PhD thesis proposal

Thesis title: Assessment of risk in the agricultural sector using statistical copulas

PhD candidate: Osama Ahmed Nasreldin Ahmed.

Director: Dr. Teresa Serra, Center for Agro-food Economy and Development (CREDA-UPC-IRTA). E-mail: teresa.serra-devesa@upc.edu.

Abstract: While risk is not exclusive of agricultural economic activities, the agricultural sector usually faces a combination of risks that is rarely found in other enterprises. In order to develop sound risk management strategies, it is important to understand the nature of risk: its origin, its distribution and correlation with other risks, and the capacity of several instruments to reduce this risk. While sound risk assessment is important in a variety of circumstances, this thesis identifies three of them which appear specially relevant and timely. First, the dependence between agricultural yields and prices, which is key for the implementation of crop revenue insurance programs. This implementation is currently being considered in Spain. Second, price volatility of food staples, which has important implications for food security. This has become specially evident after the recent food crises. Third, price volatility transmission along the food marketing chain, which has important consequences for the usually risk-averse economic agents participating in the market. Characterization of these risks requires assessing risk dependence between random variables: producer price-yield and consumer price-producer price. From a methodological perspective, this characterization is complicated by the lack of multivariate distributions in statistics. Recent research has suggested the use of statistical copulas for such purpose. Copulas are statistical instruments that combine univariate distributions to obtain a multivariate distribution with a particular dependence structure. They present the advantage of being based on univariate distributions, instead of multivariate ones. This is especially important given the scarcity of multivariate distributions available from the statistical literature. This thesis will shed light on the risks mentioned above by means of using statistical copulas.

Introduction

While risk is not exclusive of agricultural economic activities, the agricultural sector typically faces a combination of risks that is rarely found in other enterprises. Two of the major risks affecting agriculture are climatic and natural risks that can influence agricultural yields, and market risks that may lead to important agricultural price fluctuations. Recent progressive dismantling of public commodity price stabilization mechanisms[1] leading to increased dependence of prices on global markets, may have increased price risk. Price risk has also increased after the recent 2007/08 and 2010/11 global food crises, which have led to unprecedented changes in food price variability. A farm's economic viability can be challenged by these risks that can cause relevant declines in farm incomes and profits. Further, in less developed countries, price risk can also threaten households' standard of living, as well as food security.

In order to develop sound risk management strategies, it is important to understand the nature of risk: its origin, its distribution and correlation with other risks, and the capacity of several instruments to reduce this risk (Hardaker et al 1997). Several strategies can be used by the agricultural sector to manage risk. A non-exhaustive list includes marketing or production contracts, use of derivatives, agricultural insurance, or food reserves.

While sound risk assessment is important in a variety of circumstances, three of them appear specially relevant and timely. First, while agricultural insurance has traditionally covered yield risks, more recently revenue insurance has gained strong

[1] An example can be found in the progressive transformation of the Common Agricultural Policy from a policy mainly based on price supports to a policy based on direct payments.

interest. A sound implementation of this sort of insurance requires reliable assessment of price and yield dependency. The first objective of this thesis is to study this dependency in order to shed light on the possible consequences of an implementation of a revenue insurance program in Spain.

Second, price volatility of food staples is one of the most complex factors affecting food security and one of the least studied. Further, since risk may affect food producers and consumers differently it is also relevant to assess how price risk is transmitted along the food marketing chain, from producers to final food consumers. This thesis will study this issue in Niger, one of the poorest countries in the world.

Third, while food price-level transmission along the marketing chain in developed economies has been widely assessed by previous research, price volatility transmission has not received much research attention. Since the food price crisis in 2007/2008, economic research has paid substantial attention to price volatility, given the significant impact it has on the economy. The third objective of this thesis is to focus on volatility spillovers along the Spanish tomato marketing chain. Selection of the Spanish tomato sector responds to its relevance, not only within the Spanish agriculture, but also within the European and global markets, given the leadership of Spain in world tomato markets.

The objectives outlined above require assessment of dependency between variables: producer price – yield dependency and producer – consumer price dependency. From a methodological point of view, dependency assessment is complicated by the lack of multivariate distributions in statistics. Recent research has suggested the use of statistical copulas for such purpose. Copulas are statistical instruments that combine univariate distributions to obtain a multivariate distribution with a particular dependence structure. They present the advantage of being based on univariate distributions, instead of multivariate ones. This is especially important given the scarcity of multivariate distributions available from the statistical literature.

State of the art

The analysis of dependency between risks is a central area of research within the agricultural, financial and actuarial economics literature. With the launching of agricultural revenue insurance programs, which was specially relevant in the US by the end of the past century, the modeling of risk dependence between prices and yields has received increasing research attention within the agricultural economics field (USDA, 2001).

While revenue insurance schemes have been already successfully implemented in countries such as the US (through different programs such as Crop Revenue Coverage (CRC), Income Protection (IP), or Revenue Assurance (RA)), Spain is currently at the stage of considering its implementation. Agroseguro, the pool of all the Spanish agricultural insurance companies, has recently commissioned a series of studies that aim at assessing how these insurance programs should be put into practice. Yet, no published research paper has tried to shed light on the economic implications of doing so. A challenge affecting the implementation of revenue insurance programs is the computation of an actuarially fair insurance premium, or the amount charged for coverage, that should explicitly take into consideration the dependence between price and yield risks. The relevance of joint consideration of risks is manifest in that periods of low yields may be accompanied by high prices. This would lead to lower fair premium rates than if declines in both yields and prices occurred at the same time. In short, to design a revenue insurance contract it is necessary to understand the usually negative relationship between agricultural yields and prices. If this relationship is ignored, the risk will likely be over-estimated. While numerous research articles have been published on the proper modeling of agricultural yield risk, the literature focusing on price and yield risk

dependence is relatively new (Tejeda and Goodwin, 2008; Zhu et al., 2008; Woodard et al., 2010; Ghosh et al., 2011).

This thesis will assess dependence between prices and yields in the apple sector in Spain. With 533.4 thousand hectares that yield 10.7 million tons of output, apple is the most relevant fruit cultivated within the EU. It represents 9.02% of the EU's fruits harvested area and 18.24% of the EU's fruits production. Spain is a major player in the €90 billion global apple industry. It produced 646.2 thousand tons in 2010, being its per capita production the 11th highest in the world and the 4th highest in the EU (FAOSTAT, 2010). Shedding light on dependency between apple prices and yields is thus especially relevant for policy makers, insurance companies and farmers.

Food price volatility in less developed countries (LDCs) is another key research area within the risk literature. This thesis will also pay attention to study price volatility of food staples in Niger. The 2012 Human Development Index (HDI) ranked Niger 186 out of 187 countries (UNDP, 2013). Niger agriculture is overly influenced by a harsh climate and geography. Subsistence agriculture and stock rearing represent 40% of the Niger Gross Domestic Product (GDP), being the second most relevant economic activity after services and employing three-quarters of Niger labor force (Geesing and Djibo, 2006). Rough climatic conditions and market price volatility bring instability to food supply, exacerbating chronic food insecurity and poverty. Guaranteeing availability and access to food is vital in LDC economies, and can be enhanced in a number of different ways. Local food reserves, for example, have been promoted by different organizations and small producer federations with the objectives of increasing farm income and food security.

Despite their potential to promote food security, there is an important failure rate of local food reserve initiatives in LDCs largely due to market risks. Price risk remains the most complex factor affecting vulnerability of local food reserves and the least studied

(Oxfam, 2012). Lack of food price data in LDCs is the reason underlying the scarcity of studies on price behavior in these countries. This makes the contribution of the analysis proposed in this thesis a very relevant one.

Understanding price behavior along the chain is useful for economic agents when taking their decisions, as well as to policy makers and competition regulatory authorities. Hence, the link between different prices at different levels of the food marketing chain is a very interesting research topic not only in LDCs, but also in developed economies. Vertical price transmission in developed economies has been the research focus of several economic analyses. The vast majority of these studies have paid special attention to the interdependence of price levels. In contrast, price volatility interactions have received very little interest. A third objective of this thesis is to shed light on this question, by assessing vertical price transmission in the Spanish tomato sector. Tomato production is a very relevant economic activity within Spain, which increases the interest of the thesis: with more than 3.8 million tons in 2011, Spanish tomato production represents a quarter of the European Union's tomato production, ranking second behind Italy (FAOSTAT, 2012).

As noted above, the research objectives of this thesis require assessment of dependency between variables: price – yield dependency and producer – consumer price dependency. Research in this area should thus be based on flexible instruments that soundly capture the joint distribution function of the variables considered. Rather common instruments used by the literature to assess the links between variables are correlation techniques such as the Spearman's rank correlation and Kendall's tau correlation coefficient. An important limitation intrinsic to these techniques is that a single correlation coefficient is used to characterize dependence over the whole range of the distribution. Usually, this is not enough to describe dependence between random variables. For example, dependence in the extreme tails of the distribution may be

different from dependence in the central areas and may be more relevant from a risk management point of view, i.e., insurance companies might be more interested in the dependence of yields and prices during extreme weather or market events than during more frequent and less drastic events.

Recent research has suggested the use of statistical copulas to assess dependence. Copulas are statistical instruments that combine univariate distributions to obtain a joint distribution (multivariate distribution) with a particular dependence structure. They present the advantage of being based on univariate distributions, instead of multivariate ones. This is especially important given the scarcity of multivariate distributions available from the statistical literature. These multivariate distributions include the normal and the t-Student and have been shown as inappropriate to assess behavior of the type of data we intend to study.

Methodology

Copulas allow a flexible characterization of the dependence structure between random variables and are especially useful if no obvious choice for the multivariate density function exists. The use of copulas in the economics literature is rather recent and most empirical applications are found within the financial economics literature (see, for example, Patton, 2004 and 2006; or Parra and Koodi, 2006). More recently Serra and Gil (2012) have used copulas to study dependence between crude oil and biofuel prices. The agricultural economics literature has started to pay attention to copulas as a proper way of modeling dependence between prices and yields. This PhD thesis will make use of copulas to assess agricultural risks.

A copula function is a multivariate distribution function defined on the unit cube $[0, 1]^n$, with uniformly distributed marginals. Copulas are based on the Sklar's (1959) theorem that implies that, for multivariate distribution functions, the univariate margins and the multivariate dependence structure can be separated and the dependence structure represented by a copula. The Sklar's theorem allows the researcher to focus on modeling univariate distribution functions which usually leads to the construction of better models (Patton 2006). The dependence structure is fully represented by the copula. This contrasts with the use of correlation coefficients between random variables as a measure of dependence. While correlations are highly popular due to the ease with which they can be calculated, they can be very misleading if random variables are not jointly elliptically distributed.

Let $F_1,...,F_n$ be the univariate distribution functions of n random variables $(x_1,...,x_n)$. H represents the joint distribution function. According to the Sklar's theorem, there exists a copula C that can be expressed as (see Embrechts et al. 2001):

$$H(x_1,...,x_n) = C(F_1(x_1),...,F_n(x_n)) = C(u_1,...,u_n)\,. \tag{1}$$

Hence, C is an n-dimensional distribution function with uniformly distributed margins $u_i \sim Unif(0,1),\ i = 1,...,n$. The joint density function can be expressed as:

$$h(x_1,...,x_n) = \prod_{i=1}^{n} f_i(x_i)\ c(u_1,...,u_n)\,, \tag{2}$$

where c is the copula density and $f_i(x_i)$ are the univariate density functions of the random variables.

As shown by Patton (2006), consistent and asymptotically normal copula parameters can be obtained by means of a two-stage estimation procedure, in which marginal distribution models are estimated in a first stage and the copula model in the second stage. Patton's (2006) proposal will be followed.

Different copula specifications represent different dependence structures. Our empirical analysis will consider both elliptical (Gaussian and Student's t-copula) and Archimedean (Symmetrized Joe-Clayton-SJC copula) copulas. Archimedean copulas are a group of associative copulas that have the advantage of reducing dimensionality issues in the estimation process. Copulas may also be categorized as static and time-varying. A static copula relies on the assumption that parameters are constant over time, while a dynamic copula allows the parameters to change with changing environment. The range of copulas considered in the empirical analysis will be restricted by estimation convergence issues.

Data and research plan

Data collected for the purpose of conducting the thesis are described in the following lines. The research on crop insurance will be based on annual average prices and yields for apple for the period from 1954 to 2010, which yields a total of 57 observations. Apple prices are expressed in constant 2010 euros per 100 kilogram. Average apple yields in Spain are expressed in tons per hectare. Data were obtained from the Spanish Ministry of Agriculture, Food and Environment (MAGRAMA 2010).

Data to conduct the research on Niger millet markets was made available by Intermon Oxfam and consist of monthly millet producer and consumer prices in Maradi

Tillabéri, and Filingué, three relevant consumption markets, for the period from 1990 to 2011. Prices are expressed in CFA per kilo.

The third analysis will be based on weekly prices for tomatoes expressed in euro/kg, and observed from the first week of 2004 to the last week of 2011, leading a total of 416 observations. Prices at different levels of the marketing chain have been collected: the price received by producers and wholesalers and the price paid by consumers. The three series are obtained from the Food Prices Observatory (Spanish Ministry of Environment, Rural and Marine Affairs, 2012).

The thesis will be structured in a journal article format containing the following chapters:

1. **Chapter 1**: will consist of an ***Introduction*** stating aims and focus of the study, identifying its significance, and setting the frame and sequence for each of the papers that follow.
2. **Chapter 2:** will be devoted to the economic analysis of the introduction of agricultural revenue insurance contracts in Spain
3. **Chapter 3:** will focus on food price volatility of food staples. The case of millet in Niger
4. **Chapter 4:** will concentrate on vertical price volatility transmission along the food Spanish tomato marketing chain.
5. **Chapter 5:** the final chapter will provide integrative ***Conclusions***, pulling together all the work described in the core chapters of the thesis (i.e., chapter 2 to chapter 4) and relating this back to the issues raised in the Introduction.

The work plan for the development of the thesis is as follows

		First Year						Second Year						Third Year					
Tasks \ **Bimonthly**		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Ch. 1	Introduction and Literature review																		
Ch. 2	Economic analysis of the introduction of agricultural revenue insurance in Spain (conducting the research and writing up of the paper)																		
Ch. 3	Food price volatility of food staples (conducting the research and writing up of the paper)																		
Ch. 4	Vertical volatility transmission along the Spanish tomato marketing chain (conducting the research and writing up of the paper)																		
Ch. 5	Writing up the introduction and concluding remarks																		

Expected research results

Considering the objectives, the nature and project duration, the thesis will lead to two main types of academic results: scientific paper writing and congress presentation. I expect to elaborate at least three scientific articles. The articles will be presented at different national and international congresses and seminars related to Risk Economics, Food Policy, Empirical Economics, Development and Agricultural Economics. They will also be sent for publication to top referred journals in the field of economics.

References

FAOSTAT (2010). Food and Agriculture Organization of the United Nations. Dataset. http://faostat.fao.org/site/377/default.aspx#ancor. Accessed 17 June 2012.

FAOSTAT (2011). Dataset: http://faostat.fao.org/site/291/default.aspx. Accessed March, 2013.

Geesing, D., Djibo, H. (2006). Country pasture/forage resource profiles. FAO. http://www.fao.org/ag/AGP/AGPC/doc/Counprof/Niger/Niger.htm. Accessed June, 2013.

Ghosh S, Woodard JD, Vedenov DV (2011). Efficient Estimation of Copula Mixture Model: An Application to the Rating of Crop Revenue Insurance. Paper presented at the Agricultural and Applied Economics Association (AAEA) Annual Meeting, Pittsburgh, Pennsylvania, July 24-26.

Hardaker J, Huirne B, Ruud BM, Anderson JR (1997). Coping with risk in Agriculture, CAB International, Wallingford.

Oxfam (2012). First line of defense: assessing the potential of local food reserves in the Sahel. Oxfam Research Report, January 2013.

Parra H, Koodi L (2006). Using conditional copula to estimate value at risk. J Data Sci 4: 93–115.

Patton AJ (2004). On the out-of-sample importance of skewness and asymmetric dependence for asset allocation. J Financ Economet 2: 130–168.

Patton AJ (2006). Modeling asymmetric exchange rate dependence. Inter Econ Rev 47: 527–556.

Serra T, Gil JM (2012). Biodiesel as a motor fuel price stabilization mechanism. Energ Policy 50: 689–698.

Sklar, A., (1959). Fonctions de répartition à *n* dimensions et leurs marges. Publications de l'Istitut Statistique de l'Université de Paris, 8, 229-231.

Tejeda HA, Goodwin BK (2008). Modeling Crop prices through a Burr distribution and Analysis of Correlation between Crop Prices and Yields using a Copula method. Paper presented at the American Agricultural Economics Association Annual Meeting, Orlando, Florida, July 27-29..

United Nations Development Program (UNDP). (2013). Human development report 2013 – The rise on the south: human progress in a diverse world." New York: UNDP.

USDA (2001). Crop Insurance Handbook (CIH) 2002 and Succeeding Crop Years. Washington, DC: Risk Management Agency, FCIC-18010(6-01).

Woodard JD, Paulson ND, Power G (2010). Efficiency in the Modeling of Dependence Structures: An Application of Alternative Copulas to Agricultural Insurance Rating. Agr Econ 42(IS-1): 101-112.

Zhu Y, Ghosh SK, Goodwin BK (2008). Modeling Dependence in the Design of Whole Farm– A copula-based model approach. Paper presented at the American Agricultural Economics Association Annual Meeting, Orlando, Florida, July 27-29.