Excess Capital in Agricultural Production

Zhengfei Guan
Michigan State University
guanz@msu.edu

Subal C. Kumbhakar
State University of New York at Binghamton
kkar@binghamton.edu

Alfons Oude Lansink
Wageningen University
Alfons.OudeLansink@wur.nl

Selected Paper prepared for presentation at the American Agricultural Economics Association Annual Meeting, Portland, OR, July 29-August 1, 2007

Excess Capital in Agricultural Production

Abstract

In this article we propose a theoretical model for analyzing capital requirement in agricultural production and define excess capital thereupon. We develop a two-step method that allows endogenous regressors in the maximum likelihood estimation. The two-step procedure is also capably of recovering the parameters of time invariant variables in fixed effect models. The model and method are applied to a capital requirement study using data from cash crop farms in the Netherlands. Empirical results show that excess capital widely exists on the farm. The implications of excess capital are further demonstrated with a production frontier analysis.

Keywords: Agricultural production, capital requirement, endogeneity, excess capital, fixed effect, maximum likelihood estimation, stochastic frontier

In empirical analysis of production, factor demand functions are usually derived under the assumptions of profit maximization or cost minimization. The factor demand functions derived under these behavioral assumptions indicate how much inputs a producer *should* use in order to maximize profit or minimize cost given prices and the state of technology. In practice, however, the actual usage of inputs can be higher or lower than the optimal amount. The amount of an input used in actual production depends on various factors and can be studied directly from a technical perspective with an input requirement function. The input requirement function shows the minimum amount of an input that is required to produce a given level of output, given other inputs and the technology. This technical approach to studying input requirement is desirable for several reasons. First, empirical studies often reject the behavioral assumptions (Lin, Dean, and Moore 1974; Ray and Bhadra 1993; Driscoll et al. 1997; Tauer and Stefanides 1998), in which case, imposing behavioral assumptions to derive factor demand functions would result in biased and inconsistent parameter estimates (Pope and Chavas 1994). Second, the price information required for deriving factor demand functions is often unavailable which makes the traditional approaches based on profit maximization of cost minimization inapplicable. Third, a study of input actually used or technically needed yields insights on input requirement in production. This direct perspective of factor demand is particularly relevant for producers in making decisions regarding input use given resource endowment, production level, production technology adopted, and production environment, etc. The information on input requirement is also useful for policy making on resource use.

In the existing literature, primal studies on factor requirements include Diewert (1974), Kumbhakar and Hjalmarrson (1995, 1998), Battese, Heshmati, and Hjalmarsson (2000), Heshmati (2001), Kumbhakar, Heshmati, and Hjalmarsson (2002), El-Gamal and Inanoglu

(2005). This literature exclusively focuses on labor requirement in the production process. Until now, we are not aware of any study that studies capital requirement. As one of the major factors of production, capital presents not only an important but also a more complex and interesting case for research.

Capital is often found overused in agricultural production. In studies using both farm- and crop-level data, Guan et al. (2005, 2006) and Guan and Oude Lansink (2006) found that capital is overused on cash crop farms in the Netherlands. Using a nonparametric method, Guan and Oude Lansink (2003) concluded that Dutch agriculture is over-invested in capital and that capital is weakly disposable (i.e., it can not be disposed of costlessly when in excess). Because of weak disposability of capital, findings in these studies suggest that producers tend to have excess capital, which is either not used or not fully used in actual production. As capital investment is often an irreversible decision as suggested by Pindyck (1991), excess capital tends to persist.

The presence of excess capital means more than just a failure of profit maximization or cost minimization. It has serious implications for the econometric analysis of production. It leads to systematic measurement error if accounting data of capital stock are used in econometric modeling. For empirical econometric modeling of production, the “fixedness” of capital makes it “safe” to assume that capital is exogenous. Unfortunately, the exogeneity may not be as true as it seems, because the capital *actually used* in the production depends on the production levels. In agriculture, for example, a higher output level requires more capital for harvesting, processing, and storage of the output. This implies simultaneity of capital. In fact, the measurement error and simultaneity come hand in hand. Excess capital serves as a reservoir of capital supply when more capital is needed due to a higher yield; and vice versa, when less capital is used due to a low yield, excess capital appears. Measurement error and simultaneity of independent variables are fundamental sources of endogeneity that jeopardizes the econometric estimation, if not properly addressed.

To date, excess capital has not been explicitly explored in the literature. Somewhat related to excess capital, the concepts of "excess capacity" is proposed in the capacity utilization literature (Klein 1960; Fare, Grosskopf, and Kokkelenberg 1989; Morrison Paul 1999; Dupont et al. 2002; Kirkley, Morrison Paul, and Squires 2002, 2004; Felthoven and Morrison Paul 2004). In this literature, the "capacity output" is defined as the maximum or potential output that the existing capital stock, in conjunction with other inputs, can produce under normal working conditions. If the capacity output is not achieved, there exists "excess capacity". Notice that the excess capacity is an output-oriented concept and is an index measured with all inputs, whereas the excess capital is input-oriented and concerns with capital stock only. As a result, excess capacity is not a proper proxy for excess capital. Furthermore, the presence of excess capital per se would bias the measurement of excess capacity and capacity utilization.[1] This situation calls for a direct measurement of capital requirement to define excess capital.

We propose to use the stochastic frontier approach to measure excess capital. The theoretical basis of the frontier analysis dates back to 1950s from Koopmans (1951), Debreu (1951), and Shephard (1953). These studies construct a frontier and defined the distance relative to the frontier as an efficiency measure. The *stochastic* frontier approach originated in the works of Meeusen and van den Broek (1977) and Aigner, Lovell, and Schmidt (1977). Kumbhakar and Lovell (2000) gave a comprehensive overview of this literature. The Stochastic Frontier Analysis (SFA) has been widely applied in the economics literature, mainly to measure firms' efficiency. This approach uses maximum likelihood estimator to estimate the frontier function and the composed error terms (see, e.g. Kumbhakar and Lovell, 2000). The weakness of the maximum-likelihood based Stochastic Frontier Analysis is that when independent variables are endogenous the estimation is inconsistent, in which case the traditional Stochastic Frontier Analysis would fall flat and a solution is needed.

To address the issues raised, this study is aimed at both theoretical and methodological contributions to the literature. First, we define a theoretical framework of capital requirement in agricultural production from a primal, technical perspective. Based on the theoretical model, we further develop a concept of excess capital. In the methodological respect, we propose a two-step approach to solve the endogeneity and the resulting inconsistency problem in the maximum likelihood estimation and apply it to the stochastic frontier analysis. This study further analyzes the potential impact of the presence of excess capital on empirical analysis of production.

Capital requirement and excess capital

The capital requirement in agricultural production depends on many factors. Major factors include the type of product produced on the farm, the production level, resource endowment and technology used, natural and geographical condition, farm organizational arrangement, the demographical characteristics of the farmer, and other unobserved factors.

The type and mix of enterprises in farm production determine the type of buildings, machinery, and equipment and installations to be placed on the farm. For example, sowing machine and harvesting combines are often necessary for cereal production while other types of planting and harvesting machines are required for potato production. And for each type of product, a higher production level generally requires a higher capital stock.

The resource endowment of the farm and the technology adopted in production directly affect input-output combinations and the level of capital stock required. Strategically, if a farm has less land relative to (family) labor, the farmer may adopt a labor-intensive production technology which would require less capital for production. The technical substitution between capital and other inputs can also affect capital needs. Chemicals, for example, can be used for weed control and are substitute of mechanical weeding. In some circumstances, complementarity may exist, which means the use of one input requires the use of other inputs. An example is that fertilizer and pesticide application often requires machinery use.

Natural and geographical conditions that affect capital requirement include climate, weather, geographic and soil conditions, etc. Extreme weather conditions would require additional machinery in harvesting and drying and more storage spaces. For crop production, clay soil would require more capital use than sandy soil, as it is easier for machinery to work on loose soils than on sticky ones.

Organizational factors that affect capital requirement include land tenure regime,and use of contract work or outsourcing, etc. When certain operations, such as breeding, planting and soil disinfection are outsourced, capital stock to be maintained on the farm can be substantially reduced. The difference in land tenure may induce strategic difference in production technology and capital investment. Other than that, a leased farm is often equipped with some basic infrastructure, but this may not be reported in the bookkeeping due to the differences in bookkeeping rules. Other factors, such as the demographic and personal properties of the farm operator (e.g. education level and farming experience) may affect how efficiently the capital is used and therefore affect the capital required in the production.

As capital requirement can vary over time, some of the capital stock may not be used due to, for example, yearly crop rotations. In the meanwhile, a farmer may opt to maintain a high level of capital stock on the farm simply because he is risk averse and prefers to have more capital at his disposal, to guarantee timely sowing or harvesting in the case of adverse weather conditions. All these cases would result in excess capital on the farm. In the next section, we propose a theoretical model and a two-step method to study the capital requirement in agricultural production and measure excess capital.

Model and method

Theoretical model

The theoretical model of capital requirement can be formulated as follows:

(1) $$k = f(Y, X, O)e^{u}$$

where k is the capital stock maintained on the farm, $f(.)$ is the amount of capital required in production, which is a function of all the factors discussed. Y is a vector of outputs produced on the farm; X is a vector of inputs except capital used in the production; O represents all the other factors discussed in the preceding section. $u \geq 0$ represents excess capital. When u is zero, e^{u} is 1 and there is no excess capital. Thus, $u > 0$ measures the percentage of capital in excess. Random factors like weather and other nonsystematic elements that affect capital use are accommodated in the model by appending a random term v. Thus the stochastic capital requirement function is

(2) $$k = f(Y, X, O)e^{u+v}$$

where v can take both positive and negative values. The minimum amount of capital required to produce a certain level of Y given the technology $f(.)$, X, O and v is:

(3) $$k^{*} = \frac{k}{e^{u}} = f(Y, X, O)e^{v}$$

Thus, the excess capital can be measured from

(4) $$k_e = k - k^* = k - \frac{k}{e^u}$$

By taking logarithm of both sides of the equation (2), the stochastic capital requirement function can be rewritten as:

(5) $$\ln(k) = \ln f(Y, X, O) + u + v$$

We assume half-normal and normal distributions for *u* and *v*, respectively:

i) $u \sim N^+(0, \sigma_u^2),\ i.i.d.$

ii) $v \sim N(0, \sigma_v^2),\ i.i.d.$

iii) *u* and *v* are distributed independently of each other.

Based on these assumptions, the probability density function of the joint distribution of $e_0 = u + v$ is:

(6) $$P(e_0) = \frac{2}{\sigma} \cdot \phi\left(\frac{e_0}{\sigma}\right) \cdot \Phi\left(\frac{e_0 \lambda}{\sigma}\right)$$

where $\sigma = \sqrt{\sigma_u^2 + \sigma_v^2}$, $\lambda = \sigma_u / \sigma_v$, and $\phi(.)$and $\Phi(.)$ are the standard normal probability density and cumulative distribution functions, respectively (see Appendix A). With an explicit functional form of *f*(.), the capital requirement function can be estimated with the maximum likelihood (ML) estimation. Excess capital *u* for each observation can then be derived from the conditional expectation, $E(u \mid e_0)$ based on the conditional probability density function, $P(u \mid e_0)$.

Endogeneity, ML estimation, and a two-step method

Since Kumbhakar and Hjalmarrson (1995) first used the stochastic frontier to study labor use in the Swedish insurance offices, this approach has been used in several studies to model labor use efficiency (Kumbhakar and Hjalmarsson 1998; Battese, Heshmati, and Hjalmarsson 2000; Heshmati 2001; Kumbhakar, Heshmati, and Hjalmarsson 2002). A common assumption in this literature is that the level of output produced with labor, an explanatory variable of labor use, is considered exogenous when estimating the labor requirement function. Although this assumption may not pose problems in some special cases where exogeneity does hold, theoretically this is a strong assumption. The vast literature on production function models where output is modeled as a function of inputs (including labor) makes the endogeneity of output a legitimate issue to be addressed in the input requirement model.

Endogeneity poses a general problem to the maximum likelihood (ML) estimation in the stochastic frontier analysis. The main sources of endogeneity are heterogeneity, and simultaneity of and/or measurement errors in regressors. In least square (LS) based estimations, the endogeneity problem may be solved with instrumental variables (IV) method. The IV method, however, generally does not apply to ML estimation. For the ML-based stochastic frontier analysis (SFA), the problem of endogeneity points to the very foundation of this literature.

In this study we propose a two-step method to solve the problem. In the first step the model is estimated with least squares where endogeneity is addressed with instrumental variables. The residuals from the first step estimation are further regressed on a constant, and additional variables, if any, in the second step with ML estimation.

In the capital requirement model, we address the endogeneity of output *Y* in the first step and employ the ML estimation to derive the excess capital in the second step. For this purpose we assume a log-linear relationship between factors (*Y*, *X*), and *O*, and rewrite eq. (5) as,

(7) $$\ln(k) = f_1(Y, X; \alpha) + f_2(O; \beta) + u + v$$

where α and β are vectors of parameter to be estimated. We further rewrite the model as:

(8) $$\ln(k) = f_1(Y, X, \alpha) + e_1$$

where

(9) $$e_1 = f_2(O, \beta) + u + v$$

The first step is to estimate the model in (8). Notice that, the model (8) now has omitted variables *O*. The consequence of omitting other factors is that these factors are captured by the residuals e_1, which may cause the residuals to correlate with the regressors in the first-step model and result in biased and inconsistent estimates (see, e.g., Pindyck and Rubinfeld 1998, pp.185). This problem must be explicitly addressed, for which a robust estimation procedure must be used.

Using e_1 as dependent variable in (9), we estimate the effect of the *O* variables on capital use as well as excess capital for each observation. The model is estimated with ML estimator based on the joint distribution of the composed error $e_0 = u + v$ given in (6). As the dependent variable e_1 is not observable, it is replaced by

(10) $$\hat{e}_1 = \ln(k) - f_1(Y, X; \hat{\alpha})$$

After the estimation, the excess capital component *u* for each observation is obtained from $E(u \mid e_0)$.

Model specifications

The model in (8) is specified as:

(11) $$\ln(k_{it}) = c_0 + c_1 t + c_2 t^2 + \alpha_0 \ln(y_{it}) + \alpha_{00} \ln(y_{it})\ln(y_{it}) + \sum_j \alpha_{0j} \ln(y_{it})\ln(x_{jit})$$

$$+ \sum_j \alpha_j \ln(x_{jit}) + \frac{1}{2}\sum_j \sum_l \alpha_{jl} \ln(x_{jit})\ln(x_{lit}) + e_{1it}$$

where k denotes capital stock; the subscript i indexes individuals, and t indexes time periods; subscripts j and l index inputs. The variable y denotes the output level produced on the farm; the variable t and t^2 specify a quadratic time trend. The error term is defined as $e_{1it} = e_i + e_{it}$; e_i is the individual effect, and model (11) is a fixed effect model. This translog model is similar to production function models except that the capital stock and the output variables are switched.

The model in the second step regresses the residuals from (11) on other factors *O*, which affect capital requirements but are not included in the first step:

(12) $$e_{1it} = \beta_0 + \sum_{i=1}^{I-1} \beta_{0i} D_i + \sum_{m=1}^{M-1} \beta_{1m} DType_{mit} + \beta_2 DSoil_{it} + \beta_3 DTenu_{it} + \beta_4 DEdu_{it} + \beta_5 Size_{it} + \beta_6 Share_{it} + \beta_7 Contr_{it} + \beta_8 Age_{it} + e_{0it}$$

where $e_{0it} = u_{it} + v_{it}$, u_{it} and v_{it} are assumed to follow half-normal and normal distribution, respectively, as mentioned before;

DType, dummy variable for product types (0 for not being a particular product, 1 for yes),

DSoil, dummy soil type (0 for sandy soil, 1 for clay),

DTenu, dummy land tenure (0 for own land, 1 for lease),

DEdu, discrete education level (1 for primary school, 2 for non-agri education, 3 for vocational education in agriculture, 4 for higher education in agriculture),

Size, size of farm operation, in NGE (standardized Dutch Farm Unit),

Share is the share of non-arable farming operations on the farm in terms of size

Contr, the amount of contract work

Age, the age of farmer,

D_i is the farm dummy.

In this model, the dummy variables *DType* represent the type of major product on the farm, and its number *m* depends on the number of enterprises or products in the sample. *DSoil* is a dummy variable for soil; *DTenu* is a dummy for land tenure; *DEdu* represents the level of education. *Size* is measured in standardized Dutch Farm Unit (NGE), which is defined based on the scale, intensity and income generating ability of the farm operations (Van den Tempel and Giesen 1992, pp. 285-288). *Share* is the share of non-arable farming operations in terms of NGE for the case study of cash crop productions in the Netherlands. *Contr* denotes the amount of contract work. The variable *Age* of the farm operator is a proxy for experience and perhaps some other demographic characteristics as well.

The product dummy and the share of non-arable operation distinguish the capital requirements of different enterprises or product mixes. The farm and soil dummies capture the impacts of natural and geographical factors. Land tenure, amount of contract work, and the size of the farm represent the organizational factors. The education level and the age reflect the demographic differences of farm operators. The factors used in the second step cover both factors that affect the "standard" technical requirement of capital (e.g., from product or soil type) and those that cause additional "non-standard" or inefficient use of capital (e.g., education or experience). The unexplained part of the capital stock is due to white noise *v* and a one-sided error term *u* which captures excess capital.

The rationale underlying the split of model (11) and (12) is that the former addresses the basic input-output relations and the latter investigates the effect of farm characteristics on capital requirement. Technically, this split also ensures the recovery of parameters of time invariant farm characteristics (e.g. soil type) which would otherwise be impossible in the fixed effect model. This further contributes to the literature and justifies the two-step approach.

Data and Estimation

Data Description

The empirical study of capital requirement and excess capital is applied to data from the farm accountancy data network (FADN) of the Agricultural Economics Research Institute (LEI)) in the Netherlands. Panel data are available over the period 1990-1999 from 486 cash crop farms with a total of 2511 observations. The panel is unbalanced and farms stay in the sample for 5 years, on average.

The capital requirement function in the first-step is estimated with a single output and 5 inputs. The capital stock consists of buildings, machinery, equipment and installations. The output measured revenues from all products. The inputs included are land (x_1), labor (x_2), fertilizer (x_3), pesticide (x_4), and miscellaneous inputs (x_5). Land was measured in hectares, and labor in quality-corrected man-years. Miscellaneous inputs included seed, feed, energy, and services. The capital stock, output, fertilizer, pesticide, and miscellaneous inputs were deflated to 1990 prices (prices were obtained from the LEI/CBS[2]). Tornqvist price indices were calculated for capital and miscellaneous inputs. For the second step model, 7 product types were distinguished, viz., cereals, root crops, mix of cereals and root crops, mix of root and other crops, open-field vegetables, and mix of arable, horticultural and fruit production. The soil dummy takes the value 0 for sandy soil and 1 for clay soil. Land tenure distinguished own land and leased land[3] for the farm production. The education of farm operators was measured in 4 levels from low to high. The dummy soil type is time invariant for individual farms; the dummies for product type, land tenure and education have no or little variation over time. Other variables include the amount of contract work, the size of the farm, and the age of the farm operator. The summary statistics of non-dummy variables are presented in table 1.

Table 1. Summary Statistics of Cash Crop Farms in the Netherlands, 1990-1999

Variable	Unit	Mean	Std Dev.
Capital	thousand euro	229.33	180.57
Output	thousand euro	224.82	175.34
Land	hectare	64.92	43.63
Labor	man-year	1.92	1.18
Fertilizer	thousand euro	9.24	6.72
Pesticide	thousand euro	16.58	12.09
Misc.	thousand euro	47.46	39.01
Contract work	thousand euro	10.27	7.40
Size	Dutch farm unit (NGE)	114.93	79.55
Age	years	49.08	10.94
Share non-arable	ratio	0.09	0.12

Source: Dutch Agricultural Economics Research Institute (LEI)

Note: The statistics are per farm year, computed with 2511 observations from 486 farms; the monetary unit is in 1990 prices.

Estimation

Step 1: In the estimation of the panel data model in (11) there are three issues to be addressed: i) The heterogeneity across farms, ii) the simultaneity of output, and iii) the omission of other variables (i.e., those used in the second-step). The first issue points us to the fixed-effect estimation. Addressing simultaneity requires an instrumental variable method. To handle the omission of variables, a robust estimation must be used and the robustness to bias and inconsistency should be testable. We propose to use the generalized method of moment (GMM) in the estimation, which uses instrumental variables and provides the possibility of testing.

We use first-differencing to remove the individual effects. In the differenced model, the error term is $\Delta e_{1it} = e_{1it} - e_{1i,t-1}$. Thus it has a first-order autocorrelation structure. Moreover, it is correlated with the transformed variable $\Delta \ln(y_{it}) = \ln(y_{it}) - \ln(y_{i,t-1})$ since y is endogenous. To solve this problem we use a further lag, viz., $\ln(y_{i,t-2})$ as instrument, which implies the following moment condition:

(13) $$E[\ln(y_{i,t-2})\Delta e_{1it}] = 0$$

In principle all historical observations of $\ln(y_{it})$ prior to t-2 period may be used as instruments as well. As later periods in the panel have more historical values, more instruments are available thereby. For individual i, the matrix of the instruments is:

(14) $$Z_i = \begin{pmatrix} [\ln(y_{i0})] & 0 & \cdots & 0 \\ 0 & [\ln(y_{i0}), \ln(y_{i1})] & & 0 \\ \vdots & & \ddots & 0 \\ 0 & 0 & \cdots & [\ln(y_{i0}), \ldots, \ln(y_{i,T-2})] \end{pmatrix}$$

In the same way, instruments for the second order and cross terms involving y in the model can be constructed. The setup of instruments is similar to Arellano and Bond (1991).

In the first-step model, another variable that needs to be instrumented is miscellaneous inputs (x_5) as its components, energy consumption and services (e.g., storage and delivery), may depend on the capital stock of machinery and buildings. The instruments for the regressors associated with x_5 were set up the same way as in (14). In the unbalanced panel, farms stay in the sample for an average of 5 years. The number of farms that stay longer than 5 years decreases with the length of the panel, which means fewer observations are available for the moment conditions in later years. This may cause problems in the asymptotic approximations in GMM. Therefore, we restricted the instruments up to the 5th lag of the endogenous regressors. For the estimation we used a two-step GMM estimator. First, consistent estimates of the first-differenced residuals $\Delta\hat{e}_1$ are obtained from a preliminary consistent estimator. Next, the GMM weighting matrix is constructed as $W_N = \left[\frac{1}{N}(Z'\Delta\hat{e}_1\,\Delta\hat{e}_1{}'\,Z)\right]^{-1}$ and then used in the estimation (N is the sample size). As the weighting matrix in the two-step estimator depends on estimated parameters, the usual asymptotic approximations are less reliable, particularly in the case of heteroskedasticity, compared to the one-step estimator. Simulation studies suggest that standard errors for the two-step estimators tend to be too small (Arellano and Bond 1991; Blundell and Bond 1998). This study uses a finite-sample correction proposed by Windmeijer (2005).

The Hansen *J* test (Hansen 1982), available in the GMM framework, can be used to test the null hypothesis that instruments used are indeed valid. Not rejecting the null means consistent estimates of parameters and the residuals e_1 for the second-step model have been obtained, implying that the three problems discussed at the beginning of this section are solved.

Step 2: After correcting the endogeneity problem in model (11) with GMM, we proceed to estimate the second-step model in (12) with ML estimation based on the joint distribution of *u* and v given in Eq. (6). The log likelihood function to be maximized for a sample of N observations is:

(15) $$\ln L = constant - N \cdot \ln \sigma + \sum_i \sum_t \ln \Phi\left(\frac{e_{0it}\lambda}{\sigma}\right) - \frac{1}{2\sigma^2} \sum_i \sum_t e_{0it}^2$$

where i and t indexes individual farms and time periods, and N is 2,511 for this study.

Results

Estimation Results

The Hansen J test of the overidentified moment restrictions in the GMM estimation of model (11) produces a p-value of 0.175, larger than the 5% significance level, which means the null hypothesis is not rejected, implying that the estimation procedure is robust. The estimation results of the model are presented in table 2.

Table 2 suggests a significant technical change over time in the capital requirement function. Both the first- and second-order term of time trend are highly significant; the parameter estimates of c_1 and c_2 show that the capital requirement decreases over time (but at a diminishing rate), suggesting capital saving technical change in agriculture. The negative sign of the first-order terms of land, labor, fertilizer, and pesticide may suggest substitutability between capital and these inputs. However, the substitution effects are insignificant except for fertilizer.

Table 2. Estimation Results of Model (11)

Parameter	Description	Estimates	Corrected S.E.	p-value
c_0	constant	3.997**	0.862	0.000
c_1	t	-0.031**	0.007	0.000
c_2	t^2	0.002**	0.0008	0.002
α_0	ln(y)	0.846*	0.332	0.011
α_{00}	ln(y)ln(y)	-0.052	0.053	0.325
α_{01}	ln(y)ln(x_1)	-0.122	0.097	0.211
α_{02}	ln(y)ln(x_2)	0.036	0.058	0.530
α_{03}	ln(y) ln(x_3)	0.085	0.083	0.302
α_{04}	ln(y) ln(x_4)	0.057	0.098	0.561
α_{05}	ln(y) ln(x_5)	-0.030	0.073	0.679
α_1	ln(x_1)	-0.380	0.383	0.320
α_2	ln(x_2)	-0.255	0.254	0.315
α_3	ln(x_3)	-0.543**	0.202	0.007
α_4	ln(x_4)	-0.072	0.301	0.812
α_5	ln(x_5)	0.002	0.327	0.994
α_{11}	ln(x_1)ln(x_1)	0.084	0.083	0.309
α_{12}	ln(x_1)ln(x_2)	0.030	0.071	0.682
α_{13}	ln(x_1)ln(x_3)	0.018	0.065	0.782
α_{14}	ln(x_1)ln(x_4)	-0.005	0.078	0.947

α_{15}	$\ln(x_1)\ln(x_5)$	0.093	0.106	0.381
α_{22}	$\ln(x_2)\ln(x_2)$	0.006	0.014	0.670
α_{23}	$\ln(x_2)\ln(x_3)$	-0.045	0.036	0.213
α_{24}	$\ln(x_2)\ln(x_4)$	-0.041	0.049	0.409
α_{25}	$\ln(x_2)\ln(x_5)$	0.059	0.062	0.337
α_{33}	$\ln(x_3)\ln(x_3)$	0.003	0.012	0.779
α_{34}	$\ln(x_3)\ln(x_4)$	-0.042	0.053	0.428
α_{35}	$\ln(x_3)\ln(x_5)$	0.035	0.069	0.611
α_{44}	$\ln(x_4)\ln(x_4)$	0.041	0.042	0.330
α_{45}	$\ln(x_4)\ln(x_5)$	-0.080	0.081	0.319
α_{55}	$\ln(x_5)\ln(x_5)$	0.0004	0.049	0.993

Note: The dependent variable of the model is ln(k); k is capital stock. y denotes output; x_1 land; x_2 labor; x_3 fertilizer; x_4 pesticide; x_5 miscellaneous inputs. (*) and (**) indicate significant estimates at the 5% and 1% significance level, respectively.

A result that has important implications is that the output level has a significant impact on capital requirement, which is consistent with the *a priori* expectation. The significant effect of output on the capital requirement implies endogeneity of capital in the production process. The endogeneity of capital found in our study requires a consistent estimation procedure be used in the econometric analysis of production, such as production function or production frontier models, to which we will come back shortly.

From table 2 it is clear that most of the parameters are insignificant, suggesting that capital requirement does not respond significantly to most inputs in terms of production possibilities. Table 3 presents the estimation results of the second-stage model (12)[4].

Table 3. Estimation Results of Model (12) [a]

Parameter	Description	Estimates	Corrected S.E.	p-value
β_0	constant	-1.676 **	0.069	0.000
β_{11}	general arable	0.026	0.017	0.118
β_{12}	cereals	0.033	0.097	0.735
β_{13}	cereals & root crops [b]	0.044	0.034	0.194
β_{14}	root & other crops	0.025 **	0.005	0.000
β_{15}	open-field vegetables	0.354 **	0.023	0.000
β_{16}	mixed type [c]	0.004	0.030	0.900
β_2	DSoil	1.954 **	0.124	0.000
β_3	DTenu	-0.092 **	0.006	0.000
β_4	DEdu	-0.039	0.042	0.349
β_5	Size	0.001 **	0.0002	0.000
β_6	Share	0.085	0.087	0.330
β_7	Contr	-0.004 **	0.0009	0.000
β_8	Age	-0.002 *	0.0008	0.024

[a] Seven product types are distinguished by 6 product dummies and one default product type which is root crops.

[b] root crops include potato, sugar beet, fodder beet, and chicory.

[c] mixed type is a combination of arable, horticultural, and fruit production

(*) and (**) indicate significant estimates at the 5% and 1% significance level, respectively.

In table 3, β_1s are dummies that distinguish product types. The coefficients on “root and other crops” and “open-field vegetables” are positive and highly significant. Particularly, the size of the estimate of “open-field vegetables” is substantially larger than estimates of other product types, indicating that vegetable production are capital intensive and requires 35% more capital than the specialized root crop production (the default product type). Production on clay soils requires more capital than on sandy soils, as *a priori* expected. It is also expected that larger farms use more capital. Farms with leased land (*DTenu*) and outsourcing (*Contr*) require less capital, which is consistent with our *a priori* expectations. Furthermore, farm operators with more experience (*Age*) and higher education level (*DEdu*) use capital more efficiently and require less capital, but the education effect is insignificant.

After controlling for various factors that affect capital requirement, the unexplained capital stock is picked up in the residuals of (12), which are composed of the excess capital component *u* and a noise component *v*. A close look at the estimates of *u* for each observation shows that over 80 percent of observations in the sample have excess capital on the farm. The excess capital amounts to 21%, on average. After removing the excess capital, we find that the average capital is 195.39 thousand euros[5].

Implications of excess capital to econometric analysis of production

We used a standard translog production frontier model to demonstrate the impact of endogeneity of capital on empirical analysis of production (refer to Appendix B for details). The composed error structure in the frontier is $e_{it} = \varepsilon_{it} - \eta_i$, where ε_{it} is the noise component, and the nonnegative η_i is a farm-specific inefficiency component. We estimated the frontier model under the assumption of endogeneity and exogeneity of capital, respectively, to show the difference of the resulting efficiency measures. In both cases, the inefficiency component is allowed to be correlated with the regressors, and the frontier was estimated in two steps. First, we estimated the frontier model using the within estimator without imposing distribution assumptions, and derived consistent parameter estimates (except intercept). Under the assumption of endogeneity, instruments were used for the regressors associated with capital,[6] and no instruments were used under the assumption of exogeneity. Second, we used the residuals from the first step as dependent variable and regressed on an intercept as $\hat{e}_{it} = \text{constant} + \varepsilon_{it} - \eta_i$. In this step, the distribution assumptions were imposed on the composed error term and the ML estimator was used.

Results showed that farms' production efficiencies differ significantly under different assumptions and scenarios. When endogeneity is assumed, the technical efficiency (TE_{en}) of farms has a mean of 0.43. When exogeneity is assumed, however, the mean value of efficiency (TE_{ex}) is changed to 0.46.[7] The Wilcoxon signed rank test strongly rejects the null hypothesis of no difference between TE_{en} and TE_{ex} (p-value 0.000). The efficiency measures differ greatly for individual farms, as indicated by the low correlation coefficient, 0.63, between TE_{en} and TE_{ex}. The evidence from our study suggests that not recognizing the measurement error and simultaneity of capital could lead to systematically biased estimates of efficiency measures in the empirical study.

Concluding Remarks

This study proposed a theoretical framework for analyzing capital requirement in agricultural production and defined excess capital thereupon. The input-based framework also provides a primal approach compared to the output-oriented capacity utilization measurement in the literature. Methodologically, this study developed a two-step procedure that allows endogenous variables in the maximum likelihood estimation and apply it to stochastic frontier analysis. The two-step method is also capable of recovering the parameters of time invariant variables in the fixed effect models. The theoretical model and the methodology developed are applied to the capital requirement study of the cash crop farms in the Netherlands. Empirical results suggested that excess capital, 21% on average, widely exists on the farm.

This study argues that over-investment and “fixedness” of capital result in excess capital, which would cause two problems in the econometric analysis of production: measurement error and simultaneity of capital. The implications were demonstrated with a production frontier analysis. In addition to inconsistent estimates of slope parameters of production models, treating capital as exogenous resulted in overestimation of production efficiencies.

Footnotes

[1] Using accounting data of capital stock to model excess capacity would create a paradox such that "excess capacity" is defined without recognizing "excess capital".

[2] CBS denotes Central Bureau of Statistics in the Netherlands.

[3] when i) more than 2/3 of the land is owned by the farm operator, or ii) more than 1/3 is own land and the value of affiliated buildings on the land exceeds 9075 euro (20,000 guilders), the tenure is recorded as own land in the accounting system, otherwise recorded as leased land.

[4] For space considerations, the individual effects (i.e., parameter estimates of farm dummies) are not presented.

[5] Using the average for comparison, the capital is over-reported by 17.4%.

[6] The first- and second-order terms of inputs x_1 to x_5 (including cross terms), and all the factors included in model (12) were used as instruments for regressors associated with capital. For the estimation procedure for fixed-effect model with endogenous variables, refer to Baltagi (2005, p.114).

[7] The average efficiency is 0.70 if random effects are assumed for η_i and ML estimator is used directly in a single step.

References

Arellano, M., and S. Bond. 1991. "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations." *Review of Economic Studies* 58: 277-297.

Baltagi, B.H. 2005. *Econometric Analysis of Panel Data*. 3rd eds. Chichester, UK: John Wiley & Sons, Ltd.

Battese, G.E., A. Heshmati, and L. Hjalmarsson. 2000. "Efficiency of Labor Use in the Swedish Banking Industry: A Stochastic Frontier Approach." *Empirical Economics* 25: 623-640.

Blundell, R.W., and S.R Bond. 1998. "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models." *Journal of Econometrics* 87: 115-143.

Cornes, R. 1992. *Duality and Modern Economics*. Cambridge, UK: Cambridge University Press.

Debreu, G. 1951. "The Coefficient of Resource Utilization." *Econometrica* 19: 273-292.

Diewert, W.E. 1974. "Functional Forms for Revenue and Factor Requirements Functions." *International Economic Review*15: 119-130.

Driscoll, P.J., S.M. Kambhampaty, and W.D. Purcell. 1997. "Nonparametric Tests of Profit Maximization in Oligopoly with Application to the Beef Packing Industry." *American Journal of Agricultural Economics* 79: 872-879.

Dupont, D.P., R.Q. Grafton, J. Kirkley, and D. Squires. 2002. "Capacity Utilization Measures and Excess Capacity in Multi-Product Privatized Fisheries." *Resource and Energy Economics* 24: 193-210.

El-Gamal, M.A., and H. Inanoglu. 2005. "Inefficiency and Heterogeneity in Turkish Banking: 1990-2000. *Journal of Applied Econometrics* 20: 641-664.

Fare, R., S. Grosskopf, and E. Kokkelenberg. 1989. "Measuring Plant Capacity, Utilization and Technical Change: A Nonparametric Approach." *International Economic Review* 30: 655-666.

Felthoven, R.G., and C.J. Morrison Paul. 2004. "Multi-output, Nonfrontier Primal Measures of Capacity and Capacity Utilization." *American Journal of Agricultural Economics* 86: 619-633.

Guan, Z., and A. Oude Lansink. 2003. "Input Disposibility and Efficiency in Dutch Arable Farming." *Journal of Agricultural Economics* 54: 467-78.

Guan, Z., A. Oude Lansink, A. Wossink, and R. Huirne. 2005. "Damage Abating Inputs: A Comparison of Conventional and Organic Farming Systems." *European Review of Agricultural Economics* 32: 167-189.

Guan, Z., A. Oude Lansink, M. van Ittersum, and A. Wossink. 2006. "Integrating Agronomic Principles into Production Function Specification: A Dichotomy of Growth Inputs and Facilitating Inputs." *American Journal of Agricultural Economics* 88: 203-214.

Guan, Z., and A. Oude Lansink. 2006. "The source of productivity growth in Dutch agriculture: A perspective from finance." *American Journal of Agricultural Economics* 88: 644-656.

Hansen, L.P. 1982. "Large Sample Properties of Generalized Method of Moments Estimators." *Econometrica* 50: 1029-1054.

Heshmati, A. 2001. "Labor Demand and Efficiency in Swedish Savings Banks." *Applied Financial Economics* 11: 423-433.

Kirkley, J., C. J. Morrison Paul, and D. Squires. 2002. "Capacity and capacity utilization in common-pool resouce industries: Definition, Measurement, and a Comparison of Approaches." *Environmental and Resource Economics* 22: 71-97.

Kirkley, J., C.J. Morrison Paul, and D. Squires. 2004. "Deterministic and Stochastic Capacity Estimation for Fishery Capacity Reduction." *Marine Resource Economics* 19: 271-294.

Klein, L.R. 1960. "Some Theoretical Issues in the Measurement of Capacity." *Econometrica* 28: 272-286.

Koopmans, T.C. 1951. "An Analysis of Production as an Efficient Combination of Activities." in T.C. Koopmans, ed., *Activity Analysis of Production and Allocation*, Cowles Commission for Research in Economics, Monograph No. 13. New York: Wiley.

Kumbhakar, S.C., and C.A.K. Lovell. 2000. *Stochastic Frontier Analysis.* Cambridge, UK: Cambridge University Press.

Kumbhakar, S.C., and L. Hjalmarsson. 1995. "Labor-Use Efficiency in Swedish Social Insurance Offices." Journal of Applied Econometrics 10: 33-47.

Kumbhakar, S.C., and L. Hjalmarsson. 1998. "Relative Performance of Public and Private Ownership under Yardstick Competition: Electricity Retail Distribution. *European Economic Review* 42: 97-122.

Kumbhakar, S.C., A. Heshmati, and L. Hjalmarsson. 2002. " How Fast Do Banks Adjust? A Dynamic Model of Labor-Use with an Application to Swedish Banks." Journal of Productivity Analysis 18: 79-102.

Lin, W., G.W. Dean, and C.V. Moore. 1974. "An Empirical Test of Utility vs. Profit Maximization in Agricultural Production." *American Journal of Agricultural Economics* 56: 497-508.

Morrison Paul, C.J. 1999. *Cost Structure and the Measurement of Economic Performance: Productivity, Uitilization, Cost Structure, and Related Performance Indicators*. Boston, Massachusetts: Kluwer Academic Publishers.

Pindyck, R.S. 1991. "Irreversibility, Uncertainty and Investment." *Journal of Economic Literature* 29: 1110-1148.

Pindyck, R.S., and D.L. Rubinfeld. 1998. *Econometric Models and Economic Forecasts*, 4th ed. Boston, Massachusetts: Irwin/McGraw-Hill.

Pope, R.D., and J.P. Chavas. 1994. "Cost Functions Under Production Uncertainty." *American Journal of Agricultural Economics* 76: 196-204.

Ray, S.C., and D. Bhadra. 1993. "Nonparametric Tests of Cost Minimizing Behavior: A Study of Indian Farms." *American Journal of Agricultural Economics* 75: 990-999.

Shephard, R.W. 1953. *Cost and Production Functions*. Princeton: Princeton University Press.

Tauer, L, and Z. Stefanides. "Success in Maximizing Profits and Reasons for Profit Deviation on Dairy Farms." *Applied Economics* 30: 151-156.

Van den Tempel, F.C.A., and G.W.J. Giesen. 1992. *Agrarische Bedrijfseconomie*. Culemborg, the Netherlands: Educaboek BV.

Windmeijer, F. 2005. "A finite sample correction for the variance of linear efficient two-step GMM estimators." *Journal of Econometrics* 126: 25-51.

Appendix A:

The Density Function of Normal – Half Normal Joint Distribution

The composed error term is $e_0 = u + v$. The nonnegative term *u* follows a positive half normal distribution and its density function is

(A1) $$P(u) = \frac{2}{\sqrt{2\pi}\sigma_u} \cdot \exp\left(-\frac{u^2}{2\sigma_u^2}\right)$$

The noise component *v* follows a normal distribution and its density function is:

(A2) $$P(v) = \frac{1}{\sqrt{2\pi}\sigma_v} \cdot \exp\left(-\frac{v^2}{2\sigma_v^2}\right)$$

The joint density function of *u* and *v* is:

(A3) $$P(u,v) = \frac{2}{2\pi\sigma_u\sigma_v} \cdot \exp\left(-\frac{u^2}{2\sigma_u^2} - \frac{v^2}{2\sigma_v^2}\right)$$

Replacing *v* with $v = e_0 - u$, the joint density function of *u* and e_0 is:

(A4) $$P(u,e_0) = \frac{2}{2\pi\sigma_u\sigma_v} \cdot \exp\left(-\frac{u^2}{2\sigma_u^2} - \frac{(e_0 - u)^2}{2\sigma_v^2}\right)$$

Then the marginal density function of $e_0 = u + v$ is:

(A5) $$\begin{aligned} P(e_0) &= \int_0^\infty P(u,e_0)du \\ &= \int_0^\infty \frac{2}{2\pi\sigma_u\sigma_v} \cdot \exp\left(-\frac{u^2}{2\sigma_u^2} - \frac{(e_0 - u)^2}{2\sigma_v^2}\right)du \\ &= \frac{2}{\sqrt{2\pi}\sigma}\left[1 - \Phi\left(-\frac{e_0\lambda}{\sigma}\right)\right] \cdot \exp\left(-\frac{e_0{}^2}{2\sigma^2}\right) \\ &= \frac{2}{\sigma} \cdot \phi\left(\frac{e_0}{\sigma}\right) \cdot \Phi\left(\frac{e_0\lambda}{\sigma}\right) \end{aligned}$$

where $\sigma = \sqrt{\sigma_u^2 + \sigma_v^2}$, $\lambda = \sigma_u / \sigma_v$, and $\phi(.)$and $\Phi(.)$ are the standard normal probability density and cumulative distribution functions, respectively.

Appendix B:

The Specification of the Translog Production Frontier

The production frontier model is specified as:

$$\text{(B1)} \quad \ln(y_{it}) = c + \sum_{t=91}^{99} c_t D_t + \delta R_{it} + \gamma_0 \ln(k_{it}) + \gamma_{00} \ln(k_{it})\ln(k_{it}) + \sum_{j=1}^{5} \gamma_{0j} \ln(k_{it})\ln(x_{jit})$$

$$+ \sum_{j=1}^{5} \gamma_j \ln(x_{jit}) + \frac{1}{2}\sum_{j=1}^{5}\sum_{l=1}^{5} \gamma_{jl} \ln(x_{jit})\ln(x_{lit}) + e_{it}$$

where D_t is year dummy, capturing year effect which is important in agricultural production; R is the share of root crops in terms of growing area, addressing the difference in the rotation regime. All other variables in (B1) are defined as in model (11). c, δ and γ are parameters to be estimated. The error term is defined as

$$\text{(B2)} \quad e_{it} = \varepsilon_{it} - \eta_i$$

The distributions of the composed errors are assumed as:

1) $\varepsilon_{it} \sim N(0, \sigma_\varepsilon^2), \ i.i.d.$

2) $\eta_i \sim N^+(\mu, \sigma_\eta^2), \ i.i.d.$

3) ε_{it} and η_i are distributed independently of each other.

where ε_{it} is the noise component and follows a normal distribution. The nonnegative η_i is an inefficiency component and follows a normal distribution truncated below at zero.