Testing for Separation in Agricultural Household Models and Unobservable Individual Effects

Jean-Louis ARCAND and Béatrice D'HOMBRES
CERDI-CNRS, Université d'Auvergne
65, boulevard François Mitterrand, 63000 Clermont Ferrand, FRANCE
email : arcandjl@alum.mit.edu, B.D_HOMBRES@cerdi.u-clermont1.fr
tel: +33 4 73177400, fax: +33 4 73177428

Abstract

When market structure is complete, factor demands by households will be independent of their characteristics, and households will take their production decisions as if they were profit-maximizing firms. This observation constitutes the basis for one of the most popular empirical tests for complete markets, commonly known as the "separation" hypothesis. In this paper, we show that all existing tests for separation using panel data are potentially biased towards rejecting the null-hypothesis of complete markets, because of the failure to adequately control for unobservable individual effects. Since the variable on which the test for separation is based cannot be identified in most panel datasets following the usual covariance transformations, and is likely to be correlated with the individual effect, neither the within nor the variance-components procedures are able to solve the problem. We show that the Hausman-Taylor (1981) estimator, in which the impact of covariates that are invariant along one dimension of a panel can be identified through the use of covariance transformations of other included variables that are orthogonal to the individual effects as instruments, provides a simple solution. We furnish an empirical illustration using a rich Tunisian dataset in which separation -and thus the null of complete markets- is strongly rejected using the standard approach, but is not rejected once correlated unobservable individual effects are controlled for using the Hausman-Taylor instrument set.

Keywords: panel data, individual effects, household models, testing for incomplete markets, development microeconomics, Tunisia.

JEL: O120, C230, D130, D520.

1 Introduction

One of the most widely-used empirical tests for the presence of market imperfections in developing countries is provided by the so-called "separation" hypothesis. Numerous papers, including the seminal article by Benjamin (1992), have tested the hypothesis that factor demands on a given plot of land will be independent of household characteristics, when market structure is (almost) complete (Singh, Squire and Strauss, 1986; see the summary of the recent literature, as well as two typical applications to African data, in Udry, 1996).[1] Separation implies that the marginal productivity of inputs will be a function solely of plot characteristics and prices, and that households take their production decisions as if they were profit-maximizing firms. In contrast, when factor demands are a function of household characteristics, marginal productivities are not equated across plots and a deviation with respect to the first-best optimum obtains. Moreover, the production and consumption decisions of households can no longer be treated recursively.

The purpose of this note is to demonstrate that: (i) in most cases, the standard test for separation using panel data is biased towards rejecting the null-hypothesis of complete markets because of a problem of unobservable individual effects; (ii) the usual covariance transformations performed on panel data cannot solve this problem; but (iii) the Hausman-Taylor (1981) estimator can. In addition, we provide an empirical illustration using a Tunisian dataset in which the rejection of the null-hypothesis of complete markets using the standard approach is overturned once correlated individual effects are controlled for using the Hausman-Taylor estimator.

In the most common version of the test for separation, the typical equation being estimated on plot-level agronomic data is given by:

$$Y_{iht} = X_{iht}\alpha + Z_{ht}\beta + \varepsilon_{iht}, \tag{1}$$

where Y_{iht} is total labor usage (i.e., family and hired labor) on plot i, cultivated by household h, at time t, X_{iht} is a matrix of plot characteristics, Z_{ht} is a matrix of household characteristics, and ε_{iht} is a disturbance term that satisfies the usual Gauss-Markov assumptions. Separation is then associated with a simple F-test on the exclusion restriction that $\beta = 0$. In the prototypical expression of the test for separation (Benjamin, 1992), Z_{ht} is household size.

The main problem associated with this procedure is that the disturbance term ε_{iht} can in all likelihood be decomposed into a nested error component structure given by:[2]

$$\varepsilon_{iht} = \mu_t + \lambda_h + \lambda_{ht} + \eta_{iht}, \tag{2}$$

where μ_t is a shock common to all plots and households at time t, λ_h is a time-invariant household effect, λ_{ht} is a household-time effect, and η_{iht} is a disturbance term that satisfies the usual assumptions.[3] In most plot-level datasets

---

[1] A concise primer on household models is also provided by Bardhan and Udry (1999), chapter 2. Note that Benjamin (1992) used Indonesian *household*-level data and was therefore unable to control for individual effects at all.

[2] See Baltagi *et al* (2001).

[3] λ_h will represent, for exemple, unobservable productivity characteristics of the household, time-invariant whereas λ_{ht} might be given by a transitory unobserved household shock, such

used in the literature, each household cultivates several plots. This is a standard panel data framework, with one dimension being given by plots, the second by households, and the third by time. Although λ_h can be accounted for by a "within" procedure which transforms variables into deviations with respect to their household-specific means (over *all* time periods), there remains λ_{ht}. Since it is probable that λ_{ht} is correlated with Z_{ht}, the least-squares estimate of β, even after the standard "within" transformation, will be biased with, in the scalar case:

$$p\lim \hat{\beta}_w = \beta + \frac{cov[\lambda_{ht}, \hat{e}_{iht}]}{\sigma_e^2}, \tag{3}$$

where σ_e^2 is the variance of the residual $\hat{e}_{iht}$ from the auxiliary "within" regression of household size on X_{iht}.[4] If $cov[\lambda_{ht}, \hat{e}_{iht}] \neq 0$, as is likely in the context of what is essentially a labor demand equation, then *all standard tests of separation are biased towards rejecting the null-hypothesis of complete markets*, when the "true" value of β is zero. One may therefore reject the null not because market structure is necessarily incomplete, but simply because of a banal problem of unobservable heterogeneity. Another way of putting this is that, in the standard test, the rejection of separation is conditional on the maintained identifying assumption that λ_{ht} is the same across all households at a given time t. It is very likely that this assumption is violated.

The usual econometric response to a problem of unobservable individual heterogeneity in panel data is to apply one of the standard covariance transformations, such as the "within" procedure. Here, this would involve expressing all variables as deviations with respect to their household-specific means, *at a given t*. While, under the assumption of exogeneity, this does allow one to recover unbiased estimates of α, it has the regrettable side-effect of eliminating the variable(s) upon which the test for separation is based since, when one sweeps out λ_{ht}, one also sweeps out Z_{ht}. Since it is highly likely that λ_{ht} is not orthogonal to Z_{ht}, random effects are not an answer, as they too will yield biased estimates of β.

Moreover, standard instrumental variables procedures, in which one would simply instrument for Z_{ht}, are not usually implementable. This is because admissible exogenous instruments that would be correlated with Z_{ht} but are orthogonal to λ_{ht} are usually not available or, if they are, should probably already be included in Z_{ht} for theoretical reasons.

The problem, which is similar in spirit to that of consistently estimating the returns to education using panel data when schooling is correlated with the individual effects, can be solved using the Hausman-Taylor (1981, henceforth, HT)

---

as a wedding that affects simultaneously the household composition and the human capital of the household. Note that in datasets where it is possible to follow plots over time, there may also be a time-invariant plot-specific effect.

[4] Hsiao (1986), p. 64, equation (3.9.3). The corresponding matrix expression obtains when Z_{ht} involves several household characteristics.

instrumental variables estimator, which allows one to control for unobservable individual effects that are correlated with Z_{ht}, while allowing one to identify β.

2 An empirical illustration of the unwarranted rejection of the null-hypothesis of complete markets

2.1 The Hausman-Taylor instrument set

Let X_{1iht} be those elements of X_{iht} that are uncorrelated with λ_{ht}, while X_{2iht} are those that are; Z_{1ht} and Z_{2ht} are defined in a similar manner. The set of instruments proposed by HT (1981), adapted to the three-dimensional panel structure, is:

$$A_{HT} = [Q_{vt}X_{iht}; P_{vt}X_{1iht}; Z_{1ht}], \tag{4}$$

where P_{vt} and Q_{vt} are the idempotent matrices that perform the "between" and "within" transformations at time t, respectively.[5] Under the assumption that X_{iht} is uncorrelated with η_{iht}, $Q_{vt}X_{iht}$ is a legitimate set of instruments since $E[(Q_{vt}X_{iht})'\eta_{iht}] = 0$. The basic intuition behind the HT estimator is that only the λ_{ht} component of the error term is correlated with $[X_{2iht}\ Z_{2ht}]$, which allows one to use $Q_{vt}X_{2iht}$ as instruments for X_{2iht}, while $P_{vt}X_{1iht}$ furnishes the instruments for Z_{2ht}. The HT estimator therefore allows one to control for unobservable correlated individual effects, while allowing one to identify the parameters of interest (β) in the context of testing for separation. A necessary condition for identification is that the number of elements of X_{1iht} be greater than the number of elements of Z_{2ht} (HT, 1981, PROPOSITION 3.2, p. 1385).[6]

The three-dimensional nature of our dataset allows us an additional degree of freedom in terms of the definition of HT-type instruments. Above, we considered orthogonality conditions of subsets of X_{iht} and Z_{ht} with respect to the individual-time effect λ_{ht}. But the three-dimensional nature of the data also allows us to construct instruments based on orthogonality conditions with respect to variables that have been purged of their time-invariant household-specific component which is correlated with λ_h. An advantage of this procedure is that, in empirical applications, the orthogonality of $P_{vt}X_{1iht}$ with respect to the individual effects could be suspect. Purging $P_{vt}X_{1iht}$ of its component

---

[5] For simplicity of exposition, we express the instrument set as if the data were balanced. In the empirical application, the unbalanced nature of the data will, of course, be taken into account.

[6] These results have been extended by Amemiya and MaCurdy (1986) and Breusch, Mizon and Schmidt (1989) who suggest a broader set of instruments that should improve efficiency. Their approach, however, is only possible on balanced data, which is not the case in the dataset used in this paper or, for that matter, in most plot-level agronomic datasets. Notice that the HT instrument set is admissible only if exogeneity is satisfied. This is another potential source of bias in tests for separation, but which is difficult to address because of the lack of admissible plot-level instruments in most datasets.

that is correlated with the time-invariant individual effect, λ_h, should render it more palatable as a potential instrument set. In that case, the set of HT-type instruments is given by

$$\left[Q_{vt}X_{1iht}; Q_{vt}X_{2iht}; \widetilde{X}_{1iht}; Z_{1ht}\right]. \tag{5}$$

where

$$\widetilde{X}_{1iht} = Q_v(P_{vt}X_{1iht}) \tag{6}$$

denotes the matrix of explanatory variables that have been purged of their component which is correlated with λ_h. In other words, $Q_v = I - P_v$ is the idempotent projection matrix that transforms variables into deviations with respect to their household-specific means (over all time periods).[7]

2.2 An empirical illustration

In order to illustrate our fundamental point concerning the bias affecting conventional tests for separation in household models, consider the following standard procedure implemented on a typical plot-level dataset. The data come from two surveys (1993, 1995) carried out in the village of El Oulja, Tunisia (see Matoussi and Nugent, 1989, and Laffont and Matoussi, 1996, for descriptions of the village). These data display those properties discussed in the introduction: a Hausman test of random household-time effects (λ_{ht}) versus fixed effects in an empirical counterpart to equations (1) and (2) strongly rejects (with a p-value below 0.001) the null of the absence of correlation between λ_{ht} and Z_{ht}. The bias identified in equation (3) is therefore manifestly present in conventional tests of the null-hypothesis of complete markets using this panel dataset.

For the purpose of HT estimation, we divide the explanatory variables into two categories: (i) X_{1iht} variables, assumed to be uncorrelated with λ_{ht}, include four soil type dummies and a dummy variable that indicates whether the plot is irrigated or not, as well as a set of eight crop dummies;[8] (ii) X_{2iht} variables, assumed to be correlated with λ_{ht}, are given by the share of costs borne by the cultivator, divided by the share of output received, for eight different inputs, as well as log plot size in hectares.[9]

The economic rationale for allowing the variables included in X_{2iht} to be correlated with λ_{ht} is that they may, in the context of tenancy contracts (which account for 28 percent of the plots in the sample), be determined as the solution to a principal-agent relationship between a landlord and a tenant, and would

---

[7] Note that the three-dimensional nature of our dataset allows us to use some other combinations of our exogeneous explanatory variables. They are not considered here.

[8] The soil types are clay, red, sandy and barren, with mixed soil types being the excluded category; the crop dummies are other cereals, potatoes, onions, garden vegetables, tomatoes, beetroots, melon and fodder; the excluded category is wheat. We also include a year dummy.

[9] The output and cost shares both equal 1 on plots cultivated by owner-operators. Values strictly less than or greater than one of the ratio obtain on plots under share tenancy contracts.

then be functions of tenant characteristics, including those unobservable characteristics potentially captured by λ_{ht}.[10] Plot size is also assumed to be correlated with λ_{ht}, as it too may be chosen by landlords for plots under tenancy contracts. Both of these hypotheses will be subjected to a test of the corresponding overidentifying restrictions below. Our single Z_{2ht} variable is given by log household size. In line with the usual methodology, the dependent variable is log total (hired and family) labor usage on the plot, in person-days per hectare.[11] Table 1 provides summary statistics on all the aforementioned variables.

Estimation results are presented in Table 2. The standard test for separation is presented in column 1, and yields an unambiguous rejection of the null-hypothesis of complete markets in that log household size is highly significant at the usual levels of confidence (t-statistic = 4.55). In column 2, we control for time-invariant household characteristics (λ_h) using the "within" transformation: recall that the impact of household size can be identified here because we have two years of data and household size varies over the two surveys.[12] Again the null of complete markets is strongly rejected by the data (t-statistic = 2.56).[13] In column 3, we present results which allow for random household-time (λ_{ht}) effects: this specification, which also rejects the null of complete markets, can however be dismissed on the basis of the corresponding Hausman test in favor of fixed effects, as mentioned above (p-value of the Hausman test is below 0.001). Of course, household-time (λ_{ht}) *fixed* effects would not allow one to test for separation at all in that they would also sweep out the impact of household size.

In column 4, we present results corresponding to the efficient HT estimator using the instrument set given in equation (4).[14] The results are striking. In contrast to what was found in columns 1 through 3, the null of complete markets is *not* rejected at the usual levels of confidence: the point estimate of the parameter associated with household size is statistically indistinguishable from

---

[10] An additional, empirical, motivation for using the ratios of cost-shares to the output share is that the data in question come from a single village and that the only source of cross-sectional variation in effective input prices stems from heterogeneity in contractual form on plots under tenancy contracts.

[11] Note that there are *no* Z_{1ht} variables in this specification.

[12] Note, despite a substantial fall in the variance of log household size, which goes from 0.328 in levels, to 0.014 when expressed in terms of deviations with respect to household-specific means (over both periods), that the estimated standard error is still reasonably small, with the associated t-statistic being equal to 2.406. The time-invariant household fixed effects (λ_h) used here correspond to the type of specification used by Udry (1996), Table 3, column 2, for a labor demand per hectare equation estimated on the Burkina Faso ICRISAT dataset.

[13] A household-specific random effects specification (λ_h, not presented) is strongly rejected by the corresponding Hausman test.

[14] We made the assumption that the covariance matrix of the disturbance term has a household-clustering form. That covariance matrix of error term has the advantage to be flexible by allowing arbitrary intra-cluster correlation and clustering heteroskedasticity. We thus relaxe the more formal assumption assuming that the correlation within each cluster is constant and has a nested random form. Obviously, our assumption has a cost in terms of efficiency.

zero (t-statistic = 0.71).[15] Moreover, the test of the overidentifying restrictions does not lead one to reject, with a p-value equal to 0.545. In addition, the Shea's partial R^2 (0.071 for the reduced form concerning the household size) and the partial F-test (p-value below 0.001) of the joint significance of the instruments point out that we are not facing a weak instrument problem.[16] We also compute the test (developped by Bowden Turkington (1994)) based on canonical correlation. [17] We reject that the smallest canonical correlation is nul (p-value equal to 0.002).

However, the Hansen test is potentially inconclusive insofar as this test is based on the strong assumption that at least as many instruments as the number of elements of Z_{2ht} are exogenous. As the Hausman-Taylor procedure is very sensitive to the choice of the variables included in X_{1iht} and X_{2iht}, we compute the "difference Hansen" statistic which enables us to test the validity of a subsets of instruments (see Hayashi, 2000). To that end, we first implement the HT estimator using, in addition to the traditional matrix of instrument (4), the instrument set $\widetilde{X}_{1iht}$. Second, we test that the subset of instruments $P_{vt}X_{1iht}$ satisfies the orthogonality conditions. The "difference Hansen" statistic presented in Table 2 does not lead one to reject the null hypothesis that the specified variables are admissible instruments (p-value of 0.676).[18]

The upshot is that, in stark contrast to the usual approach which does not control for unobservable individual effects, HT estimation leads to the non-rejection of the null-hypothesis of complete markets. Moreover the consistency of the HT-based results presented in column 4 is ensured, in that they are not rejected by the tests of the corresponding overidentifying restrictions and the difference Hansen test.

3 Concluding Remarks

This paper has shown that the rejection of the null-hypothesis of complete markets in household models, based on the widely-used test of the exclusion restrictions implied by separation, can be entirely due to the bias stemming from uncontrolled-for unobservable individual heterogeneity.[19] Our results are particularly important for plot-level panel datasets where no time dimension is

---

[15] Note that all other point estimates presented in column 4 are fairly close to those obtained using household-specific fixed effects in column 2, except for that associated with the irrigated plot dummy and the seeds cost share.

[16] The Shea's partial R^2 takes the intercorrelations among the instruments into account.

[17] Hall et al. (1996) shows that IV estimators are not weakly identified if and only if all the canonical correlations converge to non-zero limits. They develop a likelihood ratio statistic for a null that the smallest canonical correlation is zero.

[18] Obviously, testing our subset of overidentifying restrictions is only valid if $\widetilde{X}_{1iht}$ variables are uncorrelated with time individuals effects. However, as explained above, this assumption seems to be justified in a theoretically point of view.

[19] For a very recent example of the application of the standard approach, see Bowlus and Sicular (2003).

present, since there is *no* means at all, apart from the HT estimator, of testing for separation while controlling for unobservable individual effects (i) if the latter are correlated with the household-level variable that is the focus of the test, and (ii) if no exogenous instruments are available. As was the case with the dataset considered in our empirical illustration, both of these conditions are likely to hold in practice.

The implications of our results are, moreover, suggestive, in that there may be other received results in applied microeconomics, based on panel data, to which the HT estimator could be fruitfully applied. An obvious example is constituted by tests of the precautionary savings motive, in which empirical measures of the risks faced by households are usually time invariant, and in which no attempt is made to correct for unobservable individual effects.

Our results bring the methodology of testing for separation using panel data into sharper focus. This is because we do not reject the null hypothesis of complete markets, *conditional* on λ_{ht}. If one estimates a labor demand function on US individual *firm* data, as in Griliches and Hausman (1986), one finds correlated individual firms effects, as we have found here for households. Thus, by analogy, profit-maximizing behavior by firms is not incompatible with correlated individual effects. However, in our dataset, since labor demand is a function λ_{ht}, it is not independent of household characteristics *per se*, although they are *unobservable* characteristics. Another way of putting this is that, in most panel datasets, testing for separation will undoubtedly uncover correlated individual effects. If separation is taken in its strictest sense to mean that factor demands should be independent of household characteristics, *unconditional* on λ_{ht}, then we do in fact reject the null-hypothesis of complete markets. Any structural interpretation, in terms of which market failures are binding, of the *pattern* of violations of separation based on observable household characteristics (and thus on those elements of β which are statistically different from zero) will probably, however, be biased unless unobservable individual effects are controlled for using the Hausman-Taylor estimator.

References

[1] Amemya, T., and MaCurdy, T. (1986), "Instrumental-Variable Estimation of an Error-Components Model," *Econometrica* 54:869-881.

[2] Baltagi, B.H., Song, S.H and Jung, B.C, (2001), "The unbalanced nested error component regression model ," *Journal of Econometrics*, 101(2): 357-381.

[3] Bardhan, P., and Udry, C. (1999). *Development Microeconomics* (Oxford: Oxford University Press).

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

[5] Bowlus, A.J. and Sicular, T. (2003), "Moving toward markets? Labor allocation in rural China," *Journal of Development Economics*, 71(2): 561-583.

[6] Breusch, T., Mizon, G., and Schmidt, P. (1989), "Efficient Estimation Using Panel Data," *Econometrica* 57:695-700.

[7] Griliches, Z. and Hausman, J. (1986), "Errors in Variables in Panel Data," *Journal of Econometrics* 3:93-118.

[8] Hausman, J. (1978), "Specification Tests in Econometrics," *Econometrica* 46:1251-1272.

[9] Hausman, J. and Taylor, W. (1981), "Panel Data and Unobservable Individual Effects," *Econometrica* 49:1377-1398.

[10] Hsiao, C. (1986). *Analysis of Panel Data.* Econometrics Society Monographs No. 11 (Cambridge, UK: Cambridge University Press).

[11] Laffont, J.-J., and Matoussi, M. (1996), "Moral Hazard, Financial Constraints and Sharecropping in El Oulja," *Review of Economic Studies* 62:381-399.

[12] Matoussi, M., and Nugent, J. (1989), "The Switch to Sharecropping in Medjez-El-Bab," in Nabli, M. K., and Nugent, J. B., eds., *The New Institutional Economics and Development: Theory and Applications to Tunisia* (New York: Elsevier Science).

[13] Singh, I., Squire, L., and Strauss, J., eds. (1986). *Agricultural Household Models: Extensions, Applications, and Policy* (Baltimore: Johns Hopkins University Press).

[14] Udry, C. (1996), "Efficiency and Market Structure: Testing for Profit Maximization in African Agriculture," processed, Yale University, June; available on the Yale Economic Growth Center website at: www.econ.yale.edu/∼cru2//pdf/separate.pdf.

Note on estimations

Many households did not engage in crop production in the second survey (1995) because of adverse climatic shocks; this explains why the number of household-years (ht) is much smaller than *twice* the number of households (h). Intercept, year dummy, and eight crop dummies are included in all specifications presented in Table 2 (no constant in col. (2)); random effects are rejected in

	Mean	Median	Std. dev.
Person-day labor input per hectare (Y_{iht})	190.860	119.0	253.271
Plot characteristics (X_{1iht})			
Soil type 1 (clay)	0.190	0.0	0.393
Soil type 2 (red)	0.201	0.0	0.401
Soil type 3 (sandy)	0.446	0.0	0.497
Soil type 4 (barren)	0.058	0.0	0.235
Irrigated plot	0.882	1.0	0.322
Contractual terms (X_{2iht})			
$\frac{\text{\% of costs paid by the cultivator}}{\text{\% of output accruing to the cultivator}}$ for:			
Manure	1.008	1.0	0.121
Chemical fertilizer	1.016	1.0	0.145
Irrigation	0.999	1.0	0.150
Plowing	0.984	1.0	0.250
Family labor	1.063	1.0	0.243
Hired labor	1.041	1.0	0.230
Seeds	1.008	1.0	0.094
Transportation	1.006	1.0	0.171
Surface of plot in hectares	5.615	1.5	13.535
Household size (Z_{2ht})	8.257	7.0	5.117

Table 1: Summary statistics, ElOulja, Tunisia (447 plots (i), 150 households (h), 196 household-years (ht))

favor of fixed effects in columns (2) and (3) by Hausman tests with associated p-values of less than 0.0001. The dependent variable in the estimation results presented in Table 2 is given by log person-days per hectare used on the plot; *t*-statistics in parentheses below coefficients, unless otherwise noted.

Mean of dep. var. = 3.825	Pooling (1)	Fixed effects λ_h (2)	Random effects λ_{ht} (3)	HT (efficient) λ_{ht} (5)
Plot characteristics (X_{1iht})				
Soil type 1 (clay)	−0.013 (−0.04)	0.206 (0.66)	0.184 (0.60)	0.248 (0.60)
Soil type 2 (red)	−0.467 (−1.40)	−0.304 (−1.07)	−0.471 (−1.60)	−0.510 (−0.94)
Soil type 3 (sandy)	−0.171 (−0.62)	−0.081 (−0.30)	−0.057 (−0.20)	−0.006 (0.02)
Soil type 4 (barren)	0.305 (0.84)	0.546 (1.48)	0.179 (0.46)	0.633 (1.2)
Irrigated plot	0.577 (1.92)	0.415 (1.78)	0.412 (1.86)	1.017 (2.49)
Joint signif. of plot chars. F [p-value]	2.89 [0.014]	2.33 [0.042]	13.43 [0.019]	19.92 [0.001]
Contractual terms (X_{2iht})				
$\frac{\text{\% of costs paid by the cultivator}}{\text{\% of output accruing to the cultivator}}$ for:				
Manure	0.975 (1.31)	4.120 (4.33)	1.346 (1.48)	1.295 (1.14)
Chemical fertilizer	1.070 (1.49)	−0.685 (−0.79)	1.191 (1.39)	1.979 (2.34)
Irrigation	−0.128 (−0.25)	1.826 (2.74)	0.233 (0.35)	0.539 (1.53)
Plowing	−0.080 (−0.26)	−1.542 (−3.31)	−0.408 (−1.07)	−0.864 (−1.64)
Family labor	0.776 (2.38)	−0.080 (−0.15)	0.645 (1.36)	0.170 (0.38)
Hired labor	−0.620 (−2.07)	−0.417 (−0.79)	−0.734 (−1.48)	−0.637 (−1.84)
Seeds	−0.675 (−0.83)	−2.902 (−2.71)	−0.758 (−0.81)	0.279 (0.21)
Transportation	−1.074 (−1.94)	−0.981 (−1.71)	−1.403 (−2.58)	−1.494 (−2.71)
Log surface of plot in hectares	−1.008 (−11.10)	−0.474 (−5.70)	−0.816 (−11.59)	−0.681 (−6.20)
Joint signif. of plot chars. F − [p-value]	4.96 [0.001]	4.20 [0.000]	16.00 [0.042]	29.15 [0.000]
Log household size (Z_{2ht})	0.571 (4.55)	0.943 (2.56)	0.516 (3.12)	0.404 (0.71)
R^2	0.6934	0.6558	0.6856	0.6677
Test of overid. restrictions [d.f., p-value]	*n.a.*	*n.a.*	*n.a.*	10.805 [12,0.545]
Shea Partial R^2 for log household size	*n.a.*	*n.a.*	*n.a.*	0.071
Canonical correlation	*n.a.*	*n.a.*	*n.a.*	0.053 [0.002]
Difference Hansen	*n.a.*	*n.a.*	*n.a.*	10.214 [13,0.676]

Table 2: Labor demand equations: Pooling, fixed effects, random effects, and Hausman-Taylor estimators (447 plots (i), 150 households (h), 196 household-years (ht))