ELSEVIER

Journal of Development Economics
Vol. 46 (1995) 51–84

JOURNAL OF
Development
ECONOMICS

Can unobserved land quality explain the inverse productivity relationship?

Dwayne Benjamin [1]

Department of Economics, University of Toronto, 150 St. George Street, Toronto, Canada, M5S 1A1

Received February 1992, final version received February 1994

Abstract

An inverse relationship between both farm productivity and labor intensity, and farm size, is a common empirical finding in developing country agriculture. The traditional explanation has been imperfect labor markets. Recently, it has been suggested instead that the inverse relationship is a statistical artifact resulting from omitted land quality. Using a farm-level data set from Java, I investigate whether omitted variable bias can explain the inverse relationship. I show that the inverse relationship and accompanying wage responses are inconsistent with a model of neoclassical farm behavior that ignores omitted variable bias. Instrumental variables techniques yield parameter estimates in which the inverse relationship is eliminated and the estimated wage elasticities are more in line with economic theory. Further econometric investigation hints that a model of omitted land quality may be a possible source of the inverse relationship. These results emphasize the importance of considering the sources of cross-sectional variation in estimating microeconomic relationships.

[1] This paper is a substantially revised version of Chapter 2 of my Princeton Doctoral Dissertation, *An Econometric Analysis of Farm Labor Demand in Rural Java*. Many thanks go to David Card and Angus Deaton for comments and suggestions. This revision has benefited from the comments of Loren Brandt, Michael Denny, Melvyn Fuss, Angelo Melino, and Aloysius Siow; participants in seminars at Guelph, McMaster, Princeton, Toronto, and Western Ontario; and two anonymous referees. Financial assistance from the Alfred P. Sloan Foundation, the Industrial Relations Section at Princeton, the Institute for Policy Analysis at Toronto, and the SSHRC (#410-90-0892 and #410-92-0975) is gratefully acknowledged. Of course, I am responsible for all remaining errors and shortcomings of this paper.

SSDI 0304-3878(94)00048-H

Keywords: Labor demand; Inverse relationship; Rice farming in Java; Agricultural labor markets

JEL classification: O12; Q12; J43

0. Introduction

One of the oldest empirical regularities in economics is the inverse relationship between farm productivity and farm size: small farms produce more output per acre. Closely associated is an inverse relationship between labor intensity and farm size: small farms also use more labor per acre. [2] Most researchers have focused on imperfect factor markets as the explanation of these two correlations. For example, with poorly functioning external labor markets, households have a surplus of underemployed labor available for use on the family plot. Small farms therefore cultivate more intensively. Because it is an implication of imperfect factor markets, the inverse relationship for both output and labor has been cited as indirect evidence of imperfect labor markets. More recent studies emphasize, however, that other market imperfections, such as credit constraints for small farms, may induce the inverse relationship. [3]

Instead of imperfect markets, the explanation may lie in agricultural technology. For example, decreasing returns to scale would generate the inverse relationship. Most studies, however, fail to reject that agricultural production is characterized by constant returns to scale. [4] The most commonly suggested alternative explanation is that the inverse relationship is a statistical artifact, generated by omitted land quality. Noticing that the inverse productivity relationship was more pronounced between regions than within regions, Sen (1975) suggested that the relationship could be the result of a negative correlation between farm size and unobserved land quality. [5] Carter (1984) addressed this hypothesis in a compre-

[2] Chayanov (1926) presents one of the first careful illustrations of this phenomenon. Sen (1962) is the first modern reference to the inverse output relationship. Berry and Cline (1979) provide a thorough review of empirical evidence and estimation issues regarding the inverse productivity relationship. See Booth and Sundrum (1984) and Abey et al. (1981) for a comprehensive discussion of the inverse relationship as it relates to labor absorption, especially in an Indonesian context.

[3] These studies, like Feder (1985), Eswaran and Kotwal (1986), and Carter and Wiebe (1990) also highlight the possibility that the inverse relationship would disappear, or even reverse, for larger farms. Large farms' 'disadvantages' in the labor market would be outweighed by advantages in other markets (like the credit market).

[4] Berry and Cline (1979) survey many of these studies, while Carter (1984) is an example of a careful study of the inverse relationship, where constant returns to scale appears to characterize farm production. In Benjamin (1991) I also show that decreasing returns to scale is an unlikely explanation of the inverse relationship in the context of the equations estimated in this paper.

[5] In addition to Sen, see also Khurso (1973) and Mehra (1976) for other early references to the suggestion that omitted land quality variables were responsible for the inverse relationship between yield and farm size.

hensive study of Indian farms. He controlled for unobserved land quality by including village fixed-effects. While slightly attenuated, the inverse productivity relationship survived intact. Through subsequent production function estimates, Carter established two important empirical attributes of farm production: (i) despite the inverse relationship, production was characterized by constant returns to scale, and (ii) small farms used 'too much' labor relative to the wages they faced. Carter essentially argued that the inverse productivity relationship resulted from higher intensity of labor use on small farms. This was taken as indirect evidence in favor of the imperfect labor markets hypothesis.

Bhalla (1988) and Bhalla and Roy (1988) furnish direct evidence to the contrary. Using Indian data, they find (i) that measures of land quality and farm size are negatively correlated and (ii) that including measures of farm-level soil quality eliminates the inverse output relationship. Unfortunately, they did not have data on labor inputs to address the labor allocation and imperfect labor markets question more directly. However, their remedy of the inverse relationship suggests the possibility that imperfect labor markets are not the source of the inverse relationship.

The purpose of this paper is to determine, at least for the case of rural Java, whether omitted variable biases can explain and reconcile the inverse relationships, or whether imperfect factor markets remain a plausible explanation. Unlike Bhalla and Roy, I do not have data on farm-level land quality. Instead, I have reasonably good data on labor input, and can make more direct inferences about farm labor allocation and its relationship to the labor market. The imperfect labor markets hypothesis challenges the validity of neoclassical labor demand theory: market wages do not induce labor allocation decisions as neoclassical theory suggests. Therefore, my approach is to cast farm output and labor demand decisions according to traditional labor demand theory, and then determine whether the selected empirical moments can be reconciled with the theory. This requires incorporating other moments besides the scale effect, especially wage effects. In short, my objective is to determine whether a price taking, profit-maximizing, constant returns-to-scale (CRS) farm model can be squared with the facts, through suitable allowance for omitted variable bias.

The plan of the paper is as follows. Section 1 introduces the data and establishes the empirical robustness and magnitude of the inverse relationships for a large sample of rice farmers from rural Java. Section 2 develops the 'null hypothesis' model of neoclassical, price-taking farmers. If we ignore omitted variables bias, the inverse output and labor relationships, as well as the estimated wage effects, are shown to be inconsistent with a neoclassical labor demand model. Indeed, the estimated perverse wage effects provide additional identifying information in this paper. With omitted variables, in the absence of data on the omitted variable, instrumental variable techniques must be employed. In Section 3, I endeavor to establish the validity of a set of instruments suggested by the land quality hypothesis. The quest for instruments itself yields profitable insights into the nature of labor markets in Java. The results of instrumental variables estimation suggest that omitted variable bias is a serious problem. Once this misspecification is taken into account, the neoclassical labor demand model seems to fit the data, at least for a large sample of farmers. There remain some farmers, however, for whom this is not the case, suggesting that omitted land quality does not the account for all of the inverse relationship. Section 4 places more structure on the econometric model to determine whether the omitted variable can reasonably be interpreted as land quality. A model is developed to explain the configuration of estimated biases across the OLS equations and is estimated by minimum distance. While formally rejected, the model suggests land quality may be a plausible suspect. The final section considers important issues of interpretation and offers some conclusions.

1. Demonstration of the inverse relationship

1.1. The inverse relationship

First introduced as a simple correlation between yield and farm size, the inverse relationship is usually presented as a regression of yield on farm size. The relationship can be expressed logarithmically:

$$\ln y_i = \gamma_1 + \gamma_2 \ln h_i + \sum_{j=3}^{c} \gamma_j X_{ji} + \nu_i. \tag{1}$$

y_i is a measure of output, h_i is a measure of land input, the X_{ji} are possible control variables, and ν_i is an error term. Controls for the composition of output, location of farms, per capita income of the country (if country level data are used) and measures of irrigation are commonly employed. Berry and Cline (1979) present a compendium of such regressions. The inverse relationship exists if the log land coefficient is less than one. In the labor regression, y_i represents farm labor input. In addition to output and labor equations, I also examine farm profits. There are two reasons for this. First, if land quality is the omitted variable, it should have predictable effects on farm profits. Second, by imbedding farm output and employment in a system of equations implied by a restricted profit function, the neoclassical labor demand framework can be tested.

1.2. The data: Measurement issues and compromises

The data set is based on an initial sample of 5605 wet rice farmers from Java drawn from the 1980 SUSENAS survey. These data offer several advantages over previously used data sets, especially in building a theoretically consistent micro-econometric model, the key ingredient in testing the neoclassical labor demand model. Previous studies (especially the cross country variety) use a measure of output based on aggregate value rather than physical output of a homogeneous crop. House gardens are compared to wheat farms. It is possible that this aggregation could lead to a bias toward finding the inverse relationship. If high quality land is more expensive, then in equilibrium only more expensive crops are grown on it. Using physical rice output avoids this aggregation problem. Of course, this limits the direct comparability of my results to these more aggregated studies: I seek to explain the inverse relationship that exists in rice production in Java. Especially to the degree that crop mix itself reflects a response to imperfect factor markets, my results will not generalize. For Java, however, rice is by far the single most important crop. As suggested by previous researchers (Geertz, 1963; Abey et al., 1981) rice farms have been viewed as the traditional target of surplus labor. If imperfect factor markets are driving the inverse relationship in Java, we would thus expect to see farmers' responses reflected in their rice enterprises. Furthermore, given the specificity of wet rice land, *sawah*, it is rare that farmers would switch to other crops from rice. However, for additional comparability with previous studies, I also show results for total value of all crops for these farmers. Regarding labor, previous studies have not had detailed information on person-days of both family and hired labor. For example, Abey et al. (1981) only have data on the number of persons employed on each farm. This leads to an indivisibility problem, especially on the small farms, which the authors point out could induce the inverse relationship. Again, in terms of the objective of testing the labor demand model, the use of 'rice labor' provides a relatively homogenous category of labor to work with.

A key explanatory variable is land input, the measure of productive scale. I use the actual area harvested over the previous year, which corresponds to our best measure of land input in production. This differs from previous studies which use farm size (for all agricultural purposes). To the degree that farm size is an imperfect measure of land input, using farm size rather than land input could induce measurement error and itself cause the inverse relationship. [6] While area harvested is probably the best measure of land input for the neoclassical model, this choice of productive scale is not innocuous. First, in this framework farm size is not viewed as a choice variable, and is thus treated as predetermined or exogenous. This assumption, common to most previous studies, is usually justified by the fixed nature of farm size and imperfect or non-existent land markets. Often, however, farmers can rent or sharecrop land, even if land sales are rare. Regarding Java, the SUSENAS data indicate that 85% of farmers own all of the *sawah* (wet rice land) that they cultivate. Even with land rental, since the contracts are relatively long, it is not unreasonable to view plot size as fixed. However, this

[6] Measurement error in the scale variable would bias the land coefficients from Eq. (1) below one, i.e. could cause the inverse relationship.

does *not* necessarily mean that the land input can be treated as an exogenous variable in the regressions. Due to the fortunes of location, presence of irrigation, and other factors, possibly including the state of the factor markets, some farmers are able to harvest their land more than once a year. My aggregation over the crop year warrants further discussion. First, does this aggregation over plots itself induce the inverse relationship? Actually, in this sample, the inverse relationship is estimated to be the same for multiple and single-cropping farmers. More seriously, does this aggregation hide the main margin of adjustment that labor-market constrained farmers may employ? As emphasized in several recent papers (e.g. Feder, 1985; Eswaran and Kotwal, 1986; Carter and Kalfayan, 1988; Carter and Wiebe, 1990) when faced with imperfect factor markets, farmers will adjust the intensity of their farm operation. This would entail multiple cropping where possible, and cultivating all available land. By taking the level of operation as given, I implicitly assume away this margin of adjustment. Of course, it is still an interesting question as to whether given this scale, there is evidence of inverse relationship resulting from further intensity in labor use. In these data, however, there does not seem to be evidence that intensity of land use is associated with farm size. For example, intensity as measured by the number of crops grown in a year is unrelated to farm size, controlling for irrigation and *kabupaten* (county) effects. In my 1992 paper (page 315) I show that the scale coefficient in the labor demand equation is unaffected by allowance for the endogeneity of farming (land use) intensity, given farm size. [7] Given the specificity of rice land, this is probably a reasonable result. Most adjustment to imperfect factor markets would thus occur with labor use. It must also be emphasized that the main argument of this paper is that the land input is statistically endogenous: it is correlated with the error term of regressions like those in Eq. (1): in the end we must account for this endogeneity. Finally, in order to maintain comparability with previous studies, and especially to avoid this additional source of endogeneity, I present the results of the total value regressions with a total (available) land measure as the measure of scale. [8]

Table 1 presents means of key variables for the rice farmers in our sample. Several points are worth noting. First, there is an accounting problem of the appropriate measure of total input costs, and therefore profits. One method is to impute to all non-purchased inputs the geographically nearest possible price. With

[7] In that paper, I instrument area harvested with available land (farm size) and show that this does not affect the inverse relationship. An identical result emerges for the output equation. This suggests that the inverse relationship in this sample is unaffected by variation in intensity of harvested land relative to available land.

[8] When the IR equations are estimated with total *sawah* (rice land) instead of area harvested, the IR is slightly more extreme. This is consistent with a hypothesis of measurement error. Also consistent with the measurement error hypothesis, once the equations are estimated by IV, the results for all three equations are virtually identical to those presented later in the paper.

Table 1
Sample means [a]

	Primary sample [b]	Expanded sample [c]
Total land [d] (all crops-Ha.)	0.76	0.73
	(0.01)	(0.01)
Area harvested (Ha./year)	0.71	0.65
	(0.02)	(0.02)
Total output (all crops)	2754.47	2537.13
	(49.3)	(42.8)
Rice output	2081.12	1883.65
	(42.78)	(36.93)
Pre-harvest rice labor (days/year)	84.45	76.21
	(1.70)	(1.47)
Imputed profits [e] (rice) (output less all costs)	1273.11	1155.95
	(26.77)	(23.29)
PROFITS [f] (rice) (output less labor costs)	1752.95	1586.47
	(37.17)	(32.13)
Use HYV	0.72	0.69
	(0.01)	(0.01)
Not irrigated	0.40	0.42
	(0.01)	(0.01)
% with planting wage	100	84
Sample size	4605	8486

[a] Values in Kg. of rice unless otherwise stated (standard errors in parentheses)
[b] Primary sample is the full sample, less 51 farms with negative PROFIT, and 881 farms with no planting wage.
[c] Expanded sample is the full sample less 14 with no pre-harvest labor, and 104 farms with negative PROFITS. Those farms with no planting wage are given an imputed planting wage, based on the geographically closest mean observation.
[d] Total land is the size of land operated by the farmer for all crops, whereas area harvested is the total area of rice land harvested over the year.
[e] Imputed profits (all costs) is the difference between output and all costs, where inputs are imputed a value equal to the geographically nearest price.
[f] PROFITS is the same as above, except that only pre-harvest labor costs are subtracted from output (revenue).

this method, pre-harvest labor costs comprise 42% of total costs, by far the largest category. The preferred profit measure (PROFITS) is the difference between revenue and labor costs. While it yields a less accurate estimate of profits, it fits the definition of profits used in the theoretical model to follow. Only with this measure will the shares of costs add up (there are only two inputs modeled: labor and land). [9] Furthermore, as shown in Benjamin (1989), the results do not depend on the chosen profit measure. The second issue arises in the choice of the wage, a key ingredient in the empirical model. I choose the planting wage because it yields the largest sample of farmers, and because it represents the price of the largest category of labor. As noted in my 1992 paper, the choice of wage does not effect the estimates in the labor demand equation, where the dependent variable, as is the case here, is total pre-harvest labor. [10] The bigger potential problem with choosing any wage measure is that not all farmers hire labor, so a wage is not always observed. Since sample selection would occur based on the presence of a regressor (rather than the dependent variable), selectivity bias need not arise. However, it is the farmers who do not hire labor for whom the market wage may not be the relevant price of labor. It is important to determine whether inclusion of these farmers affects the fit of the full neoclassical model. I do this by assigning these farmers an imputed planting wage based on the geographically nearest observations. Throughout the paper I will compare results with two samples: the primary sample with an intact, complete set of regressors, and an expanded sample with imputed wages. Finally, since the models will be estimated in logarithms, I use only a sample of farmers with positive ex post profits. This results in a loss of only 1% and 2% of farmers for the two samples, and is thus unlikely to result in selectivity bias.

1.3. Empirical presentation of the inverse relationship

This section establishes some of the statistical properties of the inverse relationship. The equations of interest are variations of (1), estimated on the expanded sample with controls for irrigation and the use of high yielding varieties of rice. [11] Column 1 of Table 2 shows the OLS results. The top panel shows the inverse relationship for total value with total land as the scale measure. The land coefficient is 0.70, significantly less than 1.0. The bottom three panels show the estimates for the components of the labor demand model. The coefficients for output and profits are 0.90 and 0.97, while for labor it is 0.66. At 0.01, the standard errors are quite small. The inverse relationship is therefore stronger for

[9] In my 1992 paper, I show that including other input prices (fertilizer and pesticide are the only other variable inputs with available prices) does not affect the coefficients of interest.

[10] The 1992 paper explores the bias of aggregating pre-harvest labor by separately modeling the labor done by males and females. The results in that paper support the pooling of pre-harvest labor, treating the planting wage as the price of that aggregate. The underlying reason for this is the high correlation between wages for various tasks.

[11] The irrigation control is an indicator of whether the farm has no source of irrigation (or equivalently, whether it has any irrigation), while the HYV control is an indicator of whether the farmer used a HYV of rice. More elaborate controls for irrigation, such as distinguishing between full-year and part-year availability, does not affect the results.

Table 2
Demonstration of the inverse relationship [a]

	OLS	Within	GLS
Log total value of output			
Log total land	0.72	0.82	0.78
	(0.01)	(0.01)	(0.01)
R-squared	0.49	0.60	0.78
F-test for cluster effects	4.22		
Fixed effects Hausman test	40.6		
Log output (rice)			
Log area harvested	0.90	0.87	0.89
	(0.01)	(0.01)	(0.01)
R-squared	0.69	0.74	0.85
F-test for cluster effects	4.32		
Fixed Effects Hausman Test	6.39		
Log labour (rice)			
Log area harvested	0.66	0.69	0.68
	(0.01)	(0.01)	(0.01)
R-squared	0.51	0.62	0.68
F-test for cluster effects	4.70		
Fixed effects Hausman test	8.38		
Log profits (rice)			
Log area harvested	0.97	0.94	0.95
	(0.01)	(0.01)	(0.01)
R-squared	0.60		0.75
F-test for cluster effects	3.57		
Fixed effects Hausman test	4.30		

[a] Standard errors in parentheses. Only land coefficients are shown. Sample size = 5486. Number of clusters = 2127. The value of total output includes all farm output (as opposed to rice only) and total land is the size of farm, which may not correspond to the area of rice harvested. The additional regressors in each specification are dummy variables indicating (1) the use of HYV rice and (2) no irrigation. The *F*-test is a test for the significance of the cluster fixed effects, $F(2127, 5486)$. The Hausman test is a test for equality of the land coefficients in the fixed effect and GLS specification. Under the null hypothesis, this is $\chi^2(1)$. The log(sample size) = 8.61 for the Schwarz criterion, while the 5% and 1% critical values are 3.84 and 6.63.

labor than output, and smallest for profits. The purpose of this paper is to try to reconcile or explain these coefficients. [12]

The hypothesized omitted variable can be regarded as any type of unobserved

[12] In Benjamin (1991, 1989) I use non-parametric techniques to explore the validity of the log-linear relationships that are the focus of this paper. The results of this estimation strongly support the constancy of the inverse relationship across farm sizes: the OLS coefficients accurately summarize the inverse relationship. This is quite different than studies like Carter and Wiebe (1990), where the inverse relationship reverses for large farms. The results for Java are probably different because farm sizes are so small: most farms are less than 2 Ha.

heterogeneity. Indeed, a more general objective of the paper is to estimate the labor demand model in the presence apparent unobserved heterogeneity. However, it is useful to label the heterogeneity land quality, since many researchers have suggested this as the most likely candidate. As is well known, omitting a variable from a regression is only a problem if it is correlated with the regressors. In the univariate case, the bias on the estimated land coefficient is $\gamma\rho_{hA}$, the covariance of land quality with the area harvested (ρ_{hA}), scaled by the effect of land quality on the dependent variable, γ. Since we expect land quality to vary over space, a natural means of addressing the possible omitted variables bias would be to exploit the spatial nature of the data. To determine the spatial nature of the error term f_{ij}, cluster fixed-effect regressions were estimated. If the error term has the structure

$$f_{ic} = \lambda_c + \epsilon_{ic}$$

where λ_c is a cluster or village effect, and ϵ_{ic} is a white noise error term, then a within-cluster estimator should yield consistent estimates of the land coefficients by purging the cluster level heterogeneity. [13] Columns 2 and 3 of Table 2 show the results of accounting for cluster fixed effects. For the total value equation, the land coefficient increases from 0.70 to 0.81, as one would expect if unobserved land quality was a problem. For output and profits, the land coefficients, however, become slightly smaller. As suggested earlier, the total value inverse relationship has built-in more heterogeneity than the ones based on rice output alone. Nevertheless, accounting for cluster effects brings it into line with the magnitude of the rice output coefficient. In the labor equation there is a small upward movement of the land coefficient, consistent with the a negatively correlated cluster effect. The F-tests for cluster effects suggest that there is variation of the intercept terms across Java. Note, however, that there are over 2100 clusters, some of which have only one observation, yet the F-statistics only range from 3 to 5. These statistics would be insignificant using a large sample critical value, such as the Schwarz criterion. Our primary interest, however, is whether the fixed effects (cluster land quality) are correlated with the regressors, and thus the source of the inverse relationship. The Hausman tests in Table 2 test the equality of the OLS and fixed-effects land coefficients. More precisely, in this case with the clustering of the data, a random effects estimator is more efficient than OLS, so the Hausman test compares the fixed-effects estimates to the GLS (random effects) estimates. For the total value equation, the Hausman test suggests that OLS is biased. For the other three, the Hausman tests are marginally significant at the 5% level, though only the labor coefficients are different at the 1% level. Even these coefficients are insignificantly different by the Schwarz criteria.

[13] This is precisely the approach taken in Carter (1984) in response to Sen's (1975) observation that the village level studies yielded less striking inverse relationships than the country or cross-country level studies.

The most interesting feature of these estimates, then, is not the statistical significance of the cluster effects on the land coefficients, but rather the small size of the differences. The land coefficients in the OLS and the fixed-effect specifications are very similar, within 0.03 of each other. This yields useful information on the nature of the omitted variable. Consider the univariate case for simplicity. For notational convenience, I will suppress the equation subscript, and denote ln h by h. The within-cluster and OLS estimates will be equal if

$$\operatorname{plim} \frac{1}{n}\{h'f\} = \operatorname{plim}\left\{\epsilon'\left(h - \bar{h}_c\right)\right\}.$$

Since

$$\begin{aligned} h'f &= \left(h - \bar{h}_c\right)' \lambda_c + \left(h - \bar{h}_c\right)' \epsilon + \bar{h}'_c \lambda_c + \bar{h}'_c \epsilon \\ &= \left(h - \bar{h}_c\right)' \epsilon + \bar{h}'_c f, \end{aligned}$$

the equality of the two estimators holds only if $\bar{h}'_c f = 0$, i.e. the heterogeneity that is correlated with farm size is at the *individual* rather than the cluster level. This would imply that cluster average farm size is not correlated with the omitted variable if this is the appropriate model. Only individual farm size is correlated. This suggests two possibilities. First, the simple cluster fixed effects model does not represent the heterogeneity, so that the important heterogeneity that drives the inverse relationship is at the within-cluster level. Second, the coefficients might genuinely be less than one, and are thus unaffected by fixed effect estimation. If the latter is true, this would provide evidence in favor of the imperfect labor markets hypothesis, as concluded by Carter.

2. Interpretation of the inverse relationship

2.1. A simple economic model

From above, the empirical moments I wish to explain are the elasticities of output, labor, and profits with respect to a predetermined or fixed land input. Since my focus is on labor demand, I also model the wage elasticities of each equation. Consider a CES production function $q = f(h; L)$ with elasticity of substitution σ. h is the fixed input land, while L is the variable labor input. Assume there are constant returns to scale (CRS) and competitive labor and product markets. Profit maximization implies $\partial f / \partial L = w$, where w is the real wage. Simple differentiation yields the various elasticities. We wish to predict two sets of elasticities: η_{yh}, the elasticity of y with respect to land; and η_{yw} the elasticity of y with respect to the wage, where y refers variously to output (q), profit ($q - wL$), and labor (L). Farmers are assumed to maintain their first-order conditions for profit maximization through adjustment of their use of labor. The solution to the farmer's problem yields a restricted normalized profit function $\pi(w; h)$. Hotelling's Lemma then yields the output supply and labor demand functions, and their implied elasticities:

$$\eta_{qh} = 1 \qquad \eta_{Lh} = 1 \qquad \eta_{\pi h} = 1$$

$$\eta_{qw} = -\left(\frac{s_L}{s_h}\right)\sigma \quad \eta_{Lw} = -\left(\frac{1}{s_h}\right)\sigma \quad \eta_{\pi w} = -\left(\frac{s_L}{s_h}\right)$$

where $s_L = wL/q$ and $s_h = (q - wL)/q$ are the shares of labor and land in production. It can immediately be seen that the inverse relationship ($\eta_{qh} < 1$, $\eta_{Lh} < 1$) is inconsistent with constant returns to scale for this simple model.

2.2. Testing the theory

Table 3 presents the OLS estimates of the system of equations, as well as tests of cross-equation restrictions implied by the labor demand model outlined above. The estimated equations are the same as in Table 2, except I now include the wage as a regressor.

For the regressions:

$$\ln y_{ij} = \alpha_0 + \alpha_j \ln h_i + \beta_j \ln w_i + \sum_{k=1}^{c} \delta_{jk} X_{ki} + f_{ij}$$

where y_{ij} represents q, L, and π, the following hypotheses are tested:
(1) Constant returns to scale: $\alpha_q = 1$, $\alpha_L = 1$, $\alpha_\pi = 1$;
(2) Cross-equation wage restrictions: (a) $\beta_q = s_L \beta_L$; (b) $\beta_\pi = -(s_L/s_h)$; (c) 2(a) and 2(b); (d) 1 and 2(c)

While not fully appropriate, the shares are treated as constants and following conventional practice, each share is estimated at the mean. Given the clustering of the data, and the resulting possible heteroskedasticity, all inference (standard errors and hypothesis test) is based on heteroskedasticity-consistent estimated covariance matrices. [14]

Results are presented for each sample, and two specifications: the lean specification (controls for HYV and irrigation) as well as a specification with *kabupaten* (county) level controls for climate, soil, and rainfall features. The latter specification is intended to absorb some of the *kabupaten*-level land quality variation, and is particularly important for the instrumental variables estimates that follow. Again, I also show estimates of these regressions for the total value of output, with total land as the scale variable (though I do not incorporate this specification into the system). The estimates of the land coefficients in Table 3 closely resemble those for Table 2. With the inverse relationship, constant returns to scale is likely

[14] These are standard White-corrected covariance matrices for the OLS estimates, and the analogous correction for IV estimates (see White, 1982).

Table 3
Estimates of system of equations by OLS [a]

	Sample 1				Imputed wage sample			
	Ln *TV*	Ln *Q*	Ln *L*	Ln *P*	Ln *TV*	Ln *Q*	Ln *L*	Ln *P*
Specification 1 (controls for HYV and irrigation)								
Log total land	0.70			0.71				
	(0.01)			(0.01)				
Log area harvested		0.87	0.69	0.93		0.88	0.69	0.95
		(0.01)	(0.01)	(0.01)		(0.01)	(0.01)	(0.01)
Log wage	0.20	0.18	−0.28	0.13	0.18	0.17	−0.28	0.11
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)
R-squared	0.51	0.69	0.55	0.60	0.50	0.70	0.54	0.60
Wald: C.R.S. (df = 3)		738				856		
Wald: Wage ($L-Q$) (df = 1)		194				197		
Wald: Wage ($L-P$) (df = 1)		163				146		
Wald: Both wages (df = 2)		194				201		
Wald: All restrictions (df = 5)		942				1063		
Specification 2 (Sp.1 plus Kabupaten characteristics)								
Log total land	0.71				0.71			
	(0.01)				(0.01)			
Log area harvested		0.88	0.71	0.93		0.89	0.71	0.96
		(0.01)	(0.01)	(0.01)		(0.01)	(0.01)	(0.01)
Log wage	0.20	0.17	−0.24	0.11	0.19	0.16	−0.22	0.09
	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)
R-squared	0.53	0.71	0.58	0.62	0.53	0.71	0.57	0.62
Wald: Climate (df = 3)		40	52	40		28	65	26
Wald: Rain (df = 5)		38	60	28		56	86	45
Wald: Soil (df = 4)		36	44	27		30	32	24
Wald: C.R.S. (df = 3)		680				781		
Wald: Wage ($L-Q$) (df = 1)		171				173		
Wald: Wage ($L-P$) (df = 1)		138				122		
Wald: Both wages (df = 2)		171				178		
Wald: All restrictions (df = 5)		851				953		

[a] Standard errors in parentheses. All standard errors are corrected for arbitrary forms of heteroskedasticity. Sample size for sample 1 is 4605, and for the sample with imputed wages is 5486. Ln TV is log total value of farm output, including non-rice output. Log total land is total area of farm land (sawah and non-sawah) operated by the farmer. The Ln TV equation is estimated separately from the other specifications. Kabupaten characteristics are indicators for the presence of 4 soil types, 3 climatic zones, and 5 rainfall zones in each kabupaten, and an indicator of sugar regency. All equations also include controls for irrigation and the use of HYV.

to be rejected. The Wald tests confirm (!) this suspicion. The wage coefficients are interesting as well. As the Wald tests of cross-equation restrictions indicate, the configuration of wage coefficients bears no relationship to that which would be predicted by the theoretical neoclassical model. [15] The wage coefficients in the output and profit equations are significantly positive, entirely at odds with the theory. Furthermore, the wage elasticity in the labor demand equation is consistent only with a very low elasticity of substitution of around 0.2. These wage elasticity estimates suggest that if the theory is to be rescued, there must exist a positive bias on these coefficients. If land quality is the omitted variable, it is not unreasonable to imagine that wages might be positively correlated with it. Results are very similar across specifications and samples. The wage elasticities are slightly smaller with controls for *kabupaten* characteristics, and for the sample with farmers having imputed wages. The smaller wage effects in this sample may result from measurement error implicit in the imputation of wages, or the sample may include farmers who are not behaving as price-responsively as those farmers who hired labor. With these coefficients in mind, I turn to an econometric model incorporating the possible omitted variables bias with the economic model.

3. Correcting for omitted variable bias

3.1. Omitted variable bias

Consider the standard regression equation

$$y = X\beta + \gamma z + \epsilon$$

where X is a vector of J explanatory variables, z is an omitted regressor, and ϵ is a classical error term. If we regress y on the X's alone, it is well known that

$$\text{plim } \tilde{\beta}_j = \beta_j + \gamma\rho_j$$

where $\tilde{\beta}_j$ is the OLS estimator of β_j and ρ_j is the jth element of the coefficient vector $(X'X)^{-1}X'z$, the regression of z on X. In the current context, let $\log A$ represent the omitted variable from each of the equations. This causes an omitted variable bias of $\gamma\rho_{wA}$ and $\gamma\rho_{hA}$ on the wage and land coefficients, where ρ_{wA} and ρ_{hA} are the wage and land coefficients from the regression of $\log A$ on $\log w$, $\log h$, and the other x-variables, and γ is the effect of $\log A$ on $\log y$. The parameter γ depends on how the omitted variable affects production.

The ρ_{wA} and ρ_{hA} terms embody the endogeneity of area harvested and the wage. Even if they are not choice variables for the farmer, area harvested and the wage rate are not exogenous to the equation if they are correlated with omitted

[15] The neoclassical model embodies (at least) two distinct assumptions. The first is profit maximization, or optimizing behavior. The second is market-clearing, or price taking behavior. It is primarily the second component that has been challenged by other researchers, so that departures from neoclassical behavior have been attributed to market imperfections rather than non-optimization. I maintain this interpretation of the hypothesis tests of the neoclassical model.

variables. There are many possible explanations for such correlations. Regarding farm size, if farms were subdivided through inheritance over time, egalitarian motives on the part of the benefactor would result in higher quality parcels being divided more often than low quality parcels. This would impart a negative correlation between farm size and farm quality, *particularly* at the local or village level. Elements of land quality important in rice farming are the shape and relief of the parcel, the drainage capabilities of the soil, and the location of the plot relative to the flow of irrigation or other water resources. These components can vary considerably, even within a village. Regarding wages, where marginal products of labor are higher due to land quality, one might expect real wages to be higher, depending on the model of wage determination. While the actual model of wage determination is outside the scope of this paper, it is relevant to the interpretation of my results. Significant wage variation over space or between farms is evidence against the simplest competitive model. With mobility costs, a local supply and demand model might explain the variation. In this case, estimates of the wage elasticities would be contaminated by simultaneous equations bias. Alternatively, given the limited opportunity for entry into the rice industry (imperfect land markets), it would not be surprising if there were undissipated land quality rents and therefore more complicated wage determination than can be explained by a competitive model. The objective of my paper is to test whether farmer behavior is consistent with the price-taking model outlined earlier, against an alternative of severe labor market failure that has been suggested as the explanation of the inverse relationship. It is not a conclusive test of competitive labor markets.

3.2. *Instrumental variables estimation*

The joint allocation of land, wages, and land quality in Java was not accomplished by random assignment. As a result, the effect of unobservables like land quality may be convoluted with the effects of both wages and area harvested. We need a way of artificially randomly allocating the land and wages so that the land and wage effects can be isolated from the omitted unobserved variables. As is well known, consistent estimates of β can be obtained if we can find a vector of instruments W such that

$$\operatorname{plim} \frac{1}{n}\{W'X\} \neq 0, \tag{A}$$

$$\operatorname{plim} \frac{1}{n}\{W'[\gamma z + \epsilon]\} = 0. \tag{B}$$

In estimating labor demand parameters, a natural source of instruments may be underlying labor supply factors, especially if wages are determined in local labor markets. [16] As such, the key instruments used in this paper are measures of population density. In addition to their economic interpretation these instruments exploit the spatial nature of the data. It is critical, though, to examine the two conditions necessary for the validity of these instruments.

(A) The instruments should be correlated with the area harvested and the wage. This condition should be easily satisfied. First, in terms of area harvested or farm size, one might expect that over time in more densely populated areas that average farm size is smaller. Given limited land resources and mobility costs, population pressure could result in the gradual division of farms among more and more people, particularly through inheritance. Indeed, it was this division of farm size and associated intensity of labor that Geertz observed in Java and referred to as agricultural involution:

Slowly, steadily, relentlessly, they were forced into a more and more labor-stuffed sawah pattern of the sort the 1920 figures show: tremendous populations absorbed on minuscule rice farms, particularly in areas where sugar cultivation led to improved irrigation; ....Wet rice cultivation, with its extraordinary ability to maintain levels of labor productivity by always managing to work one more man in without a serious fall in per capita income, soaked up almost the whole of the additional population that Western intrusion created, at least indirectly. It is this ultimately self defeating process that I propose to call 'agricultural involution' (Geertz, 1963, p. 80).

In Java this practice of cutting farms into smaller pieces was eventually viewed as a sufficient problem that the government imposed minimum farm size requirements as a component of its Basic Agrarian Law of 1960 (Article 17). As MacAndrews (1986) notes, however, without compulsory land registration, compliance with this component of land reform has been limited.

Related to Geertz' idea again, is the view that wages would be negatively correlated with population density. While not stated as such, his view that population pressure would result in greater labor absorption can be expressed in terms of a simple local supply and demand for labor system with mobility costs. Higher population density corresponds to shifts of the local labor supply curve, a lowering of the equilibrium wage, and higher intensity of labor use. Of course, this condition can be established empirically be examining the t-values of the instruments in the first stage equations.

(B) The instruments must themselves neither belong in the regression or be correlated with the omitted variable, land quality. This is usually the more difficult

[16] As noted in Hamermesh (1993), the supply side of the labor market is a common source of would-be reasonable instruments for the wage in labor demand equations.

condition to satisfy. Unlike condition (A), there is no definitive empirical means of establishing its validity. In the present case, however, there are several arguments which reinforce this important orthogonality assumption.

First, as was shown in the previous section, most of the variation of the omitted variable appears to be at the within-cluster level, rather than between clusters. The population density measures that I use are *kabupaten* (county) level measures which are, by construction, orthogonal to the within-village (deviation from village means) and within-county heterogeneity. Land quality covariation with population density *within* counties will thus not pose a problem for these instruments. The IV regressions therefore identify parameters principally off the between-*kabupaten* variation. The critical question is then whether the *kabupaten* level population density is correlated with *kabupaten* level omitted land quality.

As it turns out, there are extensive discussions in Horstman and Kurz (1980), and Geertz (1963) regarding the issue of whether land quality determined settlement patterns in Java. Both discussions summarize the Mohr–Kuperus debate of 1938. Mohr, a geologist and soil engineer, was particularly interested in the relationship between soil quality and population density. He postulated (quite reasonably) that "it is possible to trace and point out the correlation between the density of the population and (a) the nature of the soil, (b) hydrographic conditions, (c) the resulting use of the soil, and (d) the agricultural returns". It is principally (a) that threatens the validity of the instruments. Horstman and Kurz argue that Mohr overstated his case, that in fact his research had more accurately found empirical evidence of a correlation between population density and the rate of agriculturally used area and the rate of irrigated area. At the 13th International Geographics Congress in 1938, a heated debate arose, with Kuperus and several colleagues vehemently criticizing Mohr's conclusions. Kuperus argued that rather than soil quality, population density depended on a series of other factors. In particular (as emphasized in Geertz (1963)) the irrigated sugar estate areas were areas of high population density, as these estates historically were a leading source of employment. If these regions have more productive soils, then we should control for this feature in the regressions. Similarly, the development of cities was argued to be independent of soil quality, but important in determining population density patterns. Booth (1985) summarizes other historical evidence on the interlinkage of settlement patterns and agriculture. She cites evidence that Regency (colonial 'counties') in-migration was correlated with agricultural potential. She also notes that this linkage has probably been weakened more recently as non-agricultural opportunities have probably come to dominate migration decisions. [17] Perhaps the main consensus that arises from these discussions is that there is no definitive evidence that *kabupaten*-level population density in 1980 and

[17] Of course, flows of migration to counties need not result in a correlation of the final stocks of population and land quality, since inertia and other factors besides agricultural potential may dominate.

unobserved *kabupaten*-level land quality are correlated. This does not mean, however, that they are not, so several additional steps must be made to bolster confidence in the validity of the instruments.

To further establish the validity of the *kabupaten*-level instruments, I include broad measures of *kabupaten*-level 'land quality'. The variables are:

(1) Indicator for presence of each of four soil types in the *kabupaten*.

(2) Indicator for presence of each of three climatic zones in *kabupaten*.

(3) Indicator for presence of each of five rainfall zones in *kabupaten*.

These measures were taken from maps developed and presented in Donner (1987). In addition to these direct measures of growing conditions, following Geertz' suggestion, I include an indicator for whether the *kabupaten* was a sugar Regency, that is *kabupaten* in which sugar plantations were at one time important. The designation of these *kabupatens* is taken from Geertz (1963). Inclusion of these variables serves two purposes. First, they should directly absorb some of the residual correlation that may exist between population density and *kabupaten*-level land quality. Second, by comparing specifications with and without these *observed* land quality measures, we can speculate on the possible consequences of omitting *unobserved kabupaten* land quality measures. These variables also allow a simple exploration of the relationship between land quality and population density. Of these land quality variables, only the indicator for sugar regency is correlated with the population density measure used in this paper. [18] This suggests that 'sugar regency' is a variable that may belong in the regression, and also provides further evidence that my exclusion restrictions may be valid.

The most important step in empirically establishing the validity of the instruments is to conduct tests of the overidentification assumptions, i.e. test the exclusion restrictions employed in the 2SLS estimation. Essentially, the overidentification tests are mutual consistency tests of the instruments and are designed to test whether the instruments are correlated with the estimated residuals. These tests can only establish that necessary, not sufficient conditions for identification are met, but passing these tests indicates no evidence against the orthogonality assumptions. Indeed, failure of these tests is evidence of a problem with the exclusion restrictions in a *particular* equation, and is evidence of inconsistency for that *particular* specification.

Finally, consider the consequences of using an invalid instrument. For 2SLS estimation:

$$\text{plim}\ \hat{\beta} = \beta + \text{plim}\left(\hat{X}'X\right)^{-1}\hat{X}'u$$

[18] In a regression of the log of population density (the measure used in this paper) on the land quality variables, as well as controls for irrigation (fraction of farmers not-irrigated) the soil, rainfall, and climate variables were each (as a group) jointly insignificant at the 5% level. This is not direct evidence that the instruments are uncorrelated with unobserved land quality, only that non-correlation is plausible.

where $\hat{X}$ is the predicted X from the reduced form, and u is the error term. Let $\hat{X} = \hat{\Pi} W$, where W is the set of instruments and $\hat{\Pi}$ is the estimated first-stage parameters. The 2SLS estimator is inconsistent when some component of $\hat{X}$ is correlated with u. This would arise if one of the excluded instruments, w, is correlated with u. The sign of the resulting bias would then depend on the sign of $\pi w' u$. For the case at hand, if population density, w, and u (land quality) are positively correlated, then the bias will depend on the sign of π, the relationship between population density and the instrumented X (area harvested and the planting wage in our case). Since π is estimated to be strongly negative in these data, this suggests a negative bias would exist with the 2SLS estimates of β if population density and land quality are correlated. We can use this result when evaluating the empirical results.

For the IV estimation, the excluded (identifying) instruments are:
(1) Log of the ratio of rural population to area of sawah,
(2) Indicator for presence of city in *kabupaten* (designated Kota Madya),
(3) Number of males between 10 and 15 years of age,
(4) Number of females between 10 and 15 years of age.

The first two are *kabupaten*-level measures of population density, while the last two are household-level variables, included in particular to improve the fit of the first-stage land regressions. While my above discussion focuses on the population density instruments, the use of household-level variables themselves may not be innocuous, as they too may be correlated with the unobserved heterogeneity. This is not likely to be a problem in this case, however. First, the focus of my previous paper, Benjamin (1992) was precisely on the question of whether household demographic variables could be excluded from the labor demand equation, and the results strongly supported this conclusion. Second, the overidentification tests have the most power when the excluded variables are 'different' so that overidentification is genuine. Thus using household and regional variables is helpful in establishing identification. Passing the overidentification tests is more of an accomplishment in this case, and further reinforces the validity of the instruments. Finally, alternative measures of population density were also used, yielding essentially similar results. These are catalogued in Benjamin (1989).

Table 4 gives the first-stage equations for (log) area harvested and the (log) wage. These regressions confirm that the excluded instruments help in predict both wages and area harvested. This is important in establishing identification (condition (A)). For wages, the number of children has no effect, while both types of measures of population pressure (population density and city) have statistically significant negative effects, consistent with their labor supply interpretation. [19] In addition, HYV use has very little effect on wages, while wages are lower in

[19] The absence of a relationship between 'teenagers' and the wage further reinforces the belief that there is no relationship between unobserved productivity and the number of teenagers in the household.

Table 4
First-stage regressions for labor demand system [a]

	Specification 1				Specification 2			
	Primary sample		Expanded sample		Primary sample		Expanded sample	
	Log *W*	Log *H*	Log *W*	Log *H*	Log *W*	Log *H*	Log *W*	Log *H*
Intercept	0.71	−1.69	0.72	−1.79	0.80	−1.88	0.80	−1.91
	(0.03)	(0.05)	(0.03)	(0.05)	(0.07)	(0.10)	(0.06)	(0.09)
Log pop ratio	−0.19	−0.48	−0.19	−0.47	−0.11	−0.39	−0.11	−0.39
	(0.02)	(0.03)	(0.02)	(0.03)	(0.02)	(0.04)	(0.02)	(0.03)
City	−0.05	−0.03	−0.05	−0.05	−0.08	−0.10	−0.09	−0.11
	(0.02)	(0.03)	(0.02)	(0.03)	(0.02)	(0.03)	(0.02)	(0.03)
Males 10–15	0.01	0.13	0.01	0.13	0.01	0.12	0.01	0.12
	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)
Females 10–15	−0.01	0.15	−0.01	0.16	0.00	0.16	−0.00	0.17
	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)
HYV	−0.04	0.25	−0.04	0.30	0.01	0.27	0.00	0.28
	(0.02)	(0.03)	(0.02)	(0.03)	(0.02)	(0.03)	(0.02)	(0.03)
Not irrigated	−0.06	−0.20	−0.05	−0.27	−0.03	−0.16	−0.02	−0.17
	(0.02)	(0.03)	(0.02)	(0.03)	(0.02)	(0.03)	(0.02)	(0.03)
Wald: Climate					56	9	86	13
Wald: Rain					186	96	208	113
Wald: Soil					17	41	32	51
R-squared	0.03	0.12	0.03	0.12	0.10	0.14	0.10	0.14

[a] Standard errors in parentheses. All standard errors are corrected for arbitrary forms of heteroskedasticity.
Sample size for Primary sample is 4605, and for the Expanded sample is 5486. The Wald test statistics are $\chi^2(d)$, where $d = 3$ for climate, $d = 5$ for rain, and $d = 4$ for the soil indicators. The sample means for the excluded instruments (Primary sample) are as follows: Log pop ratio = −1.41; City = 0.24; Males 10–15 = 0.42; Females 10–15 = 0.37. For the expanded sample, the means are Log pop ratio = −1.40; City = 0.24; Males 10–15 = 0.41; Females 10–15 = 0.36.

non-irrigated areas. Regarding area harvested, we find that the number of teenagers is positively related to farm size, while measures of population density have the expected negative effect. It is also clear that HYV use and lack of irrigation have the expected significant effects on the area harvested. Including the 'land quality' measures has a small but noticeable effect on the population density coefficients.

3.3. Results

Table 5 presents the results of 3SLS estimation of the system of equations. Since the regressors and excluded instruments are the same in each equation, 3SLS is equivalent to equation by equation 2SLS. However, in order to test the cross-equation restrictions, the equations are estimated as a system. Note, though, that there are no efficiency gains from 3SLS over 2SLS. Specification tests are presented for each equation. Given the large sample size, I present results of hypothesis tests employing the more conservative Schwarz criterion which accounts for sample size, as well as the usual 5% results. The most stunning result in these tables is the elimination of the inverse relationship. The estimated land effects are virtually numerically and statistically indistinguishable from 1.0. Formal tests of constant returns to scale across equations now indicate that CRS is not rejected. That these coefficients should all simultaneously go to 1 is unlikely if the results were due to inappropriate instruments. Equally dramatic is the effect on the wage coefficients. For the output and profit equations, a significant positive elasticity has become negative or at least insignificantly different from zero. The labor demand elasticity has increased in magnitude to -1.1 or -1.5 and is still statistically significant. [20] The cross-equation wage restrictions now pass in most cases, at least at the 1% level. In the total value equation, the land coefficients also go to one, and the wage coefficients become significantly negative. This suggests that the elimination of inverse relationship is not sensitive to the output or scale measures.

Comparing results with and without controls for land quality, we see very little difference in the estimated coefficients, though the standard errors are higher in the specification with the extra regressors. The only discernible effect on the coefficient estimates is that the wage elasticities in the output and profit equations are less negative when the land quality indicators are included. In these equations at least, this suggests that a possible correlation between the instruments and the error term may manifest itself in an overstatement of the wage effects. Notice, however that the land coefficients do not change when these variables are added. Furthermore, the 2SLS land coefficients are higher than the OLS coefficients. As shown earlier, this would not happen if there was a positive correlation between population density and the unobserved land quality in these regressions. Therefore, the elimination of the inverse relationship does not seem to be an artifact of invalid instruments.

The comparison of results across samples is also very interesting. While insignificantly different from 1.00, the scale elasticity in the labor equation for the expanded sample is 0.93, which is much larger than OLS, but smaller than those in the primary sample. Furthermore, the wage elasticities are less elastic in this sample (the sample including those farmers who did not hire labor). This suggests that the tests for neoclassical labor demand behavior have statistical power. Inclusion of possibly constrained farmers who do not hire labor in the sample changes the coefficient estimates in the direction one would expect. Equivalently,

[20] The coefficient estimates of these equations with area harvested replaced by area of sawah are virtually identical when estimated by this procedure.

exclusion of these farmers results in coefficient estimates more consistent with the neoclassical, market-clearing model.

The Hausman tests test whether the instrumented equations differ significantly from their OLS counterparts. With Specification 1, these tests indicate large differences in the coefficients, emphasizing the endogeneity problems for farm size and the wage. These tests are particularly strong for the labor equation, which is not surprising given that it is in this equation that the wage and land coefficients consistently change most dramatically. The results are less dramatic for Specification 2, where only the labor equation differs significantly. In comparing the two specifications, this difference in the Hausman tests arises primarily because the standard errors in specification 2 are larger. This underlines an important caveat in interpreting the hypothesis tests in both sets of results. With large enough standard errors, any configuration of estimated coefficients would be consistent with the theory. Because most of the identification is coming from region-level variables, especially for the wage, the resulting lack of precision of the IV estimates warrants caution in the enthusiasm with which one views the elimination of the inverse relationship. Nevertheless, the important result is that the estimates move in the direction predicted by the omitted variables bias hypothesis, and furthermore, that the labor demand equation, where the inverse relationship is most acute, is statistically significantly changed and consistent with the simple model. The overidentification tests indicate that we cannot reject the validity of the instruments, even at the 5% level. Again, the labor equation performs best, with the lowest overidentification test in each specification. Furthermore the overidentification tests have an interesting interpretation: The effect of population density on production and labor demand can be accounted for entirely through its effect on the instrumented regressors: area harvested and wages.

4. An econometric model of omitted land quality

4.1. Three ways in which land quality may affect production

The only way to be sure that land quality is the omitted variable in the regressions is to include a measure of individual farm quality in the regressions, and have the inverse relationship and nasty wage elasticities disappear. That would make a simple and convincing paper. However, lacking such a measure does not commit the omitted variable to unrestrained speculation. Taking the IV estimates as truth, there is exploitable information in the relative biases of the OLS estimates. In this section I construct a very simple model of omitted land quality that places testable restrictions on the biases across the output, labor, and profit equations. The purpose of the exercise is to determine whether an omitted variable like land quality could have generated the OLS equations. I discuss three ways which land quality might affect production, as summarized by the parameter γ.

Each has the property that for a given level of inputs, more output is produced on land of higher quality. In addition, for each case there are constant returns to scale in land and labor for land of a given quality. [21] The cases differ in the way the relative marginal products of land and labor are affected by land quality and essentially mimic the models of technical change in the production literature. [22]

4.1.1. Case 1: Neutral quality effect

In this case $q = f(Ah; AL)$, which given CRS implies $q = Af(h; L)$. This is the same type of model employed by Lau and Yotopolous (1971) in their approach to comparing technical efficiency across farms. This case predicts the following elasticities of output, labor, and profits with respect to the land quality measure:

$$\gamma_q = 1 + \left[\frac{s_L}{s_h}\right]\sigma, \quad \gamma_L = \left[\frac{1}{s_h}\right]\sigma, \quad \gamma_\pi = \left[\frac{1}{s_h}\right].$$

4.1.2. Case 2: Land augmenting quality effect

In this case, $q = f(Ah; L)$. Bhalla and Roy (1988) employ this functional form in their discussion of the role of land quality in agricultural production. Output depends on effective land, Ah. The true fixed factor is Ah rather than simply h. Without data on A this fixed input is essentially mismeasured. Since there are CRS with respect to both h and Ah, we get biased estimates of the coefficient on h only if A and h are correlated. The elasticities implied by this model are:

$$\gamma_q = 1, \quad \gamma_L = 1, \quad \gamma_\pi = 1.$$

Not surprisingly, since and A and h enter identically into the production function as the 'same' fixed factor, these elasticities are the same as the land elasticities. Note that the predicted bias is the same in each equation.

A traditional errors-in-variables model yields identical predictions to this case regarding the relative biases between the labor, output and profit equations. If instead of observing the true $\log h_i$, we observe $\log h_i^* = \log h_i + v_i$, then in a simple regression of $\log y$ on $\log h_i^*$, we get the usual measurement error result:

$$\text{plim } \tilde{\beta} = \frac{\beta}{1 + \sigma_v^2 / \sigma \log_h^2} < 1$$

where $\tilde{\beta}$ is the OLS estimator, σ_v^2 is the variance of the measurement error, and $\sigma_{\log h}^2$ is the variance of the log of true farm size. This measurement error even resembles land quality in interpretation. If we under-measured the land input, it

[21] See Lichtenberg (1989) for such an example of the incorporation of land quality into production analysis of agriculture.

[22] See Fuss and McFadden (1978), especially Vol. II for a discussion of the technical change literature.

Table 5
Estimates of system of equations by 3SLS [a]

	Primary sample				Expanded sample			
	Ln *TV*	Ln *Q*	Ln *L*	Ln *P*	Ln *TV*	Ln *Q*	Ln *L*	Ln *P*
Specification 1 (controls for HYV and Irrigation)								
Log total land	0.99				0.93			
	(0.11)				(0.08)			
Log area harvested		1.00	1.00	0.95		1.00	0.93	1.00
		(0.08)	(0.11)	(0.10)		(0.07)	(0.08)	(0.08)
Log wage	−0.59	−0.47	−1.52	−0.23	−0.38	−0.34	−1.22	−0.13
	(0.25)	(0.22)	(0.30)	(0.25)	(0.19)	(0.19)	(0.23)	(0.22)
R-squared	0.35	0.23	0.06	0.22	0.42	0.28	0.09	0.25
Wald: C.R.S. (df = 3)		3.11				2.51		
Wald: Wage $(L-Q)$ (df = 1)		2.00				1.30		
Wald: Wage $(L-P)$ (df = 1)		0.14				0.00		
Wald: Both wages (df = 2)		9.19 *				10.12 *		
Wald: All restrictions (df = 5)		19.4 **				26.9 *		
Overid test (df = 2)	1.44	1.38	0.46	2.76	0.55	3.36	0.56	0.56
Hausman test (df = 2)	10.4 *	13.7 *	21.3 **	7.9 *	9.6 *	8.80 *	20.1 **	1.6
Specification 2 (Sp.1 plus Kabupaten characteristics)								
Log total land	0.97				0.92			
	(0.10)				(0.07)			
Log area harvested		1.00	0.99	0.96		1.03	0.93	1.03
		(0.07)	(0.11)	(0.09)		(0.06)	(0.08)	(0.08)
Log wage	−0.11	−0.17	−1.55	0.19	0.18	−0.13	−1.11	0.19
	(0.37)	(0.26)	(0.41)	(0.33)	(0.28)	(0.21)	(0.27)	(0.27)
R-squared	0.49	0.34	0.09	0.30	0.52	0.36	0.14	0.30
Wald: C.R.S. (df = 3)		1.81				2.67		
Wald: Wage $(L-Q)$ (df = 1)		0.00				0.00		
Wald: Wage $(L-P)$ (df = 1)		0.98				1.44		
Wald: Both wages (df = 2)		5.56				8.8 *		
Wald: All restrictions (df = 5)		7.65				19.1 *		
Overid test (df = 2)	0.41	0.46	0.46	0.92	0.10	3.36	1.12	1.17
Hausman test (df = 2)	15.4 *	2.71	10.5 *	0.73	29.8 **	5.43	11.4 *	3.93

[a] Standard errors in parentheses. All standard errors are corrected for arbitrary forms of heteroskedasticity.
Sample size for primary sample is 4605, and for the sample with imputed wages is 5486. Excluded instruments for the system of equations are log population density, a city indicator, and the number of males and females between 10 and 15 years old.
Ln *TV* is log total value of farm output, including non-rice output. Log total land is total area of farm land (sawah and non-sawah) operated by the farmer. The Ln *TV* equation is estimated separately from the other specifications. Excluded instruments for the system of equations are log population density, and the number of males and females between 10 and 15 years old. Kabupaten characteristics are indicators for the presence of 4 soil types, 3 climatic zones, and 5 rainfall zones in each kabupaten, and an indicator of sugar regency. All equations also include controls for irrigation and the use of HYV.
* Indicates rejection at 5% level.
** Indicates rejection by Schwarz criterion.

Empirical: Estimated Coefficients	Theoretical		
Consistent	Neutral	Land Augmenting	Labor Augmenting
β_{yh}	1	1	1
β_{yw}	$-\left(\frac{s_L}{s_H}\right)\sigma$	$-\left(\frac{s_L}{s_H}\right)\sigma$	$-\left(\frac{s_L}{s_H}\right)\sigma$
β_{Lh}	1	1	1
β_{Lw}	$-\left(\frac{1}{s_H}\right)\sigma$	$-\left(\frac{1}{s_H}\right)\sigma$	$-\left(\frac{1}{s_H}\right)\sigma$
$\beta_{\pi h}$	1	1	1
$\beta_{\pi w}$	$-\left(\frac{s_L}{s_H}\right)$	$-\left(\frac{s_L}{s_H}\right)$	$-\left(\frac{s_L}{s_H}\right)$

Fig. 1. Mapping between empirical moments and theoretical model consistent coefficients (ignoring intercepts and other control variables).

would appear that more output was obtained from less land, that the land was more productive. One might then erroneously conclude that smaller farms were more productive.

4.1.3. Case 3: Labor augmenting land quality effect

This case takes the form $q = f(h; AL)$. The marginal product of labor is relatively more enhanced by land quality. The implied elasticities are:

$$\gamma_q = \left[\frac{s_L}{s_h}\right]\sigma \quad \gamma_L = \left[\frac{1}{s_h}\right]\sigma - 1 \quad \gamma_\pi = \left[\frac{s_L}{s_h}\right]$$

If the elasticity of substitution is low, and labor's share is high enough, labor use could decrease with land quality. This would occur since it would take less measured labor, L, to achieve the optimal amount of effective labor, AL. Over most ranges of σ and s_L, however, increased land quality raises the marginal product of labor such that labor demand is increased.

Fig. 1 and Fig. 2 summarize the predicted coefficients for both the consistent (3SLS) estimates, and the biased (OLS) estimates, as functions of σ, s_h, s_L, 1, ρ_{hA}, and ρ_{wA} for each of the cases outlined above. If ρ_{hA}, and ρ_{wA} are the each same across equations, as they ought to be if the equations have the same

Empirical: Estimated Coefficients	Theoretical		
Biased	Neutral	Land Augmenting	Labor Augmenting
β_{yh}	$1+\rho_{hA}\left[1+\left(\frac{s_L}{s_H}\right)\sigma\right]$	$1+\rho_{hA}$	$1+\rho_{hA}\left(\frac{s_L}{s_H}\right)\sigma$
β_{yw}	$-\left(\frac{s_L}{s_H}\right)\sigma+\rho_{wA}\left[1+\left(\frac{s_L}{s_H}\right)\sigma\right]$	$-\left(\frac{s_L}{s_H}\right)\sigma+\rho_{wA}$	$-\left(\frac{s_L}{s_H}\right)\sigma+\rho_{wA}\left(\frac{s_L}{s_H}\right)\sigma$
β_{Lh}	$1+\rho_{hA}\left(\frac{1}{s_H}\right)\sigma$	$1+\rho_{hA}$	$1+\rho_{hA}\left[\left(\frac{1}{s_H}\right)\sigma-1\right]$
β_{Lw}	$-\left(\frac{1}{s_H}\right)\sigma+\rho_{wA}\left(\frac{1}{s_H}\right)\sigma$	$-\left(\frac{1}{s_H}\right)\sigma+\rho_{wA}$	$-\left(\frac{1}{s_H}\right)\sigma+\rho_{wA}\left[\left(\frac{1}{s_H}\right)\sigma-1\right]$
$\beta_{\pi h}$	$1+\rho_{hA}\left(\frac{1}{s_H}\right)$	$1+\rho_{hA}$	$1+\rho_{hA}\left(\frac{s_L}{s_H}\right)$
$\beta_{\pi w}$	$-\left(\frac{s_L}{s_H}\right)+\rho_{wA}\left(\frac{1}{s_H}\right)$	$-\left(\frac{s_L}{s_H}\right)+\rho_{wA}$	$-\left(\frac{s_L}{s_H}\right)+\rho_{wA}\left(\frac{s_L}{s_H}\right)$

Fig. 2. Mapping between empirical moments and theoretical model inconsistent coefficients (ignoring intercepts and other control variables).

regressors and the same omitted variable, then the ratios or differences in biases between the OLS equations have empirically testable restrictions. I adopt two approaches in testing these restrictions. First, the implied restrictions across the OLS and 3SLS equations are directly tested. This approach most clearly conveys the flavor of the underlying econometric model. Second, since the OLS and the 'true' coefficients are functions of the lower order parameters, minimum distance techniques are employed to test more formally the goodness of fit for each of the three competing models of land quality.

4.2. Hypothesis tests

This exercise provides the intuition behind the more formal presentation that follows. Fig. 2 lists the respective formulae for the γ as functions of σ, s_L and s_h. The strategy is to use the 3SLS estimates to glean consistent estimates of σ (from the wage elasticity), then combine this with s_L, s_h and the formulae in Fig. 2 to predict the OLS biases. Since I will not be estimating ρ until later, all that will be predicted here are the relative biases. While σ is overidentified in Fig. 2, I will use the estimate of σ derived from the labor demand equation, since this is the best behaved equation in the system. In addition, the labor equation will serve as the reference for the biases in the other equations.

The notation is cumbersome, so it is worth reviewing. β' is the omitted variable biased coefficient estimate of β, and it is 'consistently' estimated by the OLS estimate of β, $\tilde{\beta}$. As shown before, $\beta' = \beta + \gamma\rho$, where β is the 'true' vector of coefficients, γ is the effect of the omitted variable, and ρ is defined earlier. β is consistently estimated by the 3SLS estimator and its estimate is $\hat{\beta}$. These parameters are subscripted according to their equation: β_q, β_L are β_π the β's from the output, labor and profit equations. From Fig. 2 the restrictions relating the biased OLS coefficients to the 3SLS coefficients are as follows.

Neutral effect

Land coefficients:

$$(1)\quad \frac{1-\tilde{\beta}_{yh}}{1-\tilde{\beta}_{Lh}} = \frac{1-s_L\hat{\beta}_{Lw}}{-\hat{\beta}_{Lw}}, \qquad (2)\quad \frac{1-\tilde{\beta}_{Lh}}{1-\tilde{\beta}_{\pi h}} = -s_h\hat{\beta}_{Lw}.$$

Wage coefficients:

$$(1)\quad \frac{\hat{\beta}_{qw}-\tilde{\beta}_{qw}}{\hat{\beta}_{Lw}-\tilde{\beta}_{Lw}} = -\frac{\left(1-s_L\hat{\beta}_{Lw}\right)}{\hat{\beta}_{Lw}}, \qquad (2)\quad \frac{\hat{\beta}_{\pi w}-\tilde{\beta}_{\pi w}}{\hat{\beta}_{Lw}-\tilde{\beta}_{Lw}} = -\frac{1}{s_h\hat{\beta}_{Lw}}.$$

Land augmenting effect

Land coefficients:

$$(1)\quad \tilde{\beta}_{qh}-\tilde{\beta}_{Lh}=0, \qquad (2)\quad \tilde{\beta}_{\pi h}-\tilde{\beta}_{Lh}=0.$$

Wage coefficients:

$$(1)\quad \tilde{\beta}_{qw}-\tilde{\beta}_{Lw}=0, \qquad (2)\quad \tilde{\beta}_{\pi w}-\beta_{Lw}=0.$$

Labor augmenting

Land coefficients:

$$(1)\quad \frac{1-\tilde{\beta}_{Lh}}{1-\tilde{\beta}_{qh}} = \frac{1+\hat{\beta}_{Lw}}{s_L\hat{\beta}_{Lw}}, \qquad (2)\quad \frac{1-\tilde{\beta}_{Lh}}{1-\tilde{\beta}_{\pi h}} = \frac{s_L}{s_h}\left(-\hat{\beta}_{Lw}-1\right).$$

Wage coefficients:

$$(1)\quad \frac{\hat{\beta}_{Lw}-\tilde{\beta}_{Lw}}{\hat{\beta}_{qw}-\tilde{\beta}_{qw}} = \frac{1+\hat{\beta}_{Lw}}{s_L\beta_{Lw}}, \qquad (2)\quad \frac{\hat{\beta}_{Lw}-\tilde{\beta}_{Lw}}{\hat{\beta}_{\pi w}-\tilde{\beta}_{\pi w}} = -\frac{s_L}{s_h}\left(\hat{\beta}_{Lw}+1\right)$$

Table 6
Wald tests of cross-equation restrictions implied by the omitted land quality model [a]

	Primary sample		Expanded sample	
	Spec (1)	Spec (2)	Spec (1)	Spec (2)
Neutral: Land coefficient	115.5 **	105.2 **	268.7 **	298.4 **
Land Aug: Land coefficient	267.9 **	255.4 **	336.7 **	316.7 **
Labor Aug: Land coefficient	0.6	0.5	117.1 **	381.2 **
Neutral: Wage coefficient	7.1 *	9.9 *	9.6 *	13.9 *
Land Aug: Wage coefficient	7.6 *	7.2 *	7.7 *	8.7 *
Labor Aug: Wage coefficient	0.1	1.4	2.6	18.7 *

[a] Based on regressions in Tables 3 and 5. Spec (1) is the specification with controls for HYV and irrigation, while Spec (2) is specification (1) with the additional kabupaten characteristics.
* Indicates rejection at 5% level.
** Indicates rejection by the Schwarz criterion.

These pairs of restrictions, land and wage pairs for each model, form the basis for a series of non-linear Wald tests. Note these implications are not independent of the CES profit maximization model, and represent a test of the joint hypothesis of that model as well as the specification of the omitted variable bias.

The results of these Wald tests are presented in Table 6. [23] In each of the sets of equations, the Land biases of the Neutral and Land Augmenting models are resoundingly rejected. The Land Augmenting model is constrained by the fact that it predicts equal biases across the set of equations. Only the Labor Augmenting model seems capable of predicting the magnitude of the difference between the labor and other coefficients (0.2) and therefore fares better in the hypothesis tests. The omitted variable model also seems consistent with the wage biases. The Labor Augmenting model performs best here as well, though the standard errors are such that the other models do not fare as miserably as they did in explaining the land biases. The results are uniformly better with the sample containing only farmers who hired labor. As mentioned before, this is not surprising. In the sample with the non-hiring farmers, the estimated wage elasticity is significantly smaller in the labor demand equation, and the implied σ is not high enough to explain the large inverse relationship in the labor equation.

4.3. *Minimum distance estimation of the structural model*

An intuitively appealing method of testing whether the above models can explain the data is to employ minimum distance (MD) techniques. With MD we

[23] The variance–covariance matrix necessary for this Wald test, and the subsequent minimum distance estimation is derived in Benjamin (1989). This matrix links the variances and covariances of the OLS and 3SLS estimates under the null that the OLS equations are biased due to omitted land quality and the parameters are consistently estimated by 3SLS.

Table 7
Results of minimum distance estimation [a]

	Neutral		Land aug		Labor aug	
	S (1)	S (2)	S (1)	S (2)	S (1)	S (2)
Primary sample						
σ	1.20	1.18	1.03	1.01	1.38	1.39
	(0.24)	(0.32)	(0.17)	(0.23)	(0.24)	(0.32)
ρ_{hA}	−0.15	−0.14	−0.17	−0.16	−0.56	−0.52
	(0.02)	(0.02)	(0.01)	(0.01)	(0.23)	(0.29)
ρ_{hw}	0.48	0.48	0.49	0.49	2.21	2.25
	(0.05)	(0.06)	(0.07)	(0.07)	(0.51)	(0.71)
Sum of squares	0.45	0.42	0.49	0.46	0.09	0.01
Goodness of fit	676	722	290	268	116	114
Expanded sample						
σ	0.93	0.80	0.86	0.76	1.22	0.58
	(0.17)	(0.18)	(0.13)	(0.16)	(0.14)	(0.14)
ρ_{hA}	−0.14	−0.13	−0.16	−0.15	−0.74	0.60
	(0.01)	(0.01)	(0.01)	(0.01)	(0.24)	(0.23)
ρ_{hw}	0.39	0.35	0.41	0.38	2.65	−0.71
	(0.04)	(0.05)	(0.06)	(0.07)	(0.54)	(0.66)
Sum of squares	0.31	0.27	0.28	0.20	0.09	0.49
Goodness of fit	799	832	354	324	162	133

[a] Based on coefficients in Tables 3 and 5. Standard errors in parentheses. S(1) is the specification with controls for HYV and irrigation, while S(2) is the specification with additional controls for kabupaten characteristics. 'Neutral' refers to parameter estimates using the neutral effects of land quality model, while 'Land aug' and 'Labor aug' refer to estimates based on the land and labor augmenting models.

can treat the vector of estimated coefficients, $\hat{B} = (\tilde{\beta}, \hat{\beta})$ as data, or the moments to be explained, and then determine whether the smaller dimension structural model can explain them. Essentially, Figs. 1 and 2 map the function $\hat{B} = g(\theta)$, where θ is the vector of structural parameters. In my estimation, I free up the land coefficients from $\hat{\beta}$, as well as the profit equation wage coefficient from $\tilde{\beta}$. The equally weighted minimum distance estimator solves

$$\min \text{ w.r.t } \theta \quad \left(\hat{B} - g(\theta)\right)'\left(\hat{B} - g(\theta)\right).$$

Chamberlain (1984) and Newey (1985) provide details on the econometrics of minimum distance techniques. MD estimates were calculated for all of the specifications using the 3SLS and OLS results in Tables 3 and 5. The results are presented in Table 7 which show (i) the estimates of σ, ρ_{hA}, and ρ_{wA} (ii) the corresponding standard errors for these parameters; (iii) the value of the minimized function (sum of squared deviations), and (iv) the goodness of fit statistic. This statistic is distributed as a χ^2 with degrees of freedom equal to the degree of overidentification. In my case, there are eight moments which identify the three structural parameters, so the degree of overidentification is 5.

While there are a lot of numbers to digest, several broad conclusions can be drawn. None of the models' goodness of fit statistics indicate that the moments are explained by the smaller structural model. Indeed, the goodness of fit statistic is rather an 'atrocity of fit' statistic. Despite this however, the results for the Labor Augmenting model are the least discouraging. For example, looking at the results for the primary sample, full specification, The Labor Augmenting model yields an estimate of $\sigma = 1.39$. As well, this model yields large estimates of the partial correlations between the omitted land quality variable and the land and wage variables ($\rho_{hA} = -0.52$, $\rho_{wA} = 2.25$) as we would expect. The resulting deviations that enter the minimum distance objective function (not shown), differences between the actual and predicted moments are quite small, and certainly smaller for this model than the other two models. This is also indicated in the low sum of squared residuals, and the relatively low goodness of fit statistic. For example, the predicted land coefficients rarely differ from the actual by more than 0.02. This contrasts with corresponding maximum deviations of 0.15 and 0.16 for the Neutral and Land Augmenting models. In addition, with the Labor Augmenting model, none of the deviations is individually statistically significant. It is only jointly (accounting for covariances of the deviations) that the deviations are statistically significant. This contrasts sharply with the results from the other models. Of interest as well is that the performance of the MD estimates closely follows that which was predicted in the informal discussion of the biases in the above section. In general, it is the wage coefficients that have the largest deviations, but these are often statistically insignificant due to the large standard errors on the underlying wage coefficients. With the expanded sample, full specification, it is important to note that even the Labor Augmenting model performs badly. As mentioned above, this is primarily due to the low value of σ implied by the wage elasticities. Once again, this cannot be surprising: including these farmers possibly moves the sample from behaving according to the neoclassical labor demand model: the market wage may be a poor indicator of the price of labor for these households.

The conclusions to draw from the MD estimation depend on whether one views the glass as half empty or half full. Strictly speaking, the joint hypothesis of the CES production function, Constant Returns to Scale, and Omitted Land Quality Bias is rejected for each of the above models of omitted variable bias. However, given the large sample size, and the small deviations, it is not too surprising that even the Labor Augmenting model is rejected. It is a very simple model, unlikely to be literally true. Especially given the results of the previous section on hypothesis testing, I would conclude that there is some weak evidence that the error process generating the biased OLS coefficients can be characterized by omitted labor augmenting land quality, at least for the sample of farmers who participate in the market for hired labor.

5. Conclusions and implications

The objective of this paper has been to investigate the statistical robustness and economic interpretation of the inverse productivity relationship. The motivation is that the inverse relationship is a widely accepted empirical regularity often cited as prima facie evidence of severely distorted rural labor markets. The main quest has been to allow for unobserved heterogeneity and to isolate the effects of a potential omitted variable, land quality, that has been suggested as an alternative source of the inverse relationship. I demonstrate that the inverse relationship and accompanying wage responses are completely at odds with the interpretation of farmers' output and labor demand decisions as being generated by profit-maximizing, price-taking behavior, unless we account for omitted variables.

Several properties of the inverse relationship and unobserved omitted variable have been established. The inverse relationship is robust to cluster fixed-effects, suggesting the omitted variable's effect on the land coefficient is largely a within-cluster phenomena. As well, the inverse relationship is constant over the range of farm sizes in the sample. When wages are added to the regression equations defining the inverse relationship, the restrictions implied by a neoclassical model of farm behavior are strongly rejected. Selecting instruments addressed at the land quality question causes the inverse relationship to vanish. As well, most of the wage elasticities are consistent with neoclassical economic theory. Tests of overidentification indicate we cannot reject the validity of the population density measures and other variables as identifying (excluded) instruments, and reinforce confidence in the empirical results. What's more, the reversal of the inverse relationship would not be predicted even if the instruments were invalid. The relative OLS biases across the labor demand, output, and profit equations are approximately consistent with an omitted variable that can be described as labor augmenting land quality. However, more precise estimation indicates that the tightly parameterized model of farm production and omitted variable bias does not explain the joint configuration of OLS and IV estimated coefficients.

This paper suggests an alternative hypothesis to imperfect labor markets (ILM) but *does not* directly address the validity of the ILM hypothesis. However, there are several implications of this research which suggest the need for further evaluation of the ILM explanation. In essence, the ILM model implies that the omitted variable in my regressions is a farm specific shadow wage. The shadow wage would be an increasing function of farm size as well as an assortment of other variables. Where families were most underemployed (small farms) more labor would be used. This is the usual story of the inverse relationship. In addition, with ILM we would expect the labor demand elasticity to be small, though a positive wage effect in the output and profit equations would be difficult to generate. The ILM hypothesis has a more difficult time predicting the other results in the paper. First, given that local labor market conditions are likely to be correlated with the shadow wage, I would expect cluster fixed-effect regressions to remove much more of the inverse relationship. Second, I would not expect the inverse relationship to be constant over farm sizes but would expect the relationship to diminish over the larger farms, as suggested by Feder (1985) and Carter and Wiebe (1990). Third, if we take the 3SLS estimates as legitimate, then the most damaging evidence is the elimination of the inverse relationship. If the shadow wage is increasing in farm size, then the true total elasticities of output and labor with respect to land should be less than one. In fact, there is no reason to believe that the OLS estimate of the land effect is biased: it merely captures the total effect of changing scale *plus* changing the shadow wage. The elimination of the inverse relationship with instrumental variables is a prediction not made by the imperfect labor markets model. Even if the instruments were not valid, we would expect the land coefficient to *decline* with 3SLS, given the likely negative correlation between population density and the shadow wage. In addition, it is unlikely the overidentification tests would pass if the imperfect labor markets model was true. The ILM model makes no predictions (either way) regarding the changed wage elasticities. Finally, regarding ILM, the results of my earlier paper are of direct interest. We would expect the shadow wage to be correlated with measures of demographic composition. It was shown in Benjamin (1992) that for a similar sample of rice farmers, demographic composition did not affect labor demand. The overidentification tests in this paper confirm this result, at least for some demographic groups. In addition, if the shadow wage model was true, we would expect to see differing average and/or marginal efficiencies of family and hired labor. Again, I found no evidence that this was the case. On balance, there is sufficient evidence to question whether the surplus labor ILM story is the source of the IR for this sample of farmers.

In all likelihood though, the labor market is more complicated than the simple neoclassical model outlined in this paper. Simple evidence of this is found when I extend my sample to include observations on farmers who do not have wage information. While imperfect wage imputation may cause this sample to perform less cleanly, it is likely that this extended sample includes 'constrained' farmers, farmers who did not hire labor and behave like the farmers in the ILM framework. Inclusion of these farmers does not change the basic pattern of the 2SLS estimates: the IR is eliminated in each equation, though less in the labor equation. The wage elasticities are also lower, and do not line up as well with the theory. Finally, the structural model of omitted land quality performs worst in this sample. So while the primary sample suggests that the inverse relationship can be explained by farm level heterogeneity, the second sample suggests that there are some farmers for whom the traditional ILM story may be applicable. It is possible that the truth is a hybrid of these models. Many researchers, particularly those addressing questions of 'labor absorption,' the ability of agriculture to provide employment, have discussed the possible endogeneity of land quality. [24] They suggest that in response to underemployment, or other economic forces, farmers invest into farm capital: land and irrigation improvement. The investments made in drainage and irrigation facilities are regarded as labor-augmenting land improvements. To the degree that smaller farms are more affected, this could generate the inverse relationship. Of course, many of these improvements were also undertaken at the community level, since there are significant externalities in this form of capital. Given that I control for irrigation, and cluster effects, I am not sure that this is necessarily the explanation.

Whatever the true process generating the variables I have examined here, there are a two certainties. First, we cannot directly take the inverse relationship as evidence that a simple reallocation of land will increase employment and output. Second, and unfortunately more important, when estimating cross section regressions on land, prices or wages, we have to consider very carefully what generates the variation. Even if not choice variables for the individual farmer, we may not necessarily view these variables as exogenous.

References

Abey, Arun, A. Booth and R.M. Sundrum, 1981, Labour absorption in Indonesian agriculture, Bulletin of Indonesian Economic Studies 17, no. 1, 36–65.

Berry, R. Albert and William R. Cline, 1979, Agrarian structure and productivity and developing countries (Johns Hopkins University Press, Baltimore, MD).

Benjamin, D., 1989, An econometric analysis of farm labor demand in rural Java, Ph.D Dissertation (Princeton University, Princeton, NJ).

Benjamin, D., 1991, Can unobserved land quality explain the inverse productivity relationship? Working paper no. 9112 (University of Toronto, Toronto).

Benjamin, D., 1992, Household composition, labor markets, and labor demand: Testing for separation in agricultural household models, Econometrica 60, no. 2, 287–322.

Bhalla, Surjit S., 1988, Does land quality matter, Journal of Development Economics 29, 45–62.

Bhalla, Surjit S. and Prannoy Roy, 1988, Mis-specification in farm productivity analysis: The role of land quality, Oxford Economic Papers 40, 55–73.

Booth, Anne, 1985, Accommodating a growing population in Javanese agriculture, Bulletin of Indonesian Economic Studies XXI, no. 2, 115–145.

Booth, Anne and R.M. Sundrum, 1984, Labor absorption in agriculture: Theoretical analysis and empirical investigations (Oxford University Press, Delhi).

Boserup, E., 1975, The impact of population growth on agricultural output, Quarterly Journal of Economics 89, no. 2, 256–270.

Carter, Michael R., 1984, Identification of the inverse relationship between farm size and productivity: An empirical analysis of peasant agricultural production, Oxford Economic Papers 36, 131–146.

[24] See Booth and Sundrum (1984) for a thorough review of the literature, especially as it relates to rice farming. Rao (1966) and Boserup (1975) are additional leading sources for this argument.

Carter, Michael and J. Kalfayan, 1988, A general equilibrium exploration of agrarian class structure and production relations, Agricultural economics staff paper no. 279 (University of Wisconsin, Madison, WI).

Carter, Michael and Kieth Wiebe, 1990, Access to capital and its impact on agrarian structure and productivity in Kenya, American Journal of Agricultural Economics, Dec., 1147–1150.

Chamberlain, Gary, 1984, Panel data, in: Zvi Griliches and Michael Intriligator, eds., The handbook of econometrics, Vol. 2 (North-Holland, Amsterdam).

Chayanov, A.V., 1926, in: D. Thorner et al., eds., The theory of peasant economy (Irwin, Homewood, IL).

Donner, Wolf, 1987, Land use and environment in Indonesia (C. Hurst and Company, London).

Eswaran, Mukesh and A. Kotwal, 1986, Access to capital and agrarian production organization, Economic Journal 96, 482–498.

Feder, Gershon, 1985, The relationship between farm size and farm productivity: The role of family labor, supervision, and credit constraints, Journal of Development Economics 18, 297–313.

Fuss, Melvyn and D. McFadden, eds., 1978, Production economics: A dual approach to theory and applications (North-Holland, Amsterdam).

Geertz, Clifford, 1963, Agricultural involution: The process of ecological change in Indonesia (The University of California Press, Berkeley, CA).

Hamermesh, Daniel S., 1993, Labor demand (Princeton University Press, Princeton, NJ).

Horstman, Kurt and Werner Rutz, 1980, The population distribution on Java 1971 (Institute of Developing Economies, Tokyo).

Khurso, M.A., 1973, Economics of land reform and farm size in India (MacMillan, Delhi).

Kuperus, G, 1938, The relation between density of population and utilization of soil in Java, C.R. Congres Int. Geogr. Amsterdam, t2, Leiden.

Lau, Lawrence J. and Pan A. Yotopolous, 1971, A test for relative economic efficiency and an application to Indian agriculture, American Economic Review 61, 94–109.

Lichtenberg, Eric, 1989, Land quality, irrigation development, and cropping patterns in the northern high plains, American Journal of Agricultural Economics 71, 187–194.

MacAndrews, Colin, 1986, Land policy in modern Indonesia (Oelyeschlager, Grimm and Hain).

Mehra, S., 1976, Some aspects of labour use in Indian agriculture, Indian Journal of Agricultural Economics 31.

Mohr, E.C.J., 1938, The relation between soil and population density in the Netherlands East Indies, C.R. Congres Int. Geogr. Amsterdam, t.2, Leiden.

Newey, Whitney, K., 1985, Generalized method of moments specification tests, Journal of Econometrics 29, 229–256.

Rao, C.H.H., 1966, Alternative explanations of the inverse relationship between farm size and output per acre in India, Indian Economic Review 1, no. 2, 1–23.

Sen, Amartya K., 1962, An aspect of Indian agriculture, Economics Weekly, Annual Number.

Sen, Amartya K., 1975, Employment, technology, and development (Oxford University Press, London).

White, Hal, 1982, Instrumental variables regression with independent observations, Econometrica 50, no. 2, 483–499.