1.66*
Observations		5480

Omitted market is farmgate
Relative risk ratios reported
Significance of underlying parameters: * at 10%; ** at 5%; *** at 1%

Figure 3.3: The effect of transportation costs and price information on market participation (Bad road access)

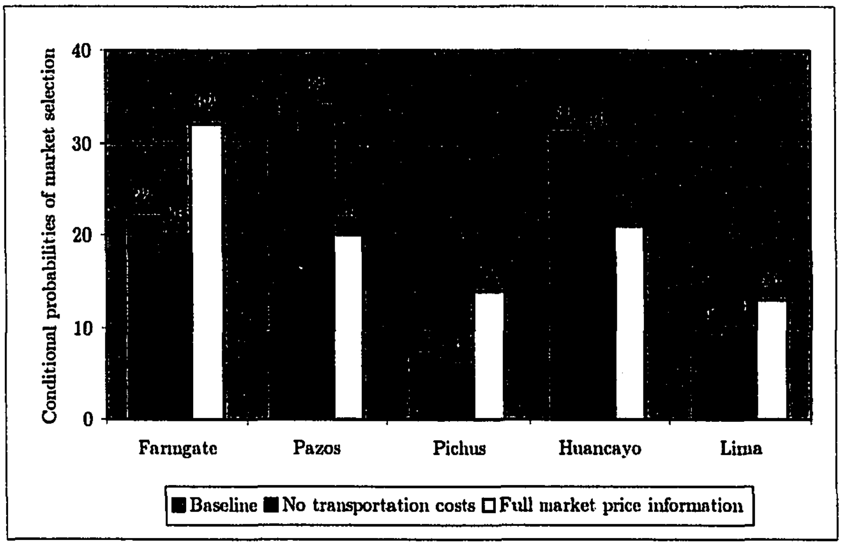


(a distant market) to 16 percent (Figure 3.3). By contrast, access to full information on prices would increase the probability to 19 percent. Similarly, we expect that lower transaction costs will decrease the conditional probabilities for selling in local markets since total costs for the distant markets would decrease. Indeed, among farmers with bad road access, the probability to sell in Pichus (a local market) would decline from 28 to 27 percent in the absence of transportation costs and to 21 percent in the presence of price information. Our overall interpretation of these figures is that fixed transaction costs in the form of information are at least as crucial for market participation as variable transaction costs.

Finally, using the predicted probabilities above, we can calculate the change in the distribution of total sales between farmgate, local and distant markets (the analog of Figure 3.2, weighting quantities with the predicted probabilities).[14] The results, illustrated

[14] Specifically, Figure 3.2 sums the total actual sales among all farmers and for each market j: $\sum_{i=1}^{N} q_{ij} * \delta$ with $\delta = 1$ if a farmer i sold in market j, 0 otherwise. By contrast, Figures 3.3 and 3.4 present the marketed surplus distribution by calculating $\sum_{i=1}^{N} q_{ij} * \widehat{p_{ij}}$ for each market j, where $\widehat{p_{ij}}$ is the conditional probabilities

Figure 3.4: The effect of transportation costs and price information on market participation (Good road access)



in Figures 3.5 and 3.6, reinforce the patterns above. For example, among farmers with bad road access, eliminating transportation costs or providing full information on market prices increases the share sold at the farmgate or in distant markets while decrease the share sold in local markets. Still, the effect of access to information is stronger.

for selecting market j.

Figure 3.5: Simulated distribution of total quantities sold due to changes of transportation costs and price information (Bad road access)

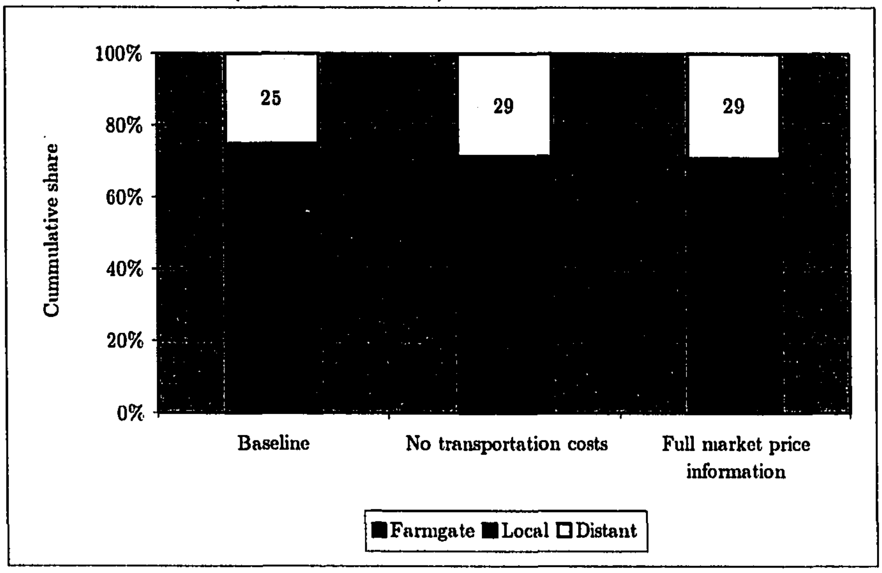
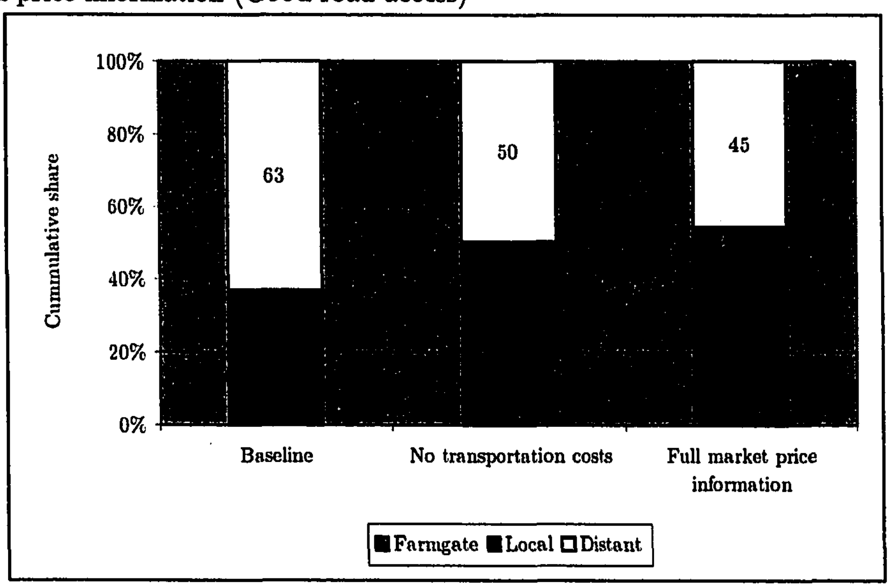


Figure 3.6: Simulated distribution of total quantities sold due to changes of transportation costs and price information (Good road access)



3.5 Conclusions

While transaction costs are difficult to measure, understanding the impact they have on behavior is crucial as it can inform policy design. This paper shows how different types of transaction costs influence decisions and outcomes for farm-households in rural Peru. We find that in addition to variable costs such as the distance to reach a market or access to good roads, transaction costs attributes like information about prices, relationships with potential buyers or bargaining abilities are also important determinants of market selection. Furthermore, bargaining explains a large share of the market specific price variation observed in the data, complimenting its role in the transaction process. Finally, simulations on the effect of reducing different types of transaction costs imply that farmers are more responsive to changes in their access to information than variable costs, offering a quantitative comparison between the two types of transaction costs.

In terms of policy, these findings suggest that if the goal is to reduce transaction costs, policy makers should address both variable and fixed costs by creating mechanisms that not only improve physical infrastructure, but also introduce tools that enhance information flows and market integration. In the case of farmers in Tayacaja, while ameliorating the existing road infrastructure is crucial, the results suggest that policies aiming to increase the availability of information and facilitate bargaining are perhaps more important. For example, the creation of cooperative-like schemes that can allow farmers market large quantities could be a feasible option to lower the fixed transaction costs associated with reaching distant markets and could also increase farmers' bargaining ability via the larger quantities offered. Improving language abilities among these farmers (in particular Spanish), could also increase the participation rates of indigenous farmers in the distant markets where indigenous dialects may be less used and reduce any existing discrimination based on language differentials. Finally, setting up local committees for collecting and disseminating market related information such as prices could be an easy mechanism to mitigate information gaps and allow farmers make informed choices.

Bibliography

[1] Javier Escobal D' Angelo. Costos de transacción en la agricultura peruana. *Working Paper 30, Grupo de análysis para el desarollo*, 2000.

[2] Christopher Barrett, Shane Sherlund, and Akinwumi Adesina. Shadow wages, allocative inefficiency and labor supply in smallholder agriculture. *Working Paper*, 2000.

[3] Michael Basch and Ricardo Paredes-Molina. Are there dual labor markets in chile?: empirical evidence. *Journal of development economics*, 50:297–312, 1996.

[4] Dwayne Benjamin. Household composition, labor markets and labor demand: testing for separation in agricultural households models. *Econometrica*, 60:287–322, 1992.

[5] David A. Binder. Estimating mixtures of normal distributions and switching regressions. comment. *Journal of the American statistical association*, 73(364):746–747, 1978.

[6] Peter Bryant. Estimating mixtures of normal distributions and switching regressions. comment. *Journal of the American statistical association*, 73(364):748–749, 1978.

[7] B.R. Clarke and C.R. Heathcote. Estimating mixtures of normal distributions and

switching regressions. comment. *Journal of the American statistical association*, 73(364):749–750, 1978.

[8] R.H. Coase. The nature of the firm. *Economica*, (4):386–405, 1937.

[9] John F. Cogan. Fixed costs labor supply. *Econometrica*, 49(4):945–963, 1981.

[10] Karen Conway and Jean Kimmel. Male labor supply estimates and the decision to moonlight. *Labor economics*, 5:135–166, 1998.

[11] C. Crissman and J. Uquillas. Seed potato systems in ecuador: a case study.

[12] Alain de Janvry, Marcel Fafchamps, and Elisabeth Sadoulet. Peasant household behavior with missing markets: some paradoxes explained. *Economic Journal*, 101(409):1400–1417, 1991.

[13] A. Dempster, N.M. Laird, and D.B. Rubin. Maximum likelihood from incomplete data via the e-m algorithm. *Journal of the Royal Statistical Society*, 39:1–38 (with discussion), 1977.

[14] E. Dimara and D. Skuras. Adoption of new tobacco varieties in greece:impacts of empirical findings on policy design. *Agricultural economics*, 19:297-307, 1998.

[15] M. Eswaran and A. Kotwal. Access to capital and agrarian production organization. *Economic Journal*, 96:482–498, 1986.

[16] M. Eswaran and A. Kotwal. Credit as insurance in agrarian economies. *Development economics*, 31:37–53, 1989.

[17] M. Eswaran and A. Kotwal. Implications of credit constraints for risk behavior in less developed countries. *Oxford economic papers*, 42:373–82, 1990.

[18] B.S. Everitt and D.J. Hand. *Finite mixture distributions.* Chapman and Hall, New York, 1981.

[19] Ellis F. Household strategies and rural livelihood diversification. *The journal of development studies*, 35(1).

[20] G. Feder. Adoption of interrelated agricultural innovations: complementarity and the impact of risk, scale and credit. *American journal of agricultural economics*, 64:94–101, 1982.

[21] G. Feder and M. Huppi. The role of groups and credit cooperatives in rural lending. *World Bank research observer*, 5(2):187–204, 1990.

[22] G. Feder, R. E. Just, and D. Zilberman. Adoption of agricultural innovations in developing countries: a survey. *Economic development and cultural change*, 33(2):255–298, 1985.

[23] G. Feder and D. L. Umali. The adoption of agricultural innovations:a review. *Technological foercasting and social change*, 43:215–239, 1993.

[24] Gershon Feder, Lawrence Lau, Justin Lin, and Xiaopeng Luo. The relationship between credit and productivity in chinese agriculture: a microeconomic model of disequilibrium. *American Journal of Agricultural Economics*, 1990.

[25] Edward B. Fowlkes. Estimating mixtures of normal distributions and switching re-

gressions. comment. *Journal of the American statistical association*, 73(364):747–748, 1978.

[26] Stephan J. Goetz. A selectivity model of household food marketing behavior in sub-sahara africa. *Journal of agricultural economics*, 74(2):444–452, 1992.

[27] Michael J. Hartley. Estimating mixtures of normal distributions and switching regressions. comment. *Journal of the American statistical association*, 73(364):738–741, 1978.

[28] D. Hiebert. Risk, learning and the adoption of fertilizer responsive seed varieties. *American journal of agricultural economics*, 56:764–768, 1974.

[29] Jill E. Hobbs. Measuring the importance of transaction costs in cattle marketing. *American Journal of Agricultural Economics*, 79:1083–1095, 1997.

[30] David W. Hosmer. Estimating mixtures of normal distributions and switching regressions. comment. *Journal of the American statistical association*, 73(364):741–744, 1978.

[31] Hanan Jacoby. Shadow wages and peasant family labor supply: an econometric application to the peruvian sierra. *Review of Economic Studies*, 60:903–21, 1993.

[32] Norman L. Johnson. Estimating mixtures of normal distributions and switching regressions. comment. *Journal of the American statistical association*, 73(364):750, 1978.

[33] A. Kaliba, A. Featherstone, and D. Norman. A stall-feeding management for im-

proved cattle in semiarid central tanzania: factors influencing adoption. *Agricultural economics*, 17:133–146, 1997.

[34] Nigel Key, Elisabeth Sadoulet, and Alain de Janvry. Transaction costs and agricultural household supply response. *American Journal of Agricultural Economics*, 82:245–59, 2000.

[35] Nicholas M. Kiefer. Discrete parameter variation: Efficient estimation of a switching regression model. *Econometrica*, 46:427–434, 1978.

[36] Nicholas M. Kiefer. Estimating mixtures of normal distributions and switching regressions. comment. *Journal of the American statistical association*, 73(364):744–745, 1978.

[37] Sylvie Lambert and Thierry Magnac. Measurement of implicit prices of family labour in agriculture: an application to cote d'ivoire. *in Agricultural household modeling anbd family economics,France Caillavet and Hervé Guyomard and Robert Lifran, eds.*, pages 9–24, 1994.

[38] Lung-Fei Lee and Robert Porter. Switching regression models with imperfect sample separation information-with an application on cartel stability. *Econometrica*, 52(2):391–418, 1984.

[39] Ramón Lopez. Estimating labor supply and production decisions of self-employed farm producers. *European Economic Review*, 24:61–82, 1984.

[40] G.S. Maddala. *Limited-dependent and qualitative variables in econometrics.* University press, Cambridge, 1983.

[41] S.K. Misra, D.H. Carley, and S.M. Fletcher. Factors influencing southern dairy farmer's choice of milk handlers. *Journal of applied econometrics*, 25:197–207, 1993.

[42] A. Monares. Analytical framework for design, and assessment of potato seed programs in developing countries.

[43] Jonathan Morduch and Hal Stern. Using mixture models to detect sex bias in health outcomes in bangladesh. *Journal of econometrics*, 77:259–276, 1997.

[44] Marc Nerlove, Stephen Vosti, and Wesley Basel. Role of farm-level diversification in the adoption of modern technology in brazil. *Research Report No.104 IFPRI*, 1996.

[45] Steven Were Omamo. Transport costs and smallholder cropping choices: an application to siaya district, kenya. *American Journal of Agricultural Economics*, 80:116–123, 1998.

[46] Richard Paap and Herman van Dijk. Distribution and mobility of wealth of nations. *European economic review*, 42:1269–1293, 1998.

[47] Richard E. Quandt and James B. Ramsey. Estimating mixtures of normal distributions and switching regressions. *Journal of the American statistical association*, 73(364):730–738, 1978.

[48] Ganesh P. Rauniyar and Frank M. Goode. Technology adoption on small farms. *World Development*, 20(2):275–282, 1992.

[49] Mitch Renkow and Daniel G. Hallstrom. Rural infrastructure, transaction costs and market participation in kenya. *Mimeo. Department of Agricultural and Resource Economics, North Carolina State University*, 2002.

[50] John Roemer. *A general theory of exploitation and class.* Harvard university press, Cambridge, 1982.

[51] Elisabeth Sadoulet, Alain de Janvry, and Catherine Benjamin. Household behavior with imperfect labor markets. *Industrial Relations*, 37(1):84–108, 1998.

[52] A. Saha, H.A. Love, and R. Schwart. Adoption of emerging technologies under output uncertainty. *American journal of agricultural economics*, 76:836–846, 1994.

[53] I. Singh, L. Squire, and J. Strauss. *Agricultural household models, extensions, applications and policy.* John Hopkins University Press, Baltimore and London, 1986.

[54] Emmanuel Skoufias. Using shadow wages to estimate labor supply of agricultural households. *American journal of agricultural economics*, 76:215–227, 1994.

[55] John Strauss, Mariza Barbosa, Sonia Teixeira, Duncan Thomas, and Raimundo Gomes Junior. Role of education and extension in the adoption of technology: a study of upland rice and soybean farmers in central-west brazil. *Agricultural economics*, 5:341–359, 1991.

[56] Graham Thiele. Informal potato seed systems in the andes: why are they important and what should we do with them? *World Development*, 27(1):83–99, 1999.

[57] D. Titterington, A. Smith, and V. Makov. *Statistical analysis of finite mixture distributions.* John Wiley and sons, New York, 1985.

[58] O.E. Williamson. *Economic organization: firms, markets and policy control.* Hertfordshire, UK: Harvester Wheatsheaf, 1986.

Appendix A

Regime Classification

In order to classify farm households in labor participation regimes, we first observe that for a farm-household that does not participate in the labor market at all, that is, for a self-sufficient (in labor) farm-household:

$$h = l^o = 0$$

Then, from the Kuhn-Tucker conditions (equations (1.10), (1.11) and (1.12)), the quantity of labor on-farm $\underline{l}^i$ is given by solving:

$$U_1(pq_L(\underline{l}^i, A), E - \underline{l}^i)\ pq_L(\underline{l}^i, A) = U_2(pq_L(\underline{l}^i, A), E - \underline{l}^i) \tag{A.1}$$

Equation (A.1) states that the household will allocate labor for on-farm activities until the value of the marginal product of an additional unit of work equates the marginal utility of income. The shadow wage $\underline{w}$ can be defined as the value of the marginal product of labor at $\underline{l}^i$:

$$\underline{w} = pq_L(\underline{l}^i, A) = \underline{w}(p, E, A) \tag{A.2}$$

Sadoulet et al. [51], show that $\underline{w}$ is an increasing function of A and a decreasing function of E :

(i) $\frac{\partial \underline{w}}{\partial A} > 0$

(ii) $\frac{\partial \underline{w}}{\partial E} < 0$

Define A_o to be the farm size such that:

$$pq_L(0, A_0) = w^o$$

Then, farm-households are classified in regimes as follows:

1. Worker

   If farm size $A < A_o$, $(i),(ii) \Rightarrow pq_L(l^i, A) < w^o$. Then, $l^o > 0, l^i = h = 0$. In this case, it is not worthwhile to engage in farm activities as the market wage $w^o$is higher. The household will only work off-farm in the labor market. If the labor constraint for this household is binding, the household will only allocate $\bar{L}$ in the market. If the constraint is not binding then it will allocate all the available labor.

2. Net seller

   If $A > A_o$ and $\underline{w} < w^o$ and using (ii): $l^o > 0$, $l^i > 0$, $h = 0$.

   For A small enough, employing all family labor in the farm would make the marginal productivity of labor lower than the effective wage w^o, so that it becomes convenient to divert part of the available family labor from farm to off-farm activities until $\underline{w} = w^o$. However, it is at this point that the maximum constraint, $\bar{L}$, becomes relevant: if the excess supply of family labor is smaller that $\bar{L}$, the household will sell in the labor market all the excess labor and the decision price will be the market price w^o. If, instead, the excess supply is larger than $\bar{L}$, the household will sell on the market all the labor it can, $\bar{L}$, and then supply more labor in the 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.

3. Self-sufficient

If $A > A_o$ and $w^o < \underline{w} < w^h$ and using (ii): $l^o = 0$, $l^i > 0$, $h = 0$.

For intermediate values of A, the household will find itself to be self sufficient in labor use. Employing all available family labor in the farm will make the marginal product of labor low enough to make unprofitable the hiring of labor at the market wage w^h, but not low enough to make it convenient to divert family labor from farm operation to market. The decision price will thus be a *shadow price*, whose value is bounded above by w^h and below by w^o.

4. Net buyer

If $A > A_o$ and $\underline{w} > w^h$ and using (ii): $l^o = 0$, $l^i > 0$, $h > 0$.

For A relatively large, the household will be a net buyer of labor. Even employing all available family labor on farm, marginal productivity of labor is still higher than the wage rate w^h, so that it is profitable to hire labor. The decision price to determine how much labor to utilize in the production activity, and thus how much to produce and how much to work will be the *market price* w^h.

Appendix B

Mixture Distributions

Consider the three equation regression model:

$$l^1 = l_1(\mathbf{x}_1;\beta) + u_1 \tag{B.1}$$

$$l^2 = l_2(\mathbf{x}_2;\gamma) + u_2 \tag{B.2}$$

$$\lambda = \lambda(\mathbf{x}_\lambda;\xi) + u_\lambda \tag{B.3}$$

where $u_1 \sim \mathrm{N}(0,\sigma_1^2)$, $u_2 \sim \mathrm{N}(0,\sigma_2^2)$ and $u_\lambda \sim \mathrm{N}(0,\sigma_\lambda^2)$ are i.i.d. disturbances, and where l^1, l^2 and λ are *latent* variables. Suppose that the *observed* variable l^i is defined by:

$$l^i = \begin{cases} l_1(\cdot) & \text{if } \lambda \leqq 0 \\ l_2(\cdot) & \text{if } \lambda > 0 \end{cases} \tag{B.4}$$

Then the problem is that of estimating the parameters $\{\beta;\gamma;\xi;\sigma_1,\sigma_2,\sigma_\lambda\}$ from the sample of N observations on $\{l^i,\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_\lambda\}$, for $i = 1,2,\ldots,N$.

Let

$$f_n(l^i) = \frac{1}{(2\pi)^{1/2}\sigma_n} \exp\left\{-\frac{1}{2}(l^i - \mathbf{x}_n{}'\tau)^2\right\}$$

for $n = 1, 2, \lambda$ and $\tau = \boldsymbol{\beta}, \gamma, \boldsymbol{\xi}$ respectively.

The joint p.d.f. of the latent variables l^1, l^2, and λ is given by:

$$g(l^1, l^2, \lambda) = f_1(l^1) \cdot f_2(l^2) \cdot f_\lambda(\lambda) \tag{B.5}$$

whereas the p.d.f. of the *observed* dependent variable l^i is given by:

$$h(l^i) = \lambda^* \cdot f_1(l^1) + (1 - \lambda^*) f_2(l^2) \tag{B.6}$$

with

$$\lambda^* = \text{Prob}[\lambda \leqq 0] = \Phi(-\mathbf{x}_\lambda{}'\boldsymbol{\xi}) \tag{B.7}$$

in which we are assuming that $\sigma_\lambda = 1$ for identification, and where $\Phi(\cdot)$ denotes the standard normal cdf.

Using the above, the *conditional* pdf of l^1, l^2 and l^λ, given the observed value of l^i is:

$$g(l^1, l^2, \lambda | l^i) = \begin{cases} g(l^i, l^2, \lambda)/h(l^i) & \text{if } \lambda \leqq 0 \\ \\ g(l^1, l^i, \lambda)/h(l^i) & \text{if } \lambda > 0 \end{cases} \tag{B.8}$$

Hartley [27] shows that the partials of the log-likelihood function:

$$L(\boldsymbol{\beta}, \gamma, \boldsymbol{\xi}, \sigma_1^2, \sigma_2^2) = \sum_{i=1}^{N} \log(h(l_i)) \tag{B.9}$$

with respect to $\boldsymbol{\beta}$ and γ can be written as:

$$\frac{\partial L}{\partial \boldsymbol{\beta}} = \frac{1}{\sigma_1^2} \sum_{i=1}^{N} w_1(l^i) \cdot (l^i - \mathbf{x}_1'\boldsymbol{\beta}) \cdot \mathbf{x}_1 = 0 \tag{B.10}$$

and

$$\frac{\partial L}{\partial \gamma} = \frac{1}{\sigma_2^2} \sum_{i=1}^{N} w_2(l^i) \cdot (l^i - \mathbf{x}_2{}'\gamma) \cdot \mathbf{x}_2 = 0 \tag{B.11}$$

while the partial of the log-likelihood function with respect to ξ can be written as:

$$\frac{\partial L}{\partial \xi} = \sum_{i=1}^{T} (E[\lambda|l^i] - \mathbf{x}_\lambda{}'\boldsymbol{\xi}) \cdot \mathbf{x}_\lambda = 0 \tag{B.12}$$

where $E[\lambda|l^i]$ is the conditional expectation of λ given the observed value of l^i, which is equal to:

$$E[\lambda|l^i] = \mathbf{x}_\lambda{}'\boldsymbol{\xi} - w_1(l^i) \cdot \frac{f_\lambda(0)}{\lambda^*} + w_2(l^i) \cdot \frac{f_\lambda(0)}{1-\lambda^*} \tag{B.13}$$

and where $w_1(l^i)$ and $w_2(l^i)$ are weights defined as:

$$w_1(l^i) \;=\; \lambda^* \cdot [f_1(l^i)/h(l^i)] \tag{B.14}$$

$$w_2(l^i) \;=\; (1-\lambda^*) \cdot [f_2(l^i)/h(l^i)] \tag{B.15}$$

Also, the partials of the log-likelihood with respect to σ_1^2 and σ_2^2 can be written as:

$$\frac{\partial L}{\partial \sigma_1^2} = -\frac{1}{\sigma_1^4} \sum_{i=1}^{N} w_1(l^i)[\sigma_1^2 - (l^i - \mathbf{x}_1{}'\boldsymbol{\beta})^2] = 0 \tag{B.16}$$

$$\frac{\partial L}{\partial \sigma_2^2} = -\frac{1}{\sigma_2^4} \sum_{i=1}^{N} w_2(l^i)[\sigma_2^2 - (l^i - \mathbf{x}_2{}'\gamma)^2] = 0 \tag{B.17}$$

The first order conditions in (B.10 – B.12) and (B.16 – B.17) can be numerically solved to find the values of $\{\boldsymbol{\beta}, \gamma, \boldsymbol{\xi}, \sigma_1, \sigma_2\}$ that maximize (B.9).

Appendix C

Derivation of the optimal input quantities

The farmer faces the following problem:

$$\underset{l_N, l_T, l_C, x_N}{Max}\ \Pi = p_N q_N(l_N, x_N, z^q) + p_T q_T(l_T, z^q) + p_C q_C(l_C, z^c, z^q) - p_{x_N} x_N \tag{C.1}$$

subject to a labor constraint:

$$\overline{L} - l_N - l_T - l_C \geq 0 \tag{C.2}$$

a cash liquidity constraint:

$$C - p_{x_N} x_N \geq 0 \tag{C.3}$$

where $C = p_C q_C(l_C, z^c, z^q)$ and the non-negativity constraints:

$$l_N \geq 0 \tag{C.4}$$

$$x_N \geq 0 \tag{C.5}$$

$$l_T \geq 0 \tag{C.6}$$

$$l_C \geq 0 \tag{C.7}$$

This problem can be solved in two parts. First, we assume that the labor constraint is binding so that

$$\overline{L} - l_N - l_T - l_C = 0$$

In addition, we focus at the case where farmers produce both the traditional and new variety i.e. $l_N > 0$, $x_N > 0$, $l_T > 0$. Substituting the labor constraint for l_T, we rewrite the maximization problem as:

$$\underset{l_C}{Max}\{\underset{l_N,x_N}{Max}\ p_N q_N(l_N, x_N, z^q) + p_T q_T(\overline{L} - l_N - l_C, z^q) + p_C q_C(l_C, z^c, z^q) - p_{x_N} x_N\} \tag{C.8}$$

subject to constraints (C.3) and (C.7) as well as (C.4), (C.5), (C.6) being strictly positive.

The first order Kuhn Tucker conditions that describe the interior problem are:

$$\frac{\partial \Pi(\cdot)}{\partial l_N} : p_N \frac{\partial q_N}{\partial l_N}(l_N, x_N, z^q) - p_T \frac{\partial q_T}{\partial l_N}(\overline{L} - l_N - l_C, z^q) = 0 \tag{C.9}$$

and

$$\frac{\partial \Pi(\cdot)}{\partial x_N} : p_N \frac{\partial q_N}{\partial x_N}(l_N, x_N, z^q) - p_{x_N}(1 + \mu) = 0,\ \mu \geq 0 \tag{C.10}$$

where μ is the multiplier associated with the cash liquidity constraint.

Intuitively, and given our setting, this multiplier represents the additional amount of the new variety input (x_N) that the farmer can purchase if he has an additional unit of cash. However, this will only be relevant if, given the optimal demand for input x_N, the farmer does not have the corresponding cash available and is thus constrained. If, on the other hand, the farmer has the required cash, the multiplier has no effect in this decision.

Formally, we can look at two cases that arise depending on whether the cash constraint is binding or not.

Unconstrained case If the cash constraint is not binding ($\mu = 0$), the Kuhn-Tucker conditions become:

$$\frac{\partial \Pi(\cdot)}{\partial l_N} : p_N \frac{\partial q_N}{\partial l_N}(l_N, x_N, z^q) - p_T \frac{\partial q_T}{\partial l_N}(\overline{L} - l_N - l_C, z^q) = 0 \tag{C.11}$$

and

$$\frac{\partial \Pi(\cdot)}{\partial x_N} : p_N \frac{\partial q_N}{\partial x_N}(l_N, x_N, z^q) - p_{x_N} = 0 \tag{C.12}$$

Given l_C, from the equations above we can solve for for the optimal quantities of l_N and x_N:

$$x_N^u = x_N^u(p_N, p_{x_N}, p_T, z^q, \overline{L}, l_C^u) \tag{C.13}$$

$$l_N^u = l_N^u(p_N, p_{x_N}, p_T, z^q, \overline{L}, l_C^u) \tag{C.14}$$

where the superscript u refers to the unconstrained case.

In addition, using the fact that:

$$\overline{L} - l_N^u - l_T^u - l_C^u = 0$$

we can solve for l_T^u :

$$l_T^u = \overline{L} - l_N(p_N, p_{x_N}, p_T, z^q, \overline{L}, l_C^u) - l_C^u = l_T^u(p_N, p_{x_N}, p_T, z^q, \overline{L}, l_C^u) \tag{C.15}$$

Notice, that the optimal levels of these inputs depends on the labor demand for the cash activity (l_C^u).

To solve for l_C^u, we first substitute l_N^u, x_N^u and l_T^u in equation (C.8). The maximization problem becomes:

$$\underset{l_C^u}{Max}\ p_N q_N(l_N^u(p_N,p_{x_N},p_T,z^q,\overline{L},l_C^u),x_N^u(p_N,p_{x_N},p_T,z^q,\overline{L},l_C^u),z^q)+$$
$$p_T q_T(\overline{L}-l_N^u(p_N,p_{x_N},p_T,z^q,\overline{L},l_C^u)-l_C^u,z^q)+ \tag{C.16}$$
$$p_C q_C(l_C^u,z^c,z^q)-p_{x_N}x_N^u(p_N,p_{x_N},p_T,z^q,\overline{L},l_C^u)$$

subject to (C.7).

The Kuhn Tucker condition is:

$$\frac{\partial\Pi(\cdot)}{\partial l_C^u}:\frac{\partial l_N^u}{\partial l_C^u}(\cdot)[p_N\frac{\partial q_N}{\partial l_N^u}(\cdot)-p_T\frac{\partial q_T}{\partial l_N^u}(\cdot)]+\frac{\partial x_N^u}{\partial l_C^u}(\cdot)[p_N\frac{\partial q_N}{\partial x_N^u}(\cdot)-p_{x_N}]+p_C\frac{\partial q_C}{\partial l_C^u}(\cdot)-p_T\frac{\partial q_T}{\partial l_C^u}(\cdot)=0 \tag{C.17}$$

which is simplified as:

$$\frac{\partial\Pi(\cdot)}{\partial l_C^u}:p_C\frac{\partial q_C}{\partial l_C^u}(l_C^u,z^c,z^q)-p_T\frac{\partial q_T}{\partial l_C^u}(p_N,p_{x_N},p_T,z^q,\overline{L},l_C^u)=0 \tag{C.18}$$

The optimal choice of l_C^u is in reduced form:

$$l_C^u=l_C^u(p_N,p_{x_N},p_T,z^q,\overline{L},p_C,z^c) \tag{C.19}$$

Constrained case If the cash constraint is binding ($\mu > 0$) then

The first order Kuhn Tucker conditions that describe the interior problem now are:

$$\frac{\partial\Pi(\cdot)}{\partial l_N}:p_N\frac{\partial q_N}{\partial l_N}(l_N,x_N,z^q)-p_T\frac{\partial q_T}{\partial l_N}(\overline{L}-l_N-l_C,z^q)=0 \tag{C.20}$$

and

$$\frac{\partial\Pi(\cdot)}{\partial x_N}:p_N\frac{\partial q_N}{\partial x_N}(l_N,x_N,z^q)-p_{x_N}(1+\mu)=0,\ \mu>0 \tag{C.21}$$

Using these two equations, given l_C (and hence C) and the observation that:

$$C - p_{x_N} x_N = 0$$

(since the cash constrained is binding), allows us to solve for the optimal quantities of l_N, x_N, and μ. In particular:

$$x_N^c = x_N^c(p_N, p_{x_N}, p_T, z^q, \overline{L}, l_C^c, C) \tag{C.22}$$

$$l_N^c = l_N^c(p_N, p_{x_N}, p_T, z^q, \overline{L}, l_C^c, C) \tag{C.23}$$

and

$$\mu = l_T^c(p_N, p_{x_N}, p_T, z^q, \overline{L}, l_C^c, C) \tag{C.24}$$

where the superscript c refers to the constrained case.

Finally, we use the fact that

$$\overline{L} - l_N^c - l_T^c - l_C^c = 0$$

to solve for l_T^c :

$$l_T^c = l_T^c(p_N, p_{x_N}, p_T, z^q, \overline{L}, l_C^c, C) \tag{C.25}$$

The important difference from the unconstrained case is that here the farmer cannot afford to purchase his optimal quantity of the new variety input (x_N). It is at this point where the income from the cash activity C becomes crucial. The ability of the farmer to adopt the new variety is directly linked with cash income: higher levels of cash income relax the cash constraint and thus enabling the farmer to adopt more. This is the reason why the level of input for the crop activities depend not only on the labor demand for the cash activity (l_C^u) but on cash income itself (C).

Finally, we substitute these quantities back in the maximization problem (equation (C.8)) in order to solve for l_C^u and hence C (since $C = p_C q_C(l_C, z^c, z^q)$). The problem is:

$$\begin{aligned} \underset{l_C^c}{Max}\ & p_N q_N(l_N^c(p_N, p_{x_N}, p_T, z^q, \overline{L}, l_C^c, C), x_N^c(p_N, p_{x_N}, p_T, z^q, \overline{L}, l_C^c, C), z^q) + \\ & p_T q_T(\overline{L} - l_N^c(p_N, p_{x_N}, p_T, z^q, \overline{L}, l_C^c, C) - l_C^c, z^q) + \\ & p_C q_C(l_C^c, z^c, z^q) - p_{x_N} x_N^c(p_N, p_{x_N}, p_T, z^q, \overline{L}, l_C^c, C) \end{aligned} \tag{C.26}$$

subject to (C.7).

The Kuhn Tucker condition is now:

$$\begin{aligned} & \tfrac{\partial \Pi(\cdot)}{\partial l_C^c} : \tfrac{\partial l_N^c}{\partial l_C^c}(\cdot)[p_N \tfrac{\partial q_N}{\partial l_N^c}(\cdot) - p_T \tfrac{\partial q_T}{\partial l_N^c}(\cdot)] + \tfrac{\partial x_N^c}{\partial l_C^c}(\cdot)[p_N \tfrac{\partial q_N}{\partial x_N^c}(\cdot) - p_{x_N}] + \\ & p_C \tfrac{\partial q_C}{\partial l_C^c}(\cdot)[p_N \tfrac{\partial q_N}{\partial l_N^c}(\cdot) \tfrac{\partial l_N^c}{\partial q_C}(\cdot) + p_N \tfrac{\partial q_N}{\partial x_N^c}(\cdot) \tfrac{\partial x_N^c}{\partial q_C}(\cdot) - p_T \tfrac{\partial q_T}{\partial l_N^c}(\cdot) \tfrac{\partial l_N^c}{\partial q_C}(\cdot) - p_{x_N} \tfrac{\partial x_N^c}{\partial q_C}(\cdot) + 1] \\ & - p_T \tfrac{\partial q_T}{\partial l_C^c}(\cdot) = 0 \end{aligned} \tag{C.27}$$

which is simplified as:

$$\frac{\partial \Pi(\cdot)}{\partial l_C^c} : p_C \frac{\partial q_C}{\partial l_C^c}(l_C, z^c, z^q) - p_T \frac{\partial q_T}{\partial l_C^c}(p_N, p_{x_N}, p_T, z^q, \overline{L}, l_C^c, C) = 0 \tag{C.28}$$

The only unknown, l_C^c, is given by:

$$l_C^u = l_C^u(p_N, p_{x_N}, p_T, z^q, \overline{L}, p_C, z^c) \tag{C.29}$$

and by definition:

$$C = p_C q_C(l_C^c(p_N, p_{x_N}, p_T, z^q, \overline{L}, p_C, z^c), z^c, z^q)) = C(p_N, p_{x_N}, p_T, z^q, \overline{L}, p_C, z^c) \tag{C.30}$$