Synthetic data. A proposed method for applied risk management

Carbajal De Nova, Carolina

Autonomous Metropolitan University

28 March 2017

Synthetic data. A proposed method for applied risk management

Working paper, March 28, 2017

Carolina Carbajal De Nova

Department of Economics, Autonomous Metropolitan University

enova@xanum.uam.mx

Abstract

The proposed method attempts to contribute towards the econometric and simulation applied risk management literature. It consists on an algorithm to construct synthetic data and risk simulation econometric models, supported by a set of behavioral assumptions. This algorithm has the advantage of replicating natural phenomena and uncertainty events in a short period of time. These features convey economically low costs besides computational efficiency. An application for wheat farmers is developed. The efficiency of this method is confirmed when its results and statistical inference converge with those generated from experimental data. Convergence is demonstrated specifically by means of information convergence and diminishing scaling variance. Modifications on the proposed algorithm regarding risk distribution parameters are not onerous. These modifications can generate diverse risk scenarios seeking to minimize and manage risk. Hence, risk sources could be anticipated, identified as well as quantified. The algorithm flexibility makes risk testing accessible to an ample variety of entrepreneurial problems *i.e.*, public health systems, farmers associations, hedge funds, insurance companies; etcetera. This method could provide grounded criteria for decision-making in order to improve management practices.

Keywords: behavioral assumptions, risk scenarios, simulation econometric models, synthetic data.

1. Introduction

In the field of applied economics there is a large literature devoted to study risk abatement in an entrepreneurial environment.[1] In this respect, utility functions for money in the context of risk premium are being widely used.[2] In particular, agriculture farming has a specialized literature devoted to output risk analysis and uncertainty modelling.[3] This is mainly due to unpredictable weather conditions, pest damages and unstable agricultural markets, besides uncertain output production to input shocks and risk attitudes.[4] In this strand, the majority of researchers use mathematical programming and numerical methods to model agricultural output risk and expected wealth.[5] This research basically implements the following methods: structural-form approach,[6] joint estimation of risk preference,[7] risk preference analysis,[8] as well as econometrics and simulation modelling.[9]

This paper belongs to the class of output risk analysis and uncertainty modelling literature, which relies on econometrics and simulation techniques. Basically, this modelling deals with probabilities measures and assessment of risk exposure by evaluating probability distributions of risky events.[10] This analysis boils down to the discrimination of sensitivity estimates and risk hypothesis, in an efficient manner. The proposed method contributes to extend this type of literature strand. This is because it synthetizes in a unique applied risk-analysis algorithm three fundamental risk management milestones: study design; agent's behavior econometric modelling and risk scenario sensitivity analysis. This algorithm addresses study design by constructing synthetic data.[11] This construction takes into account experimental data sampling properties, in order to produce "clones" out of them.[12] The production behavior is being introduced in the algorithm through a series of mathematical assumptions.[13] The

[1] For example, Bassi, Colacito and Fulghiery (2013); and Pope (1978); Melhim and Shumway (2010); Saha (1997); Saha, Malkiel and Grecu (2009); Saha, Shumway and Havenner (1997).

[2] For instance, Pratt (1964), Mehra and Prescott (1985) and Friend and Blume (1975).

[3] Saha (1993); Saha (1994); Saha, Love and Schwartz (1994); Saha, Shumway and Talpaz (1994).

[4] These points have being discussed on Chavas and Holt (1996).

[5] In this respect see Hazell (1971), and Simmons and Pomareda (1975).

[6] A comprehensive list of studies that have used the structural form approach related with the mean-variance framework could be found in Saha, Shumway and Talpaz (1994), Table 2.

[7] As in Saha, Shumway and Talpaz (1994), as well Chavas and Holt (1996).

[8] See Chavas and Holt (1990) for more information in this regard.

[9] The Farm Level Income and Policy Simulation Model (FLIPSIM) developed at the Texas A&M University does not use econometric equations. It resorts to identities and probability distributions.

[10] Chavas and Shi (2015) present a lucid discussion of these elements and how to use them to improve food sovereignty.

[11] The synthetic data term has being used broadly. In the case of Schneider and Abowd (2015) the term synthetic data nominates altered data, which protects the privacy rights of the constituent.

[12] The term "clones" is referred to the replication of distribution moments.

[13] The use of a set of behavioral assumptions is the key difference between this proposed method and current applied risk management literature. The behavioral assumptions relevance is explained in detail in Kuh, Neese and Hollinger (1985). According with Kant and Schreiber (2003), the behavioral assumptions provide a clear scaling behavior about variables properties.

simulation econometric models take into account the linear and nonlinear mathematical risk system properties.

Two distinct yet related papers that used econometrics and simulation techniques are Just and Just (2011), where risk analysis is based on revealed preference data under constant absolute risk aversion (CARA) and constant relative risk aversion (CRRA) identification problems, using the global identification method; and Kimura and Le Thi (2011) who use the vector of means and the variability matrix, along with Monte Carlo simulations to perform sensitivity analysis of the joint distribution of prices and crop yields, with the aid of the Power utility function.

This research provides converging statistic results with those obtained from experimental data, while assessing the same risk phenomena. This convergence allows similar statistical inference either by using synthetic or experimental data *i.e.*, the null hypothesis rejection of risk neutrality, in favor of Kansas farmers risk aversion.[14] In addition, its algorithm is flexible because it can customize diverse uncertainty events. Changing the corresponding parameter values and adjusting the econometric modelling accordingly achieve this flexibility. Thanks to this flexibility, the uncertainty factors values variability can generate a multiplicity of risk scenarios. Thus, risk scenarios analysis could provide grounded criteria to guide management risk decision-making.[15] That is to say, it can develop proofs for risk reliability given specific economic system uncertainty conditions and the construction of matching testing platforms. For example, these testing platforms become handy for purposes of yield insurance, before a contractual agreement takes place or is enacted.

The proposed method presented herein conveys economically low costs and computational efficiency. This is because it reduces long-time awaiting production times, which are necessary to obtain experimental data.[16] The algorithm constructs synthetic data for assessing different risk scenarios in few seconds. Therefore, study design and econometric analysis and sensitivity costs are reduced dramatically. Modifications on the theoretical framework of this method could provide risk management extensions *i.e.*, Health Policies or demand for risky assets. This method is also computationally efficient because it achieves information convergence and diminishing scaling variance.

This paper is organized as follows: in the first place, a method/algorithm application is proposed and described. Afterwards, in the third section, comparisons are provided between this method, which uses synthetic data, and the one implemented by Saha, Shumway and Talpaz (1994) (SST hereafter) who use experimental data. This exercise is useful to learn how the proposed method estimators achieve information convergence and diminishing scaling variance with respect to SST estimates. In the last section, the

[14] All parameters estimates, either synthetic or from experimental data are significant at the1% level.

[15] "Most economic problems are stochastic. There is uncertainty about the present state of the system, uncertainty about the response of the system to policy measures and uncertainty about future events." Kendrick (1981, p. xi).

[16] In the case of agriculture, some crops have one or even two cycles per year. Meanwhile, perennials need several years before the first harvest takes place.

conclusions are put forward, besides the mean findings are summarized. At the very end, insights for future research are briefly suggested.

2. An application of the propose method

In what follows, the propose method application is implemented.[17] Both, method application and SST estimates are referred to risk assessment for Kansas State wheat farmers, from 1979 to 1982. The risk scenarios are composed by each of the following estimations types: a joint estimate of Expo-Power "Join est." utility function; a Cobb-Douglas production function and exponential forms "Only CDE" and under constant absolute risk aversion "Under CARA" are used in both cases, with the aim of uncovering the farmer risks preference structure and production technology parameters.[18]

In order to conduct this method in the correct track, a set of behavioral assumptions are being taken into account. These assumptions provide support for constructing both synthetic data and risk simulation econometric models. The behavioral assumptions are represented mathematically in a set of non-negative values for prices and quantities.[19] To elaborate further, these behavioral assumptions are important, since they do not impose a specific functional form regarding the underlying utility function. In this way, no restrictions are set on risk preferences structure.[20] Although, these assumptions are not explicit in the following flow diagram, they are assumed to hold at all times.

Next, five steps in sequential order are used to describe an application of the proposed method. A flow diagram is presented in Figure 1 to illustrate these five steps.[21] A description is provided for each of these computational stages.

---

[17] For this case, the algorithm is implemented in Matlab statistic software using a personal computer.

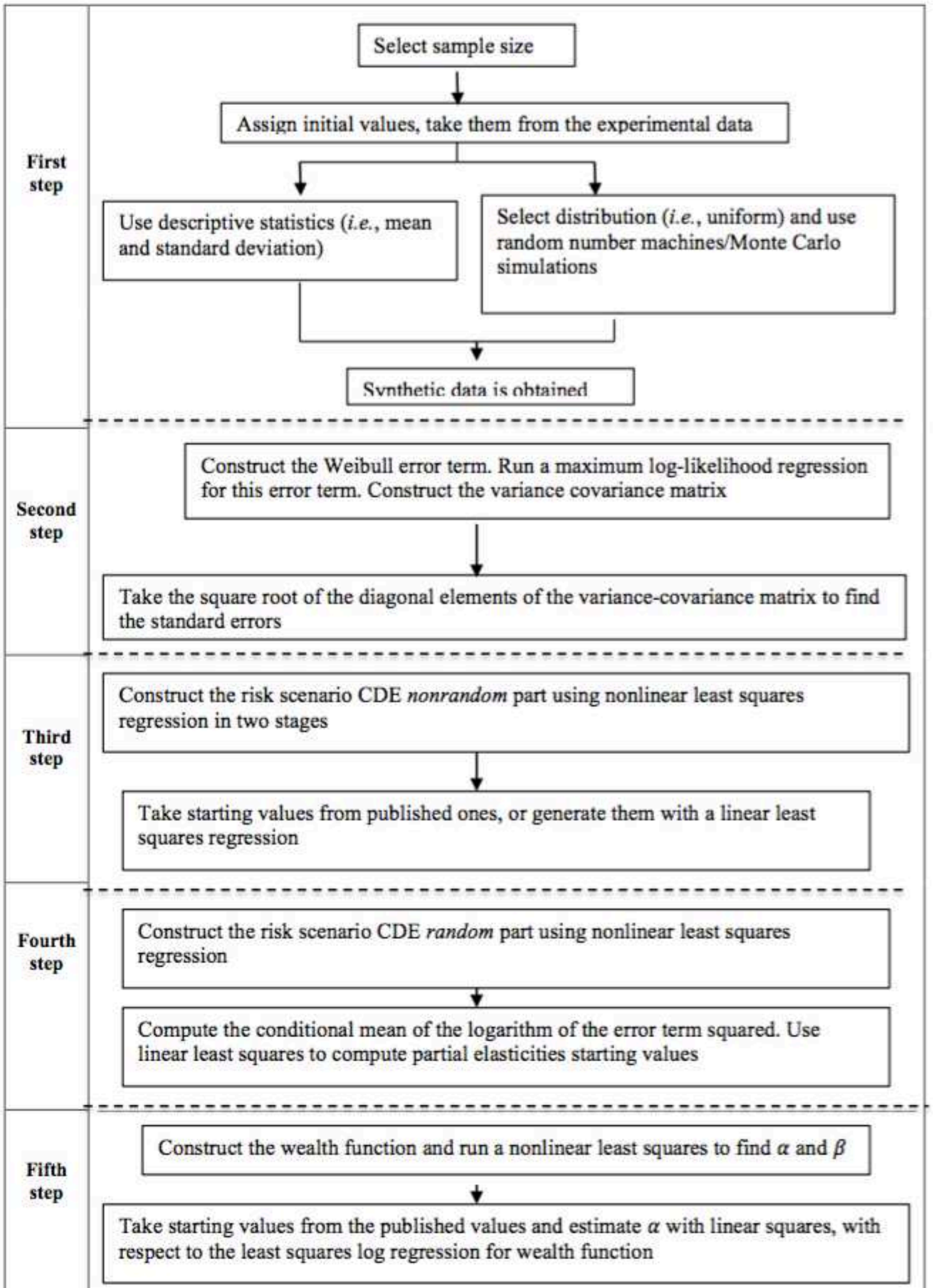
[18] The CDE is estimated with the aid of the Translog function in SST equation (17). One representation of this CDE function is $log a_t = log\{e^{b_1 c_1 + b_2 c_2 + \varepsilon_t}\}$ or alternatively $log a_t = b_1 c_1 + b_2 c_2 + \varepsilon_t$, where a, b, c can be any time series *i.e.*, output, capital and materials. In this way the degree of complementarity or substitution between c_1 and c_2 is revealed. The cross effects are assumed to be zero. The Translog flexibility allows $b_1 + b_2$ to be $\leq or \geq$ than 1. Thus materials could display complementarity or substitution.

[19] For a behavioral production assumptions mathematical proof, see Appendix 1.

[20] It has been documented in Saha (1997) that a specific utility functional form restricts the validity of alternative risk preference structure, as in Sandmo (1970); Batra and Ullah (1974), and Pope (1980). Just and Just (2011) mentioned that Saha *et al.* (1994) method is conditional to production inputs, however this conditionality does not restrict SST method flexibility.

[21] Figure 1 is constructed for the general form *i.e.*, All farms. This flow diagram could be modified to include the respective economic activity factor to construct Small and Large farms estimators.

Figure 1. A method/algorithm for farm risk management. An application



Source: Own elaboration.

First step. Synthetic data is generated on the basis of random machines,[22] to simulate SST experimental data (reported in SST Table 3).[23] Hereby, x_1 stands for capital input; x_2 for material input; p_1 for capital inputs price; p_2 for material inputs price; I is the exogenous income and Q wheat output.[24] In the algorithm the sampling properties are adjusted as follows:[25] a length of 60 observations for all farmers (15 farms for four years);[26] 36 for large farmers (nine farms for four years) and 24 for small farmers (six farms times four years).[27]

Second step. The Weibull error term is constructed with b scale parameter and c shape.[28] The initial values for the random error are taken from Table 4, column "Published." The standard errors, reported on Table 4 on the "Simulation" column are calculated based on the Levenberg-Marquardt gradient expansion algorithm.[29] Therefore, these standard errors are computed by taking the square root of the diagonal elements of this variance-covariance matrix. Figure 2 displays a representation of the random error Weibull distribution.

---

[22] The proposed method/algorithm uses Matlab integrated modules, which contain random number generators for a uniform distribution to help in the creation of synthetic data. This approach is asymptotically efficient, since sampling is taken from an infinite population: "Since the variables were generated with random number [generators], every time the code is run, the variables get updated. This allows the analyst to omit an additional step of taking draws from the population in the Monte Carlo simulation. So, the population from which the samples are being taken is infinite, instead of being bounded to 100 or 1,000 population size." Carbajal (2013, p. 4 *bracket added*). Efficiency is gained because this research variance-covariance matrix approaches asymptotically to the Cramér-Rao lower bound. More information along these lines can be found in the Central Limit Asymptotic Theorems.

[23] For convenience, from now on, equations and tables numbers reported in this document follow SST sequence. For brevity, SST equations are not reproduced here.

[24] The firm-level data used in SST are taken from the Farm Management Data Bank Documentation (Langemeier), Department of Agricultural Economics, Kansas State University. For further reference see Langemeier (1990).

[25] The synthetic data construction only needs two experimental data points: the mean and standard deviation of the relevant random variable. It is not necessary to compute r_1 (capital input price) and r_2 (material input price), because they are considered as exogenous variables since the beginning. Their mean and standard deviation are reported in Table 3. In other words, there is not need to solve equation (14) for input prices r_1 and r_2.

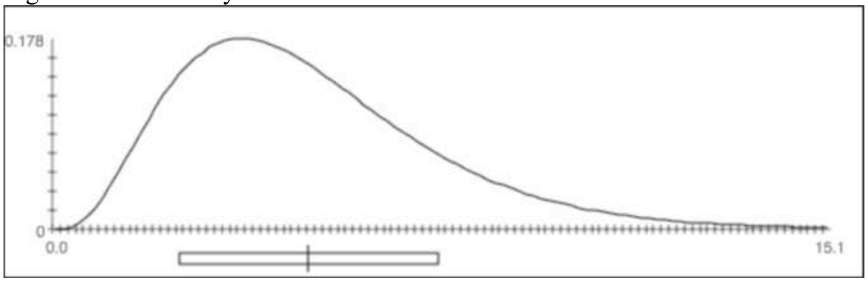
[26] In selecting the representative wheat farm size, SST mentions that "… we divided the data sample into a group of small farms and a group of large farms based on output level. Using these two data sets separately, we reestimated model parameters and computed $A(\bar{W})$ and $R(\bar{W})$." p. 181. Thus, the implied first-degree equation is solved to find the farm size for large and small categories.

[27] The common factor among farms is a multiple of 12. An alternative could involve using a different simulation factor to represent federal states, or even the whole country (*e.g.*, 1,000; 1,000,000; etc.). In this respect see Lancaster (1966, p. 135).

[28] The Weibull error term modelling follows SST theoretical assumptions. More information, about how to estimate a Weibull probability density function (pdf) can be found in Ahmed (2013). According to Hennessy (2009), the error term modelling aims to determine the systemic risk on the yield distribution.

[29] Here, the diagonal terms of the curvature matrix are increased by a factor that is optimized in each search step for optimum values. More information in this respect, are founded on Bevington and Robinson (1992); Patel, Kapadia and Owen (1976). The Fisher information matrix inverse is the asymptotic variance-covariance matrix of the parameter estimates, computed out from a maximum likelihood regression.

Figure 2. Stochastic yield variable error. Weibull distribution



Source: Own calculations and mathworks.com

In Figure 2, the random error Weibull distribution has shape of 3.8281 (c in Table 4, "Published" column) and scale 1.2894 (b in Table 4, "Published" column). The center of the boxplot at the foot of Figure 2 signals the median location.[30] The mean is smaller than the median. Thus, the distribution has a fat tail at the left side, which indicates an asymmetrical distribution.[31] This implies a higher probability of obtaining in consecutive draws, lower random error values.

Next, according with SST assumptions the complexity of the estimation of equation (14) can be substantially reduced, if parameters of equation (15), b and c are prior estimated and then used for estimating equation (11). These assumptions are accounted for on the following steps 3-4. To elaborate, equation (11) represents the risk scenario CDE (Cobb-Douglas and exponential forms), which can be generated with two nonlinear least squares stages.[32] The first stage, step 3, *nonrandom* part generates its parameter starting values, with a prior estimation of b and c using a Just-Pope modified method, to address that ϵ the stochastic yield variable error has a Weibull distribution.[33] The second stage, step 4, *random* part estimates the production function exponential part parameters.

Third step. Equation (11) first stage implements the *nonrandom* part relying in a nonlinear least squares regression based on the Levenberg-Marquardt algorithm.[34] The

---

[30] It is worth mentioning that Weibull distributions exhibit positive values only. Therefore, the error term fulfills the proposed method behavioral assumptions.

[31] For a detail skewness analysis, see Malkiel and Saha (2005).

[32] Wheat Output (Q) is assumed to be produced with two inputs: x_1 (capital) and x_2 (material).

[33] According to Just and Pope (1979), this type of estimation can describe stochastic production functions and related behavior.

[34] This algorithm is incorporated in the Matlab function handle for nonlinear least squares regression, see Seber and Wild (2003) and Prajnesgu (2008) for more information. This algorithm does not compute robust fits. In order to attain robust fits, other algorithm sequence is needed, which iteratively refits a weighted nonlinear regression (DuMouchel and O'Brien, 1989). Each iteration weights are based on the previous residual observation (Knowles and Siegmund, 1989).

The proposed algorithm conveys an iterative process. The following quote could be quite illuminating on how this process works: "The two equations in $(P)_{to}$ and $(E)_{to}$ are non-linear and therefore not easily solved by straightforward algebraic processes. An iterative or 'cut-and-try' method can quickly be applied to approximate a solution to any reasonable desired degree of accuracy. We first assume a value for $(E)_{to}$. We solve an equation for $(P)_{to}$ and substitute this value in the other equation. This implies a resulting value for $(E)_{to}$. If it's not the same as the initially assumed value, we adjust our assumption and begin again until we get a pair of values for $(E)_{to}$ and $(P)_{to}$ which mutually satisfy both equations. Once the process has been followed through, it is easy to see what sort of adjustments are needed to bring the process into rapid convergence." Klein, Ball, Hazlewood and Vandome (1961, p. 34).

output on equation (11), part Cobb-Douglas and part exponential form is needed as an independent variable for the second stage. In order to compute the dependent variable of equation (11) synthetic data is constructed.[35] The starting values for synthetic wheat output (Q) could be drawn from SST published data, or alternatively it could be generated through a linear least squares regression.[36]

Fourth step. The second stage furnishes the estimates for the *random* part of equation (11), based on a nonlinear least squares regression. A specific structure of the *random* part (*i.e.*, error term) is to be imposed using equations (19)–(21).[37] In addition, a linear regression model provides the starting values for the partial production elasticities m_1 and m_2.

Fifth step. The starting values for the Expo-Power utility parameters, α and β, can be found through a suitable grid search. This search can be executed after estimating m_1 and m_2. In this respect, the grid search could be faster than an optimization method.[38] On the other hand, optimization algorithms could be used instead of grid search, since the former represents a more thorough search. Thus, the proposed method uses the optimization to perform α and β search. The risk scenario "Joint est." is generated on steps 3-5. The corresponding estimators are reported on Table 5, in the "Simulation" column.

It is worth mentioning that SST normalize the profit function with respect to output price. Thus, output price becomes the *numéraire*.[39] This normalization makes input prices a share of output price. This procedure integrates the joint estimation in expected wealth (W).[40] Once W is computed a nonlinear least squares regression can find α and β parameters.

---

[35] It should be noted, that the absence of a closed form solution for equation (11) does not allow direct estimation of Q using the *nonrandom* part estimates.
[36] In this case a linear least square regression was implemented. No traces of endogeneity are detected in this regression. For those cases where endogeneity problems are detected, a Matlab treatment can be applied, for a useful method to counteract these problems see Carbajal (2014).
[37] According with Just and Just (2011, hereafter JJ), SST fail to identify producer risk preferences, because "parameter identification is achieved as suggested by estimated standard errors and t-ratios… then one cannot claim to have identified risk preferences for producers". Thus, JJ propose global identification as a way to overcome SST lack of identification. However, when JJ assumes that E(ε)=0, they are imposing a nonstochastic structure in the error term –a neutrality towards risk, at least on price risk distribution. Also, JJ assumption ignores the measure of precision or uncertainty that accompanies the estimates (see Rhoda (2016)). Therefore, it seems that JJ method lacks identification because the risk error structure cannot be determined from their model. Besides, Tables 2-3 does not report standard errors, precluding the reader to identify if the risk preferences for producers are or not statistically significant.
[38] Because the parameters and independent random variables needed for computing wealth are already estimated, this stage could be processed through an optimization path instead of a grid search. Performing this evaluation on a complete grid, as required by the *max* algorithm, will be much less efficient. This is because it samples a small subset of grid discrete points.
[39] For uses and implications of the *numéraire*, see Walras (2010 [1877]). In particular, when the theorem of general equilibrium is put forward (p. 185, paragraph 145).
[40] This computation integrates the modified Cobb-Douglas and exponential forms previously obtained on equation (11), in its two stages. To be more explicit about the determination of the normalized random wealth W, SST equation (5) is computed using the corresponding synthetic data and the parameters already

Summarizing, the above five computational steps provide the parameter vector ϕ with dimension: (2n+5) x 1. It is important to note that the propose method provides risk sensitivity estimates and at the same time, maximizes the expected wealth embedded in the expected utility function.[41] Therefore, in this way SST equation (6) is determined.[42]

3. A comparison. Estimates of technology, preferences and tests results

3.1 Descriptive statistics

The descriptive statistics for synthetic data "Simulation" columns are computed using the proposed method. They are presented in Table 3 (following SST table numbers), alongside with the descriptive statistics from SST experimental data "Published" columns. This display comparse method and experimental data figures in an easy way. For All farmers, convergence is evident when comparing "Simulation" and "Published" columns, because their mean and standard deviations are quite similar in magnitude and sign.[43]

---

determined, in the third and fourth optimizations steps. In addition, SST published α value is being taken as a starting value.

[41] The ϕ parameters vector dimension stands for: 2 equal to α and β; n equal 15 farms; 5 for a; b; c; m; μ. For a visualization of these parameters see column "Parameter" in Table 4 and column "Explanation" in Table 5. In this last column, A is considered as a constant, thus it is not accounted on ϕ. The simulation ϕ parameters are obtained as output from the nonlinear and linear least squares simulation econometric models. All the estimations needed in equation (14) are already simulated, due to the implementation of the proposed method. Therefore, there is no need to compute equation (14) directly, neither it is necessary to compute its left side, which is composed by input prices.

[42] The identity expressed in equation (6) on SST could be worded as follows: optimal input levels (left side) are identical to maximize the expected utility (right side). The proposed method works in the left hand side of identity (6) to find optimal input levels. Therefore, this method is capable of constructing a ϕ parameter vector. Thus, by considering the ϕ parameter vector and identity (6), the proposed method also maximizes expected utility (right side).

[43] As the size of the sample diminishes (*e.g.*, Small farmers), this resemblance seems to fade. It should be borne in mind that the Law of Large numbers assures convergence on distribution to the true estimate when the sample size is enlarged. The contrary effect is expected to happen, with a sample size reduction.

Table 3. Summary statistics

	Variable mean (and standard deviation)					
	All farmers		Small farmers		Large farmers	
	Simulation	Published	Simulation	Published	Simulation	Published
Capital inputs: x_1	50.967 (24.0887)	44.051 (23.0280)	33.716 (7.6147)	32.017 (9.8392)	56.425 (29.4501)	52.073 (25.7840)
Material inputs: x_2	69.651 (30.9575)	70.059 (28.0280)	34.975 (19.8822)	43.206 (17.7360)	84.117 (23.9502)	87.961 (25.7840)
Capital inputs price:[a] p_1	0.954 (0.0789)	0.961 (0.0663)	0.986 (0.0816)	0.966 (0.0671)	0.961 (0.0781)	0.957 (0.0664)
Material inputs price:[a] p_2	0.996 (0.0500)	0.993 (0.0455)	0.980 (0.0634)	0.991 (0.0487)	0.993 (0.0480)	0.994 (0.0438)
Exogenous income:[a] I	55.208 (21.5051)	51.652 (18.8790)	36.762 (8.9254)	41.407 (8.5820)	34.473 (18.3448)	58.482 (20.7950)
Wheat output: Q	127.547 (42.9020)	132.060 (56.3550)	80.731 (14.4280)	78.071 (25.5820)	126.849 (39.0083)	168.050 (39.9030)

[a] Prices and income are normalized by output price.
Source: Saha *et al*. (1994) and own computations.

Next, the maximum likelihood estimates and the standard errors for the Weibull error term are presented in Table 4. The “Simulation” and “Published” columns are relevant to perform a side-by-side comparison. Also, comparisons can be done by rows. For example, the mean and standard deviation estimates for the Weibull stochastic yield variable error ϵ are in rows three and five. Meanwhile, the corresponding figures for sample ϵ are in rows four and six.

Table 4. ML estimates of Weibull Parameters

Parameter	Explanation	Simulation (standard error)	Published (standard error)
b	scale parameter	1.2894 (0.0521)	1.2894 (0.0462)
c	shape parameter	3.4752 (0.3627)	3.8281 (0.3486)
$\hat{b}\Gamma\left(\frac{1+\hat{c}}{\hat{c}}\right)$	estimated mean of Weibull ϵ	1.2036	1.1658
$\bar{\epsilon}=\frac{1}{T}\sum_{t=1}^{T}\epsilon_t$	sample mean of ϵ	1.1322	1.1711
$\hat{b}\left\{\Gamma\left(\frac{2+\hat{c}}{\hat{c}}\right)-\left(\Gamma\left(\frac{1+\hat{c}}{\hat{c}}\right)^2\right)\right\}^{\frac{1}{2}}$	estimated standard deviation of Weibull ϵ	0.2875	0.3403
$\left\{\frac{1}{T-1}\sum_{t=1}^{T}(\epsilon_t-\bar{\epsilon})^2\right\}^{\frac{1}{2}}$	sample standard deviation of Weibull ϵ	0.7686	0.3125

Source: Saha *et al*. (1994) and own computations.

On Table 4 the scale b and shape parameter c from "Simulation" estimates keep a close resemblance, with respect to SST counterpart (*i.e.*, 1.2894 *vs.* 1.2894 for the scale parameter and 3.4752 *vs.* 3.8281 for the shape parameter). The estimated error mean, ϵ and its estimated standard deviation "Simulation" and "Published" estimates are also similar (*i.e.*, 1.2036 *vs.* 1.1658 and 0.2875 *vs.* 0.3403, respectively). In the same fashion, the corresponding figures for sample mean ϵ are 1.1322 *vs.* 1.1711; their standard deviations are 0.7686 *vs.* 0.3125, respectively.[44]

3.2 Risk scenarios and tests results

Comparisons are provided for the utility function and production technology estimates, along with different risk preference structure tests. The Steps 1–5 provide the logarithm sequence on how to compute the risk sensitivity estimates for "Join est." risk scenario. This scenario is reported on Table 5 in the left side, third and fourth columns, both for "Simulation" and "Published" estimates. Partial steps of this logarithm, for example steps 1-4 can produce the risk scenario CDE. By the same token, partial modifications on the method with respect to utility function parameters ($\alpha = 1$) furnish CARA estimates. Therefore, Table 5 also incorporates and reports the estimates for two additional risk scenarios: "Under CARA" in columns five and six and "Only CDE" in columns seven to eight.

[44] The method efficiency gain is expressed on lower standard errors *loc. cit.* 21, 43. The opposite takes place *i.e.*, efficiency loss, when the standard errors belong to samples that experiments size reduction, which at present is the case.

Table 5. Parameter estimates of EP utility and CDE production function

		Estimates (standard errors)					
EP utility parameter	Explanation	Joint est.[a] Simulation	Join est.[a] Published	Under CARA[b] Simulation	Under CARA[b] Published	Only CDE[c] Simulation	Only CDE[c] Published
α	$\alpha < 1 \Rightarrow DARA$	0.3654 (0.36E-11)	0.3654 (0.0294)				
β	$\beta > 0 \Rightarrow$ IRRA	2.7370 (0.01E-11)	2.7370 (0.2201)				
CDE production function parameters[d]							
A	Parameters of the non-stochastic part of CDE	1.6051 (1.51E-14)	1.6051 (0.1530)	2.114 (0.03E-13)	2.114 (0.3534)	3.0896 (0.19E-17)	3.3644 (1.2606)
a_1		0.2554 (1.45E-15)	0.2554 (0.0126)	0.2554 (0.04E-13)	0.2561 (0.0134)	0.2012 (0.21E-17)	0.2224 (0.0665)
a_2		0.7564 (1.79E-15)	0.7564 (0.0179)	0.7564 (0.52E-13)	0.7169 (0.0286)	0.6522 (0.01E-17)	0.6715 (0.0957)
m_1	$m_j<0 \Rightarrow$ jth input is "risk reducing"	0.0612 (0.04E-14)	0.0612 (0.0050)	0.0526 (0.22E-16)	0.0526 (0.0077)	1.71E-16 (0.0001)	0.0252 (0.0062)
m_2		-0.0337 (0.17E-14)	-0.0337 (0.0054)	-0.0709 (0.04E-13)	-0.0709 (0.0070)	3.17E-17 (0.0001)	0.0406 (0.0042)
SSE	Sum of squared errors[f]	2.37	2.49	4.35E-24	4.81	2.05E-28	15.9700
Partial production elasticities[e]							
μ_1	Elasticity of x_1	0.3206	0.2031	0.3722	0.2265	0.4276	0.2381
μ_2	Elasticity of x_2	0.5267	0.6016	0.5689	0.6354	0.6535	0.7047

[a] Expected utility maximization model (unrestricted);
[b] Expected utility maximization under the restriction $\alpha = 1$, implying CARA;
[c] CDE production function parameter estimates: modified Just-Pope method;
[d] Subscripts 1 refers to capital and 2 to materials;
[e] Computed at the sample mean;
[f] In simulation, MSE generated from the fists stage regression.
Source: Saha *et al*. (1994) and own computations.

Table 5 puts forward three risk scenarios: "Join est."; "Under CARA" and "Only CDE." For convenience, these risk scenarios are represented in the "Simulation" and "Published" columns, side by side. The comparison among risk scenarios underscores the risk sensitivity of their estimates under alternative parameter values and different functional forms. For instance, the "Simulation" for "Join est." scenario is reportedly the more efficient functional form when compared with "Under CARA" and "Only CDE" risk simulated scenarios. For these three simulation scenarios, the standard errors for m_2 estimates are 0.17E-14; 0.04E-13 and 0.0001, respectively.[45] Therefore, "Join est." displays efficiency gain or the smallest standard errors.[46]

The estimated standard errors of α from the "Simulation" method are smaller than the corresponding standard errors from "Published" procedure *i.e.*, 0.36E-11 *vs.* 0.0294, respectively.

Thus, "Join est." simulated risk scenario displays the smaller standard errors for α in terms of their 1) comparison among alternative risk scenarios and 2) its own comparison between "Simulation" and "Published" values. Together, these two attributes for "Join est." simulation standard errors, exhibit a diminishing scaling variance.

Convergence is achieved among α estimates of the proposed method and SST procedure.[47] The corresponding figures are reported on Table 5. For instance, "Joint est." risk scenario, a_1 capital share estimate is 0.2554 for both methods. In the same manner, a_2 material share has an elasticity of 0.7564. Similar patterns are displayed between "Simulation" and "Published" estimates for "Under CARA" and "Only CDE" risk scenarios. A possible explanation for convergence and diminishing scaling variance in α is that standard errors are estimated precisely, that the economic model of "Join est." is well identified and consistent, and the behavioral assumptions underlying the proposed method are correct.[48]

For m_1 and m_2 partial elasticities estimates, exact information convergence is also achieved for "Joint est." and "Under CARA" scenarios. For m_1 and "Joint est." the corresponding figures are 0.0612 and 0.0612, for "Simulation" and "Published" columns, respectively. The risk reducing input, m_2, is the one reported with a negative sign *i.e.*, for "Joint est." with a figure of -0.0337 for both "Simulation" and "Published" columns.[49] For "Only CDE" exact information convergence is not achieved, but the distance between simulations are less than 10%. In addition, its simulation estimates do not find

[45] "This suggests that there is indeed a substantial efficiency gain in joint estimation, corroborating similar findings by Love and Buccola." SST p. 182. This is also the case for the join simulation estimates.
[46] Efficiency is linked with the method asymptotic properties, making them more efficient in statistic terms, than their experimental data analog. Thus, the "Simulation" standard errors reach a diminishing variance.
[47] As "Simulation" and "Published" estimates differences are almost zero, they achieve convergence. This convergence is an instance of the Dennis-Moré Characterization Theorem, *i.e.*, the α for "Joint est.": "Simulation" and "Published" estimates are 0.3654 and 0.3654, respectively. It is easily seeing that by performing the difference operation between these estimates ($0.3654-0.3654\approx 0$) approaches zero (the approximation depends on the number of decimals in each magnitude). A similar operation can be performed for the rest of estimators reported on Table 5. Fletcher (1987 p. 125) provides a proof of the Dennis-Moré Characterization Theorem. This proof relies basically in the fact that convergence is achieved when the difference of the two relevant figures reach zero, as their limit.
[48] Thus asymptotic consistency is assured by convergence in probability. This implies a unique maximum at the true parameter, given asymptotic conditions for identification. More information about asymptotic theory can be found in Newey and McFadden (1994). In this case identification is guaranteed.
[49] "… it should be recalled that the materials category included a large array of inputs such as fertilizer, seed, machinery operating inputs and miscellaneous purchased inputs." SST p. 183. Chavas and Shi (2015) explain that "risk-decreasing" inputs *i.e.*, irrigation and pest control decrease output variance.

material as risk reducing, which could be a serious mistake in inference.[50] These findings rules out "Only CDE" as an efficient estimation method. Also, they support "Joint est." as the more efficient risk-scenario. It is important to remember that "Joint est." and "Under CARA" risk scenarios are more efficient, since they display materials as the risk reducing input.[51]

The partial production elasticities for capital (μ_1) and materials (μ_2) are similar across risk scenarios. For instance, materials μ_2 estimates take values for "Joint est." of 0.3206 and 0.2031 for "Simulation" and "Published" columns. The same sign is observed on μ_1 and μ_2, when contrasting "Simulation" with "Published" figures. It is important to note, that the computation of the production elasticities described on the third step algorithm, are necessary to obtain the starting values of CDE *nonrandom* part. Thus, its direct interpretation is not central. Perhaps, this is probably the reason for which their standard errors are not reported on the original SST Table 5. In what follows, Table 6 reports the Arrow-Pratt risk aversion measures, both for simulation and SST published results.[52]

Table 6. Arrow-Pratt Risk Aversion Measures

	Estimates (standard errors)[e]					
	All farmers[f] Simulation	All farmers Published	Small farmers[f] Simulation	Small farmers Published	Large farmers[f] Simulation	Large farmers Published
Estimate of EP parameter α	0.365 (1.00E-11)	0.365 (0.029)	0.405 (0.58E-09)	0.292 (0.031)	0.229 (0.004)	0.266 (0.028)
Estimated mean wealth:[a] $\overline{W}$	49.449 (45.287)	46.213 (18.483)	11.985 (10.549)	35.768 (8.422)	57.857 (31.706)	53.177 (20.114)
Absolute risk aversion:[b] A($\overline{W}$)	0.0002 (0.0001)	0.0075 (0.002)	0.0011 (0.0011)	0.0083 (0.002)	0.0008 (0.003)	0.0045 (0.023)
Relative risk aversion:[c] R($\overline{W}$)	8.522 (3.737)	5.400 (0.540)	1.137 (0.626)	3.759 (0.322)	8.028 (0.381)	4.075 (0.391)
Test statistics for the null hypothesis of risk neutrality[d]	-7.376 (1.769)	-4.693 (0.406)	-9.720 (3.719)	3.595 (-0.264)	-3.503 (0.807)	-3.724 (0.316)

[a] Estimated mean wealth: $\overline{W} \equiv \int_0^3 \{\hat{A}(\bar{x}_1)^{\bar{a}_1}(\bar{x}_2)^{\bar{a}_2} + exp\{\bar{m}_1\bar{x}_1 + \bar{m}_2\bar{x}_2 + \epsilon\} - r_1\bar{x}_1 + r_2\bar{x}_2 + I\} PDF(\epsilon)d\epsilon$;
[b] A($\overline{W}$) for EP utility: $\frac{\{1-\hat{\alpha}+(\overline{W})^{\hat{\alpha}}\}}{\overline{W}}$;
[c] R($\overline{W}$): $\{1-\hat{\alpha}+(\overline{W})^{\hat{\alpha}}\}$;
[d] H_0: $\psi = 0$, where $\hat{\psi} \equiv \hat{\alpha} - 1 - \hat{\alpha}\hat{\beta}\overline{W}^{\hat{\alpha}}$;
[e] The standard errors of A($\overline{W}$), $R(\overline{W})$ and $\hat{\psi}$ have been calculated using the delta method;
[f] Adjusted by a percentage factor.
Source: Saha *et al*. (1994) and own computations.

[50] This method is less efficient than "Joint est." and "Under CARA." Perhaps, sampling from an infinite population emphasizes "Only CDE" bias. For a discussion on econometric techniques, efficient performance comparisons and sample properties, see Saha, Havenner and Talpaz (1997).
[51] Numerous studies analyzing input level effects on output have found materials as risk reducing *i.e.*, fertilizer in Just and Pope (1979). For Tveterås (1999) labor is risk reducing.
[52] It should be mentioned that the risk sensitivities estimates depend heavily on the parameter starting values (Kantz and Schreiber, 2003 p. 65).

The first row of Table 6 provides the estimate of parameter α. Complete convergence between "Simulation" and "Published" is achieved for All farmers as well as Large farmers, with exception of Small farmers. The absolute risk aversion estimates A($\bar{W}$), for the three farmers groups, achieve convergence at two decimal places *i.e.*, All farmers "Simulation" and "Published" risk sensitivities are 0.0002 *vs.* 0.0075, respectively. Similarly, the mean wealth $\bar{W}$ values exhibit a low variance between the alternative estimates for All farmers and Large farmers.[53] Thus, decreasing absolute risk aversion (DARA) preferences are displayed for all cases: $A(\bar{W}) < 0$.

The simulation replicates the published sign results regarding risk behavior for the three types of farmers with respect to relative risk aversion $R(\bar{W})$. For example, these results for All farmers are 8.522 *vs.* 5.400 for "Simulation" and "Published" columns, respectively. Thus, the same statistical inference can be done either using synthetic or experimental data. Thus, increasing relative risk aversion (IRRA) preferences are displayed for all cases: $R(\bar{W}) > 0$.

The published or experimental data results for All farmers as well as for Large farmers, with respect to t-statistics for the null hypothesis of risk neutrality have the same sign, *i.e.*, -7.376 *vs.* -4.693 and -3.503 *vs.* -3.724. These t-statistics are statistically significant at the level of 99%. Thus, the null hypothesis of risk neutrality is rejected in both methods.

The comparisons for this subsection can be summarized as follows: the method estimates achieve convergence in sign, in all cases, when contrasted with those estimates obtained from experimental data. In a similar manner, convergence in magnitude is also achieved, with few exceptions for small farmers. These convergences in sign and magnitude validates that both data sets, synthetic and experimental can identify the same levels of optimal parameters. This is because both data sets share the same data points. The standard errors estimates from the proposed method are in all cases highly significant, when compared to those derived from experimental data. These estimates high significance reveals the propose method statistic efficiency, which is demonstrated on a diminishing scaling variance. This diminishing scaling property is derived from the propose method asymptotic properties, when sampling from an infinite population. Overall, the proposed method results provide similar statistical inference, as the one derived from experimental data. For instance, decreasing absolute risk aversion (DARA) and increasing relative risk aversion (IRRA) risk preference structure are revealed.[54] Furthermore, the method findings for all and large farmers allow the rejection of the null hypothesis of risk neutrality, in favor of Kansas farmers risk aversion. This central risk statistic inference is achieved for SST, while using experimental data.[55]

---

[53] The overbar denotes estimated means.

[54] "Arrow-Pratt estimates for both groups are consistent with DARA and IRRA. The small farmers do show a higher level of $A(\bar{W})$ and a lower level of $R(\bar{W})$ than do the larger farmers." SST p. 181.

[55] "The empirical findings clearly rejected the null hypothesis of risk neutrality in favor of risk aversion among Kansas farmers." SST p. 183.

4. Conclusions

A method is proposed to allow the researcher to compute and construct synthetic data, along with simulation econometric models based on nonlinear and linear interpretations of behavioral production relationships. This algorithm has the advantage of replicating natural phenomena and uncertainty factors within a short period of time. These features convey low financial costs, besides computational efficiency.

Five steps conform this proposed method. An application of its algorithm is developed for wheat Kansas farmers. A comparison is provided between the proposed method results and SST wheat Kansas farmer experimental results. Based on this comparison, the proposed method estimates for All and Large farmers achieve information convergence, in magnitude and sign, with respect to those results derived from experimental data. For Small farmers case convergence seems to fade, perhaps because their estimates loss asymptotic properties. The algorithm standard errors display in all farmer cases a diminishing scaling variance, with respect to their experimental data analog.

A possible explanation for convergence and diminishing scaling variance consist on correct behavioral assumptions. Estimates consistency is also implied since synthetic and experimental data sets share the same data points: mean and standard deviation.

The identification and efficiency of this proposed method are validated, when its estimates replicates the statistical inference achieve by experimental data results. The proposed method asymptotic t-tests are statistically significant at a 99% level. These statistical inference similarities between synthetic and experimental data provide the grounds for the central finding of this paper: the method herein implemented can reveal the degree and structure of risk aversion as well as experimental data do. Thus, the substitution of experimental data by synthetic data is feasible for risk analysis simulation purposes and sensitivity analysis. The proposed method could develop proofs for risk reliability given specific economic system uncertainty conditions and construct the corresponding risk testing platforms.

Therefore, the proposed method makes risk analysis accessible to policy planners, farmers associations, insurance companies, hedge funds and governments, etcetera. It is hoped that this proposed method could provide grounded criteria to improve their decision-making by providing the necessary tools to guide their management practice.

References

Ahmed, S.A. "Comparative study of four methods for estimating Weibull parameters." *International Journal of Physical Sciences* 8(2013):186-192.

Bassi, A., Colacito, R. and Fulghiery, P. "'O sole mio: an experimental analysis of weather and risk attitudes in financial decisions." *The Review of Financial Studies* 26(2013):1-29.

Batra, R.N. and Ullah, A. "The competitive firm and the theory of input demand under price uncertainty." *Journal of Political Economy* 06-07(1974):537–548.

Bevington, P.R. and Robinson, D.K. *Data reduction and error analysis for the physical sciences.* NY: McGexperimental-Hill, 1992.

Carbajal, C. "Synthetic data: an endogeneity simulation." *Working Paper, Economics Department* Universidad Autonoma Metropolitana at Iztapalapa 2014.

Carbajal, C. "How to use Matlab in a statistical application." *Unpublished mimeo.* 2013.

Chavas, J.P. and Holt, M. "Economic behavior under uncertainty: a joint analysis of risk preferences and technology." *The Review of Economics and Statistics* 78-2(1996):329–335.

Chavas, J.P. and Shi, G. "Economic analysis of risk, management, and agricultural technology." *Journal of Agricultural and Resource Economics* 40-1(2015):63-79.

DuMouchel, W.H. and O'Brien, F.L. "Integrating a robust option into a multiple regression computing environment." in *Computer Science and Statistics Proceedings of the 21st Symposium on the Interface.* Alexandria, VA: American Statistical Association, 1989.

Fletcher, R. *Practical Methods of Optimization* Padstow: Wiley (1987).

Friend, I., and Blume, M.E. "The Demand for Risky Assets." *American Economic Review* (12)1975:90-22.

Hazell, P.B.R. "A linear alternative to quadratic and semivariance programming for farm planning under uncertainty." *American Journal of Agricultural Economics* 53(1971): 53-62.

Hennessy, D.A. "Crop yield skewness and the normal distribution" *Journal of Agricultural and Resource Economics* 34(2009): 34-52.

Just, R.E. and Pope, R.D. "Stochastic specification of production functions and economic implications." *Journal of Econometrics* 7(1978):67-86.

Just, R.E. and Pope, R.D. "Production Function Estimation and Related Risk Considerations." American *Journal of Agricultural Econometrics* 61(1979):276-284.

Just, R.E. and Just, D.R. "Global identification of risk preferences with revealed preference data." *Journal of Econometrics* 162(2011):6-17.

Kantz, H. and Schreiber, T. *Nonlinear time series analysis.* Cambridge: Cambridge University Press, 2003.

Kendrick, D. *Stochastic control for economic models.* NY: McGexperimental-Hill, 1981.

Kimura, S. and Le Thi, C. "Farm level analysis of risk and risk management strategies and policies: technical note." *OECD Food, Agriculture and Fisheries Papers* 48(2011):1-37.

Klein, L.R., Ball, R.J., Hazlewood, A. and Vandome, P. *An econometric model of The United Kingdom.* Oxford: Basil Blackwell, 1961.

Knowles, M. and Siegmund, D. "On Hotelling's approach for testing for a nonlinear parameter in regression." *International Statistical Review* 57(1989):205-220.

Kuh, E., Neese, J.W. and Hollinger, P. *Structural sensitivity in econometric models*. NY: John Wiley & Sons, 1985.

Lancaster, K.J. “A new approach to consumer theory.” *The Journal of Political Economy* 74(1966):132-157.

Langemeier, L.N. “Farm management data bank documentation.” Department of Agricultural Economics, Staff Paper No. 90-10. Manhattan: Kansas State University, 1990.

Love, A. and Buccola, S.T. “Joint risk preference-technology estimation with a primal system.” *American Journal of Agricultural Economics* 73(1991):765-774.

Malkiel, B.G. and Saha, A. “Hedge Funds: Risk and Return.” *Financial Analyst Journal* 61(2005):80-88.

Mehra, R. and Prescott, E. C. “The equity premium. A puzzle.” *Journal of Monetary Economics* 15(1985):145-161.

Melhim, A. and Shumway, C. R. “Enterprise diversification in US dairy: impact of risk preferences on scale and scope economies.” *Applied Economics* 43(2010):3849-3862.

Newey, W.K. and McFadden, D. “Chapter 36. Large sample estimation and hypothesis testing.” *Handbook of Econometrics* 4(1994):2111-2245.

Patel, J.K., Kapadia, C.H. and Owen, D.B. *Handbook of statistical distributions.* NY: Marcel Dekker, 1976.

Pope, R.D. “The generalized envelope theorem and price uncertainty.” *International Economic Review* 02(1980):75-86.

Prajneshu “Fitting of Cobb-Douglas production functions: revisited.” *Agricultural Economics Research Review* 21(2008):289-292.

Pratt, J.W. “Risk aversion in the small and in the large.” *Econometrica* 32(1964):122-136.

Rhoda, D. “Visual Representation of Inferential Uncertainty.” in *STATA Conference Proceedings.* Chicago: STATA, Data Analysis and Statistical Software, 2016.

Saha, A. “Expo-power utility: a flexible form for absolute and relative risk aversion.” *American Journal of Agricultural Economics* 75(1993):905–913.

Saha, A. “Compensated optimal response under uncertainty in agricultural household models.” *Agricultural Economics* 11(1994):111–123.

Saha, A. “Risk preference estimation in the nonlinear mean standard deviation approach.” *Economic Inquiry* 35(1997):770–782.

Saha, A., Havenner, A. and Talpaz, H. “Stochastic production function estimation: small sample properties of ML versus FGLS.” *Applied Economics* 29(1997):459-469.

Saha, A., Love, H. A. and Schwartz, R. “Adoption of emerging technologies under output uncertainty.” *American Journal of Agricultural Economics* 76(1994):836–846.

Saha, A., Malkiel, B.G. and Grecu, A. “The clustering of extreme movements: stock prices and the weather.” *The Journal of Investment Management* 7(2009):20-35.

Saha, A., Shumway, C.R. and Havenner, A. “The economics and econometrics of damage control.” *American Journal of Agricultural Economics* 79(1997):773–785.

Saha, A., Shumway, C.R. and Talpaz, H. “Joint estimation of risk preference and production technology using the expo-power utility function.” *American Journal of Agricultural Economics* 76(1994):173–184.

Sandmo, A. “The effect of uncertainty on savings decisions.” *Review of Economic Studies* 07(1970):353–360.

Schneider, M. and Abowd, J.M. “A new method for protecting interrelated time series with Bayesian prior distributions and synthetic data.” *Journal of the Royal Statistical Society. Statistics in Society. Series A* 178(2015):963-975.

Seber, G.A.F. and Wild, C.J. *Nonlinear regression.* New Jersey: Wiley-Interscience, 2003.

Simmons, R.L. and Pomareda, C. "Equilibrium quantity and timing of Mexican vegetable exports." *American Journal of Agricultural Economics* 57(1975):472-479.

Tveterås, R. "Production risk and productivity growth: some findings for Norwegian Salmon Aquaculture." *Journal of Productivity Analysis* 12(1999):161-179.

Walras, L. Elements of pure economics. Abingdon: Routledge, 2010 [1877].

Wiens, T.B. "Peasant risk aversion and allocative behavior: a quadratic programming experiment." *American Journal of Agricultural Economics* 58(1976):629-635.

Internet site:
The Farm Level Income and Policy Simulation Model (FLIPSIM) https://www.afpc.tamu.edu/models/flipsim/

Appendix 1. Mathematic proof of behavioral production assumptions

Behavioral production assumption 1 (BPA1): prices are positive $\geq 0 \in \mathbb{R}_+$;
Behavioral production assumption 2 (BPA2): quantities are positive $q \geq 0 \in \mathbb{R}_+$.[56]

where p stands for prices; q for quantities and $\mathbb{R}_+$ indicates a real positive numbers realm support. $\mathbb{R}_+$ is assumed to be bounded and closed, thus implying compactness.

Proof (by contradiction)

1) Suppose that $p > c'(0)$, where $c'(0)$ represents marginal cost at the optimum and $c(q)$ is a strictly convex non-linear cost function, q is a production function twice continuously differentiable, concave in input prices and small epsilon $\epsilon > 0$. Then, consider a first order Taylor expansion near the optimum:

$c(\epsilon) = c(0) + c'(0)(\epsilon - 0)$, where $c(0) = 0$. Then:
$c(\epsilon) = c'(0)(\epsilon) + 0(\epsilon)$.

The maximization problem, then it is represented for:
$\pi(p) = \ Max\ pq - cq \geq p\epsilon - \big(c'(0)(\epsilon) + 0(\epsilon)\big) = \big(p - c'(0)\big)\epsilon + 0(\epsilon) > 0$,
Subject to $q \geq 0$, and $q = \epsilon$.

where $\pi(p)$ stands for a positive profit function.

2) Suppose that $p > c'(0)$ is not true. Then $p \leq c'(0)$. If the cost function is convex, then the next weighted linear combination should hold:

$c(\epsilon) \leq \frac{\epsilon}{q} c(q) + \left(1 - \frac{\epsilon}{q}\right) c(0) = \frac{\epsilon}{q} c(q)$.

Thus, for every $(\epsilon, q) > 0 \Rightarrow \frac{c(\epsilon)}{\epsilon} \leq \frac{c(q)}{q} \Rightarrow \lim c(\epsilon) = c'(0)$, when $\epsilon \to 0$. Applying this limit to the last inequality:

$c(q)q \geq c'(0)q$ for all $q \geq 0$, thus $c(q) \geq c'(0)$. Then profits at the maximum are bounded from above, given any q:

$pq - cq \leq pq - c'(0)q = \big(p - c'(0)\big)q \leq 0 \Rightarrow$

$\pi(p) = \ Max\ pq - cq \leq 0$,
Subject to $q \geq 0$ and $q = \epsilon$.

Contradiction: negative profits! So, it must follow that at the optimum, positive profits are positive. This imply that:

$pq \geq c'(0) \Rightarrow p \geq \frac{c'(0)}{q}$ and $q \geq \frac{c'(0)}{p}$, where $c'(0) = 0$ by invoking the envelope theorem. Thus:

$p \geq 0$ and $q \geq 0$.

Therefore, BPA1 and BPA2 hold. These attributes ensure a positive, increasing and concave production function frontier.[57] These findings complete the proof. ∎

[56] BPA1 and BPA2 can be verified empirically, by checking the corresponding database at the national statistic offices.
[57] À-la Mitscherlich-Baule yield function.