Where has all the Water Gone?
Estimation of a Production Function for the Agricultural Sector in the Region Khorezm, Uzbekistan

By

Marc Mueller

Selected paper presented at the
American Agricultural Economics Association Annual Meetings
Denver, CO-August 1-4, 2004

ZEF/UNESCO Project

Economic and Ecological Restructuring of Land- and Water Use in the Region Khorezm (Uzbekistan): A Pilot Project in Development Research

---

Where Has All The Water Gone?
Estimation of a Production Function for the Agricultural Sector in the Region Khorezm, Uzbekistan

Marc Müller

Draft Version

Marc Müller
Agricultural Economist
Centre for Development Research, University of Bonn
Walter-Flex Str. 3
53113 Bonn
Phone: +49 (0)228 73 9052
E-Mail: marc.mueller@uni-bonn.de
mmarcmueller@aol.com

Contents

1. Introduction

This study has two objectives: The first is to address a problem widely known in agricultural economics, especially in the case of developing countries: Despite the often significant importance of agriculture in regional or national economies and the relevance of agriculture for the country specific ecological systems, input-output relationships are often poorly recorded and hence, analyzed. With the exception of land allocated to certain crops it is hardly possible to assess information about quantities of water, labor, or fertilizer used for the different crops in an agricultural production system. This situation is often aggravated by a multitude of government interventions. Additionally, the analysis of input-output relations becomes more difficult if both, agricultural input and output markets are distorted and behavioral patterns of agricultural producers do not follow common assumptions such as profit maximization. In such a context it may prove difficult to analyze efficiency and productivity of input use and even more difficult to derive recommendations for more sustainable natural resource use based on sound production function analysis.

The second objective of this paper is to quantify the allocation of water for the different crops in this region and to gain information about possibilities to reduce the water demand. The case-study region Khorezm is of interest because of the reliance of the regional economy on agriculture and the significant use of irrigation water from the river Amu Darya. This supply became increasingly scarce since the mid-nineties and it is likely that the situation will deteriorate due to the recovering of crop production in more up-stream located countries, such as Afghanistan.

A set of factors determining the ability of the agricultural system in Khorezm to adapt to changes in the availability has already been identified. Most important among those are the governmental regulations concerning the production of main crops like cotton, wheat, and rice, the persistence of structures inherited from the former Soviet Union, the lack of advanced irrigation technology and the fact, that irrigation water is free of charge.

The analyses performed in this paper contribute to this discussion by estimating crop-specific input usages and marginal rates of substitution between the main inputs water, land, physical capital and labor.

2. The Black Box: Agriculture

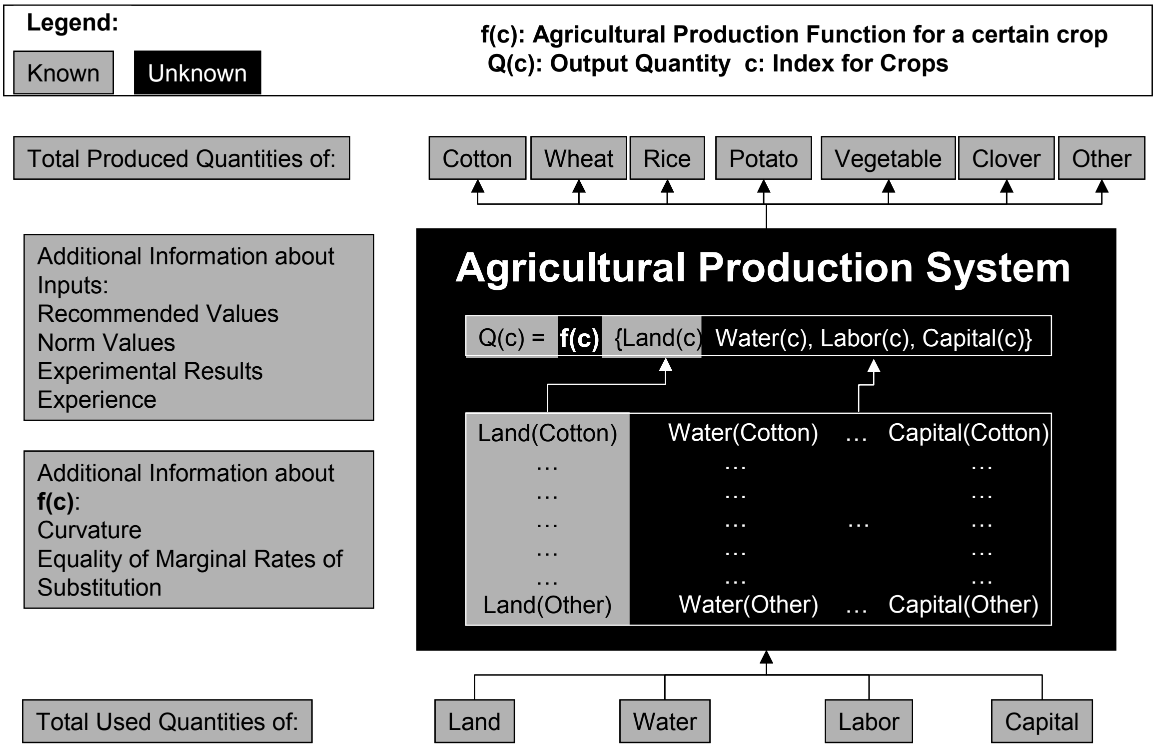
The case-study region has some special difficulties due to governmental attempts to replace the legacies inherited from the soviet era such as large-scale former collective state farms by a number of smaller, so-called private farms. The major farms currently in operation still rely on the same providers of essential inputs like machinery and fertilizer, and have to deliver a pre-determined amount of cotton, wheat and rice to state-owned processors. These state interventions continue but it is not the focus of this study to speculate about the institutional patterns that will determine the future of agriculture in this region. The entire sector will be analyzed from a Sector-wide perspective and it is assumed that the physical input-output relations of the entire region can be described with a set of primal production functions for each crop.

The problem at hand can be described with figure 1. The produced quantities of crops are known[1] as well as the total input quantities, but the quantities of inputs allocated for the production of specific crops is unknown with the exception of land. What is known about the input allocations is a set of 'norm' values that were derived during the soviet period and were used to calculate the needed quantities of water, labor-hours and diesel in the framework of a planned economy. These norm values are still in use as 'rules-of-thump' for farmers to calculate their needs for the following cropping period and represent to some extend the knowledge of farmers about their production processes.

Another source of information for the following study are assumptions about the functional forms to be estimated and behavioral patterns of the relevant actors. Those assumptions will be discussed in more detail in the following chapters.

[1] Given the general problem of data availability and reliability in the underlying case, the expression 'known' has two distinct meanings:
1. Sets of information have been made available by one source and do not contradict information from other sources.
2. Sets of information from one source do contradict information from other sources but have been confirmed by the responsible official bodies on inquiry.

Figure 1 The Black Box



3. Method

Because the lack of activity-specific input data is a widely known problem in development economics (and agricultural economics in general), this issue has been addressed already by several authors (e.g. Just 1990 and Lence, Miller 1998). While Just (1990) proposes an ordinary least squares (OLS) estimation model for the calculation of input data based on allocated area and dummy variables to capture annual and farm-specific effects, Lence and Miller (1998) suggest to use a maximum entropy (ME) approach to derive not only the input data but to estimate parameters of a production function simultaneously. The latter approach appears to match the problem discussed here and therefore will be described in more detail:

The ME approach as described in detail in Golan, Judge, and Miller (1996) allows the usage of prior information about the expected value of the variables to be estimated and a symmetric interval around it, resulting in a range of possible outcomes. Any variable can be represented by the borders of the respective assumed range and a corresponding probability for the variable to equal either of the border-values or ‘support points’[2]. E.g. if the probability of both support-points assumes 0.5 the variable equals the mean-value. The ME objective function is often specified as in equation 1 (E1) under the restrictions of E2 and E3:

E1[3]: $$H = -\sum_{l=\min}^{\max} p_l \ln(p_l)$$

E2: $$\sum_{l=\min}^{\max} p_l = 1$$

E3: $$C = \sum_{l=\min}^{\max} p_l * x_l$$

With: H: Entropy measure, to be maximized
p: Probability of each support point
x: Support point
C: Variable to be estimated
l: Index for support points (here: minimum [min] and maximum [max])

Restriction E2 assures that the probabilities add up to one; E3 links the support points of each variable with the respective probability. E1 has its global maximum in the point of equal distribution of the probabilities, as shown in the following figure 2. Consequently, the ME approach relies mainly on the quantification of the expected value of the variables to be estimated and the respective support points. The practical problem is to determine the support points based on all information available.

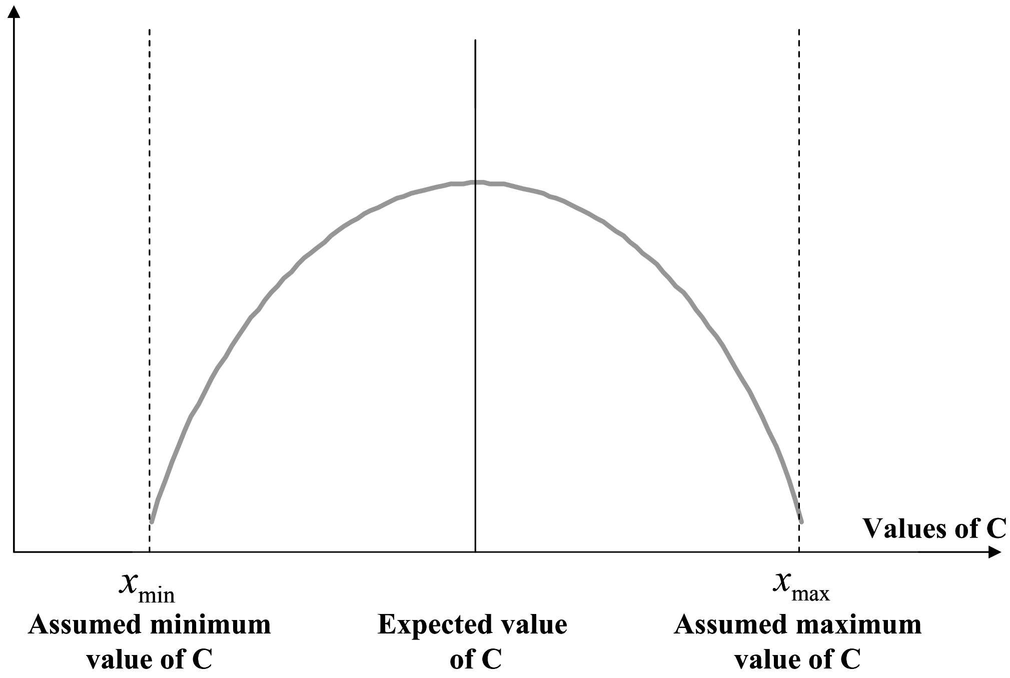
---

[2] Although it is possible to include not only border-values but also distributions within this range, these more advanced approaches are neglected in this study in order to keep the final model as simple as possible.

[3] In order to improve the readability of the formulas used in this paper, the equations are named by their type and not only by their sequence of appearance in the text:

E: Explaining formulas or side-calculations
D: Definitions
M: Equations used in the finally derived model

Figure 2 Entropy Function



Lence and Miller (1998) suggest to assume that each crop-specific input for a set of crops can be described as the respective share in the total quantity of the available input. Those shares add up to 1 and consequently fulfill the adding-up condition E2 and can be used in the ME procedure. Such a formulation would imply that all crops have an equal share in the respective total input use, because the (ME) objective function will have a maximum in the case of equal distribution. In our case, such an assumption is not realistic because the cropping system of the case-study region includes crops such as cotton and wheat on the one hand and rice on the other. Among these crops it is more likely that rice production has a much higher share in the total water usage, even if the allocated area is a small part of the total area. In the case of Khorezm, rice area is a significant part of the agricultural production system and therefore it would be unrealistic to maintain the assumption of equal input shares. This consideration leads to a variation of the estimation model proposed by Lence and Miller (1998) which will be described in the next section The question is how to incorporate assumptions about the demand of different crops on the inputs in the estimation system.

3.1. Support Points

As mentioned above, Just (1990) has suggested a method to disaggregate input usages with an OLS procedure. Although the data series used here are not as extended as in the study by Just, this model appears to be an adequate approach to the problem of the estimation of crop-specific input allocations. According to Just (1990), input allocations can be estimated by using the total available inputs on the left-hand side of a regression model and the known area allocations as well as dummies on the right-hand side (E4):

E4: $$TAI_{iuk,r,t} = \sum_{cm}\left(ccrop_{iuk,cm} + cray_{iuk,r}D_r^{ray} + cyear_{iuk,t}D_t^{year}\right) * Area_{cm,r,t} + \varepsilon_{iuk,r,t}$$

With: TAI: Total available input

Area: Allocated area

ccrop Parameter covering the crop-effects

cray: Parameter covering the regional effects

cyear: Parameter covering the annual effects

D: Dummy variables covering regional and annual effects

ε: Error term

iuk: Index for unknown inputs (All Inputs except Land)

r: Index for rayons (districts)

t: Index for time (1998 to 2001)

cm: Index for crops in the Model

ray: Index for Rayon Dummies

year: Index for year Dummies

The estimated crop-specific inputs would then be the sum of the crop effect, the annual effect, and the regional effect:

E5: $$CSI_{iuk,cm,r,t}^{est} = ccrop_{iuk,cm} + cray_{iuk,r} + cyear_{iuk,t}$$

With: CSI^{est}: Estimated crop-specific input

The estimated values for CSI are point estimates but the interest here is to derive an interval which can be used for the entropy procedure. Consequently, variance and standard deviation of CSI are derived from the estimation results:

E6:

$$
\begin{aligned}
&Var[CSI_{iuk,cm,r,t}^{est}] = \\
&Var[ccrop_{iuk,cm}] + Var[cray_{iuk,r}] + Var[cyear_{iuk,t}] + \\
&Cov[ccrop_{iuk,cm}, cray_{iuk,r}] + Cov[ccrop_{iuk,cm}, cyear_{iuk,t}] + Cov[cray_{iuk,r}, cyear_{iuk,t}]
\end{aligned}
$$

With: Var: Variance

Cov: Covariance

The support points for each CSI are then calculated with equations E7 and E8 under the assumption that in case of a normal distribution a confidence interval of 99.9% lies in the range of CSI plus and minus three times standard deviation:

E7: $$x_{cm,iuk,n,MAX}^{CSI} = CSI_{iuk,cm,r,t}^{est} + 3\sqrt{Var[CSI_{iuk,cm,r,t}^{est}]}$$

E8: $$x_{cm,iuk,n,MIN}^{CSI} = CSI_{iuk,cm,r,t}^{est} - 3\sqrt{Var[CSI_{iuk,cm,r,t}^{est}]}$$

Wtith: x^{CSI} Support point for CSI

n: Index for observations, r*t

M1: $$CSI_{iuk,cm,n} = \sum_{sx} p_{cm,ipf,n,sx}^{CSI} x_{cm,ipf,n,sx}^{CSI}$$

M2: $$\sum_{sx} p_{cm,ipf,n,sx}^{CSI} = 1$$

3.1.1. Estimation

It turned out that the results for CSI by estimating E4 with ordinary least squares were not satisfying: They did not match the available ‘norm’ values for each input and became negative in some cases, what is highly unrealistic since physical input quantities cannot have values below zero. The source that problem is the comparatively small database. This issue was addressed by employing the ‘mixed estimation method’ proposed by Theil and Goldberger (1967) which allows for the inclusion of additional information about the parameters to be estimated in the estimation procedure. The model was formulated according to Greene (2003)[4]:

E9: $$E[\boldsymbol{\beta} \mid \sigma^2, TAI, \mathbf{X}] = \left(\boldsymbol{\Sigma}_{\mathbf{0}}^{-1} + \left(\sigma^2 (\mathbf{X'X})^{-1}\right)^{-1}\right)^{-1} \left(\boldsymbol{\Sigma}_{\mathbf{0}}^{-1}\boldsymbol{\beta}_{\mathbf{0}} + \left(\sigma^2 (\mathbf{X'X})^{-1}\right)^{-1}\mathbf{b}\right)$$

E10: $$\mathbf{b} \; = \mathbf{X'X}^{-1}\mathbf{X'TAI}$$

E11: $$\mathbf{e} = TAI - \mathbf{Xb}$$

E12: $$\sigma^2 = s^2 = \frac{\mathbf{e'e}}{(n-k)}$$

With:	E[ ]	Expected value
	$\boldsymbol{\beta}$:	Vector of parameters to be estimated (ccrop, cray, cyear)
	σ^2:	Variance of $\boldsymbol{\beta}$ (obtained from OLS regression, s^2)
	X:	Matrix of Area and Dummy variables
	$\boldsymbol{\Sigma}_0$:	Prior information about variances of $\boldsymbol{\beta}$
	$\boldsymbol{\beta}_0$:	Prior information about expected values of $\boldsymbol{\beta}$
	b:	Parameter vector obtained from OLS regression
	e:	Error term of OLS regression
	k:	Number of parameters to be estimated

The crucial point of this method is to determine the prior information about expected values of the parameters ($\boldsymbol{\beta}_0$) and their variances ($\boldsymbol{\Sigma}_0$) accurately. Especially when the sample is comparatively small, the weight of that prior information in the estimation process will be very high. Consequently, $\boldsymbol{\beta}_0$ was constructed by using the 'norm' values for the inputs in the case of the parameter group covering the crop effects (ccrop). Prior information for regional effects (cray) and annual effects (cyear) are not available and set to zero. For the variances $\boldsymbol{\Sigma}_0$ it was assumed, that in the case of crop effects it has to be small enough to make negative values very unlikely, the variances of the regional and annual effects are taken from the OLS regression. The resulting prior information is shown in the following table 1 for the case of water. This table also shows the specification of the model: 12 crops, three regions and three years are covered in the estimation procedure.

[4] In order to maintain the readability of the following formulas, the variables are named according to Greene (2003)

Table 1: Prior Information for Water

		$\boldsymbol{\beta}_0$ [1000 m^3/ha]	$\boldsymbol{\Sigma}_0$
ccrop	Cotton	5,62	0,23
	Coarse grains	3,65	0,10
	Wheat	3,65	0,10
	Rice	26,20	0,10
	Maize	5,32	0,26
	Sugar beet	4,86	0,22
	Potatos	8,58	0,33
	Vegetables	8,58	0,33
	Fruit	5,19	0,12
	Melons	3,99	0,15
	Grapes	5,19	0,12
	Clover	8,42	0,29
cray	Lower Amu Darya	0	1,85
	Upper Amu Darya	0	1,52
	Not bordering Amu Darya	0	1,43
cyear	1999	0	1,63
	2000	0	2,12
	2001	0	2,02

Source: OblVodKhoz and own calculations

In order to avoid singularity of the matrix of explaining variables (X) the model was normalized for the year 1998 and the Rayon Bagat, located at the upper Amu Darya.

The results from the estimation are shown in table 2. It turns out that the crop effects do not deviate too much from the prior information what indicates that the sample has a low explanatory power in this case. The annual effects on the other hand do deviate and show plausible results with a positive value for 1999, which was a year with sufficient water supply, and negative values for 2000 and 2001, which were drought years. The regional effects show that the water usage per hectare in the regions not bordering the Amu Darya is higher than in the regions along the river. This result seems counter-intuitive but makes sense when water losses are taken into account: In the off-stream regions more water is needed per hectare in order to compensate for the losses associated with transporting water from the river to the respective regions. These results and the respective ones for diesel and labor are used to calculate the support points for the crop-specific inputs according to equations E7 and E8.

Table 2: Estimation Results for Water

		β [1000 m^3/ha]
ccrop	Cotton	5,95
	Coarse grains	3,66
	Wheat	3,68
	Rice	26,21
	Maize	5,32
	Sugar beet	4,89
	Potatos	8,59
	Vegetables	8,63
	Fruit	5,20
	Melons	4,00
	Grapes	5,19
	Clover	8,50
cray	Lower Amu Darya	5,84
	Upper Amu Darya	4,61
	Not bordering Amu Darya	8,54
cyear	1999	5,62
	2000	-1,23
	2001	-3,85

Source: Own results

3.2. Production Function

Due to the demand that the production functions for the different crops should incorporate the four different inputs land, water, labor, and physical capital, it appears not appropriate to specify them as Cobb-Douglas functions: This would imply an elasticity of substitution of one between all pairs of inputs. There is a huge variety of possible formulations of more flexible functional forms with different properties especially concerning their ability to fulfill regularity conditions such as concavity and monotonicity not only locally but globally (e.g. Diewert, Wales 1987).

The functional form chosen here is a quadratic one. This decision was made for two reasons: First, a quadratic function can be globally concave (although not globally monotonous) and second, its computational simplicity. The major disadvantage of a quadratic function is that it does not imply that certain inputs are essential: It would be possible to get positive output values even if one input value is set to zero, which is a highly unrealistic assumption particularly in the case of land. This property has to taken into account for the set-up of simulations. The functional form used in this study is specified according to Fuss et al. (1978) as follows:

M3: $$Q_{cm,n} = \sum_i \kappa^L_{cm,i} * CSI_{i,cm,n} + \frac{1}{2}\sum_i \sum_{ip} \kappa^Q_{cm,i,ip} CSI_{i,cm,n} CSI_{ip,cm,n}$$

With: Q: Output of each crop

κ^L: Parameter for the linear terms

κ^Q: Parameter for the quadratic terms (and cross terms)

i, ip: Index for all inputs

Monotonicity is imposed by restricting the first partial derivatives of Q to be non-negative in all observation points n:

M4: $$\frac{\partial Q_{cm,n}}{\partial CSI_{i,cm,n}} \geq 0$$

Concavity is imposed according to Lau (1978) by decomposing the Hessian matrix of second partial derivations (**H**) into a lower triangular unit matrix **L** and an upper triangular matrix **U** (Definitions D1 to D3 and equation M5):

D1: **H** = $\begin{bmatrix} \frac{\partial^2 Q_{cm,n}}{\partial CSI_{Land,cm,n} \partial CSI_{Land,cm,n}} & \cdots & \cdots & \frac{\partial^2 Q_{cm,n}}{\partial CSI_{Land,cm,n} \partial CSI_{Capital,cm,n}} \\ \cdots & \cdots & & \cdots \\ \cdots & & \cdots & \cdots \\ \frac{\partial^2 Q_{cm,n}}{\partial CSI_{Capital,cm,n} \partial CSI_{Land,cm,n}} & \cdots & \cdots & \frac{\partial^2 Q_{cm,n}}{\partial CSI_{Capital,cm,n} \partial CSI_{Capital,cm,n}} \end{bmatrix}$

D2: **L** = $\begin{bmatrix} 1 & 0 & 0 & 0 \\ l_{i,ip} & 1 & 0 & 0 \\ l_{i,ip} & l_{i,ip} & 1 & 0 \\ l_{i,ip} & l_{i,ip} & l_{i,ip} & 1 \end{bmatrix}$

D3: **U** = $\begin{bmatrix} u_{i,ip} & u_{i,ip} & u_{i,ip} & u_{i,ip} \\ 0 & u_{i,ip} & u_{i,ip} & u_{i,ip} \\ 0 & 0 & u_{i,ip} & u_{i,ip} \\ 0 & 0 & 0 & u_{i,ip} \end{bmatrix}$

M5: **H** = **LU**

or:

$$\frac{\partial^2 Q_{cm,n}}{\partial CSI_{i,cm,n} \partial CSI_{ip,cm,n}} = \sum_{im} l_{i,im,n} u_{im,ip,n}$$

With: im: auxiliary index for matrix multiplication, alias of i and ip

According to Lau (1978), Q is concave when **H** is negative semi-definite, which is the case, when the diagonal elements of **U** are non-positive.

M6: $u_{i,ip,n} \leq 0$, $\mathrm{i} = \mathrm{ip}$

Restriction M6 is imposed in all observation points n. Since a quadratic function is by definition globally concave when it is concave in one point, the index n becomes redundant for this model specification. In order to maintain the generality of the model description it is still included for the case that another functional form is chosen, which is not globally concave. The proposed ME procedure requires the definition of support points for the parameters and their association with respective probabilities:

M7: $$\kappa_{cm,q} = \sum_{sp} p^{\kappa}_{cm,q,sp} k^{\kappa}_{cm,q,sp}$$

M8: $$\sum_{sp} p^{\kappa}_{cm,q,sp} = 1$$

With: p^{κ}: Probability for each support point for κ

k^{κ}: Support Point for κ

q: Index for parameters of linear, quadratic terms, and cross terms

sp: Index for support points for κ

The values for k^{κ} were chosen in a way that the linear terms are always positive and the quadratic terms are always negative. Cross terms can be positive or negative. For the estimation procedure, M3 was associated with an error term with an expected value of zero. This error term is also defined in an interval between two support points:

M9: $$\varepsilon_{cm,n} = \sum_{sp} p^{\varepsilon}_{cm,n,sp} k^{\varepsilon}_{cm,n,sp}$$

M10: $$\sum_{sp} p^{\varepsilon}_{cm,n,sp} = 1$$

With: ε: Error term of the production function

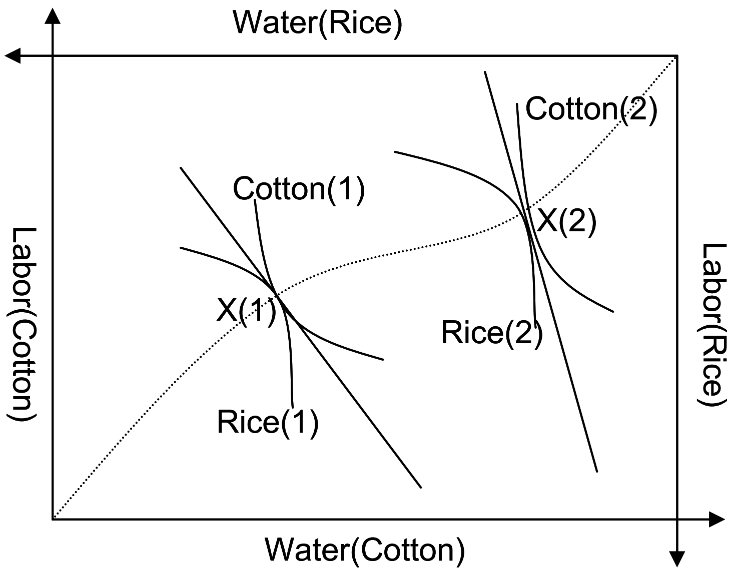
P^{z}: Probability for each support point for ε

k^{z}: Support points for ε

3.3. Imposing Rationality

The fact that agricultural production in the region Khorezm is subject to a set of governmental regulations and that there are no markets for relevant inputs like water and land and consequently no prices for those inputs, make assumptions about the behavior of the relevant actors in the system rather difficult. They are surely not profit maximizers in the sense that they can decide about optimal output quantities. But it might be realistic to assume that they produce whatever they are required by the state with a minimal cost combination of inputs. Accordingly in this study we assume that the production is efficient and available resources are not wasted. This assumption is depicted for the simplified two goods, two factor case in figure 3 with an Edgeworth Box. It shows arbitrarily chosen isoquant curves for the two products cotton and rice which are produced with two production factors only, water and land. The points x1 and x2 (where the isoquants for both products are tangent to each other) represent possible efficient combinations of input usages: The total available amounts of water and land are used in both cases. Both points also imply that the marginal rates of substitutions are equal for both production functions. This implication is realistic when the prices of land and water are equal for cotton and rice producers. The resulting equalities are summarized in definition 4. If the input prices are equal for all producers, the negative marginal rate of substitution equals the inverted ratio of input prices.

Figure 3 Efficient Production



D4: $$\frac{\dfrac{\partial Q_{cm,n}}{\partial CSI_{i,cm,n}}}{\dfrac{\partial Q_{cm,n}}{\partial CSI_{ip,cm,n}}} = -\frac{dCSI_{ip,cm,n}}{dCSI_{i,cm,n}} = -MRS_{i,ip,cm,n} = \frac{P[CSI_{i,cm,n}]}{P[CSI_{ip,cm,n}]}$$

With: MRS: Marginal rate of substitution between pairs of inputs

P[.]: (Unknown) input price

Even if the prices are unknown, their ratios can be derived. The condition of equal marginal rates of substitution is imposed in the estimation process according to equation E19:

E19: $$\frac{\dfrac{\partial Q_{cm,n}}{\partial CSI_{i,cm,n}}}{\dfrac{\partial Q_{cm,n}}{\partial CSI_{ip,cm,n}}} = \frac{\dfrac{\partial Q_{cmp,n}}{\partial CSI_{i,cm,n}}}{\dfrac{\partial Q_{cmp,n}}{\partial CSI_{ip,cm,n}}}, \ \mathrm{i} \neq \mathrm{ip}, \mathrm{cm} \neq \mathrm{cmp}$$

In order not to over-constrain the estimation process, this restriction is relaxed by defining and introducing a distortion term in E19 which has an expected value of zero. The support points are chosen such that they keep this distortion term as small as possible while maintaining the feasibility of the model. E19 is re-formulated as follows:

M10: $$\frac{\partial \mathrm{Q}_{\mathrm{cm,n}}}{\partial \mathrm{CSI}_{\mathrm{i,cm,n}}} * \frac{\partial \mathrm{Q}_{\mathrm{cmp,n}}}{\partial \mathrm{CSI}_{\mathrm{ip,cm,n}}} = \frac{\partial \mathrm{Q}_{\mathrm{cmp,n}}}{\partial \mathrm{CSI}_{\mathrm{i,cm,n}}} * \frac{\partial \mathrm{Q}_{\mathrm{cm,n}}}{\partial \mathrm{CSI}_{\mathrm{ip,cm,n}}} + \delta_{\mathrm{i,ip,cm,cmp,n}}, \ \mathrm{i} \neq \mathrm{ip}, \mathrm{cm} \neq \mathrm{cmp}$$

With: δ: Distortion term

δ is defined between two support points and the associated probabilities.

M11: $$\delta_{i,ip,cm,cmp,n} = \sum_{sp} p^{\delta}_{i,ip,cm,cmp,n,sp} k^{\delta}_{i,ip,cm,cmp,n,sp}, \ \mathrm{i} \neq \mathrm{ip}, \mathrm{cm} \neq \mathrm{cmp}$$

M12: $$\sum_{sp} p^{\delta}_{i,ip,cm,cmp,n,sp} = 1$$

With: P^{δ}: Probability for each support point for δ

k^{δ}: Support points for δ

Equation 21 completes the system that has to be complemented by an objective (ME) function. This will be described in the following chapter:

3.4. Objective Function and Implementation

The objective function contains all probabilities of the variables to be estimated (crop-specific inputs (CSI), parameters of the production function (κ), error term of the production function (ε), and distortion term of the cost-minimization condition (δ)) and is specified as shown below:

M13:

$$
\begin{aligned}
&Entropy = \\
&-\sum_{cm}\sum_{ipf}\sum_{n}\sum_{sx} p^{CSI}_{cm,ipf,n,sx} \ln(p^{CSI}_{cm,ipf,n,sx}) \\
&-\sum_{cm}\sum_{q}\sum_{sp} p^{\kappa}_{cm,q,sp} \ln(p^{\kappa}_{cm,q,sp}) \\
&-\sum_{cm}\sum_{n}\sum_{sp} p^{\varepsilon}_{cm,n,sp} \ln(p^{\varepsilon}_{cm,n,sp}) \\
&-\sum_{i}\sum_{ip}\sum_{cm}\sum_{cmp}\sum_{n}\sum_{sp} p^{\delta}_{i,ip,cm,cmp,n,sp} \ln(p^{\delta}_{i,ip,cm,cmp,n,sp})
\end{aligned}
$$

M13 is maximised subject to the constraints M1 to M12. The model was programmed with the software GAMS (General Algebraic Modeling System) and set up as a non-linear optimization problem. It was solved with the numerical solver CONOPT3.

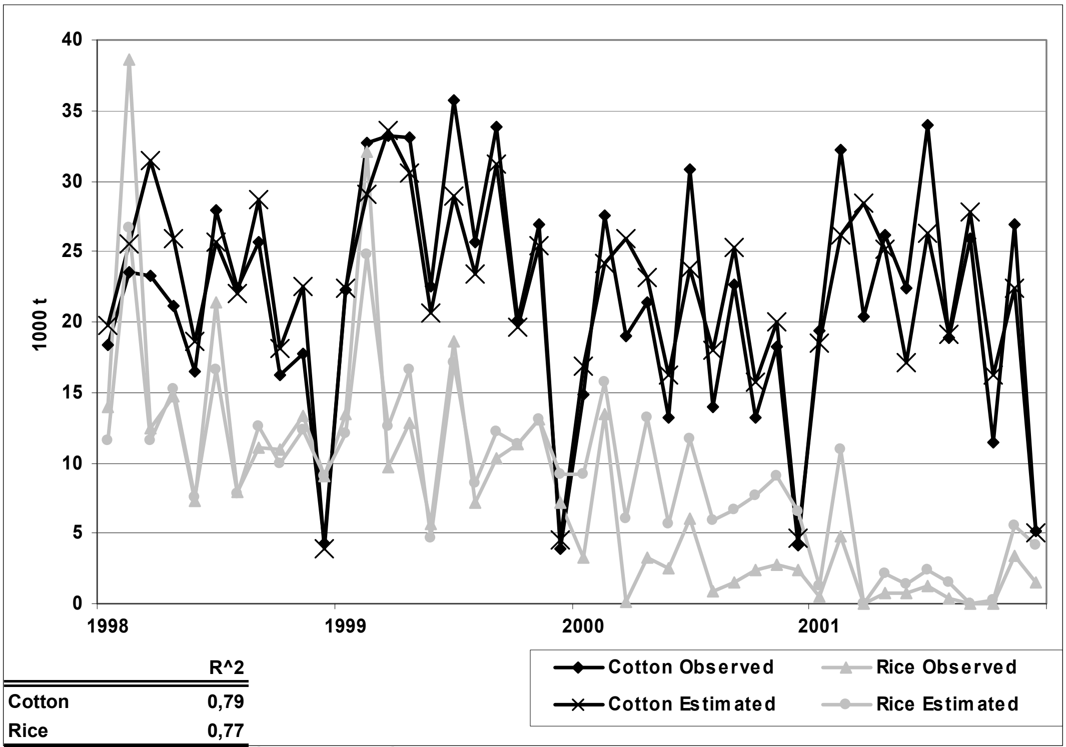
4. Results

The described model estimates the parameters of the production functions for each crop and the unknown allocated inputs water, diesel, and labor simultaneously. The accuracy of the model will be measured first according to its ability to replicate the known output quantities with the measurement of determination as indicator. The focus is here on the two main crops, cotton and rice, which together are planted on about 60% of the total crop area in Khorezm. Further relevant indicators to check the plausibility of the model are the estimated input allocations and the behavior of the production functions. Finally, the estimated price ratios of the inputs will be discussed.

4.1. The Output Side

Figure 4 compares the estimated total output results for the two main crops cotton and rice, with the respective observations. The lines show a high goodness of fit with a R^2 of 0.79 for cotton and 0.77 for rice. Only in the case of rice in 2000 a significant deviation between observed and estimated values is reported.

Figure 4: Comparison of Results



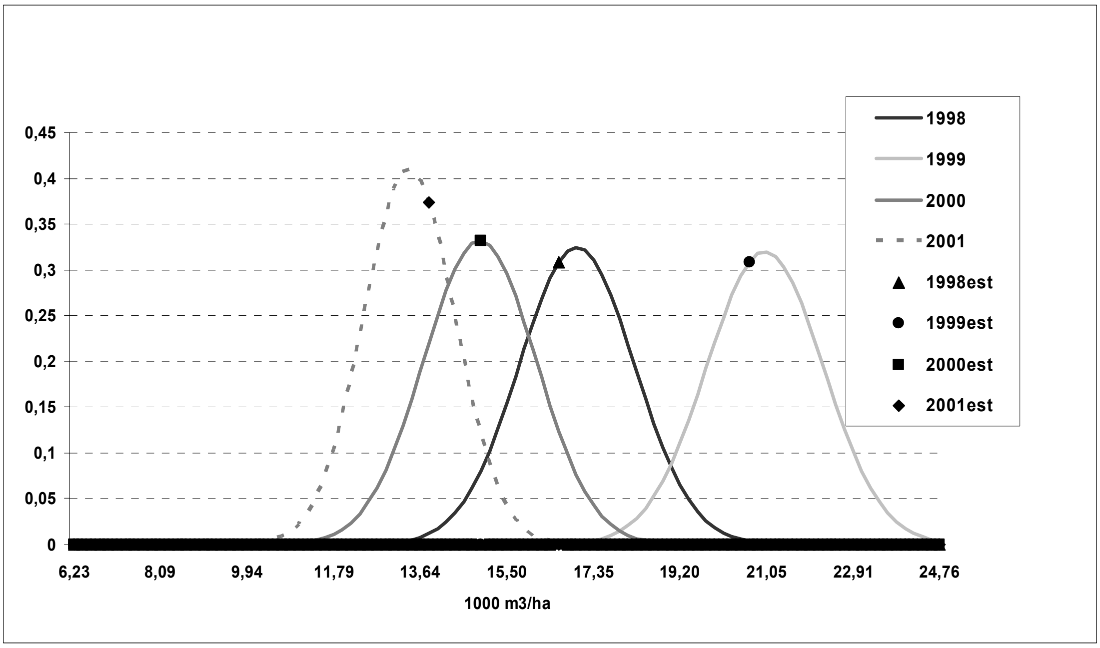
Source: OblStat and own results

The measurements of determination for other crops are not depicted for the sake of readability but are even higher in the case of wheat (0.95) and vegetable (0.94). Rice has the weakest results of all crops with the mentioned value of 0.77, which is due to the over-estimations in 2000. Taking the fit of the curves as an indicator for the ability of the model to replicate the observed values, it can be concluded that the model is well behaved.

4.2. The Input Side

Figure 5 depicts two different results for the estimation of inputs: The bell curves for the four years show the distribution functions for the water usage for cotton in the Rayon Khiva. The bell curves were derived from the results of the estimation-model E4 for expected values and variances of the input allocations. A normal distribution was assumed. The ranges of the mean values lie between 12000 m^3/ha in 2001 (the year with the highest water scarcity) and 21000 m^3/ha in 1999, which was a year with sufficient water supply. The support points for the ME estimation were chosen in a range of mean value plus and minus three times the standard deviation.

Figure 5: Water Allocations for Cotton in the Rayon Khiva [1000 m^3/ha]



Source: Own results

The black dots placed on the bell curves depict the final result for the input allocation. Neither in the case shown here nor in the other results for water, diesel, and labor do the final estimations deviate very much from the expected value of the estimated distribution function. Together with the satisfying results for the crop outputs it might be concluded that the estimates for the inputs also reveal the relatively high validity of the model.

4.3. Properties of the Production Functions

The parameters κ of the production functions are shown for the case of cotton and rice in table 3. In order to illustrate the resulting curvatures for cotton and rice, the respective isoquant curves for the inputs diesel and water are shown in figure 6. In both cases, water and diesel, which was used as proxy for physical capital usage in this study, are imperfect substitutes with a lesser substitutability in the case of rice. These results are plausible because it is possible to decrease water usage on irrigated fields with more extended use of machinery for the preparation of the fields (e.g. leveling). The latter, however, is rarely done because of high costs of capital input required while water in standard years is available in sufficient quantities.

Hence, an increased price ratio between water and capital would induce more careful usage of irrigation water.

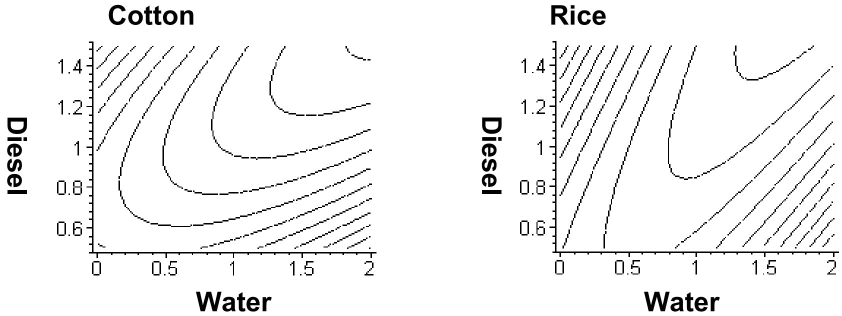
Table 3: Parameters of the Production Functions

		Cotton	Rice
Linear terms	Area	0,01	0,04
	Water	0,31	0,32
	Labor	0,04	0,02
	Diesel	0,73	0,78
Quadratic terms	Area, Area	-4,79	-3,52
	Water, Water	-1,05	-3,77
	Labor, Labor	-2,00	-4,30
	Diesel, Diesel	-2,91	-4,54
Cross terms	Area, Water	-1,58	-0,84
	Area, Labor	3,09	3,87
	Area, Diesel	3,30	0,39
	Water, Area	-1,58	-0,84
	Water, Labor	0,97	0,54
	Water, Diesel	1,61	4,10
	Labor, Area	3,09	3,87
	Labor, Water	0,97	0,54
	Labor, Diesel	-2,09	0,00
	Diesel, Area	3,30	0,39
	Diesel, Water	1,61	4,10
	Diesel, Labor	-2,09	0,00

Source: Own results

Rice depends more on water than cotton. Due to the fact, that the fields are flooded completely, proper leveling has a smaller water-saving effect. Thus, the observed difference between the isoquants appears to be plausible.

Figure 6: Isoquant Curves



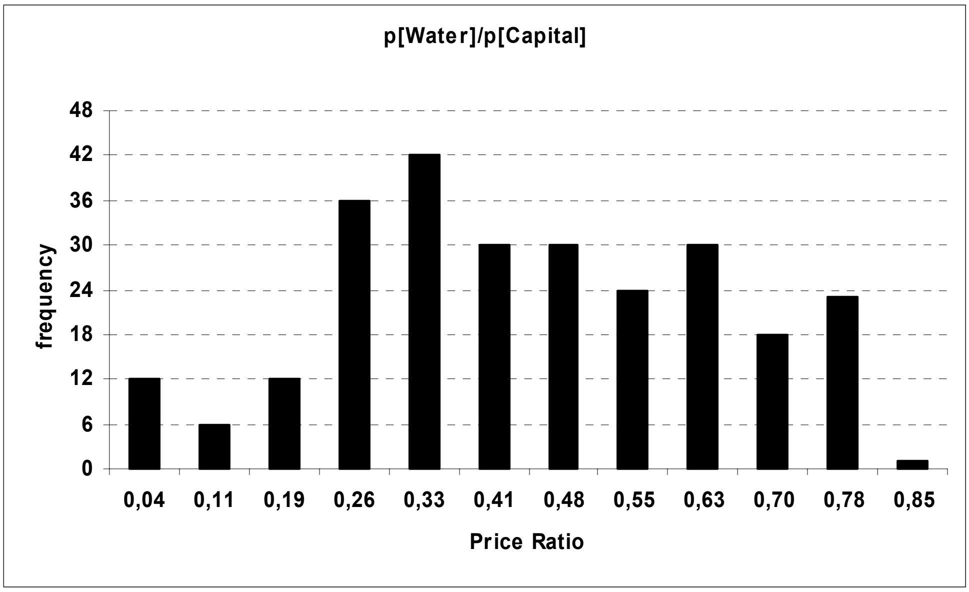
Source: Own results, input values are shown as deviations from their means

4.4. Price Ratios

Among the most interesting results of the estimation are the marginal rates of substitution between pairs of inputs, which are imposed to be equal in all observed points for all crops. Under the assumption of cost-minimizing production, these values correspond with the price ratios of the inputs. Even if there is no market price for land and water and it is difficult to find a 'price' for aggregated physical capital (for which diesel was used here as proxy), the price ratios can be interpreted as the costs the producers have to cover to get the resources needed.

Figure 7 shows the resulting price ratios for water and capital. The values range between 0.04 and 0.85 with the highest frequency for 0.33. The costs to get access to physical capital are therefore three times higher than the costs to access water, which explains the observation of excessive water use and the comparatively low machinery-use for preparing the fields prior to the irrigation process: It simply seems to be cheaper to irrigate the fields with more water than to hire a bulldozer to level it properly.

Figure 7: Price Ratios for Water and Capital



Source: Own results

The implication of this particular result is that it might be possible to reduce the water usage in Khorezm by improving the access to appropriate capital (e.g. diesel and machinery).

5. Conclusion

The study presented here described a model that allows estimating activity-specific input usage and parameters for a set of production functions simultaneously. The proposed method is a maximum entropy procedure, the needed support points for the inputs were estimated based on information about norm values, allocated area and total available quantities of water, labor and diesel.

It was possible to derive price ratios for major inputs which have no observable market price. It appeared that machinery is relatively more expensive as if compared with water. This result might be taken into account while thinking about possible policies to decrease the extensive use of irrigation water in the case-study region Khorezm: In the context of an ongoing discussion about the introduction of water prices and thus aggravating the economic pressure on farmers, it could be more helpful to think about policies that would facilitate the access to adequate machinery.

References:

Diewert, W. E., T. J. Wales (1987): Flexible Functional Forms and Global Curvature Conditions. In: Econometrica (55)

Fuss, M., D. McFadden, and Y. Mundlak (1978): A Survey of Functional Forms in the Economic Analysis of Production. In: Fuss, M. and McFadden, D. (Edts.): Production Economics: A Dual Approach to Theory and Applications. Volume I: The Theory of Production, Amsterdam.

Golan, A., G. Judge, and D. Miller (1996): Maximum entropy econometrics: robust estimation with limited data. New York.

Greene, W. H. (2003): Econometric Analysis, fifth edition, New York.

Just, R. E., D. Zilberman, E. Hochman, and Z. Bar-Shira (1990): Input Allocation in Multicrop Systems. In: American Journal of Agricultural Economics (72).

Lau, L. J. (1978): Testing and Imposing Monotonicty, Convexity and Quasi-Convexity Constraints. In: Fuss, M. and McFadden, D. (Edts.): Production Economics: A Dual Approach to Theory and Applications. Volume I: The Theory of Production, Amsterdam.

Lence, S. H., D. J. Miller (1998): Estimation of Multi-output Production Functions with Incomplete Data: A Generalised Maximum Entropy Approach, in: European Review of Agricultural Economics (25).

Data Sources:

OblStat: Department for Macro-Economics and Statistics in Khorezm Oblast: Agricultural Indicators for Khorezm Oblast, 1998-2001, Urgench, 2002.

OblSelVodKhos: Khorezm Department of Land and Water Ressources of Ministry of Agriculture and Water Resources: Data on Norms, Limits and Usage of Irrigation Water from 1997 to 2001, Urgench 2002.