Bayesian networks as a tool to assess the multiple effects of agricultural policies in rural areas

Meri Raggi[1], Laura Sardonini[2] and Davide Viaggi[2]
[1]Department of Statistics, University of Bologna, Italy
[2]Department of Agricultural Economics and Engineering, University of Bologna, Italy
Corresponding author: meri.raggi@unibo.it

Paper prepared for presentation at the 118th seminar of the EAAE
(European Association of Agricultural Economists),
'Rural development: governance, policy design and delivery'
Ljubljana, Slovenia, August 25-27, 2010

BAYESIAN NETWORKS AS A TOOL TO ASSESS THE MULTIPLE EFFECTS OF AGRICULTURAL POLICIES IN RURAL AREAS

Many economic analysis tools are used to evaluate the effects of policies on rural development. However, a number of unexplored options are still available from the literature about policy analysis and biophysical systems representations. A particularly important need concerns the representation of the complexity of rural systems either in a static or dynamic framework In this paper we use Bayesian networks, to the best knowledge of the authors, basically ignored by the literature on rural development.

The objective of this paper is to discuss the potential use of Bayesian Networks tools to represent the impact of the Common Agricultural Policies in rural areas.

KEYWORDS: Bayesian Networks (BNs), farm-household, exit.
JEL: Q1 – Agriculture, Q18 - Agricultural Policy; Food Policy

1. Introduction [1]

The development of rural areas in Europe has attracted considerable attention by both policy and research in the last decades. The Common Agricultural Policy (CAP) has played a major role in such context, both providing income for agriculture and rural households (first pillar), and supporting directly Rural Development Programs (RDP) in the second pillar. The series of reforms started at the beginning of the 1990s have progressively reinforced attention to the rural development component of agricultural policy. This attention has been strengthened during the 2003 reform and the health check. Rural development issues are also expected to be central in the next round of reforms, that will lead to post-2013 CAP.

In such context, policy analysis has been a central issue for research on rural development. This has generated a wide literature and the tools to evaluate the effects of policies on rural development components are now a very wide and heterogeneous family. The main challenge can be identified in the need to consider the complexity of the rural context, both at the micro (household, farm, other economic activities) and macro (regional, intersectorial, spatial) levels and either in a static or dynamic framework.

Attempts in this direction are available using SAM approaches or, more consistently with the need of representing multiple links in a flexible way, dynamic networks.

As an example of SAM, Thomson and Psaltopoulos (2007) (see also Balamou et al., 2008) present a combined CGE and SAM model applied to understand the interaction between different rural and urban areas.

An example of system dynamic model of agriculture and rural development was developed in the project TOPMARD (Johnson et al., 2008), that has also been used to simulate policy scenarios, e.g. in Bergman et al. (2008).

A growing stream of regional (intermediate scale) models is that of Agent-based models (AMB), such as Agripolis and RegMAS (Regional Multi Agent Simulator) (Lobianco and Esposti, 2008).

A survey of different model exercises and attempt to yield an evaluation of scientific knowledge about contribution of the CAP to regional growth, taking into account the effects of different measures and the objectives of the Lisbon agenda is provided by Esposti (2008).

[1] The research reported in this paper was funded by the European Commission within the project "Assessing the multiple Impacts of the Common Agricultural Policies (CAP) on Rural Economies" (CAP-IRE), 7th Framework Programme, contract n. 216672 (www.cap-ire.eu). However the paper does not necessarily reflect the views of the EU and in no way anticipates the Commission's future policy in this area.

Another stream of literature focus on the micro level, i.e. level of farm or farm-household. The literature include both programming and econometric models. The latter may focus on the use of secondary information or on the treatment of data coming from ad hoc surveys. The survey approach seems of particular interest in the analysis of rural development policy as direct surveys allow to gain information on a wider amount of information related to social, economic, and personal aspects, as well about specific technical and institutional features of the agents (farms, households) acting in the rural environment. However, the treatment of information derived from these studies often suffers the limitation of econometric tools. In particular, having to do with a number of interrelated variables, the elaboration often reduces to quantify the effect of single variables on one single dependent variable.

In order to deal with these limitations, further methodological options from the literature about policy analysis and biophysical systems representations may deserve to be explored.

In this paper we consider Bayesian networks, a tool that, to the best knowledge of the authors, has been basically ignored by the literature on rural development (with the exception of a previous explorative work of the same authors presented in 119th EAAE Seminar in Capri (Sardonini et al., 2010)).

The objective of this paper is to discuss the potential use of Bayesian Networks tools to represent the impact of the Common Agricultural Policies in rural areas. Within this wider objective we focus in particular on the interaction between exit decision and other structural and land allocation decisions. Also, the main focus is at the micro level, i.e. farm-household and its economic context.

In our specific application, we focus on the interpretation of data obtained through a survey of farm-household, addressing, in particular, the perspective post-2013 behavior facing different policy scenarios.

The structure of the paper is the follow: first we present the background and the methodology introducing the characteristics of Bayesian Networks, then an application to a case study in a Province located in Northern Italy is illustrated. A brief discussion concludes the paper.

2. Background and methodology

The focus of this work is the analysis of stated intention of farming exit in Italy considering a set of characteristics and determinant of the farm household and taking into account the interconnections between the exits and related farm and asset management choices. In the agricultural economics literature, the studies regarding exit behavior are not very numerous and developed (for a review see Mishra et al., 2010). Several aspects can be considered as the causes of this moderated interest; the most important of these causes regards the little data availability, particularly in the form of stated intention related to exit decision. Moreover, the process of exit from farming activity is very long and complex in terms of farmers' reaction, structural change, social conditions and its dependency from other exogenous variables.

As the intention to stop farming is driven by a complex behavior, then the study is influenced by this inner complexity. The main problems about the analysis of the stated intention to exit can be grouped as follows: i) non-linear relation between variables, ii) too many variables should be consider in the analysis with respect to the dimension of available data, iii) high correlation among variables and multiple outcomes are to be taken into account to understand the process. We try to manage these problems using the Bayesian networks tool.

Bayesian networks were developed mostly in the last few decades. In particular, the last decade of the 20th century saw an improvement in instruments for learning Bayesian networks from data. From the first development in artificial intelligence field (NASA, NOKIA software applications), Bayesian networks are increasingly being used for issues in very different areas of research. Fields of applications regard sociology (Rhodes, 2006), medical diagnosis (Kahn et al., 1997; Beinlich, 1989; Long, 1989) and environmental aspects (Marcot et al., 2006).

The methodology used in this work is based on Bayesian Networks (BNs) which "...capture the believed relation between a set of variables which are relevant to some problem" (Netica^{TM}). In theoretical terms a BNs are defined as "Direct Acyclic Graphs (DAGs) where the nodes are random variables and certain independence assumption hold" (Charniak 1991). Using the Bayesian approach compared to the "classical" (frequentist) one, the main difference is in the probability concept. In fact, the classical probability represents a physical relation empirically observed in real world, while in the Bayesian approach this is substituted by the concept of belief degree of that event, associated to a moment.

The BNs method offers some interesting advantages: a) the possibility to use incomplete and small data set avoiding dependence problems between variables because the dependencies are encoded, b) the possibility to learn from data, in fact when the causal relationships are expressed then the model can be used for an explanatory analysis, c) the combination of Bayesian statistical techniques with the domain knowledge and data, so it is possible to add some prior information that the researcher knows especially when data is insufficient or expensive, d) the over fitting of data is avoided when BNs are jointed with other types of models (Heckerman, 1996).

BNs, as the name calls to mind, are based on the Bayesian theorem and on the idea of a conditional dependence. The Bayes theorem permits to obtain the probability for an event B given event A, when the events are dependent then the probability of event B depends on the event A:

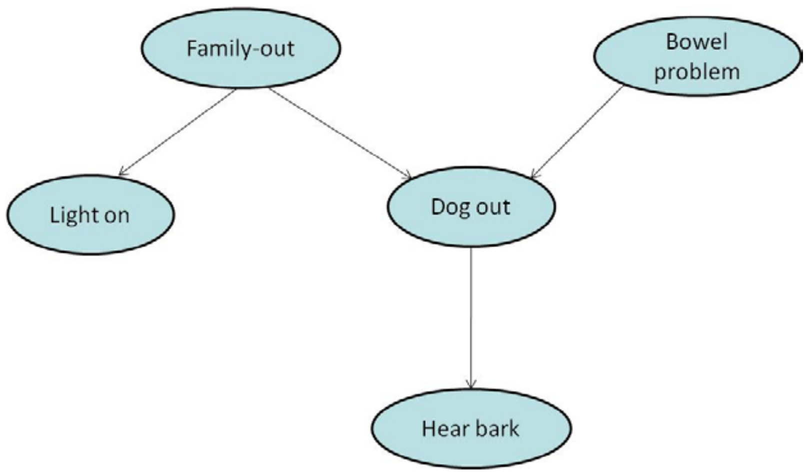
$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} = \frac{P(A \mid B)P(B)}{P(A)} \tag{1}$$

The above relation can be applied in a generalized formulation when we have more than two events. A large number of variables increases the degree of complexity in relationships and the links between variables have to be defined using the principle of the conditional dependence.

The conditional dependence consists in a selection of a subset of variables (parents) that influence other variables investigated (child).

In figure 1, an example of conditional dependences is shown (see Charniak, 1991 for more details). In the figure the circles are called nodes and the arcs represent the dependence connection between the 'parent' and 'child'. A node is called 'parent' because of its influence on node called 'child' The example is important to clarify the idea of modeling a situation in which causality plays a role but where our understanding of what is actually going on is incomplete, so we need to describe things probabilistically. BNs allow to calculate the posterior probability distribution under conditional dependence of the nodes in the network given that the values of the nodes have been observed following the Bayes' rule.

Figure 1. Example of Bayesian Networks (nodes and parents)



In general, given a set of variable X_i, where $i=1,\ldots,N$, it is possible to assume that X_i can be dependent on a subset of $pa(X)$ that $P(X_i| pa(X))$. So $pa(X)$ includes only a specified subset of (X). The reduction to a subset of variables, caused by the conditional dependence relation, implies that the dimension of the model decreases (from the full model considering all the variables) so the inference results easier and simplified. When the complexity of relationships in a net (N) of data (D) increases (i.e. when the number of links imposed are large it is not possible to directly apply the Bayes theorem but it is necessary to use the probabilistic inference, which consists in the process of calculating new beliefs for a set of variables, given some data.

The relation that identifies the probability to obtain that net given data is:

$$P(N \mid D) = \frac{P(D \mid N)P(N)}{P(D)} \qquad (2)$$

where P(N) is the prior probability to have that net, P(D) is the probability of data and P(D|N) is the likelihood which represents the probability to observe that data given a net.

The probabilistic inference is the process of finding a posterior distribution, given a prior distribution and some observations. Bayesian nets do probabilistic inference by belief updating by the data learning (parameters learning).

There are several methods for the parameters learning which determines the Conditional Probability Tables (CPTs) at each node, given the link structures and the data.

Algorithms are based on the maximization of the term P(D|N)P(N), which is the same as maximizing its logarithm: log(P(D|N)) + log(P(N)). Since the term P(D) is constant for each net so the posterior probability depends on P(D|N) P(N) as:

$$P(N \mid D) \propto P(D \mid N)P(N) \qquad (3)$$

Maximize the likelihood means to find the net which has most likely generated the data. The two terms, for the maximization, are dealing in a different way:

- P(N), in this case is considered as if each net was equally likely, so the term can simply be ignored, since it will contribute the same amount for each candidate net;
- P(D|N) (or logP(D|N) for simplifying computation) is the term to maximize using an iterative process. Starting with a candidate net and reporting its log-likelihood, then the process considers the entire case set with it to find a better net. By the nature of the algorithm the log-likelihood of the new net is always better than the previous. The process is repeated until the log-likelihood numbers are improved enough according to a specified tolerance.

In this work, the algorithm used for parameter learning is the EM algorithm that takes a Bayes net and uses it to find a better one by performing an expectation (E) step followed by a maximization (M) step. In the E step, the algorithm uses regular Bayes net inference with the existing Bayes net to compute the expected value of all the missing data, and then the M step finds the maximum likelihood Bayes net given the now extended data (i.e. original data plus expected value of missing data). The EM algorithm returns robust parameter estimation.

The result consists in the estimation of the posterior distribution for each variable defined as child. The posterior distribution is estimated considering the data evidence (likelihood). Moreover, the CPT is estimated reporting the conditional probability for each child level given all the possible parents level combinations.

3. Case study

The empirical application is based on a case study in the Province of Bologna (Emilia Romagna) using survey data from the project CAP-IRE “Assessing the multiple Impacts of the Common Agricultural Policies (CAP) on Rural Economies”, 7th Framework Programme. The network is structured in nodes based on the variables available from the data collected from 300 farm households.

The sample was selected casually from the population of the beneficiaries of the Single Farm Payment (SFP) in 2007, stratifying for altitude (plain, hill and mountain) and below or above the average amount of the SFP. The interviews were made by telephone in the beginning of 2009 and the questions concerned both the farming activity and the household in terms of: structure, innovation, chain supply, environment, social aspects and governance.

The questionnaire was intended to collect information both about the present situation of the farm and household, and about their future. Stated intentions about the future were collected under two hypothetical policy scenarios. In the first scenario called ‘Cap scenario’ (baseline) it is assumed that the CAP remains the same after 2013 and in the second one, called ‘No-Cap’, it is assumed that the CAP will be removed after 2013.

The variables used in the Bayesian network structure could be divided in three groups as household, farm and the stated intention (Table 1).

Table 1. Variables used in the analysis

Group	Variable	Description
Household	AGE	Classification of farmers by age. Young=less 40 years old, adult= between 40 and 65 years old and old=more than 65 years old
	MALE	Number of male in the household
	FEMALE	Number of female in the household
	YOUNG	Number of young in the household less than 18 years old
	LIVE_ON_FARM	Household lives on farm (yes/no/do not know)
	FULLTIME_HOUSEHOLD	Number of fulltime worker from the household
Farm	ALTITUDE	Location of the farm (plain, hill and mountain)
	HA	Dimension of the farms in terms of owned land (ha)
	RENT	Classification of rent behaviour between farmers: rent_in, out_rent, no_rent or do not answer
	SPEC	Mainly farm specialisation
	INCOME FROM FARM	Percentage of total household gross revenue coming from farming
	CAP	Policy scenario after 2013 yes= Cap (Baseline) and no= No-Cap
Stated intention	INTENTION	Intention to stay in farming activity in the next 10 years (stay, exit, other and do not know.)
	WHY EXIT	Main reason that influences the stated intention to exit (High risk of farming, No successor within the family and /or too old, Not profitable enough, Other)
	FARM FUTURE	Main plan to dispose farm in the future (I kept the farm and rent it out, I would like to sell the farm, Other)

In Table 2 and Table 3 the frequencies of the variables in the sample are reported. For the household (Table 2) we reported the frequencies of age of respondent, the number of male, female and young, percentage of those live on farm and the number of fulltime household members.

Table 2. Frequency of variables regarding the household (%)

AGE

adult	missing	old	young	Tot
36.67%	1.00%	52.00%	10.33%	100.00%

MALE

only 1	from 2 to 4	more than 4	missing	Tot
42.67%	56.00%	0.67%	0.67%	100.00%

FEMALE

only 1	from 2 to 4	more than 4	missing	Tot
57.67%	41.00%	0.33%	1.00%	100.00%

YOUNG

none	1	2	3	4	missing	Tot
74.00%	12.33%	10.67%	1.33%	0.67%	1.00%	100.00%

LIVE_ON_FARM

yes	no	do not answer	Tot
78.67%	21.00%	0.33%	100.00%

FULLTIME_HOUSEHOLD

less than 2	from 3 to 4	more than 5	missing	Tot
73.67%	6.33%	0.67%	19.33%	100.00%

For the farm (Table 3) we reported the frequencies of the altitude of the farm, land owned, main specialisation, and percentage of household income from farm activity.

Table 3. Frequency of variables regarding the farm (%)

ALTITUDE

hill	mountain	plain	Tot
29.33%	19.67%	51.00%	100.00%

HA

from 1 to 5	from 10 to 15	from 15 to 30	from 5 to 10	missing	more than 30	Tot
21.33%	12.67%	20.67%	24.33%	9.67%	11.33%	100.00%

NEW_RENT

do not answer	no_rent	out_rent	rent_in	Tot
1.50%	58.17%	5.33%	35.00%	100.00%

SPEC

missing	Mixed cropping	Mixed crops - livestock	Specialist field crops	Specialist granivores	Specialist grazing livestock	Specialist permament crops	Tot
1.33%	4.00%	2.00%	67.33%	0.67%	8.00%	16.67%	100.00%

INCOME FROM FARM

missing	less than 10%	from 10 to 49%	from 50 to 89%	more than 89%	Tot
12.67%	30.00%	21.67%	14.67%	21.00%	100.00%

The variables in Table 2 and Table 3 are considered the nodes in the net.

In Table 4, the frequencies of the stated intention about the reaction to the policy scenario of the sample are reported. In particular, in this work the intention to the policy scenarios, the motivation of exit to the farm activity and the stated intention to the farm future are considered.

Table 4. Frequency of variables regarding the stated intention (%)

	CAP	
INTENTION	**no**	**yes**
do not know	13,00%	7,33%
exit	29,00%	15,33%
other	1,00%	1,33%
stay	57,00%	76,00%
Total	100,00%	100,00%

	CAP	
WHY EXIT	**no**	**yes**
Do not answer	0,67%	0,00%
High risk of farming	1,33%	0,00%
No successor within the family	4,67%	4,67%
Not applicable	71,00%	84,67%
Not profitable enough	16,33%	4,33%
Other	6,00%	6,33%
Total	100,00%	100,00%

	CAP	
FARM FUTURE	**no**	**yes**
Do not answer	6,00%	3,00%
Do not know	5,00%	2,67%
I kept the farm and rent it out	8,00%	4,00%
I would like to sell the farm	5,33%	3,00%
Not applicable	71,00%	84,67%
Other	4,67%	2,67%
Total	100,00%	100,00%

The Bayesian network was constructed selecting as explanatory variables those regarding farm and household. The relevant connections are identified based on economic theory and the preliminary results of thematic analyses carried out within the project. The output focuses in particular on farm exits as a key issue in the future of rural areas and on the role of the CAP in preventing farm exit.

The first step for the inference process is to define the links between variables and their direction, then the conditional dependencies are estimatedexplained. Figure 2 and Figure 3 report the same graph representing the causal relationships between variables, under the two policy scenarios. The BN was computed using Netica^{TM} software (Norsys Software Corp., 2002). Once calibrated to EM learning, the model allows to simulate the effect of farm determinants on the joint decision about selected behavior variables regard the intention to quit farming activity and its linked consequences.

In both nets, the following nodes will be considered as output:

1. INTENTION which consists in the intention of farmers to stay/exit in farming under the hypothesis of policy scenarios after 2013;
2. WHY EXIT which represents the main motivation of those deciding to quit farming activity;
3. FARM FUTURE which represents the main intention of farm future of those quitting the farming activity.

The structure construction of the net is the same in both figures; however, the node CAP considers answers given under the baseline scenario in Figure 2, and answers given under the No-Cap scenario in Figure 3. The change in frequencies in the node CAP means that the two nets are estimated under two different policy scenarios.

The structure of the net is quite complex in terms of amount of variables and articulated because of links. The node INTENTION is influenced by node RENT, AGE, CAP, ALTITUDE, HA and FULLTIME_HOUSEHOLD. Three of these nodes (AGE, CAP, ALTITUDE) are parentless but the others three depend on some variables. In detail, the node RENT depends on SPEC and HA; HA depends on SPEC, INCOME_FROM_FARM, LIVE_ON_FARM (depending on ALTITUDE) and FULLTIME_HOUSEHOLD (depending on MALE, FEMALE and YOUNG). The node WHY EXIT depends on INTENTION and INCOME_FROM_FARM while node FARM FUTURE depends on WHY EXIT, RENT and INTENTION.

The information included in the CPT could not be reported in this paper as the related tables revealed too large. We however account for the main results detectable from the CPT. Specifically, the probability of stop farming activity increases in No-Cap scenario with respect to baseline; it also increases for old farmers which have a small farm and live in hilly areas. About node WHY EXIT, the reason ‘lack of successor within the family and/or too old’ captures the most likelihood motivations for those having an income from farm between 50% to 89%, the reason ‘not profitable enough’ is most likely for those having an income from farm of more than 90% and the option ‘high risk of farming’ is not so important. The node FARM FUTURE shows which will be the farm future after stopping farming activity, in particular the option to ‘keep the farm and rent it out’ is the most likelihood for those not renting land and thinking about an high risk of farming, for those renting out and having a lack of successor within the family and/or too old, for those renting-in and considering the faming activity not profitable enough. The option ‘I would like to sell the farm’ is the most likely for those not renting land and that do not have successor, for those renting-out or renting-in and thinking that the activity is not profitable enough.

Figure 2. Bayesian Networks (Cap Scenario)

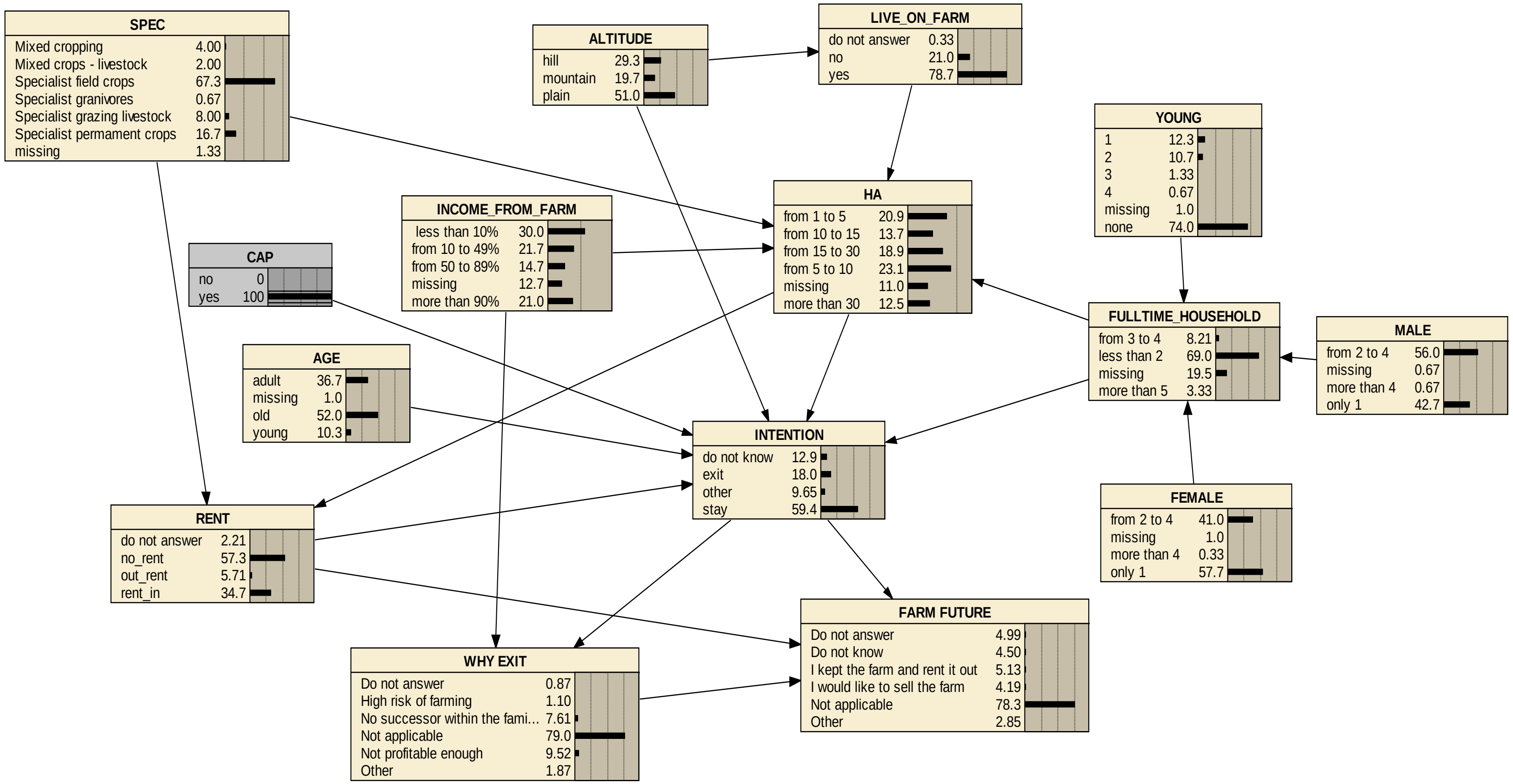
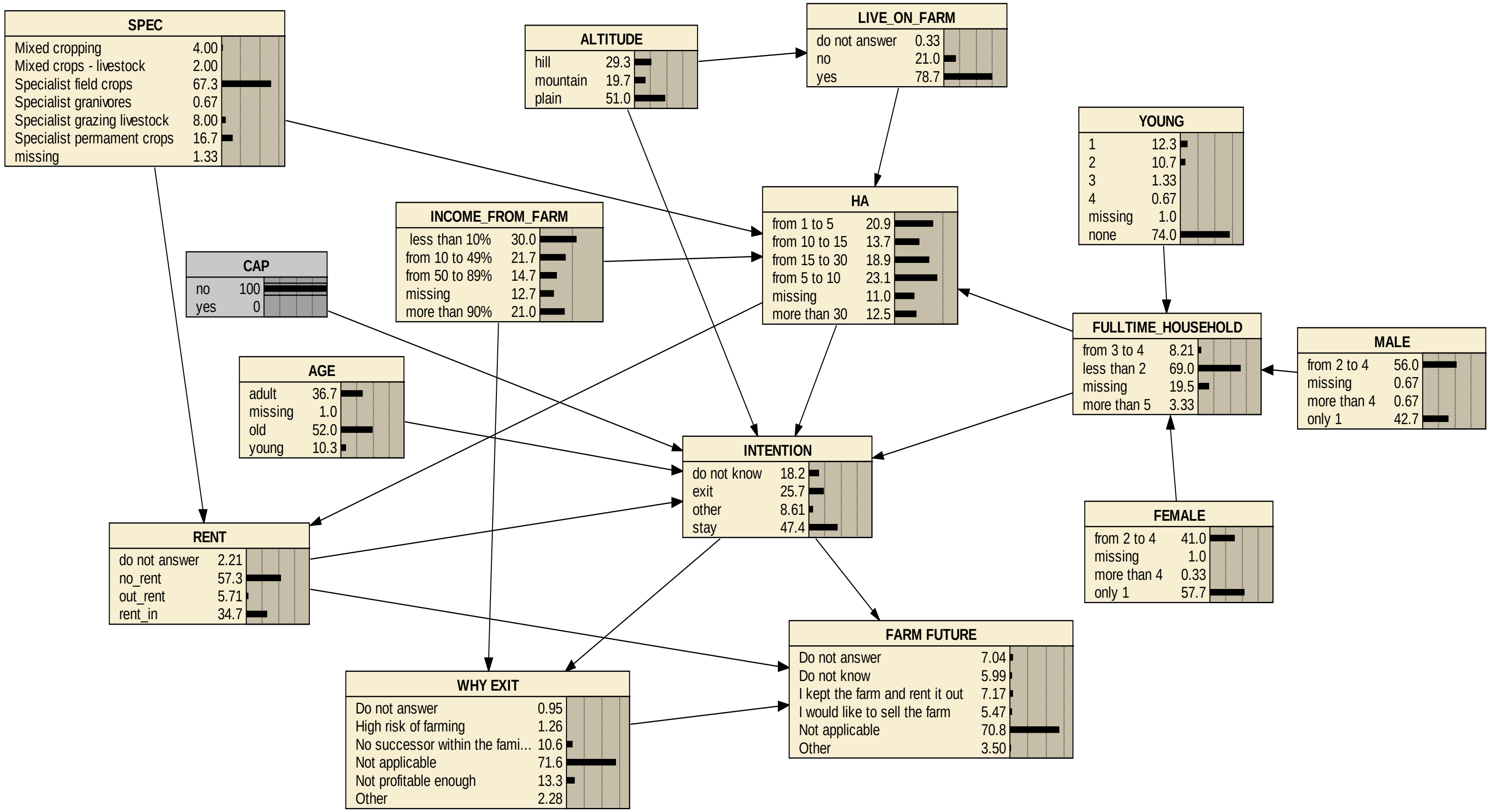


Figure 3. Bayesian Networks (No-Cap Scenario)



The effect of the Common Agricultural Policy can be identify comparing the results in Figure 2 and Figure 3 where the differences between output nodes (child) have to be associate to different policy scenarios. In fact, each net reproduces an estimation of farmers' behavior under the two policy scenarios in a probabilistic view. Comparing the node INTENTION between the two scenarios, the frequency of exit increases from Cap (18%) to No-Cap scenario (25.7%). In both scenario in child nodes WHY EXIT and FARM FUTURE, the main reason of quitting is 'not profitable enough', then 'lack of successor' and 'high risk of farming'; the future of farm for those quitting is to 'keep the farm and rent it out' and secondly 'sell the farm'.

The goodness of the net can be investigated considering the error rate for each child nodes. In Netica™ it is possible to check how the net is able to classify the data. This application can be made in two ways (Table 5): a) checking one node (variable) at a time. The node under checking is classified considering the net structure and all other variables known; b) checking all child nodes (variables) at the same time given the net structure and the other variables known.

Table 5. Error rates for child node

	I	**II**
	Error rate	Error rate
INTENTION	6,00%	19,67%
WHY EXIT	8,17%	19,00%
FARM FUTURE	12,67%	19,33%

In Table 5 the error rates are reported in both ways. As expected when the amount of knowledge increases then the error rates decrease. In the second way the increasing of error rate is to give to the few observations of those quitting farming activity, in fact the large part of the sample has intention to stay, this means that in WHY EXIT and FARM FUTURE the most frequent category chosen by the net estimation is 'not applicable'.

4. Discussion and final remarks

The results in the previous section show the capacity of the net to describe several aspects of the phenomena of farm exit as embedded in its complex determinants, co-decisions and downstream effects. One advantage consists in the simplicity of representation by a graph that describes intuitively the basis of the relationships. Data are well represented by the complexity of the net and this characteristic can solve the difficulties of interpretation of results when other methodologies are used (MNL, etc..). A complex net allows to identify and to understand multiple outcomes (child nodes) considering the structural aspects of the farm and social characteristics of the household.

This preliminary application shows the potential interest for this tool in studying complex rural development issues. The main advantages would be in the flexibility of use and in the ability to use data from different sources, with a variety of functional relationships. It also highlights the need to improve the use of this tool through more robust criteria for network design (identification of nodes and links). In fact, at this stage, the structure identification is obtained by the researcher based on preliminary analysis of individual issues carried out in the project CAP-IRE and supported by economics theory. For this reason one of the issues to develop is the structure learning procedures for the net (before parameter learning). Structure learning allows the identification of the causal relationships structure between variables. The structure learning is based on some statistical test (i.e. chi-squared or mutual information) to detect network structure from the dataset. When the conditional independence relationships among the nodes are found, these relationships are used as constraints to construct a BN. These algorithms are referred as dependency analysis-based algorithms or constraint-based algorithms (Cheng, 2002).

In addition to this, two main directions for the future development of this research should be considered: a) complementing the BN with additional information external to the basic dataset; b) using the BN to simulate systems behavior under different exogenous conditions.

About point a., the idea is to add further nodes and connections based on other source of information, to complement and extend the information available from the survey. This is in fact one of the most interesting opportunities provided by BNs, widely used in contexts in which a consistent survey-based set of data is not available (for example ecological systems). Point b. concerns the use of the model to provide simulation of exit from farming, assuming different probability distributions of one or more variables in the external parent nodes.

Changing the probability distribution and updating the net, it is possible to infer the farmers' exit behavior throughout the system. This use of BNs results particularly useful in order to extrapolate the estimated system structure and behavior to regions different from the ones from which the data was used, which could be very relevant in addressing multilevel and multiregional issues. In addition, this could potentially provide for simulation of the impact of changing structural parameters (e.g. farm size) on downstream indicators (e.g. farm future), which could be very useful as a basis for stakeholder involvement and during the policy design phase.

5. References

Balamou, E., Pouliakas, K., Reberts, D. and Psaltopoulos, D. (2008). Modeling the rural-urban effects of changes in agricultural policies: A bi-regional CGE analysis of two case study regions. 107th EAAE Seminar "Modelling of Agricultural and Rural Development Policies". Seville, Spain, January 29th - February 1st, 2008.

Beinlich, I., Suermondt, H., Chavez, R., and Cooper, G. (1989). The ALARM monitoring system: A case study with two probabilistic inference techniques for belief networks. Proceedings of the Second European Conference on Artificial Intelligence in Medicine, London, Springer Verlag, Berlin, 247-256.

Bergmann, H., Dax, T., Hocevar, V., Hovorka, G., Juvancic, L., Kröger, M. and Thomson, J.K. (2008). Reforming pillar 2 - Towards significant and sustaiable rural development? 109th EAAE Seminar " The CAP after the Fischler reform: National implementations, impact assessment and the agenda for future reforms". Viterbo, Italy., 20-21 November 2008.

Charniak, E. (1991). Bayesian Networks without tears. *AI Magazine* 12(4):, 50-63.

Cheng J., Greiner R., Kelly J., Bell D., and Liu W. (2002). Learning Bayesian networks from data: An information-theory based approach. *Artificial Magazine* 137: 43-90.

Esposti, R. (2008). Reforming the CAP: An agenda for regional growth? 109th EAAE Seminar 'The CAP after the Fischler reform: National implementations, impact assessment and the agenda for future reforms'. Viterbo, Italy. 20-21 November 2008.

Heckerman, D. (1996). A tutorial learning with Bayesian networks. Technical Report MSR-TR-95-06 Microsoft Research Advanced Technology Division. Microsoft Cooporation. One Microsoft Way. Redmont, WA 98052.

Johnson, T.G., Bryden, J., Refsgaard, K., and Lizárraga, S.A. (2008). A system dynamics model of agriculture and rural development: the TOPMARD core model. 107th EAAE Seminar, Seville, January 29th -February 1st, 2008.

Lobianco, A., and Esposti, R. (2008). The Regional Multi-Agent Simulator (RegMAS) Assessing the impact of the Health Check in an Italian region. 109th EAAE Seminar 'The CAP after the Fischler reform: National implementations, impact assessment and the agenda for future reform'. Viterbo, Italy. 20-21 November 2008.

Long W. (1989), Medical diagnosis using a probabilistic causal network, *Applied Artificial Intelligence*3: 367–383.

Marcot, B.G., Hohenloher, P.A., Morey, S., Holmes, R., Molina, R., Turley, M.C., Huff, M.H., and Laurence, J.A. (2006). Characterizing species at risk II: using Bayesian Belief Networks as decision support tools to determine species conservation categories under the northwest forest plan. *Ecology and Society* 11(2).

Mishra, A. K., Raggi, M and Viaggi, D. (2010). Determinants of farm exit: a comparison between Europe and the United States. 114th EAAE Seminar 'Structural change in agriculture' Berlin, Germany. 15-16 April 2010.

Norsys Software Corp. (1995-2010). Netica™ (www.norsys.com).

Pourett, O. (2008). Introduction to Bayesian Networks. In O. Pourett, P. Natin and B. Marcot (eds), Bayesian Networks: a practical guide to application. John Wiley & Sons, Ltd.

Rhodes, C., and Keefe, E. (2006). Social network topology: a Bayesian approach. *Journal of the Operatioanl Research Society* 58: 1605-1611.

Sardonini, L., Raggi, M., and Viaggi, D. (2010). Assessing the sustainability of agri-food systems through Bayesian networks applications: an exploratory study.. 119th EAAE Seminar 'Sustainability in the Food Sector: Rethinking the Relationship between the Agro-Food System and the Natural, Social, Economic and Institutional Environments'. Capri, Italy. June 30-July 2, 2010.

Shah, A., and Woolf, P. (2009). Python environment for Bayesian learning: inferring the structure of Bayesian Networks from knowledge and data. *Journal of Machine Learning Research* 10: 159-162.

Thomson, K.J. and Psaltopoulos, D. (2007). General Equilibrium Analysis of the Spatial Impacts of Rural Policy. 103rd EAAE Seminar 'Adding Value to the Agro-Food Supply Chain in the Future Euromediterranean Space'. Barcelona, Spain.