University of Groningen

Latent vs. Observed Variables

Tang, Jianjun; Folmer, Henk

Published in:
Journal of Agricultural Economics

DOI:
[10.1111/1477-9552.12137](https://doi.org/10.1111/1477-9552.12137)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2016

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Tang, J., & Folmer, H. (2016). Latent vs. Observed Variables: Analysis of Irrigation Water Efficiency Using SEM and SUR. *Journal of Agricultural Economics*, 67(1), 173-185. <https://doi.org/10.1111/1477-9552.12137>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Latent vs. Observed Variables: Analysis of Irrigation Water Efficiency Using SEM and SUR

Jianjun Tang and Henk Folmer¹

(Original submitted April 2014, revision received February 2015, accepted April 2015.)

Abstract

In this paper we compare conceptualising single factor technical and allocative efficiency as indicators of a single latent variable, or as separate observed variables. In the former case, the impacts on both efficiency types are analysed by means of structural equation modeling (SEM), in the latter by seemingly unrelated regression (SUR). We compare estimation results of the two approaches based on a dataset on single factor irrigation water use efficiency obtained from a survey of 360 farmers in the Guanzhong Plain, China. The main methodological findings are that SEM allows identification of the most important dimension of irrigation water efficiency (technical efficiency) via comparison of their factor scores and reliability. Moreover, it reduces multicollinearity and attenuation bias. It thus is preferable to SUR. The SEM estimates show that perception of water scarcity is the most important positive determinant of both types of efficiency, followed by irrigation infrastructure, income and water price. Furthermore, there is a strong negative reverse effect from efficiency on perception.

Keywords: *Guanzhong; irrigation water allocative efficiency; irrigation water technical efficiency; latent variable; perception of water scarcity; seemingly unrelated regressions; structural equation modeling.*

JEL classifications: *C30, C51, C52, Q15.*

1. Introduction

The definition of economic efficiency was introduced by Farrell (1957) and has triggered extensive theoretical and methodological research on its measurement ever

¹Jianjun Tang is with the Centre for Public Health/Gibson Institute for Land, Food and Environment, Queen's University Belfast, United Kingdom. E-mail: j.tang@qub.ac.uk for correspondence. Henk Folmer is with the Department of Agricultural Economics, College of Economics and Management, Northwest A&F University, Yangling, Shaanxi, China and with the Faculty of Spatial Sciences, Groningen University, Groningen, The Netherlands. The authors gratefully acknowledge the comments and suggestions of the Editor and two referees which have helped improve the quality of the paper.

since (Kopp, 1981; Kumbhakar and Lovell, 2000; Greene, 2008; among others). According to Farrell, overall efficiency, later renamed economic efficiency, is a composite concept made up of technical and allocative efficiency. The former is a decision maker's ability to produce a given amount of output using the minimum amount of inputs while the latter is the ability to choose the optimal set of inputs to produce a given output, given prices. Single factor allocative and technical efficiency are defined by analogy with the overall measures (Kopp, 1981; Reinhard *et al.*, 2000; Tang *et al.*, 2015).

From the above definitions, it follows that efficiency is a theoretical construct or latent variable that can only be measured indirectly via observed behaviour, though with measurement error (Folmer and Oud, 2008; Oud and Folmer, 2008). Most researchers tend to ignore that efficiency is a latent variable and focus their theoretical and empirical analyses on the observable indicators instead. To the best of our knowledge, only three papers have treated efficiency as a latent variable. Following Bollen (1989), Kalaitzandonakes *et al.* (1992) analysed the relationship between farm size and technical efficiency by means of a structural equation model with latent variables (SEM) of the Multiple Indicator Multiple Cause (MIMIC) type (Jöreskog and Goldberger, 1975; Bollen, 1989). The authors consider three indicators of the latent variable technical efficiency, viz. technical efficiency scores derived through (i) deterministic parameter frontier analysis, (ii) stochastic parameter frontier analysis, and (iii) non-parametric frontier analysis.² Kalaitzandonakes and Dunn (1995) used the same three indicators to analyse the relationship between education and technical efficiency. Richards and Jeffrey (2000) also applied the MIMIC model to analyse the factors that contribute to economic performance. Their indicators of the latent variable performance are allocative, technical and overall efficiency.³

These three papers did not compare the performance of SEM to alternative (conventional) estimation procedures that do not treat efficiency as a latent variable. A typical alternative approach when there are common factors influencing the disturbances in the different equations is seemingly unrelated regression (SUR). The purpose of the present paper is to compare SEM and SUR, and the underlying conceptualisations of technical and allocative efficiency as indicators of a latent variable, versus treating them as separate observed variables. For this purpose we analyse a dataset on irrigation water use efficiency obtained from a survey of 360 farmers in the Guanzhong Plain, China (Tang *et al.*, 2015). This study focuses on single factor (irrigation water) efficiency which is the ratio of the minimum feasible use of an input and actual input, keeping other inputs and output constant. Single factor irrigation water technical efficiency has been analysed by Lilienfeldt and Asmild (2007) and Frija *et al.* (2009). Single-factor allocative efficiency, however, has received little attention in the literature. It is the ratio between the optimal cost (when all inputs are technically and allocatively efficient) and the cost when the single-factor is technically efficient, keeping other inputs constant (Kopp, 1981; and Kopp and Diewert, 1982).

²The authors found that the technical efficiency rankings based on the three indicators were not robust to estimation procedure. They ascribed this finding to the fact that the alternative technical efficiency measures had been measured with errors. They turned to SEM to take measurement error into account and to reconcile the conflicting outcomes.

³Note that the three indicators measure different, though strongly related, dimensions of efficiency. Therefore, they should not be analysed in isolation from each other, but in relation to each other.

Both single-factor (irrigation water) technical and allocative efficiency are considered in this paper.

The organisation of the paper is as follows. Section 2 presents the conceptual model including a brief summary of SEM. Section 3 discusses the data and the empirical results. Section 4 concludes.

2. The Conceptual Model

Before turning to the conceptual model, we present a brief summary of SEM which can handle latent variables and their indicators within one model framework. A SEM consists of two sub-models: two measurement models (equations (1) and (2) below) and a structural model (equation (3)) (Jöreskog, 1977; Jöreskog and Sörbom, 2001). The measurement model specifies the relationship between the latent variables and their observed indicators while the structural model represents the relationships between the latent exogenous and the latent endogenous variables as well as the relationships among the latent endogenous variables mutually. Specifically:

$$\mathbf{y} = \Lambda_y \eta + \varepsilon \quad (1)$$

$$\mathbf{x} = \Lambda_x \zeta + \delta \quad (2)$$

$$\eta = \mathbf{B}\eta + \Gamma\zeta + \zeta \quad (3)$$

where $\mathbf{y}$ and $\mathbf{x}$ are $p \times 1$ and $q \times 1$ vectors of endogenous and exogenous observed variables, respectively. Λ_y and Λ_x are $p \times m$ and $q \times n$ matrices of regression coefficients or loadings on the $m \times 1$ and $n \times 1$ vectors of latent endogenous and latent exogenous variables, η and ζ , respectively. The element β_{ij} of the $m \times m$ matrix $\mathbf{B}$ represents the effect of the j th endogenous latent variable on i th endogenous latent variable, and the element γ_{ij} of the $m \times n$ matrix Γ the effect of the j th exogenous latent variable on i th endogenous latent variable. Finally, ε and δ are $p \times 1$ and $q \times 1$ vectors of measurement errors of $\mathbf{y}$ and $\mathbf{x}$, respectively,⁴ and ζ is the $m \times 1$ vector of structural errors. For identification, estimation, testing and modification indices we refer to Jöreskog and Sörbom (2001).

Folmer and Oud (2008) discuss the theoretical and empirical advantages of using SEM. They show that SEM allows a closer correspondence between theory and empirics than a model in observables only. Furthermore, SEM decomposes the variance of an observed variable into the variance of the true (latent) variable and the variance of the measurement error, as shown in equations (1) and (2). It thus reduces measurement error problems. (Note that measurement error in an explanatory variable increases the error variance of the corresponding structural equation, if it is uncorrelated with the observed explanatory variables, and to attenuation bias, if it is correlated. Furthermore, if there is measurement error in the dependent variable, the estimator is biased, if the error is correlated with the explanatory variables; if it is uncorrelated with the explanatory variables, it will lead to larger

⁴Directly observed variables can be conveniently handled in the SEM framework by specifying an identity relationship in the measurement model between an observed variable and the corresponding latent variable and fixing the measurement error at zero (see, for example, equation (5)).

Figure 1. Path diagram of the structural equation modeling

structural error variance.) The use of SEM may also reduce multicollinearity (Folmer, 1981). Finally, SEM provides information about the weight or impact of the indicators on the corresponding latent variable via the loadings and on their reliability (R^2).

The conceptual SEM and SUR of the irrigation water problem analysed in this paper are presented in Figures 1⁵ and 2, and equations (4)–(7) and (8)–(11), respectively. *IWTE* and *IWAE* are observed irrigation water technical efficiency and irrigation water allocative efficiency, respectively, obtained from a stochastic frontier analysis (Tang *et al.*, 2015). In SUR they are treated as separate endogenous variables whose disturbances are correlated. In SEM, *IWTE* and *IWAE* are taken as indicators of the latent variable *Efficiency*. In both SUR and SEM, the efficiency variables are functions of a latent variable, *Perception* (farmer's perception of water scarcity), the farmer characteristics *Age*, *Education*, *Income* and *Time* (the proportion of time a farmer spends on farming), the farm characteristics *Fragmentation* (number of different plots) and *Infrastructure* (dummy variable that takes the value 1 if the irrigation canal is cement, and 0 if it is earthen-lined) and the variables *Waterprice* and *Precipitation*. SUR models are typically presented in terms of observed variables; the use of latent variables is unusual in SUR models. We nevertheless use the latent variable *Perception* in the SUR model to maximise its

⁵A one-way arrow in a path diagram indicates a causal relationship. Hence, for the measurement models, an arrow from a latent variable to an indicator means that the former causes the indicator. It is possible to obtain the impacts of the indicators *Percep1*, *Percep2* and *Percep3* on *IWTE* and *IWAE* by specifying them as observed explanatory variables (like income and age). This approach is adopted if the interest is in the impact of an observable on an endogenous variable. If the interest primarily is in the impacts of the latent variable, the specifications in Figure 1 and equations (4)–(7) are used. The impacts of the indicators of a latent variable, say *Perception*, on an indicator, say *IWAE*, equals the product of the impact of *Perception* on *Efficiency* ($\beta_{1,2}$) and the impact of *Efficiency* on *IWAE* ($\lambda_{2,1}$).

Figure 2. Path diagram of the seemingly unrelated regression

similarity to the SEM model. Differences between SEM and SUR can thus be attributed to the way efficiency is modeled. The latent variable *Perception* is measured by the following three indicators: *Percep1* (irrigation water is scarce in my village), *Percep2* (irrigation water scarcity is worse now than it was before) and *Percep3* (irrigation water will be scarcer in the next 2 years than it is now). Each indicator is measured at a 5-point scale ranging from fully disagree to fully agree. Note that SUR is a special case of SEM. Hence, it can be specified as a SEM and estimated by SEM software packages.

In terms of equations (1)–(3), the SEM (Figure 1) is expressed as in (4)–(7) and the SUR in (8)–(11).

$$\begin{bmatrix} IWTE \\ IWAE \\ Percep1 \\ Percep2 \\ Percep3 \end{bmatrix} = \begin{bmatrix} \lambda_{1,1} & 0 \\ \lambda_{2,1} & 0 \\ 0 & \lambda_{3,2} \\ 0 & \lambda_{4,2} \\ 0 & \lambda_{5,2} \end{bmatrix} \begin{bmatrix} Efficiency \\ Perception \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \end{bmatrix} \quad (4)$$

$$\begin{bmatrix} Age \\ Time \\ Fragmentation \\ Infrastructure \\ Income \\ Education \\ Water price \\ Precipitation \end{bmatrix} = \begin{bmatrix} 10000000 \\ 01000000 \\ 00100000 \\ 00010000 \\ 00001000 \\ 00000100 \\ 00000010 \\ 00000001 \end{bmatrix} \begin{bmatrix} Age \\ Time \\ Fragmentation \\ Infrastructure \\ Income \\ Education \\ Water price \\ Precipitation \end{bmatrix} \quad (5)$$

$$\begin{aligned}
\begin{bmatrix} \text{Efficiency} \\ \text{Perception} \end{bmatrix} &= \begin{bmatrix} 0 & \beta_{1,2} \\ \beta_{2,1} & 0 \end{bmatrix} \begin{bmatrix} \text{Efficiency} \\ \text{Perception} \end{bmatrix} \\
&+ \begin{bmatrix} \gamma_{1,1} & \gamma_{1,2} & \gamma_{1,3} & \gamma_{1,4} & \gamma_{1,5} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \gamma_{2,5} & \gamma_{2,6} & \gamma_{2,7} & \gamma_{2,8} \end{bmatrix} \begin{bmatrix} \text{Age} \\ \text{Time} \\ \text{Fragmentation} \\ \text{Infrastructure} \\ \text{Income} \\ \text{Education} \\ \text{Water price} \\ \text{Precipitation} \end{bmatrix} \\
&+ \begin{bmatrix} \zeta_1 \\ \zeta_2 \end{bmatrix}
\end{aligned} \tag{6}$$

$$\Psi = \begin{pmatrix} \psi_{11} & 0 \\ 0 & \psi_{22} \end{pmatrix}. \tag{7}$$

The first two measurement equations in equation (4) link the indicators *IWTE* and *IWAE* to the latent variable *Efficiency*. In a similar vein, measurement equations 3–5 in model (4) relate *Percep1*, *Percep2* and *Percep3* to the latent variable *Perception*. Measurement model (5) is in terms of observed variables only which means that the observed variables are identical to the corresponding latent variables and that the measurement errors are equal to zero. The first structural equation in (6) specifies *Efficiency* as a function of the endogenous variable *Perception* and of the exogenous variables *Age*, *Time*, *Fragmentation*, *Infrastructure* and *Income*. Similarly, the second structural equation specifies *Perception* as a function of *Efficiency* and of the exogenous variables *Income*, *Education*, *Water price* and *Precipitation*. In matrix (7), ψ_{11} and ψ_{11} represent variances of ζ_1 and ζ_2 , respectively.

Instead of constructing a latent endogenous variable *Efficiency*, SUR assumes two observed endogenous variables, i.e. *IWTE* and *IWAE*. The first two equations in (8) indicate that the two variables are identical to their corresponding latent variables and that the measurement errors are zeros. The measurement equations 3–5 in (8) are similar to those in model (4). In similar vein, equations (9) are identical to (5). Since SUR has three endogenous variables (SEM has two), it has three equations in (10). The *IWTE* and *IWAE* are functions of the same set of variables as *Efficiency* in the SEM. Another difference between SEM and SUR is that the structural errors ζ_1 and ζ_2 are correlated so that the off-diagonal element ψ_{21} in the 3×3 covariance matrix Ψ is non-zero. It is the covariance between the errors of the structural models ζ_1 and ζ_2 .

$$\begin{bmatrix} \text{IWTE} \\ \text{IWAE} \\ \text{Percep1} \\ \text{Percep2} \\ \text{Percep3} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \lambda_{3,3} \\ 0 & 0 & \lambda_{4,3} \\ 0 & 0 & \lambda_{5,3} \end{bmatrix} \begin{bmatrix} \text{IWTE} \\ \text{IWAE} \\ \text{Perception} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \end{bmatrix} \tag{8}$$

$$\begin{bmatrix} \text{Age} \\ \text{Time} \\ \text{Fragmentation} \\ \text{Infrastructure} \\ \text{Income} \\ \text{Education} \\ \text{Water price} \\ \text{Precipitation} \end{bmatrix} = \begin{bmatrix} 10000000 \\ 01000000 \\ 00100000 \\ 00010000 \\ 00001000 \\ 00000100 \\ 00000010 \\ 00000001 \end{bmatrix} \begin{bmatrix} \text{Age} \\ \text{Time} \\ \text{Fragmentation} \\ \text{Infrastructure} \\ \text{Income} \\ \text{Education} \\ \text{Water price} \\ \text{Precipitation} \end{bmatrix} \quad (9)$$

$$\begin{bmatrix} \text{IWTE} \\ \text{IWAE} \\ \text{Perception} \end{bmatrix} = \begin{bmatrix} 0 & 0 & \beta_{1,3} \\ 0 & 0 & \beta_{2,3} \\ \beta_{3,1} & \beta_{3,2} & 0 \end{bmatrix} \begin{bmatrix} \text{IWTE} \\ \text{IWAE} \\ \text{Perception} \end{bmatrix} \\
 + \begin{bmatrix} \gamma_{1,1} & \gamma_{1,2} & \gamma_{1,3} & \gamma_{1,4} & \gamma_{1,5} & 0 & 0 & 0 \\ \gamma_{2,1} & \gamma_{2,2} & \gamma_{2,3} & \gamma_{2,4} & \gamma_{2,5} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \gamma_{3,5} & \gamma_{3,6} & \gamma_{3,7} & \gamma_{3,8} \end{bmatrix} \begin{bmatrix} \text{Age} \\ \text{Time} \\ \text{Fragmentation} \\ \text{Infrastructure} \\ \text{Income} \\ \text{Education} \\ \text{Water price} \\ \text{Precipitation} \end{bmatrix} \\
 + \begin{bmatrix} \zeta_1 \\ \zeta_2 \\ \zeta_3 \end{bmatrix} \quad (10)$$

$$\Psi = \begin{pmatrix} \psi_{11} & 0 & 0 \\ \psi_{21} & \psi_{22} & 0 \\ 0 & 0 & \psi_{33} \end{pmatrix} \quad (11)$$

3. Data and Empirical Results

We analyse a dataset of 360 farmers in the Guanzhong Plain, China. The data were collected via face-to-face interviews in October 2011. Details on sampling and descriptive statistics are shown in Table 1.

Before comparing SEM and SUR and discussing the estimates, we make the following remarks. First, a measurement scale was assigned to the latent variables (which is a prerequisite for identification) by fixing their variances (at 1). Secondly, we standardised the coefficients. Hence, a coefficient represents the standard deviation change in a dependent variable due to a standard deviation change in the corresponding explanatory variable, which makes the coefficients directly comparable. Thirdly, because of the presence of ordinal variables, we analysed a polychoric correlation matrix. Finally, we estimated the models by means of LISREL 8.8 (Jöreskog and Sörbom, 2001).

Table 1
Descriptive statistics

Variable	Unit of measurement	Min.	Max.	Mean	SD
<i>IWTE</i>	—	0.06	0.76	0.35	0.14
<i>IWAE</i>	—	0.64	0.98	0.86	0.07
<i>Percep1</i>	—	1	5	2.61	1.51
<i>Percep2</i>	—	1	5	2.77	1.31
<i>Percep3</i>	—	1	5	3.08	1.07
<i>Age</i>	Years	26	77	53.08	10.07
<i>Education</i>	Years	0	12	6.64	1.70
<i>Income</i>	Yuan	1,000	195,000	26,941	25,103
<i>Time</i>	—	1	5	3.80	1.44
<i>Water price</i>	Yuan/m ³	0.02	1.00	0.32	0.13
<i>Fragmentation</i>	—	1	15	2.74	1.75
<i>Precipitation</i>	Mm	137	220	174.31	17.07
<i>Infrastructure</i>	—	0	1	0.52	0.49

Table 2
Goodness of fit statistics

	χ^2	NFI	GFI	AGFI	RMSEA
<i>SEM</i>	46.83 (df = 34, $P = 0.07$)	0.900	0.980	0.945	0.033
<i>SUR</i>	34.54 (df = 28, $P = 0.18$)	0.926	0.985	0.951	0.026

Notes: The cut-off values for NFI, GFI, AGFI and RMSEA indicating a good fit are 0.90, 0.95, 0.90 and 0.06, respectively (Hooper *et al.*, 2008). The higher the NFI, GFI and AGFI values are and the smaller the RMSEA is, the better is the fit.

3.1. *SEM* vs. *SUR*

Model fit statistics are presented in Table 2. They indicate good fit for both models with the *SUR* statistics slightly better than the *SEM* statistics. Table 3 shows that the standardised coefficients of the indicators of *Perception* in both *SEM* and *SUR* are significant. The differences between both models are only minor. Moreover, the reliabilities (R^2) are larger than the recommended minimum level of 0.20 (Jöreskog and Sörbom, 2001) indicating that the three indicators measure *Perception* well. The most reliable indicator is *Percep1*, followed by *Percep2* and *Percep3*. Table 3 also shows that *IWTE* and *IWAE* are significant indicators of *Efficiency* with the former the most important and most reliable. Particularly, the factor loading of *IWTE* is 0.78, which is substantially larger than that of *IWAE* (0.37). This result indicates that irrigation water efficiency is more ‘quantity’ (technical efficiency) than ‘cost’ (allocative efficiency) oriented which is probably due to the sub-optimal irrigation water price (Tang *et al.*, 2015). As observed in section 2, the possibility of determining the importance of *IWTE* and *IWAE* via comparison of their loadings and reliabilities is one of the

Table 3
Measurement model (standardised coefficients)

Variables	Indicators	SEM			SUR		
		Coeff.	t-value	R ²	Coeff.	t-value	R ²
<i>Perception</i>	<i>Percep1</i>	0.82	8.07	0.69	0.83	6.68	0.66
	<i>Percep2</i>	0.54	7.19	0.28	0.53	6.17	0.29
	<i>Percep3</i>	0.48	6.70	0.22	0.47	5.83	0.23
<i>Efficiency</i>	<i>IWTE</i>	0.78	4.44	0.61	—	—	—
	<i>IWAE</i>	0.37	3.96	0.14	—	—	—

main advantages of SEM (compared to SUR and other methods that can only handle observables).⁶

We next compare the structural SEM and SUR models. Table 4 shows that *Efficiency* (SEM), and *IWTE* and *IWAE* (SUR) impact *Perception* negatively and significantly. In addition, the coefficient of *Efficiency* (−0.69) is larger than those of *IWTE* (−0.36) and *IWAE* (−0.57). This result is probably due to the fact that, in contrast to *IWTE* and *IWAE*, *Efficiency* has been purged of its measurement error, which reduces attenuation bias. For the reverse effects, from *Perception* on *Efficiency* and *IWTE*/*IWAE*, we also find a difference between SEM and SUR. Specifically, in SEM *Perception* impacts *Efficiency* positively and significantly (0.52). In SUR, the impact of *Perception* on *IWAE* is also positive and significant. However, its impact on *IWTE* is insignificant, though positive.

3.2. Determinants of Efficiency

Since SEM is more informative than SUR, we only discuss the estimated SEM structural model and total effects. The SUR model is presented for comparison; for a discussion of it we refer to Tang *et al.* (2015). Table 4 shows that the coefficient of *Perception* (0.52) is significant, indicating that the more a farmer perceives water scarcity, the more efficiently he or she uses irrigation water. *Time* has a negative and significant impact which is possibly due to excess irrigation by full-time farmers. *Fragmentation* (−0.14) impacts *Efficiency* negatively and significantly, possibly due to loss of water during conveyance. Furthermore, fragmented plots are not suitable for advanced water-saving techniques. *Infrastructure*, the presence of cement-lined canals (0.25) improves *Efficiency* because of less water leakage. Finally, *Income* increases *Efficiency* because richer farmers can afford more or more adequate water-saving techniques than poor farmers.

Efficiency has a negative and significant impact on *Perception* indicating that an efficient farmer perceives less water scarcity than an inefficient farmer. A possible explanation is that an efficient farmer expects efficiency to reduce the water scarcity problem. The positive and significant coefficient of *Water price* (0.19) indicates that price induces perception of water scarcity.

⁶Testing of the equality of the loadings of *IWTE* and *IWAE* is straightforward (Jöreskog and Sörbom, 2001).

Table 4
The estimated SEM and SUR (standardised coefficients)

Variables	SEM		SUR		
	<i>Perception</i>	<i>Efficiency</i>	<i>Perception</i>	<i>IWTE</i>	<i>IWAE</i>
<i>IWTE</i>		–	–0.36* (0.20)	–	–
<i>IWAE</i>		–	–0.57*** (0.19)	–	–
<i>Efficiency</i>	–0.69*** (0.23)	–	–	–	–
<i>Perception</i>	–	0.52** (0.24)	–	0.33 (0.25)	0.77** (0.36)
<i>Age</i>	–	0.09 (0.08)	–	0.07 (0.06)	0.09 (0.08)
<i>Education</i>	0.06 (0.07)	–	0.04 (0.07)	–	–
<i>Time</i>	–	–0.17* (0.09)	–	–0.14** (0.07)	–0.10 (0.08)
<i>Fragmentation</i>	–	–0.14* (0.08)	–	–0.11 (0.07)	–0.14* (0.08)
<i>Infrastructure</i>	–	0.25*** (0.08)	–	0.17** (0.07)	0.21** (0.09)
<i>Income</i>	–0.07 (0.08)	0.20** (0.08)	–0.10 (0.07)	0.15** (0.07)	0.16* (0.08)
<i>Waterprice</i>	0.19*** (0.07)	–	0.18*** (0.07)	–	–
<i>Precipitation</i>	–0.24*** (0.07)	–	–0.30*** (0.07)	–	–
<i>R</i> ²	0.49	0.23	0.50	0.13	0.39

Note: Standard errors are in parentheses. *Significant at 10%; **Significant at 5%; ***Significant at 1%.

Table 4 gives an incomplete overview of the impacts because the indirect effects are not taken into account. Therefore, we present in Table 5 the total effects (i.e. the sum of the direct effect – the coefficients in Table 4 – and the indirect effects) of all variables on *Perception*, *Efficiency* and *IWTE* and *IWAE*. The variable with the largest total impact (in absolute value) on *Perception* is *Efficiency* (–0.51) followed by *Precipitation* (–0.18), *Income* (–0.15), *Waterprice* (0.14) and *Infrastructure* (–0.13). *Perception* also has a large impact (–0.26) on itself because of reciprocal causation between *Perception* and *Efficiency*. The total impacts of the other variables are small or insignificant. The most important determinant of *Efficiency* is *Perception* (0.38), followed by *Infrastructure* (0.18), *Income* (0.12) and *Water price* (0.07). The main determinants of *IWTE* are *Efficiency* (0.57), *Perception* (0.30) and *Infrastructure* (0.14). The results for *IWAE* are similar, though the total effects are substantially smaller due to the smaller impact of *Efficiency*.

Table 5
Estimated total effects (standardised coefficients)

Variables	SEM				SUR		
	Perception	Efficiency	IWTE	IWAE	Perception	IWTE	IWAE
<i>IWTE</i>	–	–	–	–	–0.23* (0.14)	–0.08 (0.09)	–0.18* (0.09)
<i>IWAE</i>	–	–	–	–	–0.37*** (0.08)	–0.12 (0.07)	–0.28* (0.16)
<i>Perception</i>	–0.26* (0.18)	0.38*** (0.10)	0.30** (0.14)	0.14** (0.06)	–0.36** (0.15)	0.21 (0.13)	0.50*** (0.12)
<i>Efficiency</i>	–0.51*** (0.08)	–0.26* (0.15)	0.57*** (0.10)	0.27*** (0.06)	–	–	–
<i>Age</i>	–0.05 (0.04)	0.07 (0.07)	0.05 (0.05)	0.02 (0.02)	–0.05 (0.04)	0.05 (0.06)	0.06 (0.05)
<i>Time</i>	0.09* (0.05)	–0.13* (0.07)	–0.10* (0.05)	–0.05* (0.03)	0.07* (0.04)	–0.11** (0.06)	–0.05 (0.05)
<i>Fragmentation</i>	0.07 (0.04)	–0.10* (0.06)	–0.08* (0.05)	–0.04 (0.02)	0.08** (0.04)	–0.08 (0.05)	–0.08* (0.05)
<i>Infrastructure</i>	–0.13*** (0.05)	0.18*** (0.07)	0.14*** (0.05)	0.07** (0.03)	–0.12*** (0.04)	0.13** (0.05)	0.12** (0.05)
<i>Income</i>	–0.15** (0.06)	0.12* (0.07)	0.10* (0.05)	0.05 (0.03)	–0.15** (0.06)	0.10* (0.05)	0.04 (0.05)
<i>Education</i>	0.04 (0.05)	0.02 (0.03)	0.02 (0.02)	0.01 (0.01)	0.02 (0.04)	0.01 (0.02)	0.02 (0.03)
<i>Water price</i>	0.14** (0.06)	0.07** (0.03)	0.06* (0.03)	0.03* (0.02)	0.12** (0.05)	0.04 (0.03)	0.09** (0.04)
<i>Precipitation</i>	–0.18*** (0.06)	–0.09*** (0.03)	–0.07** (0.04)	–0.03* (0.02)	–0.19* (0.07)	–0.06 (0.04)	–0.15*** (0.05)

Notes: Standard errors are in parentheses. *Significant at 10%; **Significant at 5%; ***Significant at 1%.

4. Conclusions

The purpose of this paper is to compare two different ways of conceptualising technical and allocative efficiency – as indicators of a single latent variable or as separate observed variables and the related modeling approaches, i.e. structural equation modeling (SEM) and SUR. The main conclusion is that the conceptualisation and the modeling approach affect the estimation results. This conclusion is in line with Kalaitzandonakes *et al.* (1992) who found that technical efficiency ranking of farms may not be robust to estimation procedure. A related finding of this paper is that the estimated coefficients are not robust to efficiency measure analysed, which was also observed by Kalaitzandonakes and Dunn (1995).

Kalaitzandonakes *et al.* (1992) recommend SEM to reconcile dissimilar efficiency scores by taking them as indicators of a latent variable and to then apply SEM. The findings of this paper confirm this recommendation. We furthermore summarise several additional reasons to apply SEM including a closer correspondence between theory and empirics, accounting for measurement error and reduction of multicollinearity. SEM also allows comparison of the loadings and

reliabilities of the indicators of a latent variable which made it possible to identify the most important and reliable indicator of efficiency. We thus found that irrigation water efficiency is more 'quantity' (technical efficiency) than 'cost' (allocative efficiency) oriented in the Guanzhong Plain. On the basis of these features, we suggest that SEM is preferable to SUR as a tool to analyse the determinants of efficiency.

The total effects calculated by SEM showed that perception of water scarcity, irrigation infrastructure, water price and income increase irrigation water efficiency while time spent on farming and fragmentation decreases it. Perception of water scarcity is strongly influenced by efficiency indicating that efficient farmers expect that efficiency may reduce irrigation water scarcity.

Further research on efficiency as a latent variable is needed. Such research could, *inter alia*, shed light on technical and allocative efficiency obtained from parametric and non-parametric methods as efficiency measures.

References

- Bollen, K. A. *Structural Equations with Latent Variables* (New York: John Wiley & Sons, 1989).
- Farrell, M. J. 'The measurement of productive efficiency', *Journal of the Royal Statistical Society*, Vol. 120, (1957) pp. 253–290.
- Folmer, H. 'Measurement of the effects of regional policy instruments by means of linear structural equation models and panel data', *Environment and Planning A*, Vol. 13, (1981) pp. 1435–1448.
- Folmer, H. and Oud, J. H. L. 'How to get rid of W: A latent variables approach to modelling spatially lagged variables', *Environment and Planning A*, Vol. 40, (2008) pp. 2526–2538.
- Frija, A., Chebil, A., Speelman, S., Buysse, J. and Van Huylenbroeck, G. 'Water use and technical efficiencies in horticultural greenhouses in Tunisia', *Agricultural Water Management*, Vol. 96, (2009) pp. 1509–1516.
- Greene, W. H. 'The econometric approach to efficiency analysis', in H. O. Fried, C. A. K. Lovell and S. S. Schmidt (eds.), *The Measurement of Productive Efficiency Growth* (New York: Oxford University Press, 2008, pp. 92–250).
- Hooper, D., Coughlan, J. and Mullen, M. R. 'Structural equation modelling: Guidelines for determining model fit', *The Electronic Journal of Business Research Methods*, Vol. 6, (2008) pp. 53–60.
- Jöreskog, K. G. 'Factor analysis by least-squares and maximum-likelihood methods', in K. Enslein, A. Ralston and H. S. Wilf, (eds.), *Statistical Methods for Digital Computers* (New York: John Wiley & Sons Inc., 1977, pp. 125–153).
- Jöreskog, K. G. and Goldberger, A. S. 'Estimation of a model with multiple indicators and multiple causes of a single latent variable', *Journal of the American Statistical Association*, Vol. 75, (1975) pp. 631–639.
- Jöreskog, K. G. and Sörbom, D. *LISREL 8: User's Reference Guide* (Chicago: Scientific Software International Inc, 2001).
- Kalaitzandonakes, N. and Dunn, E. G. 'Technical efficiency, managerial ability and farmer education in Guatemalan corn production: A latent variable analysis', *Agricultural and Resource Economics Review*, Vol. 24, (1995) pp. 36–46.
- Kalaitzandonakes, N., Wu, S. and Ma, J. 'The relationship between technical efficiency and firm size revisited', *Canadian Journal of Agricultural Economics*, Vol. 40, (1992) pp. 427–442.
- Kopp, R. J. 'The measurement of productive efficiency: Reconsideration', *Quarterly Journal of Economics*, Vol. 96, (1981) pp. 477–503.

- Kopp, R. J. and Diewert, W. E. 'The decomposition of frontier cost function deviations into measures of technical and allocative inefficiency', *Journal of Econometrics*, Vol. 19, (1982) pp. 319–331.
- Kumbhakar, S. and Lovell, K. *Stochastic Frontier Analysis* (Cambridge: Cambridge University Press, 2000).
- Lilienfeldt, A. and Asmild, M. 'Estimation of excess water use in irrigated agriculture: A Data Envelopment Analysis approach', *Agricultural Water Management*, Vol. 94, (2007) pp. 73–82.
- Oud, J. H. L. and Folmer, H. 'A structural equation approach to models with spatial dependence', *Geographical Analysis*, Vol. 40, (2008) pp. 152–166.
- Reinhard, S., Lovell, C. A. K. and Thijssen, G. 'Environmental efficiency with multiple environmentally detrimental variables: Estimated with SFA and DEA', *Europe Journal of Operational Research*, Vol. 121, (2000) pp. 287–303.
- Richards, T. J. and Jeffrey, S. R. 'Efficiency and economic performance: An application of the MIMIC model', *Journal of Agricultural and Resource Economics*, Vol. 25, (2000) pp. 232–251.
- Tang, J., Folmer, H. and Xue, J. 'Technical and allocative efficiency of irrigation water use in the Guanzhong Plain, China', *Food Policy*, Vol. 50, (2015) pp. 43–52.