
**PRISE EN COMPTE DE L'HÉTÉROGÉNÉITÉ INOBSERVÉE DES EXPLOITATIONS
AGRICOLES DANS LA MODÉLISATION DU CHANGEMENT STRUCTUREL :
ILLUSTRATION DANS LE CAS DE LA FRANCE¹**

Thèse de **Legrand Dunold Fils SAINT CYR**

Analysée par Jean-Marc **BOUSSARD**

Laure LATRUFFE Chargée de recherche, INRA Rennes / directrice de thèse
Laurent PIET Ingénieur de recherche, INRA Rennes / co-encadrant

L'agriculture est hétérogène : une population d'exploitations agricoles choisies au hasard est presque toujours composée d'une multitude d'unités très différentes les unes des autres. Les sous-populations ainsi repérées évoluent dans le temps, certaines exploitations restant dans leur catégorie d'origine d'une année sur l'autre, tandis que d'autres, au contraire, « migrent » d'une catégorie dans une autre. Il est évidemment utile d'essayer de prévoir ces mouvements, en particulier lorsque ceux-ci se font en réponse à des changements de prix ou des mesures de politique agricole. C'est à l'étude statistique de ces problèmes qu'est consacrée la thèse dont il est rendu compte ici.

L'outil de base utilisé pour cela appartient à la famille des « chaînes de Markov ». Une chaîne de Markov permet d'étudier l'évolution d'une population d'objets répartie en n catégories, lorsque le passage d'un individu d'une catégorie dans une autre est soumis à une « loi de probabilité » bien définie. Cet instrument repose donc sur un tableau de dimension n sur n dont les n^2 cellules ij contiennent les probabilités p_{ij} que, d'une période sur l'autre (une période correspondant en pratique à une année), un objet de la catégorie i passe dans la catégorie j . Si l'on connaît ces probabilités, supposées constantes, connaissant les effectifs de chaque catégorie au temps t_0 , on peut évidemment prévoir, avec une erreur elle-même gouvernée par une loi de probabilité dite « binomiale » bien connue, quels seront ces effectifs à n 'importe quel moment ultérieur.

L'originalité du travail qui nous occupe est d'avoir sensiblement perfectionné cet outil. Celui-ci est manifestement un peu naïf, parce que, dans le cas des exploitations agricoles, de quelque manière que l'on définisse les catégories, il n'y a pas beaucoup de raisons de supposer que les probabilités de passage d'une catégorie dans l'autre soient constantes dans le temps et l'espace. D'abord, chaque catégorie observée peut très bien être elle-même composée de « sous-catégories » très différentes les unes des autres, avec des probabilités particulières de « migrer » vers d'autres catégories. Ensuite et surtout, les effectifs de chaque catégorie (et donc les probabilités de « migration ») doivent dépendre de circonstances extérieures telles que des changements de prix, ou de nouvelles mesures de politique agricole. Pour cette raison, si l'on veut être sérieux, il faut représenter dans un « modèle » - un ensemble de relations mathématiques entre des variables « endogènes » à déterminer et des variables « exogènes » explicatives - comment les probabilités p_{ij} dépendent de ces facteurs extérieurs.

¹ Thèse soutenue le 12-12-2016 à AGROCAMPUS OUEST sous le label de l'Université Bretagne Loire LABORATOIRE D'ACCUEIL - INRA (UMR1302). Laboratoire d'études et de recherches en économie sur les structures et marchés agricoles, ressources et territoires (SMART-LERECO)

La construction de tels modèles est réalisée par l'auteur en plusieurs étapes, chacune correspondant à un chapitre de la thèse. D'abord, il essaie de tenir compte de l'hétérogénéité interne aux catégories : il suppose pour cela qu'il existe deux sous-types d'exploitations au sein de chacune des catégories qu'il a définies : les exploitations du sous-type « fixes » ne « migreront » pratiquement pas, tandis que les autres, au contraire, les « mobiles », auront une grande soif de changement. *A priori*, cela revient seulement à dédoubler chacune des catégories initiales. Mais la difficulté, ici, est qu'on ne connaît pas les proportions de ces deux sous-catégories dans chacune des catégories principales, ce qui rend impossible l'utilisation des chaînes de Markov simples. La difficulté est contournée en calculant les effectifs de chaque catégorie chaque année tels que prévus par une chaîne de Markov, mais en laissant ces prévisions dépendre des inconnues que sont justement les proportions d'exploitations « mobiles » ou « fixes » dans chacune des catégories considérées. On peut alors déterminer ces proportions en recherchant celles qui maximisent une « fonction de vraisemblance » définie comme la probabilité d'obtenir des prévisions d'effectifs aussi proches que possible de ce qui a été observé. Le procédé - il est en fait assez classique en statistique, c'est la méthode du « maximum de vraisemblance » - permet d'améliorer significativement la qualité des estimations de probabilité de « migration » d'une catégorie dans l'autre, ce qui était l'objectif initial.

Mais cette idée de maximiser une fonction de vraisemblance pour estimer des probabilités de changement de catégorie peut s'étendre à d'autres problèmes similaires et plus intéressants. De fait, on peut utiliser la même méthode pour estimer non plus des probabilités de transfert supposées constantes, mais les paramètres des relations qui lient ces probabilités à des variables exogènes susceptibles de prendre des valeurs différentes à différents moments, comme on l'a vu plus haut. C'est l'objet des deux dernières parties de la thèse, assurément, les plus originales et les plus novatrices. Pour ne pas alourdir inutilement ce court exposé, je ne rentrerai pas dans les détails mathématiques de l'opération, en vérité assez compliquée. Je me bornerai à donner ici quelques exemples du type de résultats obtenus.

Le procédé est d'abord appliqué à un nombre très restreint de variables exogènes sur les exploitations faisant l'objet de relevés comptables au sein du RICA, le réseau d'information comptable agricole français. Cette population d'exploitations est répartie en trois catégories, « grandes », « moyenne » et « petites », sur la base de la « marge brute standard », un indicateur de l'importance économique des exploitations. On montre ainsi qu'une augmentation de 1% des aides directes du premier pilier de la PAC aura pour conséquence d'augmenter de 0,51 % les chances qu'une exploitation française de taille « moyenne » passe dans la catégorie « grande » au cours d'une année, si l'on ne tient pas compte de la distinction entre exploitations « mobiles » et « fixes » ... Il ne s'agit là que d'un exemple pris au hasard dans des tableaux beaucoup plus complets.

Mais l'auteur va encore plus loin, pour tenir compte du fait qu'une exploitation agricole n'est jamais isolée, et que les décisions de l'une peuvent agir sur celles des autres. Il applique ainsi la méthode à l'évaluation de la « probabilité de survie » de différents types d'exploitations, sur un échantillon tiré des adhérents à la MSA en Bretagne, montrant ainsi, par exemple, l'influence de la pression foncière sur les chances qu'une exploitation donnée perdure l'année suivante.

L'instrument de calcul des probabilités ainsi développé repose donc sur une grande imagination créatrice pour rendre opérationnelles des méthodes au départ assez primitives, qui sont transformées ici en un véritable outil de recherche. C'est là une contribution significative au progrès des connaissances.

Il reste - et c'est dommage - une grande naïveté dans l'analyse économique sous-jacente aux applications fournies ici à titre d'illustration. La classification des exploitations en catégories définies par des seuils de « marge brute standard » permet, en effet, de parler des « petites », des « moyennes » et des « grandes », mais c'est tout de même assez sommaire : une exploitation maraîchère de 10 ha de cultures sous serre sera classée dans la même catégorie qu'un céréalier de 500 ha... De la même façon, la classification sur la base du nombre d'hectares cultivés mettra bien le maraîcher dans une catégorie différente du céréalier, mais sans néanmoins qu'il s'agisse vraiment d'une « petite » exploitation (tandis que, évidemment, un céréalier qui ne cultiverait que 10 ha appartiendrait inévitablement à la catégorie des « petits »). En outre, aucune considération n'est accordée au problème des limites de classes : pourquoi mettre à 50 ha la limite entre les « petites » et les « moyennes », plutôt qu'à 45 ou 55 ?

En fait, dans tout ceci, le cœur du problème est la définition des catégories dans lesquelles on range les populations d'exploitation que l'on étudie. Il est curieux que l'auteur ne songe pas à discuter ce point, qui aurait mérité beaucoup de développements (et même si, là encore, des méthodes mathématiques comme les « analyses en composantes principales » peuvent rendre des services). En particulier, on ne comprend pas bien pourquoi il ne s'est pas intéressé aux « orientations technico-économiques » (OTEX) définies pour le RICA, et pour lesquelles, justement, se pose réellement le problème des passages d'une catégorie à l'autre.

Et surtout, il aurait fallu étudier les causes économiques de cette hétérogénéité qui est paradoxale aux yeux d'un économiste un peu « standard » : puisqu'elles sont toutes soumises au même environnement, les exploitations, toujours « optimales », ne devraient-elles pas être toutes semblables ? Pourtant, elles sont loin de l'être¹ ! Dès lors, la réponse à cette question aurait dû diriger l'analyse et orienter la construction du modèle sous-jacent. La virtuosité mathématique ne saurait en aucun cas remplacer l'analyse économique....

¹ Cf Boussard, J.M. : Hétérogénéité technique et structurelle dans les exploitations agricoles. *Economie rurale* 171, Nov Dec 1986, pp 5-15.