Journal of Agricultural and Resource Economics 30(3):597–608

On Production Function Estimation with Selectivity and Risk Considerations

Phoebe Koundouri and Céline Nauges

In the estimation of production functions, ignoring risk considerations can cause inefficient estimates, while biased parameter estimates arise in the presence of sample selection. In the presence of uncertainty and selection bias, the latter introduced by the endogeneity of qualitative characteristics of inputs in crop choice, we show that correcting for risk considerations (à la Just and Pope, 1978, 1979) but not selection bias, can produce incorrect inferences in terms of risk behavior. The arguments raised in this study have estimation and policy implications for stochastic production analysis applied to all goods whose qualitative characteristics can affect sample selection.

Key words: crop choice, production risk, sample selection

Introduction

Just and Pope (hereafter denoted JP) (1977, 1978) developed general models for handling production risk econometrically. Their approach has been quite popular among agricultural economists, and is still prominently used in the recent literature (e.g., Tveterås, 1999, 2000; Kumbhakar and Tsionas, 2002, among others). The basic concept introduced by JP was to construct the production function as the sum of two components, one relating to the output level, and one relating to the variability of output. This specification allows the econometrician to differentiate the impact of inputs on output and risk, and has sufficient flexibility to accommodate both positive and negative marginal risks with respect to inputs. In addition, JP show that ignoring risk in the production function can lead to wrong inferences on the technology coefficients and, in particular, can produce standard errors that are misleading by indicating much greater precision in estimation than is, in fact, obtained.

In this study, we extend JP's result with a simple empirical illustration. Using a Just-Pope production function, we demonstrate that the parameters of the risk function can be inconsistent if we fail to control for selectivity in crop choice—i.e., specific characteristics of fixed and quasi-fixed inputs, observed and unobserved (for example, quality of groundwater, type of soil, distance to town), could affect both the choice of the crop to be cultivated and the production function.

Dr. Phoebe Koundouri is a senior lecturer in Economics at the School of Business, Department of Economics, University of Reading, UK, and an Honorary Research Fellow at the Department of Economics/CSERGE, University College London, UK. Dr. Céline Nauges is a researcher at the French National Institute for Research in Agriculture (INRA, LERNA), Toulouse, France. The authors are grateful to Ben Groom for his work on preparing the data, and to Alban Thomas, Nikitas Pittis, and two anonymous referees for their insight and comments. The usual disclaimer applies.

Review coordinated by T. F. Glover.

Using a simple framework in which farmers are assumed to be risk-neutral and in which risk only enters the estimation of the crop-specific production functions, we illustrate biases that might affect the risk function parameters if one does not correct for selectivity in crop choice. Although farmers are likely to be risk-averse in actual production processes, and risk is likely to play a role in the crop-choice decision process, we do not explicitly account for this due to data limitations. However, the "correct" theoretical approach and the implied data requirements are briefly discussed here.

Selectivity, which was not an issue in Just and Pope's (1977, 1978, 1979) articles because they used experimental data, might be an issue in empirical applications using economic data. Selectivity has been acknowledged and taken into account in numerous agricultural models, but always in a risk-free environment. Among others, Moore, Gollehon, and Carey (1994) consider selectivity in modeling the crop-choice decision in their analysis of multi-output production functions. Heshmati (1994) shows that not correcting for selectivity (due to aggregation/truncation from the original sample) can bias the estimation of relevant parameters such as input elasticities and returns to scale. Selectivity has also been commonly considered in the agricultural literature on technology adoption (Dinar and Zilberman, 1991; Khanna, 2001).

The remainder of the paper is organized as follows. In the theoretical description section presented below, we model the simultaneity between the production function and sample selection introduced by crop choice, in the context of a single-output model of farmer behavior with production risk. In the empirical analysis section, this model is then applied to a Just-Pope production function with selectivity corrections using the two-step Heckman (1979) method, on a farm-level sample from Cyprus. Our results suggest that failing to correct for sample selection produces biased parameter estimates of the risk function. These results have important implications for agricultural policy, as they confirm that when selection bias is present, the impact of a policy instrument—both in terms of (foregone) expected profit from production and in terms of the change in revenue as a consequence of hedging against modified production risk—could be misleading. Concluding remarks are offered in the final section.

Theoretical Description

Ignoring production risk for the moment, consider a representative producer facing a choice set of L crops suitable for production. We focus on the case where the outcome of the producer's optimization problem is a "corner solution," representing the decision to cultivate one particular crop. To keep the analysis simple, assume the farmer is risk-neutral and hence maximizes the mathematical expectation of profit. The farmer decides which crop to cultivate, among a set of possible crops, by comparing his or her expected profit (π^*) for different cultivation choices. Using the dummy variable $D_l = 1$ when the lth crop is selected by the farmer and $D_l = 0$ otherwise, we can write:

$$D_l = \begin{cases} 1 \text{ if } \pi_l^* \succ 0 \text{ and } \pi_l^* = \max(\pi_1^*, ..., \pi_L^*) \\ 0 \text{ otherwise,} \end{cases} \tag{1}$$

where π_l^* represents the expected profit when crop l is chosen. Because π_l^* is unobservable when crop l is not grown (our data contain a single cross-section of farmers, and so we observe each farmer in a unique situation, i.e., growing a single type of crop), we model the decision of the farmer to grow crop l by using a discrete choice approach: $D_l = \mathbf{1}[g(\mathbf{z}, \lambda_l) + v_l > 0]$, where $g(\cdot)$ is an unknown function of variables $\mathbf{z}$ and unknown parameters λ_l. It is assumed that differences in terms of expected profitability will be driven by variables $\mathbf{z}$, such as the environment of the farm (quality of accessible groundwater supplies, distance from town, rainfall, total irrigated area, etc.), but also characteristics of the farmer, such as years of farming experience. As some of these characteristics may be unobservable, we specify an additive error term, v_l, with an assumed mean of zero. Once crop l has been selected, the production function corresponding to this particular crop can be defined by:

(2) $$y_l = f(\mathbf{x}, \beta_l) + u_l \quad \text{if } D_l = 1,$$

where y_l is a measure of output level for crop l, and $\mathbf{x}$ is a vector gathering inputs as well as other shifters of the production function ($\mathbf{x}$ may have some elements in common with $\mathbf{z}$); u_l is the usual econometric error term, with assumed zero mean.

With D_l and $\mathbf{z}$ observed for a random sample, but y_l observed only when $D_l = 1$, the output variable in this equation is incidentally truncated from below, on a nonpositive net profit from cultivating a particular crop (i.e., $\pi_l^* \succ 0$). Assuming that v_l and u_l have a bivariate normal distribution with zero means and correlation ρ_l:

$$E[u_l]_{D_l=1} = \sigma_l \frac{\phi(g(\mathbf{z}, \hat{\lambda}_l))}{1 - \Phi(g(\mathbf{z}, \hat{\lambda}_l))},$$

where $E[\cdot]$ denotes expectations; $\phi(\cdot)$ and $\Phi(\cdot)$ represent the standard normal probability density and cumulative distribution functions, respectively; $\sigma_l = \mathrm{cov}(v_l, u_l)$ captures simultaneity in the participation and production function equations. The so-called Mill's ratio used in the two-step Heckman (1979) correction for selection bias is represented by

$$M_l = \frac{\phi(g(\mathbf{z}, \hat{\lambda}_l))}{1 - \Phi(g(\mathbf{z}, \hat{\lambda}_l))}.$$

Therefore, for $D_l = 1$, equation (2) may be rewritten as:

(3) $$y_l = f(\mathbf{x}, \beta_l) + \sigma_l M_l + w_l,$$

where w_l is an error term with assumed zero mean.

Risk deriving from yield uncertainty is now incorporated into the production function (following JP, 1979):

(4) $$y_l = f(\mathbf{x}, \beta_l) + \sigma_l M_l + w_l,$$

with

(5) $$w_l = h(\mathbf{x}, \xi_l)\eta_l.$$

The expression $f(\cdot)$ is the deterministic portion of the production function, and is called the "mean function"; $h(\cdot)$ is the "variance" or "risk function," which captures the effects of each input on the risk of production (as measured by the variance of output). The only random component in the model is η_l, which is assumed i.i.d. $N(0, 1)$. The production disturbance, η_l, represents factors such as weather, or unpredictable variations in machine or labor performance. Consistent with Just and Pope, we assume that η_l is not known to the farmer at the time input decisions are made.

Choosing a very simplified framework, it is assumed here that production risk is the only source of risk (in particular, we assume there is no uncertainty on future input and output prices) and that the farmer is risk-neutral. For purposes of this illustration, the assumption of farmers' risk aversion would introduce undue complexity, as the discrete crop-choice decision should be modeled based on comparisons of expected utility levels. To implement such an approach, the parameters of risk preferences would need to be identified, requiring data on repeated cross-sections and the optimal input choices for alternative crops to be observed. In the absence of such data, and given our simplified framework choice, this analysis is therefore conducted under the assumption of farmer risk neutrality.

Empirical Analysis

Description of Data

A cross-section of 239 farms located in the agricultural region of Kiti in Cyprus was surveyed in 1998. Farmers were asked to give accurate information regarding production activities on representative parcels of their land.[1] Information was requested regarding expenditures on fixed and variable inputs used in the production of final outputs. Examples include pesticides, fertilizers, labor, and water.[2] Input prices, which were not observed at the farm level, are assumed to be homogeneous across farms, which is a reasonable assumption for a small agricultural region such as Kiti with many competitive farms. Output quantities and prices are crop-specific. Variability in input and output levels across farms is guaranteed by the variability in environmental (e.g., climate, soil type) and socioeconomic (e.g., education, age) conditions. Farmers were also requested to provide qualitative data, such as information regarding farm ownership, family characteristics, and access to water resources. Finally, the survey requested climatic data, specifically in the form of rainfall measurements.

Because of the wide variety of crop types represented in the sample, it was necessary to group the crops into three broad categories to overcome the sparseness of individual crop observations, namely: (*a*) vegetables (95 observations), (*b*) cereals (89 observations), and (*c*) citrus (55 observations). The reason for grouping outputs into these three categories is twofold: first, the surveyed farms specialize in either vegetable, cereal, or citrus cultivation; and second, information is available only on total expenditures and total revenues of farms (i.e., we do not have such information for each specific crop).

---

[1] For details on the construction of the questionnaire and collection of the data set, see Koundouri (2000).

[2] Data on the quantities of water used in crop production were sparse and often inconsistent. To address this problem, information regarding water requirements for the specified crops was gathered from the Cyprus Ministry of Agriculture (Agricultural Research Institute, 1998) and was used to calculate theoretical water demands for the farms based on the areas of land devoted to particular crops. This information is used when farm-specific data on irrigation water are missing. Although one of the questions in the survey concerned water use and water costs, the responses to these particular questions were sparse and did not reveal the marginal costs faced by individual farmers. In response to this, we have constructed a tariff for groundwater pumping costs based on hydrological information obtained from the Ministry of Agriculture in Nicosia.

Table 1. Descriptive Statistics by Crop Group: Vegetables and Cereals

		Vegetables Group		Cereals Group	
Description	Unit	Mean	Std. Dev.	Mean	Std. Dev.
Surface allocated (A)[a]	hectare (ha)	3.23	6.76	3.13	4.60
Gross revenue	CYP[b] /ha/annum	2,786	3,764	745	1,331
Fertilizers (F)[c] expenditure	CYP/ha/annum	243.13	444.26	119.73	325.13
Pesticides (P) expenditure	CYP/ha/annum	152.18	419.03	51.08	151.09
Labor (L)[d] expenditure	CYP/ha/annum	229.91	714.42	105.36	211.99
Water (W) expenditure	CYP/ha/annum	117.92	600.52	50.03	194.42
Number of observations		92[e]		89	

[a] Surface allocated includes both irrigated and non-irrigated areas.
[b] CYP = Cyprus pound (1 CYP was approximately $1.5 US at the time of the survey in 1998).
[c] Fertilizers include manure.
[d] Labor is casual work in production.
[e] Although there were 95 vegetable producers in our survey sample, three producers provided insufficient information on expenditures. Thus the descriptive statistics here are based on only 92 vegetable producers.

In what follows, we focus on the vegetables and cereals groups only.[3] These two crop groups possess a number of characteristics that make their cultivation structurally different. In particular, the production of vegetables is labor-intensive (in contrast to cereals production, which is not). Moreover, vegetables are by far the most profitable crops cultivated in Cyprus, as they are competitive in the international arena. Table 1 reports relevant descriptive statistics for these two crop groups.

Model Specification

The production function is assumed to be of the JP form:[4]

$$y = f(\mathbf{x}, \beta) + h(\mathbf{x}, \xi)\eta,$$

where y measures output, and $\mathbf{x}$ is the vector gathering variable inputs and extra production shifters. We assume $E(\eta) = 0$ and $V(\eta) = 1$; $f(\cdot)$ and $h(\cdot)$ are the "mean function" and the "risk function," respectively. This specification allows the input vector to influence both the mean output and the variance of output, without any requirement on the sign of these effects (i.e., inputs can be risk-increasing or risk-decreasing). A linear quadratic form is chosen for the mean output function $f(\cdot)$ for two reasons. First, it is consistent with JP postulates—there is an additive interaction between the mean and variance output functions.[5] Second, it is flexible in the sense of a second-order approximation of any unknown mean output function (Kumbhakar and Tveterås, 2003). The mean output function (f) for the representative farm is expressed as follows:

[3] We do not present the results for the subgroup of citrus producers because of identification problems and lack of fit of the production function.

[4] From this point onward, we suppress subscript l for the crop.

[5] This would not have been the case for a translog specification, which assumes a multiplicative interaction between the mean and variance functions.

$$f(\mathbf{x}, \mathbf{o}) = \beta_0 + \sum_j \beta_j x_j + \sum_j \beta_{2j} x_j^2 + \sum_j \sum_{k(k \neq j)} \beta_{jk} x_j x_k + \mathbf{o}'\gamma,$$

where the vector of inputs $\mathbf{x}$ gathers fertilizers and manure, pesticides, labor, and irrigation water; $\mathbf{o}$ denotes a vector of extra production shifters including total rainfall, total irrigated area (used here as a proxy for soil type),[6] total present value of investment in machinery, distances to the nearest town and coast, and years of experience in farming. To avoid multicollinearity, these variables are assumed to enter the mean production function linearly.

Following JP (1978, 1979), and Kumbhakar and Tveterås (2003), the variance or risk function $h(\cdot)$ is modeled as a Cobb-Douglas form:

(6) $$h(\mathbf{x}) = \xi_0 \prod_j x_j^{\xi_j},$$

where the ξ's are parameters to be estimated.

Estimation Procedure

Our purpose is to estimate the risk function for the two groups of farmers, under the assumption that their decision to grow either vegetables or cereals may also depend on observed and unobserved characteristics of the farm. If these characteristics enter the production function, then parameter estimates of the risk function may be biased. Note that because farms in our sample specialize either in vegetable or cereal cultivations, we are able to use a single (rather than a multi-output) production function. The estimation method consists of the three following steps:

- STEP 1. Two selection equations are estimated, which predict the following two probabilities, respectively: (*a*) the probability of a farmer growing vegetables against the probability of growing any other crop (i.e., cereals or citrus), and (*b*) the probability of a farmer growing cereals against the probability of growing any other crop (i.e., vegetables or citrus). For both of these selection equations, $g(\cdot)$ is assumed linear in the parameters λ, and v is assumed normally distributed. Referring back to the notation presented in equation (1), using a Probit model we estimate the probability that $D_l = 1$, where l indicates either vegetables or cereals. The choice of crop is assumed to be driven by the environment of the farm (rainfall, distances to nearest town and coast), characteristics of the farmer (such as years of experience in farming), and total irrigated area (here again used as a proxy for soil type). The Mill's ratio for each of the crop-specific production functions is computed from the estimated parameters ($\hat{\lambda}$) of the respective crop-specific Probit model.

- STEP 2. Ignoring the presence of the risk function $h(\cdot)$ in the production function model [see equation (3)], we estimate the crop-specific production functions by generalized method of moments (GMM) replacing the unknown Mill's ratios by their estimated value. As the data set is a single cross-section containing only input expenditures and no separate information for price and input quantity, the

[6] Irrigated area is determined by soil type, and for this reason it can be assumed exogenous in the model.

production functions are estimated assuming that the prices of inputs are homogeneous. For purposes of this analysis, we choose to focus on crop yield per unit of surface and not on total production.[7] Per hectare profit functions and per hectare production functions are generally used to avoid size effects (Weaver and Lass, 1992; Yaron, Dinar, and Voet, 1992; Koundouri and Pashardes, 2003).

The production function is estimated separately from the input demand equations. It can be easily shown that derived input demands do not depend on the production shock (η) when one assumes the production shock is not known to the farmer at the time input decisions are made (Zellner, Kmenta, and Drèze, 1966).[8] GMM, which is robust to heteroskedasticity, produces unbiased and efficient estimates of the set of parameters (β, γ, σ).

- STEP 3. Estimate the residuals,

$$\hat{w} = y - \left(\hat{\beta}_0 + \sum_j \hat{\beta}_j x_j + \sum_j \hat{\beta}_{2j} x_j^2 + \sum_j \sum_{k(k \neq j)} \hat{\beta}_{jk} x_j x_k + \mathbf{o}'\hat{\gamma} + \hat{\sigma}\hat{M} \right),$$

which are then used to estimate the risk function, written as in Just and Pope (1979):

$$\ln|\hat{w}| = \xi_0 + \left[\sum_j \xi_j \ln(x_j) \right] + \ln(\eta)$$

[see equations (5) and (6)]. GMM provides efficient and unbiased ξ coefficients.

Estimation Results

Results of the Probit analyses for vegetables and cereals are reported in tables 2 and 3, respectively. Findings indicate that the choice of cultivating either vegetables or cereals is influenced by the qualitative characteristics of the fixed and quasi-fixed inputs of production, as modeled above in the theoretical description section.

The probability of cultivating vegetables (cereals) is positively (negatively) affected by the proportion of the parcel of land being irrigated, since vegetable cultivation requires more water than cereals (see table 1). Distance from the coast (from the town) decreases (increases) the probability of vegetable cultivation. The opposite effects are found in the case of cereals. Finally, results suggest more experienced farmers are more likely to grow cereals and less likely to grow vegetables.

From these estimates, the Mill's ratio is computed for each of the two production groups and incorporated into each production function. In addition to the Mill's ratio, the following variables are included as determinants: variable inputs (pesticides, labor, water, fertilizers), investment in machinery, rainfall, total irrigated area, distance to

[7] We also estimated a production function incorporating absolute levels of production and inputs, along with the overall cultivated surface, in order to overcome the assumption of constant returns to scale. However, the surface variable proved to be the only significant variable in the estimation due to the presence of high collinearity in the data.

[8] There has been a large debate on the issue of input endogeneity in the estimation of the production function with risk considerations following publication of the article by Love and Buccola in 1991. These authors argue that when input choice is influenced by perceived production risk, the latter being defined in terms of the production error, correlation between inputs and the error term is likely to be present in production models. Shankar and Nelson (1999) proved that inconsistency was not an issue in the version of the JP production model in which both the mean and the variance functions have the Cobb-Douglas form.

Table 2. Probit Analysis Results: Estimation of the Probability of Growing Vegetables

Variable/Description	Coefficient	Std. Error	p-Value
Constant	1.2990	0.6135	0.034
Rainfall in 1998 (mm)	-0.0323	0.0221	0.145
Total irrigated area	0.0055	0.0049	0.268
Distance to nearest town (km)	0.0348	0.0158	0.028
Distance to nearest coast (km)	-0.0772	0.0378	0.041
Years of experience in farming	-0.0290	0.0085	0.001

Number of observations	=	239
Likelihood Ratio test statistic	=	20.66
(p-Value)	=	(0.0009)
% of correct predictions	=	63.60

Table 3. Probit Analysis Results: Estimation of the Probability of Growing Cereals

Variable/Description	Coefficient	Std. Error	p-Value
Constant	-1.4641	0.6302	0.020
Rainfall in 1998 (mm)	0.0194	0.0225	0.388
Total irrigated area	-0.0125	0.0060	0.035
Distance to nearest town (km)	-0.0710	0.0180	0.000
Distance to nearest coast (km)	0.1029	0.0385	0.007
Years of experience in farming	0.0371	0.0093	0.000

Number of observations	=	239
Likelihood Ratio test statistic	=	41.27
(p-Value)	=	(0.0000)
% of correct predictions	=	64.44

nearest town, distance to nearest coast, and years of experience in farming. The production function is estimated separately for both groups of farmers. In each case, the estimated model exhibits a satisfactorily high adjusted R^2 for cross-section data, 0.80 for the vegetable producer group and 0.83 for the cereals group. The Mill's ratio is statistically significant at the 1% level for the subgroup of vegetable growers; for the cereals producers, it is significant only at the 20% level. Estimation results of the mean production function for the two subgroups of farmers are given in appendix table A1.

Tables 4 and 5, respectively, report the parameters of the risk function $h(\cdot)$ estimated with and without selectivity correction for all variable inputs, for the vegetable growers and the cereals growers groups. Note that all standard errors have been bootstrapped using 500 replications, representing a finite sample correction on the estimated errors.

The aim of this exercise is to investigate how risk analysis is affected by ignoring selectivity bias in the estimation of the production function. As observed from tables 4 and 5, failing to correct for endogeneity in crop choice is found to bias parameter estimates. For the vegetables group (table 4), even though the signs of the effects remain

Table 4. Estimation of the Risk Function: Vegetable Growers Group

	With Selectivity Correction			Without Selectivity Correction		
Description	Coefficient	Std. Error[a]	*t*-Statistic	Coefficient	Std. Error[a]	*t*-Statistic
Fertilizers	0.449	0.257	1.749	0.350	0.266	1.319
Pesticides	0.186	0.136	1.370	0.274	0.138	1.978
Labor	-0.361	0.078	-4.617	-0.304	0.075	-4.079
Water	0.055	0.170	0.323	0.105	0.213	0.494

[a] Bootstrapped standard errors computed using 500 replications.

Table 5. Estimation of the Risk Function: Cereals Growers Group

	With Selectivity Correction			Without Selectivity Correction		
Description	Coefficient	Std. Error[a]	*t*-Statistic	Coefficient	Std. Error[a]	*t*-Statistic
Fertilizers	0.516	0.565	0.914	0.607	0.561	1.081
Pesticides	0.048	0.279	0.172	-0.132	0.304	-0.435
Labor	-0.365	0.086	-4.233	-0.541	0.105	-5.166
Water	-1.031	0.531	-1.941	-0.858	0.425	-2.017

[a] Bootstrapped standard errors computed using 500 replications.

the same, the contribution of each input to the variance is found to be different depending on whether or not we correct for selectivity. More precisely, although labor is found to be risk-increasing under either scenario, the impact of pesticides and fertilizers is different depending on whether or not we correct for selectivity. Fertilizers are shown not to affect production risk significantly when selectivity is not taken into account, but is found to have a positive and significant effect on risk when we correct for selectivity bias. In the case of pesticides, this input is found to be risk-increasing only when selectivity is taken into account.

The value of the estimated coefficients, which can be interpreted as output variance elasticities (Tveterås, 2000), also varies between selectivity-corrected and non-selectivity-corrected models. Variance elasticity with respect to labor for the vegetable growers subgroup (table 4) is estimated at -0.36 when there is a correction for selectivity, while it is estimated at -0.30 when there is no selectivity correction. The bias in the measurement of variance elasticity is larger in the case of pesticides: the unbiased elasticity is found equal to 0.19, while it is estimated at 0.27 when there is no correction for selectivity.

For the case of cereals growers (table 5), some differences are observed between estimates with and without selectivity correction, even though the Mill's ratio was almost nonsignificant in the production model for cereals producers. Labor and water are found to be risk-decreasing inputs in both models (at a high level of significance); however, the magnitude of their effects varies from one model to the other. Note also, despite its lack of significance, the estimated impact of pesticides is found to be positive in the case with selectivity correction, and negative in the case without.

We must remain cautious about the estimates of variance elasticities, as our model has been estimated under the assumption that farmers are risk-neutral. However, we argue this simple illustration shows that failing to correct for sample selection in the estimation of a stochastic production function leads to incorrect inferences with respect to the risk function. Such bias has important policy implications because it can result in misleading inferences on input-specific risk effects.

For example, consider a policy maker who is contemplating the introduction of a conservation policy aimed at reducing the use of fertilizers in order to restrict the adverse effects on the quality of soil and groundwater resources and to adhere to the relevant European Union environmental directive. If the policy maker decides to introduce a quota on the use of these inputs, then an investigation of how this policy instrument will affect agricultural revenue should be conducted. The policy instrument can affect revenue in terms of (foregone) expected profit from production alone, but can also bring about a change in revenue as the consequence of hedging against a modified production risk (the latter originating from the farmer's need to modify his or her insurance behavior). In such a case, agricultural subsidy schemes should be designated simultaneously with the environmental quota. If, however, selection bias exists in policy simulations, the policies subsequently projected will be inefficient not only with respect to technical issues but also with respect to risk hedging.

Conclusion

This study investigates selectivity bias introduced by crop choice in agricultural production under risk. To avoid estimation bias from ignoring sample selection biases, we apply Heckman's (1979) correction to a Just-Pope production function. Under the assumption of farmer risk neutrality, results of the empirical analysis suggest that failure to correct for sample selection will bias estimation of input-specific marginal risk. The arguments raised in this paper have implications for stochastic production analysis applied to all goods whose qualitative characteristics can affect sample selection.

[*Received October 2003; final revision received July 2005.*]

References

Agricultural Research Institute. "Norm Input-Output Data for the Main Crop and Livestock Enterprises of Cyprus." Agr. Econ. Rep. No. 36, Ministry of Agriculture, Natural Resources, and the Environment, Cyprus, 1998.

Dinar, A., and D. Zilberman. "The Economics of Resource-Conservation, Pollution-Reduction Technology Selection: The Case of Irrigation Water." *Resour. and Energy Econ.* 13(1991):323–348.

Heckman, J. "Sample Selection Bias as a Specification Error." *Econometrica* 47,1(1979):153–161.

Heshmati, A. "Estimating Random Effects Production Function Models with Selectivity Bias: An Application to Swedish Crop Producers." *Agr. Econ.* 11(1994):171–189.

Just, R. E., and R. D. Pope. "On the Competitive Firm Under Production Uncertainty." *Austral. J. Agr. Econ.* 21,2(1977):111–118.

———. "Stochastic Representation of Production Functions and Econometric Implications." *J. Econometrics* 7,1(1978):67–86.

———. "Production Function Estimation and Related Risk Considerations." *Amer. J. Agr. Econ.* 61,2(1979):276–284.

Khanna, M. "Sequential Adoption of Site-Specific Technologies and Its Implications for Nitrogen Productivity: A Double Selectivity Model." *Amer. J. Agr. Econ.* 83(2001):35–51.

Koundouri, P. "Three Approaches to Measuring Natural Resource Scarcity: Theory and Application to Groundwater." Unpub. Ph.D. diss., Dept. of Economics, Faculty of Economics and Politics, University of Cambridge, 2000.

Koundouri, P., and P. Pashardes. "Hedonic Price Analysis and Selectivity Bias: Water Salinity and Demand for Land." *Environ. and Resour. Econ.* 26,1(2003):45–56.

Kumbhakar, S. C., and E. G. Tsionas. "Estimation of Production Risk and Risk Preference Function—A Nonparametric Approach." Work. pap., State University of New York, 2002.

Kumbhakar, S. C., and R. Tveterås. "Risk Preferences, Production Risk, and Firm Heterogeneity." *Scandinavian J. Econ.* 105,2(2003):275–293.

Love, H. A., and S. T. Buccola. "Joint Risk Preference-Technology Estimation with a Primal System." *Amer. J. Agr. Econ.* 73(1991):765–774.

Moore, M., N. Gollehon, and M. Carey. "Multicrop Production Decisions in Western Irrigated Agriculture: The Role of Water Price." *Amer. J. Agr. Econ.* 76(1994):859–874.

Shankar, B., and C. H. Nelson. "Joint Risk Preference-Technology Estimation with a Primal System: Comment." *Amer. J. Agr. Econ.* 81(1999):241–244.

Tveterås, R. "Production Risk and Productivity Growth: Some Findings for Norwegian Salmon Aquaculture." *J. Productivity Analy.* 12,2(1999):161–179.

———. "Flexible Panel Data Models for Risky Production Technologies with an Application to Salmon Aquaculture." *Econometric Rev.* 19,3(2000):367–389.

Weaver, R. D., and D. A. Lass. "Corner Solutions in Duality Models: A Cross-Section Analysis of Dairy Production Decisions." *Amer. J. Agr. Econ.* 71(1992):1025–1040.

Yaron, D., A. Dinar, and H. Voet. "Innovations on Family Farms: The Nazareth Region in Israel." *Amer. J. Agr. Econ.* 74(1992):361–370.

Zellner, A., J. Kmenta, and J. Drèze. "Specification and Estimation of Cobb-Douglas Production Function Models." *Econometrica* 34,4(1966):784–795.

Table A1. Estimation of the Mean Production Function (including the Mill's ratio) for the Two Farmer Groups

	Vegetables Group			Cereals Group		
Description	Coefficient	Standard Error	*p*-Value	Coefficient	Standard Error	*p*-Value
Constant	-2.595	0.889	0.005	8.036	7.216	0.269
Fertilizers (F)	0.003	0.200	0.989	1.426	0.799	0.079
Pesticides (P)	0.425	0.131	0.002	1.273	0.307	0.000
Labor (L)	-0.002	0.001	0.020	0.010	0.008	0.254
Water (W)	1.357	0.288	0.000	0.069	0.329	0.835
$F \times F$	0.121	0.046	0.010	0.052	0.030	0.088
$P \times P$	0.000	0.035	0.996	0.033	0.019	0.086
$L \times L$	0.000	0.000	0.004	0.000	0.000	0.719
$W \times W$	-0.001	0.002	0.488	-0.021	0.017	0.212
$F \times P$	-0.060	0.089	0.499	-1.156	0.255	0.000
$F \times L$	0.000	0.000	0.508	-0.015	0.012	0.204
$F \times W$	-0.011	0.039	0.773	0.365	0.256	0.159
$P \times L$	0.000	0.000	0.321	0.006	0.003	0.083
$P \times W$	-0.657	0.137	0.000	0.763	0.308	0.016
$L \times W$	0.000	0.000	0.600	0.001	0.000	0.230
Investment in machinery	-0.030	0.068	0.662	0.205	0.068	0.003
Total irrigated area	-0.614	0.192	0.002	0.283	0.227	0.217
Total rainfall in 1998	11.361	3.158	0.001	-1.239	1.391	0.376
Distance to nearest town	-5.379	1.560	0.001	2.463	2.118	0.249
Distance to nearest coast	3.048	0.954	0.002	-0.809	1.194	0.500
Years of farming experience	9.463	2.782	0.001	-4.332	3.355	0.201
Mill's ratio	-16.619	4.978	0.001	-5.706	4.321	0.191
Number of observations		92			89	
Adjusted R^2		0.80			0.83	

Note: All variables have been mean-scaled.