

2004

Panel econometric evidence of Chinese agricultural household behavior in the later 1990s: production efficiency, size effects and human mobility

Zhuo Chen
Iowa State University

Follow this and additional works at: <https://lib.dr.iastate.edu/rtd>

 Part of the [Agricultural Economics Commons](#), and the [Economics Commons](#)

Recommended Citation

Chen, Zhuo, "Panel econometric evidence of Chinese agricultural household behavior in the later 1990s: production efficiency, size effects and human mobility" (2004). *Retrospective Theses and Dissertations*. 931.
<https://lib.dr.iastate.edu/rtd/931>

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.

**Panel econometric evidence of Chinese agricultural household behavior in the later
1990s: Production efficiency, size effects and human mobility**

by

Zhuo Chen

A dissertation submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Economics

Program of Study Committee:
Wallace E. Huffman, Major Professor
Joseph Herriges
Kenneth Koehler
Quinn Weninger
Jinhua Zhao

Iowa State University

Ames, Iowa

2004

Copyright © Zhuo Chen, 2004. All rights reserved.

UMI Number: 3145631

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

UMI[®]

UMI Microform 3145631

Copyright 2004 by ProQuest Information and Learning Company.

All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.

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

Graduate College
Iowa State University

This is to certify that the doctoral dissertation of

Zhuo Chen

has met the dissertation requirements of Iowa State University

Signature was redacted for privacy.

Major Professor

Signature was redacted for privacy.

For the Major Program

TABLE OF CONTENTS

1. INTRODUCTION	1
Overview	1
Institutional evolution in China's agricultural sector.....	1
Input usage and size economies	3
Objectives.....	5
Data description and econometric techniques	7
Dissertation organization.....	9
2. EFFICIENCY OF MODERN GRAIN PRODUCTION ON CHINESE FARMS	10
Introduction	10
Model specification	13
The data	20
The results	22
Conclusions	30
3. THE RELATIONSHIP BETWEEN FARM SIZE AND PRODUCTIVITY IN CHINA'S AGRICULTURE	40
Introduction	40
Econometric issues in the farm size and productivity relationship.....	43
Data and empirical results	53
Discussions and conclusions	59
4. JOB LOCATION CHOICE DECISION IN RURAL CHINA.....	66
Introduction	66
Literature review	68
The data and econometric model.....	71
Estimation and results	80
Conclusions	84
5. CONCLUSIONS	91
6. APPENDIX.....	96
A. Marginal effects of the inefficiency explanatory variables.....	96
B. Summarizing curvature conditions for flexible functional forms	97
C. Murphy-Topel type variance estimators	109
7. REFERENCES.....	112
8. ACKNOWLEDGMENTS.....	121

1. INTRODUCTION

Overview

Based on a longitudinal rural household survey in nine provinces of China, the dissertation examines the household production and employment decision in rural China. It consists of three essays that analyze the production efficiency, inverse relationship between farm size and productivity, and job location choice of rural Chinese households. The first essay estimates technical efficiency in China's grain production and suggests that there are rooms for further output and productivity growth in China's grain production, the second essay explains the empirical irregularity of inverse relationship between farm size and productivity, and the third essay examines the role of education, household size and social network in job location decision-making of these households.

Institutional evolution in China's agricultural sector

China's agriculture has been an interesting topic not only due to its sheer size but also because of its institutional evolution, which has generated a long list of literature, e.g., Lin (1992), Dong and Dow (1993), etc.

The institutional development of the People's Republic agriculture sector has gone through three stages. The first dates from 1949 to 1958, during which the state implemented land reforms to redistribute land from landlords to landless peasants, and voluntary

cooperatives were organized as well. The second stage is the communization from 1958-1978, which is featured by forced participation in the commune system. The final stage is the emergence and spread of the Household Responsibility System after 1978.

Chinn (1977, 1980) argued that the collectivization in early 1950s improved the productivity of Chinese agriculture by eliminating the fragmentation of land and scale diseconomies. Lin (1990) argued that the plunge of agricultural output after 1958 was mostly due to enforced participation in communes.

The Household Responsibility System (HRS), under which households acquire the land use right and in return contribute a fixed amount of their outputs to the state, emerged in the late 1970s. Most rural communities in rural China adopted the HRS by 1984. Lin (1987) credited twenty percent of the productivity growth to the elimination of “shirking” behavior in collective farming by HRS contracts. Lin (1992) found that de-collectivization improved total factor productivity and contributed one-half of the output growth during 1978-1984. The adjustment in state procurement prices also contributed positively to output growth mainly via the responses in input use. The effect of other market-related reforms on productivity and output growth was very small.

We quote a paragraph of Lin (1988) to end this section: “In short, the shift of institutions in Chinese agriculture was not carried out by any individual’s willingness but evolved spontaneously in response to underlying economic forces.”

Input usage and size economies

The supply and utilization of inputs are likely to be among the “underlying economic forces”. Institutional innovations may emerge as a response to certain ineffective uses of inputs. In the early 1940s, sharecropping featured Chinese agriculture. Due to civil war and severe inequality, cropping households usually had very low incentive to invest on agricultural instruments, irrigation system, and human capital. Land reform in early 1950s was implemented to change this situation via land redistribution. It succeeded to certain extent as Chinn (1980) claimed. However, when the state enforced participation in commune system and anticipated further output growth, the effort failed due to the lack of an effective mechanism to monitor labor supply, as well as price signals, to adjust input use and consumption. The Household Responsibility System reversed this process and gave households back the right to claim the output residual after fulfilling the state quota. This provided an incentive to supply more effective labor input and to invest in agricultural instruments and small pieces of machinery. Agricultural output increased significantly after the implementation of the HRS. Fertilizer and pesticide usage increased significantly shortly after. Efforts have also been put into basic agricultural research. However, new questions arise, i.e., are the input usages appropriate and does size economies exist that would justify land consolidation?

In rural China, land is scarce while labor is abundant. Intuitively, it suggests that the marginal product of land is high while that of labor is low. Sen (1960) claimed that in

developing countries the marginal product of labor use is likely to be very low. Wan & Cheng (2001) obtained negative labor elasticities for Chinese grain production. Xu (1999) suggested that the role of industrial inputs in Chinese agricultural production might have been underestimated. However, Widawsky et. al. (1998) obtained negative marginal product of pesticide and concluded that pesticides were overused in eastern China under intensive rice production systems. They suggested that host-plant resistance is an effective substitute for pesticides. In summary, current input usages are not optimized and therefore production inefficiency may exist. The extent and determinants of this inefficiency is explored in the first essay.

With the development of agricultural technology, optimal farm size also evolves. Chinn (1977) found that it was possible for prewar Chinese Agriculture to explore some extent of scale economy. After the adoption of the HRS, there did exist opinions that China need to take advantage of scale economies (Lin 1988). These opinions did not receive much attention since their advocators favored re-collectivization. Lin (1988) argued that the gain from scale economies might be outweighed by additional monitoring cost. However, while the 1950s Chinese agricultural technology and mechanism were not suitable for large-scale production, the small land plots nowadays may be a constraint to apply modern agricultural technologies that take advantage of scale economies.

Objectives

Given a growing population, declining marginal effects of variable inputs, and limited per capita resources, how should China improve its agricultural production?

Future output growth of China's agriculture relies more on technology advancement and institutional innovation rather than extensive use of inputs. Labor usage has a trivial marginal product in agricultural production. Part of the reason is that China's agricultural sector has been using excess labor in last several decades. Given limited arable land in China and the competition for land between agricultural sector and industrial/service sector, there is not much room for further increase of land usage in agricultural production. Fertilizer was found being overused in China and in some cases misused (with inappropriate ratios of N: P: K). Capital investment is hard to have significant increase when facing the competition of industrial sector and the current land ownership structure. Therefore, to improve Chinese agricultural productivity, some or all of the following changes may be made: (1) more efficient use of labor; (2) adoption of an appropriate land ownership structure; (3) efficient use of fertilizer; and (4) increased capital investment. These changes are related. More efficient use of labor/fertilizer and increased capital investment need the support of appropriate land ownership institution. On the other hand, the need of efficient use of labor motivates local communities to seek alternative land ownership structure, which resulted in the significant regional heterogeneity in China's land ownership institution as Krusekopf (2002) observed. Recent emergence of "land bank" proved this mutual dependence and

interaction between institution and economy is still evolving in China. Efficient use of labor is often accompanied with increased capital investment. Educated or well-informed labor may be able to use fertilizer more efficiently. Lastly, more efficient use of labor means more labor is available to industrial sector. Off-farm working and migration are two means of such “shift” from agricultural sector to industrial/service sectors in China.

The objective of this dissertation is to answer the following questions regarding China’s agriculture.

1. How is the grain output decided by input usages in rural China? In other words, where does the production frontier lie? Are the Chinese rural households’ grain outputs close to the frontier? What factors determine their efficiency? The first essay examines technical efficiency through the framework of stochastic production frontier with a behavioral inefficiency component. We apply Battese and Coelli (1995) methodology to a panel of 591 rural households.
2. In developing agriculture where there is a broad range of farm sizes, farm size and output per unit of land are often found to be inversely related. In China, where average farm size is small and the distribution of farm sizes is relatively compact, we find that farm size and productivity are weakly inversely related. Is this inherent to China’s agriculture so that land consolidation should not be considered? We apply instrumental variable estimation to examine whether the

inverse relationship was due to heterogeneous land quality partly introduced during the land allocation.

3. Out-migration of rural labor has accompanied the industrialization of most countries. Rural Chinese households have three employment choices with different locations, i.e., stay exclusively on farm, work partially off-farm, and have some household members that migrate out of home region seasonally. A question to be asked is that what factors affect their job location choice? What are the roles of experience, education, household size, and social network? The third essay studies these issues based on a balanced panel of 482 households in rural China.

Data description and econometric techniques

Data description

The data used in the dissertation are part of a large comprehensive survey conducted by Research Center for Rural Economy (RCRE) started in 1986 in 29 provinces of China with over 20,000 households. Panel attrition has been small. The survey was temporarily discontinued in 1992 and 1994 for financial reasons. The data set for first two studies contains 591 farm households living in 29 villages from 9 provinces in China from 1995 to 1999. We randomly selected the study's villages from the larger sample. In the third essay, we reduce the sample to a balanced panel of 482 households.

The overall survey was conducted by provincial offices under the Ministry of Agriculture. Each provincial research office first selected equal numbers of three types of counties: upper, middle and lower income; then it chose a representative village in each county. Forty to 120 households were randomly surveyed within each village. Village officers and accountants filled out a survey form on general village characteristics every year (Benjamin, Brandt, and Giles 2001). A stratified sampling technique was used during the selection of representative villages and households. A comparison of the household characteristics suggests the sample represents the national population quite well.¹

RCRE claimed that 80 percent of the households remained in the survey for the period of 1986-1997. By comparing the characteristics of those households, Chen (2001) found several new households (accounted for less than one percent of the whole sample) used the ID of an old household. He assigned new ids to these households.

Econometric techniques

The dissertation applies maximum likelihood estimation, instrumental variable estimation, and maximum simulated likelihood estimation in the three essays, respectively.

To estimate technical efficiency, the first essay uses a fixed effects stochastic frontier model with a behavior inefficiency explanatory term. We derive the close form of marginal

¹ For example, the national average rural household size is 4.2 people per household in 1990 while our sample yields average values ranging from 4.19 to 4.28 during the period of 1995-1999. In 1995, the national average grain output per unit of land is 4867 kg/Hectare while the average for this sample is 4890kg/Hectare.

effects of inefficiency terms, as well as their variance estimator. In Appendix B, we critique the usual practices of summarizing curvature conditions of flexible functional forms and propose two new methods to accomplish the goal. Our methods produce more intuitive summaries for the curvature conditions.

The second essay uses instrumental variable estimation to correct for land heterogeneity. We applied Hahn-Hausman (2002) test to examine the necessity and validity of the instrumental variable estimation. We derive the Murphy-Topel two-step estimation variance estimator in the linear case. Both village effects and household effects model are examined.

The third essay extends the dynamic discrete choice model of Wooldridge (2002a,b) to trichotomous setting and applies it to a balanced five-year panel data set. We use the maximum simulated likelihood estimation to reduce the computation burden in estimating the random effects multinomial logit model.

Dissertation organization

The dissertation is organized as follows. Chapters 2 through 4 present the three studies respectively. The last chapter concludes. Appendix A derives the estimator of the marginal effects of inefficiency explanatory variables and their variance estimator within Battese and Coelli (1995) framework. Appendix B proposes two methods to summarize curvature conditions of flexible function forms. Appendix C presents the derivation of Murphy-Topel type variance estimators in the linear regression setting.

2. EFFICIENCY OF MODERN GRAIN PRODUCTION ON CHINESE FARMS: A STOCHASTIC FRONTIER APPROACH

Introduction

As the world's largest food supplier and consumer, China's agriculture has been drawing extensive attention. Some, e.g., Brown (1995), had expressed concerns about China's ability to feed itself, especially with a growing population. Improving farm level efficiency (e.g., Abdulai and Huffman 2000; Mao and Koo 1997) and reducing fragmentation (Wan and Cheng 2001) are possible options to larger total domestic grain supplies. Evenson and Gollin (2003) shown that considerable yield potential resides in Green Revolution crop varieties and elite lines.

Hence, some real concerns exist about China's future supply of grain. From an economic perspective, we can split the issues into the location of production frontier for grain crops and proximity of current technology to the frontier or degree of inefficiency. For the 1970s and early 1980s, Lin (1987) attributed 20 percent of the productivity growth to institutional change that eliminated much of the "shirking" occurred under the collective farming. Mao and Koo (1997) and de Brauw, Huang, and Rozelle (2001) had similar conclusions. The role of increased usage of modern inputs, e.g., modern crop varieties, fertilizer, pesticide on grain production may have been underestimated (Xu 1999), but Widawsky et al. (1998) concluded that pesticides were overused in eastern China while

host-plant resistance has developed. There has been re-occurring interest in the broad topic of possible excess labor existing in China's agriculture, e.g., see Wan and Cheng (2001). China does not have private land ownership so a local village council allocates local land based on nutrition and rental needs. Hence, this village council is an institution with possible productivity effects. For example, Cheng (1998) concluded that a household having a member in the local "village council" had positive efficiency effect through better access to collectively owned farm equipment and state subsidized farm inputs. Farm size remains small in China, even under the Household Responsibility System but no consensus exists on the issue of whether the small farm sizes are a drag on productivity or efficiency, e.g., Wang, Cramer, and Wailes (1996); Wan and Cheng (2001).

For agriculture in general, (Huffman 1977) and for China, Wang, Cramer, and Wailes (1996), Yang (1997a), Cheng (1998) have shown that farmers' schooling has positive effects on technical and allocative efficiency. Yang, however, concluded that it was not really the farmer's schooling that matters for farm-level efficiency but rather the schooling of the individual in the household who had completed the most years.

Most of these studies have had only one cross-section or if they had multiple years of data, they have ignored the issue of unobserved individual household heterogeneity, which Greene (2002) has shown can affect on the estimated parameters of a behavioral model.

Farm production decisions of modern Chinese farmers fit the agricultural household model (Huffman 2001, Strauss 1986). If production decisions are separable from

consumption and labor supply decisions, the demand for inputs and supply of farm outputs can be considered separately. If input prices are readily available, a profit function could be used to examine efficiency. However, land and family labor continue to be the dominating inputs in China's agriculture and prices for them are not readily available, making a cost function examination of efficiency infeasible. An alternative route, and which we choose, is to specify and estimate a stochastic frontier function with a behavioral explanatory term for technical efficiency. The stochastic frontier has a major advantage compared to deterministic frontier models of being relatively insensitive to outliers or certain types of measurement error. The standard approach has been two-stage estimation: first to estimate the parameters of the frontier function and second to use the inefficiency scores from this function to estimate a behavior inefficiency function. Battese & Coelli (1995), Kumbhakar and Lovell (2000, p.263), and Wang and Schmidt (2002) however, suggested that this methodology leads to inefficient estimation and other problems. We apply Battese & Coelli (1995) methodology where all the parameters—frontier and technical inefficiency behavior function—are estimated jointly. We derive a new mathematical expression for marginal effects of determinants of inefficiency and their asymptotic variance estimator. Furthermore, we hypothesize that farm level inefficiency is negatively related to schooling and positively to land fragmentation.

The dataset we used is a panel of Chinese grain farms, 1995-1999. Important findings include: the translog stochastic frontier function is a good representation of the technology of

Chinese grain farms, individual household unobserved heterogeneity is shown to be statistically significant and ignoring it affects the coefficient estimates significantly. Chinese grain farms are shown to be relatively efficient— nearly 50 percent of the farms are at least 90 percent efficient. Also, farmers' schooling increases grain production efficiency, and land fragmentation reduces efficiency.

The rest of this chapter is organized as follows: Section 2 sets up the model. We describe the RCRE panel data set in Section 3 and Section 4 presents the results. The conclusions are summarized in Section 5.

Model specification

Technical efficiency is defined as minimizing input for a given output level, or maximizing output with fixed input usage. Hence, two different approaches have been followed to measure technical efficiency empirically, output-oriented and input-oriented. This study uses the output-oriented measure since output maximization is more likely in China's agriculture than input usage minimization due to the limited farmland supply. We choose to use stochastic frontier approach in this study.²

² Both the parametric approach (SPA) and nonparametric approach (DEA) have its own merits. Here we pursue the parametric approach since SPA is more robust to the measurement error and random disturbance, which is very likely in this study with large geographic and climatic variations.

Stochastic production frontier

The stochastic production frontier dates back to Meeusen and van den Broeck (1977), Aigner, Lovell, and Schmidt (1977), and Battese and Corra (1977). Their model can be summarized as: $y=f(x;\beta)\cdot\exp(v-u)$ where y is scalar output, x is a vector of inputs, β is a vector of technology parameters, v is the symmetric random disturbance, and u is the technical efficiency to be estimated. Jondow et al. (1982) extended this model by incorporating producer-specific efficiency effects. Greene (1980a, b), Stevenson (1980), and Lee (1983) proposed various specifications of efficiency distribution.

Cross section and panel data have been used in fitting stochastic frontier production functions. While most early attempts assumed time invariant firm-level technical efficiency, Cornwell, Schmidt, and Sickles (1990), Kumbhakar (1990), and Battese and Coelli (1993) relaxed this assumption. They estimated a production frontier model assuming the firm-specific effects follow an exponential function of time. However, in this study, we adopt the model of Battese and Coelli (1995), which is appropriate for a panel of short duration. The model is described in detail later.

Explaining technical efficiency

Battese and Coelli (1995) and Wang and Schmidt (2002) proposed one-step approaches based on the “location varying property” and “scaling property” of the inefficiency terms, respectively. Battese and Coelli (1993) model is similar to the Kumbhakar, Ghosh, and

McGuckin (1991) specification. They later extended Battese and Coelli (1993) model by adding exogenous variables to explain technical inefficiency. The model can be expressed as: $Y_{it} = x_{it}\beta + (V_{it} - U_{it})$, $i=1, \dots, N$, $t=1, \dots, T$, where: Y_{it} is the output (or its natural logarithm) of the i -th farm at the time t ; x_{it} is a $k \times 1$ vector of the input quantities (or their natural logarithms) of the i -th firm at the time t ; and β is the coefficient vector of x_{it} . V_{it} are random disturbance terms that are assumed to be *iid* $N(0, \sigma_v^2)$. They are incorporated in the model to reflect the random disturbance that is independent of U_{it} . The U_{it} are non-negative random disturbances assumed to account for technical inefficiency in production. They are assumed to be independently distributed and truncated at zero $N(m_{it}, \sigma_u^2)$ or as the more widely used notation $N^+(m_{it}, \sigma_u^2)$, where $m_{it} = z_{it}\delta$. z_{it} is a $p \times 1$ vector of variables which may influence farm-level efficiency; and δ is an $1 \times p$ parameter vector to be estimated. The relationship between U_{it} and the output-oriented technical efficiency TE_o is $TE_o^{it} = \exp(-U_{it})$. Battese and Coelli (1993) provided the log-likelihood function after reparameterizing σ_v^2 and σ_u^2 with $\sigma^2 = \sigma_v^2 + \sigma_u^2$ and $\gamma = \sigma_u^2 / (\sigma_v^2 + \sigma_u^2)$.

Elasticities and marginal effects

Input elasticities and their calculation have been discussed extensively in the literature (e.g., Kumbhakar and Lovell, 2000). Given a specific functional form, it is easy to obtain the corresponding marginal effects after input elasticities are calculated. The common approach used to summarize the curvature conditions is to evaluate them at a central point (mean,

geometric mean or median) of the inputs or to evaluate them at individual observations then calculate the mean or median. In Appendix B we criticize the usual approaches of summarizing the curvature conditions and propose two new methods. The new methods produce curvature condition summaries that provide policy makers a more accurate picture of how output is expected to respond to input changes.

Studies that adopted the methodologies using one-stage maximum likelihood estimation to estimate and explain technical efficiency, i.e., Battese & Coelli (1995) model, usually report the coefficient of inefficiency terms. These coefficients are directly relevant to the inefficiency but not necessarily relevant to the marginal effects of the inefficiency terms, which is more intuitive to policy makers.

Given the setting of Battese & Coelli (1995), assuming Y is the output and z is a vector of exogenous variables in the inefficiency expression, then we obtain the following results:

Theorem 1: The marginal effects of z is: $\frac{\partial E(Y)}{\partial z} = -\delta(1 + \Delta(\alpha))$, where $\alpha = \frac{z\delta}{\sigma_u}$, $\lambda(\alpha) = \frac{1-\phi(\alpha)}{1-\Phi(\alpha)}$, $\Delta(\alpha) = \lambda(\alpha)(\lambda(\alpha) - \alpha)$, and $\phi(\cdot)$ is the probability density function of standard normal distribution while $\Phi(\cdot)$ is the cumulative density function.

Theorem 2: The asymptotic variance of the vector of marginal effects $\frac{\partial E(Y)}{\partial z}$ can be estimated by: $M \text{var}(\delta, \sigma^2, \gamma) M'$

$$\text{Where } M = \begin{pmatrix} \frac{\partial m_1}{\partial \delta_1} & \cdots & \frac{\partial m_1}{\partial \delta_j} & \frac{\partial m_1}{\partial \sigma^2} & \frac{\partial m_1}{\partial \gamma} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial m_j}{\partial \delta_1} & \cdots & \frac{\partial m_j}{\partial \delta_j} & \frac{\partial m_j}{\partial \sigma^2} & \frac{\partial m_j}{\partial \gamma} \end{pmatrix},$$

$$\frac{\partial m_j}{\partial \delta_k} = \frac{\partial \left(\frac{\partial E(Y)}{\partial z_j} \right)}{\partial \delta_k} = \begin{cases} -1 - \Delta(\alpha) - \delta_j \lambda ((\lambda - \alpha)^2 + \frac{\delta_j z_j}{\sigma_u} \lambda (\lambda - \alpha) - 1) & \text{when } j = k \\ -\frac{\delta_j z_k}{\sigma_u} \lambda ((\lambda - \alpha)^2 + \lambda (\lambda - \alpha) - 1) & \text{when } j \neq k \end{cases}$$

$$\frac{\partial m_j}{\partial \sigma_u^2} = \frac{\partial \left(\frac{\partial E(Y)}{\partial z_j} \right)}{\partial \sigma_u^2} = \frac{1}{2\sigma_u^2} \delta_j \lambda \alpha ((\lambda - \alpha)^2 + \lambda (\lambda - \alpha) - 1),$$

$$\frac{\partial m_j}{\partial \sigma^2} = \frac{\partial m_j}{\partial \sigma_u^2} \gamma, \quad \frac{\partial m_j}{\partial \gamma} = \frac{\partial m_j}{\partial \sigma_u^2} \sigma^2, \quad \text{and } \text{var}(\delta, \sigma^2, \gamma) \text{ can be estimated by the asymptotic covariance}$$

estimates of $(\delta, \sigma^2, \gamma)$ obtained from maximum likelihood estimation.

When Y is the logarithm of output, the marginal effects can be estimated by multiplying the elasticity results obtained through Theorem 1 by the input quantities. Their variance estimates can be computed using Delta's method. Proofs of Theorem 1 and Theorem 2 are provided in the appendix.

The empirical model

We follow earlier studies, e.g., Battese and Coelli (1995), Guilkey, Lovell, and Sickles (1983) that have chosen a translog functional form of the production frontier, which is flexible and a second order approximation to any true functional form. It is also more flexible in permitting substitution effects among inputs than the generalized Leontief function, which is similar to fixed proportion technology.³ Furthermore, we conduct some tests to confirm

³ Performance of flexible function forms depends on the initial specification of second order curvature conditions. The Translog is preferable when the cross Allen-Uzawa elasticities of substitution (AUES) are close to unity, and the generalized Leontief is a good choice for cross AUESs close to zero. In rural China, relatively wealthy farmers tend to use commercial fertilizer to save labor usage thus the substitution effect might be strong. Same argument applies for land and labor, land and fertilizer, as well as capital and labor/fertilizer. Therefore we deem Translog as a better choice in our study.

that the translog is better than alternative specifications.

A remaining issue is how to handle the unobserved heterogeneity. If unobserved panel heterogeneity exists, it most likely biases all the estimated coefficients of the frontier function and inefficiency equation. In our setting, unobserved household effects are attributed to household managerial skills and abilities to collect information (Huffman 2001). Consider the following fixed effects model where a dummy variable is included for each household (d_{hh}):⁴

$$\text{Model 1: } \ln(Y_{it}) = \beta_0 + \sum_{l=1}^4 \beta_{d_l} d_{(year)l} + \sum_{i=1}^{590} \beta_{hh_i} d_{hh_i} + \sum_{j=1}^4 \beta_j \ln(x_{ijt}) + \sum_{j \leq k} \sum_{k=1}^4 \beta_{jk} \ln(x_{ijt}) \ln(x_{ikt}) + V_{it} - U_{it}.$$

where the subscript i indicates the household; t indexes time, and j, k index inputs used. For the i -th household at time t , x_{i1t} is sown area used of grain production; x_{i2t} is the labor employed in grain production; x_{i3t} is chemical fertilizer applied to grain crops and x_{i4t} is capital input used in grain production. The units of these inputs are Mu⁵, man-day, kilogram, and RMB Yuan⁶, respectively. The d_{year} s are the year dummy variables for 1996-1999, where 1995 is the reference year and $\beta_j, \beta_{jk}, V_{it}$, and U_{it} are defined as before.⁷

⁴ Note that there are discussions that whether fixed effects should be considered as frontier shift or inefficiency. This is a philosophical question and probably deserves further exploration but here we follow Greene (2002) approach in treating it as frontier shift.

⁵ 1 Mu=1/15 Hectare.

⁶ 1 RMB Yuan=\$0.12 approximately.

⁷ Note that the x vector in the likelihood function is all the explanatory variables in the translog functional form, including the cross-production term and the dummy variables.

To explain the inefficiency term U_{it} , we define $m_{it} = \delta_0 + \sum_{k=1}^{16} \delta_k z_k$, where z_1 - z_3 are the household schooling variables. Specifically, z_1 is one, if the highest level of schooling among household members is 5-8 years (and zero otherwise); z_2 is one if the schooling level is 9-11 years (and zero otherwise); and z_3 is one if the schooling level is greater or equal to 12 years (and zero otherwise). z_4 takes a value of one if a household member is a village officer (and zero otherwise). z_5 is a household's number of the plots under grain production (an indicator of land fragmentation)⁸. Coelli (1996) claimed that in Battese and Coelli (1995) framework explanatory variables could appear in both the production function and the inefficiency explanatory term. Therefore we avoid the potential collinearity problem by putting the number of plots in the inefficiency term. z_6 is an index of farm specialization (the ratio of land under grain production to overall land farmed by the household). z_7 through z_{10} are dummy variables representing the age of the household head. Specifically, z_7 is one when the head's age is 31- 40 year (and zero otherwise); z_8 is one when the head's age is greater than 40 years but less than 50 years (and zero otherwise); z_9 is one when the head age is between 51 and 60, zero otherwise; and z_{10} is one when the head is older than 61 years and zero otherwise. The reference age group that these households having a head who is less than 30 years old. z_{11} through z_{14} are time dummy variables and their coefficients reflect year-effects, e.g., pests, weather. z_{15} is one if the household owns or partially owns any kind of machinery for grain

⁸ Previous studies (e.g. Wan and Cheng 2001; Fleisher and Liu 1992) included the number of plots as a variable in the frontier, which may introduce linear collinearity.

production (and zero otherwise). z_{16} is one if the household is located in southern provinces (zero otherwise)⁹.

The software used in this study is FRONTIER 4.1 developed by Coelli (1996). A three-step estimation method is used in obtaining the final maximum likelihood estimates. The likelihood maximization procedure uses Davidson-Fletcher-Powell Quasi-Newton algorithm. The panel data set needs not to be balanced in this model.

The data

Dataset Description

The data for this study are a unique panel data set which is part of a large comprehensive survey conducted by Research Center for Rural Economy (RCRE), started in 1986 in 29 provinces of China and contains more than 20,000 households. Panel attrition has been small. The survey was temporarily discontinued in 1992 and 1994 for financial reasons. The overall survey was conducted by provincial offices under the Ministry of Agriculture. Each provincial research office first selected equal numbers of three types of counties: upper, middle and lower income; then it chose a representative village in each county. Forty to 120 households were randomly surveyed within each village. Village officers and accountants filled out a survey form on general village characteristics every year (Benjamin, Brandt, and

⁹ Southern provinces: Anhui, Jiangsu, Yunan, Sichuan. Northern provinces: Hebei, Shandong, Liaoning, Helongjiang, Shanxi.

Giles 2001). We randomly selected the villages used in this study from the larger sample. The data set for our study contains 591 farm households living in 29 villages from 9 provinces in China over 1995 to 1999.

RCRE claimed that 80 percent of the households remained in the survey for the period of 1986-1997. By comparing the characteristics of households in this sample, Chen (2001) found several new households (accounted for less than one percent of the whole sample) used ID of an old household. He assigned new IDs to these households.

Descriptive Statistics

Sample mean values of variables for the empirical model are reported in Table 2-1. The data set is unbalanced because some households did not engage in grain production in every year. In addition, we had to delete a few observations because of data recording mistakes or missing information. They are, however, a negligible small fraction (less than 3%) of the whole data set. Using the general retail price index, Chen (2001) converted all monetary variables such as prices, income, or expenditures into real term with 1986 as the base year.

Over the five-year period (1995-1999), household size is shrinking slowly, which is partially attributable to family planning policy and out-migration in rural area. The schooling of the highest education attained by household members is increasing over time, probably due to the efforts on illiteracy reduction. The number of plots is decreasing, which is likely the result of land consolidation efforts and land rental market activities.

Input usage and output produced for the RCRE households are summarized in lower parts of Table 2-1. Sown area of grain crops does not change much over the five-year periods. The labor usage per household fluctuates over this period, but the trend is negative. Labor input per hectare varies more during the time period, which may be due to the fact that growing population has forced Chinese farmers to put more variable inputs on the limited land. Both fertilizer usages per hectare and per household have an increasing trend over time with the exceptions of year 1997 and 1999. Capital input usage has no obvious trend. Grain output per household does not change significantly over the study period, but the average grain yield per hectare increased over 1995-98 and is lower in 1999 because of unfavorable weather conditions. Also, we have examined the ratio of the land under grain production to the overall land and find that households usually use more than eighty percent of their land for grain production. This is taken as evidence that focusing on one output—grain production is a reasonable approach.

The results

Production frontier estimates

Maximum likelihood estimates of Model 1 are presented in Table 2-2. We omit a lengthy report of household effects estimates, but report that a joint null hypothesis that all household effects are zero can be rejected using a likelihood ratio test (see Table 2-3). Also, the null hypothesis that the translog production frontier reduces to a Cobb-Douglas

production function is also rejected (see Table 2-3). Hence, we conclude that there is strong support for the particular econometric representation of grain production on Chinese farms in the late 1990s. We also tested the null hypothesis that land, labor, fertilizer, and capital individually make no statistically significant contributions to output. These hypotheses are rejected (see Table 2-3). The marginal products of the variable inputs are of importance because they show the marginal impact of an increment of an input on grain output at the frontier. For land, labor, fertilizer and capital, the marginal products evaluated at the household sample geometric mean, are 210.99 Kg/Mu, 0.02 Kg/Man-day, 0.34 Kg/Kg and 0.36 Kg/Yuan, respectively.¹⁰ Hence, at the sample geometric mean, there is no evidence of negative marginal products, or excessive input use in the sense of a negative elasticity reported by Wan and Cheng (2001) for labor and fertilizer, by Huang (2001) and Widawsky et al. (1998) for fertilizer. Huang (2001) emphasized that the nutrient composition of chemical fertilizers used in China's agriculture is about 1:0.31:0.01 for N, P and K in fertilizer uses, which is far different from the average of developed countries, 1:0.5:0.5, and world average of 1:0.45:0.36. One explanation is that chemical fertilizer being more frequently used by young farmers while older farmers use organic manures more frequently.

Our estimate of the output elasticity of land, labor, fertilizer, and capital, evaluated at the geometric mean of the inputs and output is 0.677, 0.001, 0.052 and 0.142, respectively.¹¹

¹⁰ We also evaluated them at the sample median and obtained similar results.

¹¹ The 95% confidence intervals for the estimates are (0.598, 0.756), (-0.020, 0.023), (0.035, 0.069), and (-0.026, 0.297).

Nonetheless, our estimates of the output elasticity of land confirm those reported by Fleisher and Liu (1992) using data for an earlier period. Lin (1992) used data during the transition period (1978-1984) and reported estimates of output-input elasticities as 0.49, 0.21, 0.15, and 0.06, respectively (land, labor, fertilizer and capital), which are somewhat different from our estimates. Our production elasticities show the continued important role played by land in Chinese grain production. Using an error component specification, Lanjouw (1999) obtained similar conclusion for agricultural production in an Indian village of the year 1983/1984. Her estimated land elasticity is 0.573 (standard error 0.037) and labor elasticity is 0.125 (standard error 0.087). Also, Table 2-6 presents a summary of production elasticities of variable inputs which can be compared to our results.

The scale elasticity, evaluated at the sample geometric mean of the inputs, is 0.866, and the 95% confidence interval is (0.717, 1.015), which does include unity. Hence, during the period 1995-99, we cannot reject the hypothesis of constant return to scale in grain production, given land fragmentation.

We apply the Chen and Huffman (2004) methodology, which is described in Appendix B, to summarize the curvature conditions of the production frontier. We use a K-means clustering algorithm to divide the sample into two groups and then summarize the curvature conditions separately. Table 2-4 presents these results, which show greater marginal effects and output elasticity of capital but smaller marginal effects and output elasticities of labor

and fertilizer than the traditional method. The scale elasticity is also greater using our method than that obtained through the traditional summarizing method.

Our estimates of σ^2 and γ are 0.522 and 0.9998 respectively. One might reasonably ask whether γ is statistically different from unity. However, we refrain from concluding that γ is equal to unity since the distribution of the estimates of γ is not exactly a symmetric asymptotic normal.¹² Furthermore, the deterministic frontier cannot handle outliers and measurement errors well, which is likely in such a sample that spans five years and includes nearly six hundred households. Meanwhile, it is not surprising that γ is close to unity since the fixed effects may account a large part of the symmetric variation. *This result implies that the one-sided random inefficiency component strongly dominates the measurement error and other random disturbances.*

Technical inefficiency

Technical efficiency is estimated as $\widehat{TE}_i = \exp(-\widehat{U}_i)$. The estimates for the coefficients in efficiency explanatory term are reported in the lower part of Table 2-2 and the technical efficiency distributions are summarized in Table 2-5, Figure 2-1, Figure 2-2, and Figure 2-3. At the sample mean, technical efficiency is 0.827. The average technical efficiency by year is 0.808, 0.840, 0.833, 0.883, and 0.768 for 1995 to 1999, respectively.

¹² It seems a transformation of γ might be able to produce a more reasonable asymptotic normal distribution but we would need to apply Delta's method in the computation.

Figure 2-1 presents the histogram of technical efficiency for individual years and the whole sample over all years. Figure 2-2 and Figure 2-3 present the histograms of technical efficiency for the nine provinces and twenty-nine villages. We calculated the marginal effects and the corresponding variance estimates according to Theorem 1 and 2. The marginal effects and the t-ratio tests are presented in Table 2-7.

The results show that an increase in schooling (of the person in the household with the highest level) improves technical efficiency. This result agrees with Yang (1997a,b), which concluded that collective decision-making is employed in Chinese farm households during 1990s. It is also consistent with a large amount of information summarized in Huffman (2001).

The estimated coefficient of the village officer dummy variable is positive and statistically significant, which implies loss of efficiency. This finding is different from that reported by Cheng (1998). However, since Cheng's study and ours do not overlap in time, there is no obvious contradiction. The earlier positive role of the village officer may have turned negative in modern Chinese agriculture. One reason why this might have occurred is that input markets have gained openness and hence leverage through the village council is no longer necessary for gaining access to variable inputs and capital goods in the later 1990s. The statistical significant positive effect on inefficiency may due to the time constraint on these officers as well as the reaction to wage earned through the position—officers may allocate more time in office to keep his/her position rather than working on farm.

The positive coefficient for number of plots implies reduced technical efficiency as the number of plots per farm increases—greater fragmentation. Hence, the evidence is that technical inefficiency is related to land fragmentation. Although Wan and Cheng (2001) studied the effect of land fragmentation on the production frontier while we study its effect on production efficiency, our results do not seem to be contradictory.

As a farm specializes increasingly in crops, technical efficiency increases significantly. However less diversified farmers are exposed to greater production and market risk which our study ignores.

The estimated coefficients and the marginal effects for the dummy variables for the head's age suggest rather complicated effects. Household that have a head who is in his 40s are most efficient among all households. However, those households having a head more than 60 years old are also quite efficient and more efficient than the youngest farmer group. The implications of this result need more investigation. Nonetheless, our result for Chinese farm households, 1995-99, shows that little evidence exists to support the common view that households with older heads are less efficient.

The estimated coefficient of the dummy variable for mechanized farm has a positive sign and it is statistically significant. It is plausible that mechanized farms are more efficient because they usually apply certain new technologies more effectively. Finally farms located in the northern provinces in our sample are less efficient and they may face more adverse weather conditions such as floods than farms in other areas of China.

Policy implications

Our results suggest that improving the access to rural education, removing market inefficiency and land consolidation are important to China. The Chinese government imperatively lowered the prices of major agricultural products to guarantee industrial development in 1950s. This brought distortions into both output market and labor market. The lowered output prices created an incentive for out-migration from farming but it was not openly permitted till 1980s. When the state loosened regulations on migration, many better-educated individuals left farming sector. Some were successful in finding jobs in urban areas and others worked off-farm in their home area. The marginal products of those who remained in the agricultural sector may be significantly lower than those who left due to the selectivity of migration. However, a free market of agricultural products provides farmers for a while with higher output prices and thus higher profit levels. These incentives might attract some of educated labor back to agriculture. Tian and Wan (2000) noted that the rural labor force remains largely undereducated, which suggests the need to improve the education levels of farm workers. China's access to WTO will free the agricultural market in a profound way. Although a short-run loss of welfare seems unavoidable, the improvement in the labor force and better access to technology and credit markets will definitely bring China's agriculture long-term benefits. More work should be done on integrating the production and marketing of agricultural products, which would strengthen farmers' bargaining power.

We conclude that an institutional innovation that facilitates the land consolidation might be beneficial to China's agriculture. The institutional innovation would most likely to mobilize labor to leave the farming sector. Some researchers argued that China should privatize land ownership. However, Li, Rozelle, and Brandt (1998) pointed out that, given no institutions like land courts, land registration system, and credit markets, land privatization for China at this time might have a high cost to the society. Land rental market is possibly a better solution. It provides a type of social security for households renting out lands and consolidates fragmented lands (Lohmar, Zhang, and Somwaru, 2002). However, voluntary land rental activities seem to be less successful as expected due to problems associated with negotiation of the terms of the rental contract. A new institution, land bank, emerged in rural Zhejiang province to resolve these issues and deserves further study. While the Household Responsibility System emerged in Anhui, spread all over China and brought tremendous changes to China's agriculture, the "land bank" may have such potentials to bring China's agriculture new strength by reducing land fragmentation.

We deem the following strategies as helpful to China's rural policy-makers. First, we need to improve the access of young people to elementary and junior high school education in rural area. Although college educated individuals rarely stay in rural China, exposure to agricultural science and engineering in high school might improve the farmers' productivity in the light of the Morrill Act. Second, a "land bank" might be more generally adopted for land consolidation where there are enough households to support a rental market. Even in the

case where there are not enough such households, a “land bank” may help to reduce the land fragmentation. Lastly, mechanization should be promoted where the landscape and labor/land market are suitable. We expect these strategies to improve the schoolings of rural households, eliminate certain extent of land fragmentation, and facilitate the specialization, thus to improve the production efficiency and ultimately to increase grain output.

Conclusions

In this study, a stochastic translog frontier production function with an inefficiency term has been fitted to farm-level grain production data for Chinese farms, 1995-99. The marginal products evaluated at the sample mean for land, labor, fertilizer and capital were showed to be plausible. However, there is evidence that the output elasticity of labor and fertilizers have declined over the past two decades as reforms have occurred. We cannot reject the hypothesis of constant return to scales in Chinese agriculture, 1995-99, given the existing land fragmentation. Our results reveal that Chinese farms in the late 1990s were on average relatively efficient— nearly 50 percent of the farms are at least 90 percent efficient. However, human capital in the form of highest schooling level of a household member is shown to be an important determinant of Chinese agricultural technical efficiency. This result is consistent with Yang’s findings for Chinese farms. However, the marginal product of elementary education is larger than for higher schooling completion levels. Eliminating land fragmentation, promoting grain-crop specialization, and using machinery will in general

bring significant efficiency gains for Chinese farms, which suggests a change in the land tenure system to promote larger and more specialized farms.

More research remains to be done in order to gain a better understanding of factors that affect the efficiency of Chinese agriculture, including the effects of land quality and scale elasticity. Regional disparity is an interesting topic to explore further with approaches such as spatial economics. Also, a profit function approach would enable one to examine allocative efficiency, but this requires additional information on prices of outputs and inputs for the RCRE data set. Field researches on “land bank” in Zhejiang province may shed new lights on potential institutional changes in rural China.

Table 2-1: Descriptive statistics for Chinese farm households (1995-1999: 591 households)

Variables	Unit	1995	1996	1997	1998	1999
<i>Household Characteristics</i>						
Household Size	Number	4.28 (1.34)	4.27 (1.38)	4.26 (1.36)	4.21 (1.39)	4.19 (1.36)
Labor Force	Number	2.57 (1.08)	2.59 (1.09)	2.58 (1.07)	2.53 (1.06)	2.55 (1.03)
Highest Education	Year	7.16 (2.42)	7.12 (2.42)	7.26 (2.35)	7.36 (2.28)	7.38 (2.23)
Number of Plots	Number	6.04 (4.58)	6.08 (4.68)	5.99 (4.80)	5.65 (4.52)	5.59 (4.62)
Gross Income Per Year	1000 Yuan (Real Term)	11.72 (7.62)	12.37 (8.87)	12.75 (10.15)	12.50 (8.68)	11.76 (10.33)
Agricultural labor force /agricultural land	Number/ Hectare	5.96 (5.03)	5.99 (4.68)	6.33 (6.29)	6.06 (5.16)	6.03 (4.70)
Land under Grain /Agricultural Land	Percent	0.86 (0.10)	0.87 (0.10)	0.87 (0.10)	0.87 (0.11)	0.88 (0.11)
<i>Inputs</i>						
Land Under Grain Production	Hectare	0.66 (0.59)	0.67 (0.63)	0.64 (0.59)	0.67 (0.62)	0.67 (0.60)
Labor input Per Household	Man-Day	210.5 (143.0)	222.3 (179.4)	208.7 (178.2)	202.1 (152.2)	200.8 (151.4)
Fertilizer Usage Per Household	Kilogram	476.7 (454.5)	522.3 (693.5)	488.2 (547.2)	534.1 (584.1)	498.4 (473.7)
Capital Depreciation Per Household	1000 Yuan (Real Term)	1.22 (1.11)	1.41 (3.67)	1.35 (3.63)	1.37 (3.31)	1.23 (2.83)
<i>Per Hectare Inputs</i>						
Labor input / Hectare	Man-Day/ Hectare	427.7 (400.2)	461.5 (465.2)	440.6 (404.8)	430.3 (366.8)	422.6 (345.5)
Fertilizer / Hectare	Kg/ Hectare	887.2 (663.2)	903.2 (619.8)	952.7 (770.7)	1029.2 (830.7)	982.1 (682.9)
Capital Depreciation /Hectare	Yuan/ Hectare	1.94 (0.89)	2.18 (1.31)	2.09 (1.15)	2.20 (1.52)	1.92 (1.11)
<i>Output</i>						
Grain Output Per Household	1000 Kilogram	3.04 (2.43)	3.12 (3.88)	3.19 (3.98)	3.34 (4.05)	3.12 (3.58)
Yield output/ hectare	1000 Kg/ Hectare	4.89 (1.58)	4.91 (1.55)	5.11 (3.20)	5.22 (1.69)	4.98 (1.92)
Observations		572	538	539	539	520

Table 2-2: Maximum-likelihood estimates: Panel of Chinese farm households (1995-1999)

Variables	Parameter	Estimates	<i>t</i> -value
Constant	β_0	7.198	(5.70)
Ln(land)	β_1	0.545	(2.77)
Ln(labor)	β_2	-0.057	(0.36)
Ln(fertilizer)	β_3	0.054	(0.12)
Ln(capital)	β_4	-0.339	(2.29)
(ln(land)) ²	β_{11}	-0.033	(38.06)
(ln(labor)) ²	β_{22}	0.004	(0.17)
(ln(fertilizer)) ²	β_{33}	0.002	(0.32)
(ln(capital)) ²	β_{44}	0.023	(1.53)
Ln(land)*ln(labor)	β_{12}	-0.023	(0.54)
Ln(land)*ln(fertilizer)	β_{13}	-0.013	(0.33)
Ln(land)*ln(capital)	β_{14}	0.068	(2.36)
Ln(labor)*ln(fertilizer)	β_{23}	0.002	(0.04)
Ln(labor)*ln(capital)	β_{24}	0.008	(0.21)
Ln(fertilizer)*ln(capital)	β_{34}	-0.002	(0.06)
Year 1996 Dummy	D ₉₆	-0.035	(0.64)
Year 1997 Dummy	D ₉₇	-0.022	(0.31)
Year 1998 Dummy	D ₉₈	-0.056	(0.42)
Year 1999 Dummy	D ₉₉	0.027	(2.74)
<i>Inefficiency Terms</i>			
Constant	δ_0	-0.826	(0.84)
Highest education (6-8 years)	δ_1	-1.420	(5.54)
Highest education (9-11 years)	δ_2	-1.090	(1.77)
Highest education (>12 years)	δ_3	-1.136	(1.41)
Village officer Dummy	δ_4	0.420	(2.27)
Number of plot	δ_5	0.086	(2.35)
Land Ratio	δ_6	-1.381	(2.83)
Head age (31-40 yrs old)	δ_7	-0.201	(2.91)
Head age (41-50 yrs old)	δ_8	-0.302	(0.90)
Head age (51-60 yrs old)	δ_9	-0.050	(0.19)
Head age (>=61 yrs old)	δ_{10}	-0.169	(10.46)
Mechanized (Dummy)	δ_{11}	-0.414	(2.23)
South (Dummy)	δ_{12}	-1.825	(2.37)
Year 96 Dummy	δ_{13}	-1.133	(1.79)
Year 97 Dummy	δ_{14}	-0.916	(5.63)
Year 98 Dummy	δ_{15}	-1.644	(1.06)
Year 99 Dummy	δ_{16}	0.670	(1.05)
<i>Variance Estimates</i>	σ^2	0.522	(2.46)
<i>Ln(likelihood): 1462.29</i>	Γ	0.99998	(1759.97)

Table 2-3: Likelihood ratio test for functional form and input effect

H_0	Hypothesis	Ln(lik)	λ	D.F	Critical Value ¹³	Inference
H0: $\beta_{hh}=0$	There is no fixed effect	-704.8	4334.2	590	647.6	Reject
H0:	Restrict the $\beta_1 \dots \beta_{18}$ of Model 1 as the same as these of without fixed effects	1353.5	217.6	18	28.9	Reject
H0: $\beta_{ij}=0$	Frontier is of Cobb-Douglas Form	1371.1	182.5	10	18.31	Reject
H0: $\beta_1 = \beta_{1j}=0$	Var. Land does not affect production frontier	621.0	1682.6	5	11.07	Reject
H0: $\beta_2 = \beta_{2j}=0$	Var. Labor does not affect production frontier	1405.1	114.4	5	11.07	Reject
H0: $\beta_3 = \beta_{3j}=0$	Var. Fertilizer does not affect production frontier	1412.4	99.8	5	11.07	Reject
H0: $\beta_4 = \beta_{4j}=0$	Var. Capital does not affect production frontier	1326.3	272.0	5	11.07	Reject
H1: Negation	Translog functional form	1462.3				

¹³ The critical values correspond to 5 percent level of significance.

Table 2-4: Curvature conditions estimates for RCRE grain production function

Curvature	Method		Land	Labor	Fertilizer	Capital	Scale Elasticity
Elasticity	Evaluate at	Geometric mean	0.677	0.001	0.052	0.142	0.872
		median	0.681	0.004	0.053	0.138	0.876
	Average		0.677	0.001	0.051	0.135	0.866
	Weighted	$y_i/\text{sum}(y_i)$	0.671	-0.004	0.046	0.196	0.910
K-means	Clustering	Group 1 (1614 obs)	0.688	-0.001	0.051	0.152	0.890
		(weighted) Group 1 (1614 obs)	0.679	-0.005	0.045	0.215	0.934
	Group 2 (1094 obs)	0.661	0.005	0.054	0.111	0.831	
	(weighted) Group 2 (1094 obs)	0.653	0.001	0.049	0.152	0.856	
Marginal Effects	Evaluate at	Geometric mean	210.99	0.02	0.34	0.36	
	Average		228.02	-0.06	0.42	0.38	
	Weighted	$y_i/\text{sum}(y_i)$	258.42	-0.19	0.57	0.61	
K-means	Clustering	Group 1 (1614 obs)	239.29	-0.11	0.35	0.39	
		(weighted) Group 1 (1614 obs)	247.24	-0.24	0.43	0.56	
	Group 2 (1094 obs)	211.40	0.004	0.53	0.37		
	(weighted) Group 2 (1094 obs)	283.59	-0.09	0.88	0.72		

Table 2-5: Technical efficiency distribution

Category	Group	Mean	Std. Dev.	Observations
	Overall	0.827	0.003	2708
By year:	1995	0.808	0.007	572
	1996	0.840	0.006	538
	1997	0.833	0.006	539
	1998	0.883	0.005	539
	1999	0.768	0.009	520
By province:	Anhui	0.809	0.009	294
	Hebei	0.786	0.008	659
	Helongjiang	0.847	0.006	362
	Jiangsu	0.905	0.004	292
	Liaoning	0.726	0.020	110
	Shandong	0.854	0.010	145
	Shanxi	0.817	0.011	249
	Sichuan	0.869	0.006	411
	Yunnan	0.794	0.012	186
Mechanization	No	0.820	0.004	1998
	Yes	0.845	0.005	710
Region	North	0.808	0.005	1525
	South	0.851	0.004	1183

Table 2-6: Elasticities estimates for Chinese agricultural production function

		Land	Labor	Fertilizer/ variable inputs	Capital	Scale Elasticity	Period	Data Type
Our Study	Geometric mean	0.68	0.00	0.05	0.14	0.87	95-99	Households
	Median	0.68	0.00	0.05	0.14	0.88		
Wan & Cheng (2001) ¹⁴		0.77	0.10		0.13	1	70-85	Production Team
		0.81	-0.01		0.20	1		
		0.99	-0.21		0.30	1.08		
		0.90	-0.01		0.09	0.98		
		0.80	-0.00		0.32	1.11		
Putterman (1993)		0.66	0.06	0.06	0.02		70-85	Production Team
Tian & Wan (1993)		0.66	0.06	0.06	0.02		70-85	Production Team
Fleisher & Liu (1992)		0.70	0.20	0.09	0.06	1.05	1986	Households
Kim (1990)		0.52	0.09	0.05	0.08		80-84	Production Team
Kim (1990)		0.66	0.08	-0.01	0.23		81-87	Province
Wiemer (1990)		0.54	0.20	0.11	0.09		70-79	Province
							83-85	
Lin (1992)		0.63	0.13	0.18	0.06		70-79	Province
Park (1989)		0.46	0.04	0.30	0.00		81-87	
							1985	Households

¹⁴ For Maize, Later rice, Wheat, Early rice and Tubers respectively.

Table 2-7: Marginal effects of inefficiency explanatory variables

	dlny/dx (average across individual)		dy/dx (average across individual)		dlny/dx (evaluate at the mean)		dlny/dx (evaluate at the median)	
Highest education (6-8 years)	1.494	(4.36)	4666.0	(4.29)	1.420	(5.50)	1.566	(2.36)
Highest education (9-11 years)	1.147	(1.83)	3582.5	(1.79)	1.090	(1.76)	1.202	(1.57)
Highest education (>12 years)	1.195	(1.45)	3733.1	(1.42)	1.136	(1.41)	1.253	(1.20)
Village officer Dummy	-0.442	(2.19)	-1379.2	(2.13)	-0.420	(2.27)	-0.463	(1.60)
Number of plot	-0.090	(2.20)	-281.6	(2.17)	-0.086	(2.32)	-0.094	(1.62)
Land Ratio	1.453	(2.60)	4538.1	(2.56)	1.381	(2.82)	1.523	(1.82)
Head age (31-40 yrs old)	0.211	(2.78)	660.3	(2.71)	0.201	(2.91)	0.222	(1.90)
Head age (41-50 yrs old)	0.318	(1.07)	992.9	(1.05)	0.302	(0.90)	0.333	(1.51)
Head age (51-60 yrs old)	0.053	(0.22)	164.0	(0.21)	0.050	(0.19)	0.055	(0.25)
Head age (>=61 yrs old)	0.177	(3.47)	554.1	(3.40)	0.169	(9.39)	0.186	(1.62)
Mechanized (Dummy)	0.436	(2.17)	1361.2	(2.12)	0.414	(2.23)	0.457	(1.59)

Figure 2-1: Technical efficiency, Chinese farms, 1995-1999

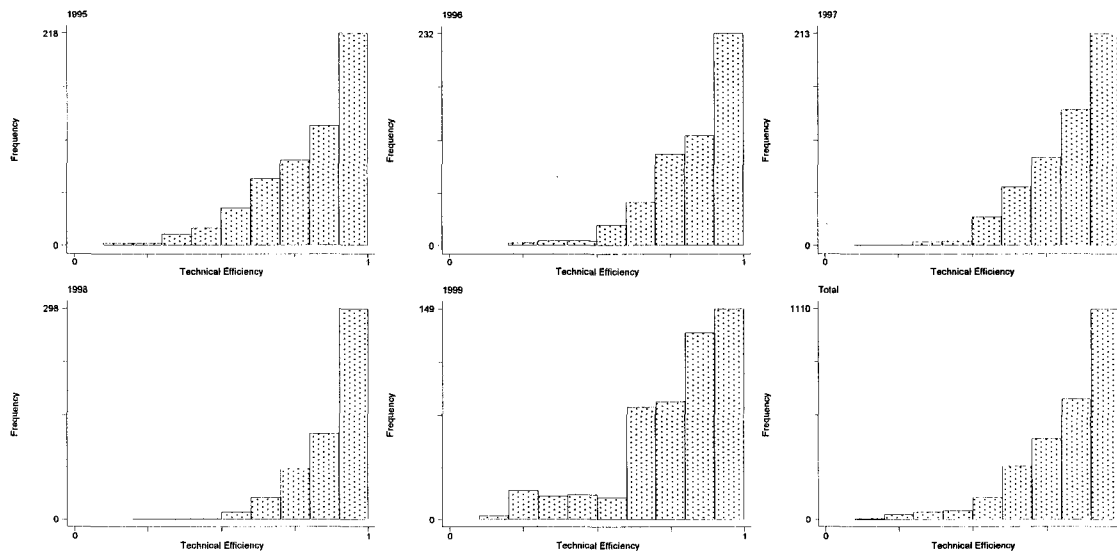


Figure 2-2: Technical efficiency by province, Chinese farms, 1995-1999

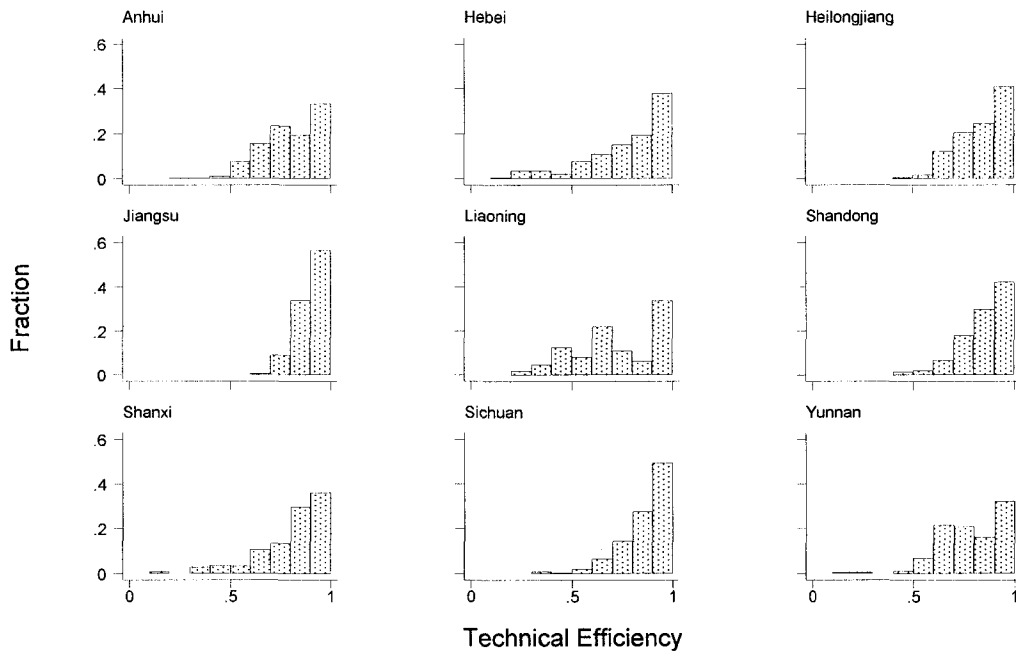
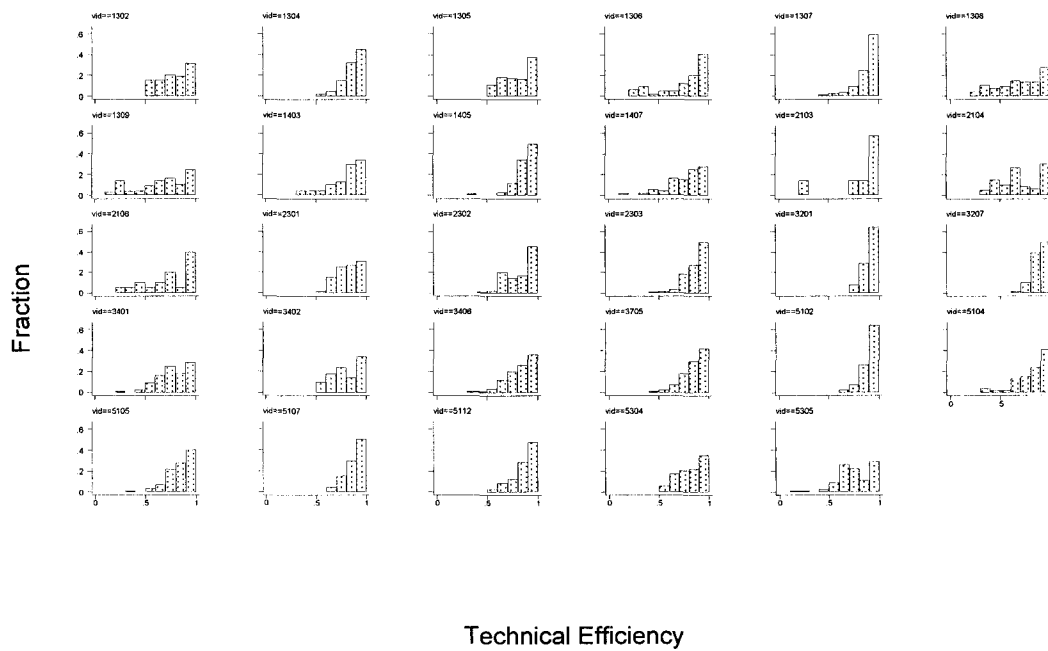


Figure 2-3: Technical efficiency by village, Chinese farms, 1995-1999



3. THE RELATIONSHIP BETWEEN FARM SIZE AND PRODUCTIVITY IN CHINA'S AGRICULTURE

Introduction

In the developing economies, farm size and productivity seem to be inversely related. The inverse relationship is formally defined as average grain yields fall when the size of farm increases.¹⁵ Chayanov (1926) is credited with first noticing this relationship in Russian agriculture, but Sen (1962) is believed to be the earliest modern reference on this subject. Berry and Cline (1979) reviewed the early empirical evidences on farm size and productivity as well as related econometric issues. In American agriculture, farm size and productivity are believed to be positively related or unrelated (Huffman and Evenson 2001; Hallam 1993). However, Ahearn, Yee and Huffman (2002) show that average farm size in the U.S. is negatively related to multifactor productivity over 1960-1996.

Sen (1962) explained the inverse relationship with labor dualism, where given the same technology, small-scale farmers have lower opportunity costs of their labor than operators of large farms. Deininger and Feder (2001) applied agency theory analysis on this subject. When a farm is small and labor markets are not functioning, small-scale farms use only family labor (Taylor and Adelman 2003). Hence, in the terminology of principal-agent theory, the principal and his family members supply all of the labor for the farm. These family

¹⁵ Benjamin (1995) regressed output on farm size and claimed that there exists inverse relationship if the coefficient of farm size is less than one.

members have a strong incentive to work because they share the farm output directly and in the long run can expect to inherit the farm. Here monitoring and incentive problems are minimal, and excess family labor may push the value of the marginal product below the off-farm wage thus result the inverse relationship. Bhalla and Roy (1988) and Benjamin (1995) suggested that unobserved land quality is positively related to farm productivity but inversely related to farm size, which might explain the inverse relationship between farm size and productivity as well.¹⁶ Heltberg (1998) claimed that Bhalla and Roy's conclusions are undermined by their use of district aggregate data. However, using farm level data obtained in Haryana, India, Carter (1984) found a significant within-village inverse relationship between farm size and productivity. Heltberg also noted that Benjamin (1995)'s first-stage regression has a very low *R*-square (0.12-0.14) and suggested that weak instruments have undermined Benjamin's analysis (Bound et al. 1995).

For China's agriculture, few studies exist on the inverse relationship between farm size and productivity. Brandt (1985) briefly reviewed the relationships between farm size, productivity, and factor markets for the pre-war (1930s) northeastern China. Later, Benjamin

¹⁶ There is another approach that we may pursue to explain the inverse relation, i.e., we can put the residual obtained in the MLE estimation of Chapter 2 to proxy land quality since we didn't control for land quality. However, there are two types of risk we have to take. The first is that the residual includes the quality effects of other input (it is fine with the regression if they are not correlated with the error term) and other random disturbances. The second is that which part of the residual should be included? There are two random terms in the stochastic frontier. Should we use the frontier residual or the efficiency residual or both? It will be a difficult decision to make. It is also hard to assess which approach (between this approach and the instrumental variable approach) is better in terms of the statistical efficiency. However, we probably can justify the instrumental variable approach since it has been adopted and well established in the development literature.

and Brandt (2002) attributed the inverse relationship between size and productivity in China's agriculture to local administrative land distribution policies and uneven off-farm work opportunities.¹⁷ Using a panel of farm households during the late 1990s, this study examines the relationship between farm size and productivity in China's agriculture, where average farm size is small and technology is only slightly dynamic. Firstly, we utilize the fact that an egalitarian land distribution policy existed under the Household Responsibility System. We then develop some instruments for farm size and apply a Hahn-Hausman specification test for the instruments. Second, Murphy and Topel (1985) claimed that the traditional form of the covariance matrix estimator for a two-step estimator is biased. Using their methodology, we identify four different approaches for estimating the variance of the instrumental variable estimate.

The rest of the chapter is organized as follows: In the second section, we describe the data and examine the relationship between farm size and grain crop yields for Chinese grain farms. In the third section, we discuss the likely effects of unobserved land quality and argue that a two-step estimation procedure should be applied. We perform the second-stage analysis and test the null hypothesis of no inverse relationship between farm size and productivity. We also derive the MT-type variance estimators in Section 3. In Section 4, discussions and conclusions are presented.

¹⁷ The areas they surveyed are in the geographic area same as or close to these of Brandt (1985).

Econometric issues in the farm size and productivity relationship

The relationship between farm size and productivity

A number of complex issues are raised when one seriously considers the relationship between farm size and productivity. In developed countries, it is well accepted that farm productivity and size are positively related (Hallam 1993), but in low-income countries it is equally well accepted that farm size and productivity are negatively correlated. However, measurement of size, quality adjustments, endogeneity, and the correct variance for the estimator are all pursued here. To related farm size and productivity, Benjamin (1995) and Heltberg (1998) considered a simple regression of logarithm of grain output on logarithm of sown area:

$$(3-1) \ln q_i = \alpha_0 + \gamma \ln l_i + \eta_i$$

where q is the real value of grain output, l is land area in grain production, η is the random disturbance, and i is the index for the observation—note we ignore the potential fixed effects at this stage and pool the panel together. Note that if γ is equal to 1 in equation (3-1) then output per unit land is unrelated to the farm size. If γ is less than one, then grain yield per unit land declines as land area increases. Finally, if γ is larger than one, then grain yield per unit of land increases with land area. Benjamin (1995) argued that using the actual area harvested rather than total farm size reduces measurement error, which otherwise could introduce a spurious inverse relationship.

Equation (3-1) is a simple regression and Benjamin (1995) and Cheng (1998) have suggested additional variables that might be included:

$$(3-2) \ln q_i = \alpha_0 + \gamma \ln l_i + \beta_{D_{rice}} D_{rice,i} + \beta_{D_{soya}} D_{soya,i} + \beta_{D_{vo}} D_{vo,i} + \eta_i$$

$$(3-3) \ln q_i = \alpha_0 + \gamma \ln l_i + \beta_{D_{rice}} D_{rice,i} + \beta_{D_{soya}} D_{soya,i} + \beta_{D_{vo}} D_{vo,i} + \sum_v \beta_{D_v} D_{v,i} + \eta_i$$

$$(3-4) \ln q_i = \alpha_0 + \gamma \ln l_i + \beta_{D_{rice}} D_{rice,i} + \beta_{D_{soya}} D_{soya,i} + \beta_{D_{vo}} D_{vo,i} + \sum_{hh} \beta_{D_{hh}} D_{hh,i} + \eta_i$$

where D_{soya} takes a value of one if the household harvests soya (zero otherwise); D_{rice} is one if the household harvests rice (zero otherwise). D_{vo} is one if a member of the farm household holds a position as a village officer (zero otherwise). D_v is one for region v (and zero otherwise, we use the first village as the reference group).¹⁸ D_{hh} has a value of 1 for household h (and zero otherwise, note we use the first household as the reference group).

In rural China, under the Household Responsibility System the majority of the arable land is owned by rural communities and managed by the village council (Agriculture Law 1993). Yao (2001) noted that the politics in rural China is characterized by a mixture of an authoritarian command system and grass-roots democracy, which establishes the legality of the egalitarian principle during land distribution and ensures its implementation. Furthermore, a recently passed law, Law of Rural Land Contracting (2002, LRLC henceforth), explicitly states the egalitarian principle.¹⁹ It stresses that women have the same right as men during land distribution (Article 6, LRLC) and forbids local governments to void the HRS contract

¹⁸ Note that we use the first region as the reference group.

¹⁹ LRLC was passed in August 2002 and has been effective since March 1, 2003.

by specifying that no reallocation is allowed during the term of the contract. Exceptions need to be ratified by at least two thirds of the village council (Article 48, LRLC).

If distributing equal amount of “effective land units” or quality adjusted land area is the objective of the local public land distribution system, then we can write:

$$(3-5) l_i = a_i h_i$$

where l is homogenous land or effective units of land, h is average land quality and a is nominal land areas. For simplicity, impose the following normalization on h , $\prod_{i=1}^n h_i = 1$. We can write equation (3-5) as:

$$(3-6) \ln l_i = \ln a_i + \ln h_i$$

Now when the “true relationship with productivity” is:

$$(3-7) \ln y_i = \alpha_L + \gamma_L (\ln a_i + \ln h_i) + \eta_{Li},$$

then if one fits equation (3-1) the estimate of γ will obviously be different from γ_L .

In particular, if denote the 2708×1 vector $\ln a$ as \mathbf{A} and $\ln h$ as \mathbf{H} , we have:

$$(3-8) \gamma = \gamma_L + \rho \hat{\gamma}_L \text{ while } \rho = (\mathbf{A}' \mathbf{A})^{-1} \mathbf{A}' \mathbf{H}$$

It is possible that $\gamma_L > 1$ but $\gamma < 1$ when a and h are negatively correlated, i.e. $\rho < 0$, which could occur under an egalitarian land distribution policy.

Benjamin (1995) used the instrumental variable approach to explain the inverse relationship with unobserved land quality. Under local egalitarian land distribution, each household is presumably allocated certain amount of land according to attributes of the household, e.g. household size, number of workers, and the extent of farm specialization in

grain production. Denote the matrix of the logarithm of those variables as \mathbf{X} and assume:

$$p \lim \frac{1}{n} \{\mathbf{X}'\mathbf{X}\} = \mathbf{Q}_{\mathbf{XX}}, \text{ which is a finite, positive definite matrix,} \quad (\text{A1})$$

$$p \lim \frac{1}{n} \{\mathbf{X}'\mathbf{A}\} = \mathbf{Q}_{\mathbf{XA}}, \text{ is a finite matrix,} \quad (\text{A2})$$

$$p \lim \frac{1}{n} \{\mathbf{X}'(\boldsymbol{\eta} - \mathbf{H})\} = 0. \quad (\text{A3})$$

A1 and A2 are quite straightforward and A3 implies that: $p \lim \frac{1}{n} \{\mathbf{X}'\mathbf{H}\} = 0$, which means that \mathbf{X} is uncorrelated with \mathbf{H} , i.e., the household's characteristics are not correlated with unobserved land quality.

Now under the egalitarian land distribution policy assume that:

$$(3-9) \ln(l)_i = \mathbf{X}_i\boldsymbol{\theta} + \eta_i$$

but due to the unobservable nature of effective land units, we fit:

$$(3-10) \ln(a)_i = \mathbf{X}_i\boldsymbol{\theta} + \eta_{Ai}$$

Hence, we have the predictor of land as: $\widehat{\mathbf{A}} = \mathbf{X}_i\boldsymbol{\theta} + \mathbf{P}_X(\boldsymbol{\eta} - \mathbf{H})$

by A3, we obtain: $p \lim \mathbf{P}_X(\boldsymbol{\eta} - \mathbf{H}) = 0$

thus, $\widehat{\mathbf{A}}$ converges to $E(\mathbf{L})$ in probability, where \mathbf{L} is the effective or constant quality land area and $E(\mathbf{L})$ is its linear best predictor. Therefore, we can use instrumental variable estimation to correct for the land quality problem. We assume that we can obtain a quality adjusted land area from the first stage. An analog is the estimation of permanent income, which is endogenous but we can correct the endogeneity problem using instrumental variable estimation. Under certain condition, instrumental variable estimation can be deemed as a

two-stage estimation. However, a two-step estimator may have better statistical properties than a two-stage estimator, which will be discussed later.

Consider the first stage regression explaining nominal land area:

$$(3-11) \ln a_i = \theta_0 + \theta_1 \ln p_i + \theta_2 \ln f_i + \theta_3 \ln k_i + \theta_4 D_{hht,i} + \theta_5 D_{rice,i} + \theta_6 D_{soya,i} + \theta_7 D_{vo,i} + \sum_v \theta_{D_v} D_{v,i} + \eta_i$$

where a is the land area for grain production; p is the number of household members; f is the number of household labor at the current year, and k is the value of agricultural equipment and machinery. D_v and D_{vo} are the village and village officer dummy variables defined previously. The θ s are unknown parameters to be estimated. The second stage equation is:

$$(3-12) \ln q_i = \beta_0 + \beta_1 \widehat{\ln a_i} + \beta_{D_{soya}} D_{soya,i} + \beta_{D_{rice}} D_{rice,i} + \beta_{D_{hht}} D_{hht,i} + \beta_{D_{vo}} D_{vo,i} + \sum \beta_{D_v} D_{v,i} + u_i$$

where q is grain output in real terms, $\widehat{\ln a}$ is the predicted logarithm of land area from equation (3-11), D_v and D_{vo} are defined as above, u_i is the random disturbance. The least-squares estimate of equation (3-12) is reported in Table 3-4.

The traditional variance estimate of the IV estimator ignores the fact that parameters were estimated in the first step and treat these parameters as constants, not random variables. This may produce a biased variance estimate in the second step (Ruud 2000). The existing literature has two versions of the instrumental variable estimator. Staiger and Stock (1997) and others assumed that:

$$y_1 = \beta y_2 + \varepsilon_1$$

$$y_2 = z\pi_2 + v_2$$

where we remove all the individual indexes and use y_2 to denote the stochastic regressor and

z the instruments. We do not consider the case where included exogenous variables exist in the first equation since we can invoke Frisch-Waugh theorem to partial them out. This model uses all the exogenous variables as instruments. However, we may use more generalized two-step procedure where we do not need to include all exogenous variables in the first step and even the linear assumption in the first step can be relaxed.

Following Hahn and Hausman (2002) and Murphy and Topel (1985), we write the equation of the first structural equation as:

$$y_1 = \beta y_2 + \varepsilon_1 = \beta(z\pi_2 + v_2) + \varepsilon_1 = \gamma z\pi_2 + v_1$$

$$y_2 = z\pi_2 + v_2$$

Although numerically identical β and γ have different interpretations. β is the coefficient for the observed value y_2 while γ is the coefficient for the unobserved $z\pi_2$. An example is the case of quality adjusted land area. The instrumental estimates of the coefficient of the observed land measure and the coefficient of the imputed homogenous land have same point estimates when all second stage exogenous variables are included in the first stage regression. They differ if there are instruments excluded in the second step. In either case, their variance estimates are different as will be shown in the following.

The usual variance estimates for the IV estimator are:

$$\begin{aligned} \text{var}(\hat{\beta}_{IV}) &= \text{var}\left(\left(\hat{y}_2' y_2\right)^{-1} \hat{y}_2' y_1\right) = \text{var}\left(\left(\hat{y}_2' y_2\right)^{-1} \hat{y}_2' (\beta y_2 + \varepsilon_1)\right) \\ &= \text{var}\left(\left(\hat{y}_2' y_2\right)^{-1} \hat{y}_2' \varepsilon_1\right) = \left(\hat{y}_2' y_2\right)^{-1} \hat{y}_2' \text{var}(\varepsilon_1) \hat{y}_2 \left(\hat{y}_2' y_2\right)^{-1} \end{aligned}$$

when assuming that no heteroscedasticity occurs, we have:

$$(3-13) \quad \text{var}(\hat{\beta}_{IV}) = \sigma_1^2 \left(\hat{y}_2' y_2 \right)^{-1} = \sigma_1^2 \left(\hat{y}_2' \hat{y}_2 \right)^{-1}$$

The commonly used estimator for σ_1^2 is:

$$(3-14) \quad \hat{\sigma}_1^2 = \frac{1}{n-k} (y_1 - \hat{\beta}_{IV} y_2)' (y_1 - \hat{\beta}_{IV} y_2)$$

However, if our main interest is to obtain the variance estimate of γ , which is coefficient of the unobserved $z\pi_2$, we need to proceed in another direction.

$$\begin{aligned} \text{var}(\hat{\gamma}_{2SE}) &= \text{var} \left(\left(\hat{y}_2' y_2 \right)^{-1} \hat{y}_2' y_1 \right) = \text{var} \left(\left(\hat{y}_2' y_2 \right)^{-1} \hat{y}_2' (\gamma z \pi_2 + v_1) \right) \\ &= \text{var} \left(\left(\hat{y}_2' y_2 \right)^{-1} \hat{y}_2' (\hat{y}_2 \gamma - \gamma z (\hat{\pi}_2 - \pi_2) + v_1) \right) \\ (3-15) \quad &= \text{var} \left(- \left(\hat{y}_2' y_2 \right)^{-1} \hat{y}_2' \gamma z (\hat{\pi}_2 - \pi_2) + \left(\hat{y}_2' y_2 \right)^{-1} \hat{y}_2' v_1 \right) \\ &= \gamma^2 \left(\hat{y}_2' y_2 \right)^{-1} \hat{y}_2' z \text{var}(\hat{\pi}_2 - \pi_2) z' \hat{y}_2 \left(\hat{y}_2' y_2 \right)^{-1} + \left(\hat{y}_2' y_2 \right)^{-1} \hat{y}_2' \text{var}(v_1) \hat{y}_2 \left(\hat{y}_2' y_2 \right)^{-1} \\ &\quad - 2 \left(\hat{y}_2' y_2 \right)^{-1} \hat{y}_2' \gamma z \text{cov}(\hat{\pi}_2 - \pi_2, v_1) \hat{y}_2 \left(\hat{y}_2' y_2 \right)^{-1} \end{aligned}$$

Obviously, equations (3-13) and (3-15) are different as Murphy and Topel (1985) noted. Although they are asymptotically equivalent when the Newey condition holds, the difference in finite samples cannot be ignored when making inferences regarding the coefficient of the stochastic regressor (Ruud 2000).

The difference between the two variance estimators is partly due to the fact that the estimate of the variance of the IV estimator omits the correlation between y_2 and the error

term. It ignores the fact that parameters were estimated in the first step, which may cause mis-specification of the sampling variance of the second stage estimator (Ruud 2000).

To extend the analysis to the variance estimation to the case when there are included exogenous variables in the second stage, we use the following matrix form:

$$(3-16) \quad L = X\theta + \varepsilon \quad (\text{Stage 1})$$

$$(3-17) \quad Y = X_2\beta + \gamma X\theta + u \quad (\text{Stage 2})$$

$$(3-18) \quad Y = X_2\beta + \gamma L + u^* \quad (\text{Stage 2'})$$

where ε , u and u^* are random disturbances.

The two-step estimator is:²⁰

$$(3-19) \quad \begin{aligned} (\hat{\beta}', \gamma)'_{2SE} &= (Z'Z)^{-1}Z'(X_2\beta + \gamma X\theta - \gamma X(\theta - \theta) + u) \\ &= (Z'Z)^{-1}Z'Z(\beta', \gamma)' - \gamma(Z'Z)^{-1}Z'X(\theta - \theta) + (Z'Z)^{-1}Z'u \\ &= (\beta', \gamma)' - \gamma(Z'Z)^{-1}Z'X(\hat{\theta} - \theta) + (Z'Z)^{-1}Z'u \end{aligned}$$

where $Z = (X_2 | P_X L)$

The usual instrumental variable estimation (or two-stage estimation) usually includes all possible instruments thus the estimator for the parameters is:

$$(3-20) \quad \begin{aligned} (\hat{\beta}', \gamma)'_{IV} &= (\hat{Z}'\hat{Z})^{-1}\hat{Z}'(X_2\beta + \gamma L + u^*) \\ &= (\beta', \gamma)' + (\hat{Z}'\hat{Z})^{-1}\hat{Z}'u^* \end{aligned}$$

where $\hat{Z} = P_X(X_2 | L)$.

²⁰ The example can be treated as a special case of Murphy and Topel (1985) Theorem 1. Here we specified a linear functional form in the first step.

From the above expressions, including $L = X\theta + \eta$, we obtain $u^* = u - \gamma\varepsilon$. The difference between the two-step estimator and instrumental variable estimator is:

$$(3-21) \quad (\hat{\beta}', \gamma)'_{2SE} - (\hat{\beta}', \gamma)'_{IV} = ((Z'Z)^{-1}Z' - (\hat{Z}'\hat{Z})^{-1}\hat{Z}')u - \gamma((Z'Z)^{-1}\hat{Z}' - (\hat{Z}'\hat{Z})^{-1}\hat{Z}')\varepsilon$$

Hence, the following proposition holds:

Proposition 1: $(\hat{\beta}', \gamma)'_{IV}$ equals to $(\hat{\beta}', \gamma)'_{2SE}$ if and only if X_2 is included in X .

The proof of this proposition is straightforward from the derivation. Two-step estimation is more general than instrumental variable estimation while two-stage estimation is a special case (most efficient without presence of heteroskedasticity) of instrumental variable estimation. However, when the condition in Proposition 1 holds, the three methods produce same result.

Variance estimates for the two-step estimation and instrumental variable estimator proceed in two different directions. The variance estimate of the instrumental estimator tries to construct a consistent estimator for u^* while the variance of the two-stage least squares estimator can be decomposed into two parts. Murphy and Topel (1985) argued that although the instrumental estimator (or two-stage least squares) yields a consistent estimator for second-stage parameters under fairly general conditions, the second-step standard errors and related test statistics based on this procedure are incorrect.

Previous econometric studies have regularly based inference on the traditional variance estimator of the instrumental variable estimator. For example, Benjamin (1995) used the robust standard error estimator to correct for arbitrary heterogeneity.

In our model, it is better to use equation (3-17) rather than equation (3-18) as the structural form. Recall that we are testing the null hypothesis that output per unit of land area is unrelated to size (land area) with the alternative hypothesis being that the output per unit of land is inversely related to farm size. We are more concerned with the effect of effective land area rather than measured land area. The estimate of γ may be biased and inconsistent if measured land area rather than effective land area is used.

We derive four variants of the variance estimators of the parameter estimates in the second stage. Due to their tie to Murphy and Topel (1985), we label them as MT variance estimates and present them in Table 3-4. Note that the adjusted standard error is much greater than the robust error given by 2SLS procedures. Therefore, using the usual variance estimates, we would be inclined to falsely conclude that there exists an inverse relationship given the land elasticity estimate is less than one. We also find that the adjusted standard error with assumption of no correlation between the two random components is identical (after rounding) to the adjusted estimates with correlation, which indicates that the random components are nearly independent, i.e., the random disturbance in the land distribution equation is independent of the random disturbance in the grain production.

Data and empirical results

The dataset

This study uses the same 591 households panel as described in the introduction and the second chapter. Summary statistics of household characteristics are presented in Table 3-1. Clearly, per capita or per household land area of rural China is very small compared to these of developed countries. The household size ranges from 1 to 12, with an average of about four members per family. Household labor ranges from 1 to 8 persons with an average of roughly 2.6 persons per household. Household agricultural productive asset ranges from RMB Yuan 0 to 65,000 with an average of RMB Yuan 1,567. The index illustrates the extent of household specialization in grain production. About 5 percent of the households actually have at least one member holding a village officer position. More detailed summary characteristics can be found in Table 2-1.

We fit equation (3-1) to the panel data set for Chinese farmers and obtain the estimate of γ as 0.890 with a standard error of 0.011 (see Table 3-2). Hence, using this methodology, we reject the null hypothesis that γ is one at the five percent significance level and conclude that output per unit of land declines as farm size increases. Excluded factors, e.g., climate, regional effects, population density, average land quality, however, could bias the size of γ (e.g., downward as Bhalla and Roy (1988) showed). Village level dummy variables can be used to control for differences due to climate, multiple cropping indexes and the regional

irrigation systems. Benjamin (1995) used a dummy variable for use of HYV (high yield varieties) versus traditional varieties.

In our dataset, grain output is measured as the aggregation value of wheat, rice, corn, and soya harvested.²¹ We observe that the wheat and corn have quite similar prices and yields per unit of land. Therefore we use dummy variables to indicate whether soya (rice) was planted to proxy the composition of grain output. Also, Cheng (1998) incorporated an indicator for a family member holding a “village official” position into the grain production function and found it is positive and statistically significant. He argued that the effects were most likely due to the local policy followed by collective ownership of large farm equipment and privileged access to state subsidized farm inputs. The above arguments justified the models used in equation (3-2), (3-3), and (3-4).

When equation (3-2) is fitted to our dataset, the estimate of γ is 0.93 with a standard error of 0.01 (see Table 3-2). Hence, the estimate of γ is significantly different from one, which supports the so-called inverse relationship. When we include village fixed effects as in equation (3-3), the estimate of γ is 0.92 and but still significantly different from one at the 5 percent level. Thus, by expanding the model from equation (3-1) to (3-2), and then (3-3), the size of the estimate of γ changes slightly—by only 0.03 and the *R*-square increases from 0.73

²¹ Pooling grain productions together is better than estimating the crops separately since households may produce a small amount of the crops that is not best suited for the local soil and climate to add some variety to their food composition. When estimating the relationship between productivity and land size, this may result in a pseudo positive relationship.

to 0.89. Since our dataset is a five-year panel, individual unobserved effects might exist. Equation (3-4), a household fixed effects model, can account for the unobserved effects. When equation (3-4) is fitted to the data, the estimate of γ is 0.84 with a standard error of 0.05 (see Table 3-4). Hence, we reject the null hypothesis that γ is equal to one and conclude that output per unit of land declines as land area increases. The R -square for this regression is 0.92. This relationship exists even after the inclusion of village or household effects, which is consistent with Carter (1984). We, however, have not controlled for land quality differences.

A role for land quality differences

Studies confirmed that the egalitarian principle has been adopted for land distribution in the majority of Chinese rural communities. Yao (2001) studied the effects of egalitarian land distribution on migration of rural men in China. Using household level data for two distinct provinces, Jiangsu and Sichuan, Burgess (2001) failed to reject the hypothesis of universal and egalitarian access to land. With an egalitarian principle in place, it is likely that average land quality and farm size are negatively correlated.

The least-squares estimate of equation (3-11) is reported in Table 3-3. The most notable feature of this equation is that the R -square is 0.79, and the partial R -square is 0.38, which means that the instruments are not “weak” (Bound et al 1995; Heltberg 1998). The least

squares estimate of equation (3-12) is reported in Table 3-4. The result suggests the inverse relationship might be explained with the unobservable land quality.²²

Some issues exist about the appropriateness of instruments, but the conditions are met in our data set.^{23 24}

²² Hall, Rudebusch, and Wilcox (1996) indicated that any relevance measure probably has little practical merit, as its use may actually exacerbate the poor finite-sample properties of the IV estimator. Their result is possibly caused by the inclusion of instruments that is Granger caused by the stochastic regressor. For the case of unobserved land heterogeneity, the objective of the first stage is to predict a homogenous measure that is free of land quality problem. As Davidson and Mackinnon (1993) pointed out, when the goal is forecasting, forecasts of the variable Y_t may be conditional on the variables X_t if Y_{t-1} does not Granger cause X_t . If the number of instruments is much less than the number of observations, including more instruments will increase the precision of prediction thus reduces the bias in the second stage. However, in finite samples, to include more instruments has two types of danger. First, it is likely to include some instruments that are Granger caused by the stochastic regressor and thus introduce correlation between the predicted values with the unobserved latent variables/random disturbance. These instruments cannot eliminate the problem of the correlation between regressor and the error term though they may have a good fit at the first stage. Second, including more instruments is at risk of constructing a linear space that is not orthogonal to the space spanned by the unobserved latent variables. An extreme example is that when we have N instruments for a dataset of N observations. If the data matrix of the N instruments is not singular, the predicted value of the stochastic regressor will be exactly same as the observed values, which means that the R -square is one and first stage random disturbances are zero. Obviously, the predicted value is still correlated with the error term in the second stage. In this study, however, we deem this danger less realistic since the first stage regression is based on our observation of China's agriculture.

²³ Hausman (1978) provided a specification test to examine whether the OLS estimate of a parameter is a consistent estimator. The test statistics H has an asymptotic chi-squared distribution with degrees of freedom being the number of regressors minus the number of instruments in the second stage, which is one in this case. For this study, the Hausman test statistics is 31.33 and the critical value of the test statistic is 3.84. Hence, the OLS estimator is not consistent and we need to use the instrumental variable estimation.

²⁴ Are the instruments appropriate? There are two aspects of this appropriateness. First, does the orthogonality condition between the instruments and the error term hold? An over-identification test can be used for this purpose (Ruud 2000, p573). The over-identification test statistics for this dataset is 3.17, which is distributed as chi-square with degree freedom 4. Hence, we cannot reject the null hypothesis that the orthogonality condition holds at 5% significance level. Second, are the instruments sufficiently correlated with the stochastic regressor? Much discussions about "weak instruments" has emerged over the past two decade, e.g., Staiger and Stock (1997) and Nelson and Startz (1990). Weak instruments may make the second stage inference invalid. Bound,

Hahn and Hausman (2002) proposed a specification test to determine whether conventional IV asymptotics are reliable. The test compares the difference of the conventional 2SLS estimate of the coefficient of the right-hand side endogenous variable with the reverse 2SLS estimate of the same unknown parameter. Under the null hypothesis that the conventional first-order asymptotics provide a reliable guide to inference, the two estimates should be very similar. The Hahn-Hausman specification test shows whether the resulting difference in the two estimates satisfies the results of second-order asymptotic theory.²⁵ The test statistic is:

$$(3-22) \quad m_1 = \frac{d_1}{\hat{w}_1^{0.5}}, \text{ where } m_1 = \sqrt{n} \left(b_{2SLS} - \frac{1}{c_{2SLS}} - \hat{B} \right), \text{ and}$$

$$(3-23) \quad \hat{w}_1 = 2 \frac{K-1}{n-K} \frac{(\sum_{i=1}^n (y_{1i} - \beta_{LIML} y_{2i})^2)^2 (y_2' P_z y_2 - \frac{K-1}{n-K} y_2' M_z y_2)^2}{(y_2' P_z y_2)^2 (y_2' P_z y_1)^2}.$$

The null hypothesis is:

Jaeger, and Baker (1995) suggested that partial R -squared and the F statistics of the identifying instruments in the first stage estimation are useful indicators of the quality of instruments. Consider the recursive model:

$$y_1 = \beta y_2 + \varepsilon_1$$

$$y_2 = z\pi_2 + v_2$$

It is straightforward to show that: $p \lim \hat{\beta}_{OLS} = \beta + \frac{\sigma_{y_2, \varepsilon_1}}{\sigma_{y_2}^2}$ and $p \lim \hat{\beta}_{IV} = \beta + \frac{\sigma_{y_2, \varepsilon_1}}{\sigma_{y_2}^2}$ which imply that:

$$\frac{p \lim \hat{\beta}_{IV} - \beta}{p \lim \hat{\beta}_{OLS} - \beta} = \frac{\sigma_{y_2, \varepsilon_1} / \sigma_{y_2, \varepsilon_1}}{R_{y_2, z}^2},$$

where $R_{y_2, z}^2$ is the R -square (the partial R -square if there are included exogenous variables) from the regression of y_2 on z . Obviously, as R -square increases, given a particular data set, the bias becomes smaller. The partial R -square for this model is 0.38, which suggests the instruments have reasonable explanatory power for effective land area.

²⁵ Please see Hahn and Hausman (2002) for details of the test.

$$H_0 : p \lim \sqrt{n} (b_{2SLS} - \frac{1}{c_{2SLS}} - \widehat{B}) = 0$$

Hahn and Hausman (2002) proved the test statistics has an asymptotic t distribution. The Hahn-Hausman test statistics for this data set is 0.35. Hence we cannot reject the null hypothesis that the 2SLS provides reliable inference.²⁶

Outlier could affect the outcome of the test, especially in small samples. One route to reduce this likelihood is to split the sample into two parts, say one sample is of households that have land areas less than or equal to 25 Mu (95% percentile). The results for this sample can be compared to that of the whole sample. Summarized in Table 3-6, they reveal a similar pattern to that of the whole sample. We also split the data into two samples and run the regressions. One of the samples is the group with farmland under grain production greater than or equal to 15 Mu and the other less. We found that the inverse relationship in the sample with larger land holdings is less severe. In fact the coefficient estimates is greater than one though not statistically significant. Land elasticity estimates for the group with larger land endowment are either similar or greater than those of the group with less land, though not statistically significant. Based on these results, we cannot reject the hypothesis that the coefficient of imputed land area is one. Hence, the empirical irregularity of the inverse

²⁶ Hahn and Hausman (2002) suggested that when the null hypothesis is rejected, a similar test based on Nagar-type estimators should be performed. If the second specification test rejects or the two Nagar-type estimators differ substantially, neither 2SLS nor LIML may provide reliable results for inference. If the null hypothesis is not rejected in the second test while the first test rejects, LIML estimator is preferred over 2SLS estimator. We only performed the first specification test. Please refer to Hahn and Hausman (2002) for details on the second specification test.

relationship between crop yields and farm size (land area) diminishes when we use the instrumental estimator with adjusted standard error.

Discussions and conclusions

Discussions

Deininger and Feder (2001) summarized several studies that confirmed the inverse relationship between farm-size and productivity. They argued that supervision cost for hired labor that comes with a larger farm is particularly large in agricultural production due to spatial dispersion and thus contribute to the inverse relationship. This can be interpreted as one reason why China's agriculture was transformed from collective farming to Household Responsibility System in the 1980s. Microeconomics theory suggests there is an optimum size for most production processes. Empirical evidence as summarized by Deininger and Feder (2001) indicate that the optimum farm size in most developing countries, given the existence of imperfect input/output/credit markets, low real wage, static agricultural technologies, and land heterogeneities, is small relative to the optimal size of farms in high wage, technically dynamic developed countries.

In China, we see a complicated picture. First, China has a very large rural population relative to the amount of arable land. The arable land per rural person is about 0.144 hectare for China, for India it is 0.221, the U.S. 2.729, and the world average is 0.426. Second, China is different from most other developing countries in that land is collectively owned by the

rural community instead of individuals. No large private-owned farms exist in China. One would expect large farms to be “specialized” and have subleased land from the community or other households. Both communities and households are more likely to sublease their less-productive plots. This contributes to the spurious inverse relationship between land productivity and farm size. Third, eastern and southern parts of China have seen an economic boost in last few decades, and a large number of rural laborers now are engaged in off-farm activities. The “land bank”s in Zhejiang province functioned as a rental market and successfully transferred lands from those households that are less relied on farming to these that are more “specialized” on farming. Similar institutional arrangement emerged in Jiangsu, Anhui, and Hunan, where rural laborers move out of agriculture sector to take local off-farm business or migrate.

The rapid economic development in China in the 1990s may have improved the function of input and output markets and most likely contributes to the weakening of inverse relationship between farm size and productivity. As China’s agriculture is mechanized and the input sector started to produce a steady stream of new technologies, larger farms will have a comparative advantage over small farms. Our results show that land heterogeneities contribute to the observed inverse relationship. This, as well as other studies (Benjamin 1995; Carter, 1984; Deininger and Feder 2001), points to an important conclusion: *the inverse relationship between farm-size and output per unit of land is not inherent to developing countries, but rather a consequence of heterogeneous land, (labor) market imperfection, and*

unobserved factors. Therefore, a public policy of breaking up large farms is not justified (Deininger and Feder 2001). The hidden unemployment problem can be improved by general economic development and investments in rural education (Huffman 2001; Huffman and Orazem, forthcoming). A mechanism that consolidates land to exploit the benefits of more advanced technology and to share such benefits between landowners and farmers is needed. The “land bank” in Zhejiang province is a result of efforts seeking such mechanism or an instrument of such mechanism. Further investigation and research are needed to judge whether these efforts are successful.

Conclusions

This study has examined the empirical relationship between farm size (measured in land area) and farm productivity (measured as grain output per unit of land) in China. Given that the local community council holds the majority of farmland and makes local land allocation decisions, we choose to use a two-step estimation procedure to examine in detail the relationship between farm size and productivity in small-scale agriculture. The data that we used are from a panel data set for Chinese farm household in the late 1990s. We find an inverse relationship between farm size and land productivity, but as we adjust for land quality and the likely endogeneity of effective land units per farm, we find that the inverse relationship is partially or completely diminished.

The study advances the methodology of variance estimators for the instrumental variable

estimation. We apply the Hahn-Hausman test (Hahn and Hausman 2002) to examine whether the two-stage least square estimator provides an appropriate estimates. We also derive MT type variance estimators for the instrumental estimates. More work, however, remains to be done to examine the impact of the emerging new land institutions on agricultural productivity in rural China.

Table 3-1: Descriptive statistics of the RCRE data set (2708 observations)

	Mean	Std. Dev	Min	Max
Land Under Grain Production (Mu)	9.866	8.772	0.3	150.0
Land Under Agricultural Use (Mu)	11.257	9.371	0.3	150.0
Household Population (People)	4.265	1.363	1	12
Household Labors (People)	2.564	1.065	1	8
Grain Production (kilogram)	3159.2	3621.6	50	77000

Table 3-2: Evidence of inverse relationship (2708 observations, OLS regression)

Equation	(3-1)	(3-2)	(3-3)	(3-4)
Ln(Land)	0.89 (0.01) ²⁷	0.93 (0.01)	0.92 (0.02)	0.84 (0.05)
Village Officer Dummy		-0.03 (0.03)	-0.06 (0.02)	0.00 (0.04)
Dummy (Soya)		-0.18 (0.01)	-0.05 (0.02)	-0.07 (0.02)
Dummy (Rice)		0.24 (0.01)	0.08 (0.03)	0.03 (0.04)
Village/Household and Time Dummies			Village Effects	Household Effects
Constant	5.96 (0.02)	5.89 (0.04)	5.88 (0.04)	6.09 (0.09)
R-square	0.73	0.77	0.89	0.92
Adjusted R-square	0.73	0.77	0.89	0.90

²⁷ The numbers in bracket are robust (White) variance estimates.

Table 3-3: First stage regression: land allocation in rural China

	Village effects model		Household effects model	
Ln(Labor)	0.13	(0.03)	0.12	(0.04)
Ln(Household size)	0.57	(0.02)	0.51	(0.05)
Ln(Agricultural Productive Asset)	0.03	(0.00)	0.03	(0.01)
Village Officer Dummy	-0.16	(0.03)	-0.04	(0.05)
Household Type (Agricultural)	0.20	(0.07)	0.01	(0.12)
Constant	0.31	(0.09)	0.48	(0.14)
Village/Household, Time, crop variety Dummies	Omitted		Omitted	
R-square	0.79		0.92	
Adjusted R-square	0.78		0.90	
Partial R-square (Excluded Instruments)	0.38			

Table 3-4: Inverse relationship: Second stage regression, village effects model

(Dependent Variable: lnY)	Estimate	2SLS Std. Err.	Robust Std. Err.	MT1 Std. Err.	MT2 std. Err.	MT3 std. Err.	MT4 std. Err.
Constant	5.735						
Fitted value of Ln(Land)	0.998	0.018	0.022	0.031	0.031	0.047	0.047
Village Officer Dummy	-0.047	0.022	0.023	0.037	0.037	0.050	0.050
Dummy (Soya)	-0.068	0.017	0.018	0.028	0.028	0.036	0.036
Dummy (Rice)	0.046	0.029	0.033	0.048	0.048	0.087	0.087
Village/Time Effects	Omitted						
Over-identification Test	Chi(4)		3.17				
Hausman Test	Chi(1)		31.33				
Hahn-Hausman Test	Asymptotic <i>t</i>		0.36				
R-square	0.89			Partial R-square (1 st Step)		0.38	

Table 3-5: Inverse relationship: Second stage Regression, household effects model

(Dependent Variable: lnY)	Estimate	2SLS Std. Err.	Robust Std. Err.	MT1 Std. Err.	MT2 std. Err.	MT3 std. Err.	MT4 std. Err.
Fitted value of Ln(Land)	0.990	0.037	0.049	0.063	0.063	0.084	0.084
Over-identification Test	Chi(4)		0.76				
Hausman Test	Chi(1)		14.53				
Hahn-Hausman Test	Asymptotic <i>t</i>		1.24				
R-square	0.90			Partial R-square (1 st Step)			0.18

Table 3-6: Inverse relationship: Reduced samples (dependent variable: lnY)

Coefficient of Land	OLS Estimate	IV Estimate	Partial R-squar e	Over-id Test	Hausman Test	Hahn- Hausman Test
Land < 25	N=2575					
Village Effects	0.91 (0.02) ²⁸	1.00 (0.05) ²⁹	0.38	4.55	33.00	-0.21
Household Effects	0.79 (0.03)	0.96 (0.11)	0.17	3.69	15.21	-0.27
Land < 15	N=2165					
Village Effects	0.90 (0.02)	0.99 (0.03) ³⁰				
Household Effects	0.80 (0.04)	0.94 (0.07)				
Land >= 15	N=543					
Village Effects	0.89 (0.05)	1.16 (0.11)				
Household Effects	0.83 (0.06)	0.97 (0.18)				

²⁸ Robust standard error.²⁹ MT1 standard error.³⁰ Here we just use the usual robust standard error.

4. JOB LOCATION CHOICE DECISION IN RURAL CHINA

Introduction

A dilemma China faced along its route to sustainable development is how to absorb the large number of agricultural labors that are being released as farmers adopt new technologies and agricultural productivity increases. If a large pool of unemployed labors develops it could become the hotbed for social unrest. Continuing with conventional agricultural production technology that employs a large percentage of rural labor force would, however, hinder China's development. An alternative strategy is to permit individuals and families to emigrate from the rural areas to the major cities. Rural labors have been emigrating from agriculture and taking local off-farm work.³¹ Treating the decision-making process of rural households as a stochastic process, Mohapatra, Rozelle and Huang (2003) studied the evolution of modes of production (including farming, non-farm activities, working in an enterprise, and migration) in rural China during 1990s. They found a systematic pattern emerging in different modes of production across space and time.

With an estimate of more than 100 million internal migrants (most from rural to urban areas), labor migration no doubt is a serious concern for Chinese policy makers and

³¹ Many rural Chinese households engage in off-farm activities, e.g. employment in TVE (town and village enterprise), transportation, construction, small business and services. However, they are classified as rural household since they still engage in agricultural production in varying extensity at the same time. It is well accepted that off-farm activities has been an important part of Chinese rural economy (Parish, Zhe, and Li, 1995; Rozelle et al., 1999).

researchers.³² Migrant laborers have brought and are bringing tremendous change to China's economy. They are building skyscrapers, preparing foods and providing domestic services in the cities. In the villages where these migrants come from, the remittance is an important component of the rural revenue. However, labor migration has been treated cautiously by Chinese officials, largely due to the social imbroglio it caused (Murphy 2002; Hare 2002). Urban and suburban areas have been troubled by increasing crime rates associated with higher population mobility. The fact that young and better-educated individuals compose the largest portion of migrants concerns agricultural researchers. They believe rural out-migration leaves aged and less educated workers to work on the farm, which may adversely affect agricultural production efficiency. Previous studies have shown the importance of education on rural households' decisions to engage in local off-farm activities—or to leave the farm, in addition to migrating—to leave the villages (Tuan, Somwaru, and Diao 2000; Song and Knight 2003).

While some researchers have focused on the population that take local off-farm work or emigrate (Knight, Song, and Jia 1999; Hare 2002; Roberts 2001), this study examines the determinants of migration in rural Chinese households. Following Huffman (2001), which listed “choosing agriculture”, “migration” and “off-farm work” as employment choices for

³² Johnson (2002) projected that inter-provincial migration during the 1990s involved somewhere between 16 million and 39 million people. The number for overall internal migration including these within province is much larger, e.g. *Migration News* June 1994 issue estimated there are 100 million internal migrants (Migration News 1994). Estimates for 1995 are about 154 million rural people engaged in off-farm activities (Rozelle et al., 1999) and 120 million (Bhattacharyya and Parker 1999).

rural households, we model Chinese rural households as having three choices: staying on the farm exclusively, staying in the village but partially engaging in local off-farm activities, and at least one household member working outside home region for a certain period. Hence, we extend the job location choice to trichotomous outcomes. This provides some advantages over the typical dichotomous choice migration studies (Zhao 1999b; Zhao 2001; Yao 2001). Our panel data provides more extensive information, such as village characteristics, than that of Tuan, Somwaru, and Diao (2000). Furthermore, the usual pooled estimator (with or without household fixed effects) ignores initial conditions that seem likely to be important and to bias the estimator when ignored (Heckman, 1981). Conditioning on initial conditions and observed values of explanatory variables, the econometric model with dynamic state dependence of Wooldridge (2002b, p493) is applied. The method leads to a random effects multinomial logit model.

This chapter is organized as follows. The second section reviews the literature. The dataset and our hypothesis are described in Section 3. An econometric model is explored in Section 4 and we present the results in Section 5. The last section concludes this chapter.

Literature review

Factors that may affect the decision-making of Chinese households include both household and village attributes. Household attributes, i.e., education of the head and household size, determine the labor supply and its quality. Village attributes reveal the

information infrastructure and possibly local labor demand. Zhao (2001) concluded that migration decisions were affected by local village attributes. Characteristics of migrants' destination are important as well. However, as Zhao (2001) observed, transportation costs are not an imminent concern for the majority of migrants. Most of the rural population in China face similar choices of destinations after controlling for migrant networks. We follow the literature on rural migration to omit destination characteristics (Tuan, Somwaru, and Diao 2000; Zhao 2001). This study focuses on household and village factors but we understand that unobserved factors may play a role in the decision process. Following the usual practice, we assume that random disturbance terms catch their effects.

Education

Based on data from ten counties randomly chosen from all over China and surveyed in 1993, Parish, Zhe, and Li (1995) concluded that the returns to education remain modest in rural China as the rural labor market had begun to emerge. However, Zhao (1999b) concluded that a negative relationship exists between schooling and the probability of a family having at least one member as a migrant worker. Her results may be explained by the fact that early migrants are mostly employed in construction industry, which does not have particular requirement on worker's educational background.

However, Tuan, Somwaru, and Diao (2000) suggested that young and well-educated generation is better prepared to work outside of agriculture. They concluded that higher

education and/or secondary school training develops the skills needed for non-agricultural activities. Zhang, Huang, and Rozelle (2002) found education increases the likelihood of an individual participating in the off-farm labor force, finding a job when he/she is unemployed, and receiving higher pay. They suggested that investments in rural education are desperately needed to improve agricultural productivity and to facilitate the demographic and economic transition of rural areas, which is essential for the economic development of China.

Using data collected from farm households in a central Chinese county in 1995, Hare (2002) applied an ordered probit model to examine wage and job location outcomes of rural migrants. He found that an individual's education, especially at low levels of schooling, has a positive impact on the earning of migrants. Hare suggested that reducing legal and other institutional barriers to migration and public investment in human capital and infrastructure are desirable to achieve efficiency and equity outcomes in labor market.

Based on a 1993 sample of migrants in Shanghai, Roberts (2001) found that illiterate migrants are more likely to choose farming in rural areas of Shanghai while individuals who completed more than junior middle school were less likely to engage in farming. Rozelle et al. (1999) surveyed 200 Chinese villages and found that younger and relatively well-educated rural residents are more inclined to migrate. Knight, Song, and Jia (1999) studied migration from the perspective of the migrants, enterprises employing migrants, and government. They found that migrants deemed vocational training to be very important even if they would reap

the benefits only later. Restructuring the rural education system might stimulate the rural human capital accumulation and economic development.

Household size and social network

Tuan, Somwaru, and Diao (2000) claimed that large households are more likely to have excess labor supply and to supply one or more members as migrant workers. Zhao (2001, 1999a) found similar results and concluded that migrants are more likely to be single young males from families having more laborers, less land and fewer dependents.

Roberts (2001) found that region of origin played an important role in sorting of rural labor migrants among occupations and sectors, which suggests the importance of social network in migration. Zhao (2001) found positive and statistically significant effects of migrant networks (measured by the number of early migrants from the village in her study) on the probability of migration. Similar results were presented in Hare (2002) where the network is measured by village proportion of households with previous migration experience.

The data and econometric model

The data

The data set used in this study is based on a panel of 591 rural Chinese households. To estimate the dynamic state dependence, we need a balanced panel, and hence, 109 households are deleted to obtain a balanced panel of 482 households. We compare the

characteristics of these being deleted from and these remained in our study and find that the differences between the two group averages are not statistically significant.³³ Therefore, we do not consider sample selection bias as a serious risk in our study.

Table 4-1 summarizes household and village attributes across the five-year period. Sample mean values of most variables change very little. Education is measured as the average schooling of the labor force in the household. We use the previous year's percentage of households within a village that had some off-farm work in the previous year as a proxy of local social network that might facilitate current emigration. No obvious trend, however, exists for this index. Table 4-2 compares attributes of households making different choices. Households engaged in full time farming tend to have lower average years of schooling than other households. Mean household size is larger for those households having emigrants.

Hypotheses

Since the structural discrete choice model has been laid out for some time and most recently summarized by Wooldridge (2002a), we can use the reduced model directly without any confusion or discontinuity. We expect education, household size and social network to be important determinants of the likelihood of a rural household having emigrants. Huffman (1979) suggested that human capital is an important factor determining whether members of

³³ We find that the household size average is 4.27 for these remained and 4.13 for these dropped, the average labor input is 208.60 for these remained and 215.31 for these dropped, and the percentage of village officers for the two groups are virtually equal.

a household work at a non-farm job. The intuition is that improved human skills, e.g., formal schooling, vocational training, and experience, shift the wage offer or labor demand curve upward.³⁴ Note that we treat education as a household attribute, and the empirical measure that we use as a regressor is the average initial period value. Most rural Chinese laborers do not return to school for more education.

Hypothesis 1: *Higher levels of initial period (average years of) schooling across a household's labor increase the likelihood of a household being a source of one or more emigrant workers.*

Second, a larger household size indicates a larger household labor supply to all uses. In rural China, the family planning policy permits households to have a second child if the first-born is a girl. Also, national family planning policy is less strictly enforced in rural areas especially those that are far from the large cities and those in which there are a high percentage of ethnic minorities. These facts, along with the tradition of extended families, ensure that household size varies across China. It also varies across time due to newborns and marriages.

Hypothesis 2: *A larger household size increases the likelihood of emigrant labor.*

³⁴ Yang (1997) argued that there are collective decision-making processes, e.g., the household member with highest level of schooling may play an important role since (s)he is likely to have better chance to access and utilize information. He claimed that the appropriate measure of human capital stock for the Chinese rural household might be the highest education achieved by any household member. We use the average years of schooling of the labor force in this study instead.

In societies where schooling levels are low, social networks substitute for own-education and facilitate successful emigration for work (Zhao 2001; Taylor 2001). The share of local village households who had at least one emigrant last period is our empirical proxy for social network. Households may obtain information regarding off-farm working from the village neighborhood. Note that in constructing this ratio we do not distinguish between emigrants and workers who choose local non-farm work due to the limitation of the village level data.

Hypothesis 3: *The last period percentage of household engaged in non-farm activities in the village increases the likelihood of a household engaging in off-farm working.*

The final issue is dynamic state dependence in job-location decisions of rural Chinese households. Clearly when a household has one or members engage in non-farm work in the previous period it is more likely to do so in subsequent time periods since household members have non-farm working experience and have better access and/or utilization of relevant information.

Hypothesis 4: *The decisions in previous time period(s) influence current decisions.*

These hypotheses can be tested in our empirical model.

Econometric Model

Researchers have tended to use different econometric models to study migration and off-farm working, e.g. hazard rate model (Huffman and Feridhanusetyawan 2003), ordered

probit (Hare 2002), and two-stage estimation (Zhu 2002). The multi-nominal logit model is also widely used (e.g., Parish, Zhe, and Li 1995; Roberts 2001). With the presence of the dynamic state dependence, we extend the method of Wooldridge (2002a) to a trichotomous setting with a random effects multinomial logit model and estimate it based on a five-year panel data set.

Dynamic dependence in discrete choice model

Three approaches, dynamic programming, semi-parametric, and parametric, have been used in the literature to incorporate and estimate state dependence when fitting discrete choice model to a panel data. Rust (1987, 1997, 2000) incorporated dynamic programming into discrete choice model for panel data. He proposed a nested fixed-point algorithm to produce conditional maximum likelihood estimates and formulated it as a problem of statistical inference while explicitly accounting for unobserved heterogeneities. Rust (1997) examined the effects of the U.S. social security and Medicare insurance systems on labor supply of older males in the presence of incomplete markets for loans, annuities, and health insurance using a dynamic programming model. However, when there are multiple state variables, the computations required to fit this model are burdensome. Furthermore, the transition nature of Chinese economy raises doubts about the applicability of dynamic programming.

An alternative approach is the semi-parametric models proposed by Honoré and Kyriazidou (2000), which used an identification strategy based on the conditional MLE. They showed how to estimate the parameters in an unobserved effects logit model with lagged dependent variables and strictly exogenous explanatory variables without making distributional assumptions about the unobserved effects. This approach is consistent but does not generally converge at the usual square root of N rate and the discrete explanatory variables (e.g., time dummy variables) must be ruled out. Furthermore, it is not possible to estimate the average partial effects (Wooldridge 2002a).

A third alternative is the parametric approach discussed in Wooldridge (2002a). The primary problem faced by parametric approach is how to handle initial conditions. Wooldridge (2002a) summarized three parametric methods. The simplest is to ignore the randomness of the initial response, which in essence is an overly strong assumption that the initial response is independent of unobserved heterogeneity. A better way is to model the initial condition as random variable with a specified distribution. However, this evokes the question of “which distribution should we use”. Some authors used a steady-state distribution but it is unlikely when there is an obvious trend, especially in transitional economies. The method proposed by Wooldridge (2002a; 2002b, p493) models the distribution of unobserved effects conditional on the initial value and any exogenous explanatory variables. Note that fixed effects models which treat the individual effects as parameters to be estimated is not preferred in the dynamic panel setting (Heckman, 1981).

For a single observation, assume \mathbf{c} is the vector of random effects, \mathbf{y}_0 is the vector of initial conditions, \mathbf{z} is the vector of explanatory variables, and $\boldsymbol{\alpha}$ is the parameter vectors, Wooldridge (2002a) proposed to use a certain distribution (e.g., normal distribution) of individual effects conditioned on \mathbf{y}_0 , $h(\mathbf{c} | \mathbf{y}_0, \mathbf{z}; \boldsymbol{\alpha})$, rather than a distribution of \mathbf{y}_0 conditioned on individual effects, $f(\mathbf{y}_0 | \mathbf{c}, \mathbf{z}; \boldsymbol{\alpha})$, to catch the dependence between \mathbf{c} and \mathbf{y}_0 . Specify a density of $(\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_T)$ conditional on \mathbf{z} and \mathbf{c} , which is denoted as $f(\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_T | \mathbf{z}, \mathbf{c}; \boldsymbol{\delta})$. Then $f(\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_T | \mathbf{z}, \mathbf{c}; \boldsymbol{\delta})$ is integrated against the density $h(\mathbf{c} | \mathbf{y}_0, \mathbf{z}; \boldsymbol{\alpha})$ to obtain the conditional density of $f(\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_T)$ given \mathbf{z} and \mathbf{y}_0 , which can be used in an MLE estimation.

State dependence in trichotomous discrete choice model

Wooldridge (2002a,b) presented an example of random effects probit models with two alternatives. For a trichotomous decision problem we have the likelihood of an observation:

$$f(\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_T | \mathbf{y}_0, \mathbf{z}, \boldsymbol{\theta}) = \int_{R^3} f(\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_T | \mathbf{y}_0, \mathbf{z}, \mathbf{c}; \boldsymbol{\delta}) h(\mathbf{c} | \mathbf{y}_0, \mathbf{z}; \boldsymbol{\alpha}) d\mathbf{c}$$

where $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_T, \mathbf{y}_0$ are 3×1 vectors with values of (0, 1) indicating whether the alternative is chosen or not. \mathbf{c} is a vector of random effects for different alternatives, i.e., $\mathbf{c} = \{c_1, c_2, c_3\}$. $\boldsymbol{\theta}$ is the parameter space which includes $\boldsymbol{\alpha}$ and $\boldsymbol{\delta}$. Following Chamberlain's (1980) approach to the static probit model with unobserved effects and conditioning on \mathbf{y}_0 as in Wooldridge (2002a), we assume independence between the random components of $\mathbf{c}_1, \mathbf{c}_2$, and \mathbf{c}_3 and specify $h(\mathbf{c}_j | \mathbf{y}_0, \mathbf{z}; \boldsymbol{\alpha})$ as a normal distribution: $\mathbf{c}_j \sim N(\psi_j + \mathbf{y}_{i0} \xi_{0j} + \mathbf{z}_i \xi_j, \sigma_{aj}^2)$. It is

equivalent to $c_j = \psi_j + \mathbf{y}_{i0}\xi_{0j} + \mathbf{z}_i\xi_j + a_{ij}$, where $a_{ij} \sim N(0, \sigma_{aj}^2)$ and independent of $(\mathbf{y}_{i0}, \mathbf{z}_i)$.

We also assume that the random effects are additive to the indirect utility function, i.e.:

$$(4-1) \quad V_{ijt} = \mathbf{z}_i\delta_j + \gamma_{2j}y_{i2,t-1} + \gamma_{3j}y_{i3,t-1} + c_{ij} + \varepsilon_{ijt}$$

which can be rewritten as:

$$(4-2) \quad V_{ijt} = \mathbf{z}_i\delta_j + \gamma_{2j}y_{i2,t-1} + \gamma_{3j}y_{i3,t-1} + \psi_j + \mathbf{y}_{i0}\xi_{0j} + \mathbf{z}_i\xi_j + a_{ij} + \varepsilon_{ijt}.$$

By the usual assumption of extreme value distribution of the disturbances ε_{ijt} , we have

the conditional probability as:

$$\text{Prob}(Y_i = j | \mathbf{y}_0, \mathbf{z}, \mathbf{c}; \delta) = \frac{e^{\mathbf{z}_i\delta_j + \gamma_{2j}y_{i2,t-1} + \gamma_{3j}y_{i3,t-1} + \psi_j + \mathbf{y}_{i0}\xi_{0j} + \mathbf{z}_i\xi_j + a_{ij}}}{\sum_{j'=1}^3 e^{\mathbf{z}_i\delta_{j'} + \gamma_{2j'}y_{i2,t-1} + \gamma_{3j'}y_{i3,t-1} + \psi_{j'} + \mathbf{y}_{i0}\xi_{0j'} + \mathbf{z}_i\xi_{j'} + a_{ij'}}} \quad \text{for } j = 1, 2, 3.$$

Due to the identification problem, we are only interested in the difference of the parameter estimates, thus we may rewrite the probabilities as:

$$\begin{aligned} \text{Prob}(Y_i = 1) &= \frac{e^{a_{i1}}}{e^{a_{i1}} + \sum_{j'=2}^3 e^{\mathbf{z}_i\delta_{j'} + \gamma_{2j'}y_{i2,t-1} + \gamma_{3j'}y_{i3,t-1} + \psi_{j'} + \mathbf{y}_{i0}\xi_{0j'} + \mathbf{z}_i\xi_{j'} + a_{ij'}}} \\ \text{Prob}(Y_i = j) &= \frac{e^{\mathbf{z}_i\delta_j + \gamma_{2j}y_{i2,t-1} + \gamma_{3j}y_{i3,t-1} + \psi_j + \xi_{0j}y_{i0} + \mathbf{z}_i\xi_j + a_{ij}}}{e^{a_{i1}} + \sum_{j'=2}^3 e^{\mathbf{z}_i\delta_{j'} + \gamma_{2j'}y_{i2,t-1} + \gamma_{3j'}y_{i3,t-1} + \psi_{j'} + \mathbf{y}_{i0}\xi_{0j'} + \mathbf{z}_i\xi_{j'} + a_{ij'}}} \quad \text{for } j = 2, 3. \end{aligned}$$

We have the following likelihood function for a single observation:

$$f(\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_T | \mathbf{y}_0, \mathbf{z}, \theta) = \int \prod_{i=1}^I \prod_{j=1}^J \{[\text{Pr}(Y_{ij} = 1)]^{y_{ij}}\} \cdot h(\mathbf{c} | \mathbf{y}_0, \mathbf{z}; \alpha) d\mathbf{c}$$

Note that $\delta_j, \gamma_{ij}, \psi_j, \xi_{0j}, \xi_j$ are the difference of the original parameters and the parameters for the baseline alternative (exclusively working on farm). However, we did not change the definition of a_{ij} s for the convenience of estimation, and we go into details on this next section. We can explicitly control on observed heterogeneity, i.e., education. The

distribution of \mathbf{c}_1 , \mathbf{c}_2 , and \mathbf{c}_3 can be specified as $N(\psi_j + \mathbf{y}_{i0}\xi_{0j} + \mathbf{z}_i\xi_j + \mathbf{x}_i\pi_j, \sigma_{aj}^2)$, in which \mathbf{x} is the measure of education and π_j is the coefficient to be estimated.

This model then can be estimated as a random effects multinomial logit model. We can test the null hypothesis that there are no random effects by restricting the variances of the random components to zero.

We examine the state dependence by testing γ_2 and γ_3 jointly equal to zero. Inference on the coefficients of the exogenous variables can be tested by asymptotic t -test or likelihood ratio tests.

Note that in this dynamic panel analysis, the household-specific variables that do not vary over time cannot be included. We use the household size to represent the labor supply and the logarithm of the village non-farm working ratio as a proxy of the social network. We focus on labor supply and social network for this dynamic study. This reduces the computation burden since we need to include four yearly values for any variable in \mathbf{z}_i and estimate the coefficients for the last two alternatives. Adding one policy variable will force us to add ten unknown parameters to maximize over. Education is treated as a household characteristic not varying over time.

Let $\mathbf{z}_i = \begin{pmatrix} \text{household size} \\ \ln(\text{village non-farm household percentage}) \end{pmatrix}$ and $\delta = \begin{pmatrix} \delta_1 \\ \delta_2 \end{pmatrix}$, the hypotheses

can be re-formulized as the standardized testable hypothesis:

Hypothesis 1: $\pi_2=\pi_3=0$, which means that increasing average years of schooling has no effect on the likelihood of a household being a source of local off-farm worker or emigrant labor.

Hypothesis 2: $\delta_{12}=\delta_{13}=0$, larger household size has no effects on the likelihood of a rural household being a source of local off-farm worker or emigrant labor.

Hypothesis 3: $\delta_{22}=\delta_{23}=0$, the past participation percentage of village households in off-farm work has no effect on current off-farm work and emigration decisions.

Hypothesis 4: $\gamma_{22}=\gamma_{23}=\gamma_{32}=\gamma_{33}=0$, no state dependence exists between a household's current off-farm work and emigration response and its previous period's response..

We also can test the null hypothesis that initial conditions are not relevant to current decisions or outcomes, which is:

Hypothesis 5: $\xi_{02}=\xi_{03}=0$, initial conditions are not relevant to current decisions.

Estimation and results

Estimation

The intercepts of the indirect utility function can be written as the difference between the original random component and the disturbance of the baseline alternative, i.e., $a'_{i2} = a_{i2} - a_{i1}$, $a'_{i3} = a_{i3} - a_{i1}$. However, this means that they are no longer independent of each other. We can use SAS PROC NLMIXED to estimate the random effect multinomial logit model but the computation is a heavy burden as Malchow-Møller & Svarer (2003) showed. In this

study, we use the mixed logit code developed by Train et. al. (1999) to estimate the random effects multinomial logit model. It requires the random components to be independent from each other. Therefore, we do not normalize the random component.

Train's code produces maximum simulated likelihood estimates. Lee (1992) and Hajivassiliou and Ruud (1994) provide the asymptotic distribution of the maximum simulated likelihood estimator. The estimator is consistent and asymptotically normal under regularity conditions. Revelt and Train (1998) pointed out that the simulated probability is an unbiased estimate of the true probability. However, the logarithm of simulated probability with fixed number of repetitions is not an unbiased estimate of the logarithm of true probability. This introduces certain levels of bias, but it proved to be decreasing as we increase the number of repetitions increases (Train, et. al, 1999; MacFadden and Train 2000).

The computation of the marginal effects of this model involves the integration over the possibility density functions of extreme value distribution and normal distribution. Close form expressions of marginal effects is intractable, and hence, we turn to simulation. We are interested in the change in the predicted probability due to the change in exogenous variables, i.e., $E\{P_k(x', \hat{\beta}) - P_k(x, \hat{\beta})\}$, where $\hat{\beta}$ is the estimated parameters. x' is the original value plus an increment, i.e., one percent or one unit of the measurement. The expectation of the marginal effects can be consistently estimated by $\frac{1}{n_1} \sum_i \frac{1}{n_2} \sum_j [P_k(x', \hat{\beta}^i, \hat{c}^j) - P_k(x, \hat{\beta}^i, \hat{c}^j)]$, where $\hat{\beta}^i$ is a draw from the estimated asymptotic

distribution of $\hat{\beta}$. \hat{c}^i is a draw from the normal distribution with the parameters generated as part of $\hat{\beta}^i$.

Results

The maximum likelihood estimates of the econometric model (random effects multinomial logit) is fitted to the balanced panel of Chinese rural households and parameter estimates are reported in Table 4-3. Tests of no state dependence and no random effects are presented in Table 4-4. Simulated marginal probability change for each of the key regressors (schoolings and rate of village members' participation in off-farm are reported in Table 4-5. Additional simulations are reported in Table 4-6.

We conclude that a household is more likely to have members emigrate for work if the household has a high level of average schooling of its members. The conclusion is consistent with the theory of Huffman (1991) and findings of previous studies on rural Chinese households, e.g., Tuan, Somwaru, and Diao (2000). As to the measurement of education, while Yang (1997) and Chen, Huffman and Rozelle (2003) found that the highest education attained is better than other measures in their production studies, the average schooling measure is preferred here. The intuition is that production decisions are collectively made but the decision to pursue a non-farm job relies more on the schooling of individuals. Household head's education is least relevant here, which is supported by the likelihood values presented in Table 4-4.

Table 4-5 provides the simulated change in probability of a household member working off-farm when the village off-farm and emigration rates increase. None of social network effects is statistically significant at the 5 percent level. Some of the partial effects of social network on likelihood of off-farm work are statistically significant but with a complicated pattern. The statistical insignificance may be due to the fact that we have 29 villages but about 500 households. Hence, major intra-village variation exists in off-farm work tendencies.

A household's size does affect the job-location decisions of its members. Controlling for the effect of the previous job-location decisions of household members, increasing a household size increases the likelihood of a household's members to pursue non-farm work.

Based on the parameter estimates, we conclude that a strong dynamic state dependence exists between current period off-farm work decisions of households and their response in previous period. The intuition is that a household that had at least a member working off-farm in the local vicinity last year is more likely to have one or more members participate in off-farm work this year. Similarly, a household that has at least one member who emigrated for work last period is more likely to have a member emigrating this period. The experience and information accumulated during the previous time periods lowers search costs and raises labor demand this period. The coefficients also revealed that previous experience as an emigrant laborer increases the likelihood of both taking local off-farm job and emigration next period. We performed a test of no dynamic state dependence after

controlling for the unobserved heterogeneity and rejected this hypothesis. The sample value of the likelihood ratio test statistic of no-state dependence is 97.34, and the critical value with 4 degrees of freedom is 9.5. Also, we test the hypothesis of no random effects by restricting the variances of the random terms to being zero. We reject this hypothesis (see Table 4-4).

In Table 4-6, we present the simulated marginal probability predictions for 1999 when varying the initial and last period decision and fixing other exogenous variables at the values in the original data set. We found that it is less likely that household members work exclusively on farm this period, if they have chosen to work off-farm in the initial period or last period. Interestingly, emigrate labor seems more likely to come from those households which experienced local non-farm work earlier, which implies that the experience of local non-farm work is beneficial to later emigration for work. Households of which an initial decision was made to have one or members emigrate are more likely to take local non-farm work. They probably were not successful in emigrant work and but the experience is useful for local off-farm work.

Conclusions

Based on recent Chinese household survey data, this study examines off-farm work and emigration decisions of rural household members. We have added to the literature a dynamic three-alternative discrete choice model as well as job-location decisions of Chinese

households. We confirm that emigrants are more likely to come from households which have more available labor and better education.

We have showed that increasing a household's average schooling increases the likelihood of its members engaging in non-farm working in a dynamic setting. Members of large households are more likely to work off-farm or to emigrate. The results of random effects multinomial logit model show that strong state dependence exists. Human capital acquired during off-farm working improves the likelihood of obtaining a non-farm job opportunity later.

The policy implications from this study are obvious. First, in order to move labor out of the agricultural sector, China needs to make large investments in elementary and junior high education. Education will help rural laborers to acquire and process labor market information and raise their productivity at non-farm jobs. Second, strong state dependency shows that the experience and information are important to job location decision. The positive effects of social networks on tendency to take off-farm work and migrate suggest that better information infrastructure may be helpful for transforming rural China. Both private and public sectors can be involved to construct an improved information infrastructure. However, we admit that the network effects in our sample are not significant due to a variety of reasons, which may deserve further surveys and researches. Finally, the effect of household attributes may help agencies to target certain households in carrying out relevant projects, e.g., large households with more male labors should be kept updated with

changes of labor market. Efforts should be made to improve large households' access to information about labor market and relevant vocational training.

Further studies will be benefited by field surveys with extensive information. We might be able to explore potential nesting structure when more alternative-specific and individual-specific information are available.

Table 4-1: Mean Household / Village attributes (1995-1999)

	1995	1996	1997	1998	1999
Observations	482	482	482	482	482
	1.703	1.973	1.951	1.848	1.793
Village Income Level (1000 Yuan)	(0.685)	(0.832)	(0.780)	(1.111)	(0.844)
	3.397	3.434	3.454	3.642	3.741
Education Achieved (Year)	(1.656)	(1.624)	(1.607)	(1.583)	(1.614)
	4.317	4.297	4.295	4.218	4.201
Household Size	(1.370)	(1.371)	(1.331)	(1.392)	(1.351)
	2.285	2.293	2.257	2.286	2.190
Land Per Capita (Mu) ³⁵	(1.767)	(1.896)	(1.866)	(1.905)	(1.905)
	0.641	0.622	0.669	0.592	0.602
Village non-farm Labor Percentage	(0.337)	(0.318)	(0.304)	(0.318)	(0.316)

Table 4-2: Mean Attributes of the three household types (1996-1999)

	Full-time farming	Local off-farm work	Emigrant labor
Observations	421	1179	328
	3.010	3.617	4.108
Education Achieved (Year)	(1.640)	(1.526)	(1.657)
	4.076	4.234	4.546
Household Size	(1.581)	(1.232)	(1.454)
	3.632	2.533	2.114
Land Per Capita (Mu)	(3.625)	(1.757)	(1.179)
	2.073	1.875	1.717
Village Income Level (1000 Yuan)	(0.835)	(0.933)	(0.839)
	0.528	0.641	0.670
Village non-farm work frequency	(0.354)	(0.309)	(0.258)

³⁵ 1 Mu=1/15 Hectare

Table 4-3: Maximum likelihood random effects multinomial logit model

	Local Off-farm Work ($j=2$)	Emigrant Labor($j=3$)
Constant	-2.268*** (0.581)	-4.309*** (0.721)
State Dependence (γ_2)	1.964*** (0.387)	1.566*** (0.454)
State Dependence (γ_3)	1.156** (0.427)	2.897*** (0.460)
y_{0a}	2.115*** (0.532)	0.986** (0.584)
y_{0b}	1.275** (0.546)	1.582** (0.647)
Average years of Schoolings	0.170*** (0.059)	0.143** (0.073)
Household size	0.182 (0.151)	0.499** (0.248)
ln(Village Non-farm work ratio)	0.163 (0.185)	-0.043 (0.292)
Partial Effects		
Household size 96	-0.188 (0.139)	-0.032 (0.137)
Household size 97	-0.253 (0.189)	-0.421 (0.198)**
Household size 98	0.138 (0.120)	0.060 (0.130)
Household size 99	0.186 (0.130)	0.038 (0.131)
ln(Village Non-farm work ratio) 96	0.170 (0.300)	0.086 (0.327)
ln(Village Non-farm work ratio) 97	1.089 (0.525)**	-0.443 (0.636)
ln(Village Non-farm work ratio) 98	-1.764*** (0.509)	0.472 (0.573)
ln(Village Non-farm work ratio) 99	0.609*** (0.199)	0.562 (0.348)*
Provincial Dummy variables	Omitted	Omitted
1997 Dummy	0.089 (0.213)	0.566** (0.270)
1998 Dummy	0.335 (0.233)	-0.053 (0.321)
1999 Dummy	1.004*** (0.253)	1.386*** (0.310)
Random Term variance estimate of the baseline alternative $var(a_1)$	0.997*** (0.330)	
Random Term variance estimate $var(a_j)$	0.841*** (0.253)	0.482 (0.383)

Note: 1)* indicates the parameter is significant at 10% significance level, ** for 5% and *** for 1%. 2) Reference group is the stay exclusively on farm.

Table 4-4: Likelihood ratio tests

H_0	Hypothesis	Ln(lik)	λ	D.F	Critical Value ³⁶	Inference
$H_0: \gamma_{22}=\gamma_{23}=\gamma_{32}=\gamma_{33}=0$	No state dependence	-1131.52	97.34	4	9.49	Reject
	H1: Negation	-1082.85				
$H_0: a_1 = a_2 = a_3 = 0$	No random effects	-1090.82	15.94	3	7.81	Reject
	H1: Negation	-1082.85				
$H_0: \pi_1 = \pi_2 = 0$	No initial conditions effects	-1104.21	42.72	2	5.99	Reject
	H1: Negation	-1082.85				
$H_0:$	No education effects	-1087.30	8.90	2	5.99	Reject
	H1: Negation	-1082.85				
Education Measure:	Head's schooling	-1087.01				
	Highest person's schooling	-1084.53				
	Average schooling	-1082.85				

³⁶ The critical values correspond to 5 percent level of significance.

Table 4-5: Simulated marginal effects

+% Probability	+ 1% Avg. year of schoolings	+1 Avg. year of Schoolings	+1% ln(vlg. non-farm work ratio)	+0.01 ln(vlg. non-farm work ratio)
Change of Prob. Of choosing Alternative 1	-0.0009 (0.0003)	-0.0154 (0.0046)	-0.0001 (0.0001)	-0.0003 (0.0004)
% Change of Prob. Of choosing Alternative 1	-0.0083 (0.0026)	-0.1278 (0.0364)	-0.0011 (0.0011)	-0.0019 (0.0020)
Change of Prob. Of choosing Alternative 2	0.0000 (0.0002)	0.0002 (0.0037)	-0.0001 (0.0002)	-0.0002 (0.0003)
% Change of Prob. Of choosing Alternative 2	0.0005 (0.0030)	0.0111 (0.0506)	-0.0014 (0.0022)	-0.0026 (0.0047)
Change of Prob. Of choosing Alternative 3	0.0009 (0.0004)	0.0151 (0.0061)	0.0002 (0.0002)	0.0005 (0.0004)
% Change of Prob. Of choosing Alternative 3	0.0025 (0.0009)	0.0435 (0.0165)	0.0006 (0.0005)	0.0016 (0.0014)

Table 4-6: Simulated probabilities for 1999

(P_1, P_2, P_3)	$y_0=1$	$y_0=2$	$y_0=3$
$y_{t-1}=1$	0.60 (0.05)	0.26 (0.06)	0.35 (0.09)
	0.07 (0.03)	0.06 (0.02)	0.14 (0.04)
	0.33 (0.04)	0.68 (0.06)	0.51 (0.08)
$y_{t-1}=2$	0.28 (0.06)	0.07 (0.02)	0.11 (0.03)
	0.10 (0.04)	0.05 (0.02)	0.15 (0.04)
	0.63 (0.06)	0.88 (0.02)	0.74 (0.05)
$y_{t-1}=3$	0.32 (0.09)	0.11 (0.04)	0.13 (0.03)
	0.33 (0.09)	0.25 (0.06)	0.47 (0.06)
	0.35 (0.07)	0.64 (0.07)	0.40 (0.06)

5. CONCLUSIONS

The three essays have analyzed production efficiency, size effects and human mobility using a panel of Chinese agricultural households in the later 1990s. The first essay estimates the stochastic frontier for grain production function and examines the marginal effects of land, labor, fertilizer, and capital, as well as the effects of education and land fragmentation on efficiency. The second essay explains the empirical irregularity of inverse relationship between farm size and productivity with land quality heterogeneity. The third essay examines the role of education, household size and social network in job location decision-making of rural Chinese households.

The first essay examines technical efficiency through the framework of a translog stochastic production frontier with a behavioral inefficiency component. The model is fitted to a panel of 591 Chinese farm households. The results reveal a decreasing trend of output elasticities with respect to labor and fertilizer. We cannot reject the hypothesis of constant return to scale in household-level Chinese agricultural production over the study period. Nearly half of the farms are shown to be ninety percent or more efficient. Schooling of farm household member and farm-level specialization are shown to have positive effect on farm-level efficiency, while land fragmentation is detrimental to efficiency. The marginal effects of inefficiency terms show significant output gains by eliminating land fragmentation, improving access to education in rural area, and promoting specialization and mechanization.

The second essay examines the relationship between farm size and productivity in China's agriculture. When we utilize the egalitarian principle during land allocation in China and use imputed quality constant land rather than actual land area in the regression, the inverse relationship between farm size and productivity disappears. Hence, the strong inverse relationship that some studies have found are undoubtedly due to a number of methodological problems, including the failure to account properly for land quality differences and the method of land distribution.

The third essay analyzes the decision-making of rural Chinese households on whether to stay exclusively on farm, take local off-farm work or migrate to another region. We observe statistically significant state dependence between the current period response and decisions of the previous time period. Simulated probability changes support our hypothesis that the average schooling of household labor and household size play important roles in job-location decision-making of rural households.

The results obtained in the three essays implied two important policy implications on institutional innovation and rural education.

Based on the results of the first essay, we cannot reject the null hypothesis of constant return to scale in the stochastic grain production frontier. The second essay stresses that the inverse relationship between farm size and productivity may be explained by measurement error and market inefficiencies. Therefore, with farmers' increasing access to modern technology, China's agricultural productivity is likely not inversely related with farm size.

Institutional innovations will allow and motivate the land consolidation in rural China thus to increase the efficient usage of labor and to motivate the long-term capital investment by reducing land fragmentation. However, while China lacks a well-constructed land registration and land court system, crop insurance, and rural medical insurance, land privatization might not be a good choice (Li, Rozelle, & Brandt, 1998). Survey results of Li, Rozelle, & Brandt (1998) confirmed that rural Chinese residents are not expecting radical land institutional change but policy promoting land rental market activity are welcomed. Moderate institutional innovations, i.e., land bank, are more likely to succeed without stimulating social turmoil.

The first essay suggests that rural education improve grain production efficiency and the third essay has found that households that have higher average schooling across household laborers are likely to have a household member working out of agricultural sector. Better-educated rural laborers face an enlarged choice set of employment. They are able to efficiently allocate their effort on agriculture and off-farm working, and use other inputs, e.g., fertilizer and pesticide, more economically. Compared to urban education, rural education has received less attention in China during last two decades. Recent statistics show that the average schooling is 6.9 years in rural China and 9.4 years in urban China. Such difference is a potential source of further economic inequality. With more educated rural labor “shifting” into industrial sector or start his/her own small business with migrant salary, the income gap between rural and urban residents may be reduced. Rural education also provides incentive for institutional innovation as well. If farmers are better informed they would be better

protected from risks during production process and from the risk in product/labor markets. Rural education also helps farmers to protect themselves from misconduct of local officials and inappropriate policies with better understanding of the social infrastructure, i.e., recent cases in rural China regarding the invalidation of excess fees imposed on farmers. In summary, improving access to education in rural area will provided Chinese farmers economical, social and political benefits.

The dissertation provides policy implications on several other issues as well, e.g., the influence of specialization, household head age, mechanization, geography on technical efficiency, the egalitarian principle during land allocation in rural China. The dissertation also advances several econometric methodologies. We provide the close form of the marginal effects of exogenous efficiency explanatory variables in Battese and Coelli (1995) framework, propose new method to summarize the curvature conditions of flexible function forms, and derive the Murphy-Topel type variance estimators for linear two-step estimation. We extend the Wooldrige (2002) approach to trichotomous setting and estimate it with maximum simulated likelihood estimation.

Not surprisingly, there are issues remained unexplored. The development of China's agriculture is related to that of industrial sector, i.e., whether the industrial/service sector can absorb the labor moving out of agriculture sector, the competition over capital and other resources between agricultural sector and industrial/service sectors. The accession to WTO pronounces the openness of China's agricultural market thus international agricultural

markets would influence China's agriculture sector. Other directions need further exploration include spatial analysis of Chinese agriculture, institutional analysis on land rental markets, as well as insurance markets. To focus on the topics we have discussed, we have decided to leave these issues out of this dissertation.

6. APPENDIX

A. Marginal effects of the inefficiency explanatory variables

Given Battese and Coelli (1995) setting, remove the subscripts and note that $Y = x\beta + V - U$, where U is distributed as truncated Normal $N^+(z\delta, \sigma_u^2)$, we have the following results.

Theorem 1: The marginal effects of z is: $\frac{\partial E(Y)}{\partial z} = -\delta(1 + \Delta(\alpha))$, where $\alpha = \frac{z\delta}{\sigma_u}$, $\lambda(\alpha) = \frac{1-\phi(\alpha)}{1-\Phi(\alpha)}$, $\Delta(\alpha) = \lambda(\alpha)(\lambda(\alpha) - \alpha)$, and $\phi(\cdot)$ is the probability density function of standard normal distribution while $\Phi(\cdot)$ is the cumulative density function.

Proof: By theorem 22.2 (Greene, 2003), we have: $E(Y) = x\beta - (z\delta + \sigma_u\lambda(\alpha))$, then we can easily obtain that: $\frac{\partial E(Y)}{\partial z} = -\delta - \sigma_u \cdot \Delta(\alpha) \cdot \frac{\delta}{\sigma_u} = -\delta(1 + \Delta(\alpha))$. Q.E.D.

Theorem 2: Replacing σ_v^2 and σ_u^2 with $\sigma^2 = \sigma_v^2 + \sigma_u^2$ and $\gamma = \sigma_u^2 / (\sigma_v^2 + \sigma_u^2)$, the asymptotic variance of the vector of marginal effects $\frac{\partial E(Y)}{\partial z}$ can be estimated by: $M \text{var}(\delta, \sigma^2, \gamma) M'$

$$\text{Where } M = \begin{pmatrix} \frac{\partial m_1}{\partial \delta_1} & \cdots & \frac{\partial m_1}{\partial \delta_j} & \frac{\partial m_1}{\partial \sigma^2} & \frac{\partial m_1}{\partial \gamma} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial m_j}{\partial \delta_1} & \cdots & \frac{\partial m_j}{\partial \delta_j} & \frac{\partial m_j}{\partial \sigma^2} & \frac{\partial m_j}{\partial \gamma} \end{pmatrix},$$

$$\frac{\partial m_j}{\partial \delta_k} = \frac{\partial (\frac{\partial E(Y)}{\partial z_j})}{\partial \delta_k} = \begin{cases} -1 - \Delta(\alpha) - \delta_j \lambda((\lambda - \alpha)^2 + \frac{\delta_j z_j}{\sigma_u} \lambda(\lambda - \alpha) - 1) & \text{when } j = k \\ -\frac{\delta_j z_k}{\sigma_u} \lambda((\lambda - \alpha)^2 + \lambda(\lambda - \alpha) - 1) & \text{when } j \neq k \end{cases}$$

$$\frac{\partial m_j}{\partial \sigma_u^2} = \frac{\partial \left(\frac{\partial E(Y)}{\partial z_j} \right)}{\partial \sigma_u^2} = \frac{1}{2\sigma_u^2} \delta_j \lambda \alpha ((\lambda - \alpha)^2 + \lambda(\lambda - \alpha) - 1) \quad , \quad \frac{\partial m_j}{\partial \sigma^2} = \frac{\partial m_j}{\partial \sigma_u^2} \gamma \quad , \quad \frac{\partial m_j}{\partial \gamma} = \frac{\partial m_j}{\partial \sigma_u^2} \sigma^2 \quad , \quad \text{and}$$

$\text{var}(\delta, \sigma^2, \gamma)$ can be estimated by the asymptotic covariance estimates of $(\delta, \sigma^2, \gamma)$ obtained from maximum likelihood estimation.

Proof: The transformation implies that $\sigma_u^2 = \gamma \sigma^2$, then apply Delta's Methods (Greene, 2003, Theorem D.21A p914), and use the fact that $\Delta'(\alpha) = \lambda((\lambda - \alpha)^2 + \lambda(\lambda - \alpha) - 1)$, we can obtain Theorem 2. Q.E.D.

B. Summarizing curvature conditions for flexible functional forms

The numerical value of the curvature conditions for nonlinear functional form usually varies across individual observations and may not be readily obtainable. In the practice of production function estimation, the marginal products and/or elasticities are calculated based on the parameter estimates of the flexible functional form. In order to examine the curvature conditions, we need to summarize these information into readable dimensions. If there are a large number of observations, even in the most restrictive case, where we compare estimates of different function forms using the same dataset, it is nearly impossible to compare the curvature conditions at all sample points. To compare curvature conditions estimated with different datasets is more complicated. Ben-Akiva and Lerman (1985) discussed in the

context of discrete choice analysis how the curvature conditions at different sample points should be summarized and presented. However, in the production economics literature, this topic has received inadequate attention.

There are two approaches to summarize curvature conditions, i.e., marginal effects, for flexible function forms in the literature (Greene, 2003). The first is to calculate the curvature conditions for individual and then present the summary statistics. The intuition underlying this approach is that these statistics describe how the (aggregate) dependent variable would respond to marginal (aggregate) changes of exogenous variables. However, this may not be true as we are going to show in the next section. The second approach is to evaluate the marginal effects/elasticities) at a sample point, e.g. mean/median/geometric mean of explanatory variables. This approach has been widely used. Diewert and Wales (1987) compared three flexible functional forms by evaluating curvature conditions at the first and last sample points. Anderson & Newell (2003) proposed a method to simplify the calculation of marginal effects at a certain data point for discrete choice models. Meanwhile, we need to note that this approach hinges on strong distribution assumptions of exogenous variables.

This study proposes two methods to address the issue of summarizing curvature conditions for flexible functional forms in the practice of production function estimation. The first approach is to improve the averaging approach by incorporating a weighting scheme according to the contribution of an individual observation. The second is to strengthen the representativeness of central points. We can either use central points that is more robust to

outliers and non-normal distribution in providing a typical individual firm/household/person or to group the data points into more than one clusters and summarize the curvature conditions for each cluster. The two new methods are more intuitive and robust to outliers and abnormal distributions of exogenous variables.

The rest of this study is organized as follows. Section 2 critiques the usual practice of summarizing the curvature conditions in the context of production study. We propose our methods in Section 3. Section 4 concludes.

Critiques on the common practices

Greene (2003) stated that: “For computing marginal effects, one can evaluate the expressions at the sample means of the data or evaluate the marginal effects at every observation and use the sample average of the individual marginal effects. The functions are continuous, so Slutsky theorem applies; in large sample they will give the same answer. But that is not so in small or moderate sized samples. Current practice favors averaging the individual marginal effects when it is possible to do so.” This statement is likely to be true when evaluating the marginal effects for discrete choice problems, where the exogenous variables are less correlated to each other and can be approximated as normal distribution in large samples. We argue, however, when we are evaluating the curvature conditions of flexible functional forms in the context of production study, this result may not hold generally. There are two reasons underlie our argument: irregularity of the input quantities distribution;

and possible correlation pattern between input usage. We describe them in the following subsections.

Averaging approach

In the production economics context, input quantities in large sample are not necessarily normally distributed. U.S. Congress, Office of Technology Assessment (1984) claimed American farm size is distributed as a bi-model. In the United States, there is an increasing trend that while average farm size is enlarged, the number of small farms (of which the purpose is for entertainment rather than income-generating) increases at the same time. In developing countries, such trend exists as well due to the limitation of resource and restriction on farm size, i.e., the Household Responsibility System in China and the Land to the Tillers Program in south Asia. The land ownership structure is consisted of large number of existing small farms and there is an increasing trend of land consolidation due to the size economies and the labor migration from agricultural sector to manufacture and service sectors. Therefore large sample theory may not apply in the context of agricultural production.

Since the literature usually apply the averaging approach in summarizing marginal effects, we focus our discussion on marginal effects henceforth in this subsection. One may argue that the purpose of the averaging approach is to provide the sample mean, as well as standard deviation, of individual marginal effects. However, since the population might be

non-normally distributed, e.g., bi-model, the mean values of marginal effects are not necessarily good representative. Neither does the sample mean converge to the “true” value in large samples. The average of individual marginal effects is not equal to the change of the aggregate dependent variable with respect to marginal change of an explanatory variable. In fact, it is only a specific realization of the change of the aggregate dependent variable (e.g., the output in agricultural production function estimation) when the marginal changes of inputs of all observations in the sample are equally weighted. In the case of production function estimation, it measures the change of aggregate output when all individual observations have equal changes in input usage. However, in finite sample, firms with different size are likely to have different levels of input changes. The averaging approach fails to summarize the marginal effects of aggregate dependent variable as illustrated henceforth.

Denote output as y , input vector as x , while x_{ji} indicates the j th input of observation i . We use $(\frac{\partial y}{\partial x_j})_i$ to indicate the marginal effect evaluated at observation i and thus the averaging approach is showed as $m_j = \frac{1}{n} \sum_{i=1}^n (\frac{\partial y}{\partial x_j})_i$.

The change of aggregated output with respect to the marginal change of the aggregated input is:

$$\frac{\partial \sum_{i=1}^n y_i}{\partial \sum_{i=1}^n x_{ji}} = \lim_{\Delta(\sum_{i=1}^n x_{ji}) \rightarrow 0} \frac{\Delta(\sum_{i=1}^n y_i)}{\Delta(\sum_{i=1}^n x_{ji})} = \lim_{\sum_{i=1}^n \Delta x_{ji} \rightarrow 0} \frac{\sum_{i=1}^n \Delta y_i}{\sum_{i=1}^n \Delta x_{ji}}$$

However, since there are n independent control variables in the denominator, we *cannot* claim that:

$$\lim_{\sum_{i=1}^n \Delta x_{ji} \rightarrow 0} \frac{\sum_{i=1}^n \Delta y_i}{\sum_{i=1}^n \Delta x_{ji}} = \frac{1}{n} \sum_{i=1}^n \lim_{\Delta x_{ji} \rightarrow 0} \frac{\Delta y_i}{\Delta x_{ji}} = \frac{1}{n} \sum_{i=1}^n \left(\frac{\partial y}{\partial x_j} \right)_i.$$

Meanwhile, we can assign a weight w_i to the input change of a single observation as its contribution to the aggregate change, i.e., $w_i = \frac{\Delta x_{ji}}{\sum_{i=1}^n \Delta x_{ji}}$, then the marginal effect of the

aggregations is:

$$\frac{\partial \sum_{i=1}^n y_i}{\partial \sum_{i=1}^n x_{ji}} = \lim_{\Delta(\sum_{i=1}^n x_{ji}) \rightarrow 0} \frac{\Delta(\sum_{i=1}^n y_i)}{\Delta(\sum_{i=1}^n x_{ji})} = \lim_{\Delta t \rightarrow 0} \frac{\Delta(\sum_{i=1}^n y_i)}{\sum_{i=1}^n w_i \Delta t}$$

Assume $f(\mathbf{x}) = \sum_{i=1}^n y_i$, we know that when there is a change of Δt in the aggregate x_i , the amount of change in x_i is $w_i \Delta t$. Take first order Taylor series approximation of $f(\mathbf{x})$, the approximate change of $\Delta(\sum_{i=1}^n y_i)$ is $\sum_{i=1}^n w_i \Delta y_i$, where Δy_i is the variation of y_i due to a change of Δt in x_i , therefore:

$$\begin{aligned} \frac{\partial \sum_{i=1}^n y_i}{\partial \sum_{i=1}^n x_{ji}} &= \lim_{\Delta t \rightarrow 0} \frac{\Delta(\sum_{i=1}^n y_i)}{\sum_{i=1}^n w_i \Delta t} = \lim_{\Delta t \rightarrow 0} \frac{\sum_{i=1}^n w_i \Delta y_i}{\Delta t \sum_{i=1}^n w_i} = \lim_{\Delta t \rightarrow 0} \frac{\sum_{i=1}^n w_i \Delta y_i}{\Delta t} \\ &= \sum_{i=1}^n \lim_{\Delta t \rightarrow 0} \frac{w_i \Delta y_i}{\Delta t} = \sum_{i=1}^n w_i \lim_{\Delta t \rightarrow 0} \frac{\Delta y_i}{\Delta t} \\ &= \sum_{i=1}^n w_i \left(\frac{\partial y}{\partial x_j} \right)_i \end{aligned}$$

We can summarize this into **Theorem 3**.

Theorem 3: Assume a weight w_i denoting the contribution of input change of observation i to the aggregate changes such that $w_i = \frac{\Delta x_{ji}}{\sum_{i=1}^n \Delta x_{ji}}$, then the marginal effect of

the aggregate amounts is:
$$\frac{\partial \sum_{i=1}^n y_i}{\partial \sum_{i=1}^n x_{ji}} = \sum_{i=1}^n w_i \left(\frac{\partial y}{\partial x_j} \right)_i$$

We also have:

Lemma 1: Assuming the contributions to input change are the same across individuals,

we have that
$$\frac{\partial \sum_{i=1}^n y_i}{\partial \sum_{i=1}^n x_{ji}} = \frac{1}{n} \sum_{i=1}^n \left(\frac{\partial y}{\partial x_j} \right)_i.$$

Greene (2003) argued that by applying large sample theory the averaged marginal effect converges to the marginal effects at a representative central points. However, with the irregularity of input distribution in the context of production study the result does not hold generally. Furthermore, as implied by Lemma 1, the averaging approach is a special case of our method with equal weight for each individual observation.

Representative individual approach

The representative individual approach relies heavily on the multivariate normal distribution of inputs. There are two issues that need further exploration. First, if input distributions deviate from normal distribution, none of mean, geometric mean and median is a good representative point. Median and geometric mean are more robust to outliers but

cannot handle the bi-model case. Second, most built-in functions in econometric computation packages ignore the potential correlation between the exogenous variables and evaluate the mean/median of the inputs separately. However, in real applications, the quantities of different inputs are possibly correlated. The centroid calculated by averaging different input quantities (as calculated by many software packages) may not represent the whole sample well since it ignored the covariance structure. It is common that inputs are constrained or exhibit certain pattern of correlation between each other, especially when there are strong substitution effects. In the practice of agricultural production function estimation, taking mean/median of inputs, i.e., land, labor, fertilizer, and capital, does not guarantee to result a good representative farm. The usual representative individual may not be a realistic approach. We provide a simple example to illustrate the failure of representative point approach.

Suppose we have two inputs: labor L and capital K , without random disturbance, the production function is characterized as:

$$y = f(L, K) = \exp(a_1 \ln L + a_2 \ln K + a_3 (\ln L)^2 + a_4 (\ln K)^2 + a_5 (\ln L)(\ln K))$$

therefore we have the output-labor elasticity as:

$$\frac{\partial \ln y}{\partial \ln L} = a_1 + 2a_3 \ln L + a_5 \ln K,$$

assume a correlation pattern between L and K is:

$$2a_3 \ln L + a_5 \ln K = C, \quad C \text{ is a constant.}$$

Then we have the elasticity is a constant for all observations, but obviously we are likely to have a different result when we evaluate at the respective means of the inputs. In this case,

geometric mean can be used and obtain the correct value. However, generally, since we do not have sufficient information while we are performing the estimation, we cannot decide which central point to use. Furthermore, the non-linearity itself can be a source of the severe bias of the marginal effect estimates. Ben Akiva and Lerman (1985) discussed such bias in detail for discrete choice models.

We conclude that the averaging and representative individual approaches used in the literature may not produce appropriate curvature condition summaries.

New ways to summarize curvature conditions

In this section we propose two methods to summarize curvature conditions of flexible functional forms in the practice of production function estimation.

Method 1

A simple solution is proposed to improve over the averaging approach. We use a predetermined weight to adjust the contribution of curvature conditions. Either the ratio of individual input usage to the aggregate sample input usage or the ratio of individual output to the aggregate output is potential candidate. Yet no theory suggests a “best” weighting scheme. These firms with small input usage may have a significant marginal effect. However, giving the market imperfection in real world, their contribution to input change may be relatively small. Which force finally dominates the curvature condition change depends on which one

is greater in magnitude. Assuming the individual contributes to the aggregation with a weight equal to the ratio of its own output to the aggregate output, multiply the weight to the individual marginal effects and we obtain the marginal effects of weighted average. This approach is more intuitive and realistic than just assigning equal weight for all individuals. It can be extended to the case of elasticities as well. Since elasticities are unit-free, output percentage as a weight may be a good choice of the weighting scheme.

Method 2

To improve the representative individual approach, we need to consider how to reduce the dimension of the inputs thus to find an appropriate representative point. One way is to calculate the distance of individual observations to a reference point, i.e., the origin or the centroid, then locate the representative individual(s) using the usual mean, median, or geometric mean. When the sample is severely clustered, the curvature conditions should be evaluated at multiple representative individuals for the existing clusters, respectively.

The distance can be defined in various ways. Two commonly used distance measures are Euclid distance and Markov distance.

Euclid distance is defined as: $d(\mathbf{x}, \mathbf{x}_0) = \sqrt{(\mathbf{x} - \mathbf{x}_0)'(\mathbf{x} - \mathbf{x}_0)}$.

Markov distance is defined as: $d(\mathbf{x}, \mathbf{x}_0) = \sqrt{(\mathbf{x} - \mathbf{x}_0)' \Lambda (\mathbf{x} - \mathbf{x}_0)}$, where $\Lambda = S^{-1}$.

Markov distance is commonly used since it is invariant to the unit of the variables under study.

A recipe of locating representative individual(s) is described as follows:

Step 1: Calculate the distance of individual points to a reference point (e.g., centroid);

Step 2: Graph the histogram of the distance for the whole sample and decide whether there are clusters according to the graph (or clustering can be applied directly, then make the corresponding judgment whether the sample appears to be clustered or not);

Step 3.1: If the distribution appears to be uni-model, then simple statistical procedure can be applied to locate the representative individual;

Step 3.2: If it appears to be a clustered sample, then apply clustering algorithm, e.g., Hierarchical Clustering Methods, to group the observations into G groups, and locate the representative points within each group. The overall summary statistics can be a weighted average of curvature conditions at these points or they can be presented directly to readers since researchers may be interested in the marginal effects of different clusters.

Note that we propose to use clustering rather than the classification procedure applied in Ben-Akiva & Lerman (1986). The difference of classification and clustering is that classification “pertain to a known number of groups, and the operational objective is to assign new observations to one of these groups” while clustering is “a more primitive technique in that no assumptions are made concerning the number of groups or the group structure” (Johnson & Wichern, 2001). In most production studies, we do not have a predetermined G , thus clustering is more applicable. Though we need to set a cut-off distance

for the dendrogram (tree diagram) to decide how many clusters we keep, the number of clusters is ex post rather than ex ante in classification problem.

Another clustering method may be used is non-hierarchical clustering method, e.g., K-means algorithm. It is computationally convenient but need a predetermined number of clusters, which is usually obtained from graphical observation, or simply intuition.

Conclusion

In this study, we criticize the usual practices of summarizing the curvature conditions of flexible functional forms and propose two new methods to accomplish that goal. The new methods produce more robust and intuitive summaries of the curvature conditions. Both methods provide policy makers a better picture of how the dependent variable may response to the marginal change of explanatory variables.

Meanwhile, when applying the first approach, alternative weighting-schemes are possible with different interpretation. In clustering algorithm, not only we need to select a distance measure, but also need to choose which point the distance measure refer to. With the importance of marginal effects in inferring policy implications, these works obviously deserve further exploration.

C. Murphy-Topel type variance estimators

For equation (3-17), we consider the usual instrumental variable estimation when X_2 is the included exogenous variables and use the methodology of Murphy and Topel (1985) to derive the estimators of the covariance matrix in the following.

Note, for simplicity, we use $(\hat{\beta}', \gamma)'$ rather than $(\hat{\beta}', \gamma)'_{2SE}$ in the following.

$$\begin{bmatrix} \hat{\beta} \\ \hat{\gamma} \end{bmatrix} = \begin{bmatrix} \beta \\ \gamma \end{bmatrix} - (Z'Z)^{-1}Z'X(\hat{\theta} - \theta)\gamma + (Z'Z)^{-1}Z'u$$

Since we have only one fitted value in the second stage, the covariance estimator is:

$$\begin{aligned} \text{var} \begin{bmatrix} \hat{\beta} \\ \hat{\gamma} \end{bmatrix} &= \gamma^2 (Z'Z)^{-1}Z'X \text{var}(\hat{\theta} - \theta)X'Z(Z'Z)^{-1} + (Z'Z)^{-1}Z' \text{var}(u)Z(Z'Z)^{-1} \\ &\quad - 2\gamma(Z'Z)^{-1}Z'XCov(\hat{\theta} - \theta, Zu)(Z'Z)^{-1} \end{aligned}$$

1. No heteroscedasticity, disturbances not correlated (MT1).

If we assume that there is no correlation between the first stage random disturbance and the second stage random disturbance, we have $\text{cov}(\hat{\theta} - \theta, Zu) = 0$. The adjusted covariance matrix is:

$$\widehat{\text{var}} \begin{bmatrix} \hat{\beta} \\ \hat{\gamma} \end{bmatrix} = \hat{\sigma}_\varepsilon^2 \gamma^2 (Z'Z)^{-1}Z'P_X Z(Z'Z)^{-1} + \hat{\sigma}_u^2 (Z'Z)^{-1}$$

where $\hat{\sigma}_\varepsilon^2 = \frac{1}{n}(L - X\theta)'(L - X\theta)$, and $\frac{1}{n}(Y - X_2\hat{\beta} - \gamma L)'(Y - X_2\hat{\beta} - \gamma L)$ is a consistently estimate of σ_u^2 (Murphy and Topel, 1985).

2. *No heteroscedasticity, disturbances correlated (MT2).*

Murphy and Topel (1985) obtained an estimator of the correlation of $\text{cov}(\frac{1}{\sqrt{n}} g(\theta), \frac{1}{\sqrt{n}} Zu)$ by the law of large numbers as: $\frac{1}{n} \sum_{i=1}^n Z_i' u_i g_i(\theta)$, similarly we have the estimator of $\text{cov}(\frac{1}{\sqrt{n}}(\theta - \theta), \frac{1}{\sqrt{n}} Zu)$ as:

$$\begin{aligned} \text{Est. cov}(\frac{1}{\sqrt{n}}(\theta - \theta), \frac{1}{\sqrt{n}} Zu) &= \text{Est. cov}((X'X)^{-1}(\frac{1}{\sqrt{n}} X' \varepsilon), \frac{1}{\sqrt{n}} Zu) \\ &= (X'X)^{-1} \frac{1}{n} \sum_{i=1}^n X_i' \hat{\varepsilon}_i Z_i \hat{u}_i \end{aligned}$$

Therefore we have:

$$\begin{aligned} \widehat{\text{var}} \begin{bmatrix} \hat{\beta} \\ \hat{\gamma} \end{bmatrix} &= \hat{\sigma}_\varepsilon^2 \gamma^2 (Z'Z)^{-1} Z' P_X Z (Z'Z)^{-1} + \hat{\sigma}_u^2 (Z'Z)^{-1} \\ &\quad - 2\gamma (Z'Z)^{-1} Z' X (X'X)^{-1} \sum_{i=1}^n \{X_i' \hat{\varepsilon}_i Z_i \hat{u}_i\} (Z'Z)^{-1} \end{aligned}$$

3. *Heteroscedasticity, disturbances not correlated (MT3).*

Considering heteroscedasticity, under some fairly general conditions, White (1980) showed that the matrix: $S_0 = \frac{1}{n} \sum_{i=1}^n e_i^2 x_i x_i'$ where e_i is the i th least squares residual, is a consistent estimator of:

$$\Sigma = \frac{1}{n} \sigma^2 X' \Omega X = \frac{1}{n} \sum_{i=1}^n \sigma_i^2 x_i x_i'$$

Hence the White estimator for the first stage covariance matrix is:

$$\text{Est. Var}(\theta) = n(X'X)^{-1} S_0 (X'X)^{-1}.$$

Similarly we have the White estimator of $(Z'Z)^{-1} Z' \text{var}(u) Z (Z'Z)^{-1}$ as:

$$n(Z'Z)^{-1} S_0^Z (Z'Z)^{-1}, \text{ where } S_0^Z = \frac{1}{n} \sum_{i=1}^n u_i^2 z_i z_i'$$

$$\widehat{\text{var}} \begin{bmatrix} \widehat{\beta} \\ \widehat{\gamma} \end{bmatrix} = n\gamma^2(Z'Z)^{-1}Z'X(X'X)^{-1}S_0(X'X)^{-1}X'Z(Z'Z)^{-1} + n(Z'Z)^{-1}S_0^Z(Z'Z)^{-1}$$

The second term on the right hand side is the usual White estimator while the first term is the variability introduced after including the fitted value as a regressor.

4. Heteroscedasticity, disturbances correlated (MT4).

If there is correlation between the first stage random disturbance and the second stage random disturbance, by using the previous estimator of $\text{cov}(\frac{1}{\sqrt{n}}(\theta - \theta), \frac{1}{\sqrt{n}}Zu)$ we have:

$$\widehat{\text{var}} \begin{bmatrix} \widehat{\beta} \\ \widehat{\gamma} \end{bmatrix} = n\gamma^2(Z'Z)^{-1}Z'X(X'X)^{-1}S_0(X'X)^{-1}X'Z(Z'Z)^{-1} + n(Z'Z)^{-1}S_0^Z(Z'Z)^{-1} \\ - 2\gamma(Z'Z)^{-1}Z'X(X'X)^{-1}\sum_{i=1}^n \{X_i'\widehat{\varepsilon}_iZ_i\widehat{u}_i\}(Z'Z)^{-1}$$

7. REFERENCES

- Abdulai, A. and W. Huffman "Structural adjustment and economic efficiency of rice farmers in northern Ghana." *Economic Development and Cultural Change* 48 (April 2000): 503-20.
- Ahearn, M., J. Yee, and W. E. Huffman "R&D, productivity, and structural change in U.S. agriculture, 1960-1996," USDA, ERS, July 2002.
- Aigner, D., C. Lovell, and P. Schmidt "Formulation and estimation of stochastic frontier production function models." *Journal of Econometrics* 6 (1977): pp.21-37.
- Anderson, S. and R. G. Newell "Simplified marginal effects in discrete choice models" *Economic Letters* 81 (2003) pp.321-326
- Battese, G. and G. S. Corra "Estimation of a production frontier model: with application to the Pastoral zone of Eastern Australia." *Australian Journal of Agricultural and Resource Economics* 21 (1977): pp.169-79.
- Battese, G. and T. Coelli "A model for technical inefficiency effects in a stochastic frontier production function for panel data." *Empirical Economics* 20 (1995): pp.325-32.
- Battese, G. and T. Coelli "A stochastic frontier production function incorporating a model for technical inefficiency effects." *Working Paper Series, No. 69* Department of Econometrics, University of New England, Armidale (1993): 22.
- Ben-Akiva M., and Lerman *Discrete choice analysis: theory and application to travel demand*, Chapter 6: Aggregate forecasting techniques, The MIT Press, Cambridge MA (1985)
- Benjamin, D., L. Brandt, and J. Giles "Income persistence and the evolution of inequality in rural China." *Paper Prepared for the NEUDC Meetings*, Boston, 2001.
- Benjamin, D. "Can unobserved land quality explain the inverse productivity relationship." *Journal of Development Economics* 46 (February 1995): pp.51-84.
- Benjamin, D. and L.Brandt "Property rights, labor markets, and efficiency in a transition economy: the case of rural China." *Canadian Journal of Economics-Revue Canadienne D Economique* 35 (November 2002): pp. 689-716.

- Berry, R.A. and W. R. Cline *Agrarian structure and productivity and developing countries*. John Hopkins University Press, Baltimore, MD, 1979.
- Bhalla, S. S. and P. Roy "Mis-specification in farm productivity analysis - the role of land quality." *Oxford Economic Papers-New Series* 40 (March 1988): pp.55-73.
- Bhattacharyya, A. and E. Parker "Labor productivity and migration in Chinese agriculture - A stochastic frontier approach." *China Economic Review* 10 (1999): pp.59-74.
- Bound, J., D. A. Jaeger, and R. M. Baker "Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak." *Journal of the American Statistical Association* 90 (June 1995): pp.443-50.
- Brandt, L. "Chinese agriculture and the international economy, 1870-1930S - A Reassessment." *Explorations in Economic History* 22 (1985): pp.168-93.
- de Brauw, A., J. Huang, and S. Rozelle "Sequencing and the success of gradualism: Empirical evidence from China's agricultural reform." *Working Paper* Dept. of Agricultural and Resource Economics, Univ. of California, Davis (2001).
- Brown, L. *Who will feed China*. Earthscan Publications, 1995.
- Burgess, R. "Land and welfare: theory and evidence from China." *Working Paper* Department of Economics and STICERD, London School of Economics, London (2001).
- Carter, M. R. "Identification of the inverse relationship between farm size and productivity: An empirical analysis of peasant agricultural production." *Oxford Economic Papers* 36(1984): 131-45.
- Chamberlain, G., "Analysis of covariance with qualitative Data", *Review of Economic Studies*, Vol 47, pp.225-238, 1980
- Chayanov, A. V. "The theory of peasant economy." D. Thorner, ed. Irwin, Homewood, IL, 1926.
- Chen, J. "Readme file for RCRE data set." *Notes* University of California, Davis (2001).
- Chen, Z., W. E. Huffman "Summarizing curvature conditions for flexible function forms" *American Agricultural Economics Association Annual Meeting*, Denver, August 2004
- Chen, Z., Wallace E. Huffman & Scott Rozelle, "Technical efficiency of modern grain

production on Chinese farms: A stochastic production frontier approach." the *AAEA Annual Meeting*, Montreal, July 27-30, 2003.

- Chen, K. and C. Brown "Addressing shortcomings in the Household Responsibility System - Empirical analysis of the Two-Farmland system in Shandong province." *China Economic Review* 12 (2001): pp.280-92.
- Cheng, E.J. "Household heads, non-economic factors & grain production in China in the 1990s." *Working Paper #98/5*, Center for Asian studies, Chinese Economy Research Unit, Univ. of Adelaide (1998).
- China " Agriculture Law." *The People's Republic of China* (1993).
- Coelli, T. "A Guide to FRONTIER version 4.1: A computer program for stochastic frontier production & cost function estimation." *Working Paper 96/07* Center for Efficiency & Productivity Analysis, Univ. of New England. Australia (1996).
- Cornwell, C., P. Schmidt, and R.C. Sickles "Production frontiers with cross-sectional and time-series variation in efficiency levels." *Journal of Econometrics* 46 (October 1990): pp.185-200.
- Davidson, R. and J. G. Mackinnon *Estimation and inference in Econometrics*. Oxford University Press, 1993.
- Deininger, K. and G. Feder "Land institutions and land markets." *Handbook of Agricultural Economics Vol. 1*. B. Gardner and G. Rausser, eds., 287-331. Elsevier Science, 2001.
- Diewert W. E. and T. J. Wales (1987) "Flexible functional forms and global curvature conditions" *Econometrica* Vol. 55, No. 1 (1987) pp.43-68
- Dong, X. Y.; Dow, G. K. "Monitoring costs in Chinese agricultural teams" *Journal of Political Economy* June 1993; Vol .101(3): pp539-53
- Evenson, R. E. and D. Gollin "Assessing the impact of the Green Revolution, 1960 to 2000." *Science* May 2, 2003: pp.758-762.
- Fleisher, B. M. and Y. H. Liu "Economies of scale, plot size, human-capital, and productivity in Chinese agriculture." *Quarterly Review of Economics and Finance* 32 (1992): pp.112-23.
- Greene, W. H. "Fixed and random effects in stochastic frontier models." *Working Paper* Stern School of Business, New York University (2002).

- Greene, W. H. "Maximum likelihood estimation of econometric frontier functions." *Journal of Econometrics* 13 (1980a): pp.27-56.
- Greene, W. H. "On the estimation of a flexible frontier production model." *Journal of Econometrics* 13 (1980b): pp.101-5.
- Greene, W. H. *Econometric analysis*, Prentice Hall, Upper Saddle River NJ, 5th ed. (2003)
- Guilkey, D. K., C. Lovell, and R. C. Sickles "A comparison of the performance of three flexible functional forms." *International Economic Review* 24(1983).
- Hahn, J. and J. Hausman "A new specification test for the validity of instrumental variables." *Econometrica* 70 (January 2002): pp.163-89.
- Hajivassiliou, V. and P. Ruud, "Classical estimation methods for LDV models using simulation," *Handbook of Econometrics*, Vol. IV, R. Engle and D. McFadden, eds., Elsevier Science B.V., New York, 1994
- Hall, A. R., G. D. Rudebusch, and D. W. Wilcox "Judging instrument relevance in instrumental variables estimation." *International Economic Review* 37 (May 1996): pp.283-98.
- Hallam, A. *Size, structure, and the changing face of American agriculture*. Westview Press, Boulder, CO. (1993)
- Hare, D. "The determinants of job location and its effect on migrants' wages: evidence from rural China." *Economic Development and Cultural Change* (2002): pp.557-79.
- Hausman, J. "Specification Tests in Econometrics." *Econometrica* 46 (1978): pp.1251-71.
- Heckman, J. "The incidental parameters problem and the problem of initial conditions in estimating a discrete time-discrete data stochastic process", in *Structural Analysis of Discrete Data with Econometric Applications*, ed. C. F. Manski and D. McFadden, Cambridge, MA, MIT Press, pp.179-195, 1981
- Heltberg, R. "Rural market imperfections and the farm size-productivity relationship: Evidence from Pakistan." *World Development* 26 (October 1998): pp. 1807-26.
- Honore, B.E.; Kyriazidou, E., "Panel data discrete choice models with lagged dependent variables", *Econometrica*, Vol 68, (2000): pp839-874
- Huang, B. "Fertilizer usage in mainland China." *Agricultural Policy and Agriculture* June(2001), Taipei

- Huffman, W.E. "Allocative efficiency: The Role of Human Capital." *Quarterly Journal of Economics* 91(1977): 59-79.
- Huffman, W. "Agricultural household models: Survey & critique", *Multiple Job-Holding among Farm Families*. M.C. Hallberg, J.L. Findeis, & D.A. Lass, eds., 79-111. Ames: Iowa State University Press, 1991.
- Huffman, W. and T. Feridhanusetyawan "Migration, location-specific amenities, and fixed costs: A hazard rate analysis." *Review of Economics of the Household Under Review* (February 2003).
- Huffman, W. and R.E. Evenson "Structural and productivity change in US agriculture, 1950-1982." *Agricultural Economics* 24 (January 2001): pp.127-47.
- Huffman, W. E., and P. F. Orazem "The role of agriculture and human capital, economic growth: farmers, schooling, and health," in *Handbook of Agricultural Economic*, Vol. III, Robert Evenson, ed., Amsterdam, The Netherlands: Elsevier Science/North Holland, forthcoming.
- Huffman, W. E., "Human capital: Education and agriculture," in *Handbook of Agricultural Economics*, Vol. IA, Bruce L. Gardner and G.C. Rausser, eds., Amsterdam, The Netherlands: Elsevier Science/North Holland, 2001, pp. 333-381.
- Johnson, R. A. & D. W. Wichern, *Applied multivariate statistical analysis*, Prentice Hall, New Jersey, 5th ed., 1995
- Johnson, D. G. "Provincial migration in China in the 1990s." *China Economic Review* 117(2002).
- Kim, S. J. "Productivity effects of economics reforms in China's agriculture" Ph.D. Dissertation, 1990 Brown University
- Knight, J., L. Song, and H. Jia "Chinese rural migrants in urban enterprises: Three perspectives." *Journal of Development Studeis* 35 (1999): pp.73-104.
- Krusekopf, C. C. "Diversity in land-tenure arrangements under the Household Responsibility System in China." *China Economic Review* 13(2002): pp.297-312.
- Kumbhakar, S. C. "Production frontiers, panel data, and time-varying technical inefficiency." *Journal of Econometrics* 46 (October 1990): pp.201-11.
- Kumbhakar, S. C. and C. Lovell, *Stochastic frontier analysis*. Cambridge University Press, 2000.

- Kumbhakar, S. C., S. Ghosh, and J. T. McGuckin "A generalized production frontier approach for estimating determinants of inefficiency in United-States dairy farms." *Journal of Business & Economic Statistics* 9 (July 1991): pp.279-86.
- Lanjouw, J. "Information and the operation of markets: tests based on a general equilibrium model of land leasing in India." *Journal of Development Economics*, 60 (1999): pp.497-527
- Lee, L., "On efficiency of methods of simulated moments and Maximum Simulated Likelihood Estimation of discrete response models," *Econometric Theory*, Vol. 8, pp. 518-552, 1992.
- Lee, L. F. "A test for distributional assumptions for the stochastic frontier functions." *Journal of Econometrics* 22 (1983): pp.245-67.
- Li, G., S. Rozelle, and L. Brandt "Tenure, land rights, and farmer investment incentives in China." *Agricultural Economics* 19 (September 1998): pp.63-71.
- Lin, J. Y. "Household farm, cooperative farm and efficiency: Evidence from rural de-collectivization in China." *Economics Growth Center Work Paper No. 533* Yale University(1987).
- Lin, J. Y. "Rural reforms and agricultural productivity growth in China." *American Economic Review*, Vol. 82, 34-51 (1992)
- Lin, J. Y. "The Household Responsibility System in China's agricultural reform: A theoretical and empirical study." *Economic Development and Cultural Change* 36(1988).
- Lohmar, B., Z. Zhang, and A. Somwaru "The role of the land rental market in China's agricultural economy." *Conference Paper* (2002).
- Malchow-Møller, Nikolaj; Michael, Svarer "Estimation of the Multinomial Logit Model with random effects" *Applied Economics Letters* Vol. 10, pp.389-392 (2003)
- Mao, W. N. and W. W. Koo "Productivity growth, technological progress, and efficiency change in Chinese agriculture after rural economic reforms: A DEA approach." *China Economic Review* 8 (1997): pp.157-74.
- McFadden, D; K. Train. "Mixed MNL models for discrete response", *Journal of Applied Econometrics*, Vol. 15, No. 5, pp. 447-470, 2000.

- Meeusen, W. and J. van den Broeck "Efficiency estimation from Cobb-Douglas production functions with composed error." *International Economic Review* 18 (1977): pp.435-44.
- Migration News "China loosens internal migration controls." *Migration News* (June 1994).
- Mohapatra, S., Rozelle, S., Huang J. K. "Evolution of modes of production during development: Evidence from China." *Working paper*, Department of Agricultural & Resource Economics, Univ. of California, Davis.
- Murphy, K. M. and R. H. Topel "Estimation and inference in two-step econometric models." *Journal of Business & Economic Statistics* 3(1985).
- Nelson, C. R. and R. Startz "The distribution of the instrumental variables estimator and its t-ratio when the instrument is a poor one." *Journal of Business* 63 (January 1990): S125-S140.
- Parish, W. L., X. Zhe, and F. Li "Nonfarm work and marketization of the Chinese countryside." *China Quarterly* 10 (1995): pp.697-730.
- Park, W. "Household production behavior and income distribution in a 1980's Chinese township: A simulation analysis, Ph.D. dissertation, 1989, Brown University
- Putterman, L. and A. F. Chiacu "Elasticities and factor weights for agricultural growth accounting: A look at the data for China." *China Economic Review* 5(1994).
- Revelt D. and K. Train, "Mixed logit with repeated choices of appliance efficiency levels." *Review of Economics and Statistics*, Vol. LXXX, No. 4, 647-657, 1998
- Roberts, K. "The determinants of job choice by rural labor migrants in Shanghai." *China Economic Review* 12(2001): 15-39.
- Rozelle, S., J. E. Taylor, and A. de Brauw "Migration, remittances, and agricultural productivity in China." *American Economic Review* 89(May 1999): 287-91.
- Rust, John, "Nested fixed point algorithm documentation manual" Staff Paper, 2000
- Rust, John, "Optimal replace of GMC bus engines: An empirical model of Harold Zurche", *Econometrica*, Vol 55, No. 5, pp.999-1033, 1987
- Rust, John; Phelan, Christopher "How social security and medicare affect retirement behavior in a world of incomplete markets", *Econometrica*, Vol 65, No. 4, pp781-831, 1997

- Ruud, P. *An introduction to classical econometric theory*. Oxford University Press, 2000.
- Sen, A. K. "An aspect of Indian agriculture." *Economic Weekly* 14 (1962).
- Song, L. and J. Knight "Chinese peasant choices: Farming, rural industry or migration." *Oxford Bulletin of Economics and Statistics* Under Review (2003).
- Staiger, D. and J. H. Stock "Instrumental variables regression with weak instruments." *Econometrica* 65 (May 1997): pp.557-86.
- Stevenson, R. E. "Likelihood functions for generalized stochastic frontier estimation." *Journal of Econometrics* 13 (1980): pp.57-66.
- Taylor, J. Edward and I. Adelman, "Agricultural household models: genesis, evolution, and extensions," *Review of Economics of the Household*, 1 (Jan./April 2003): pp.33-59.
- Taylor, J. Edward "Human Capital: Migration and Rural Population Change" (with Philip L. Martin). In B. Gardener and G. Rausser, eds., *Handbook of Agricultural Economics*, Volume I. Amsterdam: Elsevier (in press, 2001).
- Tian, W. M. and G. H. Wan "Technical efficiency and its determinants in China's grain production." *Journal of Productivity Analysis* 13 (March 2000): pp.159-74.
- Train, K., D. Revelt, P. Ruud; "Mixed Logit Estimation routine for panel data" Version 1999 <http://elsa.berkeley.edu/Software/abstracts/train0296.html>
- Tuan, F., A. Somwaru, and X. Diao "Rural labor migration, characteristics, and employment patterns: A study based on China's agricultural census." *Working Paper* (Nov. 2000).
- U.S. Congress, Office of Technology Assessment: "Technology, public policy and the changing structure of American agriculture" Washington D.C., U.S. Government Printing Office, OTA-F-285 March 1986
- Wan, G. H. and E. J. Cheng "Effects of land fragmentation and returns to scale in the Chinese farming sector." *Applied Economics* 33 (February 2001): pp.183-94.
- Wang, H., and P. Schmidt "One-step and two-step estimation of the effects of exogenous variables on technical efficiency levels" *Journal of Productivity Analysis* 18 (2002): pp.129-144
- Wang, J. R., G. L. Cramer, and E. J. Wailes "Production efficiency of Chinese agriculture: Evidence from rural household survey data." *Agricultural Economics* 15 (September 1996): pp.17-28.

- Widawsky, D.; S. Rozelle; S. Jin; J. Huang "Pesticide productivity, host-plant resistance and productivity in China." *Agricultural Economics* 19 (September 1998): pp.203-17.
- Wiemer, C. "Reform and the constraints on rural industrialization in China" Working paper, 1990, University of Hawaii
- Wooldridge, J., *Econometric analysis of cross section and panel data*, Cambridge, MA, MIT Press, pp493-496, 2002a
- Wooldridge, J., "Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity " Working Paper (2002b) Department of Economics, Michigan State University, East Lansing, Michigan, 2002b
- Xu, Y. "Agricultural productivity in China." *China Economic Review* 10 (1999): pp.108-21.
- Yang, D. T. "Education and off-farm work." *Economic Development and Cultural Change* 45 (April 1997b): pp.613-32.
- Yang, D. T. "Education in production: Measuring labor quality and management." *American Journal of Agricultural Economics* 79 (August 1997a): pp.764-72.
- Yao, Y. "Egalitarian land distribution and labor migration in rural china." *Working Paper*, China Center for Economic Research, Beijing University, Beijing (October 2001).
- Zhang, L., J. Huang, and S. Rozelle "Employment, emerging labor markets, and the role of education in rural China." *China Economic Review* 13 (2002): pp.313-28.
- Zhao, Y. "Labor migration and earnings differences: The case of rural China." *Economic Development and Cultural Change* 47 (July 1999a): pp.767-82.
- Zhao, Y. "Leaving the countryside: Rural-to-urban migration decisions in China." *American Economic Review* 89 (May 1999b): pp.281-286.
- Zhao, Y. "The Role of migrant networks in labor migration: The case of China." *Working Paper* China Center for Economic Research, Beijing University (Nov. 2001).
- Zhu, N. "The impacts of income gaps on migration decisions in China." *China Economic Review* 13 (2002): pp.213-30.

8. ACKNOWLEDGMENTS

I would like to thank all those who helped in different stages of the dissertation. I am deeply thankful to my major professor, Wallace E. Huffman for his continuous guidance and support. He has been a constant source of inspiration and encouragement during my study in Iowa State. I would also like to thank Prof. Joseph Herriges, Prof. Quinn Weninger, Prof. Jinhua Zhao, and Prof. Kenneth Koehler for their indispensable service on my committee.

In addition, I would like to express my sincere thanks to Prof. Scott Rozelle for providing the data set used in this dissertation, for his insightful comments, and for his pleasant collaboration. I have been benefited from Prof. Brent Hueth's comments and his excellent lectures on agricultural contracts. I thank Prof. Tim Coelli for providing his computation package and kindly help. I also need to thank Prof. Evi Zivot, Prof. Kenneth Train for making available to public of their Gauss routines, which proved to be invaluable for this work.

I am grateful to my parents, Yongyuan Chen, and Daoyu Zhang, for their love, encourage, and belief. I thank my brother, Meng Chen, and my sister, Wei Zhang, for acting as my example at various stages of my life. I thank all my friends in Ames for sharing their joyful time with me. Especially, I like to thank Dong Yan, Biyong Xu for their continuous help on my academic and everyday life. I am also grateful to Xin Lan, Wanhui Yang, Wei Huang for their kindly help as well.

My special thanks go to Xidong Deng, who accompanied me all the five years in the United States, though mostly through telephone lines. We have been sharing our joy, hardship, curiosity, and wishes. We are expecting to turn this into a long-term binding contract soon.