Estimation of Chinese Agricultural Production Efficiencies with Panel Data

Baiding Hu and **Michael McAleer**
Department of Economics, University of Western Australia

Abstract: Fast and steady economic growth in China during the 1990s attracted much international attention. Given the scarcity of resources, it is important for economic growth to depend on production efficiency improvement to achieve sustainability. As China is the world's second largest foreign capital recipient, foreign capital plays an important role in investment. If economic growth is fuelled by investment, an exodus or a shortage of foreign capital will render growth unsustainable. However, if growth is propelled by improvements in production efficiency, it is more likely to be sustained and to withstand reduction in production input. This paper estimates production efficiency in the agricultural sector in China with a panel data set comprising 30 provinces for the seven year period, 1991-1997. A panel data model based on the Cobb-Douglas production function is used to represent the production frontier and to compute technical efficiency at the provincial level. Individual effects are tested to determine if pooled estimation is preferred to unpooled (panel) estimation. The test confirms significant differences between the provinces, and hence warrants panel data estimation. Both fixed and random effects models are estimated, with provincial technical inefficiency specified as province-specific intercept terms for the former, and regression disturbances for the latter. Although the random effects model is rejected in favour of the fixed effects model, the latter did not produce estimates with correct signs, and is rejected on economic grounds. Using the random effects model, production efficiency has increased for most provinces, but the gap between the affluent coastal region and the hinterland in the west has increased.

1. Introduction

Fast and steady economic growth in China during the 1990s attracted much international attention, with the real GDP annual growth rate from 1991 to 1997 averaging over 10%. The first half of the 1990s, namely 1991-1995, was China's eighth five-year plan. Economic development during the eighth five-year period is widely seen as the most successful in PRC history, during which the government increased its support for agriculture: government expenditure on agriculture accounted for 8.8% of the total, an increase of 0.4% compared with the seventh five-year period, and state bank loans to agriculture also rose from 144.9 billion Yuan in 1992 to 357.2 billion Yuan in 1996 in real terms. With scarce resources, economic growth depends on production efficiency improvements to achieve sustainability. As China is the world's second largest foreign capital recipient, foreign capital plays an increasingly important role in investment. If economic growth is fuelled by investment, then an exodus or a shortage of foreign capital will render growth unsustainable. However, if growth is propelled by improvements in production efficiency, it is more likely to be sustained and to withstand reductions in production input.

Carter and Zhang (1994), Wu (1995), Kalirajan et al. (1996), Mao and Koo (1997), Li and Rozelle (2000), Yao et al. (2001) and Zhang (2002) have investigated agricultural production efficiency with sample periods preceding the high profile eighth five-year period. This paper estimates production efficiency in China's agricultural sector with a panel data set comprising 30 provinces for the seven year period 1991-1997, which encompasses the eighth five-year period. A panel data model based on the Cobb-Douglas production function is used to represent the production frontier and to compute technical efficiency at the provincial level. Individual effects are tested to determine if pooled estimation is preferred to unpooled (panel) estimation. The test confirms significant differences between the provinces, and hence warrants panel data estimation. Both fixed and random effects models are estimated, with provincial technical inefficiency specified as province-specific intercept terms for the former, and regression disturbances for latter. Although the random effects model is rejected in favour of the fixed effects model, the latter did not produce estimates with correct signs, and is rejected on economic grounds. Using the random effects model, production efficiency has increased for

most provinces, but the gap between the affluent coastal region and the hinterland in the west has increased consistently.

The plan of the paper is as follows: estimation framework and data are presented in Section 2, Section 3 discusses the empirical results, and concluding remarks are given in Section 4.

2. Estimation Framework and Data

Two distinct approaches, namely econometric and data envelope analysis (DEA), exist for estimating efficiency in production (Greene, 1993). The DEA approach uses mathematical programming methods to construct production frontiers. Lovell (1993) notes that "DEA was developed in a public sector, not-for-profit environment, in which prices were suspect at best and missing at worst. Consequently, the vast majority of DEA models use quantity data only and calculate technical efficiency only, despite the fact that the procedure is easily adapted to the calculation of economic efficiency in a situation in which prices are available and reliable."

With DEA, the relationship between productivity growth and technical efficiency cannot be characterised as easily as in the econometric approach. As this paper examines an important stage in China's economic development where profit-driven private ownership and foreign ownership of economic entities were vigorously merged the econometric approach is preferred to DEA in this paper.

Technical (in)efficiency analysis is conducted within the framework of a production frontier function, which can be represented by

$$Q = f(X, \beta) \tag{1}$$

where Q denotes production output, X a vector of inputs, and β a vector of unknown parameters. For convenience, it is assumed that f takes the form of the Cobb-Douglas production function, which is linear in the logarithms of Q and X, namely

$$q = \alpha + \beta' x \tag{2}$$

where lower case letters denote logarithms. Production does not reach the frontier because

there exists technical inefficiency, ε, which has a non-zero mean to reflect the deviation of actual output from the frontier, and is assumed to enter the frontier model additively, as follows:

$$q = \alpha + \beta' x + \varepsilon \tag{3}$$

Equation (3) is deterministic as the deviation of actual from maximum output, $\alpha + \beta' x$, is the amount of inefficiency. Such inefficiency can be controlled through means such as management and the introduction of new technology.

The deterministic approach to the specification of a production frontier fails to recognise that there are uncontrollable factors, such as an unusually high number of random equipment failures, and bad weather, that can have an impact on production efficiency. In order to overcome this limitation, the stochastic approach decomposes ε in equation (3) into u and v, whereby u is the inefficiency term and is restricted to be non-positive, while v is a white noise process that signifies uncontrollable factors. The resulting production frontier function is given by

$$q = \alpha + \beta' x + u + v \tag{4}$$

The compound disturbance term, $u + v$, has an asymmetric distribution. For panel data estimation, equation (4) becomes

$$q_{it} = \alpha_i + \beta' x_{it} + v_{it} \tag{5}$$

$$q_{it} = \alpha + \beta' x_{it} + v_{it} + u_i \tag{6}$$

for the fixed and random effects specifications, respectively. In equations (5) and (6), the subscripts i and t refer to units in the panel and the time period for which an observation was taken, respectively.

The inefficiency term in both equation (5), a_i, and equation (2.6), u_i, are assumed to be time-invariant, which is rather strong in most practical circumstances. A method for introducing time-variation (Cornwell et al. (1990)) is to specify them as functions of a deterministic time trend in both the linear and quadratic forms, so that the fixed and random effects models are given by

$$a_{it} = \gamma_i^f + \delta_i^f t + \theta_i^f t^2$$

$$u_{it} = \gamma_i^r + \delta_i^r t + \theta_i^r t^2 .$$

Denote $\hat{\varepsilon}_{it}^f$ and $\hat{\varepsilon}_{it}^r$ as the estimated residuals from equations (5) and (6), respectively. The parameters $\gamma_i^f, \delta_i^f, \theta_i^f$ are estimated from the auxiliary regression of $\hat{\varepsilon}_{it}^f$ on $(1, t, t^2)$, and $\gamma_i^r, \delta_i^r, \theta_i^r$ are estimated from the auxiliary regression of $\hat{\varepsilon}_{it}^r$ on $(1, t, t^2)$. Thus, a_{it} and u_{it} are calculated as the fitted values from the regressions. The technical efficiency of unit i at time t is computed as the difference between a_{it} or u_{it} (namely, the logarithm of its technical efficiency) and $a_{i,\max}$ or $u_{i,\max}$ (namely, the logarithm of the technical efficiency of the best performance in the panel), namely

$$TE_{it} = \exp(a_{it} - a_{i,\max}) \qquad (7)$$

$$TE_{it} = \exp(\hat{\varepsilon}_{it}^r - \hat{\varepsilon}_{i,\max}^r) \qquad (8)$$

for the fixed and random effects, respectively. The unit that performs the best is the most efficient and serves as a proxy for the production frontier. A time trend is included in equations (5) and (6) to capture simple technological progress.

The data from various issues of the China Statistical Yearbook comprises agricultural production input and output data for 1991-1997 for 30 provinces, with the subscripts i and t ranging from 1 to 30 and 1 to 7, respectively. Agricultural production output is measured as gross output value, defined as the total value of products of farming, forestry, animal husbandry and fishery. Agricultural production input includes capital and labour, the standard inputs in the context of the production function. Capital takes several forms, namely: (1) land, measured in hectares; (2) machinery, measured as total mechanical power in watts; and (3) fertiliser, measured as tonnage. Labour is the total number of people engaged in agricultural production. As output is expressed in nominal prices, it is deflated using a price deflator derived from the price indices of a number of farm products.

3. Empirical Results

Equations (5) and (6) are estimated using OLS and GLS, respectively. The coefficient estimates and diagnostic statistics are presented in Table 1. Under the assumption of random effects, the intercept term in equation (6), which varies randomly between the units in the panel, is comprised of deterministic and stochastic components. In estimation, the stochastic part is absorbed into the error term. The a parameter in Table 1 is the estimated deterministic part of the intercept. Under the assumption of fixed effects, the intercept term in equation (5) is different for different units in the panel, and is computed as the regression residual evaluated at the mean values of the dependent and explanatory variables for the unit. In the empirical analysis, there is one intercept for each of the 30 provinces.

Table 1 provides the F statistic for testing group effects in (A). The null hypothesis states that the panel units are homogeneous in terms of the relationship between the dependent and independent variables. Clearly, the null is rejected, so panel data estimation is warranted. The table also provides the coefficient estimates for both the random and fixed effects models, with asymptotic standard errors in parentheses. A Hausman test is then used to test the random effects model against the fixed effects model in (B). The random effects model is rejected for large values of the test statistic, which is distributed as chi-squared under the null. For the empirical models estimated in this paper, the Hausman test rejects the random effects model in favour of the fixed effects specification.

This empirical result is in contrast with that of Wu (1995), in which the Hausman test was in favour of the random effects model. However, an examination of the estimated coefficients under the fixed effects specification in Table 1 reveals that, apart from low t-ratios for most of the estimated coefficients, the sign of the *Machine* variable is negative, which conflicts with prior expectations. Accordingly, the random effects model is selected over the fixed effects model in order to estimate technical efficiency.

The coefficient of the time trend, η, reflects the rate of technological progress. Most of the coefficients are statistically significant at the 5% level. Land has played a significantly negative role in determining agricultural output. This is

because the total land area under cultivation has been in constant decline since the late 1970s due to the use of land for construction purposes, whereas total agricultural output has been rising due to an improvement in output per unit of land. Using the estimated coefficients, technical efficiency (TE) was calculated according to equation (8) for each province. TE measures efficiency against the benchmark of best practice in the panel. The larger is the TE for a province, the smaller is the difference between the province and best practice, and hence the less is the variability. The best practice was observed in Shanghai in 1997. This is convincing, as Shanghai has always been perceived to be the most developed area in China and the year 1997 is the latest time period in the sample.

Table 2 presents the national and regional TEs over the 7-year period. The national and regional TEs were obtained by averaging the provincial estimates. There is a general consensus that the provinces in China can be classified according to their geographical locations into three groups, namely the East, Central and West. This classification also reflects differences in the levels of economic development among these regions. The provinces are heterogenous across regions and homogeneous within regions in terms of economic development. The East region, or the eastern coastal area comprising 10 provinces, including Shanghai and Jiangsu, is endowed with a greater developed infrastructure and technology, and much higher labour force skills, as compared with the other areas in China. These favourable conditions have helped to attract foreign direct investment and technology. The West region covers 12 provinces in the interior west, while the Central region includes 8 provinces in between the East and West regions.

In Table 2, the East region clearly leads the other two regions. TE in the East region rose steadily from 71% in 1991 to 76% in 1997, and the East region led the Central by 10% and the West by 20%. The East and Central regions had upward TE trends, but the West experienced a downward trend from 56% in 1991 to 50% in 1997.

For ease of analysis, the estimated TEs within each region are given in graphical form in the remainder of the paper. Figure 1 shows the movements in TE for the ten provinces in the East region. Most of the provinces in the East region generally attained a TE of between 60 to 80% throughout the seven years. Guangdong

constantly remained above 80%, reflecting in part the fact that, for the first five years of the sample period, its average growth rate of agricultural output was 5.8%, and the trade proportion of its total agriculture output rose from 67% to 72%. In the first two years, Guangdong led the rest of the provinces. This changed in the next three years when Hainan took over the leading position due to its average growth rate of 17.9% in the eighth five-year period, which was well above the national average of 12%, and due to the provincial government's efforts to develop tropical agriculture as one of its three polar industries. The continuing strong TE growth, coupled with the beginning of rural urbanisation in Shanghai, made it overtake the other provinces in 1996 and 1997. A decline in TE was recorded over the period for Beijing and Tianjin. This outcome arose as non-agricultural activities accounted for an increasingly higher proportion of overall rural activities, thereby competing for scarce resources with agriculture.

In the Central region, all the provinces exhibited smooth and generally upward changes in TE, as can be seen from Figure 2. Heilongjiang achieved a level of TE that is significantly and consistently higher than the other provinces in the region, starting at a level of 74% in 1991 and finishing at 90% in 1997. This compares with a range of 46% to 60% at the beginning of the sample period for the other provinces, and to a range of 52% to 69% at the end of the period. This empirical finding reflects the fact that Heilongjiang has always been a national base for grain, bean, sugar, livestock and milk cows throughout history. Associated with its importance in agricultural and geographical attributes, Heilongjiang has also commanded a high degree of machine usage. While Heilongjiang has been the most technically efficient province, Henan, a well-known poor province, had been at the other end of the spectrum, with its TE remaining below 50% for much of the sample period.

Not surprisingly, the West region was behind the East and Central regions in TE. Figure 3 shows that, for most of the provinces in the region, TE was between 30% and 60%. There seemed to be a tendency of a downward TE movement over the period for all the provinces except Guangxi, Gansu and Shanxi, which managed increases of 8%, 2% and 3%, respectively, over the full sample period. The biggest drop was recorded in

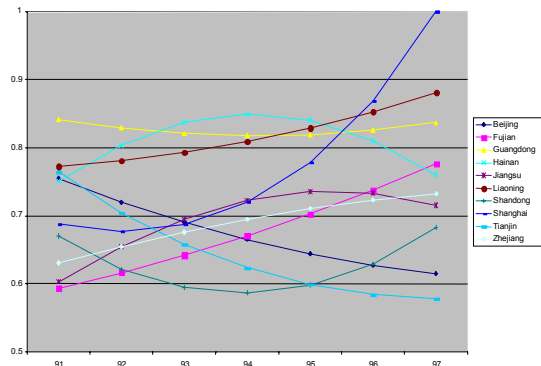Figure 1. TE by province in the East region

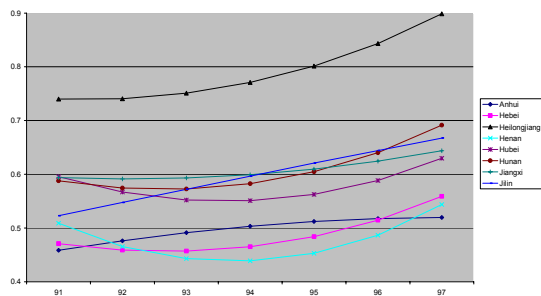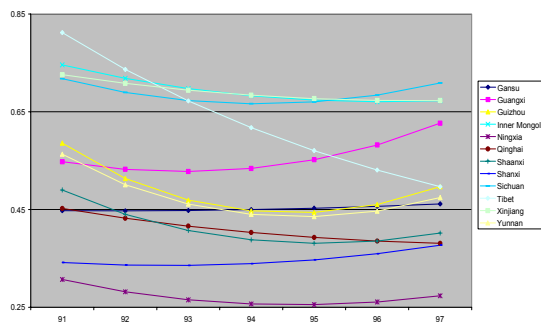Figure 2. TE by province in the Central region

Figure 3. TE by province in the West region

Tibet, where TE plummeted from 82% in 1991 to almost 50% in 1997. While the magnitude of the fall may not be accurate due to the quality of the data that are available in Tibet, the declining trend was consistent with the outcomes in the majority of the provinces in the region. The leading provinces in the region were Sichun, Xinjiang and Inner Mongolia, which have unique characteristics compared with some of the other provinces. For example, Sichun has the largest

provincial population in China, while Xinjiang and Inner Mongolia have the greatest land areas.

4. Concluding Remarks

A panel data set consisting of 30 provinces in China over a period of 7 years was used to evaluate agricultural production efficiency. Technical efficiency was defined as the difference between the observed output and the maximum possible output, namely output derived from the production frontier, which was based on the framework of the Cobb-Douglas production function. The panel nature of the data set was tested to determine if the pooled (OLS) or unpooled (panel) estimation procedure was preferred empirically. Both the random effects and fixed effects models were estimated for the Cobb-Douglas production function, where the production frontier was approximated by the best estimated performance in the panel. On the basis of the F test for group effects, panel data estimation was preferred. Although the Hausman test supported the fixed effects over the random effects model, the latter was chosen on economic grounds to estimate technical efficiency.

At the national level, technical efficiency generally rose consistently over the 7-year period. A regional breakdown indicated that the provinces in the East region were the primary source of greater technical efficiency. Regional comparisons revealed that the East region was much more efficient than the other two regions. Moreover, the efficiency gap between the East and the West regions increased during the sample. The Central region had a higher growth rate in technical efficiency towards the end of the period, which helped reduce the difference between the Central and East regions. Overall, the West region experienced a fall in efficiency over time. A graphical analysis summarised the temporal movements in technical efficiency for each of the provinces over time.

Acknowledgements

Table 1. Diagnostic Tests and Parameter Estimates

(A) Group Effects Test of H_o: Pooled versus H_1: Unpooled , $F(29,175) =14.812$					
Estimation under Random Effects					
a	η	Labour	Machine	Fertiliser	Land
0.9264	0.0122	0.2347	0.1976	0.5290	-0.1347
(1.1057)	(0.0092)	(0.0777)	(0.0979)	(0.0934)	(0.0705)
Estimation under Fixed Effects					
	0.0500	0.4447	-0.0935	0.1880	-0.0825
	(0.0131)	(0.2845)	(0.1774)	(0.1424)	(0.1057)
(B) Hausman Test of H_o: Random Effects versus H_1: Fixed Effects, $\chi^2(4) =17.70$					

Note: Asymptotic standard errors are given in parentheses.

Table 2. Estimates of Technical Efficiency at the National and Regional Levels

Year	National	East	Central	West
1991	0.61	0.71	0.56	0.56
1992	0.59	0.71	0.55	0.53
1993	0.59	0.71	0.55	0.51
1994	0.59	0.72	0.56	0.49
1995	0.59	0.73	0.58	0.49
1996	0.60	0.74	0.61	0.49
1997	0.63	0.76	0.64	0.50

References

Carter, A.C. and B. Zhang (1994), Agricultural Efficiency Gains in Centrally Planned Economies, Journal of Comparative Economics, 18 (3), 314-328.

Cornwell, C., P. Schmidt and R.C. Sickles (1990), Production Frontiers with Cross-Sectional and Time-Series Variation in Efficiency Levels, Journal of Econometrics, 46, 185-200.

Greene, W. (1993), The Econometric Approach to Efficiency Analysis, in H.O. Fried, C.A.K. Lovell and S. Schmidt (eds.), The Measurement of Productivity Efficiency: Techniques and Applications, 69-119, New York, Oxford University Press.

Li, H. and S. Rozelle (2000), Saving or Stripping Rural Industry: An Analysis of Privatization and Efficiency in China, Agricultural Economics, 23 (3), 241-252.

Lovell, C.A.K. (1993), Production frontiers and productive efficiency, in H. Fried, C.A.K. Lovell and S. Schmidt (eds), The Measurement of Productive Efficiency: Techniques and Applications, Oxford University Press, New York.

Kalirajan, K.P., M.B. Obwona and S. Zhao (1996), A Decomposition of Total Factor Productivity Growth: The Case of Chinese Agricultural Growth Before and After Reforms, American Journal of Agricultural Economics, 78 (2), 331-338.

Mao, W. and W.W. Koo (1997), Productivity Growth, Technological Progress, and Efficiency Change in Chinese Agriculture After Rural Economic Reforms: A DEA Approach, China Economic Review, 8 (2), 157-74.

State Statistical Bureau , Statistical Yearbook of China, 1992-1998, Statistical Publishing House of China, Beijing.

Wu, Y. (1995), Productivity Growth, Technological Progress, and Technical Efficiency Change in China: A Three-Sector Analysis, Journal of Comparative Economics, 21, 207-229.

Yao, S., Z. Liu and Z. Zhang (2001), Spatial Differences of Grain Production Efficiency in China, 1987-1992, Economics of Planning, 34 (1-2), 139-157.

Zhang, Y. (2002), The Impacts of Economic Reform on the Efficiency of Silviculture: A Non-parametric Approach, Environment & Development Economics, 7 (1), 107-22.