HAL open science

Hétérogénéité inobservée et solutions en coin dans les modèles micro-économétriques de choix de production multiculture

Obafemi Philippe Koutchade

To cite this version:

Obafemi Philippe Koutchade. Hétérogénéité inobservée et solutions en coin dans les modèles micro-économétriques de choix de production multiculture. Géométrie algorithmique [cs.CG]. Agrocampus Ouest, 2018. Français. NNT : 2018NSARE048 . tel-01937310

HAL Id: tel-01937310

https://theses.hal.science/tel-01937310v1

Submitted on 28 Nov 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITE BRETAGNE LOIRE

AGRO CAMPUS OUEST

Obafèmi Philippe KOUTCHADE • 19 janvier 2018

Thèse AGROCAMPUS OUEST
sous le label de l'Université européenne de Bretagne
pour obtenir le grade de
DOCTEUR D'AGROCAMPUS OUEST
Spécialité Sciences économiques

ÉCOLE DOCTORALE • Sciences économiques et sciences de gestion (EDGE)

LABORATOIRE D'ACCUEIL • UMR INRA-AO Structures et marchés agricoles, ressources et territoires – Laboratoire d'études et de recherche en économie (SMART-LERECO)

Hétérogénéité inobservée et solutions en coin dans les modèles micro-économétriques de choix de production multiculture

Raja CHAKIR
Directrice de recherche, INRA / *présidente*

Jean-Paul CHAVAS
Professeur, Université du Wisconsin / *rapporteur*

Alban THOMAS
Directeur de recherche, INRA / *rapporteur*

Salima El-Kolej
Responsable de la filière génie statistique, ENSAI / *examinatrice*

Alain Carpentier
Directeur de recherche, INRA / *directeur de thèse*

RÉSUMÉ *ABSTRACT*

Hétérogénéité inobservée et solutions en coin dans les modèles micro-économétriques de choix de production multiculture

Dans cette thèse, nous nous intéressons aux questions de l'hétérogénéité inobservée et des solutions en coin dans les modèles de choix d'assolements. Pour répondre à ces questions, nous nous appuyons sur un modèle de choix de production multicultures avec choix d'assolement de forme NMNL, dont nous proposons des extensions. Ces extensions conduisent à des problèmes spécifiques d'estimation, auxquels nous apportons des solutions.

La question de l'hétérogénéité inobservée est traitée en considérant une spécification à paramètres aléatoires. Ceci nous permet de tenir compte des effets de l'hétérogénéité inobservée sur l'ensemble des paramètres du modèle. Nous montrons que les versions stochastiques de l'algorithme EM sont particulièrement adaptées pour estimer ce type de modèle. Nos résultats d'estimation et de simulation montrent que les agriculteurs réagissent de façon hétérogène aux incitations économiques et que ne pas tenir compte de cette hétérogénéité peut conduire à des effets simulés de politiques publique biaisés.

Pour tenir compte des solutions en coin dans les choix d'assolement, nous proposons une modélisation basée sur les modèles à changement de régime endogène avec coûts fixes associés aux régimes. Contrairement aux approches basées sur des systèmes de régression censurées, notre modèle est cohérent d'un point de vue micro-économique. Nos résultats montrent que les coûts fixes associés aux régimes jouent un rôle important dans le choix des agriculteurs de produire ou non certaines cultures et qu'ils constituent, à court terme, un déterminant important des choix d'assolement.

Mots-clés : choix d'assolement NMNL, hétérogénéité inobservée, solutions en coin, modèle à paramètres aléatoires, modèle à changement de régime endogène, algorithme EM

Unobserved Heterogeneity and Corner Solution in Micro-econometrics Multicrops Production choice models

In this thesis, we are interested in questions of unobserved heterogeneity and corner solutions in acreage choice models. To answer these questions, we rely on a NMNL acreage share multi-crop models, of which we propose extensions. These extensions lead to specific estimation problems, to which we provide solutions.

The question of unobserved heterogeneity is dealt with by considering a random parameter specification. This allows us to take into account the effects of the unobserved heterogeneity on all the parameters of the model. We show that the stochastic versions of the EM algorithm are particularly suitable for estimating this type of model. Our estimation and simulation results show that farmers react heterogeneously to economic incentives and that ignoring this heterogeneity can lead to biased simulated effects of public policies.

In order to take account of the corner solutions in acreage choices, we propose modelling based on endogenous regime switching models with regime fixed costs. Unlike approaches based on censored regression systems, our model is "fully" consistent from a micro-economic viewpoint. Our results show that the regime fixed costs play an important role in farmers' choice to produce or not some crops and they are, in the short term, an important determinant of acreage choices.

Keywords: *NMNL acreage choice, unobserved heterogeneity, corner solutions, random parameter models, endogenous regime switching models, EM algorithm*

AGROCAMPUS OUEST • Institut supérieur des sciences agronomiques, agroalimentaires, horticoles et du paysage
65 rue de Saint-Brieuc – CS84215 – F-35042 Rennes Cedex
Tél. : 02 23 48 50 00
www.agrocampus-ouest.fr

Remerciements

Je voudrais tout d'abord remercier Alain Carpentier, mon directeur de thèse, pour m'avoir accueilli au sein de l'UMR SMART-LERECO et m'avoir fait confiance afin d'entreprendre ces travaux de thèse. Il est le principal initiateur des questions de recherche traitées dans cette thèse et ces questions n'auraient pu être menées à bien sans son expérience et ses compétences. Merci pour ta disponibilité, tes conseils et les connaissances théoriques et techniques que tu m'as apporté.

Aussi je tiens remercier Fabienne Féménia pour avoir co-encadré cette thèse. Merci de m'avoir consacré de ton temps, que ce soit pendant la réalisation des travaux de recherche et lors de la rédaction du manuscrit, et aussi pour toutes les connaissances théoriques et techniques que tu m'as apporté. Je suis ravi d'avoir travaillé en votre compagnie, Alain et Fabienne.

Cette thèse a été cofinancée par le département SAE2 de l'INRA et Arvalis Institut du végétal à qui j'adresse mes remerciements.

Je tiens également à remercier Jean Paul Chavas et Alban Thomas d'avoir accepté d'être rapporteur, ainsi que Raja Chakir, Salima El-Kolei d'avoir accepté de faire partir du jury.

Un grand merci à tous les membres de mon comité de thèse qui m'ont apporté de précieux conseils, et permis le bon déroulement de mes travaux de recherches : Valérie Leveau, Bénoit Pagès François Piraux et Alexandre Gohin.

Enfin, un grand merci à tous les membres de l'UMR SMART-LERECO de Rennes, qui ont rendu mon séjour au sein cette unité agréable à vivre tout au long de cette thèse.

Sommaire

Remerciements **v**

Sommaire **vi**

Introduction Générale **1**

Partie I. Hétérogénéité inobservée et choix de cultures dans les systèmes de grandes cultures : modélisation et estimation **21**

Chapitre 1. Modèle de choix de production multiculture avec hétérogénéité inobservée **25**

1.1. Modèles de choix de production multiculture 26

1.1.1. Fondements théoriques et structure générale 28

1.1.2. Fonctions d'offre de rendement et de demande d'intrants variables 32

1.1.3. Modèles de choix d'assolement MNL 34

1.1.4. Modèles micro-économétriques de choix de production multiculture : formes fonctionnelles 41

1.1.5. Modèles micro-économétriques de choix de production multiculture : spécifications paramétriques et fonctions de vraisemblance 48

1.2. Modèles micro-économétriques de choix de production multiculture à paramètres aléatoires 53

1.2.1. Modèles micro-économétriques de choix de production multiculture à paramètres aléatoires : intérêts 54

1.2.2. Modèles micro-économétriques de choix de production multiculture à paramètres aléatoires : estimation semi-paramétrique ? 60

1.2.3. Modèles micro-économétriques de choix de production multiculture à paramètres aléatoires : spécification paramétrique 63

1.2.4. Modèles micro-économétriques de choix de production multiculture à paramètres aléatoires : vraisemblance 71

1.2.5. Modèles micro-économétriques de choix de production multiculture à paramètres aléatoires : les estimateurs du MVS 77

Chapitre 2. Modèles micro-économétriques de choix de production multiculture à paramètres aléatoires : Estimation via les algorithmes MCEM et SAEM **81**

2.1. Structure et convergence des algorithmes (G)EM 85

2.2. Algorithme EM et estimation des modèles à paramètres aléatoires 90

2.3. Algorithmes ECM et GEM et estimations des modèles multivariés 91

2.4. Algorithmes MCEM et simulations par échantillonnage préférentiel 93

2.5. Mise en œuvre de l'algorithme MCEM avec échantillonnage préférentiel 97

2.6. Algorithme SAEM 99

2.7. Distribution asymptotique des estimateurs calculés 104

2.8. Calibration des paramètres individuels et diagnostics 105

2.9. Remarques conclusives 109

Chapitre 3. Modèle de choix d'assolement, choix de cultures et solutions en coin **111**

3.1. Modèles de choix d'assolement et régimes de productions endogènes 116

3.1.1. Modèles économétriques fondés sur les conditions de KKT caractérisant les choix d'assolement 119

3.1.2. Coûts fixes d'assolement et de régime de production 125

3.1.3. Modèles de choix d'assolement MNL, choix de cultures et coûts fixes de régime 129

3.2. Modèles micro-économétriques de choix de production multiculture à paramètres aléatoires et avec choix de cultures : spécification 134

3.2.1. Modèles micro-économétriques de choix de production multiculture à paramètres aléatoires et avec choix de cultures : formes fonctionnelles 135

3.2.2. Modèles micro-économétriques de choix de production multiculture à paramètres aléatoires et avec choix de cultures : spécification paramétrique 141

3.3. Arbitrage : richesse de la spécification versus difficulté d'estimation 143

Chapitre 4. Modèles micro-économétriques de choix de production multiculture à paramètres aléatoires et avec choix de cultures : estimation via les algorithmes MCEM et SAEM **145**

4.1. Modèles micro-économétriques de choix de production multiculture à paramètres aléatoires et avec choix de cultures : vraisemblance 147

4.1.1. Notations 148

4.1.2. Vraisemblances 151

4.2. Cadre général de mise en œuvre des algorithmes de type (G)EM 155

4.2.1. Intégration par simulations sur la distribution des paramètres aléatoires 158

4.2.2. Etape M généralisée 159

4.3. Rendements, quantités d'intrants et assolements « manquants » 164

4.4. Algorithmes MCEM 168

4.5. Algorithme SAEM 170

4.6. Quelques remarques sur la mise en œuvre pratique des procédures d'estimation présentées 175

Partie II. Hétérogénéité inobservée et choix de cultures dans les systèmes de grandes cultures : applications 177

Chapitre 5. On the heterogeneity in farmers' responsiveness to economic incentives: Random parameter multi-crop models and the effects of EU biofuel support 179

Introduction 180

5.1. Unobserved heterogeneity and random parameter models 186

5.1.1. Behavioral model and "kernel" likelihood function 187

5.1.2. Mixing probability distribution 191

5.1.3. Likelihood functions and "statistical calibration" of individual parameters 192

5.2 RP acreage choice models: theoretical interest and empirical issues 195

5.3. The RP multi-crop model 202

5.3.1. Acreage share choice models 202

5.3.2. Yield supply and variable input demand models 205

5.3.3. Expected crop returns 207

5.4. The F/RP multi-crop model 211

5.5. Estimation of the RP model and calibration of the Sim-RP model 212

5.5.1. Parametric specification of the RP multi-crop model and exogeneity assumptions 212

5.5.2. Sketch of the estimation procedure 214

5.5.3. The Sim-RP model and its "statistical" calibration 225

5.6. Empirical application: impact of the UE support to biofuels on acreage choices 228

5.6.1. Data 228
5.6.2. Modeling details 231
5.6.3. Statistical performances of the RP model 231
5.7. Estimated distribution of the random parameter vector of the RP model 237
5.7.1. Estimation of the impact of EU biofuel support: simulations with the Sim-RP model 241
5.7.2. Comparison of the simulations obtained from the RP and F/RP models 243
5.7.3. Lagged acreage shares as additional control variables 248
Concluding remarks 249

Chapitre 6. Dealing with corner solutions in multi-crop micro-econometric models: an endogenous regime switching approach with regime fixed costs. 253
Introduction 254
6.1. Endogenous regime switching acreage choices with regime fixed costs 257
6.2. An ERS micro-econometric multi-crop model with regime fixed costs 264
6.2.1. Yield supply and variable input demand models 265
6.2.2. Acreage share choice models 267
6.2.3. Production regime choice model 270
6.3. Parametric specification and estimation procedure 273
6.3.1. Distributional assumptions 273
6.3.2. Estimation 276
6.3.3. Calibration of the simulation model 281
6.4. Empirical application 282
6.4.1. Data and empirical specification of the model 282
6.4.2. Estimation results 288
6.4.3. Simulation results 294
Concluding remarks 299

Conclusion Générale 301

Bibliographie 310

Introduction Générale

Le secteur agricole fait face à de nombreux enjeux. Comprendre les choix des agriculteurs est dès lors important pour déterminer les meilleurs solutions privées et publiques pour les résoudre.

Les modèles micro-économétriques de choix de production agricoles proposés dans la littérature ont longtemps ignoré ou contourné de manière plus ou moins satisfaisante certaines questions importantes d'un point de vue empirique, telles que l'hétérogénéité inobservée des agriculteurs et de leurs conditions de production et les problèmes posés par la présence de solutions en coin dans les données (Carpentier *et al*., 2015; Gohin *et al*., 2015; Letort, 2009). Cette thèse s'intéresse à ces questions, notamment pour l'analyse de leurs choix d'assolement, c'est-à-dire les choix d'allocation des parts de surface aux cultures. Ces questions sont traitées, en cohérence avec la théorie micro-économique des choix de production agricoles, dans le cadre d'un modèle de choix de production multiculture. Nous mobilisons pour cela des techniques de modélisation et d'estimation développées dans d'autres champs de l'économie et de la statistique computationnelle ces deux dernières décennies. L'utilisation de ces techniques, rendue possible grâce au développement d'outils de calcul puissants et la disponibilité de bases de données de plus en plus riches, constitue un apport méthodologique en économie de la production agricole. Nous contribuons aussi plus généralement à la littérature économique en démontrant qu'il est possible d'estimer des modèles économétriques à paramètres aléatoires multivariés intégrant une dimension panel. La contribution de cette thèse est donc avant tout méthodologique.

Afin de mieux situer les enjeux de la thèse, une première partie de ce chapitre introductif est

consacrée à une revue de la littérature sur la prise en compte de l'hétérogénéité inobservée et des solutions en coin dans les modèles micro-économétriques utilisés en économie de la production agricole et dans d'autres champs de l'économie. Dans une seconde partie, nous resituons les travaux menés dans la thèse par rapport à cette littérature en présentant les contributions que constituent les approches proposées pour traiter ces questions. Une dernière partie est consacrée à la présentation de l'organisation du rapport de thèse.

Prise en compte de l'hétérogénéité inobservée et des solutions en coin dans les modèle micro-économétriques

La question de l'hétérogénéité inobservée est importante en économie de la production agricole. En effet, les agriculteurs et les exploitations agricoles sont hétérogènes. Cette hétérogénéité est présente même au sein de régions relativement petites et d'exploitations spécialisées dans les mêmes types de cultures. On peut citer, parmi les principales sources d'hétérogénéités entre exploitations, la qualité du sol, le capital d'exploitation, les aptitudes et aspirations des exploitants et leurs pratiques agricoles. Mais ces aspects, indispensables pour différencier les agriculteurs, sont mal renseignés dans les bases de donnée usuelles, et sont donc « inobservés » pour l'économètre. Dans les modèles de choix de production agricole, ces effets de l'hétérogénéité inobservée, lorsqu'ils sont pris en compte, ne sont introduits qu'au niveau de la constante du modèle par le biais de termes individuels additifs (Lacroix and Thomas, 2011). Or ces effets peuvent aussi affecter d'autres paramètres du modèle tels que les paramètres associés aux prix. Les agriculteurs peuvent en effet réagir de

façon différente à un même changement de prix. Ne pas en tenir compte pourrait conduire à des modèles biaisés, et donc à des résultats de simulation de politique publique biaisés. Ce point a été discuté par Keane (2009) qui montre que dans un modèle linéaire, les effets individuels additivement séparables ne sont pas suffisants pour tenir compte des effets de l'hétérogénéité inobservée car les effets des variables explicatives du modèle sur la variable à expliquer peuvent être corrélés avec les effets des caractéristiques inobservées des individus. Pour traiter ce problème, Keane (2009) préconise l'utilisation de modèles à paramètres individuels.

Malgré son intérêt, la prise en compte des effets de l'hétérogénéité inobservée est encore aujourd'hui largement ignorée dans les modèles micro-économétriques de choix de production agricole. Il existe en revanche des modèles micro-économétriques appliqués tenant compte de cette hétérogénéité dans d'autres champs de l'économie, tels que l'économie du travail, l'économie industrielle empirique ou le commerce international (Allenby *et al*., 1998; Keane, 2009; Mannering *et al*., 2016). Dans cette littérature, deux approches sont souvent utilisées pour prendre en compte l'hétérogénéité inobservée: des modèles à paramètres aléatoires et des modèles intégrant des indices latents. Très utilisés en micro-économétrie appliquée et en statistique computationnelle, les modèles à paramètres aléatoires (Hsiao et Pesaran, 2004; Train, 2007) permettent de ne pas seulement se limiter à la constante du modèle, comme seul paramètre affecté par l'hétérogénéité inobservée. Les paramètres associés aux variables d'intérêt du modèle, telles que les prix, sont supposés varier d'un individu à un autre. Les vecteurs de paramètres individuels du modèle sont alors supposés issus d'un même processus générateur de données, caractérisé par une distribution

dont on cherche à estimer les paramètres. En économétrie, ce type de modèle est généralement estimé à l'aide de méthodes de Maximum de Vraisemblance Simulée, alors que dans la littérature statistique c'est l'algorithme EM – pour Expectation-Maximization – (Dempster *et al*., 1977; McLachlan and Krishnan, 2008) ainsi que ces versions stochastiques (MCEM, SAEM) (Delyon *et al*., 1999; Wei and Tanner, 1990) qui sont le plus souvent utilisés pour estimer les modèles à paramètres aléatoires. Les modèles intégrant des indices latents (voir, *e.g*., Eaton *et al*, 2011), quant à eux, peuvent être considérés comme des modèles à paramètres aléatoires particuliers, dans lesquels une structure spécifique est imposée sur les paramètres aléatoires. L'avantage de cette approche est qu'elle permet de réduire le nombre de paramètres à estimer, elle peut toutefois s'avérer restrictive lorsque la spécification théorique des modèles ne se prête pas à l'imposition de structures particulières sur l'ensemble de ces paramètres. Ce type de modèle est fréquemment utilisé en micro-économétrie, rarement en statistique computationnelle, et est généralement estimé par la méthode des moments simulés (McFadden, 1989).

La littérature économique et la statistique computationnelle fournissent donc des outils permettant de tenir de compte de l'hétérogénéité inobservée dans les modèles économétriques. Il faut toutefois souligner que, quelque soit l'approche utilisée, modèles à paramètres aléatoires ou à indices latents, les modèles estimés dans la littérature ne se composent généralement que d'une seule équation. Or, les modèles micro-économétriques multiculture de choix de production auxquels nous nous intéressons ici présentent une structure plus complexe, notamment parce qu'il s'agit de systèmes d'équations dépendantes les unes des autres.

La seconde question à laquelle nous nous intéressons dans cette thèse concerne les solutions en coin, c'est-à-dire les valeurs nulles, observées dans les choix d'assolement. Il s'agit d'un aspect important de la modélisation des choix de production agricole. Ce problème est récurrent car les agriculteurs produisent rarement tous les mêmes cultures chaque année. Les données à l'échelle des exploitations contiennent donc des valeurs nulles correspondant aux parts de surface, intrants et rendements des cultures non produites. Tenir compte de ce problème permet notamment de comprendre pourquoi les agriculteurs produisent certaines cultures et d'autres pas, ce qui est crucial dans l'étude des politiques publiques visant la diversification des assolements, source de bénéfices économique et environnementale. Les systèmes de régression censurée sont souvent utilisés dans les modèles de choix d'assolement pour prendre en compte les solutions en coin (Fezzi and Bateman, 2011; Platoni *et al*., 2012). Or, cette approche permet de traiter les parts de surface nulle d'un point de vue statistique, mais n'est pas cohérente avec la théorie micro-économique de la production agricole, *e.g*., avec cette approche, le prix d'une culture non produite impacte l'assolement des cultures produites.

Wales et Woodland (1983) et Lee et Pitt (1986) ont été les premiers à donner une interprétation économique aux valeurs nulles dans un modèle micro-économétrique. L'approche de Wales et Woodland (1983) est basée sur la théorie primale. Ces auteurs utilisent les conditions de Kuhn-Tucker issues de la maximisation d'une fonction objective sous des contraintes de budget et de non-négativité. Lee et Pitt (1986) proposent une approche basée sur la théorie duale, équivalente à celle de Wales et Woodland (1983). Ils utilisent une approche de prix virtuels pour caractériser les régimes de biens (sous ensemble

de biens consommés). Les modèles micro-économétriques qui dérivent de ces cadres théoriques sont estimables par la méthode du Maximum de Vraisemblance dont l'estimateur possède de bonnes propriétés (Arndt *et al*., 1999). La mise en œuvre de ces estimations est toutefois compliquée, et ces modèles sont rarement estimés dans la littérature dans le cas de plusieurs biens/plusieurs cultures, en raison de leur complexité liée à l'absence de formes analytiques explicites des modèles et aux difficultés computationnelles qu'implique leur estimation. Ainsi, la prise en compte des solutions en coin dans les modèles économétriques se base le plus souvent, non pas sur les approches proposées par Wales et Woodland (1983) et Lee et Pitt (1986), mais sur des extensions de l'approche en deux étapes proposée par Heckman (1979) (Moore et Negri, 1992; Perali et Chavas, 2000; Yen, 2005; Yen et Lin, 2006). Ces approches conduisent à des modèles de forme réduite, estimables mais pas nécessairement cohérents du point de vue de la théorie économique (Arndt *et al*., 1999).

Approches proposées et contributions de la thèse

Comme nous l'avons vu, la prise en compte de l'hétérogénéité inobservée des exploitations et le traitement des solutions en coin dans les modèles de choix de production agricole soulèvent des problèmes de spécification et d'estimation complexes, ce qui explique pourquoi ces questions ont souvent été ignorées dans les applications empiriques en économie agricole. Pour traiter ces problèmes de façon cohérente, il est dans un premier temps nécessaire de considérer un cadre théorique de base adapté, permettant de dériver des modèles micro-économétriques flexibles et estimables. La modélisation micro-

économétrique des choix de production agricole fait l'objet d'une littérature abondante (Carpentier et Letort, 2011; Just et Pope, 2001, 2003). Nous nous appuyons ici sur les travaux de Carpentier and Letort (2014) qui ont proposé un modèle de choix de production agricole intégrant un modèle de choix d'assolement de forme Logit Multinomial (MNL). Ce modèle micro-économétrique multiculture permet de décrire les décisions de court terme des agriculteurs dans un cadre statique. En plus des choix d'assolement, les décisions d'utilisation d'intrants variables et de rendement objectif y sont modélisées. Ce modèle repose sur des formes fonctionnelles dont les paramètres ont des interprétations agronomiques aisément compréhensible pour les non-économistes et combine à la fois les concepts de la programmation mathématique positive (Heckelei *et al*., 2012; Howitt, 1995) et de la modélisation micro-économétrique (Just *et al*., 1983) avec utilisation de la terre comme un facteur fixe allouable. Pour dériver leur modèle, Carpentier et Letort (2014) supposent que chaque agriculteur maximise son profit total espéré sous les contraintes d'utilisation totale de la terre, de non négativité des surfaces allouées aux cultures et d'une fonction de production quadratique. Le profit total espéré est la somme des marges brutes espérées pondérées par les parts de surfaces de chaque culture, moins des coûts implicites liés à la gestion des assolements. La résolution de ce programme d'optimisation conduit à un système d'équations représentant, pour chaque culture, les choix optimaux de rendements, d'utilisation d'intrants, et de ce fait les marges brutes optimales, ainsi que les assolements optimaux. Dans leur modèle, Carpentier et Letort (2014) considèrent une fonction de coûts de gestion des assolements de forme entropique, de ce fait les équations de choix d'assolement dérivées du modèle ont une forme logit multinomiale. Ce modèle de choix de production agricole avec assolement MNL présente plusieurs avantages. D'une part,

les parts de surfaces allouées aux cultures sont par construction (strictement) comprises entre zéro et un, ce qui facilite la mise en œuvre de l'approche que nous proposons pour traiter du problème des solutions en coin. D'autre part, ce modèle est assez parcimonieux en paramètres, ce qui peut en faciliter l'estimation, mais il n'est pas complètement flexible. Pour accroitre la flexibilité du modèle, il est possible de considérer une structure d'emboitement des cultures dans la fonction de coûts de gestion des assolements. Les équations de choix d'assolement qui en dérivent sont alors de forme Logit Multinomial Emboîté ou Nested Multinomial Logit (NMNL). Cette structure d'emboitement permet d'introduire de l'asymétrie dans les substitutions entre les cultures, pour tenir compte de certaines contraintes de gestion de l'assolement, telle que la structure des rotations entre cultures. On construit donc un schéma d'assolement où les cultures sont regroupées en groupes, en sous-groupes et ainsi de suite selon le nombre de niveaux d'emboîtement retenu par le modélisateur, sachant qu'une augmentation du nombre de niveaux permet d'accroitre la flexibilité du modèle mais augmente également le nombre de paramètres à estimer.

Les travaux présentés dans cette thèse s'appuient donc sur le modèle multiculture avec choix d'assolement NMNL proposés par Carpentier et Letort (2014) et en proposent des extensions pour répondre aux objectifs fixés : la prise en compte l'hétérogénéité inobservée des agriculteurs et de leur conditions de production dans le modèle et le traitement des solutions en coin dans les choix d'assolement de façon cohérente du point de vue de la théorie économique. Cette thèse apporte trois contributions principales à la littérature sur la modélisation des choix de production agricole.

La prise en compte de l'hétérogénéité inobservée par la spécification des modèles micro-

économétriques de choix de production à paramètres aléatoires (Train, 2007) est la première contribution de cette thèse. Elle vise à exploiter les possibilités offertes par la spécification des modèles à paramètres aléatoires, notamment le fait que l'ensemble des paramètres du modèle, et non uniquement la constante, peuvent être affectés par les effets de l'hétérogénéité inobservée. A notre connaissance, cette approche n'est pas utilisée dans les modèles de choix de production agricoles, et quand elle est utilisée habituellement, ce n'est pas sur des modèles avec une structure aussi complexe que le nôtre. En effet, les modèles de choix de production agricole que nous considérons se composent de plusieurs équations liées entre elles du fait que les marges brutes des cultures intervenant dans les choix d'assolement sont déterminées par les décisions de rendements et de demande d'intrants. De plus, ces modèles reposent sur des formes fonctionnelles non linéaires ou intègrent des variables endogènes dans leur version log-linéaire, du fait que dans ce cas les parts de surface de chaque culture dépendent des parts de surface des autres cultures. Enfin, certaines variables explicatives du modèle sont potentiellement corrélées aux termes aléatoires. Les marges brutes anticipées des cultures, notamment, qui sont les déterminants principaux des choix d'assolement, peuvent être impactées par certaines caractéristiques non observées des exploitations ayant aussi un impact sur les comportements des agriculteurs et donc les paramètres aléatoires du modèle. Tous ces éléments peuvent significativement accroitre la complexité de l'estimation de notre modèle et nous ont amené, comme nous le verrons plus loin, à développer des procédures spécifiques.

Une deuxième contribution de ce travail de thèse est le traitement des solutions en coin dans

les choix d'assolement à partir d'un cadre théorique cohérent d'un point de vue micro-économique et estimable de façon relativement simple. Nous utilisons pour cela une approche à changement de régime endogène (Heckman, 1976; Lee, 1978; Maddala and Nelson, 1975), un régime de production étant défini ici comme le sous-ensemble de cultures effectivement produites par l'agriculteur. Ce modèle micro-économétrique à changement de régimes endogènes avec coûts fixes associés aux régimes repose sur un cadre théorique cohérent car il dérive d'un programme de maximisation du profit total espéré sous contrainte de non-négativité des parts de surface et de l'utilisation totale de la terre. Nous supposons qu'un coût fixe, permettant de tenir compte des coûts de commercialisation et/ou des contraintes liées à la production simultanée de différentes cultures, est associé à chaque régime. C'est, à notre connaissance, la première fois que ce type de coûts est pris en compte dans un modèle à changement de régime endogène. On suppose que les agriculteurs choisissent leur régime en comparant les profits associés à chaque régime potentiel, c'est-à-dire la somme des marges brutes associées à chaque culture produite dans le régime pondérées par les parts de surfaces moins les coûts de gestion de l'assolement, modulo les coûts fixes qui sont associés aux régimes. Les fonctions de coûts associés aux régimes ne dépendent pas des parts de surfaces et introduisent donc une discontinuité dans la fonction du profit total espéré. Au sein de chaque régime potentiel, les choix d'assolement, de rendement objectif et d'utilisation variable reposent sur le modèle de Carpentier et Letort (2014). Comme nous l'avons vu précédemment, la structure NMNL du modèle de choix d'assolement impose implicitement que, pour chaque culture appartenant au régime considéré, les parts de surfaces soient strictement positives, ce qui garantit l'appartenance d'un assolement au régime dans lequel il a été optimisé. L'optimisation d'un problème de choix d'assolement NMLN dans un

régime conduira donc toujours à des assolements optimaux appartenant à ce régime. De ce fait, le problème de choix d'assolement initial du producteur peut être décomposé en deux sous problèmes imbriqué l'un dans l'autre de façon à avoir une solution unique (un régime optimal unique et donc un assolement optimal unique): un problème de choix de régime ne dépendant que des profits des régimes et des coûts fixes associées aux régimes, et un problème de choix d'assolement NMNL dans chaque régime tenant compte implicitement des conditions de non-négativité des parts de surface, les parts de surfaces étant strictement comprises entre zéro et un pour les cultures produites dans le régime.

Notre modèle est donc principalement composé de trois parties : (a) un modèle de choix discrète décrivant le choix de régime, (b) pour chaque régime potentiel, un système d'équation de parts de surface d'écrivant le choix d'assolement pour les cultures produites dans ce régime, et (c) pour chaque régime un système d'équations de rendement et de demande d'intrant pour les cultures produites dans le régime. Là encore, le modèle à changement de régime endogène que nous considérons est donc plus complexe que ceux généralement estimés dans la littérature, dans la mesure où nous avons un sous-système d'équations pour chaque régime de production considéré. De plus, les choix de régime et d'assolement sont supposés simultanés et la probabilité de choix de régime dépend du choix d'assolement à travers les profits associés à chaque régime potentiel. Enfin, les caractéristiques inobservées des choix de production (niveaux de rendement et d'utilisation des intrants variables des cultures non produites) peuvent être corrélées avec la décision de choix de régime. L'un des avantages de l'approche à paramètres aléatoires que nous utilisons dans la thèse est qu'elle permet de tenir compte de cette endogénéité potentielle entre les

décisions de régime et d'assolement.

La troisième contribution de cette thèse est liée aux approches d'estimation proposées reposant sur des outils spécifiques d'estimation, économétriques et statistiques, encore non utilisés (ou peu utilisés) en économie de la production agricole. Les modèles que nous considérons sont paramétriques, ce qui présente plusieurs avantages pour répondre aux problématiques qui nous intéressent dans la thèse. En effet, comme nous l'avons évoqué précédemment, les modèles de choix de production que nous considérons présentent un certain nombre de caractéristiques, telles que des corrélations entre variables explicatives et paramètres aléatoires, la présence de variables endogènes dans l'équation de choix d'assolement, la non linéarité du modèle ou encore l'endogénéité entre choix de régime et d'assolement, qui impliquent que la distribution des paramètres aléatoires ne peut être estimer directement sur la base d'hypothèses simples et d'adaptations de modèles semi-paramétriques (voir, par exemple, Arellano and Bonhomme, 2011a, 2011b; Keane, 2009; Murtazashvili and Wooldridge, 2008, 2016). Nous montrons que ces problèmes peuvent être surmontés (a) en considérant la structure complète du modèle de choix de production agricole de Carpentier and Letort (2014), où les marges brutes espérées sont calculées à partir des équations de rendement et d'utilisation d'intrant variable, dépendant des prix, des paramètres aléatoires et des aléas, et (b) en adoptant une approche paramétrique, comme le préconise Keane (2009), spécifiant la distribution des paramètres aléatoires avec une matrice de variance-covariance complète et la distribution des termes d'erreurs du modèle. Cette approche permet de définir explicitement les structures de corrélation et d'endogénéité. De

plus, les hypothèses sur les distributions des variables aléatoires permettent d'utiliser des approches basées sur le Maximum de Vraisemblance. En plus de faciliter, d'un point de vue empirique, l'estimation des modèles à paramètres aléatoires, le fait d'utiliser une approche paramétrique et de poser des hypothèses sur la distribution des paramètres présente un avantage important lorsqu'on souhaite utiliser ces modèles à des fins de simulation. Il est en effet relativement aisé, une fois la distribution des paramètres aléatoires estimée, d'obtenir les valeurs des paramètres du modèle spécifique à chaque agriculteur de l'échantillon considéré, sur la base des choix observés de ces agriculteurs. Cette propriété des modèles paramétriques à paramètres aléatoires, qui est maintenant utilisée de façon routinière dans les analyses économétriques basées sur les modèles de choix discrets à paramètres aléatoires (Greene, 2004; Train, 2009), peut être exploitée pour « calibrer statistiquement» des modèles multiculture de choix de production propres à chaque agriculteur, afin de tenir compte de l'hétérogénéité de leurs comportements dans les simulations de politiques publiques conduites à l'aide de ces modèles.

Ensuite, les approches d'estimation que nous proposons sont basées sur l'algorithme EM et non sur le Maximum de Vraisemblance Simulé (MVS) souvent utilisé par les micro-économètres. L'approche du MVS est en effet difficilement applicable dans notre cas en raison de la dimension du modèle multiculture que nous considérons qui complexifie grandement l'écriture de sa fonction de vraisemblance et est susceptible de générer des problèmes de stabilité numérique lors de la maximisation de cette vraisemblance.

L'algorithme EM permet, quant à lui, de décomposer un problème de maximisation complexe, tel que celui auquel nous sommes confronté, en plusieurs sous problèmes moins

complexes que le problème original et c*et al*gorithme est stable numériquement (Lange, 1995). Combiner avec les modèles à paramètres aléatoires, l'algorithme EM permet de bien exploiter les spécificités de la micro-économétrie structurelle, notamment de traiter les informations non observées par l'économètre. Nous contribuons également donc à la littérature micro-économétrique appliquée en montrant que l'utilisation des méthodes de simulation de Monte Carlo combinées à l'algorithme EM (MCEM) (voir, par exemple, McLachlan et Krishnan 2008; Lavielle 2014) - qui sont habituellement utilisées par les statisticiens appliqués mais rarement considérés par les économètres appliqués (voir, par exemple, Jank and Booth (2003)) - peut considérablement atténuer les problèmes de calcul liés à l'estimation de larges modèles multivariés à paramètres aléatoires.

Enfin, une dernière contribution méthodologique concerne le traitement des problèmes d'estimation dans le modèle à changement de régime endogène que nous considérons pour traiter du problème des solutions en coin. Nous utilisons une approximation par la méthode de Laplace (Harding et Hausman, 2007) pour approximer les probabilités de choix de régime. Ceci permet de réduire le temps de calcul et de tenir compte du fait que la probabilité de choix d'un régime dépend des performances non observées des cultures non produites dans ce régime. L'algorithme EM est également mis à contribution pour estimer les paramètres des distributions des termes d'erreur du modèle (rendements, intrants et assolements) qui ne sont « observés » que pour les cultures produites et dont la matrice de variance-covariance dépend du régime considéré, ce qui complique son estimation directe. Ainsi, nous tenons compte de la dépendance entre les variables dépendantes du modèle, ce qui, à notre connaissance n'est jamais fait dans les modèles à changement de régime dans l'écriture de

leur vraisemblance.

Organisation du rapport de thèse

Le reste de ce rapport de thèse est organisé en six chapitres répartis en deux parties. La première partie se décompose en quatre chapitres et présente les apports méthodologiques de la thèse. La seconde partie se décompose en deux chapitres, correspondant à deux applications empiriques des méthodes présentées dans la première partie.

Le premier chapitre de la thèse présente le modèle micro-économétrique de choix de production multiculture avec choix assolement MNL et sa version à paramètres aléatoires. Nous y présentons les formes fonctionnelles sur lesquelles repose le modèle ainsi que les avantages de la spécification à paramètres aléatoires. Nous montrons également dans ce premier chapitre que l'estimation de ce modèle est difficilement envisageable dans un cadre non entièrement paramétrique, justifiant l'utilisation d'une approche paramétrique. Nous montrons enfin que les estimateurs du Maximum de Vraisemblance Simulée (MVS), fréquemment utilisés par les économètres, ne sont pas appropriés pour l'estimation de notre modèle.

Le deuxième chapitre présente l'approche que nous proposons pour l'estimation de notre modèle de choix de production multiculture à paramètres aléatoires. Nous commençons par

présenter l'algorithme EM correspondant à notre modèle. Nous montrons ensuite comment certaines extensions de cet algorithme, tel que l'algorithme *Expectation-Conditional-Maximization* (ECM) (Meng et Rubin, 1993), l'algorithme *Monte-Carlo-EM* (MCEM) (Weï et Tanner, 1990) et l'algorithme *Stochastic-Approximation-EM* (SAEM) permettent de faciliter la mise en œuvre concrète de nos calculs d'estimation.

Le troisième chapitre présente les fondements théoriques du modèle à changement de régime endogène que nous proposons pour tenir compte pour traiter du problème des solutions en coin dans les choix d'assolement. Nous mettons tout d'abord en évidence les limites des modélisations micro-économétriques proposées jusqu'à présent pour tenir compte de ces choix de cultures de manière cohérente par rapport à la théorie économique. Nous présentons ensuite l'extension du modèle de choix d'assolement MNL à paramètres aléatoires, à changement de régime endogène et avec coûts fixes de choix de régime proposé.

.

Dans le quatrième chapitre nous abordons les difficultés d'estimation du modèle micro-économétrique à changement de régime endogène avec coûts fixes associés à chaque régime. Nous présentons notamment les difficultés posées par le calcul des probabilités de choix de régime et l'estimation des paramètres de distribution des termes d'erreur des modèles de choix d'assolement, d'offre de rendement et de demande d'intrant.

Le cinquième chapitre est la version longue d'un article accepté pour publication (Koutchadé

et al, 2017a). Il s'agit d'une application empirique de l'approche présentée dans les Chapitres 1 et 2 pour tenir compte de l'hétérogénéité inobservée des exploitations agricoles et des agriculteurs. Considérant le modèle de choix de production agricoles avec assolement NMNL à 2 niveaux d'emboîtement, Nous estimons le système d'offre de rendements, de demande d'intrants variables, et de choix d'assolement sur un panel de producteurs français de grandes cultures dans La Meuse et mettons en évidence la présence d'une hétérogénéité significative des exploitations agricoles et des agriculteurs au sein de ce territoire relativement peu étendue. Nous montrons également qu'ignorer l'hétérogénéité des processus de choix des agriculteurs peut avoir d'importants impacts sur les résultats de simulations conduites à partir des modèles estimés, avec une application sur les effets de la hausse du prix du colza induite par la politique européenne sur les biocarburants. Nos résultats de simulation sur les effets du soutien de l'Union Européenne aux bio-carburants montrent notamment que l'utilisation des modèles usuels, *i.e.*, des modèles à paramètres fixes ou semi-fixes (dans lesquels seules les constantes sont aléatoires), peut conduire à sous-estimer ou surestimer les effets de la politique. Nous montrons également que le modèle à paramètres aléatoires leur est statistiquement préférable.

Le sixième chapitre repose sur un article à soumettre pour publication (Koutchadé et al, 2017b). Nous proposons dans ce chapitre, une application empirique de l'approche présentée dans les chapitres 3 et 4 thèse pour tenir compte des solutions en coin. Cette application considère le modèle de choix de production agricole à changements de régime endogènes avec coûts fixes de régime et choix d'assolement NMNL à 3 niveaux d'emboîtement. Nous

considérons sept cultures et les huit régimes de production les plus représentés dans notre échantillon de producteurs de grandes cultures dans la Marne. Pour chaque régime, les coûts fixes sont supposés égaux à la somme des coûts par cultures. Dans cette application, nous montrons principalement que les coûts fixes associés aux régimes jouent un rôle important dans le choix des agriculteurs de produire ou non une culture et que les marges brutes espérées ne sont pas les seuls déterminants de ces choix. Pour cela, nous simulons dans un premier temps le choix des régimes à partir des résultats d'estimation avec ou sans les coûts fixes associés aux régimes et comparons les résultats aux choix observés. Nous utilisons également nos résultats d'estimation pour calculer les élasticités prix des surfaces allouées aux cultures. La décomposition de ces élasticités, basée sur une généralisation de la méthode proposée par McDonald et Moffitt (1980), montre que, pour les cultures représentant une petite part de l'assolement, une grande partie des effets du prix sur les surfaces est en fait du au choix de produire ou non la culture. Enfin, une simulation de l'impact d'un changement du prix du pois sur les assolements nous permet de mettre en évidence des effets de seuil, dus à des changements de régime de production, dans les impacts sur les surfaces.

Partie I

Hétérogénéité inobservée et choix de cultures dans les systèmes de grandes cultures : modélisation et estimation

Introduction

L'objectif de cette thèse est d'explorer deux voies d'amélioration possible de la modélisation micro-économétrique des choix de producteurs de grandes cultures. Il s'agit de de prendre en compte de l'hétérogénéité inobservée des agriculteurs et de leurs conditions de production d'une part, et des décisions d'assoler ou non certaines cultures. Ces deux objectifs soulèvent des questions de recherche intéressantes, tant dans le domaine de la modélisation micro-économique des choix de production agricole que dans le domaine de la statistique appliquée.

Cette première partie décrit les aspects méthodologiques des travaux de micro-économétrie appliquée présentés dans cette thèse. Elle présente des revues de littérature, en économie agricole et en statistique appliquée, relatives aux enjeux scientifiques des questions recherches explorées. Elle présente également les modèles et les outils d'inférence statistique choisis pour répondre aux mieux aux objectifs poursuivis dans cette thèse. Les modèles et approches d'estimation présentés dans cette partie sont, évidemment, ceux mobilisés dans les applications empiriques regroupées dans la seconde partie.

Nous portons une attention particulière aux aspects liés à l'estimation des modèles proposés dans cette thèse. En effet, les modèles micro-économétriques de choix de production multiculture sont des modèles statistiques multivariés de grande dimension structurés de manière relativement complexe par la théorie économique. Une grande partie des travaux de recherche dont les résultats sont présentés ici a consisté à adapter les outils d'inférence statistique proposés en statistique computationnelle à l'estimation de ces modèles. Il s'agit là d'une des principales originalités de cette thèse.

Cette première partie est composée de deux séries de deux chapitres. Les chapitres 1 et 2 traitent des questions relatives à la prise en compte de l'hétérogénéité inobservée des agriculteurs et des conditions de production. Le chapitre 1 présente le modèle de choix de production multiculture choisi pour servir de base à l'ensemble des travaux menés dans cette thèse et en propose une extension capable de rendre compte d'une hétérogénéité inobservée substantielle entre exploitations. Le chapitre 2 présente les outils d'inférence statistique employés dans l'ensemble de la thèse, avec un focus particulier sur l'estimation des modèles spécifiés dans le chapitre 1.

Les chapitres 3 et 4 abordent les questions liées à la modélisation des décisions des agriculteurs d'assoler ou non certaines cultures. Le chapitre 3 se concentre sur la modélisation économique de ces décisions. Il présente également une version « avec choix de cultures » du modèle micro-économétrique proposé dans le chapitre 1. Le chapitre 4 présente comment les outils d'inférence statistique présentés dans le chapitre 2 peuvent être adaptés pour estimer le modèle spécifié dans le chapitre 3.

Chapitre 1

Modèle de choix de production multiculture avec hétérogénéité inobservée

1.1. Modèles de choix de production multiculture

Nous présentons dans cette section les modèles de choix de production multiculture proposés par Carpentier and Letort (2014). Ces modèles sont à la base des travaux empiriques réalisés dans le cadre de cette thèse. Ils définissent des modèles micro-économétriques des choix de production des agriculteurs à partir d'un système d'équations composé d'un sous-système de fonctions de choix (de parts) d'assolements et d'un sous-système de fonctions d'offre de rendement et de demande d'intrants variables.

Ces modèles ont été choisis pour trois raisons principales. Ils sont parcimonieux en paramètres et aisément interprétables. Les modèles de choix d'assolement proposés par Carpentier et Letort (2014), les modèles de choix d'assolement « Logit MultiNomiaux » (*MultiNomial Logit*, MNL), ont des propriétés spécifiques que nous exploitons dans notre travail sur la modélisation économétrique des solutions en coin dans les modèles de choix d'assolement (voir Chapitre 3).

L'originalité de ces modèles est triple.

(*a*) L'approche de modélisation proposée par Carpentier et Letort (2014) repose à la fois sur les concepts de la programmation mathématique positive (Howitt, 1995 ; Heckeleï et Wolff, 2003 ; Heckeleï et al, 2012) et des modèles économétriques multiculture considérant la terre comme un facteur fixe allouable (Chambers et Just, 1989 ; Moore et Negri, 1992 ; Guyomard et al, 1996 ; Lacroix et Thomas, 2011 ; Platoni et al, 2012 ; Bayramoglu et Chakir, 2016) et le modèle multiculture qui en dérive est cohérent d'un point de vue micro-économétrique.

(*b*) Carpentier et Letort (2014) proposent une famille de modèles de choix d'assolement originaux, les modèles de choix d'assolement « Logit MultiNomiaux » (*MultiNomial Logit*, MNL). Le nom de ces modèles tien à leurs formes fonctionnelles. En effet, les modèles de Carpentier et Letort (2014) décrivent les parts d'assolement à partir de modèles dont la forme fonctionnelle générale est celle des probabilités de choix des modèles de choix discret Logit Multinomiaux (voir, *e.g.*, Train, 2009). Ceci relie ces modèles de choix d'assolement aux modèles de choix de culture définis à l'échelle d'une parcelle qui sont de plus en plus employés (voir, *e.g.*, Wu et al, 2004 ; Langpap et Wu, 2011 ; Claassen et al, 2017). Ces modèles de choix discret permettent d'exploiter les données d'utilisation du sol qui sont maintenant disponibles sur de grandes aires géographiques et sur longue période. Les modèles MNL les plus simples sont les modèles MNL « standards » (*Standard MultiNomial Logit*, SMNL) alors que les plus élaborés, et les plus flexibles, sont les modèles MNL « avec emboîtements » (*Nested MultiNomial Logit*, NMNL). Ces modèles ont la particularité d'expliciter le rôle des marges des cultures dans les choix d'assolement, ce qu'on retrouve dans tous les modèles de programmation mathématique mais dans peu de modèles économétriques multiculture, les modèles dérivés des travaux de Chavas et Holt (1990) faisant exception (voir, *e.g.*, Chavas et Holt, 1996 ; Holt, 1999).

(*c*) Les modèles de choix de production multiculture proposés par Carpentier et Letort (2014) sont cohérents dans leurs parties déterministes et aléatoires tout en considérant que les agriculteurs décident de leurs assolements avant leurs utilisations d'intrants variables tels que les pesticides ou les engrais.

Nous commençons par présenter les fondements théoriques et la structure générale de ces

modèles. Puis nous présentons leurs sous-systèmes de fonctions de rendement et de demande d'intrants variables. Nous présentons ensuite en détail le sous-système des modèles de choix d'assolement MNL et ses principales propriétés. Enfin, nous décrivons les modèles micro-économétriques de choix de production multiculture combinant les éléments présentés précédemment, et donnons leur vraisemblance.

1.1.1. Fondements théoriques et structure générale

Considérons $\mathcal{K} \equiv \{1, \ldots, K\}$, l'ensemble des cultures produites par l'agriculteur *i* l'année *t*. L'assolement en *t* de cet agriculteur est décrit par les parts de sa surface totale, $s_{k,it}$, dédiées à chaque culture $k \in \mathcal{K}$ avec $\mathbf{s}_{it} = (s_{k,it} : k \in \mathcal{K})$. Nous supposons que l'agriculteur utilise l'ensemble de sa surface de terre arable, *i.e.* que $\sum_{k \in \mathcal{K}} s_{k,it} = 1$. Pour une culture *k*, l'agriculteur obtient un rendement $y_{k,it}$ en utilisant une quantité $x_{k,it}$, agrégée ici par souci de simplicité, d'intrants variables (engrais, semences et pesticides dans nos applications empiriques).[1]

Le choix d'assolement $\mathbf{s}_{it}$ est décidé par l'agriculteur avant ses utilisations d'intrants variables $\mathbf{x}_{it} = (x_{k,it} : k \in \mathcal{K})$. De fait, les utilisations d'intrants variables des agriculteurs

[1] Voir Carpentier et Letort (2014) pour le cas à plusieurs intrants.

s'adaptent aux conditions climatiques et sanitaires de la campagne en cours, voire à l'évolution des prix des cultures. Les prix des intrants variables, $\mathbf{v}_{it} = (v_{k,it} : k \in \mathcal{K})$, sont connus au moment des décisions d'utilisation d'intrants. Nous supposerons qu'ils sont également connus au moment des décisions d'assolement. En revanche, les prix de vente $\tilde{\mathbf{p}}_{it} = (\tilde{p}_{k,it} : k \in \mathcal{K})$ et les rendements $\mathbf{y}_{it} = (y_{k,it} : k \in \mathcal{K})$ des cultures ne sont connus qu'à la fin du processus de production, après que l'agriculteur ait pris toutes ses décisions. Nous supposerons dans la suite que l'agriculteur fonde ses choix d'assolement et ses décisions sur les mêmes anticipations des prix des cultures, $\mathbf{p}_{it} = (p_{k,it} : k \in \mathcal{K})$.

Dans la suite, l'ensemble d'information disponible au moment du choix d'assolement sera noté ω_{it}^{s}. Il est supposé contenir $\mathbf{p}_{it}$ et $\mathbf{v}_{it}$. L'ensemble d'information disponible au moment des utilisations d'intrants variables de la culture *k* d'assolement sera noté $\omega_{k,it}^{x}$. Il contient ω_{it}^{s}, et donc $p_{k,it}$ et $v_{k,it}$.

La spécification du modèle de Carpentier and Letort (2014) suppose que l'agriculteur est neutre face au risque et que la technologie de production de chaque culture ne dépend pas de la surface sur laquelle elle est produite, *i.e.* que cette technologie est à rendement de la terre constant. Ces hypothèses ont deux implications principales.

D'une part, l'agriculteur décide de ses quantités d'intrants variables culture par culture, sans avoir à tenir compte de son choix d'assolement, et de sorte à maximiser l'espérance de la marge brute de chaque culture. La technologie de production de la culture *k* est décrite sous la forme de la fonction de production (aléatoire) :

(1) $y_{k,it} = f_{k,it}(x_{k,it})$,

qui ne dépend pas de l'assolement $\mathbf{s}_{it}$. La fonction de production $f_{k,it}(x_k)$ est supposée non croissante et concave en x_k (en tout point de son support). La marge brute de la culture *k* obtenue par l'agriculteur est donnée par $\tilde{p}_{k,it} y_{k,it} - v_{k,it} x_{k,it}$. Elle n'est connue qu'à fin du processus de production. L'agriculteur choisit ses quantités d'intrants variables $x_{k,it}$ sur la base de ses anticipations du prix et du rendement de la culture *k* et compte tenu des conditions climatiques et sanitaires prévalant au moment du choix. Il est à noter que l'anticipation du rendement d'une culture dépend de l'anticipation des conditions climatiques et sanitaires affectant ce rendement à partir du choix d'intrants variables jusqu'à la récolte. L'espérance de marge brute de la culture *k* anticipée par l'agriculteur *i* l'année *t* est donnée par $E[\tilde{p}_{k,it} f_{k,it}(x_k) - v_{k,it} x_k \mid \omega^x_{k,it}, x_k]$ s'il envisage d'utiliser l'intrant variable agrégé en quantité x_k. Nous supposons ici que les rendements et les prix des cultures sont indépendants, ce qui est justifié par le fait que les marchés des cultures considérées ici (les principales grandes cultures) sont des marchés mondiaux. Aussi, l'utilisation d'intrant variable de l'agriculteur *i* l'année *t* est modélisée comme la solution, supposée unique, d'un problème de maximisation de la marge brute espérée, avec :

(2) $x_{k,it} = \arg\max_{x_k \geq 0} \left\{ p_{k,it} E[f_{k,it}(x_k) \mid x_k, \omega^x_{k,it}] - v_{k,it} x_k \right\}$

pour la culture *k*.

D'autre part, l'agriculteur choisit son assolement à partir de son anticipation de la marge brute de chaque culture $\boldsymbol{\pi}_{it} = (\pi_{k,it} : k \in \mathcal{K})$ et, le cas échéant, des aides qu'il peut percevoir par unité de surface de culture. Ces aides sont notés $a_{k,it}$ avec $\mathbf{a}_{it} = (a_{k,it} : k \in \mathcal{K})$. Sous les

hypothèses précédentes, la marge brute de la culture *k* telle qu'elle est anticipée au moment du choix d'assolement de l'agriculteur *i* pour l'année *t* est donnée par :

(3) $$\pi_{k,it} = E[p_{k,it} y_{k,it} - v_{k,it} x_{k,it} \mid \omega_{it}^{s}] + a_{k,it} .$$

Les termes $y_{k,it}$ et $x_{k,it}$ sont aléatoires du point de vue de l'agriculteur au moment du choix d'assolement. De fait, ces termes font le lien entre les choix réalisés pour chaque culture, les utilisations d'intrants variables (et les rendements objectifs associés), et pour l'ensemble des cultures, le choix d'assolement.

Si l'agriculteur n'était pas contraint dans son choix d'assolement, il choisirait de ne cultiver que la culture la plus rentable. Or l'agriculteur doit gérer le planning de ses utilisations de machine (tracteurs, pulvérisateurs, épandeurs, charrues, ...) et de sa main d'œuvre. De même, il doit également considérer les cycles biologiques de ses cultures de sorte à avoir récolté la culture précédente sur la parcelle où il envisage de semer la culture envisagée lors de la campagne en cours. Ces questions de planning et de compatibilité des cycles biologiques des cultures imposent des contraintes sur les choix des cultures et des surfaces qui leur sont dédiées qui incitent les agriculteurs à diversifier leurs assolements.

La fonction de « coût implicite de gestion de l'assolement » de Carpentier et Letort (2014), noté ici $C_{it}(\mathbf{s})$ pour un assolement **s**, résume les effets de ces contraintes. Elle contient également des coûts de production des cultures non pris en compte dans les calculs des marges brutes ainsi que les effets des conditions de culture au moment des choix d'assolement. Puisqu'elle représente essentiellement les effets de motifs de diversification des assolements, cette fonction de coût implicite est supposée convexe en **s**. Les coûts

représentés par cette fonction sont essentiellement liés à la main d'œuvre et à l'utilisation des machines. Ces intrants pouvant être loués dans le cadre d'acquisition de « travaux par tiers », cette fonction est supposée dépendre du prix de ces travaux (dans nos applications empiriques l'ensemble des prix sont déflatés par l'indice des prix des travaux par tiers).

Sous les hypothèses précédentes, l'agriculteur décide son assolement $\mathbf{s}$ en arbitrant entre d'une part le revenu qu'il retire de cet assolement, $\sum_{k\in\mathcal{K}} s_k \pi_{k,it}$, et d'autre part le coût implicite de gestion de cet assolement, $C_{it}(\mathbf{s})$. La différence entre ces deux termes définit la fonction de profit en $\mathbf{s}$ de l'agriculteur (définie ici par unité de surface) :

(4) $$\Pi_{it}(\mathbf{s};\boldsymbol{\pi}_{it}) = \sum_{k\in\mathcal{K}} s_k \pi_{k,it} - C_{it}(\mathbf{s})$$

Lorsque la fonction $C_{it}(\mathbf{s})$ est strictement convexe, le choix d'assolement de l'agriculteur *i* en *t* est modélisé comme la solution unique d'un problème de maximisation de profit, avec :

(5) $$\mathbf{s}_{it} = \arg\max_{\mathbf{s}\geq\mathbf{0}} \left\{ \sum_{k\in\mathcal{K}} s_k \pi_{k,it} - C_{it}(\mathbf{s}) \text{ s.c. } \sum_{k\in\mathcal{K}} s_k = 1 \right\}.$$

Les choix de formes fonctionnelles des fonctions de rendements et de la fonction de coût implicite de gestion de l'assolement définissent la forme du modèle de choix de production multiculture. Nous choisissons ici celle de Carpentier et Letort (2014).

1.1.2. Fonctions d'offre de rendement et de demande d'intrants variables

La fonction de rendement quadratique « normalisée » :

(6) $$f_{k,it}(x_k) = \beta^y_{k,it} - (1/2)(\alpha^x_k)^{-1}(\beta^x_{k,it} - x_k)^2 \text{ avec } \alpha^x_k > 0$$

conduit à des fonctions d'offre de rendement et de demandes d'intrants de formes simples et aisément interprétables. Le paramètre α_k^x définit la courbure de la fonction $f_{k,it}(x_k)$. Il est supposé strictement positif pour assurer la stricte concavité de cette fonction de rendement. Sous cette hypothèse, le terme $\beta_{k,it}^y$ s'interprète comme le niveau de rendement potentiel de l'agriculteur i l'année t pour la culture k, alors que le terme $\beta_{k,it}^x$ s'interprète comme le niveau d'intrant variable agrégé nécessaire pour atteindre le niveau de rendement potentiel $\beta_{k,it}^y$. Les termes $\beta_{k,it}^y$ et $\beta_{k,it}^x$ sont supposés aléatoires dans modèles économétriques considérés dans cette thèse.

Nous supposons que l'agriculteur observe $\beta_{k,it}^x$ au moment du choix de $x_{k,it}$. Ceci implique que la solution du problème de la maximisation en x_k de la marge brute espérée $p_{k,it}E[f_{k,it}(x_k) \mid x_k, \omega_{k,it}^x] - v_{k,it}x_k$ est donné par :

(7) $$x_{k,it}^o(v_{k,it}, p_{k,it}) = x_{k,it} = \beta_{k,it}^x - \alpha_k^x q_{k,it} \quad \text{où} \quad q_{k,it} \equiv v_{k,it}(p_{k,it})^{-1}.$$

Cette équation définit la fonction de demande d'intrant variable pour la culture *k* de l'agriculteur *i* l'année *t*. Substituer cette fonction de demande d'intrant dans la fonction de rendement $f_{k,it}(x_k)$ donne la fonction d'offre de rendement pour la culture *k* de l'agriculteur *i* l'année *t* :

(8) $$y_{k,it}^o(v_{k,it}, p_{k,it}) = y_{k,it} = \beta_{k,it}^y - \alpha_k^x q_{k,it}^2 / 2 \ .$$

La forme des fonctions d'offre de rendement $y_{k,it}^o(v_{k,it}, p_{k,it})$ et de demande d'intrants $x_{k,it}^o(v_{k,it}, p_{k,it})$ montrent que le paramètre de courbure détermine l'élasticité de ces fonctions en fonction des prix des cultures, $p_{k,it}$, et des intrants, $v_{k,it}$. Comme cela sera vu plus bas, la

linéarité de ces fonctions en les termes aléatoires $\beta^x_{k,it}$ et $\beta^y_{k,it}$ s'avère pratique pour la spécification des modèles micro-économétriques de fonctions d'offre de rendement et de demande d'intrants. Cette propriété de linéarité provient de la normalisation spécifique de la fonction de rendement quadratique $f_{k,it}(x_k)$, cette dernière étant par ailleurs à l'origine des interprétations simples, à partir de concepts employés par les agronomes, des termes $\beta^x_{k,it}$ et $\beta^y_{k,it}$.

Ces fonctions d'offre de rendement et de demande d'intrants permettent de calculer les marges brutes obtenues par les agriculteurs et leur anticipation de ces marges brutes au moment des choix d'assolement :

(9) $$\pi^o_{k,it}(v_{k,it}, p_{k,it}, a_{k,it}) = \pi_{k,it} = p_{k,it}E[\beta^y_{k,it} \mid \omega^s_{it}] - v_{k,it}E[\beta^x_{k,it} \mid \omega^s_{it}] + \alpha^x_k p_{k,it} q^2_{k,it} / 2 + a_{k,it}.$$

Ces marges brutes anticipées sont des déterminants essentiels des choix d'assolement des agriculteurs.

1.1.3. Modèles de choix d'assolement MNL

Le modèle de choix d'assolement SMNL est obtenu à partir de la spécification de la fonction implicite de coût de gestion des assolements suivante :

(10) $$C_{it}(\mathbf{s}) = D_{it} + \sum_{k\in\mathcal{K}} s_k \beta^s_{k,it} + (\alpha^s)^{-1} \times \sum_{k\in\mathcal{K}} s_k \ln s_k \quad \text{avec } \alpha^s > 0.$$

Le terme D_{it} définit un coût fixe spécifique à l'ensemble de culture $\mathcal{K}$. Ce terme n'est pas identifiable lorsque les agriculteurs considérés produisent tous les mêmes cultures mais il

joue un rôle important lorsque les agriculteurs ont des ensembles de cultures différents, *i.e.* lorsqu'on travaille sur des modèles avec solutions en coin dans les choix d'assolement. Le terme $\beta_{k,it}^{s}$ s'interprète comme un coût par unité de surface de la culture *k*. Il capte les coûts de production de la culture *k* non pris en compte dans de le calcul de la marge brute. Il capte également les effets des conditions de culture, telles que l'état des sols ou la présence d'adventices, qui peuvent impacter les choix d'assolement des agriculteurs. Les termes $\beta_{k,it}^{s}$ sont considérés comme aléatoires du point de vue de l'économètre mais supposés connus des agriculteurs au moment de leurs choix d'assolement.

La partie entropique, $\sum_{k\in\mathcal{K}} s_k \ln s_k$, de la fonction de coût $C_{it}(\mathbf{s})$ porte l'effet des motifs de diversification de l'assolement. Ce terme est strictement convexe en **s** dès lors que le paramètre α^s est strictement positif. Le paramètre α^s joue ici un rôle de pondération entre la partie « marges et coûts des cultures », $\sum_{k\in\mathcal{K}} s_k(\pi_{k,it} - \beta_{k,it}^{s})$, et la partie « motif de diversification », $\sum_{k\in\mathcal{K}} s_k \ln s_k$, de la fonction de profit $\Pi_{it}(\mathbf{s};\boldsymbol{\pi}_{it})$. Plus α^s est faible, plus l'agriculteur a tendance à choisir un assolement selon lequel la surface de terre arable est distribuée de manière uniforme aux différentes cultures produites.

Carpentier et Letort (2014) proposent une normalisation alternative de la fonction de coût SMNL considérée ici. De fait, il est aisé de réécrire la fonction $C_{it}(\mathbf{s})$ sous la forme

(11) $$C_{it}(\mathbf{s}) = B_{it} + (\alpha^s)^{-1} \times \sum_{k\in\mathcal{K}} s_k(\ln s_k - \ln r_{k,it})$$

avec :

(12) $$\sum_{k\in\mathcal{K}} r_{k,it} = 1 \text{ et } r_{k,it} \in (0,1) \text{ pour } k \in \mathcal{K}.$$

Dans ce cas, le terme $\mathbf{r}_{it} = (r_{k,it} : k \in \mathcal{K})$ peut être interprété comme un assolement de référence, celui qui minimise les coûts implicites de gestion de l'assolement de l'agriculteur considéré, *i.e.* l'assolement pour lequel cet agriculteur minimise les pics de travail et d'utilisation des machines tout tenant compte des coûts d'implantation des cultures de l'année considérée. Cette formulation de la fonction de coût $C_{it}(\mathbf{s})$ n'est pas retenue ici malgré son interprétation intuitive. La formulation présentée précédemment s'avère être plus « pratique » pour la spécification de modèles micro-économétriques.

Maximiser en $\mathbf{s}$ la fonction de profit $\Pi_{it}(\mathbf{s};\boldsymbol{\pi}_{it})$ permet de modéliser l'assolement choisi par l'agriculteur *i* en *t* à partir d'un modèles de choix (de parts) d'assolement SMNL :

(13) $$s_{k,it} = s^{o}_{k,it}(\boldsymbol{\pi}_{it}) = \frac{\exp\left(\alpha^{s} \times (\pi_{k,it} - \beta^{s}_{k,it})\right)}{\sum_{\ell\in\mathcal{K}} \exp\left(\alpha^{s} \times (\pi_{\ell,it} - \beta^{s}_{\ell,it})\right)} \text{ pour } k \in \mathcal{K}.$$

Ce modèle garantit des parts de surface strictement comprises en 0 et 1. Plus le paramètre α^{s} est petit plus les $s_{k,it}$ s'approchent de 1/*K*. Plus ce paramètre est grand plus $s_{k,it}$ s'approche de 1 si la culture *k* est la plus rentable (*i.e.*, celle qui procure la marge $\pi_{k,it} - \beta^{s}_{k,it}$ maximale), plus $s_{k,it}$ tend vers 0 sinon. Il est à noter que les termes $\beta^{s}_{k,it}$ ne sont identifiables qu'à une constante additive près, justifiant la contrainte de normalisation $\beta^{s}_{K,it} = 0$ retenue dans la suite.

Comme cela sera développé plus bas, ce modèle a trois avantages importants pour la

modélisation des assolements nuls. Tout d'abord, il fournit une forme explicite de la fonction de profit indirecte. Calculer la fonction de profit $\Pi_{it}(\mathbf{s};\boldsymbol{\pi}_{it})$ en $s_k = s^o_{k,it}(\boldsymbol{\pi}_{it})$ pour $k \in \mathcal{K}$ permet d'obtenir le profit de l'agriculteur en fonction des marges brutes espérées $\boldsymbol{\pi}_{it}$, des aides aux cultures $\mathbf{a}_{it}$ et des paramètres de la fonction de coût $C_{it}(\mathbf{s})$, *i.e.* la fonction de profit indirecte :

(14) $$\Pi^o_{it}(\boldsymbol{\pi}_{it}) = (\alpha^s)^{-1} \ln \sum_{\ell \in \mathcal{K}} \exp\left(\alpha^s \times (\pi_{\ell,it} - \beta^s_{\ell,it})\right) - D_{it}.$$

Par ailleurs, on a $\frac{\partial}{\partial \pi_\ell} \Pi^o_{it}(\boldsymbol{\pi}_{it}) = s^o_{k,it}(\boldsymbol{\pi}_{it})$ conformément au lemme de Hotelling. Ensuite, si la part de surface de la culture *k* est toujours strictement positive, le terme $s^o_{k,it}(\boldsymbol{\pi}_{it})$ est très proche de 0 lorsque la marge de la culture *k*, $\pi_{k,it} - \beta^s_{k,it}$, est suffisamment petite par rapport aux marges des autres cultures. Enfin, les fonctions $s^o_{k,it}(\boldsymbol{\pi}_{it})$ sont continûment différentiables à « volonté » en $\boldsymbol{\pi}_{it}$.

Les modèles de choix d'assolement SMNL décrits ci-dessus ont cependant un défaut majeur. Ils contraignent significativement les substitutions entre les cultures dans l'assolement. Par exemple, les effets des marges espérées des cultures $\pi_{k,it}$ sur les choix d'assolement sont essentiellement gouvernés par le seul paramètre α^s. Plus ce paramètre est petit, plus ces effets sont petits.

De fait, toutes les cultures sont considérées de manière symétrique dans la fonction de coût implicite SMNL. Or, les cultures peuvent être plus ou moins concurrentes pour les utilisations de facteurs fixes dans la construction de l'assolement de l'agriculteur, selon leurs

cycles biologiques et leurs places dans le schéma de rotation suivi par l'agriculteur.

Carpentier and Letort (2014) ont défini la fonction de coût implicite NMNL pour pallier à ces défauts des modèles de choix d'assolement SMNL standards. Les cultures sont regroupées selon une structure hiérarchique (mutuellement exclusive) – en groupes, en sous-groupes et ainsi de suite selon le nombre de niveaux retenu par le modélisateur – en fonction de leur « degré » de concurrence. Par exemple, dans une structure d'emboîtement à deux niveaux – cultures et groupes de cultures – les groupes de cultures regroupent les cultures qui sont les plus concurrentes entre elles dans la construction de l'assolement. On peut par exemple, regrouper les têtes d'assolement, les cultures d'hiver, *etc*.

Supposons que les K cultures de l'ensemble $\mathcal{K}$ sont regroupées, de manière mutuellement exclusive, en G groupes définissant l'ensemble de groupes $\mathcal{G} = \{1,...,G\}$. Le groupe g est défini par le sous-ensemble $\mathcal{K}(g)$ de $\mathcal{K}$ et contient $K(g)$ cultures. Pour rendre compte de cette structure à deux niveaux, Carpentier and Letort (2014) ont proposé de modéliser les choix d'assolement des agriculteurs à partir de la fonction de coût NMNL suivante :

(11)
$$C_{it}(\mathbf{s}) = D_{it} + \sum_{k\in\mathcal{K}} s_k \beta^s_{k,it} + (\alpha^s)^{-1} \sum_{g\in\mathcal{G}} s_{(g)} \ln s_{(g)} + \sum_{g\in\mathcal{G}} s_{(g)} (\alpha^s_{(g)})^{-1} \sum_{k\in\mathcal{K}(g)} s_{k|(g)} \ln s_{k|(g)}$$

où le terme $s_{(g)}$ représente la somme des parts de surface des cultures du groupe g :

(15)
$$s_{(g)} = \sum_{\ell\in\mathcal{K}(g)} s_\ell$$

et $s_{k|(g)}$ la part de surface de la culture k dans les parts de surface des cultures du groupe g :

(16) $$s_{k|(g)} = s_k / s_{(g)}$$

Les termes α^s et $\alpha^s_{(g)}$ déterminent la flexibilité du choix d'assolement. Les conditions

(17) $$0 < \alpha^s \leq \alpha^s_{(g)} \text{ pour } g \in \mathcal{G}$$

et :

(18) $$\alpha^s_{(g)} = \alpha^s \text{ si le groupe } g \text{ est un singleton}$$

assurent la convexité en **s** de la fonction $C_{it}(\mathbf{s})$. La forme fonctionnelle de cette fonction reflète la structure d'emboîtement hiérarchique des cultures : les coûts de gestion de l'assolement **s** dépendent d'abord de la manière dont la surface de terre arables est allouée aux groupes de cultures, puis dont les surfaces dédiées aux groupes de cultures sont allouées aux cultures de ces groupes.

Maximiser en **s** la fonction de profit $\Pi_{it}(\mathbf{s};\boldsymbol{\pi}_{it})$ construite avec la fonction de coût NMNL $C_{it}(\mathbf{s})$ permet de modéliser l'assolement choisi par l'agriculteur *i* en *t* à partir d'un modèle de choix (de parts) d'assolement MNL « avec emboîtements » :

(19) $$s^o_{k,it}(\boldsymbol{\pi}_{it}) = s_{k,it} = \frac{\exp\left(\alpha^s_{(g)}(\pi_{k,it} - \beta^s_{k,it})\right)}{\sum_{\ell\in\mathcal{K}(g)} \exp\left(\alpha^s_{(g)}(\pi_{\ell,it} - \beta^s_{\ell,it})\right)} \frac{\left(\sum_{\ell\in\mathcal{K}(g)} \exp\left(\alpha^s_{(g)}(\pi_{\ell,it} - \beta^s_{\ell,it})\right)\right)^{\alpha^s(\alpha^s_{(g)})^{-1}}}{\sum_{h\in\mathcal{G}}\left(\sum_{\ell\in\mathcal{K}(h)} \exp\left(\alpha^s_{(h)}(\pi_{\ell,it} - \beta^s_{\ell,it})\right)\right)^{\alpha^s(\alpha^s_{(h)})^{-1}}}$$

pour $k \in \mathcal{K}(g)$ et $g \in \mathcal{G}$. Le terme :

(20) $$s^o_{k|(g),it}(\boldsymbol{\pi}_{it}) = s_{k|(g),it} = \frac{\exp\left(\alpha^s_{(g)}(\pi_{k,it} - \beta^s_{k,it})\right)}{\sum_{\ell\in\mathcal{K}(g)} \exp\left(\alpha^s_{(g)}(\pi_{\ell,it} - \beta^s_{\ell,it})\right)}$$

donne le modèle de la part de surface de la culture k dans la surface de son groupe g alors que le terme :

$$s^{o}_{(g),it}(\boldsymbol{\pi}_{it}) = s_{(g),it} = \frac{\left(\sum_{\ell\in\mathcal{K}(g)} \exp\left(\alpha^{s}_{(g)}(\pi_{\ell,it} - \beta^{s}_{\ell,it})\right)\right)^{\alpha^{s}(\alpha^{s}_{(g)})^{-1}}}{\sum_{h\in\mathcal{G}}\left(\sum_{\ell\in\mathcal{K}(h)} \exp\left(\alpha^{s}_{(h)}(\pi_{\ell,it} - \beta^{s}_{\ell,it})\right)\right)^{\alpha^{s}(\alpha^{s}_{(h)})^{-1}}} \tag{21}$$

donne la part de surface du groupe g. Les modèles de choix d'assolement NMNL sont bien plus flexibles que leur contrepartie SMNL. La structure du terme $s^{o}_{k,it}(\boldsymbol{\pi}_{it})$, avec $s^{o}_{k,it}(\boldsymbol{\pi}_{it}) = s^{o}_{k|(g),it}(\boldsymbol{\pi}_{it})s^{o}_{(g),it}(\boldsymbol{\pi}_{it})$, reflète la structure d'emboîtement hiérarchique des cultures définissant la fonction NMNL $C_{it}(\mathbf{s})$. Les substitutions entre cultures dépendent des substitutions entre groupes de cultures, et, le cas échéant, entre cultures d'un même groupe. Elles sont gouvernées par le vecteur de paramètres $\boldsymbol{\alpha}^{s}$ qui regroupe le paramètre α^{s} et les paramètres $\alpha^{s}_{(g)}$ pour $g \in \mathcal{G}$. Le paramètre α^{s} gouverne les substitutions entre groupes, le paramètre $\alpha^{s}_{(g)}$ gouverne les substitutions au sein du groupe g.

De fait, les fonctions de choix d'assolement NMNL conservent les propriétés de leurs contreparties SMNL tout en étant plus flexibles quant aux substitutions entre cultures et le modèle SMNL est le cas particulier du modèle NMNL dans lequel $\alpha^{s}_{(g)} = \alpha^{s}$ pour $g \in \mathcal{G}$. Les seules différences substantielles tiennent aux formes fonctionnelles des fonctions de parts d'assolement et des fonctions de profit indirecte associées. La fonction de profit indirecte associée à la fonction de coût implicite NMNL décrite ci-dessus est donnée par :

$$\Pi^{o}_{it}(\boldsymbol{\pi}_{it}) = (\alpha^{s})^{-1} \ln \sum_{h\in\mathcal{G}}\left(\sum_{\ell\in\mathcal{K}(h)} \exp\left(\alpha^{s}_{(h)}(\pi_{\ell,it} - \beta^{s}_{\ell,it})\right)\right)^{\alpha^{s}(\alpha^{s}_{(h)})^{-1}} - D_{it}\,. \tag{22}$$

La généralisation des résultats précédents aux cas de structures d'emboîtements à plus de deux niveaux est aisée, quoique relativement pénible des points de vue calculatoire et notationnel. Le Chapitre 3 repose sur des modèles de choix d'assolement NMNL à trois niveaux.

1.1.4. Modèles micro-économétriques de choix de production multiculture : formes fonctionnelles

Les modèles micro-économétriques de choix de production multiculture considérés ici sont définis comme des systèmes d'équations décomposés en deux sous-systèmes. Le premier regroupe les équations de rendement $y_{k,it}$ et d'utilisations d'intrant $x_{k,it}$ pour chaque culture dérivées des modèles d'offre de rendement $y^o_{k,it}(v_{k,it}, p_{k,it})$ et de demande d'intrant $x^o_{k,it}(v_{k,it}, p_{k,it})$ présentés ci-dessus :

(23)
$$\begin{cases} x_{k,it} = \beta^x_{k,it} - \alpha^x_k q_{k,it} \\ y_{k,it} = \beta^y_{k,it} - \alpha^x_k q^2_{k,it} / 2 \, . \\ k \in \mathcal{K} \end{cases}$$

Le second est constitué des équations décrivant les parts d'assolement des cultures $s_{k,it}$ dérivées des modèles de choix d'assolement $s^o_{k,it}(\boldsymbol{\pi}_{it})$ décrits ci-dessus. Nous ne considérons ici que le cas NMNL à deux niveaux qui donne des modèles des parts d'assolement de la forme :

$$(24) \qquad \begin{cases} s_{k,it} = \dfrac{\exp\left(\alpha^s_{(g)}(\pi_{k,it} - \beta^s_{k,it})\right)\left(\sum_{\ell \in \mathcal{K}(g)} \exp\left(\alpha^s_{(g)}(\pi_{\ell,it} - \beta^s_{\ell,it})\right)\right)^{\alpha^s(\alpha^s_{(g)})^{-1}-1}}{\sum_{h \in \mathcal{G}}\left(\sum_{\ell \in \mathcal{K}(h)} \exp\left(\alpha^s_{(h)}(\pi_{\ell,it} - \beta^s_{\ell,it})\right)\right)^{\alpha^s(\alpha^s_{(h)})^{-1}}} \\ k \in \mathcal{K}_{-K} \end{cases},$$

où $\mathcal{K}_{-K} = \mathcal{K} \setminus \{K\}$. Le cas SMNL se déduit simplement du cas NMNL en imposant les contraintes $\alpha^s_{(g)} = \alpha^s$ pour $g \in \mathcal{G}$. Ces deux sous-systèmes d'équations sont liées par les équations de marges brutes espérées des cultures $\pi_{k,it}$ dérivées des modèles $\pi^o_{k,it}(v_{k,it}, p_{k,it}, a_{k,it})$ avec :

$$(25) \qquad \begin{cases} \pi_{k,it} = p_{k,it} E[\beta^y_{k,it} \mid \omega^s_{it}] - v_{k,it} E[\beta^x_{k,it} \mid \omega^s_{it}] + \alpha^x_k q^2_{k,it} p_{k,it} / 2 + a_{k,it} \\ k \in \mathcal{K} \end{cases}$$

Dans ces systèmes d'équations les choix des agriculteurs $(\mathbf{y}_{it}, \mathbf{x}_{it}, \mathbf{s}_{it})$ sont exprimés comme des fonctions des prix des cultures et des intrants et des aides aux cultures $(\mathbf{p}_{it}, \mathbf{v}_{it}, \mathbf{a}_{it})$. Il reste ici à définir plus précisément les termes de rendements potentiels $\boldsymbol{\beta}^y_{it} \equiv (\beta^y_{k,it} : k \in \mathcal{K})$, de niveaux d'intrants nécessaires pour atteindre ces rendements potentiels $\boldsymbol{\beta}^x_{it} \equiv (\beta^x_{k,it} : k \in \mathcal{K})$ et des paramètres de coût par culture des fonctions de coût implicite de gestion de l'assolement $\boldsymbol{\beta}^s_{it} \equiv (\beta^s_{k,it} : k \in \mathcal{K}_{-K})$.

L'ensemble de ces termes captent les effets de conditions de culture. Certaines de ces conditions sont observables – telles que les conditions climatiques alors que d'autres ne le sont pas – telles de l'exploitation ou l'état sanitaire des cultures. Ces termes dépendent également en partie de caractéristiques de l'exploitation et de l'exploitant. Là encore,

certaines de ces caractéristiques sont observées – telles que la taille de l'exploitation ou son niveau de capital matériel – alors que d'autres ne le sont pas – telles que les aptitudes de l'agriculteur ou que la qualité des sols de l'exploitation. Dans les modèles considérés par la suite, ces termes sont modélisés sous la forme suivante:

$$(26) \qquad \begin{cases} \beta_{k,it}^{y} = \beta_{k}^{y} + (\boldsymbol{\delta}_{k}^{y})'\mathbf{c}_{k,it}^{y} + \varepsilon_{k,it}^{y} \\ \beta_{k,it}^{x} = \beta_{k}^{x} + (\boldsymbol{\delta}_{k}^{x})'\mathbf{c}_{k,it}^{x} + \varepsilon_{k,it}^{x} \\ \qquad k \in \mathcal{K} \end{cases}, \quad \begin{cases} \beta_{k,it}^{s} = \beta_{k}^{s} + (\boldsymbol{\delta}_{k}^{s})'\mathbf{c}_{k,it}^{s} - \varepsilon_{k,it}^{s} \\ \qquad k \in \mathcal{K}_{-K} \end{cases} \text{ et } \beta_{K,it}^{s} = 0$$

avec :

$$(27) \qquad \begin{cases} E[\varepsilon_{k,it}^{y} \mid \omega_{it}^{s}] = 0 \\ E[\varepsilon_{k,it}^{x} \mid \omega_{it}^{s}] = 0 \\ \qquad k \in \mathcal{K} \end{cases} \text{ et } \begin{cases} E[\varepsilon_{k,it}^{s}] = 0 \\ k \in \mathcal{K}_{-K} \end{cases}.$$

Dans ces modèles, les termes $\boldsymbol{\beta}^{s} \equiv (\beta_{k}^{s} : k \in \mathcal{K}_{-K})$, $\boldsymbol{\delta}^{s} \equiv (\boldsymbol{\delta}_{k}^{s} : k \in \mathcal{K}_{-K})$, $\boldsymbol{\beta}^{yx} \equiv ((\beta_{k}^{y}, \beta_{k}^{x}) : k \in \mathcal{K})$ et $\boldsymbol{\delta}^{yx} \equiv ((\delta_{k}^{y}, \delta_{k}^{x}) : k \in \mathcal{K})$ sont des vecteurs de paramètres à estimer. Les termes $\mathbf{c}_{k,it}^{w}$ pour $w \in \{y, x, s\}$ sont des vecteurs de variables de contrôle intégrant des variables climatiques et les caractéristiques observées de l'exploitation (surface arable totale et capital matériel dans nos applications empiriques). Les variables climatiques contenues dans les vecteurs $\mathbf{c}_{k,it}^{x}$, respectivement $\mathbf{c}_{k,it}^{s}$, sont définies en fonction de l'information mobilisable par l'agriculteur lors de ses choix d'intrants variables, respectivement d'assolement. Les termes $\boldsymbol{\varepsilon}_{it}^{s} \equiv (\boldsymbol{\varepsilon}_{k,it}^{s} : k \in \mathcal{K}_{-K})$, respectivement $\boldsymbol{\varepsilon}_{it}^{w} \equiv (\boldsymbol{\varepsilon}_{k,it}^{w} : k \in \mathcal{K})$ pour $w \in \{y, x\}$, sont des vecteurs de termes d'erreur captant les effets sur les termes $\boldsymbol{\beta}_{it}^{s}$, respectivement $\boldsymbol{\beta}_{it}^{w}$ pour $w \in \{y, x\}$, de leurs déterminants non observés. Les conditions

$E[\boldsymbol{\varepsilon}_{it}^{y} \mid \omega_{it}^{s}] = E[\boldsymbol{\varepsilon}_{it}^{y} \mid \omega_{it}^{s}] = \mathbf{0}$ et $E[\boldsymbol{\varepsilon}_{k,it}^{s}] = \mathbf{0}$ sont les contraintes de normalisation usuelles des termes d'erreur des modèles linéaires avec terme constant.

Les variables climatiques contenues dans les vecteurs $\mathbf{c}_{k,it}^{y}$ sont celles susceptibles d'expliquer les niveaux de rendement observés $y_{k,it}$. Le calcul des marges brutes espérées $\pi_{k,it}$ doit tenir compte de ce que l'agriculteur n'a pas connaissance de l'ensemble des variables climatiques contenues dans $\mathbf{c}_{k,it}^{y}$ et $\mathbf{c}_{k,it}^{x}$. Compte tenu des modèles de $\beta_{k,it}^{y}$ et $\beta_{k,it}^{x}$, les anticipations des agriculteurs de ces termes au moment de leur choix d'assolement sont données par :

(28) $$E[\beta_{k,it}^{y} \mid \omega_{it}^{s}] = \beta_{k}^{y} + (\boldsymbol{\delta}_{k}^{y})' E[\mathbf{c}_{k,it}^{y} \mid \omega_{it}^{s}] \text{ et } E[\beta_{k,it}^{x} \mid \omega_{it}^{s}] = \beta_{k}^{x} + (\boldsymbol{\delta}_{k}^{y})' E[\mathbf{c}_{k,it}^{x} \mid \omega_{it}^{s}].$$

Dans nos applications empiriques, les variables climatiques des termes $E[\mathbf{c}_{k,it}^{y} \mid \omega_{it}^{s}]$ et $E[\mathbf{c}_{k,it}^{x} \mid \omega_{it}^{s}]$ sont approchées, lorsque nécessaire, par les moyennes des variables correspondantes sur les cinq années précédant *t*, avec :

(29) $$\mathbf{c}_{k,it}^{ys} \simeq E[\mathbf{c}_{k,it}^{y} \mid \omega_{it}^{s}] \text{ et } \mathbf{c}_{k,it}^{xs} \simeq E[\mathbf{c}_{k,it}^{x} \mid \omega_{it}^{s}].$$

Nous sommes maintenant en mesure de présenter la forme fonctionnelle des modèles économétriques multiculture qui seront considérés dans la suite. Le modèle avec choix d'assolement NMNL à deux niveaux est donné par :

$$
(30)\quad \begin{cases}
\begin{cases} y_{k,it} = \beta_k^y + (\boldsymbol{\delta}_k^y)'\mathbf{c}_{k,it}^y + \varepsilon_{k,it}^y - \alpha_k^x q_{k,it}^2 / 2 \\ x_{k,it} = \beta_k^x + (\boldsymbol{\delta}_k^x)'\mathbf{c}_{k,it}^x + \varepsilon_{k,it}^x - \alpha_k^x q_{k,it} \\ k \in \mathcal{K} \end{cases} \\
\begin{cases} s_{k,it} = \dfrac{\exp\left(\alpha_{(g)}^s(\pi_{k,it} - \beta_{k,it}^s)\right)\left(\sum_{\ell \in \mathcal{K}(g)} \exp\left(\alpha_{(g)}^s(\pi_{\ell,it} - \beta_{\ell,it}^s)\right)\right)^{\alpha^s(\alpha_{(g)}^s)^{-1}-1}}{\sum_{h \in \mathcal{G}}\left(\sum_{\ell \in \mathcal{K}(h)} \exp\left(\alpha_{(h)}^s(\pi_{\ell,it} - \beta_{\ell,it}^s)\right)\right)^{\alpha^s(\alpha_{(h)}^s)^{-1}}} \\ \beta_{k,it}^s = \beta_k^s + (\boldsymbol{\delta}_k^s)'\mathbf{c}_{k,it}^s - \varepsilon_{k,it}^s \\ k \in \mathcal{K}_{-K} \\ \beta_{K,it}^s = 0 \end{cases} \\
\begin{cases} \pi_{k,it} = p_{k,it}\left(\beta_k^y + (\boldsymbol{\delta}_k^y)'\mathbf{c}_{k,it}^{ys}\right) - v_{k,it}\left(\beta_k^x + (\boldsymbol{\delta}_k^x)'\mathbf{c}_{k,it}^{xs}\right) + \alpha_k^x q_{k,it}^2 p_{k,it} / 2 + a_{k,it} \\ k \in \mathcal{K} \end{cases}
\end{cases}
$$

Les paramètres de ces modèles sont définis par (a) les vecteurs des paramètres constants $\boldsymbol{\beta} \equiv (\boldsymbol{\beta}^{yx}, \boldsymbol{\beta}^s)$ et des coefficients des variables de contrôle $\boldsymbol{\delta} \equiv (\boldsymbol{\delta}^{yx}, \boldsymbol{\delta}^s)$ des rendements potentiels, des utilisations d'intrants requises pour atteindre ces rendements potentiels et des coûts implicites des assolements, (b) les paramètres de flexibilité des choix d'intrants collectés dans le vecteur $\boldsymbol{\alpha}^x \equiv (\alpha_k^x : k \in \mathcal{K})$ et (c) les paramètres de flexibilité des choix d'assolement collectés dans le vecteur $\boldsymbol{\alpha}^s \equiv \left(\alpha^s, (\alpha_{(g)}^s : g \in \mathcal{G})\right)$ dans le cas des modèles NMNL. Les paramètres de flexibilité des choix de production sont regroupés dans le vecteur $\boldsymbol{\alpha}$, avec $\boldsymbol{\alpha} \equiv (\boldsymbol{\alpha}^x, \boldsymbol{\alpha}^s)$.

Il est à noter que les différentes équations des systèmes d'équations considérés ici partagent de nombreux paramètres. D'une part, les équations de rendements et d'utilisations d'intrants partagent les paramètres de flexibilité des choix d'intrants $\boldsymbol{\alpha}^x$. D'autre part, la

définition des marges brutes espérées $\boldsymbol{\pi}_{it}$ implique que les équations d'assolement dépendent de l'ensemble des paramètres – *i.e.* $\boldsymbol{\beta}^{yx}$, $\boldsymbol{\delta}^{yx}$ et $\boldsymbol{\alpha}^{x}$ – des équations de de rendements et d'utilisations d'intrants.

Les termes d'erreurs de ces modèles sont les termes aléatoires des équations de rendements et d'utilisations d'intrants $\boldsymbol{\varepsilon}_{it}^{yx} \equiv (\boldsymbol{\varepsilon}_{k,it}^{y}, \boldsymbol{\varepsilon}_{it}^{x})$, et ceux des équations de surface $\boldsymbol{\varepsilon}_{it}^{s}$. Ils sont collectés dans le vecteur $\boldsymbol{\varepsilon}_{it} \equiv (\boldsymbol{\varepsilon}_{k,it}^{yx}, \boldsymbol{\varepsilon}_{it}^{s})$.

Les variables explicatives de ces modèles sont (a) les aides aux cultures $\mathbf{a}_{it}$, (*b*) les prix des cultures $\mathbf{p}_{it}$ et des intrants $\mathbf{v}_{it}$, et (*c*) les variables de contrôle employées dans les différentes équations des systèmes considérés – *i.e.* les termes $\mathbf{c}_{k,it}^{y}$, $\mathbf{c}_{k,it}^{x}$, $\mathbf{c}_{k,it}^{s}$, $\mathbf{c}_{k,it}^{ys}$ et $\mathbf{c}_{k,it}^{xs}$ – et collectées dans le vecteur $\mathbf{c}_{it}$.

Carpentier et Letort (2014) emploient les modèles micro-économétriques de choix de production multiculture présentés ci-dessus en considérant que les paramètres $\boldsymbol{\beta}$, $\boldsymbol{\alpha}^{x}$ et $\boldsymbol{\alpha}^{s}$ de ces modèles sont fixes, *i.e.* communs à l'ensemble des agriculteurs de l'échantillon considérés. Ceci leur permet d'envisager l'estimation de ces paramètres dans un cadre d'inférence semi-paramétrique – *i.e.* sans choix *a priori* de la distribution des termes d'erreur $\boldsymbol{\varepsilon}_{it}$ de ces modèles – à partir d'estimateurs de la Méthode des Moments Généralisée (MMG).

Dans ce contexte, la log-linéarisation des modèles de choix de part d'assolement s'avère particulièrement utile car elle permet de définir des systèmes d'équations standards, *i.e.* à termes d'erreur additivement séparables, avec :

$$(31) \qquad \left\{ \begin{array}{c} \ln s_{k,it} - \ln s_{K,it} = \alpha^s \times \left\{ \begin{array}{l} (\pi_{k,it} - \pi_{K,it}) - (\beta_k^s + (\boldsymbol{\delta}_k^s)' \mathbf{c}_{k,it}^s) \\ + \left((\alpha^s)^{-1} - (\alpha_{(g)}^s)^{-1} \right) \times \ln s_{k|(g),it} \\ - \left((\alpha^s)^{-1} - (\alpha_{(G)}^s)^{-1} \right) \times \ln s_{K|(G),it} \end{array} \right\} + \alpha^s \times \varepsilon_{k,it}^s \\ k \in \mathcal{K}_{-K} \end{array} \right.$$

dans le cas NMNL (Berry, 1994). Ces techniques de log-linéarisation sont également employées ici, mais moins pour linéariser les modèles de choix d'assolement que pour obtenir les termes d'erreur $\varepsilon_{k,it}^s$ en fonction des assolement observés, de leurs déterminants observés et des paramètres des modèles considérés.

Les conditions de moment estimantes considérées par Carpentier et Letort sont des conditions d'orthogonalité construites en croisant les élément du vecteur de terme d'erreurs $\boldsymbol{\varepsilon}_{it}$ et des instruments construits comme des fonctions des variables explicatives – les variables d'aide $\mathbf{a}_{it}$, de prix $\mathbf{p}_{it}$ et $\mathbf{v}_{it}$, et les variables de contrôle $\mathbf{c}_{it}$ – des modèles de choix de production considérés.

La validité de ces conditions de moment repose sur l'hypothèse selon laquelle ces variables, collectées dans le vecteur $\mathbf{z}_{it} \equiv (\mathbf{a}_{it}, \mathbf{p}_{it}, \mathbf{v}_{it}, \mathbf{c}_{it})$, sont exogènes par rapport au vecteur de termes d'erreur des systèmes d'équation considérés avec :

$$(32) \qquad E[\boldsymbol{\varepsilon}_{it} \mid \mathbf{z}_{it}] = \mathbf{0}.$$

Cette hypothèse d'exogénéité est usuelle. Elle est justifiée par l'exogénéité par rapport aux phénomènes modélisés des processus générateurs des aides aux cultures $\mathbf{a}_{it}$, des prix de marché pour $\mathbf{p}_{it}$ et $\mathbf{v}_{it}$ et des évènements climatiques pour les variables climatiques de $\mathbf{c}_{it}$.

Les caractéristiques des exploitations contenues dans $\mathbf{c}_{it}$ sont largement prédéterminées par rapport aux choix modélisés ici, les utilisations d'intrants et les choix d'assolement résultant de décisions de court terme des agriculteurs. De fait, la construction des variables de prix en tant qu'indices de prix construits avec une base 1 pour la moyenne des prix constatés en $t = 1$ implique que les éléments de qualité des productions sont captés par les niveaux de rendements tels qu'ils sont construits pour les estimations. Une hypothèse d'exogénéité similaire sera employée dans la suite, mais dans une version plus forte.

1.1.5. Modèles micro-économétriques de choix de production multiculture : spécifications paramétriques et fonctions de vraisemblance

Les utilisations des formes fonctionnelles des modèles micro-économétriques de choix de production multiculture présentés ci-dessus nécessitent de considérer des versions « entièrement paramétriques » de ces modèles, *i.e.* de choisir une forme paramétrique de la distribution de $\boldsymbol{\varepsilon}_{it}$ conditionnelle en $\mathbf{z}_{it}$. En effet, nous allons employer ces formes fonctionnelles pour spécifier (*a*) des modèles à paramètres aléatoires et (*b*) des modèles à paramètres aléatoires et à changement de régime endogène dont l'estimation est difficile à envisager dans un cadre d'estimation semi-paramétrique, et ce d'autant plus qu'on considère ici l'estimation de systèmes d'équations de relativement grande dimension.

Nous supposons ici que (*a*) les termes d'erreur $\boldsymbol{\varepsilon}_{it}$ sont indépendants et identiquement distribués selon une loi normale multivariée et (*b*) que les vecteurs de variables explicatives

$\mathbf{z}_{it}$ sont strictement exogènes par rapport à ces termes d'erreurs. Formellement, nous supposons que :

(33) $$\boldsymbol{\varepsilon}_{it} \mid \mathbf{z}_{it} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Psi}_0) \text{ ou } \begin{bmatrix} \boldsymbol{\varepsilon}_{it}^{yx} \\ \boldsymbol{\varepsilon}_{it}^{s} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \boldsymbol{\Psi}_0^{yx} & \mathbf{C}_0 \\ \mathbf{C}_0' & \boldsymbol{\Psi}_0^{s} \end{bmatrix} \right)$$

où $\boldsymbol{\varepsilon}_{it}^{yx} \equiv (\boldsymbol{\varepsilon}_{it}^{y}, \boldsymbol{\varepsilon}_{it}^{x})$. L'hypothèse de normalité est usuelle. L'hypothèse d'indépendance entre les termes $\boldsymbol{\varepsilon}_{it}$ et $\mathbf{z}_{it}$ est plus restrictive que l'hypothèse d'exogénéité invoquée par Carpentier et Letort (2014). Elle implique également l'homoscédasticité conditionnelle en $\mathbf{z}_{it}$ des termes d'erreur $\boldsymbol{\varepsilon}_{it}$. Cette dernière peut s'avérer discutable pour ce qui concerne les variables de contrôle – variables climatiques ou caractéristiques des exploitations – contenues dans $\mathbf{z}_{it}$. En particulier, les conditions climatiques impactent l'état sanitaire des cultures qui en retour impacte la variabilité des rendements et utilisations d'intrants.

Cette hypothèse de normalité permet d'écrire les vraisemblances en les paramètres $(\boldsymbol{\beta}, \boldsymbol{\alpha}, \boldsymbol{\delta}, \boldsymbol{\Psi})$ des observations des choix de production modélisés $(\mathbf{y}_{it}, \mathbf{x}_{it}, \mathbf{s}_{it})$ conditionnelles en $\mathbf{z}_{it}$ à partir de fonctions de densité de lois normales multivariées. Ces vraisemblances sont définies à partir des fonctions de résidu :

(34) $$\varepsilon_{k,it}^{y}(\boldsymbol{\beta}^{y}, \boldsymbol{\alpha}^{x}, \boldsymbol{\delta}^{y}) \equiv y_{k,it} - \beta_k^{y} - (\boldsymbol{\delta}_k^{y})'\mathbf{c}_{k,it}^{y} + \alpha_k^{x} q_{k,it}^{2} / 2$$

pour les modèles de rendement,

(35) $$\varepsilon_{k,it}^{x}(\boldsymbol{\beta}^{y}, \boldsymbol{\alpha}^{x}, \boldsymbol{\delta}^{y}) \equiv x_{k,it} - \beta_k^{x} - (\boldsymbol{\delta}_k^{x})'\mathbf{c}_{k,it}^{x} + \alpha_k^{x} q_{k,it}$$

pour les modèles d'utilisations d'intrants et :

$$(36) \qquad \varepsilon_{k,it}^{s}(\boldsymbol{\beta},\boldsymbol{\alpha},\boldsymbol{\delta}) = (\alpha^{s})^{-1} \times (\ln s_{k,it} - \ln s_{K,it}) - \left\{ \begin{array}{c} \pi_{k,it}(\boldsymbol{\beta}^{yx},\boldsymbol{\alpha}^{x},\boldsymbol{\delta}^{yx}) - \pi_{K,it}(\boldsymbol{\beta}^{yx},\boldsymbol{\alpha}^{x},\boldsymbol{\delta}^{yx}) \\ -(\beta_{k}^{s} + (\boldsymbol{\delta}_{k}^{s})'\mathbf{c}_{k,it}^{s}) \\ +\left((\alpha^{s})^{-1} - (\alpha_{(g)}^{s})^{-1}\right) \times \ln s_{k|(g),it} \\ -\left((\alpha^{s})^{-1} - (\alpha_{(G)}^{s})^{-1}\right) \times \ln s_{K|(G),it} \end{array} \right\}$$

pour les modèles d'assolement NMNL. Les fonctions de marge brute espérée $\pi_{k,it}(\boldsymbol{\beta}^{yx},\boldsymbol{\alpha}^{x},\boldsymbol{\delta}^{yx})$ sont données par :

$$(37) \qquad \pi_{k,it}(\boldsymbol{\beta}^{yx},\boldsymbol{\alpha}^{x},\boldsymbol{\delta}^{yx}) \equiv p_{k,it}\left(\beta_{k}^{y} + (\boldsymbol{\delta}_{k}^{y})'\mathbf{c}_{k,it}^{ys}\right) - v_{k,it}\left(\beta_{k}^{x} + (\boldsymbol{\delta}_{k}^{x})'\mathbf{c}_{k,it}^{xs}\right) + \alpha_{k}^{x} q_{k,it}^{2} p_{k,it} / 2 + a_{k,it}$$

pour $k \in \mathcal{K}$.

La vraisemblance en $(\boldsymbol{\beta}^{yx},\boldsymbol{\alpha}^{x},\boldsymbol{\delta}^{yx},\boldsymbol{\Psi}^{yx})$ de l'observation des rendements et utilisations d'intrant $(\mathbf{y}_{it},\mathbf{x}_{it})$ conditionnelle en $\mathbf{z}_{it}$, notée $\ell_{it}^{ys}(\boldsymbol{\beta}^{yx},\boldsymbol{\alpha}^{x},\boldsymbol{\delta}^{yx},\boldsymbol{\Psi}^{yx})$, est de forme standard dans la mesure où $(\mathbf{y}_{it},\mathbf{x}_{it}) \mid \mathbf{z}_{it}$ suit une loi normale multivariée dont la matrice de variance-covariance est constante. En notant $\varphi(\mathbf{u};\mathbf{B})$ la fonction de densité de probabilité de la loi normale $\mathcal{N}(\mathbf{0},\mathbf{B})$ en $\mathbf{u}$, elle est donnée par :

$$(38) \qquad \ell_{it}^{ys}(\boldsymbol{\beta}^{yx},\boldsymbol{\alpha}^{x},\boldsymbol{\delta}^{yx},\boldsymbol{\Psi}^{yx}) = \varphi\left(\boldsymbol{\varepsilon}_{it}^{yx}(\boldsymbol{\beta}^{yx},\boldsymbol{\alpha}^{x},\boldsymbol{\delta}^{yx});\boldsymbol{\Psi}^{yx}\right)$$

où :

$$(39) \qquad \boldsymbol{\varepsilon}_{it}^{yx}(\boldsymbol{\beta}^{yx},\boldsymbol{\alpha}^{x},\boldsymbol{\delta}^{yx}) \equiv \left(\left(\varepsilon_{k,it}^{y}(\boldsymbol{\beta}^{y},\boldsymbol{\alpha}^{x},\boldsymbol{\delta}^{y}), \varepsilon_{k,it}^{x}(\boldsymbol{\beta}^{y},\boldsymbol{\alpha}^{x},\boldsymbol{\delta}^{y})\right) : k \in \mathcal{K}\right).$$

Cette vraisemblance est déduite de notre modèle paramétrique de la fonction de densité de $(\mathbf{y}_{it},\mathbf{x}_{it})$ conditionnelle en $\mathbf{z}_{it}$, fonction de densité qui est ici notée $f(\mathbf{y}_{it},\mathbf{x}_{it} \mid \mathbf{z}_{it})$.

La vraisemblance $\ell_{it}^{ys}(\boldsymbol{\beta}^{yx},\boldsymbol{\alpha}^{x},\boldsymbol{\delta}^{yx},\boldsymbol{\Psi}^{yx})$ est celle d'un système de régressions linéaires

empilées Gaussien avec des contraintes inter-équations, celles liées au fait que les paramètres α_k^x sont partagés par les équations d'offre de rendement et de demande d'intrants de la culture *k*.

La vraisemblance en $(\boldsymbol{\beta}, \boldsymbol{\alpha}, \boldsymbol{\delta}, \boldsymbol{\Psi}^s)$ de l'observation des choix d'assolement $\mathbf{s}_{it}$ conditionnelle en $\mathbf{z}_{it}$, notée $\ell_{it}^s(\boldsymbol{\beta}, \boldsymbol{\alpha}, \boldsymbol{\delta}, \boldsymbol{\Psi}^s)$, est de forme plus particulière en raison de la forme spécifique des modèles de choix d'assolement MNL. Bien que pénible à calculer,[2] elle s'avère être de forme simple, avec :

(40) $$\ell_{it}^s(\boldsymbol{\beta}, \boldsymbol{\alpha}, \boldsymbol{\delta}, \boldsymbol{\Psi}^s) = (\alpha^s)^{1-G} \left(\prod_{g \in \mathcal{G}} (\alpha_{(g)}^s)^{1-K_{(g)}} \right) \left(\prod_{k \in \mathcal{K}} s_{k,it}^{-1} \right) \times \varphi\left(\boldsymbol{\varepsilon}_{it}^s(\boldsymbol{\beta}, \boldsymbol{\alpha}, \boldsymbol{\delta}); \boldsymbol{\Psi}^s \right)$$

dans le cas NMNL où :

(41) $$\boldsymbol{\varepsilon}_{it}^s(\boldsymbol{\beta}, \boldsymbol{\alpha}, \boldsymbol{\delta}) \equiv \left(\varepsilon_{k,it}^s(\boldsymbol{\beta}, \boldsymbol{\alpha}, \boldsymbol{\delta}) : k \in \mathcal{K}_{-K} \right).$$

Cette vraisemblance est déduite de notre modèle paramétrique de la fonction de densité de $\mathbf{s}_{it}$ conditionnelle en $\mathbf{z}_{it}$, $f(\mathbf{s}_{it} \mid \mathbf{z}_{it})$.

La vraisemblance en $(\boldsymbol{\beta}, \boldsymbol{\alpha}, \boldsymbol{\delta}, \boldsymbol{\Psi})$ de l'observation des choix de production $(\mathbf{y}_{it}, \mathbf{x}_{it}, \mathbf{s}_{it})$

---

[2] Cette vraisemblance s'obtient en utilisant le Théorème de changement de variable appliquée aux fonctions de densité multivariée et utilisant les identités

$\det(\mathbf{A} + \gamma \times \mathbf{u}') = (1 + \gamma \times \boldsymbol{\iota}'\mathbf{A}^{-1}\boldsymbol{\iota}) \times \det \mathbf{A}$ et $(\mathbf{A} + \gamma \times \mathbf{u}')^{-1} = \mathbf{A}^{-1} - \gamma \times \mathbf{A}^{-1}\mathbf{u}'\mathbf{A}^{-1} \times (1 + \gamma \times \boldsymbol{\iota}'\mathbf{A}^{-1}\boldsymbol{\iota})^{-1}$

où γ est un scalaire, $\mathbf{A}$ une matrice symétrique non singulière et $\boldsymbol{\iota}$ le vecteur unitaire de la dimension de $\mathbf{A}$.

conditionnelle en $\mathbf{z}_{it}$ est notée $\ell_{it}(\boldsymbol{\beta},\boldsymbol{\alpha},\boldsymbol{\delta},\boldsymbol{\Psi})$. Elle correspond à modèle paramétrique de la fonction de densité de $(\mathbf{y}_{it},\mathbf{x}_{it},\mathbf{s}_{it})$ conditionnelle en $\mathbf{z}_{it}$, $f(\mathbf{y}_{it},\mathbf{x}_{it},\mathbf{s}_{it}\mid\mathbf{z}_{it})$. Elle est obtenue par la loi de Bayes, avec $f(\mathbf{y}_{it},\mathbf{x}_{it},\mathbf{s}_{it}\mid\mathbf{z}_{it})=f(\mathbf{s}_{it}\mid\mathbf{z}_{it})f(\mathbf{y}_{it},\mathbf{x}_{it}\mid\mathbf{z}_{it},\mathbf{s}_{it})$:

(42) $$\ell_{it}(\boldsymbol{\beta},\boldsymbol{\alpha},\boldsymbol{\delta},\boldsymbol{\Psi})=\ell_{it}^{s}(\boldsymbol{\beta},\boldsymbol{\alpha},\boldsymbol{\delta},\boldsymbol{\Psi}^{s})\ell_{it}^{yx|s}(\boldsymbol{\beta},\boldsymbol{\alpha},\boldsymbol{\delta},\boldsymbol{\Psi})$$

où :

(43) $$\ell_{it}^{yx|s}(\boldsymbol{\beta},\boldsymbol{\alpha},\boldsymbol{\delta},\boldsymbol{\Psi})=\varphi\left(\boldsymbol{\varepsilon}_{it}^{yx}(\boldsymbol{\beta}^{yx},\boldsymbol{\alpha}^{x},\boldsymbol{\delta}^{yx})-(\boldsymbol{\Psi}^{yx})^{-1}\mathbf{C}\boldsymbol{\varepsilon}_{it}^{s}(\boldsymbol{\beta},\boldsymbol{\alpha},\boldsymbol{\delta},\boldsymbol{\Psi});\boldsymbol{\Psi}^{yx}-\mathbf{C}'(\boldsymbol{\Psi}^{s})^{-1}\mathbf{C}\right).$$

Le terme $\ell_{it}^{yx|s}(\boldsymbol{\beta},\boldsymbol{\alpha},\boldsymbol{\delta},\boldsymbol{\Psi})$ correspond à la vraisemblance en $(\boldsymbol{\beta},\boldsymbol{\alpha},\boldsymbol{\delta},\boldsymbol{\Psi})$ de l'observation des rendements et utilisations d'intrants $(\mathbf{y}_{it},\mathbf{x}_{it})$ conditionnelle en $(\mathbf{z}_{it},\mathbf{s}_{it})$, *i.e.* en les variables explicatives exogènes du modèle et en l'observation des choix d'assolement $\mathbf{s}_{it}$. L'expression de ce terme illustre l'intérêt – ou le côté « confortable » – de l'hypothèse de normalité jointe des termes d'erreur de systèmes d'équations.[3]

La forme de $\ell_{it}^{yx|s}(\boldsymbol{\beta},\boldsymbol{\alpha},\boldsymbol{\delta},\boldsymbol{\Psi})$ indique que l'expression de la vraisemblance de $(\mathbf{y}_{it},\mathbf{x}_{it},\mathbf{s}_{it})$ se simplifie significativement lorsque lorsque les termes d'erreur $\boldsymbol{\varepsilon}_{it}^{yx}$ et $\boldsymbol{\varepsilon}_{it}^{s}$ peuvent être supposés indépendants, *i.e.* $\mathbf{C}_0$ est nulle. Dans ce cas, la vraisemblance de $(\mathbf{y}_{it},\mathbf{x}_{it},\mathbf{s}_{it})$ est

---

[3] Le terme $\boldsymbol{\varepsilon}_{it}\equiv(\boldsymbol{\varepsilon}_{it}^{yx},\boldsymbol{\varepsilon}_{it}^{s})$ suivant la loi normale multivariée donnée par l'équation (39), le terme $\boldsymbol{\varepsilon}_{it}^{yx}\mid(\boldsymbol{\varepsilon}_{it}^{s}=\boldsymbol{\varepsilon}^{s})$ suit également la loi une normale multivariée $\mathcal{N}\left((\boldsymbol{\Psi}_0^{yx})^{-1}\mathbf{C}_0\boldsymbol{\varepsilon}^{s},\boldsymbol{\Psi}_0^{yx}-\mathbf{C}_0'(\boldsymbol{\Psi}_0^{s})^{-1}\mathbf{C}_0\right)$ dont la fonction de densité en $\boldsymbol{\varepsilon}^{s}$ s'écrit $\varphi\left(\boldsymbol{\varepsilon}^{ys}-(\boldsymbol{\Psi}_0^{yx})^{-1}\mathbf{C}_0\boldsymbol{\varepsilon}^{s},\boldsymbol{\Psi}_0^{yx}-\mathbf{C}_0'(\boldsymbol{\Psi}_0^{s})^{-1}\mathbf{C}_0\right)$.

simplement le produit des vraisemblances de $(\mathbf{y}_{it}, \mathbf{x}_{it})$ et $\mathbf{s}_{it}$, avec $\ell_{it}(\boldsymbol{\beta}, \boldsymbol{\alpha}, \boldsymbol{\delta}, \boldsymbol{\Psi}) = \ell_{it}^{s}(\boldsymbol{\beta}, \boldsymbol{\alpha}, \boldsymbol{\delta}, \boldsymbol{\Psi}^{s})\ell_{it}^{ys}(\boldsymbol{\beta}^{yx}, \boldsymbol{\alpha}^{x}, \boldsymbol{\delta}^{yx}, \boldsymbol{\Psi}^{yx})$. Cette propriété sera utilisée par la suite.

1.2. Modèles micro-économétriques de choix de production multiculture à paramètres aléatoires

Carpentier et Letort (2014) emploient les modèles micro-économétriques de choix de production multiculture présentés ci-dessus en considérant que les paramètres $\boldsymbol{\beta}$, $\boldsymbol{\alpha}^{x}$ et $\boldsymbol{\alpha}^{s}$ de ces modèles sont fixes, *i.e.* communs à l'ensemble des agriculteurs de l'échantillon considérés.

Ces auteurs notent néanmoins que considérer des versions de ces modèles où les termes $\boldsymbol{\beta}$, $\boldsymbol{\alpha}^{x}$ et $\boldsymbol{\alpha}^{s}$ seraient considérés comme spécifiques à chaque exploitation du panel considéré serait particulièrement utile. L'objectif de cette section est de développer cette idée en mettant en avant les propriétés des modèles micro-économétriques de choix de production multiculture à paramètres aléatoires tout en mettant en évidence les défis liés à leur estimation.

Nous présentons dans un premier temps une version « à paramètres aléatoires » des modèles de choix de production multiculture présentés dans la section précédente. Nous en présentons les avantages et précisons leur interprétation. En particulier, nous discutons le fait qu'un modèle de structure essentiellement statique peut prendre en compte les aspects dynamiques les plus persistants de la production agricole et des choix des agriculteurs.

Nous montrons ensuite que l'estimation d'un modèle de choix de production multiculture

à paramètres aléatoires est très difficile dans un cadre de modélisation non entièrement paramétrique. Ceci nous conduit à proposer une version paramétrique du modèle à la base de nos travaux.

Enfin, nous abordons la question de l'estimation du modèle considéré ici à partir d'estimateurs du Maximum de Vraisemblance Simulée (MVS), les estimateurs privilégiés par les micro-économètres dans ce contexte. Nous montrons en particulier que ces estimateurs sont peu appropriés, en tout cas moins appropriés que d'autres estimateurs, pour des modèles tels que des modèles de choix de production multiculture à paramètres aléatoires.

1.2.1. Modèles micro-économétriques de choix de production multiculture à paramètres aléatoires : intérêts

Un des principaux avantages des données de panel est de permettre l'estimation de modèles rendant compte de la variabilité de déterminants inobservés – dite hétérogénéité inobservée – des phénomènes modélisés (voir, *e.g.*, Woodridge, 2002 ; Arellano et Bonhomme, 2011). Il s'agit ici des caractéristiques inobservées des exploitants (*e.g.*, aptitudes, motivations) et des exploitations (*e.g.*, qualité des sols, distribution spatiale du parcellaire, matériel disponible) qui ne varient pas ou varient peu sur la période considérée.

Dans des modèles tels que celui décrit par le système des équations de rendement et d'utilisation d'intrants considéré ici, l'approche usuelle consiste à spécifier les paramètres constants de chacune des équations comme dépendant des exploitations (voir, *e.g.*, Lacroix et Thomas, 2011 ; Platoni et al, 2012 ; Bayramoglu et Chakir, 2016). Il suffit de considérer

que les paramètres constants des rendements potentiels, β_k^y, et ceux des quantités d'intrants nécessaires à l'obtention de ces rendements potentiels, β_k^x, sont spécifiques à chaque exploitation avec :

$$(44) \qquad \begin{cases} y_{k,it} = \beta_{k,i}^y + (\boldsymbol{\delta}_{k,0}^y)'\mathbf{c}_{k,it}^y - \alpha_{k,0}^x q_{k,it}^2 / 2 + \varepsilon_{k,it}^y \\ x_{k,it} = \beta_{k,i}^x + (\boldsymbol{\delta}_{k,0}^x)'\mathbf{c}_{k,it}^x - \alpha_{k,0}^x q_{k,it}^2 + \varepsilon_{k,it}^x \\ k \in \mathcal{K} \end{cases}$$

Dans ce modèle, les coefficients des variables de prix ou de contrôle sont considérés comme fixes, *i.e.* comme étant communs à l'ensemble des exploitations de l'échantillon considéré. Ces modèles à « constantes individuelles additivement séparables », les termes $\beta_{k,i}^y$ et $\beta_{k,i}^x$, sont standards en économétrie des données de panel (voir, *e.g.*, Arellano et Bonhomme, 2011). Leur estimation est relativement aisée (voir, *e.g.*, Woodridge, 2002 ; Arellano et Bonhomme, 2012).

Employer cette approche de modélisation pour le système des équations d'assolement de notre modèle de choix de production multiculture pose des problèmes techniques inhabituels en ce qui concerne leur estimation. La version log-linéarisée des équations d'assolement NMNL à « constantes individuelles additivement séparables » est un système d'équations linéaires à coefficients fixes (à quelques transformations de paramètres et de variables près) :

$$(45) \qquad \begin{cases} \ln s_{k,it} - \ln s_{K,it} = \alpha_0^s \times \begin{Bmatrix} \pi_{k,it} - \pi_{K,it} - \beta_{k,i}^s - (\boldsymbol{\delta}_{k,0}^s)'\mathbf{c}_{k,it}^s \\ + \left((\alpha_0^s)^{-1} - (\alpha_{(g),0}^s)^{-1}\right) \times \ln s_{k|(g),it} \\ - \left((\alpha_0^s)^{-1} - (\alpha_{(G),0}^s)^{-1}\right) \times \ln s_{K|(G),it} \end{Bmatrix} + \alpha_0^s \varepsilon_{k,it}^s \\ k \in \mathcal{K}_{-K} \end{cases}.$$

dans lequel les termes $\beta_{k,i}^s$ jouent le rôle des effets individuels des modèles linéaires de

données de panel usuels. Les variables de parts de surface des cultures $s_{k|(g),it}$ dans leurs groupes sont endogènes par rapport au terme d'erreur des équations où elles apparaissent, par construction. En effet, $s_{k|(g),it}$ est une fonction, entre autres termes, de $\varepsilon^{s}_{k,it}$. Ce problème d'endogénéité peut être surmonté (voir, *e.g.*, Woodridge, 2002). Mais le système de choix d'assolement décrit ci-dessus n'est standard en économétrie des données de panels que lorsque les marges brutes $\pi_{k,it}$ espérées des cultures sont observées. Les termes $\pi_{k,it}$ sont ici construits à partir des éléments des modèles de rendement et d'utilisation d'intrants et dépendent donc directement des termes individuels $\beta^{y}_{k,i}$ et $\beta^{x}_{k,i}$.[4] Ce type de contraintes inter-équations, lié à la structure du modèle considéré ici, est inhabituel en économétrie des données de panel.

Néanmoins, notre recours à des modèles de choix de production multiculture à paramètres aléatoires ne s'inscrit pas dans une logique d'extension « technique » des modèles économétriques de données de panel usuel. Nous souhaitons employer ces modèles pour

[4] Ce problème n'apparaît pas dans les modèles de choix de production multiculture usuels car les modèles de choix d'assolement intégrés à ces modèles sont construits sans utiliser les notions de marges brutes (voir, *e.g.*, Lacroix et Thomas, 2011 ; Platoni et al, 2012 ; Bayramoglu et Chakir, 2016). Dans ces modèles, les choix d'assolement dépendent directement des prix des cultures et des intrants. Pour Carpentier et Letort (2014), le fait que leurs modèles de choix d'assolement MNL soient conçus à partir des notions de marges brutes est un atout pour des travaux de recherche interdisciplinaires. Les agronomes et les conseillers des agriculteurs manipulent régulièrement les notions de marges brutes et les modèles de choix d'assolement MNL sont des fonctions relativement simples des marges brutes espérées des cultures de l'assolement.

pallier à certaines limites des fondements des modèles présentés ci-dessus. De fait, ajouter des termes spécifiques aux exploitations à chacune des équations des modèles de choix de production multiculture de Carpentier et Letort (2014) s'avère insuffisant pour rendre compte de l'hétérogénéité potentielle des choix d'assolement des agriculteurs.

Nous considérons explicitement nos modèles comme des approximations – structurées par des hypothèses sur la rationalité économique des agriculteurs et sur la forme de la technologie de production multiculture – des logiques de choix de production employées par les agriculteurs. Ces derniers étant hétérogènes dans leurs aptitudes et aspirations et faisant face à des conditions et contraintes hétérogènes, nos modèles doivent être interprétés comme des approximations « locales », *i.e.* adaptées à chaque exploitation de l'échantillon considéré, des logiques de choix de production employées par les agriculteurs.

La construction des modèles de rendement et d'utilisation d'intrants de notre modèle de choix de production multiculture repose sur des hypothèses restrictives (*e.g.* formes quadratiques des fonctions de production des cultures et neutralité face au risque des agriculteurs). Ces hypothèses ont le mérite de structurer nos modèles d'offre de rendement et de demande d'intrants. Supposer que les niveaux de rendements potentiels $\beta_{k,i}^{y}$ et les niveaux utilisations intrants associés $\beta_{k,i}^{x}$ sont spécifiques à chaque agriculteurs permet de rendre compte de l'effet de l'hétérogénéité inobservée des déterminants de choix de production des agriculteurs sur leurs niveaux d'utilisation d'intrants et les niveaux de rendement qu'ils obtiennent.

Néanmoins, les déterminants inobservés des choix de production des agriculteurs peuvent également impacter la manière dont ces derniers ajustent leurs utilisations d'intrant, et donc

leurs niveaux de rendement objectif, aux évolutions des prix des cultures et des intrants. Dans le modèle précédent, ces ajustements dépendent de la valeur des paramètres $\alpha_{k,0}^{x}$ qui est supposée commune à l'ensemble des agriculteurs de l'échantillon. Or, les agriculteurs peuvent répondre différemment à ces évolutions des prix selon leurs pratiques de production qui peuvent être plus ou moins intensives, leur attitude face au risque de rendement ou encore leur attitude vis-à-vis des questions agri-environnementales. Cette variabilité des réactions des agriculteurs peut être prise en compte dans nos modèles d'offre de rendement et de demande d'intrants en spécifiant des paramètres de flexibilité des choix d'intrants spécifiques à chaque exploitation, *i.e.* des paramètres $\alpha_{k,i}^{x}$. Cette extension du modèle décrit dans l'équation (46) fait de ce système d'équations un système de modèles linéaires à paramètres – constantes et coefficients – aléatoires

$$(47) \qquad \begin{cases} y_{k,it} = \beta_{k,i}^{y} + (\boldsymbol{\delta}_{k,0}^{y})'\mathbf{c}_{k,it}^{y} - \alpha_{k,i}^{x} q_{k,it}^{2} / 2 + \varepsilon_{k,it}^{y} \\ x_{k,it} = \beta_{k,i}^{x} + (\boldsymbol{\delta}_{k,0}^{x})'\mathbf{c}_{k,it}^{x} - \alpha_{k,i}^{x} q_{k,it} + \varepsilon_{k,it}^{x} \\ k \in \mathcal{K} \end{cases} .$$

Chacune des équations de ce système est relativement standard mais le fait que les paramètres $\alpha_{k,i}^{x}$ soient toujours partagés par deux équations du systèmes introduit des contraintes inhabituelles.

Introduire de manière très générale les effets de l'hétérogénéité inobservée des déterminants des choix des agriculteurs est certainement encore plus important pour les modèles de choix d'assolement de nos modèles multiculture. En effet, le concept de « fonction de coût de gestion de l'assolement » sur lequel repose la construction des modèles de choix

d'assolement MNL est un concept de nature très générique, pour ne pas dire « fourre-tout ». De fait, la forme et la valeur des paramètres des fonctions $C_{it}(\mathbf{s})$ présentées dans la section précédente déterminent entièrement les choix d'assolement des agriculteurs en fonction des marges brutes espérées des cultures. En réalité, ces choix dépendent de nombreux facteurs dont les effets sont captés par les paramètres de la fonction $C_{it}(\mathbf{s})$ lors de l'estimation de ces paramètres. Nombre de ces facteurs sont spécifiques aux exploitants et à leurs exploitations. Ils incluent la composition – en quantité et en qualité – de leur parc de machines, les schémas de rotation culturale sur lesquels ils s'appuient ou leur stratégie de diversification des risques (voir, *e.g.*, Carpentier et Gohin, 2015), sans compter leur attitude vis-à-vis des questions agri-environnementales. Si les variations des termes $\beta_{k,i}^{s}$ permettent de rendre compte des effets de ces facteurs sur les choix d'assolement, elles ne permettent pas de rendre compte de l'effet des facteurs inobservés spécifiques aux exploitations et aux exploitants sur les réponses de ces choix d'assolement aux variations des marges brutes espérées des cultures.

Introduire des effets de l'hétérogénéité inobservée des exploitants sur ces réponses nécessite de considérer que les paramètres de flexibilité des choix d'assolement $\boldsymbol{\alpha}^{s}$ sont eux aussi spécifiques aux exploitants. Ceci revient à considérer la fonction de coût de gestion de l'assolement suivante :

$$
\begin{aligned}
C_{it}(\mathbf{s}) = D_{it} &+ \sum\nolimits_{k\in\mathcal{K}} s_k \left(\beta_{k,i}^{s} + (\boldsymbol{\delta}_{k,0}^{s})'\mathbf{c}_{k,it}^{s} + \varepsilon_{k,it}^{s}\right) \\
&+ (\alpha_i^{s})^{-1} \sum\nolimits_{g\in\mathcal{G}} s_{(g)} \ln s_{(g)} + \sum\nolimits_{g\in\mathcal{G}} s_{(g)} (\alpha_{(g),i}^{s})^{-1} \sum\nolimits_{k\in\mathcal{K}(g)} s_{k|(g)} \ln s_{k|(g)}
\end{aligned}
\tag{48}
$$

et donc à considérer des modèles de choix d'assolement NMNL à paramètres aléatoires dont la version log-linéarisée est donnée par :

$$
(49) \qquad \begin{cases} \ln s_{k,it} - \ln s_{K,it} = \alpha_i^s \times \begin{Bmatrix} \pi_{k,it} - \pi_{K,it} - \beta_{k,i}^s - (\boldsymbol{\delta}_{k,0}^s)'\mathbf{c}_{k,it}^s \\ +\left((\alpha_i^s)^{-1} - (\alpha_{(g),i}^s)^{-1}\right) \times \ln s_{k|(g),it} \\ -\left((\alpha_i^s)^{-1} - (\alpha_{(G),i}^s)^{-1}\right) \times \ln s_{K|(G),it} \end{Bmatrix} + \alpha_i^s \varepsilon_{k,it}^s \\ k \in \mathcal{K}_{-K} \end{cases}.
$$

Ce système d'équations définit un système de modèles linéaires (à quelques transformations près) à paramètres – constantes et coefficients – aléatoires.

1.2.2. Modèles micro-économétriques de choix de production multiculture à paramètres aléatoires : estimation semi-paramétrique ?

Considérer la version log-linéarisée des modèles de choix d'assolement NMNL de notre modèle de choix de production multiculture à paramètres aléatoires permet d'obtenir un système d'équations avec des termes d'erreur additivement séparables dans chacune des équations :

$$
(50) \qquad \begin{cases} \begin{cases} y_{k,it} = \beta_{k,i}^y + (\boldsymbol{\delta}_{k,0}^y)'\mathbf{c}_{k,it}^y - \alpha_{k,i}^x q_{k,it}^2 / 2 + \varepsilon_{k,it}^y \\ x_{k,it} = \beta_{k,i}^x + (\boldsymbol{\delta}_{k,0}^x)'\mathbf{c}_{k,it}^x - \alpha_{k,i}^x q_{k,it} + \varepsilon_{k,it}^x \\ k \in \mathcal{K} \end{cases} \\ \begin{cases} \ln s_{k,it} - \ln s_{K,it} = \alpha_i^s \times \begin{Bmatrix} \pi_{k,it} - \pi_{K,it} - \beta_{k,i}^s - (\boldsymbol{\delta}_{k,0}^s)'\mathbf{c}_{k,it}^s \\ +\left((\alpha_i^s)^{-1} - (\alpha_{(g),i}^s)^{-1}\right) \times \ln s_{k|(g),it} \\ -\left((\alpha_i^s)^{-1} - (\alpha_{(G),i}^s)^{-1}\right) \times \ln s_{K|(G),it} \end{Bmatrix} + \alpha_i^s \varepsilon_{k,it}^s \\ k \in \mathcal{K}_{-K} \end{cases} \\ \begin{cases} \pi_{k,it} = p_{k,it}\left(\beta_{k,i}^y + (\boldsymbol{\delta}_{k,0}^y)'\mathbf{c}_{k,it}^{ys}\right) - v_{k,it}\left(\beta_{k,i}^x + (\boldsymbol{\delta}_{k,0}^x)'\mathbf{c}_{k,it}^{xs}\right) + \alpha_{k,i}^x q_{k,it}^2 p_{k,it} / 2 + a_{k,it} \\ k \in \mathcal{K} \end{cases} \end{cases}
$$

Les modèles linéaires à coefficients aléatoires sont utilisés depuis longtemps en économétrie

des données de panel (voir, *e.g.*, Hsiao et Pesaran, 2008). En particulier, Swamy (1970) a proposé des estimateurs de la MM pour les espérance et matrice de variance-covariance du vecteur de coefficient de tels modèles sous deux conditions : les variables explicatives du modèle ne sont corrélées ni à son terme d'erreur, ni à ses coefficients aléatoires. Aucune de ces conditions n'est vérifiée pour les équations d'assolement de notre modèle. En particulier, les termes $s_{k|(g),it}$ dépendent, par construction, des termes d'erreur $\varepsilon^{s}_{k,it}$ et des coefficients $\alpha^{s}_{(g),i}$. Même si les économètres se sont récemment beaucoup penchés sur les questions posées par l'estimation semi-paramétrique des modèles à coefficients aléatoires sous des hypothèses moins restrictives que celles de Swamy (1970), les résultats disponibles à ce jour ne nous permettent pas d'envisager l'estimation du modèle décrit par l'équation (50) sans hypothèses de distribution paramétrique sur ses termes d'erreur et ses paramètres aléatoires. De fait, l'estimation des modèles linéaires à paramètres aléatoires est généralement considérée dans le cas d'une seule équation et dans des cas particuliers lorsque des questions d'endogénéité des variables explicatives sont abordées (voir, *e.g.*, Angrist et Pischke, 2009 ; Arellano et Bonhomme, 2011, 2012 ; Wooldridge, 2005, 2010 ; Murtazashvili et Wooldridge, 2008, 2016).

Or notre modèle de choix de production multiculture à paramètres aléatoires (*a*) est multivarié, (*b*) est constitué d'équations entretenant de nombreuses relations et (*c*) contient – dans sa version log-linéarisée – des variables explicatives endogènes par rapport au terme d'erreur de leur équation et corrélées à leur coefficient aléatoire. De fait, ces problèmes d'estimation majeurs peuvent être gérés en considérant une version entièrement paramétrique

de notre modèle de choix de production multiculture à paramètres aléatoires, *i.e.* en imposant des hypothèses de distribution paramétrique sur ses termes d'erreur et ses termes aléatoires.

Un modèle à paramètres aléatoires entièrement paramétrique décrit explicitement, *i.e.* dans une logique « en information complète », les relations – structurelles ou statistiques – à l'origine des problèmes d'endogénéité et de corrélation entre paramètres aléatoires et variables explicatives le caractérisant. Ces relations sont explicitement prises en compte dans l'écriture de la vraisemblance du modèle et, par conséquent, sont gérées de manière efficace par les estimateurs du Maximum de Vraisemblance (MV) des paramètres du modèle.

De fait, nous faisons face ici à un arbitrage habituel. Alors que les micro-économètres sont réticents à employer des modèles entièrement paramétriques, les spécificités des modèles qui permettent d'aborder certaines questions de recherche ne peuvent être « gérées » empiriquement qu'en considérant des versions entièrement paramétriques de ces modèles. Par exemple, la plupart des modèles de choix discrets (voir, *e.g.*, Train, 2009), des modèles de frontières stochastiques (voir, *e.g.*, Greene, 2008) ou des modèles à variables dépendantes limitées (voir, *e.g.*, Wooldridge, 2002) employés en micro-économétrie appliquée sont des modèles entièrement paramétriques.

De manière plus spécifique, Keane (2009) considère l'estimation d'un modèle multivarié à paramètres aléatoires – combinant modèles de choix de production et de marchés d'échanges internationaux dans son cas – similaire à ceux considérés dans nos applications des Chapitres 5 et 6. Pour lui, l'utilisation d'hypothèses de distributions paramétriques sur les termes aléatoires de son modèle est un prix raisonnable à payer pour être en mesure d'estimer un modèle de choix complexe rendant compte de l'hétérogénéité inobservée des

individus considérés dans l'échantillon utilisé. Nous souscrivons à ce point de vue.

Dans notre cas l'estimation de modèles à paramètres aléatoires entièrement paramétriques présente un avantage supplémentaire. Les hypothèses paramétriques sur les termes aléatoires – termes d'erreur et paramètres aléatoires – de notre modèle de choix de production nous permettent de calibrer aisément les valeurs des paramètres individuels de notre modèle exploitation par exploitation afin de construire un modèle de simulation des choix de production composé d'un ensemble de modèles spécifiques à chaque exploitation de l'échantillon. Cette procédure de calibration détermine les paramètres spécifiques à chaque exploitation à partir d'une logique qui peut être résumée par : « Dis-moi ce que tu fais, je te dirai qui tu es ». En effet, le modèle que j'ai estimé (*a*) m'indique comment ce que tu es détermine ce que tu fais et (*b*) me donne des idées *a priori* sur qui tu peux être ».

1.2.3. Modèles micro-économétriques de choix de production multiculture à paramètres aléatoires : spécification paramétrique

Comme cela sera vu dans la suite, la log-linéarisation du modèle de choix d'assolement NMNL de notre modèle multiculture n'a que peu d'intérêt lorsque ce modèle est présenté sous une forme entièrement paramétrique. Aussi, dans cette sous-section, nous présentons la version paramétrique du modèle de choix de production multiculture à paramètres aléatoires dans sa forme non-linéaire, avec :

$$
(51) \qquad \begin{cases}
\begin{cases} y_{k,it} = \beta^{y}_{k,i} + (\boldsymbol{\delta}^{y}_{k,0})'\mathbf{c}^{y}_{k,it} - \alpha^{x}_{k,i} q^{2}_{k,it} / 2 + \varepsilon^{y}_{k,it} \\ x_{k,it} = \beta^{x}_{k,i} + (\boldsymbol{\delta}^{x}_{k,0})'\mathbf{c}^{x}_{k,it} - \alpha^{x}_{k,i} q_{k,it} + \varepsilon^{x}_{k,it} \\ k \in \mathcal{K} \end{cases} \\
\begin{cases} s_{k,it} = \dfrac{\exp\left(\alpha^{s}_{(g),i}(\pi_{k,it} - \beta^{s}_{k,it})\right)\left(\sum_{\ell \in \mathcal{K}(g)} \exp\left(\alpha^{s}_{(g),i}(\pi_{\ell,it} - \beta^{s}_{\ell,it})\right)\right)^{\alpha^{s}_{i}(\alpha^{s}_{(g),i})^{-1}-1}}{\sum_{h \in \mathcal{G}}\left(\sum_{\ell \in \mathcal{K}(h)} \exp\left(\alpha^{s}_{(h),i}(\pi_{\ell,it} - \beta^{s}_{\ell,it})\right)\right)^{\alpha^{s}_{i}(\alpha^{s}_{(h),i})^{-1}}} \\ \beta^{s}_{k,it} = \beta^{s}_{k,i} + (\boldsymbol{\delta}^{s}_{k})'\mathbf{c}^{s}_{k,it} + \varepsilon^{s}_{k,it} \\ k \in \mathcal{K}_{-K} \\ \beta^{s}_{K,it} = 0 \end{cases} \\
\begin{cases} \pi_{k,it} = p_{k,it}\left(\beta^{y}_{k,i} + (\boldsymbol{\delta}^{y}_{k})'\mathbf{c}^{ys}_{k,it}\right) - v_{k,it}\left(\beta^{x}_{k,i} + (\boldsymbol{\delta}^{x}_{k})'\mathbf{c}^{xs}_{k,it}\right) + \alpha^{x}_{k,i} q^{2}_{k,it} p_{k,it} / 2 + a_{k,it} \\ k \in \mathcal{K} \end{cases}
\end{cases} .
$$

De fait, la non-linéarité du modèle des $s_{k,it}$ considéré ici et les problèmes d'endogénéité et de corrélation entre paramètres aléatoires et variables explicatives de sa version log-linéarisée sont les deux faces d'une même pièce.

Spécifier une version paramétrique de ce modèle consiste à spécifier (*a*) la distribution des vecteurs de paramètres aléatoires $\boldsymbol{\beta}^{yx}_{i}$, $\boldsymbol{\beta}^{s}_{i}$, $\boldsymbol{\alpha}^{x}_{i}$ et $\boldsymbol{\alpha}^{s}_{i}$, (*b*) la distribution des vecteur des termes d'erreur $\boldsymbol{\varepsilon}_{it}$ et (*c*) les relations de ces termes aléatoires entre eux et avec les variables explicatives $\mathbf{z}_{it}$.

Comme dans la section précédente, nous supposons que le vecteur des termes d'erreur $\boldsymbol{\varepsilon}_{it}$ du modèle suit une loi normale multivariée :

$$
(52) \qquad \boldsymbol{\varepsilon}_{it} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Psi}_{0})
$$

Cette hypothèse de normalité est usuelle dans ce contexte.

Nous supposons également que les termes d'erreur $\boldsymbol{\varepsilon}_{it}$ sont indépendants entre eux, tant dans la dimension individuelle que dans la dimension temporelle, *i.e.* que :

(53) les $\boldsymbol{\varepsilon}_{it}$ sont indépendants pour $t = 1,...,T$ et $i = 1,...,N$.[5]

L'hypothèse d'indépendance est usuelle dans la dimension individuelle. Elle est d'autant plus justifiée ici que les paramètres individuels du modèle captent l'essentiel des effets qui peuvent lier les choix des individus. L'absence de corrélation des termes d'erreur $\boldsymbol{\varepsilon}_{it}$ dans la dimension temporelle est une hypothèse plus discutable *a priori*. Elle renvoie à l'interprétation des paramètres aléatoires de ces modèles.

De fait, les modèles de choix de production considérés ici sont, tout au moins en apparence, essentiellement statiques. En effet, les choix passés d'un agriculteur n'affectent pas ses choix courants dans ces modèles. De même, supposer que les termes d'erreur ne sont pas corrélés dans le temps suggère que les déterminants inobservés des performances et choix passés ne sont pas corrélés dans le temps.

Ces conclusions sont à nuancer ici en raison de la spécification de paramètres aléatoires spécifiques à chaque exploitation dans le modèle. En effet, ces paramètres sont spécifiques à chaque exploitation. Ils peuvent donc capter les effets dynamiques les plus persistants des

[5] Nous supposons ici que nous disposons de données issues d'un panel cylindré. Les calculs et formules reportées ici s'adaptent aisément au cas d'un panel non cylindré dès lors que, comme c'est le cas ici, les observations des exploitations sont sélectionnées de manière exogène.

processus de production agricole et des choix des agriculteurs. De fait, la relative stabilité dans le temps des choix d'assolement des agriculteurs suggère que ces derniers suivent, dans une large mesure, des schémas de rotations de leurs cultures relativement stables dans le temps. Une remarque similaire s'applique à leurs choix d'intrants. Ces derniers sont relativement stables dans le temps et suggèrent que les agriculteurs suivent des pratiques de production très proches d'une année sur l'autre, avec des ajustements limités des utilisations d'intrants dans le temps. Cette stabilité des choix de production implique en retour une stabilité des processus dynamiques de la production qui peuvent en grande partie être captés par les paramètres aléatoires du modèle.

Les estimations et tests présentés dans le Chapitre 5 de cette thèse suggère effectivement que les agriculteurs suivent des schémas de rotation des cultures relativement stables dans le temps avec des ajustements à la marge en fonction des aléas climatiques et des évolutions de prix.[6] La structure des modèles estimés est donc essentiellement statique. Mais ces modèles

[6] En outre, spécifier et estimer des modèles de choix de production multiculture réellement dynamiques s'avère très difficile. D'une part, il est maintenant bien établi qu'il est difficile de distinguer les effets dynamiques des effets de l'hétérogénéité individuelle inobservée affectant un processus lorsque les effets dynamiques considérés sont persistants (voir, *e.g.*, Angrist et Pischke, 2009 ; Arellano et Bonhomme, 2011, 2012) comme cela semble être le cas pour les choix de production multiculture des producteurs français de grandes cultures. D'autre part, spécifier des modèles de choix dynamiques suppose de spécifier les aspects dynamiques de la production agricole à l'origine de l'intérêt des aspects dynamiques de ces choix et d'inscrire la modélisation des choix des agriculteurs dans un cadre – d'optimisation – inter-temporel. En particulier, considérer que les choix d'assolement des agriculteurs sont des processus intrinsèquement dynamiques nécessite de spécifier explicitement les effets des successions culturales sur les performances productives des cultures et les contraintes inter-temporelles sur les choix d'assolement qui résultent de la prise en compte de ces effets (voir, *e.g.*, Carpentier et Gohin, 2015).

peuvent être interprétés comme des modèles de choix de production routiniers générant et s'inscrivant dans une dynamique de production relativement stable.

Bien entendu, il n'est pas ici question d'estimer (tout au moins directement) les vecteurs de paramètres $\boldsymbol{\beta}_i^{yx}$, $\boldsymbol{\beta}_i^{s}$, $\boldsymbol{\alpha}_i^{x}$ et $\boldsymbol{\alpha}_i^{s}$ pour chacune des exploitations de l'échantillon. La dimension temporelle nécessaire pour obtenir des estimations fiables est trop importante eu égard à la dimension temporelle des panels habituellement disponibles. Dans les échantillons employés pour les estimations présentées dans les Chapitres 5 et 6, les choix des agriculteurs sont observés sur une période allant de 4 à 7 ans. L'objet de l'estimation est d'estimer la distribution jointe de ces paramètres dans la population d'agriculteurs représentée par l'échantillon disponible.

Assez classiquement, nous supposons que le vecteur composé de l'ensemble de ces paramètres (à une transformation près) suit, conditionnellement en $\mathbf{z}_i$, une loi normale multivariée, avec :

(54) $$\boldsymbol{\gamma}_i \sim \mathcal{N}(\boldsymbol{\mu}_0, \boldsymbol{\Omega}_0) \quad \text{où} \quad \boldsymbol{\gamma}_i \equiv (\boldsymbol{\beta}_i^{yx}, \ln \boldsymbol{\alpha}_i^{x}, \boldsymbol{\beta}_i^{s}, \ln \boldsymbol{\alpha}_i^{s})$$

La log-transformation des paramètres de flexibilité $\boldsymbol{\alpha}_i^{x}$ et $\boldsymbol{\alpha}_i^{s}$ implique que ces termes sont supposés log-normaux et assurés d'être positifs. La non-négativité de ces paramètres de flexibilité est nécessaire pour que le modèle de choix considéré se « comporte correctement » d'un point de vue économique. Dans le Chapitre 5, d'autres transformations de la loi normale jointe ont été utilisées pour différents éléments de $\boldsymbol{\gamma}_i$ afin d'examiner la robustesse des résultats obtenus vis-à-vis du choix de la distribution des paramètres aléatoires du modèle.

Aucune restriction n'est imposée sur la matrice de variance-covariance $\boldsymbol{\Omega}_0$ afin de

pouvoir examiner les relations entre les différents paramètres aléatoires du modèle. Cette approche a l'inconvénient d'accroître significativement la dimension du problème d'estimation considéré.

Il est possible de spécifier l'hétérogénéité inobservée de nos modèles à partir d'indices latents spécifiques (voir, *e.g.*, Eaton et al, 2011) afin de réduire la dimension du problème d'estimation considéré. Par exemple, les paramètres de rendement potentiel $\beta_{k,i}^{y}$ peuvent être spécifiés comme des fonctions linéaires dépendant d'un « indice latent de productivité » β_{i}^{y} propre à chaque exploitation. Avec $\beta_{k,i}^{y} = \mu_{k,0} + b_{k,0}\beta_{i}^{y} + u_{k,i}^{y}$, où les termes d'espérance nulle β_{i}^{y} et $u_{k,i}^{y}$ sont supposés mutuellement indépendants pour $k \in \mathcal{K}$, le nombre de paramètres caractérisant l'hétérogénéité inobservée du modèle diminue significativement. Cette hypothèse de modélisation revient de fait à imposer des restrictions sur les paramètres de covariance de la matrice $\mathbf{\Omega}_0$. Cette approche de modélisation peut être étendue au cas des paramètres d'utilisation d'intrants $\beta_{k,i}^{x}$, avec un « indice latent d'intensité des pratiques culturales » β_{i}^{x} propre à chaque exploitant potentiellement lié à l'indice de productivité β_{i}^{y}. Néanmoins, cette approche paraît peu adaptée pour les paramètres aléatoires des choix d'assolement. En effet, les éléments des paramètres $\boldsymbol{\beta}_{i}^{s}$ et $\boldsymbol{\alpha}_{i}^{s}$ ont des interprétations moins directes que les paramètres de rendement potentiel et d'utilisation d'intrants $\beta_{k,i}^{y}$ et $\beta_{k,i}^{x}$.

Les estimations calculées avec des indices latents tels que $\beta_{k,i}^{y}$ et $\beta_{k,i}^{x}$ se sont avérées encourageantes dans certains cas, plus décevantes dans d'autres. Nous avons privilégié les approches « sans restrictions » dans nos travaux empiriques pour nous concentrer sur les

problèmes qui se posent pour l'estimation des modèles de choix de production multiculture à paramètres aléatoires, que ces restrictions soient imposées ou non.

Nous supposons que :

(55) les termes $\mathbf{z}_i$, $\boldsymbol{\gamma}_i$ et $\boldsymbol{\varepsilon}_i$ sont mutuellement indépendants.

où $\mathbf{z}_i \equiv (\mathbf{z}_{it} : t = 1,..,T)$ et $\boldsymbol{\varepsilon}_i \equiv (\boldsymbol{\varepsilon}_{it} : t = 1,...,T)$. Cette hypothèse implique une forme d'exogénéité forte des variables explicatives $\mathbf{z}_i$ ainsi que l'absence de corrélation entre ces variables explicative et les paramètres aléatoires $\boldsymbol{\gamma}_i$.

L'indépendance entre les termes $\boldsymbol{\varepsilon}_i$ et $\boldsymbol{\gamma}_i$ impliquent l'indépendance entre les termes $\boldsymbol{\varepsilon}_{it}$ et $\mathbf{z}_i$. Si elle implique l'hypothèse d'exogénéité invoquée par Carpentier et Letort (2014), elle impose une hypothèse d'exogénéité plus forte puisqu'elle implique l'hypothèse d'exogénéité stricte des $\mathbf{z}_i$ par rapport aux $\boldsymbol{\varepsilon}_{it}$, *i.e.* $E[\boldsymbol{\varepsilon}_{it} \mid \mathbf{z}_i] = \mathbf{0}$. Cette hypothèse est raisonnable compte-tenu du contenu de $\mathbf{z}_i$. Ce vecteur ne contient que des variables de prix et des variables de contrôle purement exogènes – les variables de conditions climatiques – et des variables prédéterminées – les caractéristiques des exploitations. L'hypothèse d'indépendance entre les termes $\boldsymbol{\varepsilon}_{it}$ et $\mathbf{z}_i$ implique également l'homoscédasticité conditionnelle en $\mathbf{z}_i$ des termes d'erreur $\boldsymbol{\varepsilon}_{it}$. Comme indiqué précédemment, cette hypothèse peut s'avérer discutable pour ce qui concerne les variables climatiques contenues dans $\mathbf{z}_i$. Ces dernières peuvent impacter l'état sanitaire des cultures et, donc, la variabilité des rendements et des utilisations d'intrants.

L'hypothèse d'indépendance des paramètres aléatoires $\boldsymbol{\gamma}_i$ et des variables explicatives $\mathbf{z}_i$

du modèle est standard lorsque ces variables contiennent essentiellement des variables de contrôle et de prix, les paramètres aléatoires captant les effets de déterminants non captés par les éléments de $\mathbf{z}_i$.[7] L'hypothèse d'indépendance entre les paramètres aléatoires $\boldsymbol{\gamma}_i$ et les termes d'erreur $\boldsymbol{\varepsilon}_i$ est standard dans ce contexte.

Enfin, nous supposons également que les termes d'erreur des modèles de rendement et d'utilisation d'intrants et les termes d'erreur des modèles de choix d'assolement sont indépendants:

(56) les termes $\boldsymbol{\varepsilon}_{it}^{yx}$ et $\boldsymbol{\varepsilon}_{it}^{s}$ sont indépendants

avec :

(57) $$\boldsymbol{\varepsilon}_{it}^{yx} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Psi}_0^{yx}) \text{ et } \boldsymbol{\varepsilon}_{it}^{s} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Psi}_0^{s}).$$

Cette hypothèse simplifie singulièrement nombre des calculs liés à la mise en œuvre des procédures d'estimations de nos modèles. Elle n'est pas imposée dans l'application présentée dans le Chapitre 5 mais elle est imposée dans le Chapitre 6. Les tests réalisés dans le Chapitre 5 tendent à montrer que les termes d'erreur $\boldsymbol{\varepsilon}_{it}^{yx}$ et $\boldsymbol{\varepsilon}_{it}^{s}$ sont indépendants dans les modèles de choix de production multiculture à paramètres aléatoires considérés ici. Cette hypothèse n'est

---

[7] Il aurait été possible d'introduire les effets des variables de contrôle invariantes dans le temps dans l'espérance des paramètres aléatoires. Il est à noter que ces effets sont déjà captés par la spécification des $\boldsymbol{\beta}_{it}$. Cette approche aurait simplement enrichi la spécification des paramètres de flexibilité $\boldsymbol{\alpha}_i$. Néanmoins, les résultats obtenus lors des essais estimations des Chapitres 5 et 6 tendent à montrer que, de manière générale, le pouvoir explicatif des variables de contrôle utilisées dans nos applications est très limité.

jamais imposée dans les modèles de choix de production multiculture usuels.

Deux points peuvent être mis en avant pour justifier cette hypothèse et expliquer les résultats obtenus dans le Chapitre 5. Tout d'abord, et peut-être surtout, les caractéristiques des exploitations qui peuvent expliquer des relations systématiques entre les rendements, utilisations d'intrants et assolements des agriculteurs sont prises en compte dans les variables de contrôle $\mathbf{z}_i$ ou les paramètres aléatoires $\boldsymbol{\gamma}_i$. Ceci implique que les termes d'erreur contiennent essentiellement les effets d'évènements aléatoires – climatiques et sanitaires – impactant les choix de production et les performances des agriculteurs. Ensuite, les termes $\boldsymbol{\varepsilon}_{it}^{yx}$ contiennent les effets d'évènements aléatoires affectant les niveaux de rendement et d'utilisation d'intrants des agriculteurs. La plupart de ces évènements surviennent bien après ceux dont les effets sont captés par les termes $\boldsymbol{\varepsilon}_{it}^{s}$. En effet, les termes $\boldsymbol{\varepsilon}_{it}^{s}$ contiennent les effets d'évènements survenant avant ou au moment des choix d'assolement, i.e. les effets d'évènements affectant l'état des parcelles (*e.g.*, structure, humidité, état sanitaire ou contenu en éléments nutritifs des sols). Soit les agriculteurs adaptent leurs assolements, voire leurs choix de cultures, à l'état de leurs parcelles en modifiant peu leurs utilisations d'intrants, soit ils adaptent leurs utilisations d'intrants (*e.g.*, apports d'azote, travail du sol) ou certaines techniques (*e.g.*, date des semis) sans modifier leurs assolements. Dans les deux cas ils adoptent les solutions qui leur permettent d'espérer des niveaux des rendements satisfaisants pour les cultures assolées.

1.2.4. Modèles micro-économétriques de choix de production multiculture à paramètres aléatoires : vraisemblance

L'objectif de cette section est double. Il s'agit tout d'abord de dériver la vraisemblance du modèle de choix de production multiculture à paramètres aléatoires présenté ci-dessus, puis de présenter les difficultés liées aux estimateurs du Maximum de Vraisemblance Simulée (MVS) pour les paramètres de ce modèle.

Estimer le modèle de choix de production multiculture à paramètres aléatoires présenté ci-dessus consiste à estimer ses paramètres « fixes ». Ces derniers sont de trois types.

Le premier type regroupe les coefficients des variables de contrôle $\mathbf{c}_{it}$ du modèle. Ces coefficients sont collectés dans le vecteur $\boldsymbol{\delta}_0 \equiv (\boldsymbol{\delta}_0^{yx}, \boldsymbol{\delta}_0^{s})$. Le second type de paramètres regroupe les éléments non nuls de la matrice de variance-covariance $\boldsymbol{\Psi}_0$ du vecteur des termes d'erreur $\boldsymbol{\varepsilon}_{it}$ du modèle, *i.e.* de $\boldsymbol{\Psi}_0^{yx}$ pour $\boldsymbol{\varepsilon}_{it}^{yx}$ et de $\boldsymbol{\Psi}_0^{s}$ pour $\boldsymbol{\varepsilon}_{it}^{s}$. Le troisième type de paramètres regroupe les paramètres de la distribution des paramètres aléatoires $\boldsymbol{\gamma}_i \equiv (\boldsymbol{\gamma}_i^{yx}, \boldsymbol{\gamma}_i^{s})$ où $\boldsymbol{\gamma}_i^{yx} \equiv (\boldsymbol{\beta}_i^{yx}, \ln \boldsymbol{\alpha}_i^{x})$ et $\boldsymbol{\gamma}_i^{s} \equiv (\boldsymbol{\beta}_i^{s}, \ln \boldsymbol{\alpha}_i^{s})$, *i.e.* l'espérance $\boldsymbol{\mu}_0$ et la matrice de variance-covariance $\boldsymbol{\Omega}_0$.

Les observations d'une même exploitation étant liées par le vecteur de paramètres aléatoires spécifiques à cette exploitation $\boldsymbol{\gamma}_i$, il convient ici de déterminer la vraisemblance des choix de production observés pour chaque exploitation sur l'ensemble de la période d'observation, *i.e.* du vecteur $(\mathbf{y}_i, \mathbf{x}_i, \mathbf{s}_i) \equiv \left((\mathbf{y}_{it}, \mathbf{x}_{it}, \mathbf{s}_{it}) : t = 1, ..., T\right)$. Cette vraisemblance s'obtient à partir de la fonction de densité de $(\mathbf{y}_i, \mathbf{x}_i, \mathbf{s}_i)$ conditionnelle en les variables exogènes $\mathbf{z}_i$, $f(\mathbf{y}_i, \mathbf{x}_i, \mathbf{s}_i \mid \mathbf{z}_i)$. Les paramètres aléatoires $\boldsymbol{\gamma}_i$ structurant le modèle considéré, cette fonction s'écrit sous la forme de l'espérance de $(\mathbf{y}_i, \mathbf{x}_i, \mathbf{s}_i)$ conditionnelle en les variables exogènes et

les paramètres aléatoires $(\mathbf{z}_i, \boldsymbol{\gamma}_i)$ intégrée sur la distribution de probabilité des paramètres aléatoires $\boldsymbol{\gamma}_i$ conditionnelle en $\mathbf{z}_i$. Les termes $\boldsymbol{\gamma}_i$ et $\mathbf{z}_i$ étant supposés indépendants on obtient la fonction de densité $f(\mathbf{y}_i, \mathbf{x}_i, \mathbf{s}_i \mid \mathbf{z}_i)$ sous la forme de l'intégrale multivariée suivante:

(58) $$f(\mathbf{y}_i, \mathbf{x}_i, \mathbf{s}_i \mid \mathbf{z}_i) = \int f(\mathbf{y}_i, \mathbf{x}_i, \mathbf{s}_i \mid \mathbf{z}_i, \boldsymbol{\gamma}) f(\boldsymbol{\gamma}) d\boldsymbol{\gamma} \ .$$

Les termes d'erreur $\boldsymbol{\varepsilon}_{it}$ étant supposés indépendants de $(\mathbf{z}_i, \boldsymbol{\gamma}_i)$ et indépendants entre eux dans la dimension temporelle, les vecteurs $(\mathbf{y}_{it}, \mathbf{x}_{it}, \mathbf{s}_{it})$ sont indépendants conditionnellement en $\boldsymbol{\gamma}_i$ et $\mathbf{z}_i$, impliquant que[8] $f(\mathbf{y}_i, \mathbf{x}_i, \mathbf{s}_i \mid \mathbf{z}_i, \boldsymbol{\gamma}_i) = \prod_{t=1}^{T} f(\mathbf{y}_{it}, \mathbf{x}_{it}, \mathbf{s}_{it} \mid \mathbf{z}_{it}, \boldsymbol{\gamma}_i)$ et donc que :

(59) $$f(\mathbf{y}_i, \mathbf{x}_i, \mathbf{s}_i \mid \mathbf{z}_i) = \int \prod_{t=1}^{T} f(\mathbf{y}_{it}, \mathbf{x}_{it}, \mathbf{s}_{it} \mid \mathbf{z}_{it}, \boldsymbol{\gamma}) f(\boldsymbol{\gamma}) d\boldsymbol{\gamma} \ .$$

La fonction de densité des paramètres aléatoires $\boldsymbol{\gamma}_i$, $f(\boldsymbol{\gamma}_i)$, est désignée sous le terme de « densité mélangeante » du modèle par Train (2009).[9] On a :

(60) $$f(\boldsymbol{\gamma}_i) = \varphi(\boldsymbol{\gamma}_i - \boldsymbol{\mu}_0; \boldsymbol{\Omega}_0)$$

par l'hypothèse de normalité des $\boldsymbol{\gamma}_i$ avec $\boldsymbol{\gamma}_i \sim \mathcal{N}(\boldsymbol{\mu}_0, \boldsymbol{\Omega}_0)$. Le terme $\varphi(\boldsymbol{\gamma}_i - \boldsymbol{\mu}; \boldsymbol{\Omega})$ définit la vraisemblance de $\boldsymbol{\gamma}_i$ calculée en $(\boldsymbol{\mu}, \boldsymbol{\Omega})$.

---

[8] On utilise ici également le fait que le modèle de $(\mathbf{y}_{it}, \mathbf{x}_{it}, \mathbf{s}_{it})$ ne dépend de $\mathbf{z}_i$ que *via* $\mathbf{z}_{it}$.

[9] Les modèles paramétriques à paramètres aléatoires étant parfois désignés sous le nom de modèles à loi mélangées dans la littérature statistique.

La fonction de densité des observations $(\mathbf{y}_{it}, \mathbf{x}_{it}, \mathbf{s}_{it})$ conditionnelle en $(\mathbf{z}_{it}, \boldsymbol{\gamma}_i)$, $f(\mathbf{y}_{it}, \mathbf{x}_{it}, \mathbf{s}_{it} \mid \mathbf{z}_{it}, \boldsymbol{\gamma}_i)$, est désignée sous le terme de « densité noyau » du modèle par Train (2009). Les termes d'erreur $\boldsymbol{\varepsilon}_{it}^{yx}$ et $\boldsymbol{\varepsilon}_{it}^{s}$ étant supposés indépendants, la fonction de densité $(\mathbf{y}_{it}, \mathbf{x}_{it}, \mathbf{s}_{it})$ conditionnelle en $(\mathbf{z}_{it}, \boldsymbol{\gamma}_i)$ s'écrit comme le produit des fonctions de densité de $(\mathbf{y}_{it}, \mathbf{x}_{it})$ et $\mathbf{s}_{it}$ conditionnelles en $(\mathbf{z}_{it}, \boldsymbol{\gamma}_i)$:

(61) $$f(\mathbf{y}_{it}, \mathbf{x}_{it}, \mathbf{s}_{it} \mid \mathbf{z}_{it}, \boldsymbol{\gamma}_i) = f(\mathbf{y}_{it}, \mathbf{x}_{it} \mid \mathbf{z}_{it}, \boldsymbol{\gamma}_i) f(\mathbf{s}_{it} \mid \mathbf{z}_{it}, \boldsymbol{\gamma}_i)$$

D'après les résultats présentés dans la section précédente, la densité des rendements et utilisations d'intrants observés $(\mathbf{y}_{it}, \mathbf{x}_{it})$ conditionnelle en les variables exogènes et paramètres aléatoires $(\mathbf{z}_{it}, \boldsymbol{\gamma}_i)$ est donnée par :

(62) $$f(\mathbf{y}_{it}, \mathbf{x}_{it} \mid \mathbf{z}_{it}, \boldsymbol{\gamma}_i) = \varphi\left(\boldsymbol{\varepsilon}_{it}^{yx}(\boldsymbol{\gamma}_i^{yx}, \boldsymbol{\delta}_0^{yx}); \boldsymbol{\Psi}_0^{yx}\right)$$

alors que celle des assolements observés $\mathbf{s}_{it}$ conditionnelle en $(\mathbf{z}_{it}, \boldsymbol{\gamma}_i)$ est donnée par:

(63) $$f(\mathbf{s}_{it} \mid \mathbf{z}_{it}, \boldsymbol{\gamma}_i) = (\alpha_i^s)^{1-G} \left(\prod_{g \in \mathcal{G}} (\alpha_{(g),i}^s)^{1-K_{(g)}}\right)\left(\prod_{k \in \mathcal{K}} s_{k,it}^{-1}\right) \times \varphi\left(\boldsymbol{\varepsilon}_{it}^{s}(\boldsymbol{\gamma}_i, \boldsymbol{\delta}_0); \boldsymbol{\Psi}_0^{s}\right).$$

Compte-tenu des formes fonctionnelles des modèles d'offre de rendement, de demande d'intrants et de choix d'assolement, les termes résiduels $\boldsymbol{\varepsilon}_{it}^{yx}(\boldsymbol{\gamma}_i^{yx}, \boldsymbol{\delta}^{yx})$ et $\boldsymbol{\varepsilon}_{it}^{s}(\boldsymbol{\gamma}_i, \boldsymbol{\delta})$ peuvent s'écrire sous la forme :

(64) $$\boldsymbol{\varepsilon}_{it}^{yx}(\boldsymbol{\gamma}_i^{yx}, \boldsymbol{\delta}^{yx}) = \varsigma_{it}^{yx}(\boldsymbol{\gamma}_i^{yx}) - (\mathbf{Z}_{it}^{yx})'\boldsymbol{\delta}^{yx} \text{ et } \boldsymbol{\varepsilon}_{it}^{s}(\boldsymbol{\gamma}_i, \boldsymbol{\delta}) = \varsigma_{it}^{s}(\boldsymbol{\gamma}_i) - (\mathbf{Z}_{it}^{s})'\boldsymbol{\delta}$$

où les fonctions $\varsigma_{it}^{yx}(.)$ dépendent des variables $(\mathbf{y}_{it}, \mathbf{x}_{it}, \mathbf{z}_{it})$, les fonctions $\varsigma_{it}^{s}(.)$ des variables $(\mathbf{s}_{it}, \mathbf{z}_{it})$ et les matrices $\mathbf{Z}_{it}^{yx}$ et $\mathbf{Z}_{it}^{s}$ des variables $\mathbf{z}_{it}$. Bien entendu, on a également :

(65) $$\boldsymbol{\varepsilon}_{it}(\boldsymbol{\gamma}_i, \boldsymbol{\delta}) = \varsigma_{it}(\boldsymbol{\gamma}_i) - \mathbf{Z}_{it}'\boldsymbol{\delta}$$

avec :

(66) $$\varsigma_{it}(\tilde{\gamma}_{i,r}) \equiv \begin{bmatrix} \varsigma_{it}^{yx}(\gamma_i^{yx}) \\ \varsigma_{it}^{s}(\gamma_i) \end{bmatrix} \text{ et } \mathbf{Z}_{it} \equiv \begin{bmatrix} \mathbf{Z}_{it}^{yx} & \mathbf{Z}_{it}^{yxs} \\ \mathbf{0} & \mathbf{Z}_{it}^{s} \end{bmatrix}$$

où la matrice $\mathbf{Z}_{it}^{yxs}$ dépend des variables $\mathbf{z}_{it}$

Ces termes résiduels permettent de définir la vraisemblance de $(\mathbf{y}_{it}, \mathbf{x}_{it}, \mathbf{s}_{it})$ conditionnelle en $(\mathbf{z}_{it}, \boldsymbol{\gamma}_i)$ calculée en $(\boldsymbol{\delta}, \boldsymbol{\Psi})$ sous la forme :

(67) $$\ell_{it}(\boldsymbol{\delta}, \boldsymbol{\Psi} \mid \boldsymbol{\gamma}_i) = \ell_{it}^{yx}(\boldsymbol{\delta}^{yx}, \boldsymbol{\Psi}^{yx} \mid \boldsymbol{\gamma}_i)\ell_{it}^{s}(\boldsymbol{\gamma}_i, \boldsymbol{\delta}, \boldsymbol{\Psi}^{s} \mid \boldsymbol{\gamma}_i)$$

avec :

(68) $$\ell_{it}^{yx}(\boldsymbol{\delta}^{yx}, \boldsymbol{\Psi}^{yx} \mid \boldsymbol{\gamma}_i) = \varphi\left(\boldsymbol{\varepsilon}_{it}^{yx}(\boldsymbol{\gamma}_i^{yx}, \boldsymbol{\delta}^{yx}); \boldsymbol{\Psi}^{yx}\right)$$

et :

(69) $$\ell_{it}^{s}(\boldsymbol{\delta}, \boldsymbol{\Psi}^{s} \mid \boldsymbol{\gamma}_i) = (\alpha_i^{s})^{1-G}\left(\prod_{g\in\mathcal{G}}(\alpha_{(g),i}^{s})^{1-K_{(g)}}\right)\left(\prod_{k\in\mathcal{K}} s_{k,it}^{-1}\right)\times\varphi\left(\boldsymbol{\varepsilon}_{it}^{s}(\boldsymbol{\gamma}_i, \boldsymbol{\delta}); \boldsymbol{\Psi}^{s}\right).$$

Finalement, la vraisemblance de l'observation des choix de production $(\mathbf{y}_i, \mathbf{x}_i, \mathbf{s}_i)$ conditionnelle en les variables exogènes calculée en $(\boldsymbol{\delta}, \boldsymbol{\Psi}, \boldsymbol{\mu}, \boldsymbol{\Omega})$ est donnée par l'intégrale :

(70) $$\ell_i(\boldsymbol{\delta}, \boldsymbol{\Psi}, \boldsymbol{\mu}, \boldsymbol{\Omega}) = \int\left(\prod_{t=1}^{T}\ell_{it}(\boldsymbol{\delta}, \boldsymbol{\Psi} \mid \boldsymbol{\gamma})\right)\varphi(\boldsymbol{\gamma} - \boldsymbol{\mu}; \boldsymbol{\Omega})d\boldsymbol{\gamma}\ .$$

L'estimateur du MV de $\boldsymbol{\theta}_0 \equiv (\boldsymbol{\delta}_0, \boldsymbol{\Psi}_0, \boldsymbol{\mu}_0, \boldsymbol{\Omega}_0)$ est alors défini comme la solution en $\boldsymbol{\theta}$ du programme de maximisation de la log-vraisemblance de l'échantillon considéré, avec:

(71) $$\hat{\boldsymbol{\theta}}_N \equiv (\hat{\boldsymbol{\delta}}_N, \hat{\boldsymbol{\Psi}}_N, \hat{\boldsymbol{\mu}}_N, \hat{\boldsymbol{\Omega}}_N) \equiv \arg\max_{(\boldsymbol{\delta}, \boldsymbol{\Psi}, \boldsymbol{\mu}, \boldsymbol{\Omega})}\sum_{i=1}^{N}\ln\ell_i(\boldsymbol{\delta}, \boldsymbol{\Psi}, \boldsymbol{\mu}, \boldsymbol{\Omega})$$

Comme tout estimateur du MV, cet estimateur est convergent et asymptotiquement efficace si le modèle considéré est correctement spécifié et si les conditions assurant l'identification de $\boldsymbol{\theta}_0$ sont assurées, ce que nous supposerons ici (ces conditions pouvant être, en partie, vérifiées empiriquement). Cet estimateur est asymptotiquement normal avec :

(72) $$\sqrt{N}(\hat{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_0) \xrightarrow[N \to +\infty]{L} \mathcal{N}(\mathbf{0}, \mathbf{J}_0^{-1}) \ ,$$

sa matrice-de variance-covariance asymptotique étant définie par l'inverse de la matrice d'information de Fisher. Cette dernière est ici définie comme l'opposé de l'espérance de la matrice Hessienne en $\boldsymbol{\theta} \equiv (\boldsymbol{\delta}, \boldsymbol{\Psi}, \boldsymbol{\mu}, \boldsymbol{\Omega})$ des log-vraisemblances individuelles calculée en $\boldsymbol{\theta}_0$. La propriété d'égalité des matrices d'information garantit ici que cette matrice d'information est égale à l'espérance du produit croisé des scores individuels calculés en $\boldsymbol{\theta}_0$:

(73) $$\mathbf{J}_0 = -E\left[\tfrac{\partial^2}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'} \ln \ell_i(\boldsymbol{\theta}_0)\right] = E\left[\tfrac{\partial}{\partial\boldsymbol{\theta}} \ln \ell_i(\boldsymbol{\theta}_0) \tfrac{\partial}{\partial\boldsymbol{\theta}'} \ln \ell_i(\boldsymbol{\theta}_0)\right] \ .$$

Les fonctions de vraisemblance $\ell_i(\boldsymbol{\theta})$ n'ayant pas de solution analytique en raison de l'intégrale multidimensionnelle nécessaire à son calcul, l'estimateur du MV de $\boldsymbol{\theta}_0$ ne peut être réalisée en pratique. Dans ce contexte, l'estimation de $\boldsymbol{\theta}_0$ est généralement abordée en calculant les termes $\ell_i(\boldsymbol{\theta})$ avec des méthodes d'intégration par simulations. L'intégration numérique de $\ell_i(\boldsymbol{\theta})$ n'est possible que lorsque le vecteur de paramètres aléatoires $\boldsymbol{\gamma}_i$ est de petite dimension, ce qui n'est pas le cas avec notre modèle de choix de production multiculture.

1.2.5. Modèles micro-économétriques de choix de production multiculture à paramètres aléatoires : les estimateurs du MVS

L'estimation par le MVS est la méthode privilégiée par les micro-économètres pour l'estimation de modèles paramétriques à paramètres aléatoires (voir, *e.g.*, Train, 2009 ; Keane, 2009). L'approche par le MVS est une extension directe de l'approche par le MV : il s'agit d'approcher les fonctions de log-vraisemblance des observations de l'échantillon par des approximations obtenues à l'aide de méthodes d'intégration par simulation (voir, *e.g.*, McFadden, 1989 ; Gourieroux et Monfort, 1990, 1993 ; Hajivassiliou et Ruud, 1994 ; Hajivassiliou, 2000).

Les paramètres $\boldsymbol{\gamma}_i$ étant supposées suivre la loi normale multivariée $\mathcal{N}(\boldsymbol{\mu}_0, \boldsymbol{\Omega}_0)$, ils peuvent s'écrire sous la forme $\boldsymbol{\gamma}_i = \boldsymbol{\mu}_0 + \mathbf{L}_0 \mathbf{u}_i$ où $\mathbf{u}_i$ suit la loi normale multivariée centrée réduite $\mathcal{N}(\mathbf{0}, \mathbf{I})$ et $\mathbf{L}_0$ est la racine de la décomposition de Cholesky de $\boldsymbol{\Omega}_0$ (avec $\boldsymbol{\Omega}_0 = \mathbf{L}_0 \mathbf{L}_0'$). Il est relativement aisé d'obtenir des tirages (pseudo-)aléatoires indépendants de la loi $\mathcal{N}(\mathbf{0}, \mathbf{I})$. Une approximation (non biaisée) par simulation de $\ell_i(\boldsymbol{\delta}, \boldsymbol{\Psi}, \boldsymbol{\mu}, \boldsymbol{\Omega})$ peut être obtenue à partir de *N* ensembles (pour $i = 1, ..., N$) de *R* de ces tirages, notés $\tilde{\mathbf{u}}_{i,r}$ pour $r = 1, ..., R$, pourvu que *R* soit suffisamment grand. La fonction de vraisemblance simulée :

(74) $$\tilde{\ell}_{i,R}(\boldsymbol{\delta}, \boldsymbol{\Psi}, \boldsymbol{\mu}, \mathbf{L}) = R^{-1} \sum_{r=1}^{R} \left(\prod_{t=1}^{T} \ell_{it}(\boldsymbol{\delta}, \boldsymbol{\Psi} \mid \boldsymbol{\mu} + \mathbf{L}\tilde{\mathbf{u}}_{i,r}) \right)$$

converge presque sûrement en *R* vers $\ell_i(\boldsymbol{\delta}, \boldsymbol{\Psi}, \boldsymbol{\mu}, \boldsymbol{\Omega})$.

L'estimateur du MVS de $(\boldsymbol{\delta}_0, \boldsymbol{\Psi}_0, \boldsymbol{\mu}_0, \boldsymbol{\Omega}_0)$ est donné par la solution en $(\boldsymbol{\delta}, \boldsymbol{\Psi}, \boldsymbol{\mu}, \mathbf{L})$ du problème de maximisation de la log-vraisemblance simulée de l'échantillon d'exploitations considéré, avec:

(75) $$\hat{\boldsymbol{\theta}}_{N,R}^{MVS} \equiv \arg\max_{(\boldsymbol{\delta},\boldsymbol{\Psi},\boldsymbol{\mu},\mathbf{L})} \sum_{i=1}^{N} \ln \tilde{\ell}_{i,R}(\boldsymbol{\delta}, \boldsymbol{\Psi}, \boldsymbol{\mu}, \mathbf{L}) \text{ et } \hat{\boldsymbol{\Omega}}_{N,R} = \hat{\mathbf{L}}_{N,R} \hat{\mathbf{L}}'_{N,R}.$$

L'estimateur du MVS $\hat{\boldsymbol{\theta}}_{N,R}^{MVS}$ a des propriétés asymptotiques équivalentes à celles de l'estimateur du MV correspondant si R croît plus vite que $\sqrt{N}$, *i.e.* si le nombre de tirages employés est suffisamment grand et est d'autant plus grand que l'échantillon employé est grand. Cette propriété s'applique à l'ensemble des estimateurs considérés ici.

Bien entendu, employer un grand nombre de simulations pour assurer les « bonnes propriétés » de l'estimateur du MVS rend le calcul de cet estimateur gourmand en temps et en puissance de calcul. Ceci dit, toutes les approches d'estimation de modèles à paramètres aléatoires se heurtent à ces problèmes de temps et de puissance de calcul.

L'estimation par le MVS s'avère cependant particulièrement complexe à mettre en œuvre pour notre modèle de choix multiculture à paramètres aléatoires, pour deux raisons principales (qui sont liées). Tout d'abord, la fonction de log-vraisemblance simulée de l'échantillon correspondant à notre modèle est particulièrement « affreuse ». La log-vraisemblance simulée :

(76) $$\sum_{i=1}^{N} \ln \tilde{\ell}_{i,R}(\boldsymbol{\delta}, \boldsymbol{\Psi}, \boldsymbol{\mu}, \mathbf{L}) = \sum_{i=1}^{N} \ln \left[R^{-1} \sum_{r=1}^{R} \left(\prod_{t=1}^{T} \ell_{it}(\boldsymbol{\delta}, \boldsymbol{\Psi} \mid \boldsymbol{\mu} + \mathbf{L}\tilde{\mathbf{u}}_{i,r}) \right) \right]$$

est une somme de N transformations logarithmiques de moyennes de R produits de T fonctions de densité de lois multivariées. Elle présente des non-linéarités très importantes en

les paramètres à optimiser. Ensuite, le problème de maximisation de cette log-vraisemblance simulée est de grande dimension. Il s'agit d'optimiser simultanément l'ensemble des paramètres du modèle, ceux de la loi noyau et ceux de la loi mélangeante. Le problème de maximisation de la log-vraisemblance simulée ne peut être décomposé en un ensemble de sous-problèmes plus petits car la log-transformation des vraisemblances individuelles ne permet pas de séparer l'ensemble des termes composant ces vraisemblances.

Ce problème de maximisation présente donc un réel défi en pratique, et ce même si de nombreux algorithmes très efficaces sont disponibles. La plupart de ces algorithmes ont des propriétés de convergence locale. Ils se « comportent correctement » lorsqu'ils sont initiés avec des valeurs des paramètres proches des valeurs optimales. Obtenir de bonnes valeurs initiales s'avère très complexe lorsque le nombre de paramètres considérés est très élevé comme dans les estimations considérées ici. En outre, la conception de ces algorithmes s'appuie sur des approximations quadratiques de la fonction à maximiser qui peuvent s'avérer peu précises dans le cas de la log-vraisemblance considérée ici.

Face à ces difficultés de mise en œuvre de l'approche par le MVS, nous nous sommes tournés vers des approches développées en statistique computationnelle. Comme nous l'expliquons dans le chapitre suivant, certaines de ces approches s'avèrent particulièrement adaptées à l'estimation de nos modèles de choix de production multiculture à paramètres aléatoires.

Chapitre 2

Modèles micro-économétriques de choix de production multiculture à paramètres aléatoires :

Estimation via les algorithmes MCEM et SAEM

Les statisticiens ont développé un ensemble impressionnant d'approches pour estimer des modèles paramétriques à paramètres aléatoires (voir, *e.g.*, MacLachlan et Krishnan, 2008 ; Lavielle, 2014). La plupart de ces approches sont conçues comme des extension de l'algorithme Espérance-Maximisation (*Expectation-Maximisation*, EM) initialement proposé par Dempster et al (1977). Cet algorithme exploite les propriétés des fonctions de log-vraisemblance pour décomposer le problème de leur maximisation en une suite de problèmes de maximisation qui, dans le cas des modèles à paramètres aléatoires, s'avère bien plus simples à résoudre que le problème initial. Il a en outre des propriétés de convergence globales intéressantes (Wu, 1983).

Les algorithmes EM itèrent deux étapes jusqu'à convergence numérique. L'étape E consiste à calculer l'espérance conditionnelle en les observations de la fonction de log-vraisemblance dite complète du modèle. L'étape M consiste à maximiser sur les paramètres à estimer l'espérance obtenue à l'étape E précédente. La littérature de statistique computationnelle propose deux grands types d'extension des algorithmes EM « de base » (voir, *e.g.*, MacLachlan et Krishnan, 2008 ; Lavielle, 2014).

Certaines extensions visent à simplifier l'étape M, à l'instar des algorithmes *Expectation-Conditional-Maximization* (ECM) proposés par Meng et Rubin (1993) que nous employons dans la suite. Elles visent en général à décomposer le problème de maximisation de l'étape M en un ensemble de problèmes de maximisation plus simples à résoudre. Elles s'avèrent particulièrement intéressantes dans le cas de l'estimation de modèles multivariés de grande dimension comme celui considéré ici.

D'autres extensions de l'algorithme EM visent à permettre de réaliser l'étape E dans les

cas où l'espérance à calculer est difficile à intégrer. Nos approches d'estimation s'appuient plus particulièrement sur deux de ces extensions. Dans le cas des modèles à paramètres aléatoires suivant une distribution continue, comme c'est le cas de nos modèles, l'espérance à calculer lors de l'étape E peut rarement être intégrée analytiquement ou numériquement. Weï et Tanner (1990) ont proposé des extensions des algorithmes EM « déterministes », les algorithmes *Monte-Carlo-EM* (MCEM) dans lesquelles l'espérance à calculer lors de l'étape E est intégrée par des méthodes simulées. Delyon et al (1999) combinent les idées à l'origine des algorithmes MCEM avec les méthodes d'approximation stochastique (*Stochastic Approximation*, SA) initialement proposées par Robins et Monro (1951) et Kiefer et Wolfowitz (1952) pour définir les algorithmes *Stochastic-Approximation-EM* (SAEM). Ces algorithmes SAEM apparaissent particulièrement efficaces pour exploiter les simulations des paramètres aléatoires du modèle et les propriétés originales des algorithmes EM.

De fait, la statistique computationnelle propose de très nombreux outils pour calculer des estimateurs analogues aux estimateurs du MVS. Les estimateurs calculés à partir d'algorithmes MCEM et les estimateurs du MVS convergent vers l'estimateur du MV correspondant dès lors que le nombre de simulations employées tend vers l'infini. En outre, les estimateurs calculés à partir d'algorithmes MCEM s'avèrent plus efficaces numériquement que les estimateurs du MVS correspondants pour des modèles à paramètres aléatoires (Jank et Booth, 2003). Enfin, Train (2009) rapporte que l'estimation des paramètres de covariance des paramètres aléatoires (gaussiens) des modèles de choix de discrets est difficile avec les estimateurs du MVS, ce qui conduit souvent les économètres à supposer que ces paramètres sont nuls en pratique. Comme cela sera vu dans la suite, l'estimation de

ces paramètres est, numériquement, très aisée avec les algorithmes MCEM. Ces estimateurs sont construits comme des covariances empiriques.

Curieusement, les extensions de l'algorithme EM sont largement ignorées par les micro-économètres qui leur préfèrent les estimateurs du MVS pour l'estimation de modèles paramétriques complexes. Ces derniers ont volontiers recours à l'algorithme EM pour l'estimation de modèles à paramètres aléatoires suivant une loi discrète – *i.e.* les modèles à classes latentes (voir, *e.g.*, Train (2009) pour les modèles de choix discret) – ou l'estimation de modèles à variables dépendantes limitées (Ruud, 1991). Mais ils n'utilisent pas ou très peu les algorithmes MCEM ou SAEM. A notre connaissance, Lichtenberg et Smith-Ramirez (2011) sont les seuls à avoir employé un algorithme MCEM pour estimer un modèle micro-économétrique (de choix de production agricole à changement de régime endogène en l'occurrence).[10]

Ce chapitre présente progressivement l'approche que nous proposons pour l'estimation de notre modèle de choix de production multiculture à paramètres aléatoires. Nous commençons

[10] Une anecdote illustre de manière évidente « l'étanchéité » des travaux des économètres vis-à-vis des travaux en statistique computationnelle sur l'estimation des modèles à paramètres aléatoires à loi mélangeante continue. K.E. Train a proposé en 2008 une procédure d'estimation – qu'il nomme estimateur récursif – pour l'estimation de modèles de choix discrets de forme Logit Multinomial à paramètres aléatoires (Train, 2008, 2009). Cette procédure est de fait un algorithme MCEM construit avec un simulateur d'échantillonnage préférentiel. Les seules références en statistique computationnelle citées par cet auteur renvoient aux articles fondateurs de Dempster et al (1977) et de Wu (1983).

par présenter l'algorithme EM qui correspond à notre modèle. Nous montrons ensuite qu'étendre cet algorithme en un algorithme ECM peut faciliter la mise en œuvre concrète de nos calculs d'estimation.

Nous abordons ensuite les questions soulevées par l'intégration des espérances qui définissent l'étape M de notre procédure d'estimation. Nous montrons alors que les méthodes de simulation par échantillonnage préférentiel s'avèrent particulièrement pratiques dans le cadre d'algorithmes MCEM puis d'algorithmes SAEM. Nous abordons ensuite brièvement les questions liées au tirage des simulations nécessaires à la mise en œuvre de nos procédures d'estimation, avec une attention particulière à la réduction de la variance des simulateurs des espérances que nous souhaitons approcher.

Enfin, nous présentons brièvement comment les calculs associés aux algorithmes MCEM et SAEM peuvent être parallélisés afin de réduire les temps de calculs, et en particulier d'exploiter les possibilités offertes par les ordinateurs à processeurs multiples avec le logiciel SAS/IML® qui nous avons employé pour l'ensemble de nos calculs empiriques.

2.1. Structure et convergence des algorithmes (G)EM

L'algorithme EM et ses extensions s'avèrent particulièrement utiles pour estimer des modèles paramétriques caractérisés par des problèmes de données manquantes. Dans le contexte de notre modèle de choix de production multiculture, les paramètres aléatoires $\{\boldsymbol{\gamma}_i : i = 1,...,N\}$ sont les données manquantes, les choix de production observés $\{(\mathbf{y}_i, \mathbf{x}_i, \mathbf{s}_i) : i = 1,...,N\}$ sont les données observées et l'ensemble des deux

$\{(\mathbf{y}_i, \mathbf{x}_i, \mathbf{s}_i, \boldsymbol{\gamma}_i) : i = 1, ..., N\}$ forme les données complètes. Les variables exogènes $\mathbf{z}_i$ de notre modèle conditionnent l'inférence statistique conduite ici.

Deux fonctions de log-vraisemblance sont considérées pour la mise en œuvre de l'algorithme EM. Dans notre cas, la première est la fonction de log-vraisemblance en $\boldsymbol{\theta} \equiv (\boldsymbol{\delta}, \boldsymbol{\Psi}, \boldsymbol{\mu}, \boldsymbol{\Omega})$ de l'échantillon considéré :

(1) $$\ln L_N(\boldsymbol{\theta}) \equiv \sum_{i=1}^{N} \ln \ell_i(\boldsymbol{\theta}) = \sum_{i=1}^{N} \ln \int \left(\prod_{t=1}^{T} \ell_{it}(\boldsymbol{\delta}, \boldsymbol{\Psi} \mid \boldsymbol{\gamma}) \right) \varphi(\boldsymbol{\gamma} - \boldsymbol{\mu}; \boldsymbol{\Omega}) d\boldsymbol{\gamma} .$$

L'estimateur du MV de $\boldsymbol{\theta}_0 \equiv (\boldsymbol{\delta}_0, \boldsymbol{\Psi}_0, \boldsymbol{\mu}_0, \boldsymbol{\Omega}_0)$ maximise en $\boldsymbol{\theta}$ cette fonction de log-vraisemblance.

La seconde est la log-vraisemblance « complète » de l'échantillon selon le modèle considéré. La vraisemblance « complète » de l'observation *i* est obtenue à partir de la fonction de densité de $(\mathbf{y}_i, \mathbf{x}_i, \mathbf{s}_i, \boldsymbol{\gamma}_i)$ conditionnelle en $\mathbf{z}_i$ qui est simplement donnée par :

(2) $$f(\mathbf{y}_i, \mathbf{x}_i, \mathbf{s}_i, \boldsymbol{\gamma}_i \mid \mathbf{z}_i) = \left(\prod_{t=1}^{T} \ell_{it}(\boldsymbol{\delta}_0, \boldsymbol{\Psi}_0 \mid \boldsymbol{\gamma}_i) \right) \varphi(\boldsymbol{\gamma}_i - \boldsymbol{\mu}_0; \boldsymbol{\Omega}_0)$$

pour notre modèle. Aussi la log-vraisemblance de l'observation *i* en $\boldsymbol{\theta}$ est donnée par :

(3) $$\ln \ell_i^C(\boldsymbol{\theta}; \boldsymbol{\gamma}_i) \equiv \sum_{t=1}^{T} \ln \ell_{it}(\boldsymbol{\delta}, \boldsymbol{\Psi} \mid \boldsymbol{\gamma}_i) + \ln \varphi(\boldsymbol{\gamma}_i - \boldsymbol{\mu}; \boldsymbol{\Omega})$$

et la log-vraisemblance de l'échantillon par :

(4) $$\ln L_N^C(\boldsymbol{\theta}) \equiv \sum_{i=1}^{N} \ln \ell_i^C(\boldsymbol{\theta}; \boldsymbol{\gamma}_i) = \sum_{i=1}^{N} \sum_{t=1}^{T} \ln \ell_{it}(\boldsymbol{\delta}, \boldsymbol{\Psi} \mid \boldsymbol{\gamma}_i) + \sum_{i=1}^{N} \ln \varphi(\boldsymbol{\gamma}_i - \boldsymbol{\mu}; \boldsymbol{\Omega}).$$

L'algorithme EM vise à calculer un estimateur du MV en itérant, jusqu'à convergence numérique, une étape E et une étape M. L'itération *m* de l'algorithme fournit une évaluation $\hat{\boldsymbol{\theta}}_{(m)} \equiv (\hat{\boldsymbol{\delta}}_{(m)}, \hat{\boldsymbol{\Psi}}_{(m)}, \hat{\boldsymbol{\mu}}_{(m)}, \hat{\boldsymbol{\Omega}}_{(m)})$ de $\boldsymbol{\theta}_0 \equiv (\boldsymbol{\delta}_0, \boldsymbol{\Psi}_0, \boldsymbol{\mu}_0, \boldsymbol{\Omega}_0)$

L'étape E de l'algorithme consiste à calculer l'espérance de la log-vraisemblance complète de l'échantillon en $\boldsymbol{\theta}$ sur la distribution des variables manquantes $\{\boldsymbol{\gamma}_i : i = 1,...,N\}$ conditionnelle en les variables observées $\{(\mathbf{y}_i, \mathbf{x}_i, \mathbf{s}_i) : i = 1,...,N\}$ telle que cette distribution peut être évaluée à la suite de l'étape M précédente. Par la loi de Bayes, la fonction de densité de $\boldsymbol{\gamma}_i$ conditionnelle en $(\mathbf{y}_i, \mathbf{x}_i, \mathbf{s}_i, \mathbf{z}_i)$ est donnée par :

(5) $$f(\boldsymbol{\gamma}_i \mid \mathbf{y}_i, \mathbf{x}_i, \mathbf{s}_i, \mathbf{z}_i) = \frac{f(\mathbf{y}_i, \mathbf{x}_i, \mathbf{s}_i, \boldsymbol{\gamma}_i \mid \mathbf{z}_i)}{f(\mathbf{y}_i, \mathbf{x}_i, \mathbf{s}_i \mid \mathbf{z}_i)}$$

et peut s'écrire à partir des fonctions de vraisemblance du modèle avec :

(6) $$f(\boldsymbol{\gamma}_i \mid \mathbf{y}_i, \mathbf{x}_i, \mathbf{s}_i, \mathbf{z}_i) = g_i(\boldsymbol{\theta}_0; \boldsymbol{\gamma}_i) \equiv \frac{\ell_i^C(\boldsymbol{\theta}_0; \boldsymbol{\gamma}_i)}{\ell_i(\boldsymbol{\theta}_0)}.$$

L'étape E de l'itération $m+1$ de l'algorithme consiste donc à calculer l'espérance conditionnelle des fonctions de log-vraisemblances « complètes » $\ln \ell_i^C(\boldsymbol{\theta}; \boldsymbol{\gamma}_i)$:

(7) $$E_{(m),i}[\ln \ell_i^C(\boldsymbol{\theta}; \boldsymbol{\gamma}_i)] = \int \ln \ell_i^C(\boldsymbol{\theta}; \boldsymbol{\gamma}) g_i(\hat{\boldsymbol{\theta}}_{(m)}; \boldsymbol{\gamma}) d\boldsymbol{\gamma}$$

pour $i = 1,...,N$.

Bien entendu, cette espérance n'a pas de forme analytique dans notre cas, pour deux raisons. Cette espérance est obtenue à partir d'une intégrale qui ne peut être résolue analytiquement. Cette espérance est également intégrée sur une fonction de densité, $g_i(\hat{\boldsymbol{\theta}}_{(m)}; \boldsymbol{\gamma})$, qui n'a pas de forme analytique. Ces questions seront abordées par suite afin de nous concentrer ici sur la description de l'algorithme EM.

L'étape M de l'itération $m+1$ de l'algorithme consiste ensuite à maximiser en $\boldsymbol{\theta}$ l'espérance conditionnelle en les observations de la vraisemblance complète de l'échantillon obtenue lors de l'étape E :

(8) $$Q_N(\boldsymbol{\theta};\hat{\boldsymbol{\theta}}_{(m)}) \equiv \sum_{i=1}^{N} E_{(m),i}[\ln \ell_i^C(\boldsymbol{\theta};\boldsymbol{\gamma}_i)] = \sum_{i=1}^{N} \int \ln \ell_i^C(\boldsymbol{\theta};\boldsymbol{\gamma}) g_i(\hat{\boldsymbol{\theta}}_{(m)};\boldsymbol{\gamma}) d\boldsymbol{\gamma}$$

pour obtenir une valeur actualisée de $\boldsymbol{\theta}_0$:

(9) $$\hat{\boldsymbol{\theta}}_{(m+1)} = \arg\max_{\boldsymbol{\theta}} Q_N(\boldsymbol{\theta};\hat{\boldsymbol{\theta}}_{(m)})$$

De fait, la fonction $Q_N(\boldsymbol{\theta};\hat{\boldsymbol{\theta}}_{(m)})$ joue le rôle de proxy de la fonction de log-vraisemblance $\ln L_N(\boldsymbol{\theta})$ dans laquelle la distribution des variables manquantes, les paramètres aléatoires $\boldsymbol{\gamma}_i$ ici, est approchée par leur distribution conditionnelle en les variables observées telle qu'elle est définie par le modèle et estimée au cours du processus d'estimation.

L'algorithme EM a des propriétés de convergence globale liées aux fait que la fonction $Q_N(\boldsymbol{\theta};\hat{\boldsymbol{\theta}}_{(m)})$ est croissante en *m* (Wu, 1983). Quelle que soit la valeur de $\boldsymbol{\theta}$ choisie pour initier l'algorithme, cet algorithme converge vers un maximum de $\ln L_N(\boldsymbol{\theta})$.

De fait, cette propriété est conservée lorsque le problème maximisation de l'étape M est remplacée par une simple recherche de valeur de $\boldsymbol{\theta}$ qui accroît $Q_N(\boldsymbol{\theta};\hat{\boldsymbol{\theta}}_{(m)})$, *i.e.* $\hat{\boldsymbol{\theta}}_{(m+1)}$ peut simplement être définie à partir de la condition :

(10) $$\hat{\boldsymbol{\theta}}_{(m+1)} \in \left\{\boldsymbol{\theta} : Q_N(\boldsymbol{\theta};\hat{\boldsymbol{\theta}}_{(m)}) > Q_N(\hat{\boldsymbol{\theta}}_{(m)};\hat{\boldsymbol{\theta}}_{(m)})\right\}.$$

Dempster et al (1977) ont nommé algorithmes EM Généralisés (*Generalized Expectation-Maximization*, GEM) les algorithmes reposant sur une telle étape M.

La vérification de la convergence numérique des algorithmes (G)EM repose sur des critères de distance entre les valeurs de $\hat{\boldsymbol{\theta}}_{(m)}$ et $\hat{\boldsymbol{\theta}}_{(m+1)}$ et/ou de $Q_N(\hat{\boldsymbol{\theta}}_{(m+1)};\hat{\boldsymbol{\theta}}_{(m+1)})$ et $Q_N(\hat{\boldsymbol{\theta}}_{(m)};\hat{\boldsymbol{\theta}}_{(m)})$.

L'examen des conditions de convergence de l'algorithme EM permet de faire le lien entre les estimateurs calculés avec cet algorithme et l'estimateur du MVS correspondant. Ces conditions sont également utiles pour interpréter la procédure de simulation que nous employons pour résoudre le problème posé par l'intégration des fonctions de log-vraisemblance approchées $Q_N(\boldsymbol{\theta};\hat{\boldsymbol{\theta}}_{(m)})$.

L'estimateur de $\boldsymbol{\theta}_0$ résultant de la convergence de l'algorithme EM est noté $\hat{\boldsymbol{\theta}}_N$. Il satisfait les conditions du premier ordre du problème de maximisation de l'étape M avec $\hat{\boldsymbol{\theta}}_N = \hat{\boldsymbol{\theta}}_{(m)}$ et $\hat{\boldsymbol{\theta}}_N = \boldsymbol{\theta}$:

(11) $$\sum_{i=1}^{N} \int \frac{\partial}{\partial \boldsymbol{\theta}} \ln \ell_i^C(\hat{\boldsymbol{\theta}}_N;\boldsymbol{\gamma}) g_i(\hat{\boldsymbol{\theta}}_N;\boldsymbol{\gamma}) d\boldsymbol{\gamma} = \mathbf{0}$$

Etant donnée la forme de la fonction de densité conditionnelle :

(12) $$g_i(\boldsymbol{\theta};\boldsymbol{\gamma}) \equiv \frac{\ell_i^C(\boldsymbol{\theta};\boldsymbol{\gamma})}{\ell_i(\boldsymbol{\theta})} = \frac{\ell_{it}(\boldsymbol{\delta},\boldsymbol{\Psi} \mid \boldsymbol{\gamma})}{\int \ell_{it}(\boldsymbol{\delta},\boldsymbol{\Psi} \mid \boldsymbol{\gamma}) \varphi(\boldsymbol{\gamma}-\boldsymbol{\mu};\boldsymbol{\Omega}) d\boldsymbol{\gamma}} \varphi(\boldsymbol{\gamma}-\boldsymbol{\mu};\boldsymbol{\Omega}),$$

Ces conditions peuvent être réécrites sous la forme :

(13) $$\sum_{i=1}^{N} \frac{\partial}{\partial \boldsymbol{\theta}} \ln \ell_i(\hat{\boldsymbol{\theta}}_N) = \mathbf{0} \ .$$

Cette expression montre que l'algorithme EM converge vers une racine des conditions du

premier ordre caractérisant l'estimateur du MV.

2.2. Algorithme EM et estimation des modèles à paramètres aléatoires

Dans le cas des modèles à paramètres aléatoires, l'avantage des algorithmes EM et GEM provient de la simplicité de l'étape M. Les fonctions de log-vraisemblance complètes $\ln \ell_i^C(\boldsymbol{\theta};\boldsymbol{\gamma}_i)$ et $\ln L_N^C(\boldsymbol{\theta})$ séparent de manière additive les fonctions de log-vraisemblance correspondant aux variables observées conditionnelles en les paramètres aléatoires $\ln \ell_{it}(\boldsymbol{\delta},\boldsymbol{\Psi} \mid \boldsymbol{\gamma}_i)$ d'une part, et aux paramètres aléatoires eux-mêmes $\ln \varphi(\boldsymbol{\gamma}_i - \boldsymbol{\mu};\boldsymbol{\Omega})$ d'autre part. Cette décomposition additive des vraisemblances complètes permet une décomposition équivalente de de l'étape M. En effet, le problème de maximisation :

(14) $$\hat{\boldsymbol{\theta}}_{(m+1)} = \arg\max_{\boldsymbol{\theta}} \sum_{i=1}^{N} \int \ln \ell_i^C(\boldsymbol{\theta};\boldsymbol{\gamma}) g_i(\hat{\boldsymbol{\theta}}_{(m)};\boldsymbol{\gamma}) d\boldsymbol{\gamma}$$

peut être décomposé en deux sous-étapes M équivalentes :

(15) $$(\hat{\boldsymbol{\delta}}_{(m+1)}, \hat{\boldsymbol{\Psi}}_{(m+1)}) = \arg\max_{(\boldsymbol{\delta},\boldsymbol{\Psi})} \sum_{i=1}^{N} \sum_{t=1}^{T} \int \ln \ell_{it}(\boldsymbol{\delta},\boldsymbol{\Psi} \mid \boldsymbol{\gamma}) g_i(\hat{\boldsymbol{\theta}}_{(m)};\boldsymbol{\gamma}) d\boldsymbol{\gamma}$$

et :

(16) $$(\hat{\boldsymbol{\mu}}_{(m+1)}, \hat{\boldsymbol{\Omega}}_{(m+1)}) = \arg\max_{(\boldsymbol{\mu},\boldsymbol{\Omega})} \sum_{i=1}^{N} \int \ln \varphi(\boldsymbol{\gamma} - \boldsymbol{\mu};\boldsymbol{\Omega}) g_i(\hat{\boldsymbol{\theta}}_{(m)};\boldsymbol{\gamma}) d\boldsymbol{\gamma}$$

Il sera vu dans la suite que les termes $(\hat{\boldsymbol{\mu}}_{(m+1)}, \hat{\boldsymbol{\Omega}}_{(m+1)})$ sont très faciles à calculer dans les cas considérés ici. Sans être compliquée à mettre en œuvre, la procédure d'actualisation des paramètres $(\boldsymbol{\delta},\boldsymbol{\Psi})$ décrite par le problème de maximisation de l'équation (15) est plus gourmande en temps de calcul. Cette sous-étape M peut être significativement simplifiée sans affecter les propriétés de convergence de l'algorithme considéré. Il suffit d'exploiter la structure multivariée du modèle considéré et les propriétés des algorithmes ECM (et donc

GEM).

2.3. Algorithmes ECM et GEM et estimations des modèles multivariés

L'algorithme EM et ses extensions s'avèrent particulièrement utiles pour estimer des modèles à paramètres. La structure multivariée de notre modèle de choix de production agricole multiculture peut également être exploitée pour simplifier l'actualisation de l'estimation du sous-vecteur de paramètres $(\boldsymbol{\delta}_0, \boldsymbol{\Psi}_0)$. En effet, la log-vraisemblance en $(\boldsymbol{\delta}, \boldsymbol{\Psi})$ des choix de production observés conditionnelle en les paramètres aléatoires, $\ell_{it}(\boldsymbol{\delta}, \boldsymbol{\Psi} \mid \boldsymbol{\gamma})$, se décompose en la somme de la log-vraisemblance en $(\boldsymbol{\delta}^{yx}, \boldsymbol{\Psi}^{yx})$ des niveaux de rendement et d'utilisation d'intrants et de la log-vraisemblance en $(\boldsymbol{\delta}, \boldsymbol{\Psi}^s) = (\boldsymbol{\delta}^{yx}, \boldsymbol{\delta}^s, \boldsymbol{\Psi}^s)$ des choix d'assolement :

(17) $$\ln \ell_{it}(\boldsymbol{\delta}, \boldsymbol{\Psi} \mid \boldsymbol{\gamma}) = \ln \ell_{it}^{yx}(\boldsymbol{\delta}^{yx}, \boldsymbol{\Psi}^{yx} \mid \boldsymbol{\gamma}) + \ln \ell_{it}^{s}(\boldsymbol{\delta}, \boldsymbol{\Psi}^s \mid \boldsymbol{\gamma}).$$

Le fait que $\ln \ell_{it}^{s}(\boldsymbol{\delta}, \boldsymbol{\Psi}^s \mid \boldsymbol{\gamma})$ et $\ln \ell_{it}^{yx}(\boldsymbol{\delta}^{yx}, \boldsymbol{\Psi}^{yx} \mid \boldsymbol{\gamma})$ dépendent toutes deux de $\boldsymbol{\delta}^{yx}$ ne permet pas une décomposition simple de l'étape de réactualisation de $(\boldsymbol{\delta}, \boldsymbol{\Psi})$ en deux procédures de maximisation séparées.

En outre, les log-vraisemblances individuelles $\ln \ell_{it}^{s}(\boldsymbol{\gamma}_i, \boldsymbol{\delta}, \boldsymbol{\Psi}^s \mid \boldsymbol{\gamma})$ et $\ln \ell_{it}^{yx}(\boldsymbol{\delta}^{yx}, \boldsymbol{\Psi}^{yx} \mid \boldsymbol{\gamma})$ sont celles de systèmes de régressions empilées. Les estimateurs des paramètres de ces modèles sont généralement calculés de manière itérative selon un algorithme de Gauss-Seidel qui alterne deux étapes jusqu'à convergence numérique. Dans le cas considéré ces étapes sont (*a*) la maximisation de la log-vraisemblance en $\boldsymbol{\delta}$ à $\boldsymbol{\Psi}$ donné et (*b*) la maximisation de la log-vraisemblance en $\boldsymbol{\Psi}$ à $\boldsymbol{\delta}$ donné (cette procédure de maximisation se réduisant à un simple

calcul de matrice de variance-covariance empirique).

De fait, un modèle sans variables de contrôle (ou avec des coefficients des variables de contrôle aléatoires) aurait été beaucoup plus simple à estimer. Les étapes M des algorithmes présentés ici auraient consisté en des calculs de moyennes, variances et covariances empiriques pondérées.

Néanmoins, l'utilisation des propriétés des algorithmes ECM de Meng et Rubin (1993) permet de simplifier l'actualisation de $(\boldsymbol{\delta}, \boldsymbol{\Psi})$ en deux étapes simples de maximisation « conditionnelles » successives. Nous avons choisi les étapes de maximisation conditionnelles suivantes pour remplacer la sous-étape M décrite par l'équation (82) :

$$(18) \qquad \begin{cases} \hat{\boldsymbol{\delta}}_{(m+1)} = \arg\max_{\boldsymbol{\delta}} \sum_{i=1}^{N} \sum_{t=1}^{T} \int \ln \ell_{it}(\boldsymbol{\delta}, \hat{\boldsymbol{\Psi}}_{(m)} \mid \boldsymbol{\gamma}) g_i(\hat{\boldsymbol{\theta}}_{(m)}; \boldsymbol{\gamma}) d\boldsymbol{\gamma} \\ \hat{\boldsymbol{\Psi}}^{yx}_{(m+1)} = \arg\max_{\boldsymbol{\Psi}^{yx}} \sum_{i=1}^{N} \sum_{t=1}^{T} \int \ln \ell^{yx}_{it}(\hat{\boldsymbol{\delta}}^{yx}_{(m+1)}, \boldsymbol{\Psi}^{yx} \mid \boldsymbol{\gamma}) g_i(\hat{\boldsymbol{\theta}}_{(m)}; \boldsymbol{\gamma}) d\boldsymbol{\gamma} \\ \hat{\boldsymbol{\Psi}}^{s}_{(m+1)} = \arg\max_{(\boldsymbol{\delta}^s, \boldsymbol{\Psi}^s)} \sum_{i=1}^{N} \sum_{t=1}^{T} \int \ln \ell^{s}_{it}(\boldsymbol{\gamma}_i, \hat{\boldsymbol{\delta}}_{(m+1)}, \boldsymbol{\Psi}^s \mid \boldsymbol{\gamma}) g_i(\hat{\boldsymbol{\theta}}_{(m)}; \boldsymbol{\gamma}) d\boldsymbol{\gamma} \end{cases}$$

Cette suite d'étapes M conditionnelles correspond à une itération de l'algorithme de Gauss-Seidel habituellement employé pour l'estimation des paramètres de systèmes de régressions empilées.

Lorsqu'elle est combinée avec la procédure de calcul de $(\hat{\boldsymbol{\mu}}_{(m+1)}, \hat{\boldsymbol{\Omega}}_{(m+1)})$ cette suite de problèmes de maximisation conduit à une valeur de $(\hat{\boldsymbol{\delta}}_{(m+1)}, \hat{\boldsymbol{\Psi}}_{(m+1)})$ vérifiant la condition $Q_N(\hat{\boldsymbol{\theta}}_{(m+1)}; \hat{\boldsymbol{\theta}}_{(m)}) > Q_N(\hat{\boldsymbol{\theta}}_{(m)}; \hat{\boldsymbol{\theta}}_{(m)})$. L'utilisation de ces étapes M conditionnelles transforme donc l'algorithme EM considéré initialement en un algorithme GEM (Meng et Rubin, 1993).

2.4. Algorithmes MCEM et simulations par échantillonnage préférentiel

Jusqu'à présent nous avons ignoré un problème essentiel lié à la mise en œuvre concrète d'algorithmes (G)EM pour l'estimation des paramètres de modèle de choix de production multiculture à paramètres aléatoires. Nous avons laissé de côté le problème posé par l'intégration de la fonction $Q_N(\boldsymbol{\theta};\hat{\boldsymbol{\theta}}_{(m)}) = \sum_{i=1}^{N} E_{(m),i}[\ln \ell_i^C(\boldsymbol{\theta};\boldsymbol{\gamma}_i)]$ pour nous concentrer sur les décompositions de l'étape M permises par la structure de notre modèle. Or ce problème d'intégration est difficile. Il s'agit ici d'intégrer les log-vraisemblances individuelles complètes $\ln \ell_i^C(\boldsymbol{\theta};\boldsymbol{\gamma})$ sur la distribution des paramètres aléatoires conditionnellement en les variables observées $g_i(\hat{\boldsymbol{\theta}}_{(m)};\boldsymbol{\gamma})$, *i.e.* de calculer les termes :

(19) $$E_{(m),i}[\ln \ell_i^C(\boldsymbol{\theta};\boldsymbol{\gamma}_i)] = \int \ln \ell_i^C(\boldsymbol{\theta};\boldsymbol{\gamma}) g_i(\hat{\boldsymbol{\theta}}_{(m)};\boldsymbol{\gamma}) d\boldsymbol{\gamma}$$

pour $i = 1,...,M$ à chaque itération de l'algorithme.

L'expression de l'intégrale $\int \ln \ell_i^C(\boldsymbol{\theta};\boldsymbol{\gamma}) g_i(\hat{\boldsymbol{\theta}}_{(m)};\boldsymbol{\gamma}) d\boldsymbol{\gamma}$ suggère immédiatement une approche par simulation des paramètres aléatoires $\boldsymbol{\gamma}_i$ suivant la densité $g_i(\hat{\boldsymbol{\theta}}_{(m)};\boldsymbol{\gamma})$. C'est la logique de construction des algorithmes MCEM proposés par Weï et Tanner (1990). Obtenir directement des tirages des $\boldsymbol{\gamma}_i$ suivant la densité $g_i(\hat{\boldsymbol{\theta}}_{(m)};\boldsymbol{\gamma})$ est possible en utilisant des méthodes de Monte Carlo à chaînes de Markov, telles les méthodes d'échantillonnage de

Gibbs (McCulloch, 1994) ou de Metropolis-Hastings (McCulloch, 1997). Ces méthodes sont cependant très coûteuses en puissance et/ou temps de calcul.

Booth et Hobert (1999) proposent d'employer des méthodes d'échantillonnage préférentiel (*Importance Sampling*, IS) afin de réduire le temps de calcul lié au tirage des valeurs simulées de $\boldsymbol{\gamma}_i$.[11] Nous avons adopté cette approche car (*a*) elle est simple à mettre en œuvre et (*b*) nous disposons d'une densité de proposition « naturelle ». La forme de la densité $g_i(\hat{\boldsymbol{\theta}}_{(m)};\boldsymbol{\gamma})$ s'écrivant sous la forme :

$$(20)\ g_i(\hat{\boldsymbol{\theta}}_{(m)};\boldsymbol{\gamma}) \equiv \frac{\ell_i^C(\hat{\boldsymbol{\theta}}_{(m)};\boldsymbol{\gamma})}{\ell_i(\hat{\boldsymbol{\theta}}_{(m)})} = \frac{\prod_{t=1}^T \ell_{it}(\hat{\boldsymbol{\delta}}_{(m)},\hat{\boldsymbol{\Psi}}_{(m)} \mid \boldsymbol{\gamma})}{\int \prod_{t=1}^T \ell_{it}(\hat{\boldsymbol{\delta}}_{(m)},\hat{\boldsymbol{\Psi}}_{(m)} \mid \boldsymbol{\gamma})\varphi(\boldsymbol{\gamma}-\hat{\boldsymbol{\mu}}_{(m)};\hat{\boldsymbol{\Omega}}_{(m)})d\boldsymbol{\gamma}}\varphi(\boldsymbol{\gamma}-\hat{\boldsymbol{\mu}}_{(m)};\hat{\boldsymbol{\Omega}}_{(m)}),$$

elle suggère d'intégrer $\int \ln \ell_i^C(\boldsymbol{\theta};\boldsymbol{\gamma})g_i(\hat{\boldsymbol{\theta}}_{(m)};\boldsymbol{\gamma})d\boldsymbol{\gamma}$ en employant $\varphi(\boldsymbol{\gamma}-\hat{\boldsymbol{\mu}}_{(m)};\hat{\boldsymbol{\Omega}}_{(m)})$, *i.e.* la densité de la loi $\mathcal{N}(\hat{\boldsymbol{\mu}}_{(m)},\hat{\boldsymbol{\Omega}}_{(m)})$, comme densité de proposition et une fonction $w_i(\hat{\boldsymbol{\theta}}_{(m)};\boldsymbol{\gamma})$ proportionnelle à :

$$(21)\qquad \frac{\prod_{t=1}^T \ell_{it}(\hat{\boldsymbol{\delta}}_{(m)},\hat{\boldsymbol{\Psi}}_{(m)} \mid \boldsymbol{\gamma})}{\int \prod_{t=1}^T \ell_{it}(\hat{\boldsymbol{\delta}}_{(m)},\hat{\boldsymbol{\Psi}}_{(m)} \mid \boldsymbol{\gamma})\varphi(\boldsymbol{\gamma}-\hat{\boldsymbol{\mu}}_{(m)};\hat{\boldsymbol{\Omega}}_{(m)})d\boldsymbol{\gamma}}$$

et sommant à 1 sur $r = 1,...,R$ comme fonction d'importance (ou de ratio de vraisemblance) normalisée. Il suffit alors (*a*) de réaliser *R* tirages, les $\tilde{\boldsymbol{\gamma}}_{i,r}$ pour $r = 1,...,R$, de $\boldsymbol{\gamma}_i$ suivant la

[11] Ces dernières pouvant être combinées avec des méthodes à chaînes de Markov (Levine et Casella, 2001).

distribution $\mathcal{N}(\hat{\boldsymbol{\mu}}_{(m)}, \hat{\boldsymbol{\Omega}}_{(m)})$, (*b*) d'intégrer le dénominateur de la fonction d'importance $w_i(\hat{\boldsymbol{\theta}}_{(m)}; \boldsymbol{\gamma})$ par simulation directe pour obtenir un simulateur convergent (mais baisé) de cette fonction :

(22) $$\tilde{w}_{i,r,(m)} \equiv \frac{\prod_{t=1}^{T} \ell_{it}(\hat{\boldsymbol{\delta}}_{(m)}, \hat{\boldsymbol{\Psi}}_{(m)} \mid \tilde{\boldsymbol{\gamma}}_{i,r})}{\sum_{r=1}^{R} \prod_{t=1}^{T} \ell_{it}(\hat{\boldsymbol{\delta}}_{(m)}, \hat{\boldsymbol{\Psi}}_{(m)} \mid \tilde{\boldsymbol{\gamma}}_{i,r})}$$

pour finalement (*c*) calculer l'approximation par simulations de $\int \ln \ell_i^C(\boldsymbol{\theta}; \boldsymbol{\gamma}) g_i(\hat{\boldsymbol{\theta}}_{(m)}; \boldsymbol{\gamma}) d\boldsymbol{\gamma}$ avec :

(23) $$\tilde{E}_{(m),i,R}[\ln \ell_i^C(\boldsymbol{\theta}; \boldsymbol{\gamma}_i)] = \sum_{r=1}^{R} \tilde{w}_{i,r,(m)} \ln \ell_i^C(\boldsymbol{\theta}; \tilde{\boldsymbol{\gamma}}_{i,r}) .$$

Le simulateur IS $\tilde{E}_{(m),i,R}[\ln \ell_i^C(\boldsymbol{\theta}; \boldsymbol{\gamma}_i)]$ converge presque sûrement en *R* vers $E_{(m),i}[\ln \ell_i^C(\boldsymbol{\theta}; \boldsymbol{\gamma}_i)]$. Ceci garantit que $\tilde{E}_{(m),i,R}[\ln \ell_i^C(\boldsymbol{\theta}; \boldsymbol{\gamma}_i)]$ est une bonne approximation de $E_{(m),i}[\ln \ell_i^C(\boldsymbol{\theta}; \boldsymbol{\gamma}_i)]$ dès lors que *R* est « suffisamment » élevé. Cependant, choisir un nombre de tirages très élevé augmente très rapidement les temps de calculs. Plusieurs méthodes, dites de réduction de la variance peuvent être employées pour améliorer l'efficacité des simulateurs IS.

Choisir une autre densité de proposition que $\varphi(\boldsymbol{\gamma} - \hat{\boldsymbol{\mu}}_{(m)}; \hat{\boldsymbol{\Omega}}_{(m)})$ pourrait améliorer l'efficacité de l'intégration de $E_{(m),i}[\ln \ell_i^C(\boldsymbol{\theta}; \boldsymbol{\gamma}_i)]$ (voir, *e.g.*, Cappé et al, 2008 ; Hoogerheide *et al*, 2012b ; Douc *et al*, 2007). Mais là encore, la mise en œuvre en œuvre des approches permettant de choisir des densités de proposition efficaces sont gourmandes en temps de calcul (et de codage).

Dans tous les cas, il est nécessaire que les tirages des $\tilde{\boldsymbol{\gamma}}_{i,r}$ soient efficaces. Nous avons

employé des séquences de Halton (1960), pour leur capacité à couvrir l'intervalle unitaire, que nous avons ensuite transformées en tirages de lois normales $\mathcal{N}(0,1)$. Ces tirages ont été employés tout au long de l'algorithme de sorte à garantir au mieux la convergence de l'algorithme MCEM. Ces tirages de lois normales ont ensuite été combinés à chaque étape M pour obtenir des tirages de lois normales multivariées $\mathcal{N}(\hat{\boldsymbol{\mu}}_{(m)}, \hat{\boldsymbol{\Omega}}_{(m)})$. Les séquences de Halton employées ont été brouillées afin de garantir au mieux leur indépendance (Bhat, 2003). Ce brouillage est d'autant plus important dans notre cas que nous considérons des vecteurs de paramètres aléatoires de dimension conséquente et que nous souhaitions maintenir des choix de R raisonnables. Nous avons transformé les séquences de Halton initiales en séquences de Halton modifiées (Atanassov, 2004) à partir de l'algorithme proposé par Atanassov et Durchova (2002) pour obtenir un brouillage « intelligent ».

Nous avons également tenté de réduire les problèmes de dégénérescence de l'algorithme MCEM en augmentant l'efficacité du simulateur IS $\tilde{E}_{(m),i,R}[\ln \ell_i^C(\boldsymbol{\theta};\boldsymbol{\gamma}_i)]$ à partir d'une méthode d'échantillonnage préférentiel avec ré-échantillonnage préférentiel (*Importance Sampling Resampling*, SIR). La méthode SIR vise à augmenter l'occurrence des valeurs des $\tilde{\boldsymbol{\gamma}}_{i,r}$ avec des poids $\tilde{w}_{i,r,(m)}$ élevés de sorte à éviter qu'une trop grande part des poids $\tilde{w}_{i,r,(m)}$ deviennent pratiquement nuls à mesure que l'algorithme progresse, conduisant ainsi à sa dégénérescence. Il s'agit ici de ré-échantillonner les valeurs simulées de $\tilde{\boldsymbol{\gamma}}_{i,r}$ en effectuant R tirages indépendants avec remise dans la collection $\{\tilde{\boldsymbol{\gamma}}_{i,r} : r = 1,...,R\}$, ces tirages étant définis de telle sorte que la probabilité de tirer de $\tilde{\boldsymbol{\gamma}}_{i,r}$ soit égale à $\tilde{w}_{i,r,(m)}$. Ceci dit, ces techniques de

ré-échantillonnage ne doivent être employées que lorsque le risque de dégénérescence de l'algorithme est élevé. Ce risque est élevé lorsque l'indicateur ESS (*Effective Sampling Size*) donné ici par $\left(\sum_{r=1}^{R}(\tilde{w}_{i,r,(m)})^2\right)^{-1}$ est faible. Nous avons décidé de ne mettre en œuvre des techniques de ré-échantillonnage que lorsque l'indicateur ESS descend sous un certain seuil.

2.5. Mise en œuvre de l'algorithme MCEM avec échantillonnage préférentiel

L'algorithme MCEM obtenu en combinant l'étape E décrite ici et l'étape M décrite dans les sections précédentes est particulièrement simple à mettre en œuvre.

L'étape E de l'itération $m+1$ consiste simplement à calculer les termes $\tilde{w}_{i,r,(m)}$ pour $r=1,...,R$ et $i=1,...,N$ en employant N séries de R tirages de $\boldsymbol{\gamma}_i$ suivant la distribution $\mathcal{N}(\hat{\boldsymbol{\mu}}_{(m)},\hat{\boldsymbol{\Omega}}_{(m)})$.

L'étape M de l'itération $m+1$ consiste alors à résoudre le problème de maximisation définissant la valeur réactualisée de l'estimateur de $\boldsymbol{\theta}_0$, *i.e.* à calculer :

(24) $$\hat{\boldsymbol{\theta}}_{(m+1)} = \arg\max_{\boldsymbol{\theta}} \sum_{i=1}^{N}\sum_{r=1}^{R} \tilde{w}_{i,r,(m)} \ln \ell_i^C(\boldsymbol{\theta};\tilde{\boldsymbol{\gamma}}_{i,r}) .$$

Compte-tenu de la structure de notre modèle de choix de production multiculture à paramètres aléatoires, le problème de maximisation précédent peut être simplifié en une série de problèmes dont la solution permet de réactualiser les différents éléments du vecteur de paramètres $\boldsymbol{\theta}$, avec :

$$
(25) \quad \begin{cases} (\hat{\boldsymbol{\mu}}_{(m+1)}, \hat{\boldsymbol{\Omega}}_{(m+1)}) = \arg\max_{(\boldsymbol{\mu},\boldsymbol{\Omega})} \sum_{i=1}^{N} \sum_{r=1}^{R} \tilde{w}_{i,r,(m)} \ln \varphi(\tilde{\boldsymbol{\gamma}}_{i,r} - \boldsymbol{\mu}; \boldsymbol{\Omega}) \\ \hat{\boldsymbol{\delta}}_{(m+1)} = \arg\max_{\boldsymbol{\delta}} \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{r=1}^{R} \tilde{w}_{i,r,(m)} \ln \ell_{it}(\boldsymbol{\delta}, \hat{\boldsymbol{\Psi}}_{(m)} \mid \tilde{\boldsymbol{\gamma}}_{i,r}) \\ \hat{\boldsymbol{\Psi}}^{yx}_{(m+1)} = \arg\max_{\boldsymbol{\Psi}^{yx}} \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{r=1}^{R} \tilde{w}_{i,r,(m)} \ln \ell^{yx}_{it}(\hat{\boldsymbol{\delta}}^{yx}_{(m+1)}, \boldsymbol{\Psi}^{yx} \mid \tilde{\boldsymbol{\gamma}}_{i,r}) \\ \hat{\boldsymbol{\Psi}}^{s}_{(m+1)} = \arg\max_{\boldsymbol{\Psi}^{s}} \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{r=1}^{R} \tilde{w}_{i,r,(m)} \ln \ell^{s}_{it}(\hat{\boldsymbol{\delta}}_{(m+1)}, \boldsymbol{\Psi}^{s} \mid \tilde{\boldsymbol{\gamma}}_{i,r}) \end{cases}
$$

La définition du modèle implique que le terme $\hat{\boldsymbol{\theta}}_{(m)}$ peut s'écrire à partir de formes analytiques simples.

L'hypothèse de normalité des paramètres aléatoires $\boldsymbol{\gamma}_i$ implique que les termes $\hat{\boldsymbol{\mu}}_{(m+1)}$ et $\hat{\boldsymbol{\Omega}}_{(m+1)}$ sont définis sous forme analytique, avec :

$$
(26) \quad \hat{\boldsymbol{\mu}}_{(m+1)} = N^{-1} \sum_{i=1}^{N} \sum_{r=1}^{R} \tilde{w}_{i,r,(m)} \tilde{\boldsymbol{\gamma}}_{i,r}
$$

et :

$$
(27) \quad \hat{\boldsymbol{\Omega}}_{(m+1)} = N^{-1} \sum_{i=1}^{N} \sum_{r=1}^{R} \tilde{w}_{i,r,(m)} \tilde{\boldsymbol{\gamma}}_{i,r} \tilde{\boldsymbol{\gamma}}'_{i,r} - \hat{\boldsymbol{\mu}}_{(m+1)} \hat{\boldsymbol{\mu}}'_{(m+1)}
$$

i.e. sous la forme d'une moyenne empirique pondérée pour $\hat{\boldsymbol{\mu}}_{(m+1)}$ et d'une matrice de variance-covariance empirique pour $\hat{\boldsymbol{\Omega}}_{(m+1)}$.

De manière similaire, l'hypothèse de normalité des termes d'erreur de ces modèles implique que les termes $\hat{\boldsymbol{\Psi}}^{yx}_{(m+1)}$ et $\hat{\boldsymbol{\Psi}}^{s}_{(m+1)}$ sont définis sous forme analytique, avec :

$$
(28) \quad \hat{\boldsymbol{\Psi}}^{yx}_{(m+1)} = (NT)^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{r=1}^{R} \tilde{w}_{i,r,(m)} \boldsymbol{\varepsilon}^{yx}_{it}(\tilde{\boldsymbol{\gamma}}_{i,r}, \hat{\boldsymbol{\delta}}^{yx}_{(m+1)}) \boldsymbol{\varepsilon}^{yx}_{it}(\tilde{\boldsymbol{\gamma}}_{i,r}, \hat{\boldsymbol{\delta}}^{yx}_{(m+1)})'
$$

et :

$$
(29) \quad \hat{\boldsymbol{\Psi}}^{s}_{(m+1)} = (NT)^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{r=1}^{R} \tilde{w}_{i,r,(m)} \boldsymbol{\varepsilon}^{s}_{it}(\tilde{\boldsymbol{\gamma}}_{i,r}, \hat{\boldsymbol{\delta}}_{(m+1)}) \boldsymbol{\varepsilon}^{yx}_{it}(\tilde{\boldsymbol{\gamma}}_{i,r}, \hat{\boldsymbol{\delta}}_{(m+1)})'
$$

i.e. sous forme de matrices de variance-covariance empirique dans les deux cas.

Les formes fonctionnelles du modèle impliquant que le termes résiduels $\boldsymbol{\varepsilon}_{it}(\tilde{\boldsymbol{\gamma}}_{i,r},\boldsymbol{\delta})$ du modèle peuvent s'écrire sous la forme $\boldsymbol{\varepsilon}_{it}(\tilde{\boldsymbol{\gamma}}_{i,r},\boldsymbol{\delta}) = \varsigma_{it}(\tilde{\boldsymbol{\gamma}}_{i,r}) - \mathbf{Z}'_{it}\boldsymbol{\delta}$, les hypothèses de normalité des termes d'erreur $\boldsymbol{\varepsilon}_{it}$ et d'exogénéité des termes $\mathbf{Z}_{it}$ impliquent que le terme $\hat{\boldsymbol{\delta}}_{(m+1)}$ est défini sous forme analytique, avec :

(30) $$\hat{\boldsymbol{\delta}}_{(m+1)} = \left(\sum_{i=1}^{N}\sum_{t=1}^{T}\mathbf{Z}_{it}\hat{\boldsymbol{\Psi}}_{(m)}^{-1}\mathbf{Z}'_{it}\right)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\mathbf{Z}_{it}\hat{\boldsymbol{\Psi}}_{(m)}^{-1}\sum_{r=1}^{R}\tilde{w}_{i,r,(m)}\varsigma_{it}(\tilde{\boldsymbol{\gamma}}_{i,r})$$

i.e., sous la forme d'un estimateur des Moindres Carrés Quasi-Généralisé (MCQG).

Malheureusement, cet algorithme s'est avéré peu performant en pratique. Il converge difficilement quel que soit le nombre de simulations employé. Ces résultats négatifs nous ont amené à l'utilisation des algorithmes SAEM proposés par Delyon *et al* (1999). Dans la section suivante nous mobilisons de nombreux résultats obtenus à propos de la mise en œuvre de l'algorithme MCEM pour présenter l'algorithme SAEM que nous employons dans nos applications.

2.6. Algorithme SAEM

L'étape M de l'itération $m+1$ de l'algorithme MCEM présenté ci-dessus vise à réactualiser l'estimation de $\boldsymbol{\theta}_0$ obtenu à l'itération *m*, $\hat{\boldsymbol{\theta}}_{(m)}$, par :

(31) $$\hat{\boldsymbol{\theta}}_{(m+1)} = \arg\max_{\boldsymbol{\theta}} \sum_{i=1}^{N}\tilde{Q}_{N,R}(\boldsymbol{\theta};\hat{\boldsymbol{\theta}}_{(m)}).$$

où $\tilde{Q}_{N,R}(\boldsymbol{\theta};\hat{\boldsymbol{\theta}}_{(m)}) \equiv \sum_{i=1}^{N}\sum_{r=1}^{R}\tilde{w}_{i,r,(m)}\ln \ell_i^C(\boldsymbol{\theta};\tilde{\boldsymbol{\gamma}}_{i,r})$ est une approximation par des méthodes de

simulation de la fonction objectif $Q_N(\boldsymbol{\theta};\hat{\boldsymbol{\theta}}_{(m)}) \equiv \sum_{i=1}^{N} E_{(m),i}[\ln \ell_i^C(\boldsymbol{\theta};\boldsymbol{\gamma}_i)]$ de l'algorithme EM déterministe sous-jacent (mais inutilisable en pratique). En combinant les idées à l'origine de l'algorithme MCEM et les méthodes d'approximation stochastique (Robbins et Monro, 1951 ; Kiefer et Wolfowitz, 1952), Delyon et al (1999) proposent une approche alternative du calcul de l'estimateur du MV de $\boldsymbol{\theta}$. Ils proposent de maximiser une approximation stochastique de $Q_N(\boldsymbol{\theta};\hat{\boldsymbol{\theta}}_{(m)})$, reposant sur une autre utilisation des simulateurs des termes $E_{(m),i}[\ln \ell_i^C(\boldsymbol{\theta};\boldsymbol{\gamma}_i)]$.

Cet algorithme s'avère être globalement convergent sous des hypothèses très générales et très efficace en pratique (Allassonnière e*t al*, 2010 ; Lavielle, 2014). En outre, il s'avère aussi simple à mettre en œuvre, et à coder, que l'algorithme MCEM présenté ci-dessus.

L'algorithme SAEM est construit sur le principe des algorithmes (G)EM. Il utilise une étape M identique à celle des algorithmes (G)EM mais remplace l'étape E par deux étapes : une étape de simulation, l'étape S, et une étape d'approximation, l'étape SA. Ces étapes S et SA servent à construire l'approximation stochastique de $Q_N(\boldsymbol{\theta};\boldsymbol{\theta})$.

L'étape S de l'itération $m+1$ de l'algorithme SAEM consiste à construire une approximation par simulation des termes $E_{(m),i}[\ln \ell_i^C(\boldsymbol{\theta};\boldsymbol{\gamma}_i)]$ pour $i=1,...,N$. Il peut s'agir d'un seul tirage de $\boldsymbol{\gamma}_i$ suivant la densité $g_i(\hat{\boldsymbol{\theta}}_{(m)};\boldsymbol{\gamma})$, d'une moyenne de tels tirages ou d'un simulateur de $E_{(m),i}[\ln \ell_i^C(\boldsymbol{\theta};\boldsymbol{\gamma}_i)]$ tel que le simulateur IS présenté ci-dessus. Il est cependant important de

noter que ce simulateur n'a pas besoin d'être construit à partir d'un nombre important de simulations.

Nous avons utilisé des simulateurs IS de $E_{(m),i}[\ln \ell_i^C(\boldsymbol{\theta};\boldsymbol{\gamma}_i)]$ décrits précédemment, *i.e.* les moyennes pondérées $\sum_{r=1}^{R} \tilde{w}_{i,r,(m)} \ln \ell_i^C(\boldsymbol{\theta};\tilde{\boldsymbol{\gamma}}_{i,r})$.

L'étape SA de l'itération $m+1$ de l'algorithme SAEM construit l'approximation stochastique de $Q_N(\boldsymbol{\theta};\boldsymbol{\theta})$, $\tilde{H}_{(m)}(\boldsymbol{\theta})$, à partir de l'approximation construite lors de l'itération précédente $\tilde{H}_{(m)}(\boldsymbol{\theta})$ et des simulateurs $\sum_{r=1}^{R} \tilde{w}_{i,r,(m)} \ln \ell_i^C(\boldsymbol{\theta};\tilde{\boldsymbol{\gamma}}_{i,r})$ construits lors de l'étape S. La fonction $\tilde{H}_{(m+1)}(\boldsymbol{\theta})$ est construite à partir de la formule récursive:

$$\tilde{H}_{(m+1)}(\boldsymbol{\theta}) \equiv \tilde{H}_{(m)}(\boldsymbol{\theta}) + \vartheta_{m+1}\left(N^{-1}\sum_{i=1}^{N}\sum_{r=1}^{R} \tilde{w}_{i,r,(m)} \ln \ell_i^C(\boldsymbol{\theta};\tilde{\boldsymbol{\gamma}}_{i,r}) - \tilde{H}_{(m)}(\boldsymbol{\theta})\right) \tag{32}$$

où $\{\vartheta_m\}$ est une séquence positive décroissante telle que $\vartheta_1 = 1$, que $\sum_{m=1}^{+\infty} \vartheta_m = +\infty$ et que $\sum_{m=1}^{+\infty} \vartheta_m^2 < +\infty$.

Le paramètre ϑ_m définit le pas de l'approximation stochastique. Il détermine les poids relatifs des termes $\tilde{H}_{(m)}(\boldsymbol{\theta})$ et $N^{-1}\sum_{i=1}^{N}\sum_{r=1}^{R} \tilde{w}_{i,r,(m)} \ln \ell_i^C(\boldsymbol{\theta};\tilde{\boldsymbol{\gamma}}_{i,r})$ dans le calcul de $\tilde{H}_{(m+1)}(\boldsymbol{\theta})$. Plus ϑ_m est grand, plus l'algorithme SAEM atteint rapidement le voisinage d'un maximum de la log-vraisemblance mais plus il implique de bruit de simulation (Jank, 2006). Ce dernier peut être réduit en accroissant le nombre de tirages, R, employé lors de l'étape S. L'approche généralement adoptée a été proposée par Kuhn et Lavielle (2005), elle consiste à fixer à l le pas des m_1 premières itérations puis à diminuer progressivement ce pas. Nous avons choisi

des pas de la forme :

(33) $$\vartheta_m = \begin{cases} 1 \text{ pour } m_1 \geq m \geq 1 \\ (m - m_1 + 1)^{3/4} \text{ pour } m \geq m_1 \end{cases}$$

L'étape M de l'itération $m+1$ de l'algorithme SAEM consiste, à l'instar des algorithmes (G)EM, à réactualiser la valeur de $\boldsymbol{\theta}$ par maximisation de la fonction $\tilde{H}_{(m+1)}(\boldsymbol{\theta})$ ou en cherchant une valeur de $\boldsymbol{\theta}$ permettant d'accroître $\tilde{H}_{(m+1)}(\boldsymbol{\theta})$ à partir de $\tilde{H}_{(m+1)}(\hat{\boldsymbol{\theta}}_{(m)})$.

L'étape M de l'algorithme SAEM que nous employons dans nos applications reprend la structure de l'étape de l'algorithme MCEM présenté dans la section précédente. En effet, la structure de notre modèle multivarié à paramètres aléatoires implique que la fonction $\tilde{H}_{(m)}(\boldsymbol{\theta})$ peut être décomposée de manière additivement séparable avec :

(34) $$\tilde{H}_{(m)}(\boldsymbol{\theta}) = \tilde{H}^{\gamma}_{(m)}(\boldsymbol{\mu}, \boldsymbol{\Omega}) + \tilde{H}^{yxs}_{(m)}(\boldsymbol{\delta}, \boldsymbol{\Psi}) \quad \text{et} \quad \tilde{H}^{yxs}_{(m)}(\boldsymbol{\delta}, \boldsymbol{\Psi}) = \tilde{H}^{yx}_{(m)}(\boldsymbol{\delta}^{yx}, \boldsymbol{\Psi}^{yx}) + \tilde{H}^{s}_{(m)}(\boldsymbol{\delta}, \boldsymbol{\Psi}^{s}).$$

L'étape M de notre algorithme exploite cette décomposition en définissant les différents éléments de $\hat{\boldsymbol{\theta}}_{(m+1)}$ à partir de la série de problèmes de maximisation suivante :

(35) $$\begin{cases} (\hat{\boldsymbol{\mu}}_{(m+1)}, \hat{\boldsymbol{\Omega}}_{(m+1)}) = \arg\max_{(\boldsymbol{\mu}, \boldsymbol{\Omega})} \tilde{H}^{\gamma}_{(m+1)}(\boldsymbol{\mu}, \boldsymbol{\Omega}) \\ \hat{\boldsymbol{\delta}}_{(m+1)} = \arg\max_{\boldsymbol{\delta}} \tilde{H}^{yxs}_{(m+1)}(\boldsymbol{\delta}, \hat{\boldsymbol{\Psi}}_{(m+1)}) \\ \hat{\boldsymbol{\Psi}}^{yx}_{(m+1)} = \arg\max_{\boldsymbol{\Psi}^{yx}} \tilde{H}^{yx}_{(m+1)}(\hat{\boldsymbol{\delta}}^{yx}_{(m+1)}, \boldsymbol{\Psi}^{yx}) \\ \hat{\boldsymbol{\Psi}}^{s}_{(m+1)} = \arg\max_{\boldsymbol{\Psi}^{s}} \tilde{H}^{yx}_{(m+1)}(\hat{\boldsymbol{\delta}}_{(m+1)}, \boldsymbol{\Psi}^{s}) \end{cases}$$

L'étape M des algorithmes SAEM est particulièrement simple lorsque les densités noyau et mélangeante du modèle à paramètres aléatoires considéré sont exponentielles, ce qui est le cas de notre modèle. Dans ce cas, il est aisé de montrer que $\hat{\boldsymbol{\theta}}_{(m+1)}$ s'obtient sous forme

analytique en employant les statistiques suffisantes relatives aux vraisemblances impliquées dans le calcul de $\tilde{H}_{(m)}(\boldsymbol{\theta})$. Avec :

$$
(36)\quad \begin{cases}
\mathbf{h}^{\mu}_{(m+1)} \equiv \mathbf{h}^{\mu}_{(m)} + \vartheta_{m+1}\left(\sum_{i=1}^{N}\sum_{r=1}^{R}\tilde{w}_{i,r,(m)}\tilde{\boldsymbol{\gamma}}_{i,r} - \mathbf{h}^{\mu}_{(m)}\right) \\
\mathbf{H}^{\Omega}_{(m+1)} \equiv \mathbf{H}^{\Omega}_{(m)} + \vartheta_{m+1}\left(\sum_{i=1}^{N}\sum_{r=1}^{R}\tilde{w}_{i,r,(m)}\tilde{\boldsymbol{\gamma}}_{i,r}\tilde{\boldsymbol{\gamma}}'_{i,r} - \mathbf{H}^{\Omega}_{(m)}\right) \\
\mathbf{h}^{\delta}_{(m+1)} \equiv \mathbf{h}^{\delta}_{(m)} + \vartheta_{m+1}\left(\sum_{r=1}^{R}\tilde{w}_{i,r,(m)}\varsigma_{it}(\tilde{\boldsymbol{\gamma}}_{i,r}) - \mathbf{h}^{\delta}_{(m)}\right) \\
\mathbf{H}^{\Psi yx}_{(m+1)} \equiv \mathbf{H}^{\Psi yx}_{(m)} + \vartheta_{m+1}\left(\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{r=1}^{R}\tilde{w}_{i,r,(m)}\boldsymbol{\varepsilon}^{yx}_{it}(\tilde{\boldsymbol{\gamma}}_{i,r},\hat{\tilde{\boldsymbol{\delta}}}^{yx}_{(m+1)})\boldsymbol{\varepsilon}^{yx}_{it}(\tilde{\boldsymbol{\gamma}}_{i,r},\hat{\tilde{\boldsymbol{\delta}}}^{yx}_{(m+1)})' - \mathbf{H}^{\Psi yx}_{(m)}\right) \\
\mathbf{H}^{\Psi s}_{(m+1)} \equiv \mathbf{H}^{\Psi s}_{(m)} + \vartheta_{m+1}\left(\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{r=1}^{R}\tilde{w}_{i,r,(m)}\boldsymbol{\varepsilon}^{s}_{it}(\tilde{\boldsymbol{\gamma}}_{i,r},\hat{\tilde{\boldsymbol{\delta}}}_{(m+1)})\boldsymbol{\varepsilon}^{s}_{it}(\tilde{\boldsymbol{\gamma}}_{i,r},\hat{\tilde{\boldsymbol{\delta}}}_{(m+1)})' - \mathbf{H}^{\Psi s}_{(m)}\right)
\end{cases}
$$

on obtient simplement :

$$
(37)\quad \begin{cases}
\hat{\boldsymbol{\mu}}_{(m+1)} = N^{-1}\mathbf{h}^{\mu}_{(m+1)} \\
\hat{\boldsymbol{\Omega}}_{(m+1)} = N^{-1}\mathbf{H}^{\Omega}_{(m+1)} - \hat{\boldsymbol{\mu}}'_{(m+1)}\hat{\boldsymbol{\mu}}'_{(m+1)} \\
\hat{\boldsymbol{\delta}}_{(m+1)} = \left(\sum_{i=1}^{N}\sum_{t=1}^{T}\mathbf{Z}_{it}\hat{\boldsymbol{\Psi}}^{-1}_{(m)}\mathbf{Z}'_{it}\right)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\mathbf{Z}_{it}\hat{\boldsymbol{\Psi}}^{-1}_{(m)}\mathbf{h}^{\delta}_{(m+1)} \\
\hat{\boldsymbol{\Psi}}^{yx}_{(m+1)} = (NT)^{-1}\mathbf{H}^{\Psi yx}_{(m+1)} \\
\hat{\boldsymbol{\Psi}}^{s}_{(m+1)} = (NT)^{-1}\mathbf{H}^{\Psi s}_{(m+1)}
\end{cases}
$$

Le calcul des statistiques suffisantes $\mathbf{h}^{\mu}_{(m+1)}$, $\mathbf{H}^{\Omega}_{(m+1)}$, $\mathbf{h}^{\delta}_{(m+1)}$, $\mathbf{H}^{\Psi yx}_{(m+1)}$ et $\mathbf{H}^{\Psi s}_{(m+1)}$ constitue de fait l'étape SA de l'itération $(m+1)$.

Nous utilisons la règle d'arrêt standard de l'algorithme EM (Booth et Hobert, 1999). Cette dernière est basée sur la variation relative des valeurs des paramètres entre deux itérations consécutives:

$$
(38)\qquad e(k) = \max_{j}\left(\frac{\left|\hat{\theta}_{j,(m)} - \hat{\theta}_{j,(m-1)}\right|}{\left|\hat{\theta}_{j,(m-1)}\right| + \sigma_1}\right) < \sigma_2 ,
$$

les paramètres σ_1 et σ_2 prenant de petites valeurs. Nous stoppons l'algorithme lorsque la condition $e(m) < \sigma_2$ est satisfaite pour trois itérations consécutives. Il existe toujours un risque que cette condition soit satisfaite alors qu'un maximum de la log-vraisemblance n'est pas atteint (Gu et Zhu, 2001). Pour cette raison, il est conseillé d'examiner les équations de score lors de l'arrêt de l'algorithme.

2.7. Distribution asymptotique des estimateurs calculés

Les estimateurs de $\boldsymbol{\theta}_0$ calculés à partir des algorithmes MCEM et SAEM présentés ci-dessus ont des propriétés similaires à celle de son estimateur MVS. Leurs propriétés asymptotiques sont équivalentes à celle de l'estimateur du MV (infaisable en pratique) correspondant si R croît plus vite que $\sqrt{N}$, *i.e.* si le nombre de tirages employés est suffisamment grand et est d'autant plus grand que l'échantillon employé est grand. Ces estimateurs, notés $\hat{\boldsymbol{\theta}}_N$ ici, sont donc asymptotiquement normaux (et asymptotiquement efficaces), avec :

(39) $$\sqrt{N}(\hat{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_0) \xrightarrow[N \to +\infty]{L} \mathcal{N}(\mathbf{0}, \mathbf{J}_0^{-1}).$$

La propriété d'égalité des matrices d'information garantit ici que la matrice d'information de Fisher $\mathbf{J}_0$ est égale à l'espérance du produit croisé des scores individuels calculés en $\boldsymbol{\theta}_0$:

(40) $$\mathbf{J}_0 \equiv E\left[\tfrac{\partial}{\partial \boldsymbol{\theta}} \ln \ell_i(\boldsymbol{\theta}_0) \tfrac{\partial}{\partial \boldsymbol{\theta}'} \ln \ell_i(\boldsymbol{\theta}_0)\right]$$

Louis (1982) a montré que le score de l'observation i en $\boldsymbol{\theta}_0$ est donné par :

(41) $$\tfrac{\partial}{\partial\boldsymbol{\theta}}\ln\ell_i(\boldsymbol{\theta}_0)=\int\tfrac{\partial}{\partial\boldsymbol{\theta}}\ln\ell_i^C(\boldsymbol{\theta}_0;\boldsymbol{\gamma})g_i(\boldsymbol{\gamma};\boldsymbol{\theta}_0)d\boldsymbol{\gamma}\,.$$

Ces scores peuvent être estimés en utilisant le simulateur IS utilisé pour les procédures d'estimation calculé en $\hat{\boldsymbol{\theta}}_N$. Il suffit d'obtenir R tirages de $\boldsymbol{\gamma}_i$ selon la loi $\mathcal{N}(\hat{\boldsymbol{\mu}}_N,\hat{\boldsymbol{\Omega}}_N)$, les $\tilde{\boldsymbol{\gamma}}_{i,r}$ pour $r=1,...,N$ et de calculer les poids correspondants :

(42) $$\hat{w}_{i,r}\equiv\frac{\prod_{t=1}^{T}\ell_{it}(\hat{\boldsymbol{\delta}}_N,\hat{\boldsymbol{\Psi}}_N\mid\hat{\boldsymbol{\gamma}}_{i,r})}{\sum_{r=1}^{R}\prod_{t=1}^{T}\ell_{it}(\hat{\boldsymbol{\delta}}_N,\hat{\boldsymbol{\Psi}}_N\mid\hat{\boldsymbol{\gamma}}_{i,r})}\,.$$

Le termes $\sum_{r=1}^{R}\hat{w}_{i,r}\tfrac{\partial}{\partial\boldsymbol{\theta}}\ln\ell_i^C(\hat{\boldsymbol{\theta}}_N;\hat{\boldsymbol{\gamma}}_{i,r})$ sont des estimateurs convergents (en N et R) des scores $\tfrac{\partial}{\partial\boldsymbol{\theta}}\ln\ell_i(\boldsymbol{\theta}_0)$. Ils permettent de calculer un estimateur convergent de $\mathbf{J}_0$ sous la forme:

(43) $$\hat{\mathbf{J}}_N\equiv N^{-1}\sum_{i=1}^{N}\Big(\sum_{r=1}^{R}\hat{w}_{i,r}\tfrac{\partial}{\partial\boldsymbol{\theta}}\ln\ell_i^C(\hat{\boldsymbol{\theta}}_N;\hat{\boldsymbol{\gamma}}_{i,r})\Big)\Big(\sum_{r=1}^{R}\hat{w}_{i,r}\tfrac{\partial}{\partial\boldsymbol{\theta}'}\ln\ell_i^C(\hat{\boldsymbol{\theta}}_N;\hat{\boldsymbol{\gamma}}_{i,r})\Big).$$

2.8. Calibration des paramètres individuels et diagnostics

Obtenir des estimateurs des paramètres individuels $\boldsymbol{\gamma}_i$ pour chaque agriculteur permet d'atteindre deux objectifs. Ces estimateurs permettent de construire des modèles originaux de simulation de politiques agri-environnementales à partir des modèles de choix multiculture utilisés ici. Les modèles de simulation ainsi obtenus sont originaux car définis comme un échantillon de modèles hétérogènes. La construction de ces modèles de simulation est discutée plus en détail dans les Chapitres 5 et 6, chapitres dans lesquels sont présentés nos applications.

Les estimateurs de $\boldsymbol{\gamma}_i$ peuvent également être employés pour examiner la qualité

d'estimation du modèle. Cette utilisation d'estimations de $\boldsymbol{\gamma}_i$ sera discutée ici, à la suite de la présentation des principales approches disponibles pour obtenir ces estimations.

Bien entendu, il n'est pas question ici d'estimer les termes $\boldsymbol{\gamma}_i$ directement à partir des données. Il s'agit de définir un estimateur des $\boldsymbol{\gamma}_i$ et de calculer les estimations correspondantes de ces termes à partir des données disponibles et du modèle estimé, *i.e.* de sa forme paramétrique et de l'estimation disponible de $\boldsymbol{\theta}_0$.

Le titre de cette section fait référence à une procédure de calibration plutôt qu'à une procédure d'estimation pour souligner le fait que les estimations des paramètres $\boldsymbol{\gamma}_i$ présentées ici reposent sur un ensemble d'information limitée par rapport à l'exploitation *i*. Si les estimations obtenues ont des fondements statistiques solides, chaque évaluation de $\boldsymbol{\gamma}_i$ présentée ici repose principalement sur la « petite » série de choix observés de l'agriculteur dans les données disponibles.

L'estimateur de $\boldsymbol{\gamma}_i$ le plus couramment employé est son espérance conditionnelle en les données disponibles. Cette espérance conditionnelle est le meilleur prédicteur de $\boldsymbol{\gamma}_i$ au sens de l'erreur quadratique (de prédiction) espérée. Etant données les hypothèses définissant notre modèle de choix de production, cet estimateur est donné par :

(44) $$\hat{\boldsymbol{\gamma}}_i \equiv E[\boldsymbol{\gamma}_i \mid \mathbf{y}_i, \mathbf{x}_i, \mathbf{s}_i, \mathbf{z}_i] = \int \boldsymbol{\gamma} g_i(\boldsymbol{\gamma}; \boldsymbol{\theta}_0) d\boldsymbol{\gamma}$$

Une estimation de $\hat{\boldsymbol{\gamma}}_i$ peut être obtenue à partir du simulateur IS utilisé dans les algorithmes MCEM et SAEM en l'estimation de $\boldsymbol{\theta}_0$ disponible, $\hat{\boldsymbol{\theta}}_N$. On peut ici réutiliser les tirages de

$\boldsymbol{\gamma}_i$ et leurs poids tels qu'ils ont été construits pour l'estimation de la matrice de variance-covariance asymptotique de l'estimateur $\hat{\boldsymbol{\theta}}_N$. Cette estimation est simplement donnée par :

(45) $$\hat{\boldsymbol{\gamma}}_i(\hat{\boldsymbol{\theta}}_N) \equiv \sum_{r=1}^{R} \hat{w}_{i,r} \hat{\boldsymbol{\gamma}}_{i,r} \ .$$

Définir d'estimateur de $\boldsymbol{\gamma}_i$ selon le principe du MV consiste à définir cet estimateur comme le mode de la distribution de $\boldsymbol{\gamma}_i$ conditionnelle en $(\mathbf{y}_i, \mathbf{x}_i, \mathbf{s}_i, \mathbf{z}_i)$. Une estimation de ce mode conditionnel peut être obtenue à partir des tirages de $\boldsymbol{\gamma}_i$ suivant la loi $\mathcal{N}(\hat{\boldsymbol{\mu}}_N, \hat{\boldsymbol{\Omega}}_N)$, les $\hat{\boldsymbol{\gamma}}_{i,r}$ pour $r = 1,...,R$, et les poids correspondants, les $\hat{w}_{i,r}$ pour $r = 1,...,R$. En effet, le tirage de $\boldsymbol{\gamma}_i$ ayant le poids le plus élevé est une estimation convergente du mode de la distribution de $\boldsymbol{\gamma}_i$ conditionnelle en $(\mathbf{y}_i, \mathbf{x}_i, \mathbf{s}_i, \mathbf{z}_i)$.

Ce mode conditionnel peut également être obtenu en maximisant une estimation de la fonction de densité de $\boldsymbol{\gamma}_i$ conditionnelle en $(\mathbf{y}_i, \mathbf{x}_i, \mathbf{s}_i, \mathbf{z}_i)$, $g_i(\boldsymbol{\gamma}; \boldsymbol{\theta}_0)$. Cet estimateur du mode conditionnel est nommé Maximum *a posteriori* (MAP) de $g_i(\boldsymbol{\gamma}; \boldsymbol{\theta}_0)$. En utilisant la définition de $g_i(\boldsymbol{\gamma}; \boldsymbol{\theta}_0)$, il est aisé de montrer que cet estimateur de $\boldsymbol{\gamma}_i$ est donné par:

(46) $$\hat{\boldsymbol{\gamma}}_i = \arg\max_{\boldsymbol{\gamma}} \ln \ell_i^C(\boldsymbol{\theta}_0 \mid \boldsymbol{\gamma}) \ ,$$

i.e. le maximum en $\boldsymbol{\gamma}$ de la log-vraisemblance complète de l'observation *i*. Ceci implique que le terme $\hat{\boldsymbol{\gamma}}_i(\hat{\boldsymbol{\theta}}_N) = \arg\max_{\boldsymbol{\gamma}} \ln \ell_i^C(\hat{\boldsymbol{\theta}}_N \mid \boldsymbol{\gamma})$ fournit une estimation convergente de $\hat{\boldsymbol{\gamma}}_i$.

Dans tous les cas, ces estimations de $\boldsymbol{\gamma}_i$ sont peu précises lorsque l'information disponible

sur l'exploitation *i* dans les données est limitée, *i.e.* lorsque les données sont peu informatives au niveau individuel. Dans de tels cas, les estimations de l'espérance conditionnelle ou du mode conditionnel de $\boldsymbol{\gamma}_i$ tendent à se concentrer autour de l'estimation de l'espérance (non conditionnelle) de $\boldsymbol{\gamma}_i$, $\hat{\boldsymbol{\mu}}_N$, selon un effet dit de « $\boldsymbol{\gamma}_i$ – shrinkage » (Savic and Karlsson, 2009).

Lorsque ce phénomène apparaît, une meilleure estimation de $\boldsymbol{\gamma}_i$ peut être obtenue par simulation à partir d'une approximation par la loi normale de la distribution de $\boldsymbol{\gamma}_i$ conditionnelle en $(\mathbf{y}_i, \mathbf{x}_i, \mathbf{s}_i, \mathbf{z}_i)$, *i.e.* en utilisant un tirage de $\boldsymbol{\gamma}_i$ selon la loi normale approchée $\mathcal{N}(E[\boldsymbol{\gamma}_i \mid \mathbf{y}_i, \mathbf{x}_i, \mathbf{s}_i, \mathbf{z}_i], V[\boldsymbol{\gamma}_i \mid \mathbf{y}_i, \mathbf{x}_i, \mathbf{s}_i, \mathbf{z}_i])$. Bien entendu, l'espérance et la matrice de variance-covariance de cette distribution peuvent être estimées *via* le simulateur IS, avec $\sum_{r=1}^{R} \hat{w}_{i,r} \hat{\boldsymbol{\gamma}}_{i,r}$ pour l'espérance et $\sum_{r=1}^{R} \hat{w}_{i,r} \hat{\boldsymbol{\gamma}}_{i,r} \hat{\boldsymbol{\gamma}}'_{i,r} - \left(\sum_{r=1}^{R} \hat{w}_{i,r} \hat{\boldsymbol{\gamma}}_{i,r}\right)\left(\sum_{r=1}^{R} \hat{w}_{i,r} \hat{\boldsymbol{\gamma}}'_{i,r}\right)$ pour la matrice de variance-covariance. Il est également possible d'employer directement les tirages $\hat{\boldsymbol{\gamma}}_{i,r}$ et leurs poids $\hat{w}_{i,r}$ pour $r = 1, ..., R$. L'estimation de $\boldsymbol{\gamma}_i$ est alors obtenue comme un tirage d'un élément dans la collection des $\hat{\boldsymbol{\gamma}}_{i,r}$ selon une loi multinomiale telle que $P[\hat{\boldsymbol{\gamma}}_i(\hat{\boldsymbol{\theta}}_N) = \hat{\boldsymbol{\gamma}}_{i,r}] = \hat{w}_{i,r}$.

La justification de cette dernière méthode d'estimation de $\boldsymbol{\gamma}_i$ provient de ce que la distribution des $\boldsymbol{\gamma}_i$ dans la population doit être identique à celle des prédicteurs $\hat{\boldsymbol{\gamma}}_i$ des $\boldsymbol{\gamma}_i$ dans la population. Cette égalité fournit un critère de qualité pour le modèle estimé. Généralement, la distribution des $\hat{\boldsymbol{\gamma}}_i(\hat{\boldsymbol{\theta}}_N)$ dans l'échantillon est comparée graphiquement à cette de $\mathcal{N}(\hat{\boldsymbol{\mu}}_N, \hat{\boldsymbol{\Omega}}_N)$, la distribution estimée des $\boldsymbol{\gamma}_i$. Il est également possible de comparer

les moyennes et variances empiriques des $\hat{\boldsymbol{\gamma}}_i(\hat{\boldsymbol{\theta}}_N)$ dans l'échantillon à $\hat{\boldsymbol{\mu}}_N$ et $\hat{\boldsymbol{\Omega}}_N$ (Train, 2009).

2.9. Remarques conclusives

Les procédures décrites ici ont été employées directement pour conduire les estimations présentées dans le Chapitre 5. Elles servent de base à celles utilisées pour conduire les estimations présentées dans le Chapitre 6. Dans les deux cas, l'algorithme SAEM s'est avéré performant et relativement robuste pour estimer nos modèles de choix de production multiculture à paramètres aléatoires.

Néanmoins, s'ils se réfèrent ici à l'estimation de notre modèle de choix de production multiculture, les travaux présentés dans ce chapitre ont, selon nous, une portée plus générale. Ils montrent que les outils d'inférence développés en statistique computationnelle peuvent être très utiles aux micro-économètres appliqués qui considèrent des modèles multivariés à paramètres aléatoires.

Tout exercice de modélisation micro-économétrique implique un arbitrage entre, schématiquement, le « réalisme » du modèle considéré d'une part, et la difficulté de son estimation. Dans cette partie nous avons opté pour le « réalisme » du modèle. Par exemple, supprimer les variables de contrôle du modèle (qui s'avèrent peu informatives selon les résultats obtenus dans le Chapitre 6) aurait significativement simplifié son estimation. Cet arbitrage s'avère bien plus délicat pour les choix de production modélisés dans le chapitre suivant.

Chapitre 3

Modèle de choix d'assolement, choix de cultures et solutions en coin

Tout micro-économètre de la production agricole s'intéressant à la modélisation des choix d'assolement des agriculteurs est confronté au problème des solutions en coin affectant ces choix, *i.e.* à la prise en compte des assolements nuls pour certaines cultures (Carpentier *et al*, 2015). En effet, les agriculteurs produisent rarement les mêmes cultures, les choix de cultures différant entre agriculteurs et pouvant évoluer dans le temps pour un agriculteur donné.

Par exemple, l'échantillon duquel celui utilisé dans l'application présentée dans le Chapitre 6 est tiré contient 6001 observations. En se concentrant sur les 8 cultures principales, il s'avère que 64 combinaisons de cultures différentes, ce que dans la suite nous nommerons les « régimes de production », apparaissent dans les observations (sur les 127 combinaisons possibles). De même, la quasi-totalité des agriculteurs adoptent des combinaisons de cultures différentes sur une période de 3 années consécutives, abandonnant et/ou introduisant des cultures dans leur assolement certaines années.

De fait, plus que la question des assolements nuls de certaines cultures, la question d'intérêt ici est celle des agriculteurs de choisir d'assoler ou non certaines cultures. En effet, le choix des cultures assolées renvoie à la notion de diversification des cultures, qui elle-même tend à devenir centrale dans la construction des politiques agri-environnementales, notamment dans l'Union Européenne. Par exemple, une mesure dite de diversification a récemment été introduite dans la Politique Agricole Commune (PAC). Dorénavant, le versement des aides directes de la PAC aux producteurs de grandes cultures et cultures industrielles européens est conditionné au respect de critères de diversification de leurs assolements. Ces critères sont très peu contraignants à l'heure actuelle. Ils pourraient devenir plus contraignants à l'avenir. En effet, la diversification des cultures est un des principaux leviers agronomiques permettant

de réduire l'utilisation d'intrants chimiques – pesticides ou engrais minéraux – voire de réduire les besoins en eau d'irrigation. Cette diversification est la clé de voûte des systèmes de production agricole agro-écologiques pour les cultures annuelles en général, et pour les grandes cultures et cultures industrielles en particulier.

Aussi, modéliser de manière adéquate les choix de cultures des agriculteurs maintenant est un prérequis pour analyser de manière pertinente des questions de politiques agri-environnementales de plus en plus pressantes. Ceci dit, tenir compte des décisions de produire ou non certaines cultures est de nature à améliorer significativement les analyses des décisions de production des agriculteurs en général. Par, exemple les résultats d'estimation présentés dans le Chapitre 6 de cette thèse tendent à montrer que les effets de la décision de produire ou non compte pour environ 50% de l'élasticité prix propre agrégée des surfaces des cultures qui ne sont pas produites par tous les agriculteurs de la population considérée.

Or les économistes de la production agricole ont peu analysé la question des choix de cultures dans un contexte multiculture. Si l'utilisation des modèles de choix de cultures à l'échelle d'une parcelle se développe très rapidement depuis une dizaine d'années (voir, *e.g.*, Wu and Brorsen, 1995 ; Wu, 2000 ; Wu et al, 2004 ; Ji et al, 2014 ; Fezzi et al, 2014 ; Livingston et al, 2015 ; Costinot et al, 2016), ces modèles considèrent les choix d'un même agriculteur comme indépendants entre eux. Cette hypothèse est discutable. En particulier, elle va à l'encontre des fondements des modèles de choix de production multiculture considérant la terre comme un facteur fixe allouable (voir, *e.g.*, Carpentier et Letort, 2014).

Les économistes analysant les choix d'assolement des agriculteurs ont abordé la question

des choix de cultures de manière superficielle lorsqu'ils ne l'éliminaient pas *a priori*. Ils ont abordé cette question essentiellement à travers la gestion du problème d'estimation posé par les assolements nuls. Ils ont essentiellement employé deux approches. La première consiste à agréger les cultures de sorte à éliminer les assolements nuls tout en considérant des agrégations les plus cohérentes possibles. Bien entendu, cette procédure d'agrégation réduit le niveau de précisions des analyses conduites et élimine définitivement le problème du choix de cultures de ces analyses. La seconde approche adoptée consiste à spécifier les choix d'assolement des agriculteurs comme des systèmes de régressions censurées de sorte à tenir compte explicitement de la masse des assolements nuls dans le modèle micro-économétrique considéré (voir, *e.g.*, Sckokaï et Moro, 2006, 2012 ; Lacroix et Thomas ; 2011 ; Fezzi et Bateman 2011 ; Fezzi *et al*, 2014).

Les modèles employés dans ces travaux s'inspirent des travaux pionniers de Shonkwiler et Yen[12] (1990) sur l'estimation des systèmes de régressions censurées. Or, les modèles de choix de production avec solutions en coin ne devraient pas être spécifiés comme des systèmes de régressions censurées, mais comme des systèmes d'équations à changement de régime endogène à l'instar de ceux proposés par Wales et Woodland (1980) ou Lee et Pitt

[12] Ainsi que des nombreux articles subséquents de S.T. Yen. Contrairement à celle des systèmes à changement de régime endogène, l'estimation des systèmes de régressions censurées peut être conduite en deux étapes relativement simples selon la logique de la correction pour la sélection endogène de Heckman, *i.e.* avec l'ajout dans l'estimation d'un terme de « ratio de Mills » estimé en première étape. Cette approche d'estimation par étapes explique certainement le succès de la modélisation par des systèmes de régressions censurées des choix avec solutions en coin.

(1986, 1987) pour la modélisation des systèmes de fonctions de demande avec demandes nulles. Les critiques de Arndt, Liu et Preckel (1999) à propos des systèmes de fonctions de demande avec solutions en coin employés par les économistes agricoles s'appliquent aux modèles de choix de production : les systèmes de régressions censurées ne sont que des approximations assez frustes des systèmes d'équations à changement de régime endogène qui devraient être employés (voir également, e.g., Arndt, 1999 ; Carpentier *et al* 2015). Un exemple simple suffit pour s'en convaincre : dans les systèmes de régressions censurées le prix d'une culture non produite affecte le choix d'assolement des cultures produites.

L'objectif de ce chapitre est tout d'abord d'examiner les fondements théoriques des choix de cultures des agriculteurs. Il s'agit ensuite de mettre en évidence les limites des modélisations micro-économétriques proposées jusqu'à présent pour tenir compte de ces choix de cultures de manière cohérente par rapport à la théorie économique. En particulier, nous présenterons brièvement une adaptation des modèles de consommation de Wales et Woodland (1980) au cas des choix d'assolement.

Enfin, nous présenterons des modèles de choix d'assolement avec choix de cultures et coûts fixes de mise en culture. Ces modèles sont conçus comme des extensions des modèles de choix d'assolement MNL proposés par Carpentier et Letort (2014) et utilisés dans les deux chapitres précédents. Nous présenterons ces modèles d'un point de vue théorique. En particulier, nous montrerons en quoi les spécificités des modèles de choix d'assolement MNL les rendent particulièrement adaptés à la modélisation simultanée des choix de cultures et des choix d'assolement. Nous présenterons ensuite des versions micro-économétriques de ces modèles. Ces derniers sont conçus comme des modèles à paramètres aléatoires, à changement

de régime endogène et avec coûts fixes de choix de régime. Ils sont donc très originaux. Le modèle employé dans l'application dont les résultats sont présentés dans le Chapitre 6 de cette thèse appartient à cette famille de modèles. La question de l'estimation de ces modèles est abordée dans le chapitre suivant.

3.1. Modèles de choix d'assolement et régimes de productions endogènes

Nous considérons une version générique des modèles de choix d'assolement considérés dans les chapitres précédents. Les agriculteurs considérés sont supposés pouvoir allouer leur surface de terre arable à un ensemble de cultures définissant l'ensemble $\mathcal{K} \equiv \{1,\ldots,K\}$. L'assolement en *t* de l'agriculteur *i* est décrit par les parts de sa surface totale, $s_{k,it}$, dédiées à chaque culture $k \in \mathcal{K}$. Les agriculteurs sont supposés utiliser l'ensemble de sa surface de terre arable, *i.e.* que $\sum_{k\in\mathcal{K}} s_{k,it} = 1$ pour $i = 1,...,N$ et $t = 1,...,T$. L'espérance en *t* de la marge brute de la culture *k* de l'agriculteur *i* est mesurée par $\pi_{k,it}$, avec $\boldsymbol{\pi}_{it} = (\pi_{k,it} : k \in \mathcal{K})$.

Nous supposons que tous les agriculteurs produisent la culture *K* chaque année, *i.e.* que $s_{K,it} > 0$ pour $i = 1,...,N$ et $t = 1,...,T$. La culture *K* servira de culture de référence dans les modèles utilisés ici. L'ensemble $\mathcal{K}_{-K} \equiv \{1,\ldots,K-1\}$ désigne l'ensemble des cultures considérées, culture *K* exclue. L'assolement des agriculteurs est décrit par les vecteurs $\mathbf{s}_{it} = (s_{k,it} : k \in \mathcal{K}_{-K})$, avec $s_{K,it} = 1 - \sum_{k\in\mathcal{K}_{-K}} s_{k,it}$.

Les agriculteurs sont supposés être neutres face au risque et décider leur choix d'assolement

sur la base d'un problème de maximisation d'une fonction de profit espéré définie par :

(1) $$\Pi_{it}(\mathbf{s};\boldsymbol{\pi}_{it}) = \sum_{k\in\mathcal{K}} s_k \pi_{k,it} - C_{it}(\mathbf{s}) \text{ avec } \mathbf{s} = (s_k : k \in \mathcal{K}_{-K}) \text{ et } s_K = 1 - \sum_{k\in\mathcal{K}_{-K}} s_k$$

où $C_{it}(\mathbf{s})$ désigne un exemple de fonction de « coût implicite de gestion de l'assolement » telle que ces fonctions ont été définies par Carpentier et Letort (2012, 2014). Cette fonction est supposée strictement convexe en **s** sur :

(2) $$\mathcal{S} \equiv \left\{ \mathbf{s} \in \mathbb{R}^{K-1} \mid \sum_{k\in\mathcal{K}_{-K}} s_k \leq 1 \text{ et } s_k \geq 0,\ k \in \mathcal{K}_{-K} \right\}.$$

Elle définit de ce fait les coûts implicites de gestion de l'assolement comme des motifs de diversification des choix d'assolement des agriculteurs.

Compte tenu de la définition de **s** et de l'hypothèse d'utilisation totale de la terre, la fonction de profit peut s'écrire sous la forme :

(3) $$\Pi_{it}(\mathbf{s};\boldsymbol{\pi}_{it}) = \sum_{k\in\mathcal{K}_{-K}} s_k (\pi_{k,it} - \pi_{K,it}) - C_{it}(\mathbf{s}) + \pi_{K,it} \text{ avec } s_K = 1 - \sum_{k\in\mathcal{K}_{-K}} s_k .$$

Cette expression montre que les choix d'assolement optimaux ne dépendent pas que des différences de marges brutes espérées. Le vecteur $\bar{\boldsymbol{\pi}}_{it}$ désigne le vecteur des différences de marges brutes utilisant celle de la culture K en référence, avec $\bar{\boldsymbol{\pi}}_{it} = (\pi_{k,it} - \pi_{K,it} : k \in \mathcal{K}_{-K})$. Nous supposons en outre que la fonction de coût $C_{it}(\mathbf{s})$ peut être décomposée en :

(4) $$C_{it}(\mathbf{s}) = \sum_{k\in\mathcal{K}_{-K}} s_k \beta^s_{k,it} + C^s_i(\mathbf{s}) \text{ avec } s_K = 1 - \sum_{k\in\mathcal{K}_{-K}} s_k$$

où la fonction $C^s_i(\mathbf{s})$ est, par construction, convexe en **s** sur $\mathcal{S}$. C'est donc cette fonction qui porte spécifiquement le motif de diversification de l'assolement.

Sous les hypothèses posées ici, le choix d'assolement de l'agriculteur i en t résulte de la

maximisation de $\Pi_{it}(\mathbf{s};\boldsymbol{\pi}_{it})$ en $\mathbf{s}$, *i.e.* le vecteur $\mathbf{s}_{it}$ est définit par :

$$\mathbf{s}_{it} = \arg\max_{\mathbf{s}} \left\{ \begin{array}{c} \sum_{k\in\mathcal{K}_{-K}} s_k(\pi_{k,it} - \pi_{K,it} - \beta^s_{k,it}) - C^s_i(\mathbf{s}) \\ \text{s.c.} \\ \sum_{k\in\mathcal{K}_{-K}} s_k \leq 1 \text{ et } s_k \geq 0 \text{ pour } k \in \mathcal{K}_{-K} \end{array} \right\}. \tag{5}$$

Les contraintes imposées sur $\mathbf{s}$ dans ce problème sont équivalentes à $\mathbf{s} \in \mathcal{S}$. Elles sont explicitées ici parce que leur interprétation est importante. La fonction $\Pi_{it}(\mathbf{s};\boldsymbol{\pi}_{it})$ étant strictement concave en $\mathbf{s} \in \mathcal{S}$, le vecteur $\mathbf{s}_{it}$ est l'unique solution en $\mathbf{s}$ du problème de maximisation considéré.

Puisque les agriculteurs assolent toujours la culture K (et $\mathbf{s} \geq \mathbf{0}$), nous savons que la contrainte $\sum_{k\in\mathcal{K}_{-K}} s_k \leq 1$ n'est jamais saturée à l'optimum. En revanche les assolements des autres cultures peuvent être nuls, *i.e.* certaines des solutions en les s_k peuvent être des solutions en coin saturant les contraintes de non-négativité correspondantes. Dans la suite nous désignons par $\mathcal{K}^+_{it}$, respectivement par $\mathcal{K}^0_{it}$, le sous-ensemble de $\mathcal{K}$ contenant les cultures produites, respectivement non produites, par l'agriculteur i en t. Le terme $\mathcal{K}^+_{-K,it}$ désigne l'ensemble, culture K exclue, des cultures produites par l'agriculteur i en t.

Sous les hypothèses considérées, les conditions de Kuhn-Karush-Tucker (KKT) caractérisent le choix d'assolement $\mathbf{s}_{it}$ avec :

$$\left\{ \begin{array}{l} \pi_{k,it} - \pi_{K,it} - \beta^s_{k,it} - \frac{\partial}{\partial s_k} C^s_i(\mathbf{s}_{it}) = 0 \text{ et } s_{k,it} > 0 \text{ pour } k \in \mathcal{K}^+_{-K,it} \\ \pi_{k,it} - \pi_{K,it} - \beta^s_{k,it} - \frac{\partial}{\partial s_k} C^s_i(\mathbf{s}_{it}) < 0 \text{ et } s_{k,it} = 0 \text{ pour } k \in \mathcal{K}^0_{it} \end{array} \right. \tag{6}$$

Ces conditions de KKT indiquent simplement que l'agriculteur n'assole pas la culture k si sa marge brute espérée ne compense pas les coûts de gestion que son assolement génère.

3.1.1. Modèles économétriques fondés sur les conditions de KKT caractérisant les choix d'assolement

L'approche proposée par Wales et Woodland (1980) pour la modélisation micro-économétrique de systèmes de fonctions de demande peut être adaptée au cas des choix d'assolement considérés ici. Nous développons cette idée pour illustrer les limites des modèles fondés sur les conditions de KKT mais également afin de présenter des concepts et mécanismes importants dans un cadre d'analyse simple.[13]

Définir un modèle de choix d'assolement « à la Wales et Woodland » consiste simplement à choisir une forme fonctionnelle quadratique déterministe pour $C_i^s(\mathbf{s})$ et à supposer que les termes $\boldsymbol{\beta}_{it}^s$ suivent une loi normale multivariée. Nous considérons ici le modèle le plus simple possible. La partie convexe de la fonction de coût implicite $C_i^s(\mathbf{s})$ est donnée par $C_i^s(\mathbf{s}) = (1/2) \times \mathbf{s}'\mathbf{H}_0\mathbf{s}$, la matrice $\mathbf{H}_0$ étant supposée définie positive. Le vecteur des coûts spécifiques à chaque culture $\boldsymbol{\beta}_{it}^s \equiv (\beta_{k,it}^s : k \in \mathcal{K}_{-K})$ est supposé suivre, conditionnellement en $\bar{\boldsymbol{\pi}}_{it}$, la loi normale multivariée $\mathcal{N}(\mathbf{b}_0, \boldsymbol{\Psi}_0^s)$. L'objet du problème d'inférence statistique considéré ici est l'estimation du terme $(\mathbf{H}_0, \mathbf{b}_0, \boldsymbol{\Psi}_0^s)$ selon le principe du MV.[14]

---

[13] L'approche de Lee et Pitt (1986), la plus utilisée en micro-économétrie appliquée, aurait également pu être présenté dans le cadre de modélisation. Mais cette approche duale suppose des calculs et l'introduction de concepts supplémentaires.

[14] De fait, le modèle développé dans la suite peut être considéré comme une extension du modèle de choix d'assolement « sans solutions en coin » proposé par Carpentier et Letort (2012).

Dans la suite le terme r_{it} désigne le « régime de production » de l'agriculteur *i* en l'année *t*, *i.e.* l'ensemble des cultures produites par cet agriculteur l'année considérée. Cette notion joue un rôle ici dans la mesure où le régime de production détermine deux éléments importants du modèle de choix d'assolement. Tout d'abord, r_{it} détermine l'ensemble des cultures non assolées par *i* en *t*. Ensuite, et c'est là un point essentiel, le régime r_{it} détermine la forme du modèle des choix d'assolement des cultures assolées par *i* en *t*. En effet, si l'agriculteur produit la culture *k*, son choix d'assolement de *k* réagira de manière différente aux variations de $\bar{\boldsymbol{\pi}}_{it}$ selon le régime de production qu'il a choisi. L'assolement de *k* ne dépend pas des marges brutes espérées des cultures non assolées, *i.e.* de $\bar{\boldsymbol{\pi}}_{it}(r_{it}^0)$. De même, le choix d'assolement de la culture *k* apparaît, en général, d'autant moins réactif aux variations des marges des cultures produites que ces cultures sont peu nombreuses.

L'écriture de la vraisemblance en $(\mathbf{H}, \mathbf{b}, \boldsymbol{\Psi})$ de $\mathbf{s}_{it}$ conditionnelle en $\bar{\boldsymbol{\pi}}_{it}$ nécessite d'introduire quelques notations.

Le vecteur $\mathbf{s}_{it}(r_{it}^+)$ est le sous-vecteur de $\mathbf{s}_{it}$ ne contenant que les éléments non nuls de $\mathbf{s}_{it}$, et $\mathbf{s}_{it}(r_{it}^0)$ est le sous-vecteur de $\mathbf{s}_{it}$ ne contenant que les éléments nuls de $\mathbf{s}_{it}$. De manière similaire, le vecteur $\boldsymbol{\beta}_{it}^s(r_{it}^+)$ est le sous-vecteur de $\boldsymbol{\beta}_{it}^s$ ne contenant que les éléments correspondant aux éléments non nuls de $\mathbf{s}_{it}$, et $\boldsymbol{\beta}_{it}^s(r_{it}^0)$ est le sous-vecteur de $\boldsymbol{\beta}_{it}^s$ ne contenant que les éléments correspondant aux éléments nuls de $\mathbf{s}_{it}$. Les vecteurs de paramètres $\mathbf{b}(r_{it}^+)$ et $\mathbf{b}(r_{it}^0)$ sont obtenus de manière similaire à partir du vecteur **b**, de même que les vecteurs

de marges brutes $\overline{\boldsymbol{\pi}}_{it}(r_{it}^{+})$ et $\overline{\boldsymbol{\pi}}_{it}(r_{it}^{0})$ le sont à partir de $\overline{\boldsymbol{\pi}}_{it}$.

$\mathbf{H}(r_{it}^{++})$ est la matrice de paramètres obtenue en éliminant de $\mathbf{H}$ les lignes et colonnes correspondant aux éléments nuls de $\mathbf{s}_{it}$. $\mathbf{H}(r_{it}^{00})$ est la matrice obtenue en éliminant de $\mathbf{H}$ les lignes et colonnes correspondant aux éléments non nuls de $\mathbf{s}_{it}$. $\mathbf{H}(r_{it}^{0+})$ est la matrice obtenue en éliminant de $\mathbf{H}$ les lignes correspondant aux éléments non nuls de $\mathbf{s}_{it}$ et les colonnes correspondant aux éléments nuls de $\mathbf{s}_{it}$, et $\mathbf{H}(r_{it}^{+0}) = \mathbf{H}(r_{it}^{0+})'$. Les matrices $\boldsymbol{\Psi}^{s}(r_{it}^{++})$, $\boldsymbol{\Psi}^{s}(r_{it}^{00})$, $\boldsymbol{\Psi}^{s}(r_{it}^{0+})$ et $\boldsymbol{\Psi}^{s}(r_{it}^{+0})$ sont définies de manière similaire à partir de la matrice $\boldsymbol{\Psi}^{s}$.

Munis de ces notations, nous pouvons réécrire les conditions de KKT sous la forme

$$(7) \qquad \begin{cases} \mathbf{s}_{it}(r_{it}^{+}) = \mathbf{H}_{0}(r_{it}^{++})^{-1}\left(\overline{\boldsymbol{\pi}}_{it}(r_{it}^{+}) - \boldsymbol{\beta}_{it}^{s}(r_{it}^{+})\right) > \mathbf{0} \\ \overline{\boldsymbol{\pi}}_{it}(r_{it}^{+}) - \mathbf{H}_{0}(r_{it}^{0+})\mathbf{s}_{it}(r_{it}^{+}) < \boldsymbol{\beta}_{it}^{s}(r_{it}^{0}) \text{ et } \mathbf{s}_{it}(r_{it}^{0}) = \mathbf{0} \end{cases}$$

La forme du modèle de $\mathbf{s}_{it}(r_{it}^{+})$ illustre un point important évoqué plus haut : le modèle de choix d'assolement des cultures produites dépend du régime de production r_{it} dans lequel ces cultures sont produites. En particulier, l'assolement des cultures produites $\mathbf{s}_{it}(r_{it}^{+})$ ne dépend pas des marges brutes espérées des cultures non assolées, *i.e.* de $\overline{\boldsymbol{\pi}}_{it}(r_{it}^{0})$. De même, ces conditions de KKT montrent que le choix du régime de production est endogène, au sens économique du terme parce que ce choix dépend des paramètres économiques que sont les marges brutes , mais également au sens économétrique du terme dans la mesure où ce choix dépend des effets des facteurs inobservables captés par le terme $\boldsymbol{\beta}_{it}$.

La densité de $\mathbf{s}_{it}$ conditionnelle en $\overline{\boldsymbol{\pi}}_{it}$, $f(\mathbf{s}_{it} \mid \overline{\boldsymbol{\pi}}_{it})$, fournit la base de la détermination de

la vraisemblance en $(\mathbf{H}, \mathbf{b}, \mathbf{\Psi}^s)$ de $\mathbf{s}_{it}$ conditionnelle en $\overline{\boldsymbol{\pi}}_{it}$. La loi de Bayes et le point de masse de $\mathbf{s}_{it}^0$ impliquent que :

(8) $$f(\mathbf{s}_{it} \mid \overline{\boldsymbol{\pi}}_{it}) = f(\mathbf{s}_{it}(r_{it}^+) \mid \overline{\boldsymbol{\pi}}_{it}) P[\mathbf{s}_{it}(r_{it}^0) = \mathbf{0} \mid \overline{\boldsymbol{\pi}}_{it}, \mathbf{s}_{it}^+].$$

Les KKT caractérisant $\mathbf{s}_{it}$ et la loi de $\boldsymbol{\beta}_{it} \mid \overline{\boldsymbol{\pi}}_{it}$ donnent que :

(9) $$\mathbf{s}_{it}(r_{it}^+) \mid \overline{\boldsymbol{\pi}}_{it} \sim \mathcal{N}\left(\mathbf{H}_0(r_{it}^{++})^{-1}\left(\overline{\boldsymbol{\pi}}_{it}(r_{it}^+) - \mathbf{b}_0(r_{it}^+)\right), \mathbf{H}_0(r_{it}^{++})^{-1}\mathbf{\Psi}_0^s(r_{it}^{++})^{-1}\mathbf{H}_0(r_{it}^{++})^{-1}\right)$$

et :

(10) $$\boldsymbol{\beta}_{it}(r_{it}^0) \mid (\overline{\boldsymbol{\pi}}_{it}, \mathbf{s}_{it}(r_{it}^+)) \sim \mathcal{N}\left(\mathbf{b}_0(r_{it}^{0|+}), \mathbf{\Psi}_0^s(r_{it}^{0|+})\right)$$

avec :

(11) $$\mathbf{b}_0(r_{it}^{0|+}) \equiv \mathbf{b}_0(r_{it}^0) + \mathbf{\Psi}_0^s(r_{it}^{0+})\mathbf{\Psi}_0^s(r_{it}^{++})^{-1}\mathbf{b}_0(r_{it}^+)$$

et :

(12) $$\mathbf{\Psi}_0^s(r_{it}^{0|+}) \equiv \mathbf{\Psi}_0^s(r_{it}^{00}) - \mathbf{\Psi}_0^s(r_{it}^{0+})\mathbf{\Psi}_0^s(r_{it}^{++})^{-1}\mathbf{\Psi}_0^s(r_{it}^{+0}).$$

Le modèle de $\mathbf{s}_{it}(r_{it}^+)$ ne dépend pas directement de $\boldsymbol{\beta}_{it}(r_{it}^0)$, la partie aléatoire des coûts spécifiques aux cultures non assolées dans le régime r_{it}, mais ces variables aléatoires sont statistiquement liées dès lors que les termes $\boldsymbol{\beta}_{it}(r_{it}^0)$ et $\boldsymbol{\beta}_{it}(r_{it}^+)$ ne sont pas indépendants. Ceci illustre un point essentiel. Il est impossible d'analyser les choix d'assolement des cultures produites sans tenir compte de ce que les agriculteurs ont décidé de les assoler. En particulier, toute sélection d'un sous-échantillon sur les décisions de produire ou non certaines cultures, *i.e.* selon les régimes de production, est nécessairement endogène. En effet, aucun argument ne peut être avancé pour supposer *a priori* que les termes $\boldsymbol{\beta}_{it}(r_{it}^0)$ et $\boldsymbol{\beta}_{it}(r_{it}^+)$ sont

indépendants.[15] Ceci implique qu'une analyse pertinente des choix d'assolement nécessite de tenir compte, de manière adéquate, des décisions d'assoler ou non les cultures qui font partie de son ensemble de choix, *i.e.* de l'endogénéité du choix de régime r_{it}.

Ces résultats nous permettent finalement d'écrire la vraisemblance en $(\mathbf{H},\mathbf{b},\mathbf{\Psi}^s)$ de $\mathbf{s}_{it}$ conditionnelle en $\overline{\boldsymbol{\pi}}_{it}$, $\ell^s_{it}(\mathbf{H},\mathbf{b},\mathbf{\Psi}^s)$, sous la forme :

$$\begin{aligned}\ell^s_{it}(\mathbf{H},\mathbf{b},\mathbf{\Psi}) = \varphi\Big(&\mathbf{s}_{it}(r^+_{it}) - \mathbf{H}(r^{++}_{it})^{-1}(\overline{\boldsymbol{\pi}}_{it}(r^+_{it}) - \mathbf{b}(r^+_{it})); \mathbf{H}(r^{++}_{it})^{-1}\mathbf{\Psi}^s(r^{++}_{it})^{-1}\mathbf{H}(r^{++}_{it})^{-1}\Big) \\ &\times\Big[1 - \Phi\Big(\overline{\boldsymbol{\pi}}_{it}(r^0_{it}) - \mathbf{H}(r^{0+}_{it})\mathbf{s}_{it}(r^+_{it}) - \mathbf{b}(r^{0|+}_{it}); \mathbf{\Psi}^s(r^{0|+}_{it})\Big)\Big]\end{aligned} \tag{13}$$

où $\varphi(\mathbf{u};\mathbf{\Omega})$ désigne la fonction de densité en $\mathbf{u}$ de la loi normale multivariée $\mathcal{N}(\mathbf{0},\mathbf{\Omega})$ et $\Phi(\mathbf{u};\mathbf{\Omega})$ sa fonction de répartition. La forme fonctionnelle de $\ell^s_{it}(\mathbf{H},\mathbf{b},\mathbf{\Psi}^s)$ appelle trois remarques.

Tout d'abord, elle est relativement « pénible » à calculer car sa forme fonctionnelle dépend fortement du régime de production choisi par l'agriculteur. Ceci implique qu'il est nécessaire de définir autant de formes fonctionnelles de fonction de vraisemblance qu'il existe de régimes de production dans l'échantillon considéré. C'est en fait la principale différence qui existe entre les modèles de choix d'assolement définis comme des modèles à changement de régime endogène, comme celui décrit ici, et ceux définis comme des modèles de régressions censurées. Dans les deux cas le modèle tient compte de l'existence de régime de production,

[15] Ce qui conduirait in fine à supposer que les éléments de $\boldsymbol{\beta}_{it}$ sont mutuellement indépendants entre eux.

mais un modèle de régression censurée ignore l'impact du régime de production sur la forme des choix d'assolement des cultures assolées.

Ensuite, le terme $\Phi\left(\bar{\boldsymbol{\pi}}_{it}(r_{it}^{0}) - \mathbf{H}(r_{it}^{0+})\mathbf{s}_{it}(r_{it}^{+}) - \mathbf{b}(r_{it}^{0|+}); \boldsymbol{\Psi}^{s}(r_{it}^{0|+})\right)$ est la valeur d'une fonction de répartition de loi normale multivariée dont la dimension correspond au nombre de cultures non assolées. De telles valeurs ne peuvent être intégrées numériquement que pour des lois normales de dimension 3 au maximum. Des techniques d'intégration par simulations sont nécessaires dès lors que la dimension dépasse 3. Malgré l'augmentation des puissances de calcul et l'efficacité des simulateurs GHK[16], l'évaluation de termes de ce type a tendance à sérieusement compliquer l'estimation de modèles multivariés à changement de régime endogène tels que le modèle de choix d'assolement présenté ici. De fait, les modèles multivariés de dimension importante ont rarement été employés, les travaux de Kao *et al* (2001) sur les systèmes de fonctions de demandes de biens de consommation et de Chakir *et al* (2004) sur ceux d'intrants faisant ici figure d'exceptions.[17] Il est également intéressant de noter que l'estimation de systèmes de régressions censurées de dimension importante requiert l'évaluation de ce type de termes (voir, *e.g.*, Sckokaï et Moro, 2006, Lacroix et Thomas, 2011 ; Fezzi et Bateman, 2011, Platoni *et al*, 2012).

---

[16] Pour Geweke, Hajivassiliou et Keane, les trois auteurs qui l'ont proposé, simultanément mais séparément, ce simulateur de Gibbs exploitant l'additivité des lois normales et les propriétés de la décomposition de Cholesky des matrices définies-positives. Voir, *e.g.*, Train (2009) pour une présentation très claire de ce simulateur.

[17] Lichtenberg et Smith-Ramirez (2011) estiment un autre type de modèle multivarié à changement de régime endogène par des algorithmes MCEM, les régimes considérés étant dichotomiques.

Enfin, ce type de modélisation ne permet pas de déterminer, aisément tout au moins, les probabilités de choix de régime de production. Les conditions de KKT caractérisent les choix d'assolement dans un régime donné. Ceci peut expliquer l'utilisation des systèmes de régressions censurées. La construction de ces modèles repose en grande partie sur l'existence d'une fonction de probabilité de choix de régime.

3.1.2. Coûts fixes d'assolement et de régime de production

Les conditions de KKT caractérisent les choix d'assolement selon une analyse marginale classique. Le modèle décrit repose sur une fonction de profit continue, et « lisse », en **s**. Il autorise des choix de surface potentiellement très petits pourvu que leur profit marginal soit nul. Or, tout au moins dans les systèmes de grandes cultures et cultures industrielles, les agriculteurs dédient des surfaces conséquentes à la plupart des cultures de leur assolement.

Ceci peut s'expliquer par le fait qu'une culture impose des coûts et des contraintes dès lors qu'elle est assolée. En effet, quelle que soit la surface de la culture considérée, l'agriculteur doit s'assurer de disposer de suffisamment de main d'œuvre et de matériel aux moments clés du cycle de production de cette culture. Il doit également se procurer les intrants nécessaires et s'occuper de la commercialisation des quantités produites. Ces coûts sont fixes par rapport à la surface assolée dans la mesure où ils apparaissent dès que l'agriculteur a décidé de produire la culture considérée. Il est d'ailleurs à noter que les coûts fixes liés à la production d'une culture peuvent dépendre du régime de production dans lequel cette production est envisagée. En effet, introduire une culture dans l'assolement ne pose pas de problème

spécifique de disponibilité de main d'œuvre ou de matériel lorsque le cycle biologique de cette culture est décalé par rapport aux autres cultures de l'assolement.

Il est important de distinguer les couts fixes de régime considérés ici des coûts d'investissement liés à la production d'une nouvelle culture. L'agriculteur subit les coûts fixes de régime chaque année. Ce sont des coûts de mise en œuvre.

De fait, nous supposons ici, et c'est une hypothèse relativement forte, que chaque agriculteur de l'échantillon considéré sait produire et dispose des facteurs de production quasi-fixes – parce qu'il les possède ou peut les louer – nécessaires à chaque culture de l'ensemble $\mathcal{K}$. Nous supposons donc que le choix de produire ou non une culture de l'ensemble $\mathcal{K}$ est une décision de court terme. La très grande variabilité des régimes de production, entre agriculteurs et dans le temps pour un même agriculteur, constatée dans les données procure des arguments en faveur de cette hypothèse dans l'application présentée dans le Chapitre 6.

Introduire ce type de coûts fixes de régime dans le modèle considéré jusqu'à présent s'avère relativement compliqué. En effet, cela suppose d'introduire un coût fixe pour chaque régime dans le problème de maximisation de profit espéré, ce que nous présentons ici.

L'ensemble des régimes de production potentiels du problème de choix d'assolement considéré est défini par $\mathcal{R} = \{1,...,R\}$. Les coûts fixes des régimes sont définis par les termes $D_{it}(r)$ pour $r \in \mathcal{R}$ et définissent la collection $\mathcal{D} = \{D_{it}(1),...,D_{it}(R)\}$. Enfin, la fonction $\rho(\mathbf{s})$ définit le régime de l'assolement $\mathbf{s}$. Cette fonction permet de sélectionner le coût du régime

de $\mathbf{s}$ dans la collection de coûts $\mathcal{D}$. Le terme $\mathcal{K}^{+}_{-K}(r)$ définit l'ensemble des cultures produites, culture K exclue, dans le régime r, et $\mathcal{K}^{0}(r)$ l'ensemble des cultures non produites dans le régime r.

Tenir compte des coûts fixes de régime dans le modèle « à la Wales et Woodland » considéré jusqu'ici consiste à considérer qu'en t l'agriculteur i choisit son assolement $\mathbf{s}_{it}$ en résolvant le problème suivant :

$$\text{(14)} \qquad \max_{\mathbf{s}} \left\{ \begin{array}{c} \sum_{k \in \mathcal{K}_{-K}} s_k (\pi_{k,it} - \pi_{K,it} - \beta^{s}_{k,it}) - C^{s}_{i}(\mathbf{s}) - D_{it}(\rho(\mathbf{s})) \\ \text{s.c.} \\ \sum_{k \in \mathcal{K}_{-K}} s_k \leq 1 \text{ et } s_k \geq 0 \text{ pour } k \in \mathcal{K}_{-K} \end{array} \right\}.$$

La « fonction de coût fixe de régime » $D_{it}(\rho(\mathbf{s}))$ introduit une discontinuité saillante en $\mathbf{s}$ dans la fonction objectif de ce problème de maximisation. La valeur de $D_{it}(\rho(\mathbf{s}))$ passe d'un élément de $\mathcal{D}$ à un autre dès qu'un élément de $\mathbf{s}$ n'est plus nul ou n'est plus strictement positif. Par conséquent, on ne peut pas de caractériser les solutions en $\mathbf{s}$ de ce problème par des conditions de KKT usuelles. De fait, le terme $D_{it}(\rho(\mathbf{s}))$ introduit des éléments de choix discret dans le choix d'assolement. De plus, il est difficile de résoudre, voire de caractériser de manière simple, la solution en $\mathbf{s}$ de ce problème.

On peut être tenté de résoudre R problèmes de maximisation de profit en imposant dans chacun d'entre eux une contrainte imposant que le choix d'assolement optimal appartient à un régime prédéterminé, *i.e.* la contrainte $\rho(\mathbf{s}) = r$ pour $r \in \mathcal{R}$. La résolution de ces

problèmes permet de définir le régime optimal en comparant les niveaux de profit obtenus pour chaque régime de production imposé. Cette approche repose sur une décomposition « naturelle » du problème de maximisation du profit avec coûts fixes de régime, avec :

$$\text{(15)} \qquad \max_{r\in\mathcal{R}}\left\{\max_{\mathbf{s}}\left\{\begin{array}{c}\sum_{k\in\mathcal{K}_{-K}} s_k(\pi_{k,it}-\pi_{K,it}-\beta^s_{k,it})-C^s_i(\mathbf{s})\\ \text{s.c.}\\ \sum_{k\in\mathcal{K}_{-K}} s_k\leq 1 \text{ et } s_k=0 \text{ pour } k\in\mathcal{K}^0(r)\\ s_k>0 \text{ pour } k\in\mathcal{K}^+_{-K}(r)\end{array}\right\}-D_{it}(r)\right\}$$

Cette décomposition permet d'éliminer le problème de discontinuité lié au terme $D_{it}(\rho(\mathbf{s}))$, le coût fixe de régime n'apparaissant que dans le problème du choix de régime, en second étape.

Néanmoins, cette décomposition s'avère peu opérationnelle en pratique, en particulier pour spécifier des modèles économétriques de choix d'assolement. Tout d'abord, imposer la « contrainte de régime » $\rho(\mathbf{s})=r$ suppose d'imposer les contraintes $s_k=0$ pour $k\in\mathcal{K}^0(r)$ d'une part, et $s_k>0$ pour $k\in\mathcal{K}^+_{-K}(r)$ d'autre part. Cette dernière série de contraintes est composée des contraintes de positivité stricte qu'il est impossible d'imposer en pratique, sauf à utiliser des astuces numériques – telles que l'ajout de termes de perturbation – qui peuvent sérieusement compliquer la résolution du problème. De fait, dans certaines configurations imposer un régime de production revient à forcer un agriculteur à allouer une quantité infinitésimale de terre à une culture qu'il n'a pas envie d'assoler. Ce problème est une manifestation des discontinuités sévères de $D_i(\rho(\mathbf{s}))$ en $\mathbf{s}$.

Ensuite, quelle que soit l'approche adoptée pour les obtenir, les solutions des problèmes de maximisation de profit avec contraintes de régime n'ont en général pas de solution

analytique simple. Ceci pose de sérieux problèmes lorsqu'on envisage d'utiliser le modèle considéré pour spécifier un modèle micro-économétrique de choix d'assolement.

Dans la suite de cette section nous montrons que les modèles de choix d'assolement MNL de Carpentier et Letort (2014) offrent des solutions simples aux problèmes mis en évidence ici.

3.1.3. Modèles de choix d'assolement MNL, choix de cultures et coûts fixes de régime

L'objectif de cette sous-section est de montrer que les propriétés spécifiques des modèles de choix d'assolement MNL permettent de spécifier des modèles simples mais suffisamment riches pour tenir compte des décisions de produire ou non certaines cultures et de coûts fixes de régime. Ces propriétés sont dues aux termes d'entropie composant les fonctions de coût implicite de gestion de l'assolement sous-tendant les modèles de choix MNL.

Nous considérons ici des modèles de choix d'assolement SMNL afin de simplifier la présentation. Mais l'approche de modélisation présentée avec les modèles SMNL s'étend aux modèles NMNL, quoiqu'avec des calculs (et des notations) plus pénibles et des expressions plus compliquées. Le Chapitre 6 présente une application de l'approche de modélisation présentée ici avec un modèle de choix NMNL à trois niveaux d'emboîtement.

Le modèle considéré sera construit à partir de la fonction de coût implicite :

(16) $$C_{it}(\mathbf{s}) = D_{it}(\rho(\mathbf{s})) + \sum_{k\in\mathcal{K}} s_k \beta^s_{k,it} + (\alpha^s_i)^{-1} \sum_{k\in\mathcal{K}} s_k \ln s_k \quad \text{où} \quad s_K = 1 - \sum_{k\in\mathcal{K}_{-K}} s_k \,,$$

la contrainte de normalisation $\beta^s_{K,it} \equiv 0$ étant imposée. Nous supposons que la culture K est

produite dans tous les régimes de $\mathcal{R}$. Malgré la présence de termes en logarithme de parts de surface, imposer des contraintes d'assolements nuls dans cette fonction ne pose pas de problèmes. En effet, la continuité sur $\mathbb{R}_*^+$ des termes $s_k \ln s_k$ peut être étendue en 0, cette extension de continuité justifiant par ailleurs la convention usuelle donnée par la formule « $0 \times \ln 0 = 0$ ». En utilisant la variable indicatrice :[18]

(17) $$\iota_k(r) = \begin{cases} 1 \text{ si } k \in \mathcal{K}^+(r) \\ 0 \text{ si } k \in \mathcal{K}^0(r) \end{cases}$$

où le terme $\mathcal{K}^+(r)$ définit l'ensemble des cultures produites, culture K inclue, dans le régime r, la fonction peut être réécrite sous la forme :

(18) $C_{it}(\mathbf{s}) = D_{it}\left(\rho(\mathbf{s})\right) + \sum_{k \in \mathcal{K}} s_k \beta_{k,it}^s + (\alpha_i^s)^{-1} \sum_{k \in \mathcal{K}} \iota_k\left(\rho(\mathbf{s})\right) s_k \ln s_k$ où $s_K = 1 - \sum_{k \in \mathcal{K}_{-K}} s_k$.

Le modèle de choix d'assolement MNL avec choix de régime et coûts fixes de régime est défini par la solution en $\mathbf{s}$ du problème de maximisation de profit :

(19) $$\max_{\mathbf{s} \geq 0} \left\{ \begin{array}{c} \sum_{k \in \mathcal{K}_{-K}} s_k (\pi_{k,it} - \pi_{K,it} - \beta_{k,it}^s) - (\alpha_i^s)^{-1} \sum_{k \in \mathcal{K}} \iota_k\left(\rho(\mathbf{s})\right) s_k \ln s_k - D_{it}(\rho(\mathbf{s})) \\ \text{s.c.} \\ s_K = 1 - \sum_{k \in \mathcal{K}_{-K}} s_k \text{ et } s_k \geq 0 \text{ pour } k \in \mathcal{K} \end{array} \right\} .$$

Ce problème peut être décomposé en :

(20) $$\max_{r \in \mathcal{R}} \left\{ \Pi_{it}^o(\boldsymbol{\pi}_{it}; r) - D_{it}(r) \right\}$$

---

[18] Qui s'avère très pratique dans les cas NMNL.

où :

$$(21) \qquad \Pi_{it}^{o}(\boldsymbol{\pi}_{it};r) \equiv \max_{\mathbf{s}\geq 0}\left\{\begin{array}{c} \sum_{k\in\mathcal{K}_{-K}} s_k(\pi_{k,it}-\pi_{K,it}-\beta_{k,it}^{s})-(\alpha_i^{s})^{-1}\sum_{k\in\mathcal{K}} \iota_k(r)s_k \ln s_k \\ \text{s.c.} \\ s_K = 1-\sum_{k\in\mathcal{K}_{-K}} s_k \ \text{ et } \ s_k = 0 \text{ pour } k\in\mathcal{K}^{0}(r) \\ s_k \geq 0 \text{ pour } k\in\mathcal{K}^{+}(r) \end{array}\right\}$$

La forme entropique du terme de diversification de l'assolement $\sum_{k\in\mathcal{K}} \iota_k(r)s_k \ln s_k$, assure que les contraintes de positivité des parts de surface des cultures assolées dans le régime *r*, *i.e.* les contraintes $s_k \geq 0$ pour $k\in\mathcal{K}^{+}(r)$, ne sont jamais saturées.[19] Dit autrement, maximiser une fonction de profit MNL en imposant un régime de production conduit à un choix d'assolement optimal dans ce régime de production. La solution en **s** de ce problème est donnée par :

$$(22) \qquad s_{k,it}^{o}(\boldsymbol{\pi}_{it};r) = \frac{\exp\left(\alpha_i^{s}(\pi_{k,it}-\beta_{k,it}^{s})\right)}{\sum_{\ell\in\mathcal{K}} \iota_\ell(r)\exp\left(\alpha_i^{s}(\pi_{\ell,it}-\beta_{\ell,it}^{s})\right)} \ \text{pour} \ \ k\in\mathcal{K}^{+}(r)$$

et :

$$(23) \qquad s_{k,it}^{o}(\boldsymbol{\pi}_{it};r) = 0 \ \ \text{pour} \ \ k\in\mathcal{K}^{0}(r).$$

Les cultures produites dans le régime *r* ont une forme de choix d'assolement SMNL, les surfaces allouées aux autres cultures sont nulles par définition du régime. De même, les fonctions de profit indirect $\Pi_{it}^{o}(\boldsymbol{\pi}_{it};r)$ ont une forme analytique simple, avec :

---

[19] Ceci provient de ce que la dérivée de $s_k \ln s_k$ en s_k tend vers $-\infty$ lorsque s_k tend vers 0^{+}.

(24) $$\Pi_{it}^{o}(\boldsymbol{\pi}_{it};r) = (\alpha_i^s)^{-1} \ln \sum_{\ell \in \mathcal{K}} \iota_\ell(r) \exp\left(\alpha_i^s(\pi_{\ell,it} - \beta_{\ell,it}^s)\right).$$

Ces fonctions de profit indirectes sont lisses dans leurs arguments, et il est aisé de vérifier que :

(25) $$\tfrac{\partial}{\partial \pi_k} \Pi_{it}^{o}(\boldsymbol{\pi}_{it};r) = \iota_k(r) s_{k,it}^{o}(\boldsymbol{\pi}_{it};r).$$

Cette propriété est utile dans la procédure d'estimation présentée dans le chapitre suivant (lors de la mise en œuvre de l'approximation de Laplace des probabilités de choix de régime).

Le problème du choix de régime, $\max_{r \in \mathcal{R}} \left\{\Pi_{it}^{o}(\boldsymbol{\pi}_{it};r) - D_{it}(r)\right\}$, est un simple problème de choix discret, celui du choix du régime de production le plus profitable, coûts fixes de régime inclus.

Ces résultats nous permettent d'obtenir le modèle de choix d'assolement SMNL avec choix de régime de production endogènes et coûts fixes de régime suivant :

(26) $$\begin{cases} r_{it} = \arg\max_{r \in \mathcal{R}} \left\{(\alpha_i^s)^{-1} \ln \sum_{\ell \in \mathcal{K}} \iota_\ell(r) \exp\left(\alpha_i^s(\pi_{\ell,it} - \beta_{\ell,it}^s)\right) - D_{it}(r)\right\} \\ s_{k,it} = \dfrac{\exp\left(\alpha_i^s(\pi_{k,it} - \beta_{k,it}^s)\right)}{\sum_{\ell \in \mathcal{K}} \iota_\ell(r_{it}) \exp\left(\alpha_i^s(\pi_{\ell,it} - \beta_{\ell,it}^s)\right)} \quad \text{pour } k \in \mathcal{K}^+(r_{it}) \end{cases}$$

ou encore, en forme log-linéarisée :

(27) $$\begin{cases} r_{it} = \arg\max_{r \in \mathcal{R}} \left\{(\alpha_i^s)^{-1} \ln \sum_{\ell \in \mathcal{K}} \iota_\ell(r) \exp\left(\alpha_i^s(\pi_{\ell,it} - \beta_{\ell,it}^s)\right) - D_{it}(r)\right\} \\ \ln s_{k,it} - \ln s_{K,it} = \alpha_i^s(\pi_{k,it} - \pi_{K,it} - \beta_{k,it}^s) \quad \text{pour } k \in \mathcal{K}_{-K}^+(r_{it}) \end{cases}$$

Le fait que $s_{k,it} = 0$ pour $k \in \mathcal{K}^0(r)$ est une information apportée par le contenu du régime r_{it}. Ce modèle peut servir de base à des modèles micro-économétriques de choix de

production multiculture à choix de régime de production endogène et avec coûts fixes de production.

La version NMNL à deux niveaux d'emboîtement est donnée par :

$$
(28)\quad \begin{cases}
r_{it} = \arg\max_{r\in\mathcal{R}} \left\{ \Pi_{it}^{o}(\boldsymbol{\pi}_{it};r) - D_{it}(r) \right\} \\[2ex]
s_{k,it} = \dfrac{\exp\left(\alpha_{(g),i}^{s}(\pi_{k,it} - \beta_{k,it}^{s})\right)\left(\sum_{\ell\in\mathcal{K}(g)} \iota_{\ell}(r_{it})\exp\left(\alpha_{(g),i}^{s}(\pi_{\ell,it} - \beta_{\ell,it}^{s})\right)\right)^{\alpha_i^s(\alpha_{(g),i}^s)^{-1}-1}}{\exp\left(\alpha_i^s \Pi_{it}^{o}(\boldsymbol{\pi}_{it};r_{it})\right)} \\
\text{pour } k\in\mathcal{K}^{+}(r_{it}) \text{ et } k\in\mathcal{K}_{(g)} \\[2ex]
\Pi_{it}^{o}(\boldsymbol{\pi}_{it};r) = (\alpha_i^s)^{-1}\ln\sum_{h\in\mathcal{G}} \iota_{(g)}(r)\left(\sum_{\ell\in\mathcal{K}(h)} \iota_{\ell}(r)\exp\left(\alpha_{(h),i}^{s}(\pi_{\ell,it} - \beta_{\ell,it}^{s})\right)\right)^{\alpha_i^s(\alpha_{(h),i}^s)^{-1}} \\
\text{pour } r\in\mathcal{R}
\end{cases}
$$

ou encore, en forme log-linéarisée :

$$
(29)\quad \begin{cases}
r_{it} = \arg\max_{r\in\mathcal{R}} \left\{ \Pi_{it}^{o}(\boldsymbol{\pi}_{it};r) - D_{it}(r) \right\} \\[2ex]
\ln s_{k,it} - \ln s_{K,it} = \alpha_i^s \times \left\{ \begin{array}{c} \pi_{k,it} - \pi_{K,it} - \beta_{k,i}^{s} - (\boldsymbol{\delta}_{k,0}^{s})'\mathbf{c}_{k,it}^{s} \\ +\left((\alpha_i^s)^{-1} - (\alpha_{(g),i}^{s})^{-1}\right)\times \ln s_{k|(g),it} \\ -\left((\alpha_i^s)^{-1} - (\alpha_{(G),i}^{s})^{-1}\right)\times \ln s_{K|(G),it} \end{array} \right\} + \alpha_i^s \varepsilon_{k,it}^{s} \\
\text{pour } k\in\mathcal{K}_{-K}^{+}(r_{it}) \text{ et } k\in\mathcal{K}_{(g)} \\[2ex]
\Pi_{it}^{o}(\boldsymbol{\pi}_{it};r) = (\alpha_i^s)^{-1}\ln\sum_{h\in\mathcal{G}} \iota_{(g)}(r)\left(\sum_{\ell\in\mathcal{K}(h)} \iota_{\ell}(r)\exp\left(\alpha_{(h),i}^{s}(\pi_{\ell,it} - \beta_{\ell,it}^{s})\right)\right)^{\alpha_i^s(\alpha_{(h),i}^s)^{-1}} \\
\text{pour } r\in\mathcal{R}
\end{cases}
$$

Dans la section suivante nous utilisons ces extensions des choix d'assolement MNL pour spécifier des modèles micro-économétriques paramétriques de choix de production multiculture avec paramètres aléatoires et choix de cultures. Dans le chapitre suivant nous proposons un algorithme SAEM pour l'estimer dans le cadre du MV. Ces modèles à paramètres aléatoires et ces algorithmes sont des extensions, avec choix de régime de production endogène et avec coûts fixes de production, de ceux présentés dans les chapitres précédents.

3.2. Modèles micro-économétriques de choix de production multiculture à paramètres aléatoires et avec choix de cultures : spécification

Dans cette section nous « couplons » les modèles d'offre de rendement et d'utilisation d'intrants présentés dans les chapitres précédents d'une part, et les modèles de choix d'assolement avec choix de régime de production endogène présentés dans la section précédente d'autre part, pour définir des modèles de choix de production multiculture à paramètres aléatoires, à choix de régime de production endogène et avec coûts fixes de régime.

Nous présentons d'abord la forme fonctionnelle de ces modèles, puis leur spécification

paramétrique et enfin leur vraisemblance. Les modèles présentés ici s'appuient largement sur des éléments donnés et décrits dans les chapitres et la section précédente. La vraisemblance des modèles considérés fera l'objet d'une attention particulière dans la mesure où elle est de forme assez complexe et pose des problèmes spécifiques de maximisation.

3.2.1. Modèles micro-économétriques de choix de production multiculture à paramètres aléatoires et avec choix de cultures : formes fonctionnelles

Nous reprenons ici les formes des modèles d'offre de rendement et de demande d'intrants à paramètres aléatoires considérés dans la partie précédente :

(30) $$\begin{cases} y_{k,it} = \beta^y_{k,i} + (\boldsymbol{\delta}^y_{k,0})'\mathbf{c}^y_{k,it} - \alpha^x_{k,i} q^2_{k,it} / 2 + \varepsilon^y_{k,it} \\ x_{k,it} = \beta^x_{k,i} + (\boldsymbol{\delta}^x_{k,0})'\mathbf{c}^x_{k,it} - \alpha^x_{k,i} q^2_{k,it} + \varepsilon^x_{k,it} \end{cases} \text{ pour } k \in \mathcal{K}$$

qui nous permettent de définir les marges brutes espérées :

(31) $\pi_{k,it} = p_{k,it}\left(\beta^y_{k,i} + (\boldsymbol{\delta}^y_{k,0})'\mathbf{c}^{ys}_{k,it}\right) - v_{k,it}\left(\beta^x_{k,i} + (\boldsymbol{\delta}^x_{k,0})'\mathbf{c}^{xs}_{k,it}\right) + \alpha^x_{k,i} q^2_{k,it} p_{k,it} / 2 + a_{k,it}$ pour $k \in \mathcal{K}$,

lesquelles sont utilisées en tant qu'arguments des modèles de choix d'assolement et de régime de production.

Il est important de rappeler ici que les anticipations des termes d'erreur $\varepsilon^y_{k,it}$ et $\varepsilon^x_{k,it}$ sont supposées être égales à 0 au moment des choix d'assolement, et donc des choix de régime.

Les modèles de choix d'assolement MNL considérés dans la section suivante sont complétés par la spécification des termes $\beta^s_{k,it}$ utilisée dans les chapitres précédents :

(32) $$\beta^s_{k,it} = (\boldsymbol{\delta}^s_{k,0})'\mathbf{c}^s_{k,it} + \varepsilon^s_{k,it} \text{ pour } k \in \mathcal{K}_{-K} \text{ et } \beta^s_{K,it} \equiv 0.$$

Il est également important de rappeler ici que les termes (d'erreur dans le modèle économétrique) $\varepsilon^{s}_{k,it}$ et $\varepsilon^{r}_{r,it}$ sont supposés connus par les agriculteurs au moment de leurs choix d'assolement, et donc de leur choix de régime.

Le modèle de choix de régime est défini comme un modèle probabiliste en définissant les modèles de coûts de régime par :

(33) $$D_{it}(r) = d_{r,i} - \sigma_i^{-1} e_{r,it} \text{ avec } \sigma_i > 0 \text{ pour } r \in \mathcal{R} \text{ et } d_{R,i} \equiv 0$$

i.e. en considérant des sommes d'un terme spécifique à chaque exploitation, $d_{r,i}$, et d'un terme d'erreur, $\sigma_i^{-1} e_{r,it}$, dans lequel le terme σ_i^{-1} joue le rôle de facteur d'échelle spécifique à chaque exploitant. Ce facteur d'échelle est important dans la mesure où la spécification paramétrique du modèle fixe la distribution jointe des $e_{r,it}$ (à la manière d'un modèle de choix discret de type Logit Multinomial). Il mesure également l'impact des facteurs économiques représentés par les termes de profit et de coût dans les choix de régime de production des agriculteurs. Plus σ_i^{-1} est petit, plus ces facteurs jouent un rôle important dans les décisions de choix de régime de l'agriculteur *i*. La condition $d_{R,i} \equiv 0$ est une contrainte de normalisation définissant le régime *R* comme le régime de référence dans la définition des coûts fixes de régime. Les termes $e_{r,it}$ sont évidemment supposés connus par les agriculteurs au moment de leurs choix d'assolement, et donc de leur choix de régime. Ces termes d'erreur captent les effets de déterminants inobservés et variables d'une année à l'autre des choix de régime.

En considérant un modèle de choix NMNL à deux niveaux et à choix de régime de production

endogène, ces modèles donnent la forme fonctionnelle du modèle de choix de production multiculture considéré dans la suite de cette section :

$$
(34)\quad \begin{cases}
\begin{cases} y_{k,it} = \beta^{y}_{k,i} + (\boldsymbol{\delta}^{y}_{k,0})'\mathbf{c}^{y}_{k,it} - \alpha^{x}_{k,i} q^{2}_{k,it} / 2 + \varepsilon^{y}_{k,it} \\ x_{k,it} = \beta^{x}_{k,i} + (\boldsymbol{\delta}^{x}_{k,0})'\mathbf{c}^{x}_{k,it} - \alpha^{x}_{k,i} q_{k,it} + \varepsilon^{x}_{k,it} \\ \text{pour } k \in \mathcal{K}^{+}(r_{it}) \end{cases} \\
\begin{cases} s_{k,it} = \dfrac{\exp\left(\alpha^{s}_{(g),i}(\pi_{k,it} - \beta^{s}_{k,it})\right)\left(\sum_{\ell \in \mathcal{K}(g)} \iota_{\ell}(r_{it}) \exp\left(\alpha^{s}_{(g),i}(\pi_{\ell,it} - \beta^{s}_{\ell,it})\right)\right)^{\alpha^{s}_{i}(\alpha^{s}_{(g),i})^{-1}-1}}{\exp\left(\alpha^{s}_{i}\Pi_{it}(\boldsymbol{\gamma}_{i}, \boldsymbol{\delta}_{0}, \boldsymbol{\varepsilon}^{s}_{it}; r_{it})\right)} \\ \text{pour } k \in \mathcal{K}^{+}(r_{it}) \text{ et } k \in \mathcal{K}_{(g)} \end{cases} \\
r_{it} = \arg\max_{r \in \mathcal{R}} \left\{ \Pi_{it}(\boldsymbol{\gamma}_{i}, \boldsymbol{\delta}_{0}, \boldsymbol{\varepsilon}^{s}_{it}; r) - d_{r,i} + \sigma_{i}^{-1} e_{r,it} \right\} \\
\begin{cases} \pi_{k,it} = p_{k,it}\left(\beta^{y}_{k,i} + (\boldsymbol{\delta}^{y}_{k,0})'\mathbf{c}^{ys}_{k,it}\right) - v_{k,it}\left(\beta^{x}_{k,i} + (\boldsymbol{\delta}^{x}_{k,0})'\mathbf{c}^{xs}_{k,it}\right) + \alpha^{x}_{k,i} q^{2}_{k,it} p_{k,it} / 2 + a_{k,it} \\ \text{pour } k \in \mathcal{K} \end{cases} \\
\begin{cases} \beta^{s}_{k,it} = \beta^{s}_{k,i} + (\boldsymbol{\delta}^{s}_{k})'\mathbf{c}^{s}_{k,it} + \varepsilon^{s}_{k,it} \\ k \in \mathcal{K}_{-K} \\ \beta^{s}_{K,it} = 0 \end{cases} \\
\begin{cases} \Pi_{it}(\boldsymbol{\gamma}_{i}, \boldsymbol{\delta}_{0}, \boldsymbol{\varepsilon}^{s}_{it}; r) = (\alpha^{s}_{i})^{-1} \ln \sum_{h \in \mathcal{G}} \iota_{(g)}(r)\left(\sum_{\ell \in \mathcal{K}(h)} \iota_{\ell}(r) \exp\left(\alpha^{s}_{(h),i}(\pi_{\ell,it} - \beta^{s}_{\ell,it})\right)\right)^{\alpha^{s}_{i}(\alpha^{s}_{(h),i})^{-1}} \\ \text{pour } r \in \mathcal{R} \end{cases}
\end{cases}
$$

Ce modèle décrit :

- le régime choisi r_{it},
- les niveaux de rendement obtenus $\mathbf{y}^{+}_{it} \equiv \left(y_{k,it} : k \in \mathcal{K}(r_{it})\right)$,
- les quantités d'intrants utilisées $\mathbf{x}^{+}_{it} \equiv \left(x_{k,it} : k \in \mathcal{K}(r_{it})\right)$,

- les assolements choisis $\mathbf{s}_{it}^{+} \equiv (x_{k,it} : k \in \mathcal{K}_{-K}(r_{it}))$, étant entendu que $s_{K,it} > 0$

pour chaque observation (i,t) de l'échantillon considéré à partir d'un vecteur de variables explicatives $\mathbf{z}_{it}$ composé :

- des prix des cultures $\mathbf{p}_{it} \equiv (p_{k,it} : k \in \mathcal{K})$,
- des prix des intrants $\mathbf{v}_{it} \equiv (v_{k,it} : k \in \mathcal{K})$,
- des rapports de prix correspondants $\mathbf{q}_{it} \equiv (q_{k,it} : k \in \mathcal{K})$ avec $q_{k,it} \equiv v_{k,it} p_{k,it}^{-1}$,
- des variables de contrôle des niveaux de rendement $\mathbf{c}_{k,it}^{y}$, d'utilisation d'intrants $\mathbf{c}_{k,it}^{x}$ et des assolements $(\mathbf{c}_{k,it}^{s}, \mathbf{c}_{k,it}^{ys}, \mathbf{c}_{k,it}^{xs})$ collectées dans $\mathbf{c}_{it} \equiv \left((\mathbf{c}_{k,it}^{y}, \mathbf{c}_{k,it}^{x}, \mathbf{c}_{k,it}^{s}, \mathbf{c}_{k,it}^{ys}, \mathbf{c}_{k,it}^{xs}) : k \in \mathcal{K}\right)$.

Les prix des cultures et des intrants des cultures non produites par un agriculteur ne sont pas observés. Ils ont été remplacés par les moyennes annuelles correspondantes calculées à partir des prix observés. Bien entendu, cette procédure d'attribution des prix manquants pourrait être améliorée.

La spécification de ce modèle repose sur un vecteur de paramètres aléatoires $\boldsymbol{\gamma}_i$ spécifique à chaque agriculteur composé :

- du sous-vecteur $\boldsymbol{\gamma}_i^{yx} \equiv \left((\beta_{k,i}^{y}, \beta_{k,i}^{x}, \ln \alpha_{k,i}^{x}) : k \in \mathcal{K}\right)$ qui collecte les paramètres de rendements potentiels $\beta_{k,i}^{y}$, d'utilisations d'intrants associées $\beta_{k,i}^{x}$ et de flexibilité des choix d'intrants $\ln \alpha_{k,i}^{x}$,
- du sous-vecteur $\boldsymbol{\gamma}_i^{s} \equiv (\boldsymbol{\beta}_i^{s}, \ln \boldsymbol{\alpha}_i^{s})$ qui collecte les paramètres de coût linéaire des

assolements $\boldsymbol{\beta}_i^s \equiv (\beta_{k,i}^s : k \in \mathcal{K}_{-K})$ et les paramètres de flexibilité des choix d'assolement $\ln \boldsymbol{\alpha}_i^s \equiv \left(\ln \alpha_i^s, (\ln \alpha_{(g),i}^s : g \in \mathcal{G})\right)$ entre groupes de cultures, $\ln \alpha_i^s$, et au sein des groupes de cultures, $\ln \alpha_{(g),i}^s$,

- du sous-vecteur $\boldsymbol{\gamma}_i^r \equiv \left(\ln \sigma_i, (d_{r,i} : r \in \mathcal{R} \setminus \{R\})\right)$ qui collecte les coûts fixes de régime, $d_{r,i}$, et le paramètre d'échelle, σ_i, des termes d'erreur du modèle probabiliste de choix de régime

ainsi que sur un vecteur de paramètres fixes $\boldsymbol{\delta}_0$, commun à l'ensemble des agriculteur, composé :

- du vecteur des coefficients des variables de contrôle des modèles de niveaux de rendement et d'utilisation d'intrants, $\boldsymbol{\delta}_{k,0}^{yx} \equiv (\delta_{k,0}^y, \delta_{k,0}^x)$, collectés dans $\boldsymbol{\delta}_0^{yx} \equiv (\boldsymbol{\delta}_{k,0}^{yx} : k \in \mathcal{K})$
- des variables de contrôle des modèles de choix d'assolement $\boldsymbol{\delta}_0^s \equiv (\boldsymbol{\delta}_{k,0}^s : k \in \mathcal{K})$.

Les paramètres de flexibilité sont considérés en logarithme ici pour assurer leur positivité.

Ce modèle contient un vecteur de termes d'erreur $(\boldsymbol{\varepsilon}_{it}, \mathbf{e}_{it})$ composé :

- du vecteur des termes d'erreur $\boldsymbol{\varepsilon}_{it}^{yx} \equiv (\boldsymbol{\varepsilon}_{k,it}^{yx} : k \in \mathcal{K})$ des équations d'offre de rendement et de demandes d'intrants des cultures, $\boldsymbol{\varepsilon}_{k,it}^{yx} \equiv (\varepsilon_{it}^y, \varepsilon_{it}^x)$
- du vecteur des termes d'erreur $\boldsymbol{\varepsilon}_{it}^s \equiv \left(\varepsilon_{it}^s : k \in \mathcal{K}(r_{it})\right)$ des équations d'offre de rendement et de demandes d'intrants
- du vecteur des termes d'erreur $\mathbf{e}_{it} \equiv (e_{r,it} : r \in \mathcal{R})$ du modèle de choix de régime.

Le vecteur $\boldsymbol{\varepsilon}_{it}$ est défini par $\boldsymbol{\varepsilon}_{it} \equiv (\boldsymbol{\varepsilon}_{it}^{yx}, \boldsymbol{\varepsilon}_{it}^{s})$.

Le modèle considéré a une structure récursive, le modèle de choix de régime dépendant des modèles de choix d'assolement, *via* les profits par régime $\Pi_{it}(\boldsymbol{\gamma}_i, \boldsymbol{\delta}_0, \boldsymbol{\varepsilon}_{it}^{s}; r)$ pour $r \in \mathcal{R}$, les modèles de choix d'assolement dépendant eux-mêmes des modèles d'offre de rendement et de demande d'intrants *via* le vecteur des marges brutes $\boldsymbol{\pi}_{it}$. Aussi, les vecteurs de paramètres aléatoires $\boldsymbol{\gamma}_i^{yx}$ interviennent dans les modèles de choix d'assolement et le vecteur $\boldsymbol{\gamma}_i$ dans son ensemble intervient dans le modèle de choix de régime. De même, le vecteur de paramètres fixes $\boldsymbol{\delta}_0^{yx}$ intervient dans les modèles de choix d'assolement et le vecteur $\boldsymbol{\delta}_0$ dans son ensemble intervient dans le modèle de choix de régime.

Les termes d'erreur $\varepsilon_{k,it}^{s}$ correspondant aux cultures non produites dans le régime r_{it} ne peuvent pas être reconstruits à partir du modèle et des variables observées. Ils déterminent les choix du régime r_{it} *via* leur impact dans les profits de régime $\Pi_{it}(\boldsymbol{\gamma}_i, \boldsymbol{\delta}_0, \boldsymbol{\varepsilon}_{it}^{s}; r)$ puisqu'ils sont supposés connus des agriculteurs au moment de leurs choix d'assolement. Le fait que certains des éléments de $\boldsymbol{\varepsilon}_{it}^{s}$ ne puissent être reconstruits simplement pose des problèmes d'intégration des fonctions de probabilité de choix de régime qui nous ont conduit à utiliser des approximations de Laplace pour ces fonctions. Ce constat implique également que le vecteur des termes d'erreur des équations d'assolement des cultures non produites, $\boldsymbol{\varepsilon}_{it}^{s,0} \equiv \left(\varepsilon_{k,it}^{s} : k \in \mathcal{K}^{0}(r_{it})\right)$, doit être considéré comme une variable manquante. Celui des termes d'erreur des équations d'assolement des cultures produites, $\boldsymbol{\varepsilon}_{it}^{s,+} \equiv \left(\varepsilon_{k,it}^{s} : k \in \mathcal{K}_{-K}^{+}(r_{it})\right)$, sera reconstruit en cours d'estimation.

Le vecteur des termes d'erreur $\boldsymbol{\varepsilon}_{k,it}^{yx}$ des équations de rendement et d'utilisation d'intrants des cultures non produites, collectés dans $\boldsymbol{\varepsilon}_{it}^{yx,0} \equiv \left(\boldsymbol{\varepsilon}_{k,it}^{yx} : k \in \mathcal{K}^0(r_{it})\right)$, sera également considéré comme une variable manquante, mais pour des raisons techniques liées à la mise en œuvre des algorithmes considérés pour l'estimation des paramètres du modèle. Les termes d'erreur $\boldsymbol{\varepsilon}_{k,it}^{yx}$ des équations de rendement et d'utilisation d'intrants des cultures produites sont collectées dans le vecteur $\boldsymbol{\varepsilon}_{it}^{yx,+} \equiv \left(\boldsymbol{\varepsilon}_{k,it}^{yx} : k \in \mathcal{K}^+(r_{it})\right)$.

3.2.2. Modèles micro-économétriques de choix de production multiculture à paramètres aléatoires et avec choix de cultures : spécification paramétrique

Les hypothèses sur les termes aléatoires – paramètres aléatoires et termes d'erreur – et leurs relations avec les variables explicatives du modèle considéré reprennent largement celles du Chapitre 1.

Les termes d'erreur $(\boldsymbol{\varepsilon}_{it}, \mathbf{e}_{it})$ sont supposés indépendants et équidistribués pour $i = 1,...,N$ et $t = 1,...,T$. Les hypothèses d'équidistribution et d'indépendance entre agriculteurs sont standard. L'hypothèse d'absence de dépendance dans la dimension temporelle est plus discutable mais repose sur l'idée que la dynamique des processus modélisés est suffisamment persistante et stable pour être captée en grande partie par les paramètres aléatoires du modèle.

Les paramètres $\boldsymbol{\gamma}_i$ aléatoires sont supposés indépendants et équidistribués pour pour $i = 1,...,N$. Cette hypothèse est standard. Elle implique que la distribution commune des paramètres spécifiques à chaque agriculteur est également celle de ces paramètres au sein de la population d'agriculteurs représentée par l'échantillon disponible.

Les termes $(\boldsymbol{\varepsilon}_i, \mathbf{e}_i) = \big((\boldsymbol{\varepsilon}_{it}, \mathbf{e}_{it}) : t = 1,...,T\big))$, $\mathbf{z}_i = (\mathbf{z}_{it} : t = 1,...,T)$ et $\boldsymbol{\gamma}_i$ sont supposés mutuellement indépendants. L'hypothèse d'indépendance entre $(\boldsymbol{\varepsilon}_i, \mathbf{e}_i)$ et $\boldsymbol{\gamma}_i$ est standard, les termes $(\boldsymbol{\varepsilon}_{it}, \mathbf{e}_{it})$ étant sensés capter les effets de phénomènes variant dans le temps. L'indépendance de $\mathbf{z}_i$ avec $(\boldsymbol{\varepsilon}_i, \mathbf{e}_i, \boldsymbol{\gamma}_i)$ constitue une hypothèse d'exogénéité des variables explicatives relativement forte.

On supposera également que les termes d'erreur $\boldsymbol{\varepsilon}_{it}^{yx}$, $\boldsymbol{\varepsilon}_{it}^{s}$ et $\mathbf{e}_{it}$ sont mutuellement indépendants. Cette hypothèse d'indépendance est raisonnable étant donné la présence des paramètres aléatoires dans le modèle. L'indépendance entre $\boldsymbol{\varepsilon}_{it}^{s}$ et $\mathbf{e}_{it}$ repose essentiellement sur le fait que ces termes d'erreurs captent, de par la construction du modèle, les effets de facteurs très différents. L'indépendance entre $\boldsymbol{\varepsilon}_{it}^{yx}$ et $(\boldsymbol{\varepsilon}_{it}^{s}, \mathbf{e}_{it})$ repose essentiellement sur le fait que ces termes d'erreurs captent les effets d'évènements aléatoires décalés dans le temps, ceux affectant $\boldsymbol{\varepsilon}_{it}^{yx}$ survenant après ceux affectant $(\boldsymbol{\varepsilon}_{it}^{s}, \mathbf{e}_{it})$.

L'hypothèse d'indépendance entre $\boldsymbol{\varepsilon}_{it}^{yx}$ et $(\boldsymbol{\varepsilon}_{it}^{s}, \mathbf{e}_{it})$ implique que les rendements et quantités d'intrants $(\mathbf{y}_{it}^{+}, \mathbf{x}_{it}^{+})$ observés sont indépendants des choix d'assolement et de régime $(\mathbf{s}_{it}^{+}, r_{it})$ conditionnellement en $(\mathbf{z}_i, \boldsymbol{\gamma}_i)$. Cette implication s'avère très utile pour simplifier la forme de la vraisemblance du modèle.

Nous supposons, de manière standard, que les termes d'erreur $(\boldsymbol{\varepsilon}_{it}^{yx}, \boldsymbol{\varepsilon}_{it}^{s})$ sont normaux, avec :

(35) $$(\boldsymbol{\varepsilon}_{it}^{yx}, \boldsymbol{\varepsilon}_{it}^{s}) \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Psi}_0) \text{ avec } \boldsymbol{\Psi}_0 = \begin{bmatrix} \boldsymbol{\Psi}_0^{yx} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Psi}_0^{s} \end{bmatrix}$$

et que les paramètres aléatoires $\boldsymbol{\gamma}_i$ sont également normaux, avec :

(36) $$\boldsymbol{\gamma}_i \sim \mathcal{N}(\boldsymbol{\mu}_0, \boldsymbol{\Omega}_0).$$

Aucune restriction n'est imposée sur les relations entre les différents éléments de $\boldsymbol{\gamma}_i$. Les paramètres de flexibilité des choix d'assolement $\boldsymbol{\alpha}_i^s$ et des choix d'intrants $\boldsymbol{\alpha}_i^x$ apparaissent en logarithme dans $\boldsymbol{\gamma}_i$. Ils suivent donc une loi log-normale jointe, et sont donc positifs, ce qui est nécessaire pour que le modèle de choix de production considéré se « comporte correctement » d'un point de vue économique. Une remarque similaire s'applique pour le paramètre d'échelle σ_i.

Enfin, nous supposons que les éléments de $\mathbf{e}_{it}$ sont indépendants et suivent une loi de Gumbel standard. Ce choix de distribution fait du modèle du choix de régime un modèle de choix discret probabiliste de type Logit Multinomial à paramètres aléatoires (McFadden et Train, 2000 ; Train, 2009).

3.3. Arbitrage : richesse de la spécification *versus* difficulté d'estimation

La spécification du modèle considéré est relativement complexe puisque ce modèle combine paramètres aléatoires et changement de régime endogène. Ceci-dit nous verrons dans la suite que les problèmes liés à l'estimation de ce modèle proviennent principalement de sa structure avec choix de régime. Nous avons évité, dans la mesure du possible, d'imposer des hypothèses peu réalistes mais qui auraient permis de simplifier l'estimation du modèle. En

particulier, supprimer l'ensemble des variables de contrôle du modèle aurait singulièrement simplifié son estimation. Ce point a déjà été mentionné dans le chapitre 2 de la thèse. Son importance est accrue ici. De même, certaines hypothèses quant au processus de choix de régime aurait également singulièrement simplifié la procédure d'estimation. Ce point est discuté dans le chapitre suivant.

Dans cette partie de la thèse, comme dans les chapitres précédents, nous avons pris l'option de spécifier le modèle de la manière la plus satisfaisante possible quant à notre interprétation des processus de choix des agriculteurs, quitte à devoir recourir à des procédures d'estimation relativement complexe. Ce choix est assumé.

Chapitre 4

Modèles micro-économétriques de choix de production multiculture à paramètres aléatoires et avec choix de cultures : estimation *via* les algorithmes MCEM et SAEM

L'objectif de ce chapitre est de présenter une approche basée sur les algorithmes MCEM et SAEM pour l'estimation des modèles de choix de production multiculture présentés dans le chapitre précédent dans le cadre du MV. Nous utiliserons l'approche générale proposée dans le Chapitre 2 pour l'estimation des modèles à régime de production unique. Aussi, nous nous appuierons sur de nombreux éléments et résultats de ce chapitre, en ne rappelant que les plus délicats ou les moins usuels.

Bien entendu, la gestion des régimes de production pose des problèmes spécifiques que nous présentons progressivement. Déterminer et interpréter la forme de la vraisemblance des modèles de choix de production multiculture à changement de régime endogène et à paramètres aléatoires considérés ici est l'objet de la première section de ce chapitre. La structure des modèles à changement de régime endogène étant relativement complexe, la vraisemblance est définie progressivement.

Dans la seconde section nous présentons la structure générale des algorithmes MCEM et SAEM considérés pour calculer un estimateur du MV des paramètres du modèle. Nous nous appuierons sur les éléments et concepts introduits dans le Chapitre 4. Néanmoins, nous montrons dans cette section que la structure avec choix de régime du modèle complique singulièrement la mise en œuvre des algorithmes MCEM et SAEM. De fait, que le modèle soit à paramètres aléatoires ou non a peu d'impact sur les principales difficultés de mise en œuvre des algorithmes de maximisation de la vraisemblance du modèle considéré. Ces difficultés proviennent de la structure récursive du modèle : les choix de régimes dépendent des paramètres des choix d'assolement qui dépendent eux-mêmes des niveaux de rendements et d'utilisation d'intrants espérés. Cette structure rend difficile la décomposition des étapes

M des algorithmes MCEM et SAEM. Une solution aisée à mettre en œuvre a été définie pour les modèles à régime unique considéré dans les deux premiers chapitres de cette thèse. La solution adoptée ici est plus difficile à mettre en œuvre.

La troisième section aborde le problème lié au fait que les rendements, utilisations d'intrants et assolements des cultures non produites ne sont pas observés. Il s'agit ici d'une question de maximisation de vraisemblance de forme fonctionnelle complexe. Nous montrons que l'utilisation d'un algorithme EM proposé par Ruud (1991) s'avère utile, et ce d'autant plus qu'il s'insère de manière naturelle dans les algorithmes MCEM et SAEM que nous avons utilisés dans nos travaux empiriques. Ces algorithmes sont présentés dans les quatrième et cinquième sections.

4.1. Modèles micro-économétriques de choix de production multiculture à paramètres aléatoires et avec choix de cultures : vraisemblance

La forme de la vraisemblance du modèle est pénible à définir du point de vue des notations. La forme de la vraisemblance d'une observation dépend explicitement du régime de production de cette observation. Nous avons tenté, dans la mesure du possible, de choisir des notations intuitives. Néanmoins, le volume de notations introduit reste conséquent, mais nécessaire. Nous avons décidé de définir les principales notations employées au début de cette section, avant de définir la vraisemblance du modèle proprement dit.

Afin d'alléger les notations, nous ignorons ici le fait que les termes $\boldsymbol{\varepsilon}_{it}^{yx}$ et $\boldsymbol{\varepsilon}_{it}^{s}$ sont supposés indépendants. En effet cette hypothèse d'indépendance complique singulièrement les

notations à mobiliser pour décrire les procédures d'estimation considérées dans ce chapitre. Elle est imposée dans nos travaux empiriques car elle est raisonnable et permet d'économiser des paramètres à estimer.

4.1.1. Notations

Exprimer les termes d'erreur des équations de rendement, d'utilisation d'intrants et d'assolement des cultures produites à partir des autres éléments de ces équations permet de définir une fonction de résidus utile pour déterminer la vraisemblance du modèle. Il est aisé d'écrire $\boldsymbol{\varepsilon}_{it}^{yx,+}$ sous la forme :[20]

(1) $$\boldsymbol{\varepsilon}_{it}^{yx,+} = \boldsymbol{\varepsilon}_{it}^{yx,+}(\boldsymbol{\gamma}_i^{yx}, \boldsymbol{\delta}_0^{yx} r_{it})$$

avec :

(2) $$\boldsymbol{\varepsilon}_{it}^{yx,+}(\boldsymbol{\gamma}^{yx}, \boldsymbol{\delta}^{yx}) = \boldsymbol{\varsigma}_{k,it}^{yx,+}(\boldsymbol{\gamma}^{yx}) - (\mathbf{Z}_{k,it}^{yx,+})'\boldsymbol{\delta}^{yx}.$$

De même, on montre aisément que $\boldsymbol{\varepsilon}_{it}^{s,+}$ peut s'écrire sous la forme :[21]

---

[20] Les des équations de rendement et d'intrants donnent que :

$$\begin{bmatrix} \varsigma_{it}^{y}(\boldsymbol{\gamma}_{k,i}^{yx}) \\ \varsigma_{it}^{x}(\boldsymbol{\gamma}_{k,i}^{yx}) \end{bmatrix} = \begin{bmatrix} y_{k,it} - \beta_{k,i}^{y} + \alpha_{k,i}^{x} q_{k,it}^{2} / 2 \\ x_{k,it} - \beta_{k,i}^{x} + \alpha_{k,i}^{x} q_{k,it}^{2} \end{bmatrix} = \begin{bmatrix} (\mathbf{c}_{k,it}^{y})' & \mathbf{0} \\ \mathbf{0} & (\mathbf{c}_{k,it}^{x})' \end{bmatrix} \begin{bmatrix} \boldsymbol{\delta}_{k,0}^{y} \\ \boldsymbol{\delta}_{k,0}^{x} \end{bmatrix} + \begin{bmatrix} \varepsilon_{k,it}^{y} \\ \varepsilon_{k,it}^{x} \end{bmatrix}$$

[21] En effet, les expressions des marges brutes $\pi_{k,it}$ et la forme log-linéarisée de l'équation de choix d'assolement donnent que :

(3) $$\boldsymbol{\varepsilon}_{it}^{s,+} = \boldsymbol{\varepsilon}_{it}^{s,+}(\boldsymbol{\gamma}_i, \boldsymbol{\delta}_0)$$

avec :

(4) $$\boldsymbol{\varepsilon}_{it}^{s,+}(\boldsymbol{\gamma}, \boldsymbol{\delta}) = \boldsymbol{\varsigma}_{k,it}^{s,+}(\boldsymbol{\gamma}) - (\mathbf{Z}_{k,it}^{s,+})'\boldsymbol{\delta}\,.$$

Enfin, ces différents éléments sont ensuite empilés pour écrire $\boldsymbol{\varepsilon}_{it}^{+} \equiv (\boldsymbol{\varepsilon}_{it}^{xy,+}, \boldsymbol{\varepsilon}_{it}^{s,+})$ sous la forme :

(5) $$\boldsymbol{\varepsilon}_{it}^{+} = \boldsymbol{\varepsilon}_{it}^{+}(\boldsymbol{\gamma}_i, \boldsymbol{\delta}_0)$$

avec :

(6) $$\boldsymbol{\varepsilon}_{it}^{+}(\boldsymbol{\gamma}, \boldsymbol{\delta}) = \boldsymbol{\varsigma}_{k,it}^{+}(\boldsymbol{\gamma}) - (\mathbf{Z}_{k,it}^{+})'\boldsymbol{\delta}\,.$$

Les termes d'erreur $\boldsymbol{\varepsilon}_{it}^{+}$ et $\boldsymbol{\varepsilon}_{it}^{0} \equiv (\boldsymbol{\varepsilon}_{it}^{yx,0}, \boldsymbol{\varepsilon}_{it}^{s,0})$, et la distribution de la variable $\boldsymbol{\varepsilon}_{it}^{0}$ conditionnelle en $\boldsymbol{\varepsilon}_{it}^{+}$, sont utiles pour décrire les algorithmes employés pour l'estimation des paramètres du modèle. Nous introduisons des notations similaires, avec :

(7) $$\boldsymbol{\Psi}_{0,it}^{++} \equiv V[\boldsymbol{\varepsilon}_{it}^{+}],\ \boldsymbol{\Psi}_{0,it}^{00} \equiv V[\boldsymbol{\varepsilon}_{it}^{0}],\ \boldsymbol{\Psi}_{0,it}^{0+} \equiv Cov[\boldsymbol{\varepsilon}_{it}^{0}, \boldsymbol{\varepsilon}_{it}^{+}] \text{ et } \boldsymbol{\Psi}_{0,it}^{+0} \equiv Cov[\boldsymbol{\varepsilon}_{it}^{+}, \boldsymbol{\varepsilon}_{it}^{0}].$$

Munis de ces notations nous avons :

$$\left\{\begin{array}{c} (\alpha_i^s)^{-1}\ln s_{(g),it} + (\alpha_{(g),i}^s)^{-1}\ln s_{k|(g),it} \\ -(\alpha_i^s)^{-1}\ln s_{(G),it} - (\alpha_{(G),i}^s)^{-1}\ln s_{K|(G),it} \\ +\beta_{k,i}^s - a_{k,it} + a_{k,it} \\ -(p_{k,it}\beta_{k,i}^y - v_{k,it}\beta_{k,i}^x) - \alpha_{k,i}^x q_{k,it}^2 p_{k,it}/2 \\ +(p_{K,it}\beta_{K,i}^y - v_{K,it}\beta_{K,i}^x) + \alpha_{K,i}^x q_{K,it}^2 p_{K,it}/2 \end{array}\right\} = \begin{bmatrix} (p_{k,it}\mathbf{c}_{k,it}^{ys})' & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -(v_{k,it}\mathbf{c}_{k,it}^{xs})' & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & -(p_{K,it}\mathbf{c}_{K,it}^{ys})' & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & (v_{K,it}\mathbf{c}_{k,it}^{xs})' & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & (\mathbf{c}_{k,it}^{s})' \end{bmatrix} \begin{bmatrix} \boldsymbol{\delta}_{k,0}^y \\ \boldsymbol{\delta}_{k,0}^x \\ \boldsymbol{\delta}_{K,0}^y \\ \boldsymbol{\delta}_{K,0}^x \\ \boldsymbol{\delta}_{k,0}^s \end{bmatrix} + \varepsilon_{k,it}^s$$

(8) $$\boldsymbol{\varepsilon}_{it}^{0} \mid \boldsymbol{\varepsilon}_{it}^{+} \sim \mathcal{N}(\boldsymbol{\Gamma}_{0,it}^{0|+}\boldsymbol{\varepsilon}_{it}^{s,+}, \boldsymbol{\Psi}_{0,it}^{0|+})$$

où :

(9) $$\boldsymbol{\Gamma}_{0,it}^{0|+} \equiv \boldsymbol{\Psi}_{0,it}^{0+}(\boldsymbol{\Psi}_{0,it}^{++})^{-1} \text{ et } \boldsymbol{\Psi}_{0,it}^{0|+} = \boldsymbol{\Psi}_{0,it}^{00} - \boldsymbol{\Psi}_{0,it}^{0+}(\boldsymbol{\Psi}_{0,it}^{++})^{-1}\boldsymbol{\Psi}_{0,it}^{+0}.$$

Déterminer la vraisemblance du modèle de choix de régime repose sur la distribution de $\boldsymbol{\varepsilon}_{it}^{s,0}$ conditionnelle en $\boldsymbol{\varepsilon}_{it}^{+}$. Etant donné que $\boldsymbol{\varepsilon}_{it}$ suit une la loi normale multivariée $\mathcal{N}(\mathbf{0}, \boldsymbol{\Psi}_0)$, la distribution de $\boldsymbol{\varepsilon}_{it}^{s,0}$ conditionnelle en $\boldsymbol{\varepsilon}_{it}^{+}$ est également normale et peut être définie à partir des éléments de $\boldsymbol{\Psi}_0$. Cette distribution est ici définie par :

(10) $$\boldsymbol{\varepsilon}_{it}^{s,0} \mid \boldsymbol{\varepsilon}_{it}^{+} \sim \mathcal{N}(\boldsymbol{\Gamma}_{0,it}^{s,0|+}\boldsymbol{\varepsilon}_{it}^{s,+}, \boldsymbol{\Psi}_{0,it}^{s,0|+}).$$

Les termes $\boldsymbol{\Gamma}_{0,it}^{s,0|+}$ et $\boldsymbol{\Psi}_{0,it}^{s,0|+}$ ne sont pas décrit en détail ici car ce n'est pas utile par la suite.

Enfin, il apparaîtra utile d'écrire explicitement $\boldsymbol{\varepsilon}_{it}^{s}$ à partir des éléments de $\boldsymbol{\varepsilon}_{it}^{s,+}$ et $\boldsymbol{\varepsilon}_{it}^{s,0}$ « dans le bon ordre ». Pour ce faire nous définissons les matrices de sélection $\mathbf{H}_{it}^{s,0}$ et $\mathbf{H}_{it}^{s,+}$ de telle sorte que les conditions $(\mathbf{H}_{it}^{s,0})'\boldsymbol{\varepsilon}_{it}^{s} = \boldsymbol{\varepsilon}_{it}^{s,0}$ et $(\mathbf{H}_{it}^{s,+})'\boldsymbol{\varepsilon}_{it}^{s} = \boldsymbol{\varepsilon}_{it}^{s,+}$ soient satisfaites. On a alors :

(11) $$\boldsymbol{\varepsilon}_{it}^{s} = \mathbf{H}_{it}^{s,0}\boldsymbol{\varepsilon}_{it}^{s,0} + \mathbf{H}_{it}^{s,+}\boldsymbol{\varepsilon}_{it}^{s,+}.$$

Les matrices de sélection $\mathbf{H}_{it}^{0}$ et $\mathbf{H}_{it}^{+}$ sont définies de la même manière pour $\boldsymbol{\varepsilon}_{it}$ à partir des éléments de $\boldsymbol{\varepsilon}_{it}^{+}$ et $\boldsymbol{\varepsilon}_{it}^{0}$, avec :

(12) $$\boldsymbol{\varepsilon}_{it} = \mathbf{H}_{it}^{0}\boldsymbol{\varepsilon}_{it}^{0} + \mathbf{H}_{it}^{+}\boldsymbol{\varepsilon}_{it}^{+}.$$

4.1.2. Vraisemblances

Les observations d'une même exploitation étant liées par le vecteur de paramètres aléatoires spécifiques à cette exploitation $\boldsymbol{\gamma}_i$, il convient ici de déterminer la vraisemblance des choix de production observés pour chaque exploitation sur l'ensemble de la période d'observation, *i.e.* du vecteur $(\mathbf{y}_i^+, \mathbf{x}_i^+, \mathbf{s}_i^+, \mathbf{r}_i) \equiv \left((\mathbf{y}_{it}^+, \mathbf{x}_{it}^+, \mathbf{s}_{it}^+, r_{it}) : t = 1,...,T\right)$. Cette vraisemblance s'obtient à partir de la fonction de densité de $(\mathbf{y}_i^+, \mathbf{x}_i^+, \mathbf{s}_i^+, \mathbf{r}_i)$ conditionnelle en les variables exogènes $\mathbf{z}_i$, $f(\mathbf{y}_i^+, \mathbf{x}_i^+, \mathbf{s}_i^+, \mathbf{r}_i \mid \mathbf{z}_i)$. Les paramètres aléatoires $\boldsymbol{\gamma}_i$ structurant le modèle considéré, cette fonction s'écrit sous la forme :

(13) $$f(\mathbf{y}_i^+, \mathbf{x}_i^+, \mathbf{s}_i^+, \mathbf{r}_i \mid \mathbf{z}_i) = \int f(\mathbf{y}_i^+, \mathbf{x}_i^+, \mathbf{s}_i^+, \mathbf{r}_i \mid \mathbf{z}_i, \boldsymbol{\gamma}) f(\boldsymbol{\gamma}) d\boldsymbol{\gamma} .$$

Nous utilisons ici l'hypothèse selon laquelle les termes $\boldsymbol{\gamma}_i$ et $\mathbf{z}_i$ sont indépendants. Les termes d'erreur $\boldsymbol{\varepsilon}_{it}$ étant supposés indépendants de $(\mathbf{z}_i, \boldsymbol{\gamma}_i)$ et indépendants entre eux dans la dimension temporelle, les vecteurs $(\mathbf{y}_{it}^+, \mathbf{x}_{it}^+, \mathbf{s}_{it}^+, r_{it})$ sont indépendants conditionnellement en $\boldsymbol{\gamma}_i$ et $\mathbf{z}_i$. On obtient alors que :[22]

(14) $$f(\mathbf{y}_i^+, \mathbf{x}_i^+, \mathbf{s}_i^+, \mathbf{r}_i \mid \mathbf{z}_i) = \int \prod_{t=1}^{T} f(\mathbf{y}_{it}^+, \mathbf{x}_{it}^+, \mathbf{s}_{it}^+, r_{it} \mid \mathbf{z}_{it}, \boldsymbol{\gamma}_i) f(\boldsymbol{\gamma}) d\boldsymbol{\gamma} .$$

La fonction de densité des paramètres aléatoires $\boldsymbol{\gamma}_i$, $f(\boldsymbol{\gamma}_i)$, est donné par :

---

[22] On utilise ici également le fait que le modèle de $(\mathbf{y}_{it}^+, \mathbf{x}_{it}^+, \mathbf{s}_{it}^+, r_{it})$ ne dépend de $\mathbf{z}_i$ que *via* $\mathbf{z}_{it}$.

(15) $$f(\boldsymbol{\gamma}_i) = \varphi(\boldsymbol{\gamma}_i - \boldsymbol{\mu}_0; \boldsymbol{\Omega}_0)$$

puisque $\boldsymbol{\gamma}_i \sim \mathcal{N}(\boldsymbol{\mu}_0, \boldsymbol{\Omega}_0)$. Le terme $\varphi(\boldsymbol{\gamma}_i - \boldsymbol{\mu}; \boldsymbol{\Omega})$ définit la vraisemblance de $\boldsymbol{\gamma}_i$ en $(\boldsymbol{\mu}, \boldsymbol{\Omega})$.

Les termes d'erreur $\boldsymbol{\varepsilon}_{it}$ et $\mathbf{e}_{it}$ sont supposés indépendants mais la structure récursive du modèle indique que le modèle de choix de régime r_{it} dépend explicitement de $\boldsymbol{\varepsilon}_{it}^{s}$. Les termes d'erreur $\boldsymbol{\varepsilon}_{it}^{yx}$ et $\boldsymbol{\varepsilon}_{it}^{s}$ étant potentiellement liés, aucune hypothèse d'indépendance conditionnelle ne peut être invoquée pour décomposer simplement la fonction de densité $f(\mathbf{y}_{it}^{+}, \mathbf{x}_{it}^{+}, \mathbf{s}_{it}^{+}, r_{it} \mid \mathbf{z}_{it}, \boldsymbol{\gamma}_i)$. Nous utilisons simplement la loi de Bayes pour écrire cette fonction sous la forme :

(16) $$f(\mathbf{y}_{it}^{+}, \mathbf{x}_{it}^{+}, \mathbf{s}_{it}^{+}, r_{it} \mid \mathbf{z}_{it}, \boldsymbol{\gamma}_i) = f(\mathbf{y}_{it}^{+}, \mathbf{x}_{it}^{+}, \mathbf{s}_{it}^{+} \mid \mathbf{z}_{it}, \boldsymbol{\gamma}_i, \mathbf{s}_{it}^{+}) f(r_{it} \mid \mathbf{z}_{it}, \boldsymbol{\gamma}_i, \mathbf{y}_{it}^{+}, \mathbf{x}_{it}^{+}, \mathbf{s}_{it}^{+}).$$

Notons simplement ici que le modèle de choix de régime r_{it} dépend directement de $\boldsymbol{\varepsilon}_{it}^{s}$ et indirectement, via les relations statistiques potentielles entre $\boldsymbol{\varepsilon}_{it}^{yx}$ et $\boldsymbol{\varepsilon}_{it}^{s}$, de $\boldsymbol{\varepsilon}_{it}^{yx}$.

D'après les résultats présentés dans la partie précédente et les notations introduites ici, la densité jointe des rendements, utilisations d'intrants et assolement observés $(\mathbf{y}_{it}^{+}, \mathbf{x}_{it}^{+}, \mathbf{s}_{it}^{+})$ conditionnelle en les variables exogènes et paramètres aléatoires $(\mathbf{z}_{it}, \boldsymbol{\gamma}_i)$ est donnée par :

(17) $$\begin{aligned} f(\mathbf{y}_{it}^{+}, \mathbf{x}_{it}^{+}, \mathbf{s}_{it}^{+} \mid \mathbf{z}_{it}, \boldsymbol{\gamma}_i) = (\alpha_i^{s})^{1-G^{+}(r_{it})} \left(\prod_{g \in \mathcal{G}^{+}(r_{it})} (\alpha_{(g),i}^{s})^{1-K_{(g)}^{+}(r_{it})} \right) \left(\prod_{k \in \mathcal{K}^{+}(r_{it})} s_{k,it}^{-1} \right) \\ \times \varphi\left(\boldsymbol{\varepsilon}_{it}^{+}(\boldsymbol{\gamma}_i, \boldsymbol{\delta}_0); \boldsymbol{\Psi}_{0,it}^{++} \right) \end{aligned}$$

où, pour le régime r_{it}, $\mathcal{G}^{+}(r_{it})$ désigne l'ensemble de groupes de cultures assolés, $G^{+}(r_{it})$ le

nombre de groupes de cultures assolées et $K^{+}_{(g)}(r_{it})$ le nombre de cultures produites dans le groupe *g*.

L'hypothèse posée sur les termes d'erreur $\mathbf{e}_{it}$ du modèle de choix de régime implique que ce modèle de choix discret probabiliste est un modèle de forme Logit Multinomial standard conditionnellement en $(\mathbf{z}_{it}, \boldsymbol{\gamma}_i, \boldsymbol{\varepsilon}^s_{it})$, avec :

$$(18) \qquad f(r_{it} \mid \mathbf{z}_{it}, \boldsymbol{\gamma}_i, \boldsymbol{\varepsilon}^s_{it}) = \frac{\exp\left(\sigma_i\left(\Pi_{it}(\boldsymbol{\gamma}_i, \boldsymbol{\delta}_0, \boldsymbol{\varepsilon}^s_{it}; r_{it}) - d_{r_{it},i}\right)\right)}{\sum_{r\in\mathcal{R}} \exp\left(\sigma_i\left(\Pi_{it}(\boldsymbol{\gamma}_i, \boldsymbol{\delta}_0, \boldsymbol{\varepsilon}^s_{it}; r) - d_{r,i}\right)\right)}.$$

Un problème apparaît ici. Il nous faut calculer $f(r_{it} \mid \mathbf{z}_{it}, \boldsymbol{\gamma}_i, \mathbf{y}^+_{it}, \mathbf{x}^+_{it}, \mathbf{s}^+_{it})$ et non pas $f(r_{it} \mid \mathbf{z}_{it}, \boldsymbol{\gamma}_i, \boldsymbol{\varepsilon}_{it})$. Or, le conditionnement par $\boldsymbol{\varepsilon}^s_{it}$ est plus informatif que le conditionnement par $\mathbf{s}^+_{it}$. On sait que $\boldsymbol{\varepsilon}^s_{it} = \mathbf{H}^{s,0}_{it}\boldsymbol{\varepsilon}^{s,0}_{it} + \mathbf{H}^{s,+}_{it}\boldsymbol{\varepsilon}^{s,+}_{it}$.[23] Le conditionnement par $\mathbf{s}^+_{it}$ permet de reconstruire $\boldsymbol{\varepsilon}^{s,+}_{it}$ sous la forme $\boldsymbol{\varepsilon}^{s,+}_{it} = \boldsymbol{\varepsilon}^s_{it}(\boldsymbol{\gamma}_i, \boldsymbol{\delta}_0)$. Mais aucune information ne permet de reconstruire « entièrement » $\boldsymbol{\varepsilon}^{s,0}_{it}$, le terme d'erreur du paramètre de coût des cultures non assolées dans le régime observé r_{it}. La seule information mobilisable est donnée par $\boldsymbol{\varepsilon}^+_{it}$ dont la corrélation avec $\boldsymbol{\varepsilon}^{s,0}_{it}$ peut permettre de prédire $\boldsymbol{\varepsilon}^{s,0}_{it}$. Aussi la fonction de probabilité $f(r_{it} \mid \mathbf{z}_{it}, \boldsymbol{\gamma}_i, \mathbf{y}^+_{it}, \mathbf{x}^+_{it}, \mathbf{s}^+_{it})$ est obtenue en intégrant la fonction de probabilité $f(r_{it} \mid \mathbf{z}_{it}, \boldsymbol{\gamma}_i, \boldsymbol{\varepsilon}^s_{it})$ sur la distribution du terme manquant $\boldsymbol{\varepsilon}^{s,0}_{it}$ conditionnelle en $\boldsymbol{\varepsilon}^+_{it}$, ce dernier étant le seul terme susceptible de contenir de l'information sur $\boldsymbol{\varepsilon}^{s,0}_{it}$. On a alors :

---

[23] où les matrices $\mathbf{H}^{s,+}_{it}$ et $\mathbf{H}^{s,0}_{it}$ sont de simples matrices de sélection.

$$(19) \qquad f(r_{it} \mid \mathbf{z}_{it}, \boldsymbol{\gamma}_i, \mathbf{y}_{it}^+, \mathbf{x}_{it}^+, \mathbf{s}_{it}^+) = \int \left\{ \begin{array}{c} f\left(r_{it} \mid \mathbf{z}_{it}, \boldsymbol{\gamma}_i, \mathbf{H}_{it}^{s,0}\boldsymbol{\varepsilon}_{it}^{s,0} + \mathbf{H}_{it}^{s,+}\boldsymbol{\varepsilon}_{it}^{s,+}(\boldsymbol{\gamma}_i, \boldsymbol{\delta}_0)\right) \\ \times\varphi(\boldsymbol{\varepsilon}_{it}^{s,0} - \boldsymbol{\Gamma}_{0,it}^{s,0+}\boldsymbol{\varepsilon}_{it}^{+}(\boldsymbol{\gamma}_i, \boldsymbol{\delta}_0); \boldsymbol{\Psi}_{0,it}^{s,0|+}) \end{array} \right\} d\boldsymbol{\varepsilon}_{it}^{s,0} .$$

Les résultats précédents permettent de définir la vraisemblance de $(\mathbf{y}_{it}^+, \mathbf{x}_{it}^+, \mathbf{s}_{it}^+, r_{it})$ conditionnelle en $(\mathbf{z}_{it}, \boldsymbol{\gamma}_i)$ calculée en $(\boldsymbol{\delta}, \boldsymbol{\Psi})$ sous la forme :

$$(20) \qquad \ell_{it}(\boldsymbol{\delta}, \boldsymbol{\Psi} \mid \boldsymbol{\gamma}_i) = \ell_{it}^+(\boldsymbol{\delta}, \boldsymbol{\Psi} \mid \boldsymbol{\gamma}_i)\ell_{it}^r(\boldsymbol{\delta}, \boldsymbol{\Psi} \mid \boldsymbol{\gamma}_i)$$

avec :

$$(21) \qquad \begin{array}{c} \ell_{it}^+(\boldsymbol{\delta}, \boldsymbol{\Psi} \mid \boldsymbol{\gamma}_i) = (\alpha_i^s)^{1-G^+(r_{it})}\left(\prod_{g\in\mathcal{G}^+(r_{it})}(\alpha_{(g),i}^s)^{1-K_{(g)}^+(r_{it})}\right)\left(\prod_{k\in\mathcal{K}^+(r_{it})} s_{k,it}^{-1}\right) \\ \times\varphi\left(\boldsymbol{\varepsilon}_{it}^+(\boldsymbol{\gamma}_i, \boldsymbol{\delta}); \boldsymbol{\Psi}_{it}^{++}\right) \end{array} .$$

et :

$$(22) \qquad \ell_{it}^r(\boldsymbol{\delta}, \boldsymbol{\Psi} \mid \boldsymbol{\gamma}_i) = \int \mathcal{P}_{it}(r_{it} \mid \boldsymbol{\varepsilon}_{it}^{s,0}; \boldsymbol{\gamma}_i, \boldsymbol{\delta}, \boldsymbol{\Psi})\varphi(\boldsymbol{\varepsilon}_{it}^{s,0} - \boldsymbol{\Gamma}_{it}^{s,0+}\boldsymbol{\varepsilon}_{it}^{+}(\boldsymbol{\gamma}_i, \boldsymbol{\delta}_{it}); \boldsymbol{\Psi}_{it}^{s,0|+})d\boldsymbol{\varepsilon}_{it}^{s,0}$$

où :

$$(23) \qquad \mathcal{P}_{it}(r_{it} \mid \boldsymbol{\varepsilon}_{it}^s; \boldsymbol{\gamma}_i, \boldsymbol{\delta}, \boldsymbol{\Psi}) = \frac{\exp\left(\sigma_i\left(\Pi_{it}(\boldsymbol{\gamma}_i, \boldsymbol{\delta}, \mathbf{H}_{it}^{s,0}\boldsymbol{\varepsilon}_{it}^{s,0} + \mathbf{H}_{it}^{s,+}\boldsymbol{\varepsilon}_{it}^{s,+}(\boldsymbol{\gamma}_i, \boldsymbol{\delta}); r_{it}) - d_{r_{it},i}\right)\right)}{\sum_{r\in\mathcal{R}}\exp\left(\sigma_i\left(\Pi_{it}(\boldsymbol{\gamma}_i, \boldsymbol{\delta}, \mathbf{H}_{it}^{s,0}\boldsymbol{\varepsilon}_{it}^{s,0} + \mathbf{H}_{it}^{s,+}\boldsymbol{\varepsilon}_{it}^{s,+}(\boldsymbol{\gamma}_i, \boldsymbol{\delta}); r) - d_{r,i}\right)\right)} .$$

La fonction de probabilité du choix de régime $\ell_{it}^r(\boldsymbol{\delta}, \boldsymbol{\Psi} \mid \boldsymbol{\gamma}_i)$ est intégrée en utilisant une approximation de Laplace, suivant une approche proposée par Harding et Hausman (2007) pour l'intégration des probabilités de modèles de choix discret de forme Logit Multinomial et à paramètres aléatoires. Cette approche d'intégration par approximation s'est montrée performante, quoique gourmande en temps de calculs. L'idée de l'approximation de Laplace est celle des approximations de Taylor au second ordre avec un choix optimal du point

d'approximation.

Aussi, nous considérerons que cette vraisemblance est « intégrée » dans la suite de ce chapitre.

Finalement, la vraisemblance de l'observation des choix de production $(\mathbf{y}_{it}^{+}, \mathbf{x}_{it}^{+}, \mathbf{s}_{it}^{+}, r_{it})$ conditionnelle en les variables exogènes calculée en $(\boldsymbol{\delta}, \boldsymbol{\Psi}, \boldsymbol{\mu}, \boldsymbol{\Omega})$ est donnée par l'intégrale :

(24) $$\ell_i(\boldsymbol{\delta}, \boldsymbol{\Psi}, \boldsymbol{\mu}, \boldsymbol{\Omega}) = \int \ell_i(\boldsymbol{\delta}, \boldsymbol{\Psi} \mid \boldsymbol{\gamma}) \varphi(\boldsymbol{\gamma} - \boldsymbol{\mu}; \boldsymbol{\Omega}) d\boldsymbol{\gamma} = \int \ell_i^{+}(\boldsymbol{\delta}, \boldsymbol{\Psi} \mid \boldsymbol{\gamma}) \ell_i^{r}(\boldsymbol{\delta}, \boldsymbol{\Psi} \mid \boldsymbol{\gamma}) \varphi(\boldsymbol{\gamma} - \boldsymbol{\mu}; \boldsymbol{\Omega}) d\boldsymbol{\gamma}$$

avec :

(25) $$\ell_i^{+}(\boldsymbol{\delta}, \boldsymbol{\Psi} \mid \boldsymbol{\gamma}) \equiv \prod_{t=1}^{T} \ell_{it}^{+}(\boldsymbol{\delta}, \boldsymbol{\Psi} \mid \boldsymbol{\gamma}) \text{ et } \ell_i^{r}(\boldsymbol{\delta}, \boldsymbol{\Psi} \mid \boldsymbol{\gamma}) \equiv \prod_{t=1}^{T} \ell_{it}^{r}(\boldsymbol{\delta}, \boldsymbol{\Psi} \mid \boldsymbol{\gamma})$$

et :

(26) $$\ell_i(\boldsymbol{\delta}, \boldsymbol{\Psi} \mid \boldsymbol{\gamma}) \equiv \ell_i^{+}(\boldsymbol{\delta}, \boldsymbol{\Psi} \mid \boldsymbol{\gamma}) \ell_i^{r}(\boldsymbol{\delta}, \boldsymbol{\Psi} \mid \boldsymbol{\gamma}).$$

4.2. Cadre général de mise en œuvre des algorithmes de type (G)EM

L'estimateur du MV de $\boldsymbol{\theta}_0 \equiv (\boldsymbol{\delta}_0, \boldsymbol{\Psi}_0, \boldsymbol{\mu}_0, \boldsymbol{\Omega}_0)$ est alors défini comme la solution en $\boldsymbol{\theta}$ du programme de maximisation de la log-vraisemblance de l'échantillon considéré, avec:

(27) $$\hat{\boldsymbol{\theta}}_N \equiv (\hat{\boldsymbol{\delta}}_N, \hat{\boldsymbol{\Psi}}_N, \hat{\boldsymbol{\mu}}_N, \hat{\boldsymbol{\Omega}}_N) \equiv \arg\max_{(\boldsymbol{\delta}, \boldsymbol{\Psi}, \boldsymbol{\mu}, \boldsymbol{\Omega})} \sum_{i=1}^{N} \ln \ell_i(\boldsymbol{\delta}, \boldsymbol{\Psi}, \boldsymbol{\mu}, \boldsymbol{\Omega})$$

Comme tout estimateur du MV, cet estimateur est convergent et asymptotiquement efficace si le modèle considéré est correctement spécifié et si les conditions assurant l'identification de $\boldsymbol{\theta}_0$ sont assurées, ce que nous supposerons ici (ces conditions pouvant être, en partie, vérifiées empiriquement).

Le calcul de l'estimateur du MV de $\boldsymbol{\theta}_0$ est impossible dans ce cas car deux problèmes d'intégration sont à résoudre. Tout d'abord, il est nécessaire de calculer les probabilité de choix de régime $\ell_{it}^{r}(\boldsymbol{\delta},\boldsymbol{\Psi}\,|\,\boldsymbol{\gamma}_i)$, en intégrant $\mathcal{P}_{it}(r_{it}\,|\,\boldsymbol{\varepsilon}_{it}^{s};\boldsymbol{\gamma}_i,\boldsymbol{\delta},\boldsymbol{\Psi})$ sur la distribution de $\boldsymbol{\varepsilon}_{it}^{s,0}$ conditionnelle en $\boldsymbol{\varepsilon}_{it}^{s,+}$. Ce problème d'intégration ne peut être résolu, ni numériquement, ni analytiquement. Cette question sera abordée plus bas. Nous supposons simplement pour l'instant qu'il peut être résolu.

Ensuite, le calcul direct de l'estimateur du MV de $\boldsymbol{\theta}_0$ requiert le calcul préalable des fonctions de vraisemblance $\ell_i(\boldsymbol{\delta},\boldsymbol{\Psi},\boldsymbol{\mu},\boldsymbol{\Omega})$. Ce calcul repose lui aussi sur la résolution d'un problème d'intégration, celui de l'intégration des vraisemblances conditionnelles en les paramètres aléatoires, $\ell_i(\boldsymbol{\delta},\boldsymbol{\Psi}\,|\,\boldsymbol{\gamma})$, sur la distribution de ces paramètres aléatoires. Ce problème d'intégration est celui que nous avons résolu, ou plutôt contourné, dans les deux premiers chapitres en ayant recours à des algorithmes MCEM ou SAEM. Nous adoptons la même approche ici.

Nous considérons d'abord l'algorithme EM adapté à notre problème d'estimation afin de définir la structure des algorithmes utilisés par la suite. La log-vraisemblance « complète » en $\boldsymbol{\theta}$ de l'observation $(\mathbf{y}_i^{+},\mathbf{x}_i^{+},\mathbf{s}_i^{+},\mathbf{r}_i,\boldsymbol{\gamma}_i)$ conditionnelle en $\mathbf{z}_i$ est simplement donnée par :

(28) $$\ln \ell_i^{C}(\boldsymbol{\theta};\boldsymbol{\gamma}_i) \equiv \ln \ell_i(\boldsymbol{\delta},\boldsymbol{\Psi}\,|\,\boldsymbol{\gamma}_i) + \ln \varphi(\boldsymbol{\gamma}_i - \boldsymbol{\mu};\boldsymbol{\Omega}) = \sum_{t=1}^{T} \ln \ell_{it}(\boldsymbol{\delta},\boldsymbol{\Psi}\,|\,\boldsymbol{\gamma}_i) + \ln \varphi(\boldsymbol{\gamma}_i - \boldsymbol{\mu};\boldsymbol{\Omega})$$

pour notre modèle. Aussi la log-vraisemblance en $\boldsymbol{\theta}$ de l'échantillon par :

(29) $$\ln L_N^C(\boldsymbol{\theta}) \equiv \sum_{i=1}^N \ln \ell_i^C(\boldsymbol{\theta};\boldsymbol{\gamma}_i) = \sum_{i=1}^N \ln \ell_i(\boldsymbol{\delta},\boldsymbol{\Psi} \mid \boldsymbol{\gamma}_i) + \sum_{i=1}^N \ln \varphi(\boldsymbol{\gamma}_i - \boldsymbol{\mu};\boldsymbol{\Omega}).$$

L'étape E de l'algorithme consiste à calculer l'espérance de la log-vraisemblance complète de l'échantillon en $\boldsymbol{\theta}$ sur la distribution des variables manquantes $\{\boldsymbol{\gamma}_i : i = 1,...,N\}$ conditionnelle en les variables observées $\{(\mathbf{y}_i^+, \mathbf{x}_i^+, \mathbf{s}_i^+, \mathbf{r}_i) : i = 1,...,N\}$ telle que cette distribution peut être évaluée à la suite de l'étape M précédente. La fonction de densité de $\boldsymbol{\gamma}_i$ conditionnelle en $(\mathbf{y}_i^+, \mathbf{x}_i^+, \mathbf{s}_i^+, \mathbf{r}_i)$ peut s'écrire à partir des fonctions de vraisemblance du modèle avec :

(30) $$f(\boldsymbol{\gamma}_i \mid \mathbf{y}_i^+, \mathbf{x}_i^+, \mathbf{s}_i^+, \mathbf{r}_i) = g_i(\boldsymbol{\theta}_0;\boldsymbol{\gamma}_i) \equiv \frac{\ell_i^C(\boldsymbol{\theta}_0;\boldsymbol{\gamma}_i)}{\ell_i(\boldsymbol{\theta}_0)}.$$

L'étape E de l'itération $m+1$ de l'algorithme consiste donc à calculer l'espérance conditionnelle des fonctions de log-vraisemblances « complètes » $\ln \ell_i^C(\boldsymbol{\theta};\boldsymbol{\gamma}_i)$:

(31) $$E_{(m),i}[\ln \ell_i^C(\boldsymbol{\theta};\boldsymbol{\gamma}_i)] = \int \ln \ell_i^C(\boldsymbol{\theta};\boldsymbol{\gamma}) g_i(\hat{\boldsymbol{\theta}}_{(m)};\boldsymbol{\gamma}) d\boldsymbol{\gamma}$$

pour $i = 1,...,N$. Bien entendu, cette espérance n'a pas de forme analytique dans notre cas. Cette question sera abordée par suite afin de nous concentrer ici sur la description de l'algorithme EM.

L'étape M de l'itération $m+1$ de l'algorithme EM consiste ensuite à maximiser en $\boldsymbol{\theta}$ l'espérance conditionnelle en les observations de la vraisemblance complète de l'échantillon obtenue lors de l'étape E :

(32) $$Q_N(\boldsymbol{\theta};\hat{\boldsymbol{\theta}}_{(m)}) \equiv \sum_{i=1}^N E_{(m),i}[\ln \ell_i^C(\boldsymbol{\theta};\boldsymbol{\gamma}_i)] = \sum_{i=1}^N \int \ln \ell_i^C(\boldsymbol{\theta};\boldsymbol{\gamma}) g_i(\hat{\boldsymbol{\theta}}_{(m)};\boldsymbol{\gamma}) d\boldsymbol{\gamma}$$

pour obtenir une valeur actualisée de $\boldsymbol{\theta}_0$:

(33) $$\hat{\boldsymbol{\theta}}_{(m+1)} = \arg\max_{\boldsymbol{\theta}} Q_N(\boldsymbol{\theta};\hat{\boldsymbol{\theta}}_{(m)})$$

Dans les algorithmes GEM, le problème d'optimisation utilisé pour réactualise l'estimation de $\boldsymbol{\theta}_0$ est remplacé par une simple recherche de valeur de $\boldsymbol{\theta}$ qui accroît $Q_N(\boldsymbol{\theta};\hat{\boldsymbol{\theta}}_{(m)})$, i.e. $\hat{\boldsymbol{\theta}}_{(m+1)}$ peut simplement être définie à partir de la condition :

(34) $$\hat{\boldsymbol{\theta}}_{(m+1)} \in \left\{\boldsymbol{\theta} : Q_N(\boldsymbol{\theta};\hat{\boldsymbol{\theta}}_{(m)}) > Q_N(\hat{\boldsymbol{\theta}}_{(m)};\hat{\boldsymbol{\theta}}_{(m)})\right\}.$$

Les algorithmes EM et GEM ont les propriétés de convergence globale.

4.2.1. Intégration par simulations sur la distribution des paramètres aléatoires

Le problème de l'intégration de l'espérance de la log-vraisemblance complète des observations, $E_{(m),i}[\ln \ell_i^C(\boldsymbol{\theta};\boldsymbol{\gamma}_i)]$, est relativement facile à résoudre. Comme dans les deux premiers chapitres, nous utilisons ici un simulateur par échantillonnage préférentiel (IS) avec $\mathcal{N}(\hat{\boldsymbol{\mu}}_{(m)},\hat{\boldsymbol{\Omega}}_{(m)})$ pour distribution de proposition.[24] Ce faisant nous réalisons l'étape M de l'itération $m+1$ d'un algorithme MCEM ou l'étape S d'un algorithme SAEM.

Concrètement, pour chaque individu i considéré nous réalisons J tirages, les $\tilde{\boldsymbol{\gamma}}_{i,j}^{(m)}$ pour

[24] Le cas échéant amélioré par un ré-échantillonnage préférentiel (SIR) afin de limiter le risque de dégénérescence de l'algorithme.

$j = 1,...,J$, de $\boldsymbol{\gamma}_i$ suivant la distribution $\mathcal{N}(\hat{\boldsymbol{\mu}}_{(m)}, \hat{\boldsymbol{\Omega}}_{(m)})$. Puis nous calculons les poids $\tilde{w}_{i,j}^{(m)}$ attribués à chacun ces tirages dans le calcul de l'approximation de $E_{(m),i}[\ln \ell_i^C(\boldsymbol{\theta};\boldsymbol{\gamma}_i)]$. Ces poids sont donnés par :

(35) $$\tilde{w}_{i,j}^{(m)} \equiv \frac{\ell_i(\hat{\boldsymbol{\delta}}_{(m)}, \hat{\boldsymbol{\Psi}}_{(m)} \mid \tilde{\boldsymbol{\gamma}}_{i,j}^{(m)})}{\sum_{j=1}^{J} \ell_i(\hat{\boldsymbol{\delta}}_{(m)}, \hat{\boldsymbol{\Psi}}_{(m)} \mid \tilde{\boldsymbol{\gamma}}_{i,j}^{(m)})}.$$

Enfin, le terme d'espérance de log-vraisemblance complète $E_{(m),i}[\ln \ell_i^C(\boldsymbol{\theta};\boldsymbol{\gamma}_i)]$ est approximé par :

(36) $$\tilde{E}_{(m),i,J}[\ln \ell_i^C(\boldsymbol{\theta};\boldsymbol{\gamma}_i)] = \sum_{j=1}^{J} \tilde{w}_{i,j}^{(m)} \ln \ell_i^C(\boldsymbol{\theta};\tilde{\boldsymbol{\gamma}}_{i,j}^{(m)}).$$

L'étape M d'un algorithme MCEM est complétée par le calcul de l'approximation du terme $Q_N(\boldsymbol{\theta};\hat{\boldsymbol{\theta}}_{(m)})$, avec:

(37) $$\tilde{Q}_{N,J}(\boldsymbol{\theta};\hat{\boldsymbol{\theta}}_{(m)}) \equiv \sum_{j=1}^{J} \sum_{i=1}^{N} \tilde{w}_{i,j}^{(m)} \ln \ell_i^C(\boldsymbol{\theta};\tilde{\boldsymbol{\gamma}}_{i,j}^{(m)}).$$

4.2.2. Etape M généralisée

La mise en œuvre de l'étape M d'un algorithme type EM est difficile dans notre cas en raison de la forme et de la structure de la log-vraisemblance du modèle considéré. Pour notre modèle, le terme $\tilde{Q}_{N,J}(\boldsymbol{\theta};\hat{\boldsymbol{\theta}}_{(m)})$ peut être décomposé en une somme de trois termes, avec :

(38) $$\begin{aligned}\tilde{Q}_{N,J}(\boldsymbol{\theta};\hat{\boldsymbol{\theta}}_{(m)}) = &\sum_{i=1}^{N} \sum_{j=1}^{J} \tilde{w}_{i,j}^{(m)} \ln \varphi(\tilde{\boldsymbol{\gamma}}_{i,j}^{(m)} - \boldsymbol{\mu};\boldsymbol{\Omega}) \\ &+ \sum_{i=1}^{N} \sum_{j=1}^{J} \tilde{w}_{i,j}^{(m)} \ln \ell_i^{+}(\boldsymbol{\delta},\boldsymbol{\Psi} \mid \tilde{\boldsymbol{\gamma}}_{i,j}^{(m)}) \\ &+ \sum_{i=1}^{N} \sum_{j=1}^{J} \tilde{w}_{i,j}^{(m)} \ln \ell_i^{r}(\boldsymbol{\delta},\boldsymbol{\Psi} \mid \tilde{\boldsymbol{\gamma}}_{i,j}^{(m)})\end{aligned}$$

Mais le problème de maximisation de $\tilde{Q}_{N,J}(\boldsymbol{\theta};\hat{\boldsymbol{\theta}}_{(m)})$ en $\boldsymbol{\theta}$ ne peut être décomposé qu'en deux sous-problèmes, le problème lié à la réactualisation de l'estimateur des paramètres $(\boldsymbol{\mu}_0, \boldsymbol{\Omega}_0)$ de la distribution des paramètres aléatoires du modèle :

(39) $$(\hat{\boldsymbol{\mu}}_{(m+1)}, \hat{\boldsymbol{\Omega}}_{(m+1)}) = \arg\max_{(\boldsymbol{\mu},\boldsymbol{\Omega})} \sum_{i=1}^{N} \sum_{j=1}^{J} \tilde{w}_{i,j}^{(m)} \ln \varphi(\tilde{\boldsymbol{\gamma}}_{i,j}^{(m)} - \boldsymbol{\mu};\boldsymbol{\Omega})$$

et le problème lié à réactualisation des paramètres $(\boldsymbol{\delta}_0, \boldsymbol{\Psi}_0)$ de la densité « noyau » du modèle :

(40) $$(\hat{\boldsymbol{\delta}}_{(m+1)}, \hat{\boldsymbol{\Psi}}_{(m+1)}) = \operatorname{argmax}_{(\boldsymbol{\delta},\boldsymbol{\Psi})} \left\{ \begin{array}{l} \sum_{i=1}^{N} \sum_{j=1}^{J} \tilde{w}_{i,j}^{(m)} \ln \ell_i^{+}(\boldsymbol{\delta},\boldsymbol{\Psi} \mid \tilde{\boldsymbol{\gamma}}_{i,j}^{(m)}) \\ + \sum_{i=1}^{N} \sum_{j=1}^{J} \tilde{w}_{i,j}^{(m)} \ln \ell_i^{r}(\boldsymbol{\delta},\boldsymbol{\Psi} \mid \tilde{\boldsymbol{\gamma}}_{i,j}^{(m)}) \end{array} \right\}$$

La log-vraisemblance $\ln \varphi(\boldsymbol{\gamma} - \boldsymbol{\mu};\boldsymbol{\Omega})$ étant standard, le calcul de $(\hat{\boldsymbol{\mu}}_{(m+1)}, \hat{\boldsymbol{\Omega}}_{(m+1)})$ est facile. Il ne repose que sur des calculs d'espérances et de matrices de variance-covariance empiriques des $\tilde{\boldsymbol{\gamma}}_{i,j}^{(m)}$ pondérées par les $\tilde{w}_{i,j}^{(m)}$. Les expressions des termes sont données lors de la description des algorithmes MCEM et SAEM testés empiriquement.

Le calcul de $(\hat{\boldsymbol{\delta}}_{(m+1)}, \hat{\boldsymbol{\Psi}}_{(m+1)})$ s'avère lui très compliqué dans notre cas. Tout d'abord, le paramètre $(\boldsymbol{\delta}, \boldsymbol{\Psi})$ intervient dans les fonctions de log-vraisemblance conditionnelles des modèles d'offre de rendement, de demande d'intrants et de choix d'assolement $\ln \ell_i^{+}(\boldsymbol{\delta},\boldsymbol{\Psi} \mid \boldsymbol{\gamma})$ et celle du choix de régime $\ln \ell_i^{r}(\boldsymbol{\delta},\boldsymbol{\Psi} \mid \boldsymbol{\gamma})$. De même, le paramètre $\boldsymbol{\Psi}^{s}$ intervient dans les fonctions de log-vraisemblance des choix d'assolement et des choix de régime. Ensuite, ces fonctions de log-vraisemblance ne sont pas triviales à maximiser en leurs paramètres

respectifs.

Les fonctions de log-vraisemblance $\ln \ell_i^+(\boldsymbol{\delta}, \boldsymbol{\Psi} \mid \boldsymbol{\gamma})$ des modèles d'offre de rendement, de demande d'intrants et de choix d'assolement sont des log-vraisemblances de systèmes de régressions empilées avec variables dépendantes manquantes. Ceci-dit, comme cela sera vu plus bas, ce problème peut être résolu. Nous supposons dans cette section que c'est le cas.

En revanche, la fonction de vraisemblance des choix de régime étant particulièrement « affreuse », il est difficile d'envisager de l'employer en tant que fonction objectif pour un problème de maximisation. C'est la raison pour laquelle il ne nous a pas été possible de décomposer le problème de maximisation définissant $(\hat{\boldsymbol{\delta}}_{(m+1)}, \hat{\boldsymbol{\Psi}}_{(m+1)})$ en une séquence de sous-problèmes simples (selon la logique de construction des étapes M des algorithmes ECM) garantissant que la valeur de $\tilde{Q}_{N,J}(\hat{\boldsymbol{\theta}}_{(m+1)}; \hat{\boldsymbol{\theta}}_{(m)})$ soit supérieure à celle de $\tilde{Q}_{N,J}(\hat{\boldsymbol{\theta}}_{(m+1)}; \hat{\boldsymbol{\theta}}_{(m)}) > \tilde{Q}_{N,J}(\hat{\boldsymbol{\theta}}_{(m)}; \hat{\boldsymbol{\theta}}_{(m)})$. La complexité des fonctions de log-vraisemblance du modèle de choix de régime $\ell_{it}^r(\boldsymbol{\delta}, \boldsymbol{\Psi} \mid \boldsymbol{\gamma})$ ne nous permet pas d'intégrer ces dernières dans un problème d'optimisation.

De fait, nous avons été contraints d'ignorer ces fonctions $\ln \ell_{it}^r(\boldsymbol{\delta}, \boldsymbol{\Psi}^s \mid \boldsymbol{\gamma})$ dans les étapes M de nos algorithmes. Ceci implique que les probabilités de choix de régime jouent un rôle « passif » dans ces algorithmes. Elles n'interviennent que dans le calcul des poids $\tilde{w}_{i,j,(m)}$ des simulations des paramètres aléatoires lors des étapes M.

Nous prenons ici l'exemple de l'étape M d'un algorithme MCEM pour décrire la logique de construction des algorithmes, MCEM ou SAEM, que nous avons employés. L'étape M de

l'itération $m+1$ est composée des séquences de problèmes :

$$(41)\qquad \begin{cases} (\hat{\boldsymbol{\mu}}_{(m+1)}, \hat{\boldsymbol{\Omega}}_{(m+1)}) = \arg\max_{(\boldsymbol{\mu},\boldsymbol{\Omega})} \sum_{i=1}^{N}\sum_{j=1}^{J} \tilde{w}_{i,j}^{(m)} \ln \varphi(\tilde{\boldsymbol{\gamma}}_{i,j}^{(m)} - \boldsymbol{\mu}; \boldsymbol{\Omega}) \\ (\hat{\boldsymbol{\delta}}_{(m+1)}, \hat{\boldsymbol{\Psi}}_{(m+1)}) = \operatorname{argmax}_{(\boldsymbol{\delta},\boldsymbol{\Psi})} \left\{ \sum_{i=1}^{N}\sum_{j=1}^{J} \tilde{w}_{i,j}^{(m)} \ln \ell_i^{+}(\boldsymbol{\delta}, \boldsymbol{\Psi} \mid \tilde{\boldsymbol{\gamma}}_{i,j}^{(m)}) \right\} \end{cases}$$

dans une version classique d'une étape M ou :

$$(42)\qquad \begin{cases} (\hat{\boldsymbol{\mu}}_{(m+1)}, \hat{\boldsymbol{\Omega}}_{(m+1)}) = \arg\max_{(\boldsymbol{\mu},\boldsymbol{\Omega})} \sum_{i=1}^{N}\sum_{j=1}^{J} \tilde{w}_{i,j}^{(m)} \ln \varphi(\tilde{\boldsymbol{\gamma}}_{i,j}^{(m)} - \boldsymbol{\mu}; \boldsymbol{\Omega}) \\ \text{Calculer une valeur de } (\hat{\boldsymbol{\delta}}_{(m+1)}, \hat{\boldsymbol{\Psi}}_{(m+1)}) \text{ vérifiant la condition :} \\ \sum_{i=1}^{N}\sum_{j=1}^{J} \tilde{w}_{i,j}^{(m)} \ln \ell_i^{+}(\hat{\boldsymbol{\delta}}_{(m+1)}, \hat{\boldsymbol{\Psi}}_{(m+1)} \mid \tilde{\boldsymbol{\gamma}}_{i,j}^{(m)}) > \sum_{i=1}^{N}\sum_{j=1}^{J} \tilde{w}_{i,j}^{(m)} \ln \ell_i^{+}(\hat{\boldsymbol{\delta}}_{(m)}, \hat{\boldsymbol{\Psi}}_{(m)} \mid \tilde{\boldsymbol{\gamma}}_{i,j}^{(m)}) \end{cases}$$

dans une version généralisée d'étape M. Bien entendu, l'étape M généralisée est moins gourmande en temps de calcul. Mais, aucune de ces versions ne garantit que la condition de convergence des algorithmes considérés vers un maximum de la vraisemblance. En effet, si ces étapes M garantissent, par construction, les conditions:

$$(43)\qquad \sum_{i=1}^{N}\sum_{j=1}^{J} \tilde{w}_{i,j}^{(m)} \ln \varphi(\tilde{\boldsymbol{\gamma}}_{i,j}^{(m)} - \hat{\boldsymbol{\mu}}_{(m+1)}; \hat{\boldsymbol{\Omega}}_{(m+1)}) > \sum_{i=1}^{N}\sum_{j=1}^{J} \tilde{w}_{i,j}^{(m)} \ln \varphi(\tilde{\boldsymbol{\gamma}}_{i,j}^{(m)} - \hat{\boldsymbol{\mu}}_{(m)}; \hat{\boldsymbol{\Omega}}_{(m)})$$

et :

$$(44)\qquad \sum_{i=1}^{N}\sum_{j=1}^{J} \tilde{w}_{i,j}^{(m)} \ln \ell_i^{+}(\hat{\boldsymbol{\delta}}_{(m+1)}, \hat{\boldsymbol{\Psi}}_{(m+1)} \mid \tilde{\boldsymbol{\gamma}}_{i,j}^{(m)}) > \sum_{i=1}^{N}\sum_{j=1}^{J} \tilde{w}_{i,j}^{(m)} \ln \ell_i^{+}(\hat{\boldsymbol{\delta}}_{(m)}, \hat{\boldsymbol{\Psi}}_{(m)} \mid \tilde{\boldsymbol{\gamma}}_{i,j}^{(m)})$$

elles ne garantissent pas la condition $\tilde{Q}_{N,J}(\hat{\boldsymbol{\theta}}_{(m+1)}; \hat{\boldsymbol{\theta}}_{(m)}) > \tilde{Q}_{N,J}(\hat{\boldsymbol{\theta}}_{(m)}; \hat{\boldsymbol{\theta}}_{(m)})$ qui est nécessaire pour le « fonctionnement correcte » d'un algorithme de type GEM car on peut avoir :

$$(45)\qquad \sum_{i=1}^{N}\sum_{j=1}^{J} \tilde{w}_{i,j}^{(m)} \ln \ell_i^{r}(\hat{\boldsymbol{\delta}}_{(m+1)}, \hat{\boldsymbol{\Psi}}_{(m+1)} \mid \tilde{\boldsymbol{\gamma}}_{i,j}^{(m)}) < \sum_{i=1}^{N}\sum_{j=1}^{J} \tilde{w}_{i,j}^{(m)} \ln \ell_i^{r}(\hat{\boldsymbol{\delta}}_{(m)}, \hat{\boldsymbol{\Psi}}_{(m)} \mid \tilde{\boldsymbol{\gamma}}_{i,j}^{(m)}).$$

Dit autrement, il est possible que la réactualisation des probabilités des choix de régime aboutissent à une diminution de $\tilde{Q}_{N,J}(\boldsymbol{\theta}; \hat{\boldsymbol{\theta}}_{(m)})$ par rapport à $\tilde{Q}_{N,J}(\hat{\boldsymbol{\theta}}_{(m)}; \hat{\boldsymbol{\theta}}_{(m)})$.

Nous avons employé une solution heuristique simple pour contourner ce type de problème.[25] Cependant, cette heuristique n'a été que très peu activée dans nos travaux empiriques. Plusieurs raisons peuvent expliquer ce constat. Nos résultats d'estimation suggèrent que le contenu informatif de nos variables de contrôle est assez limité. De ce fait, les variations des estimations de $\boldsymbol{\delta}_0$ affectent peu le calcul des probabilités de choix de régime, tout comme elles affectent assez peu les autres parties de la vraisemblance de nos modèles. Ensuite, les valeurs des estimations de $\boldsymbol{\Psi}_0$ n'affectent la probabilité de choix de régime que *via* leur impact sur les valeurs prédites du sous-vecteur des termes d'erreur $\boldsymbol{\varepsilon}_{it}^{s,0}$. Cette « porte d'entrée » des effets des variations des estimations de $\boldsymbol{\Psi}_0$ sur les $\ln \ell_i^r(\boldsymbol{\delta}, \boldsymbol{\Psi} \mid \tilde{\boldsymbol{\gamma}}_{i,j}^{(m)})$ est relativement petite, et ce d'autant plus que ce vecteur est petit dans les régimes dans lesquels beaucoup de cultures sont assolées. Enfin, nos résultats d'estimation suggèrent que la distribution des paramètres aléatoires tend à dominer les autres éléments de spécification des

---

[25] Si la condition $\tilde{Q}_{N,J}(\boldsymbol{\theta}_{(m+1)};\hat{\boldsymbol{\theta}}_{(m)}) < \tilde{Q}_{N,J}(\hat{\boldsymbol{\theta}}_{(m)};\hat{\boldsymbol{\theta}}_{(m)})$ survient alors (a) nous imposons $\hat{\boldsymbol{\theta}}_{(m+1)} = \hat{\boldsymbol{\theta}}_{(m)}$ de sorte à « refuser » la valeur réactualisée $\hat{\boldsymbol{\theta}}_{(m+1)}$, (*b*) nous accroissons le nombre *J* de tirages utilisés pour calculer $\tilde{Q}_{N,J}(\boldsymbol{\theta};\hat{\boldsymbol{\theta}}_{(m)})$ et (*c*) nous réitérons l'étape M pour obtenir une seconde valeur réactualisée de $\boldsymbol{\theta}$, $\hat{\boldsymbol{\theta}}_{(m+1),2}$. Si $\tilde{Q}_{N,J}(\boldsymbol{\theta}_{(m+1),2};\hat{\boldsymbol{\theta}}_{(m)}) > \tilde{Q}_{N,J}(\hat{\boldsymbol{\theta}}_{(m)};\hat{\boldsymbol{\theta}}_{(m)})$ nous reprenons l'exécution de l'algorithme à partir de $\hat{\boldsymbol{\theta}}_{(m+1),2}$. Dans le cas contraire nous réitérons les étapes (*a*)-(*b*) et obtenons une troisième valeur de $\boldsymbol{\theta}$, $\hat{\boldsymbol{\theta}}_{(m+1),3}$. Si $\tilde{Q}_{N,J}(\boldsymbol{\theta}_{(m+1),3};\hat{\boldsymbol{\theta}}_{(m)}) > \tilde{Q}_{N,J}(\hat{\boldsymbol{\theta}}_{(m)};\hat{\boldsymbol{\theta}}_{(m)})$ nous reprenons l'exécution de l'algorithme à partir de $\hat{\boldsymbol{\theta}}_{(m+1),3}$. Dans le cas contraire nous reprenons l'exécution de l'algorithme à partir de $\hat{\boldsymbol{\theta}}_{(m+1)}$.

modèles estimés. Dans une certaine mesure, les estimations des paramètres $(\boldsymbol{\delta}_0, \boldsymbol{\Psi}_0)$ s'adaptent plus à celles de $(\boldsymbol{\mu}_0, \boldsymbol{\Omega}_0)$ que l'inverse. Ce constat est cohérent avec les résultats des analyses descriptives des données employées pour nos estimations. Dans ces données, la variabilité inter-individuelle des choix de production, et des choix de régime de production en particulier, est nettement plus important que leur variabilité intertemporelle.[26]

Nous abordons dans la section suivante la question de l'actualisation des paramètres $(\boldsymbol{\delta}, \boldsymbol{\Psi})$. En effet, l'approche adoptée pour cette actualisation conditionne la mise en œuvre des algorithmes MCEM et SAEM considérés par la suite.

4.3. Rendements, quantités d'intrants et assolements « manquants »

Comme indiqué précédemment, les fonctions de log-vraisemblance conditionnelle en $(\boldsymbol{\gamma}_i, \mathbf{z}_{it})$ des rendements, quantités d'intrants et choix d'assolement observés $(\mathbf{y}_{it}^+, \mathbf{x}_{it}^+, \mathbf{s}_{it}^+)$ ont des log-vraisemblances d'un modèle bien connu. En effet, on a :

(46) $$\ln \ell_{it}^+(\boldsymbol{\delta}, \boldsymbol{\Psi} \mid \boldsymbol{\gamma}) = b_{it}(\boldsymbol{\gamma}) + \ln \varphi\left(\boldsymbol{\varepsilon}_{it}^+(\boldsymbol{\gamma}, \boldsymbol{\delta}); \boldsymbol{\Psi}_{it}^{++}\right)$$

avec :

(47) $$b_{it}(\boldsymbol{\gamma}) = \left(1 - G^+(r_{it})\right) \ln \alpha^s + \left(1 - K_{(g)}^+(r_{it})\right) \sum_{g \in \mathcal{G}^+(r_{it})} \ln \alpha_{(g)}^s - \sum_{k \in \mathcal{K}^+(r_{it})} \ln s_{k,it} \ .$$

---

[26] Ce constat étant d'ailleurs à l'origine de nos travaux sur la prise en compte de l'hétérogénéité inobservée dans les modèles de choix de production agricole.

Le terme $\ln \varphi\left(\boldsymbol{\varepsilon}_{it}^{+}(\boldsymbol{\gamma},\boldsymbol{\delta});\boldsymbol{\Psi}_{it}^{++}\right)$ correspond à la log-vraisemblance d'un système de régressions linéaires ordinaires Gaussiennes à données manquantes (voir, *e.g.*, Ruud, 1991), celui donné par l'équation :

(48) $$\varsigma_{it}^{+}(\boldsymbol{\gamma}) = (\mathbf{Z}_{it}^{+})'\boldsymbol{\delta} + \boldsymbol{\varepsilon}_{it}^{+}(\boldsymbol{\gamma},\boldsymbol{\delta}),$$

le mécanisme de sélection des données observées ou non étant défini par le choix du régime de production. Seules les « variables dépendantes » $\varsigma_{it}^{+}(\boldsymbol{\gamma})$ sont manquantes ici. Les « variables explicatives » $\mathbf{Z}_{it}^{+}$ peuvent être construites pour tout régime de production et pour toute observation (i,t).

La forme fonctionnelle des termes $\ln \varphi\left(\boldsymbol{\varepsilon}_{it}^{+}(\boldsymbol{\gamma},\boldsymbol{\delta},r_{it});\boldsymbol{\Psi}_{it}^{++}\right)$ implique que la log-vraisemblance:

(49) $$\begin{aligned}\ln \tilde{L}_{N,J}(\boldsymbol{\delta},\boldsymbol{\Psi}) &\equiv \sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{j=1}^{J}\tilde{w}_{i,j}^{(m)} \ln \ell_{it}^{+}(\boldsymbol{\delta},\boldsymbol{\Psi} \mid \tilde{\boldsymbol{\gamma}}_{i,j}^{(m)}) \\ &= \sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{j=1}^{J}\tilde{w}_{i,j}^{(m)} \ln \varphi\left(\boldsymbol{\varepsilon}_{it}^{+}(\boldsymbol{\gamma},\boldsymbol{\delta});\boldsymbol{\Psi}_{it}^{++}\right)\end{aligned}$$

est difficile à maximiser ou à « augmenter » pour obtenir directement $(\hat{\boldsymbol{\delta}}_{(m+1)},\hat{\boldsymbol{\Psi}}_{(m+1)})$.[27]

Ruud (1991) a montré que l'utilisation d'un algorithme EM facilite la maximisation de la log-vraisemblance $\ln \tilde{L}_{N,J}(\boldsymbol{\delta},\boldsymbol{\Psi})$ dès lors que, dans notre cas, les choix de production – qui auraient été décidés par les agriculteurs – relatifs aux cultures non assolées sont considérés

[27] Nous omettons ici les termes $b_{it}(\boldsymbol{\gamma})$ puisqu'ils ne dépendent pas des paramètres $(\boldsymbol{\delta},\boldsymbol{\Psi})$.

comme des données manquantes. Cette approche est décrite ici.

Néanmoins, il convient d'être prudent dans cette interprétation en termes de données manquantes. En effet, s'il est possible d'imaginer des rendements et des quantités d'intrants qui auraient pu être obtenus et employées par les agriculteurs, le choix d'assolement des agriculteurs est ce qu'il est.

De fait plus qu'un problème de données de choix de production manquants, il s'agit ici de traiter un problème de termes d'erreurs manquants : les termes $\boldsymbol{\varepsilon}_{it}^{0}$. Or ces termes ont une interprétation essentielle. Le terme $\boldsymbol{\varepsilon}_{it}^{s,0}$ est connu des agriculteurs et impacte directement leur choix d'assolement et de régime de production. Le terme $\boldsymbol{\varepsilon}_{it}^{yx,0}$ définit les effets d'évènements aléatoires qui se sont produits et qui auraient impacté les rendements obtenus et les quantités d'intrants utilisées par l'agriculteur *i* s'il avait assolé en *t* les cultures de $\mathcal{K}^{0}(r_{it})$.

Techniquement, considérer les termes $\boldsymbol{\varepsilon}_{it}^{0}$ comme des variable manquantes revient à écrire la log-vraisemblance $\ln\varphi\left(\boldsymbol{\varepsilon}_{it}^{+}(\boldsymbol{\gamma},\boldsymbol{\delta});\boldsymbol{\Psi}_{it}^{++}\right)$ sous la forme :

(50) $$\varphi\left(\boldsymbol{\varepsilon}_{it}^{+}(\boldsymbol{\gamma},\boldsymbol{\delta});\boldsymbol{\Psi}_{it}^{++}\right)=\int\varphi\left(\mathbf{H}_{it}^{0}\boldsymbol{\varepsilon}_{it}^{0}+\mathbf{H}_{it}^{+}\boldsymbol{\varepsilon}_{it}^{+}(\boldsymbol{\gamma},\boldsymbol{\delta});\boldsymbol{\Psi}\right)d\boldsymbol{\varepsilon}_{it}^{0}$$

en utilisant le fait que :

(51) $$\boldsymbol{\varepsilon}_{it}=\mathbf{H}_{it}^{0}\boldsymbol{\varepsilon}_{it}^{0}+\mathbf{H}_{it}^{+}\boldsymbol{\varepsilon}_{it}^{+}=\mathbf{H}_{it}^{0}\boldsymbol{\varepsilon}_{it}^{0}+\mathbf{H}_{it}^{+}\boldsymbol{\varepsilon}_{it}^{+}(\boldsymbol{\gamma}_{i},\boldsymbol{\delta}_{0})\quad\text{et}\quad\boldsymbol{\varepsilon}_{it}\sim\mathcal{N}(\mathbf{0},\boldsymbol{\Psi}_{0}).$$

L'algorithme EM, en considérant $\boldsymbol{\varepsilon}_{it}^{0}$ comme une donnée manquante dans le problème de la maximisation de $\ln\tilde{L}_{N,J}(\boldsymbol{\delta},\boldsymbol{\Psi})$, s'applique en itérant les étapes E en M suivantes.

L'étape E de l'itération de l'itération $n+1$ consiste à calculer, pour tout triplet (i,t,j), la log-vraisemblance des données complètes, $\ln\varphi\left(\mathbf{H}_{it}^{0}\boldsymbol{\varepsilon}_{it}^{0}+\mathbf{H}_{it}^{+}\boldsymbol{\varepsilon}_{it}^{+}(\boldsymbol{\gamma},\boldsymbol{\delta});\boldsymbol{\Psi}\right)$, sur la distribution des variables manquantes $\boldsymbol{\varepsilon}_{it}^{0}$ conditionnelle en les variables observées, *i.e.* en $\boldsymbol{\varepsilon}_{it}^{+}$ ici.[28] Puisqu'on sait que $\boldsymbol{\varepsilon}_{it}^{0}\mid\boldsymbol{\varepsilon}_{it}^{+}\sim\mathcal{N}(\boldsymbol{\Gamma}_{0,it}^{0|+}\boldsymbol{\varepsilon}_{it}^{s,+},\boldsymbol{\Psi}_{0,it}^{0|+})$ et $\boldsymbol{\varepsilon}_{it}^{+}(\boldsymbol{\gamma},\boldsymbol{\delta})=\boldsymbol{\varsigma}_{it}^{+}(\boldsymbol{\gamma})-(\mathbf{Z}_{it}^{+})'\boldsymbol{\delta}$, cette espérance est donnée, à un terme constant près, par:

$$
\begin{aligned}
&E_{(n)}\left[\ln\varphi\left(\mathbf{H}_{it}^{0}\boldsymbol{\varepsilon}_{it}^{0}+\mathbf{H}_{it}^{+}\boldsymbol{\varepsilon}_{it}^{+}(\tilde{\boldsymbol{\gamma}}_{i,j}^{(m)},\boldsymbol{\delta});\boldsymbol{\Psi}\right)\right]\\
&\quad=-1/2\times\left\{\ln\det\boldsymbol{\Psi}+tr\left\{\begin{array}{c}\boldsymbol{\Psi}^{-1}\left[\hat{\boldsymbol{\Delta}}_{(n),it}\boldsymbol{\varepsilon}_{it}^{+}(\tilde{\boldsymbol{\gamma}}_{i,j}^{(m)},\boldsymbol{\delta})\boldsymbol{\varepsilon}_{it}^{+}(\tilde{\boldsymbol{\gamma}}_{i,j}^{(m)},\boldsymbol{\delta})'(\hat{\boldsymbol{\Delta}}_{(n),it})'\right]\\+\boldsymbol{\Psi}^{-1}\mathbf{H}_{it}^{0}\hat{\boldsymbol{\Psi}}_{(n),it}^{0|+}(\mathbf{H}_{it}^{0})'\end{array}\right\}\right\}
\end{aligned}
\tag{52}
$$

où :

$$
\hat{\boldsymbol{\Delta}}_{(n),it}=\mathbf{H}_{it}^{+}+\mathbf{H}_{it}^{0}\hat{\boldsymbol{\Gamma}}_{(n),it}^{0|+}\quad\text{et}\quad\boldsymbol{\varepsilon}_{it}^{+}(\tilde{\boldsymbol{\gamma}}_{i,j}^{(m)},\boldsymbol{\delta})=\boldsymbol{\varsigma}_{it}^{+}(\tilde{\boldsymbol{\gamma}}_{i,j}^{(m)})-(\mathbf{Z}_{it}^{+})'\boldsymbol{\delta}
\tag{53}
$$

L'étape M de l'itération $n+1$ consiste à calculer :

$$
\hat{\boldsymbol{\Psi}}_{(n+1)}=\arg\max_{\boldsymbol{\Psi}}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{j=1}^{J}\tilde{w}_{i,j}^{(m)}E_{(n)}\left[\ln\varphi\left(\mathbf{H}_{it}^{0}\boldsymbol{\varepsilon}_{it}^{0}+\mathbf{H}_{it}^{+}\boldsymbol{\varepsilon}_{it}^{+}(\tilde{\boldsymbol{\gamma}}_{i,j}^{(m)},\hat{\boldsymbol{\delta}}_{(n)});\boldsymbol{\Psi}\right)\right]
\tag{54}
$$

puis :

$$
\hat{\boldsymbol{\delta}}_{(n+1)}=\arg\max_{\boldsymbol{\delta}}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{j=1}^{J}\tilde{w}_{i,j}^{(m)}E_{(n)}\left[\ln\varphi\left(\mathbf{H}_{it}^{0}\boldsymbol{\varepsilon}_{it}^{0}+\mathbf{H}_{it}^{+}\boldsymbol{\varepsilon}_{it}^{+}(\tilde{\boldsymbol{\gamma}}_{i,j}^{(m)},\boldsymbol{\delta});\hat{\boldsymbol{\Psi}}_{(n+1)}\right)\right].
\tag{55}
$$

Etant donnée la linéarité en $\boldsymbol{\delta}$ du modèle considéré, les termes $\hat{\boldsymbol{\Psi}}_{(n+1)}$ et $\hat{\boldsymbol{\delta}}_{(n+1)}$ sont obtenus sous forme analytique. Ces dernières seront donnés ci-dessous. Les termes $\hat{\boldsymbol{\Psi}}_{(n+1)}$ et $\hat{\boldsymbol{\delta}}_{(n+1)}$

---

[28] Sous les hypothèses posées pour le modèle considérée le terme $\boldsymbol{\varepsilon}_{it}$ est indépendant des autres composantes des log-vraisemblances considérées ici.

peuvent également être obtenus à partir des statistiques suffisantes du problème d'estimation considéré, cette approche s'avérant être utile pour les algorithme SAEM.

L'algorithme décrit ici sera désigné comme l'algorithme de Ruud dans la suite. Ces résultats peuvent être employés pour définir plusieurs types d'algorithme MCEM ou SAEM, ce qui fait l'objet des sections suivantes.

4.4. Algorithmes MCEM

Deux types d'algorithmes peuvent être conçus sur la base des résultats présentés ci-dessus, en fonction de la position donnée aux $\boldsymbol{\varepsilon}_{it}^{0}$ manquants. Il est possible de considérer les $\boldsymbol{\varepsilon}_{it}^{0}$ manquants dans la définition de l'algorithme MCEM lui-même ou seulement pour maximiser la log-vraisemblance des observations à chaque itération de l'algorithme MCEM principal.

Si l'algorithme MCEM est défini pour gérer le problème des paramètres aléatoires, alors l'algorithme de Ruud est exécuté en boucle interne avec :

Etape E, itération $m+1$

1. Tirage des $\tilde{\boldsymbol{\gamma}}_{i,j}^{(m)}$ suivant $\mathcal{N}(\hat{\boldsymbol{\mu}}_{(m)}, \hat{\boldsymbol{\Omega}}_{(m)})$ et calcul des $\tilde{w}_{i,j}^{(m)}$ pour $j = 1,...,J$ et $i = 1,...,N$
2. Calcul des $\tilde{w}_{i,j}^{(m)}$ pour $j = 1,...,J$ et $i = 1,...,N$
3. Calcul implicite de $\sum_{i=1}^{N} \sum_{j=1}^{J} \tilde{w}_{i,j}^{(m)} \ln \varphi(\tilde{\boldsymbol{\gamma}}_{i,j}^{(m)} - \boldsymbol{\mu}; \boldsymbol{\Omega})$
4. Calcul implicite de $\sum_{i=1}^{N} \sum_{j=1}^{J} \tilde{w}_{i,j}^{(m)} \ln \ell_i^{+}(\boldsymbol{\delta}, \boldsymbol{\Psi} \mid \tilde{\boldsymbol{\gamma}}_{i,j}^{(m)})$

Etape M :

1. Calcul de $\hat{\boldsymbol{\mu}}_{(m+1)} = N^{-1}\sum_{i=1}^{N}\sum_{j=1}^{J}\tilde{w}_{i,j}^{(m)}\tilde{\boldsymbol{\gamma}}_{i,j}^{(m)}$

2. Calcul de $\hat{\boldsymbol{\Omega}}_{(m+1)} = N^{-1}\sum_{i=1}^{N}\sum_{j=1}^{J}\tilde{w}_{i,j}^{(m)}\tilde{\boldsymbol{\gamma}}_{i,j}^{(m)}(\tilde{\boldsymbol{\gamma}}_{i,j}^{(m)})' - \hat{\boldsymbol{\mu}}_{(m+1)}\hat{\boldsymbol{\mu}}'_{(m+1)}$

3. Calcul de $(\hat{\boldsymbol{\delta}}_{(m+1)}, \hat{\boldsymbol{\Psi}}_{(m+1)}) = \text{argmax}_{(\boldsymbol{\delta},\boldsymbol{\Psi})}\sum_{i=1}^{N}\sum_{j=1}^{J}\tilde{w}_{i,j}^{(m)}\ln \ell_i^{+}(\boldsymbol{\delta},\boldsymbol{\Psi} \mid \tilde{\boldsymbol{\gamma}}_{i,j}^{(m)})$ *via* l'algorithme de Ruud en boucle interne.

Cet algorithme est gourmand en temps de calcul même si l'algorithme de Ruud (1991) est rapide.

Considérer les $\boldsymbol{\varepsilon}_{it}^{0}$ manquants dans la conception de l'algorithme MCEM revient à conduire les réactualisations de l'ensemble des paramètres simultanément avec :

Etape E :

1. Tirage des $\tilde{\boldsymbol{\gamma}}_{i,j}^{(m)}$ suivant $\mathcal{N}(\hat{\boldsymbol{\mu}}_{(m)}, \hat{\boldsymbol{\Omega}}_{(m)})$ et calcul des $\tilde{w}_{i,j}^{(m)}$ pour $j = 1,...,J$ et $i = 1,...,N$

2. Calcul des $\tilde{w}_{i,j}^{(m)}$ pour $j = 1,...,J$ et $i = 1,...,N$

3. Calcul implicite de $\sum_{i=1}^{N}\sum_{j=1}^{J}\tilde{w}_{i,j}^{(m)}\ln\varphi(\tilde{\boldsymbol{\gamma}}_{i,j}^{(m)} - \boldsymbol{\mu};\boldsymbol{\Omega})$

4. Calcul implicite de $\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{j=1}^{J}\tilde{w}_{i,j}^{(m)}E_{(m)}\left[\ln\varphi\left(\mathbf{H}_{it}^{0}\boldsymbol{\varepsilon}_{it}^{0} + \mathbf{H}_{it}^{+}\boldsymbol{\varepsilon}_{it}^{+}(\tilde{\boldsymbol{\gamma}}_{i,j}^{(m)}, \hat{\boldsymbol{\delta}}_{(m)}); \boldsymbol{\Psi}\right)\right]$

Etape M :

1. Calcul de $\hat{\boldsymbol{\mu}}_{(m+1)} = N^{-1}\sum_{i=1}^{N}\sum_{j=1}^{J}\tilde{w}_{i,j}^{(m)}\tilde{\boldsymbol{\gamma}}_{i,j}^{(m)}$

2. Calcul de $\hat{\boldsymbol{\Omega}}_{(m+1)} = N^{-1}\sum_{i=1}^{N}\sum_{j=1}^{J}\tilde{w}_{i,j}^{(m)}\tilde{\boldsymbol{\gamma}}_{i,j}^{(m)}(\tilde{\boldsymbol{\gamma}}_{i,j}^{(m)})' - \hat{\boldsymbol{\mu}}_{(m+1)}\hat{\boldsymbol{\mu}}'_{(m+1)}$

3. Calcul de $(\hat{\boldsymbol{\delta}}_{(m+1)}, \hat{\boldsymbol{\Psi}}_{(m+1)})$ selon une itération de l'algorithme de Ruud, avec :

$$\hat{\mathbf{\Psi}}_{(n+1)} = (NT)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{j=1}^{J}\tilde{w}_{i,j}^{(m)}\hat{\mathbf{\Delta}}_{(n),it}\boldsymbol{\varepsilon}_{it}^{+}(\tilde{\boldsymbol{\gamma}}_{i,j}^{(m)},\hat{\boldsymbol{\delta}}_{(n)})\boldsymbol{\varepsilon}_{it}^{+}(\tilde{\boldsymbol{\gamma}}_{i,j}^{(m)},\hat{\boldsymbol{\delta}}_{(n)})'(\hat{\mathbf{\Delta}}_{(n),it})' + (NT-C)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\mathbf{H}_{it}^{0}\hat{\mathbf{\Psi}}_{(n),it}^{0|+}(\mathbf{H}_{it}^{0})' \tag{56}$$

et :

$$\hat{\boldsymbol{\delta}}_{(n+1)} = \left(\sum_{i=1}^{N}\sum_{t=1}^{T}\mathbf{Z}_{it}^{+}(\hat{\mathbf{\Delta}}_{(n),it})'(\hat{\mathbf{\Psi}}_{(n+1)})^{-1}\hat{\mathbf{\Delta}}_{(n),it}(\mathbf{Z}_{it}^{+})'\right)^{-1} \times\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{j=1}^{J}\tilde{w}_{i,j}^{(m)}\mathbf{Z}_{it}^{+}(\hat{\mathbf{\Delta}}_{(n),it})'(\hat{\mathbf{\Psi}}_{(n+1)})^{-1}\hat{\mathbf{\Delta}}_{(n),it}\boldsymbol{\varsigma}_{it}^{+}(\tilde{\boldsymbol{\gamma}}_{i,j}^{(m)}). \tag{57}$$

Dans l'expression de $\hat{\mathbf{\Psi}}_{(n+1)}$ le terme C est le nombre d'observations caractérisées par le régime c, le régime « complet » dans lequel les agriculteurs assolent toutes les cultures de $\mathcal{K}$. On impose $\mathbf{H}_{it}^{0} = \mathbf{0}$ et $\hat{\mathbf{\Psi}}_{(n),it}^{0|+} = \mathbf{0}$ pour $r_{it} = c$.

Les performances de ces deux algorithmes ont été assez médiocres lors de nos essais, avec de fréquents problèmes de convergence. L'algorithme SAEM présenté dans la section suivante s'est montré bien plus performant.

4.5. Algorithme SAEM

Jusqu'à présent nous avons adopté une approche d'estimation par les algorithmes MCEM. Dans cette section nous présentons un algorithme de type SAEM (Delyon *et al*, 1999 ; Lavielle, 2014). L'algorithme SAEM est construit sur le principe des algorithmes (G)EM et utilise des simulations des paramètres aléatoires à l'instar des algorithmes MCEM. Mais les deux algorithmes utilisent différemment les simulations. Le principe des algorithmes MCEM est d'intégrer la fonction $Q_N(\boldsymbol{\theta};\hat{\boldsymbol{\theta}}_{(m)})$ par simulations et d'utiliser l'approximation obtenue de

cette fonction objectif comme celle d'un algorithme (G)EM déterministe.

Le principe des algorithmes SAEM est de maximiser une approximation stochastique de la fonction $Q_N(\boldsymbol{\theta}) = Q_N(\boldsymbol{\theta};\boldsymbol{\theta})$. De fait, les algorithmes SAEM et MCEM génèrent et utilisent différemment les simulations des paramètres aléatoires. Un grand nombre de tirages des paramètres aléatoires doit être réalisé (ou réutilisé) pour l'étape E d'un algorithme MCEM. Il s'agit ici d'approximer correctement $Q_N(\boldsymbol{\theta};\hat{\boldsymbol{\theta}}_{(m)})$ à chaque itération par une nouvelle procédure d'intégration par simulation. Un algorithme SAEM n'utilise pas d'étape E à proprement parler. Un algorithme SAEM ne comporte pas d'étape E, mais une étape S et une étape SA.

L'étape S consiste à réaliser des tirages des paramètres aléatoires, mais en nombre plus limité que pour une étape E d'algorithme MCEM. Ces tirages sont utilisés pour construire de nouvelles approximations des $E_{(m),i}[\ln \ell_i^C(\boldsymbol{\theta};\boldsymbol{\gamma}_i)]$ qui seront ensuite utilisées dans l'étape SA afin d'améliorer l'approximation stochastique de $Q_N(\boldsymbol{\theta})$ dont la construction repose sur les tirages réalisés lors des itérations précédentes et se poursuit tout au long de l'exécution de l'algorithme SAEM.

Ici nous considérons une étape S qui consiste à construire approximation par simulation des termes $E_{(m),i}[\ln \ell_i^C(\boldsymbol{\theta};\boldsymbol{\gamma}_i)]$ à partir de J tirages, les $\tilde{\boldsymbol{\gamma}}_{i,j}^{(m)}$ pour $j = 1,...,J$, de $\boldsymbol{\gamma}_i$ suivant la distribution $\mathcal{N}(\hat{\boldsymbol{\mu}}_{(m)},\hat{\boldsymbol{\Omega}}_{(m)})$. Comme pour les algorithmes MCEM, nous utilisons les simulateurs IS $\sum_{j=1}^{J} \tilde{w}_{i,j}^{(m)} \ln \ell_i^C(\boldsymbol{\theta};\tilde{\boldsymbol{\gamma}}_{i,j}^{(m)})$.

L'étape SA de l'itération $m+1$ de l'algorithme SAEM construit l'approximation stochastique de $Q_N(\boldsymbol{\theta})$, $\tilde{Q}_{(m+1)}(\boldsymbol{\theta})$, à partir de l'approximation construite lors de l'itération précédente selon la formule récursive:

(58) $$\tilde{Q}_{(m+1)}(\boldsymbol{\theta}) \equiv \tilde{Q}_{(m)}(\boldsymbol{\theta}) + \vartheta_{m+1}\left(N^{-1}\sum_{i=1}^{N}\sum_{j=1}^{J}\tilde{w}_{i,j}^{(m)} \ln \ell_i^C(\boldsymbol{\theta};\tilde{\boldsymbol{\gamma}}_{i,j}^{(m)}) - \tilde{Q}_{(m)}(\boldsymbol{\theta})\right)$$

où :

(59) $$\vartheta_m = \begin{cases} 1 \text{ pour } m_1 \geq m \geq 1 \\ (m - m_1 + 1)^{3/4} \text{ pour } m \geq m_1 \end{cases}$$

Le paramètre ϑ_m définit le pas de l'approximation stochastique. Plus il est élevé moins $\tilde{Q}_{(m)}(\boldsymbol{\theta})$ a de poids dans $\tilde{Q}_{(m+1)}(\boldsymbol{\theta})$, avec un poids nul si $\vartheta_m = 1$. Choisir des ϑ_m élevés au début de l'exécution de l'algorithme garantit que le voisinage d'un maximum de la log-vraisemblance est atteint rapidement. Réduire les ϑ_m ensuite permet de réduire le bruit de simulation.

L'étape M de l'itération $m+1$ de l'algorithme SAEM consiste, comme dans les algorithmes (G)EM ou MCEM, à réactualiser la valeur de $\boldsymbol{\theta}$ par maximisation de la fonction $\tilde{Q}_{(m+1)}(\boldsymbol{\theta})$ ou en cherchant une valeur de $\boldsymbol{\theta}$ permettant d'accroître $\tilde{Q}_{(m+1)}(\boldsymbol{\theta})$ à partir de $\tilde{Q}_{(m+1)}(\hat{\boldsymbol{\theta}}_{(m)})$. La structure de notre modèle implique que la fonction $\tilde{Q}_{(m)}(\boldsymbol{\theta})$ peut être décomposé de manière additivement séparable avec :

(60) $$\tilde{Q}_{(m)}(\boldsymbol{\theta}) = \tilde{Q}_{(m)}^{\gamma}(\boldsymbol{\mu},\boldsymbol{\Omega}) + \tilde{Q}_{(m)}^{+}(\boldsymbol{\delta},\boldsymbol{\Psi}) + \tilde{Q}_{(m)}^{r}(\boldsymbol{\delta},\boldsymbol{\Psi})$$

où :

$\tilde{Q}^{\gamma}_{(m)}(\boldsymbol{\mu},\boldsymbol{\Omega})$ est une somme pondérée des $\sum_{i=1}^{N}\sum_{j=1}^{J}\tilde{w}_{i,j}^{(n)}\ln\varphi(\tilde{\boldsymbol{\gamma}}_{i,j}^{(n)}-\boldsymbol{\mu};\boldsymbol{\Omega})$,

$\tilde{Q}^{+}_{(m)}(\boldsymbol{\delta},\boldsymbol{\Psi})$ est une somme pondérée des $\sum_{i=1}^{N}\sum_{j=1}^{J}\tilde{w}_{i,j}^{(n)}\ln\ell_i^{+}(\boldsymbol{\delta},\boldsymbol{\Psi}\mid\tilde{\boldsymbol{\gamma}}_{i,j}^{(n)})$,

$\tilde{Q}^{r}_{(m)}(\boldsymbol{\delta},\boldsymbol{\Psi})$ est une somme pondérée des $\sum_{i=1}^{N}\sum_{j=1}^{J}\tilde{w}_{i,j}^{(n)}\ln\ell_i^{r}(\boldsymbol{\delta},\boldsymbol{\Psi}\mid\tilde{\boldsymbol{\gamma}}_{i,j}^{(n)})$

pour $n=1,....,m$. Les poids associés aux termes générés lors de l'étape S de l'itération *n* décroissant à mesure que *m* s'écarte de *n*.

L'étape M de notre algorithme SAEM exploite la décomposition de $\tilde{Q}_{(m+1)}(\boldsymbol{\theta})$ en définissant les différents éléments de $\hat{\boldsymbol{\theta}}_{(m+1)}$ à partir de la série de problèmes suivante :

(61)
$$\begin{cases} (\hat{\boldsymbol{\mu}}_{(m+1)},\hat{\boldsymbol{\Omega}}_{(m+1)}) = \arg\max_{(\boldsymbol{\mu},\boldsymbol{\Omega})}\tilde{Q}^{\gamma}_{(m+1)}(\boldsymbol{\mu},\boldsymbol{\Omega}) \\ \text{Calculer une valeur de } (\hat{\boldsymbol{\delta}}_{(m+1)},\hat{\boldsymbol{\Psi}}_{(m+1)}) \text{ vérifiant la condition :} \\ \tilde{Q}^{+}_{(m+1)}(\hat{\boldsymbol{\delta}}_{(m+1)},\hat{\boldsymbol{\Psi}}_{(m+1)}) > \tilde{Q}^{+}_{(m+1)}(\hat{\boldsymbol{\delta}}_{(m)},\hat{\boldsymbol{\Psi}}_{(m)}) \end{cases}$$

Bien entendu, cette étape M ne garantit pas que $\hat{\boldsymbol{\theta}}_{(m+1)}$ augmente $\tilde{Q}_{(m+1)}(\boldsymbol{\theta})$ à partir de $\tilde{Q}_{(m+1)}(\hat{\boldsymbol{\theta}}_{(m)})$. Mais, comme indiqué précédemment, ce type de problème est rarement survenu lors de nos travaux empiriques.

Dans le cas de notre modèle le terme $\hat{\boldsymbol{\theta}}_{(m+1)}$ s'obtient sous forme analytique en employant les statistiques suffisantes relatives aux vraisemblances impliquées dans le calcul de $\tilde{Q}_{(m+1)}(\boldsymbol{\theta})$. Avec les statistiques suffisantes :

(62)
$$\begin{cases} \mathbf{a}^{\gamma}_{(m+1)} = \mathbf{a}^{\gamma}_{(m)} + \vartheta_m \sum_{i=1}^{N} \sum_{j=1}^{J} \tilde{w}^{(m)}_{i,j} \tilde{\boldsymbol{\gamma}}^{(m)}_{i,j} \\ \mathbf{A}^{\gamma}_{(m+1)} = \mathbf{A}^{\gamma}_{(m)} + \vartheta_m \sum_{i=1}^{N} \sum_{j=1}^{J} \tilde{w}^{(m)}_{i,j} \tilde{\boldsymbol{\gamma}}^{(m)}_{i,j} \tilde{\boldsymbol{\gamma}}^{(m)\prime}_{i,j} \end{cases}$$

et :

(63)
$$\begin{cases} \mathbf{b}^{+}_{(m+1),it} = \mathbf{b}^{+}_{(m),it} + \vartheta_m \left(\sum_{j=1}^{J} \tilde{w}^{(m)}_{i,j} \hat{\boldsymbol{\Delta}}_{(n),it} \boldsymbol{\varsigma}^{+}_{it} (\tilde{\boldsymbol{\gamma}}^{(m)}_{i,j}) - \mathbf{b}^{+}_{(m),it} \right) \\ \mathbf{B}^{+}_{(m+1)} = \mathbf{B}^{+}_{(m)} + \vartheta_m \left(\sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{j=1}^{J} \tilde{w}^{(m)}_{i,j} \boldsymbol{\varsigma}^{+}_{it} (\tilde{\boldsymbol{\gamma}}^{(m)}_{i,j})' (\hat{\boldsymbol{\Delta}}_{(n),it})' \hat{\boldsymbol{\Delta}}_{(n),it} \boldsymbol{\varsigma}^{+}_{it} (\tilde{\boldsymbol{\gamma}}^{(m)}_{i,j}) - \mathbf{B}^{+}_{(m)} \right) \end{cases}$$

on obtient :

(64)
$$\hat{\boldsymbol{\mu}}_{(n+1)} = N^{-1} \mathbf{a}^{\gamma}_{(m+1)}$$

(65)
$$\hat{\boldsymbol{\Omega}}_{(n+1)} = N^{-1} \mathbf{A}^{\gamma}_{(m+1)} - \hat{\boldsymbol{\mu}}_{(n+1)} (\hat{\boldsymbol{\mu}}_{(n+1)})'$$

(66)
$$\begin{aligned} \hat{\boldsymbol{\Psi}}_{(n+1)} = {} & (NT)^{-1} \mathbf{B}^{+}_{(m+1)} + (NT - C)^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} \mathbf{H}^{0}_{it} \hat{\boldsymbol{\Psi}}^{0|+}_{(n),it} (\mathbf{H}^{0}_{it})' \\ & + (NT)^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} (\hat{\boldsymbol{\Delta}}_{(n),it} (\mathbf{Z}^{+}_{it})' \hat{\boldsymbol{\delta}}_{(n)} - \mathbf{b}^{+}_{(m+1),it}) (\hat{\boldsymbol{\Delta}}_{(n),it} (\mathbf{Z}^{+}_{it})' \hat{\boldsymbol{\delta}}_{(n)} - \mathbf{b}^{+}_{(m+1),it})' \\ & - (NT)^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} \mathbf{b}^{+}_{(m+1),it} (\mathbf{b}^{+}_{(m+1),it})' \end{aligned}$$

et :

(67)
$$\begin{aligned} \hat{\boldsymbol{\delta}}_{(n+1)} = {} & \left(\sum_{i=1}^{N} \sum_{t=1}^{T} \mathbf{Z}^{+}_{it} (\hat{\boldsymbol{\Delta}}_{(n),it})' (\hat{\boldsymbol{\Psi}}_{(n+1)})^{-1} \hat{\boldsymbol{\Delta}}_{(n),it} (\mathbf{Z}^{+}_{it})' \right)^{-1} \\ & \times \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{j=1}^{J} \tilde{w}^{(m)}_{i,j} \mathbf{Z}^{+}_{it} (\hat{\boldsymbol{\Delta}}_{(n),it})' (\hat{\boldsymbol{\Psi}}_{(n+1)})^{-1} \mathbf{b}^{+}_{(m+1),it} \end{aligned}$$

Il est possible d'itérer plusieurs fois le calcul des termes $\hat{\boldsymbol{\Psi}}_{(n+1)}$ et $\hat{\boldsymbol{\delta}}_{(n+1)}$ au sein d'une même étape M de sorte à accroître la valeur de $\tilde{Q}^{+}_{(m+1)}(\hat{\boldsymbol{\delta}}_{(m+1)}, \hat{\boldsymbol{\Psi}}_{(m+1)})$. En particulier, exécuter l'algorithme de Ruud jusqu'à son terme lors de l'étape M de chaque itération aurait transformé cette étape M en la série de problèmes de maximisation suivante :

$$(68) \qquad \begin{cases} (\hat{\boldsymbol{\mu}}_{(m+1)}, \hat{\boldsymbol{\Omega}}_{(m+1)}) = \arg\max_{(\boldsymbol{\mu},\boldsymbol{\Omega})} \tilde{Q}^{\gamma}_{(m+1)}(\boldsymbol{\mu},\boldsymbol{\Omega}) \\ (\hat{\boldsymbol{\delta}}_{(m+1)}, \hat{\boldsymbol{\Psi}}_{(m+1)}) = \arg\max_{(\boldsymbol{\delta},\boldsymbol{\Psi})} \tilde{Q}^{+}_{(m+1)}(\boldsymbol{\delta},\boldsymbol{\Psi}) \end{cases}$$

Néanmoins, nous n'avons pas utilisé cette option gourmande en temps de calcul.

4.6. Quelques remarques sur la mise en œuvre pratique des procédures d'estimation présentées

L'estimation de modèles multivariés à choix de régime endogène est difficile. De fait, considérer des modèles à paramètres aléatoires ne complique pas significativement le problème d'estimation pourvu qu'on utilise des algorithmes EM appropriés.

L'utilisation de l'approximation de Laplace proposée par Harding et Hausman (2007) pour l'intégration des probabilités de choix de régime sur les termes d'erreur $\boldsymbol{\varepsilon}_{it}^{s,0}$ des coûts implicites des cultures non assolées nous a grandement facilité la tâche. Dans une certaine mesure, elle nous a permis de résoudre un problème équivalent à celui de l'intégration des fonctions de répartition de lois normales multivariées de grande dimension que rencontrent la plupart des économètres travaillant avec des modèles multivariés à changement de régime endogène ou avec des systèmes de régressions censurées. Cette approximation est gourmande en temps de calcul mais elle nous a évité d'avoir recours à des techniques par simulations que nous utilisons assez massivement pour l'intégration sur la distribution des paramètres aléatoires du modèle.

S'il repose sur des calculs moins directs et intuitifs que ceux des algorithmes MCEM, l'algorithme SAEM s'est avéré être le plus performant dans nos travaux empiriques. Les

estimations présentées dans le Chapitre 6 ont été obtenues avec cet algorithme. Il s'est avéré être bien plus performant et robuste que ses contreparties MCEM.

Partie II

Hétérogénéité inobservée et choix de cultures dans les systèmes de grandes cultures : applications

Introduction

Nous avons, dans la première partie de cette thèse présenté les aspects de modélisation et d'estimation nécessaires pour aborder d'un point de vue empirique les questions posées dans cette thèse, à savoir la prise en compte de l'hétérogénéité inobservée (traitée dans les chapitres 1 et 2) et le traitement des solutions en coin (traité dans les chapitres 3 et 4) dans les modèles de choix de production multiculture. Il ressort de cette première partie qu'il est possible de traiter ces questions à partir de modèles cohérents d'un point de vue micro-économique et d'approches d'estimation disponibles en « théorie ». Pour vérifier s'ils sont applicables en pratique, il est donc nécessaire de confronter ces modèles et approches d'estimation à des données réelles. Ainsi, cette deuxième partie de la thèse complète la première partie en présentant deux applications empiriques se basant, bien sûr, sur les résultats de la première partie.

La deuxième partie se décompose en deux chapitres : les chapitres 5 et 6. Le chapitre 5 présente l'application empirique sur la prise en compte de l'hétérogénéité inobservée dans les modèles de choix de production multiculture. L'échantillon d'application concerne les agriculteurs en grandes cultures de la région de Meuse. Ce chapitre utilise les approches présentées dans les chapitres 1 et 2. Le chapitre 6 présente une application sur la prise en compte des problèmes de solution en coin dans les modèles de choix de production multiculture. L'échantillon d'application est la Marne. Les approches présentées dans les chapitres 3 et 4 sont ici mises à contribution.

Chapitre 5

On the heterogeneity in farmers' responsiveness to economic incentives: Random parameter multi-crop models and the effects of EU biofuel support[29]

[29] Ce chapitre est la version longue d'un article co-écrit avec Alain Carpentier et Fabienne Femenia, et accepté pour publication à l'American Journal of Agricultural Economist (AJAE)

Introduction

Farmers face different production conditions due to heterogeneous farm conditions such as soil quality, micro variations in climate or farm machinery. Farmers also differ because of heterogeneous personal characteristics such as skills, personal attitude toward the environment or aversion toward risk. Controlling for these differences when investigating farm-level production decisions is challenging because typically only a small number of farm and farmer-related variables are available. Following the pioneering work of Mundlak (1961), micro-econometric methods have been developed for coping with unobserved heterogeneity in panel data econometric models. Standard approaches account for the effects of unobserved heterogeneity on the levels of the modeled choices through the specification of additively separable random individual effects (Wooldridge 2010).

Accounting for the effects of unobserved heterogeneity on how economic phenomena respond to their drivers has also become an important research topic during the last two decades (see, *e.g.*, Wooldridge 2005, 2010; Arellano and Bonhomme 2011, 2012). The most common modeling approach for accounting for such effects consists of specifying individual specific random slope parameters for the variables describing the drivers of interest in the considered model. The related empirical applications generally demonstrate substantial

heterogeneity in the responses of firms, consumers or workers to homogeneous incentives.[30]

These potential impacts of unobserved heterogeneity have largely been ignored in panel data agricultural production choice models. These models account for unobserved heterogeneity effects on farmers' choice, typically by considering additively separable random farm effects but they generally assume that farmers' choices homogeneously respond to economic incentives. The effects of netput prices or of crop returns are typically specified as (fixed) slope parameters that are common to all farms.

Accounting for the variability of farmers' reactions to their economic drivers may be necessary to avoid substantial biases in estimation and simulation results, notwithstanding that characterizing this variability is of interest in itself. When farmers' responses to economic incentives exhibit significant unobserved heterogeneity that could be captured by random slope parameters, standard panel data models allow for consistently estimating the mean effects of these incentives only under specific conditions, including linearity of the considered effects. Estimation biases occur even in linear models when the incentive variables and their effects on farmers' choices are correlated across farms. Such biases are likely to be limited with price variables, but they can be significant with crop return variables. Prices are largely determined by factors that are external to individual farms, while crop returns strongly depend on factors that are farm specific.

---

[30] Examples include applications in labor economics (Heckman 2001; Angrist and Pischke 2009), empirical industrial organization (Ackerberg et al 2007), trade economics (Keane 2009; Eaton et al 2011) or production economics (Greene 2005).

The main objectives of this article are threefold. First, it provides empirical results demonstrating that unobserved heterogeneity can substantially impact the way farmers respond to economic incentives. Second, it shows that ignoring these features of farmers' choices can have important consequences for agricultural policy analyses based on micro-econometric models. Third, this article presents original tools – modeling, estimation and simulation approaches – that are specifically required for accounting for rich unobserved heterogeneity effects in panel data micro-econometric multi-crop models.

The empirical application we propose focuses on the production choices of a sample of French grain crop producers. Specifically, we examine how these choices were impacted by the rapeseed price increase that was induced by enhanced EU support of biofuels.[31] This case study, which spans the years 2006 to 2011, works well because the policy-induced price increase was particularly significant. Our estimation results demonstrate substantial variability in the responses of variable input uses and of yield levels to netput prices as well as in the responses of acreage choices to expected crop returns. Our simulation results demonstrate that ignoring this variability leads to significantly (*i*) overestimating farmers' yield supply and input use responses to EU biofuel support, (*ii*) underestimating the variability of farmers' acreage responses to this price increase and (*iii*) overestimating (underestimating) the small (large) increases in the rapeseed acreage.

[31] The EU's objective was to increase the use of biofuels in transportation fuels. In France, this policy was implemented using a tax that strongly encouraged the incorporation of rapeseed oil in diesel fuel. At the farm level, this policy basically led to an increase in the price of rapeseed.

The theoretical background of our analysis is a micro-econometric multi-crop model proposed by Carpentier and Letort (2014). It is defined as a theoretically consistent system of variable input demand, yield supply and acreage share choice equations. This model was chosen for the easy interpretation and parsimony of its parameters and for its explicitly considering expected crop gross margins, a key feature linking this model to the now widely used (parcel level) discrete crop choice models. We extend this model by allowing for random parameters as a way to account for unobserved heterogeneity. First, the intercept parameter of each equation is specified as the usual additively separable farm specific random term that is featured in standard panel data models. Second, the parameters that describe how farmers' choices respond to economic incentives are defined as farm specific random slope parameters.

This panel data random parameter (RP) multi-crop model is the first multi-crop micro-econometric model featuring random slope coefficients for describing the effects of the economic drivers on farmers' production choices. It extends the recent panel data multi-crop models of Lacroix and Thomas (2011), Platoni et al (2012) or Bayramoglu and Chakir (2016) by allowing for heterogeneous responses to economic incentives. In fact, models for describing acreage choices (see, *e.g.*, Fezzi and Bateman 2011; Carpentier and Letort 2012; 2014) or crop choices at the parcel level (see, *e.g.*, Wu et al 2004; Langpap and Wu 2011) mostly assume that these choices homogenously respond to economic drivers. The recent RP discrete crop choice model of Claassen et al (2017) is a notable exception in this respect. This RP model is an application of the latent class discrete choice model initially proposed by Greene and Hensher (2003) and widely used since then. It allows for heterogeneous

responses to economic incentives, but its random parameters and its explanatory variables, those describing the crop returns in particular, are assumed to be independent.

Challenging issues arise when considering the estimation of the RP acreage share model that is part of our RP multi-crop model and require specific modeling and estimation approaches. This RP acreage share model contains explanatory variables – the crop expected gross margins – that may be correlated with its random parameters. It also features non-linear functional forms in its original version and contains endogenous explanatory variables in its log-linearized version. These characteristics imply that the joint probability distribution of the random parameters of our RP acreage share model cannot be estimated based on simple adaptations of existing semi-parametric methods despite considerable recent advances on these estimation issues (see, *e.g.*, Murtazashvili and Wooldridge 2008, 2016; Keane 2009; Arellano and Bonhomme 2011, 2012).

Following Keane (2009), we show that these estimation issues can be overcome by (*a*) considering that the RP acreage share model is part of the theoretically consistent RP multi-crop model and (*b*) by adopting a parametric approach specifying the probability distributions of its random elements – farm specific parameters and error terms – up to a set of (deep) parameters to be estimated. Indeed, this full information approach explicitly defines the structure of the considered correlation and endogeneity issues. In particular, in the considered RP multi-crop model, the expected crop gross margins used in the acreage share sub-model are constructed based on elements of the yield supply and input demand sub-models. Distributional assumptions then enable to use maximum likelihood (ML) estimators that, by

construction, fully and efficiently take on the considered estimation problems.[32]

Apart from enabling the specification of an empirically tractable RP multi-crop model, distributional assumptions have two other important advantages for the purpose of this study. First, estimates of the farm specific parameters of a parametric multi-crop RP model can easily be obtained for each sampled farm. This property of parametric RP models, which is now routinely used in econometric analyses based on RP discrete choice models (Train 2009), can be exploited for calibrating sets of farm specific simulation models based on the farm sample and on the statistical background provided by estimated parametric RP multi-crop models. Second, micro-econometricians mostly rely on the simulated maximum likelihood (SML) estimation framework for estimating parametric RP models. However, computing SML estimators for RP multi-crop models involves solving intimidating simulated likelihood maximization problems. We also contribute to the applied econometrics literature by showing that the use of Monte Carlo expectation-maximization (MCEM) algorithms (see, *e.g.*, McLachlan and Krishnan 2008; Lavielle 2014) – which are routinely used by applied statisticians but rarely considered by applied econometricians (see, *e.g.*, Jank and Booth 2003) – can substantially alleviate the computational burden related to the estimation of large panel data multivariate RP.

The rest of the article is organized as follows. The general features of RP models are presented in the first section. The second section presents the theoretical interest of RP

[32] Whether the RP acreage share model is log-linearized.

acreage choice models and the related estimation issues. The third section describes the multi-crop econometric models considered for estimation: the RP multi-crop model and its versions with "homogeneity" restrictions that are used for comparison. Estimation and identification issues of the RP multi-crop model are briefly discussed in the fourth section. The estimation results and the investigation of the effects of EU biofuel support on French farmers' choices are provided in the fifth section.

5.1. Unobserved heterogeneity and random parameter models

This section presents the main features of RP models. It also introduces important elements to be used in the presentation of the estimation issues. The multi-crop model considered in the empirical section serves as an example. Its estimation makes use of a panel data set with observations are indexed by $i = 1,...,N$ for farms and by $t = 1,...,T$ for years.[33] We consider short run production choices of farmers – *i.e.* an acreage (share) choice system and a yield supply system – and we take for granted that farmers' choices rely on heterogeneous determinants.

A RP model is composed of two parts. The first part, the "behavioral model", formally describes the process of interest and defines its statistical characteristics conditional on the

[33] If panel data sets are not necessary for estimating RP models, identification of the probability distribution of the farm specific parameters is greatly facilitated with multiple observations of the sampled farms.

considered random parameters and on the exogenous variables. Our micro-econometric multi-crop model considers a sequence of acreage share vectors, $\mathbf{s}_i \equiv (\mathbf{s}_{it} : t = 1,...,T)$ where $\mathbf{s}_{it} \equiv (s_{k,it} : k \in \mathcal{K})$, the corresponding sequence of yield level vectors, $\mathbf{y}_i \equiv (\mathbf{y}_{it} : t = 1,...,T)$ where $\mathbf{y}_{it} \equiv (y_{k,it} : k \in \mathcal{K})$ and the corresponding sequence of input use level vectors, $\mathbf{x}_i \equiv (\mathbf{x}_{it} : t = 1,...,T)$ where $\mathbf{x}_{it} \equiv (x_{k,it} : k \in \mathcal{K})$. The term $\mathcal{K}$ denotes the considered crop set with $\mathcal{K} \equiv \{1,...,K\}$. These choices and outcomes are simultaneously modeled, with $\mathbf{w}_i \equiv (\mathbf{w}_{it} : t = 1,...,T)$ where $\mathbf{w}_{it} \equiv (\mathbf{y}_{it}, \mathbf{x}_{it}, \mathbf{s}_{it})$. The second part of the model defines the distribution characteristics of the random parameters conditionally on the exogenous variables.

5.1.1. Behavioral model and "kernel" likelihood function

Equation (1) describes the acreage choices and yield levels $\mathbf{c}_{it}$ as a known response function $\mathbf{r}$ to $(\mathbf{z}_{it}, \mathbf{e}_{it})$, the determinants of these production choices and outcomes.

(1) $$\mathbf{w}_{it} = \mathbf{r}(\mathbf{z}_{it}, \mathbf{e}_{it}; \boldsymbol{\delta}, \boldsymbol{\rho}_i)$$

The term $\mathbf{z}_{it}$ contains price and climatic variables. It is observed by the econometrician. The term $\mathbf{e}_{it}$ contains unobserved the unobserved determinants of $\mathbf{w}_{it}$. The response function $\mathbf{r}$ is parameterized by a fixed parameter vector $\boldsymbol{\delta}$ and a farmer specific parameter vector $\boldsymbol{\rho}_i$. Equation (1) describes how the choices of farmer *i* and their outcomes, $\mathbf{w}_{it}$, are determined by $(\mathbf{z}_{it}, \mathbf{e}_{it})$ up to the characteristics of this farmer and of his farm, $\boldsymbol{\rho}_i$. This equation can be

any agricultural production choice model where part of the usual fixed parameters are replaced by farmer specific parameter vector $\boldsymbol{\rho}_i$.

In a short run production choice context the random parameter vector $\boldsymbol{\rho}_i$ mainly captures the effects of the farms' natural or quasi-fixed factor endowments, of the production technologies used by the farmer and of farmers' characteristics. Our application considers short-run crop production choices and relies on a short panel data set, *i.e.* with $T = 6$. Since farms and farmers' production technology generally evolves slowly over time, we assume that the parameters $\boldsymbol{\rho}_i$ of the production choice model are constant during the considered time period.

Importantly, specification of the role of $\boldsymbol{\rho}_i$ in the model of $\mathbf{w}_{it}$ depends on how unobserved heterogeneity effects are expected to affect farmers' choices and outcomes. Standard panel data models generally assume that the effects of $\boldsymbol{\rho}_i$ and of $\mathbf{e}_{it}$ are additively separable in $\mathbf{r}$.[34] In this case the so-called "individual effect" $\boldsymbol{\rho}_i$ does not affect the effect of $\mathbf{z}_{it}$ on $\mathbf{w}_{it}$, implying relatively homogeneous responses of $\mathbf{w}_{it}$ to changes in $\mathbf{z}_{it}$. Keane (2009) highlights a basic trade-off related to this issue. Econometric models with additively separable random terms, *i.e.* error or parameters, are relatively easily estimated without specifying the parametric distribution of the error term. But standard "individual effect"

[34] *E.g.* they assume that $\mathbf{r}(\mathbf{z}_{it}, \mathbf{e}_{it}; \boldsymbol{\delta}, \boldsymbol{\rho}_i) = \mathbf{r}_0(\mathbf{z}_{it}; \boldsymbol{\delta}) + \boldsymbol{\rho}_i + \mathbf{e}_{it}$ for some function $\mathbf{r}_0$.

models are unsuitable when the effect of $\mathbf{z}_{it}$ on $\mathbf{w}_{it}$ actually depends on unobserved characteristics of farmer *i* or of its farm.[35] The RP framework allow for interactions between $\mathbf{z}_{it}$ and $\boldsymbol{\rho}_i$ in $\mathbf{r}(\mathbf{z}_{it}, \mathbf{e}_{it}; \boldsymbol{\delta}, \boldsymbol{\rho}_i)$. However the estimation of RP models appear to be difficult without parametric assumptions on the probability distribution of the random terms $(\mathbf{e}_{it}, \boldsymbol{\rho}_i)$. It remains uneasy even under such assumptions.

Keane (2009) argues that the use of fully parametric mixed models and of relatively involved inference tools is a reasonable price for buying the opportunity to introduce rich unobserved heterogeneity effects in micro-econometric models. Of course, this trade-off is an empirical issue and is likely to significantly depend on the modeled processes. Our viewpoint is that the empirical evidences accumulated in other applied economics fields suggest that it is worth investigating for agricultural production choice modeling. This is the main topic of this article with a specific focus on the effects of unobserved heterogeneity on farmers' responses to economic incentives.

Equation (1) is completed by statistical assumptions in order to define the "behavioral model". It is assumed here that $(\mathbf{z}_{it}, \boldsymbol{\rho}_i)$, and $\mathbf{e}_{it}$ are independent. This assumption is restrictive but it is standard, at least for reduced form models where $\mathbf{z}_{it}$ describes external factors – such as market prices or climatic events – affecting $\mathbf{w}_{it}$. The random parameter $\boldsymbol{\rho}_i$

[35] *I.e.* when $\frac{\partial}{\partial \mathbf{z}} \mathbf{r}(\mathbf{z}_{it}, \mathbf{e}_{it}; \boldsymbol{\delta}, \boldsymbol{\rho}_i)$ actually depends on $\mathbf{q}_i$.

is assumed to capture the persistent unobserved characteristics of farm/farmer *i* affecting the choice and outcome variable $\mathbf{w}_{it}$. The term $\mathbf{e}_{it}$ is assumed to represent the effects of idiosyncratic shocks on $\mathbf{w}_{it}$, *i.e.* $\mathbf{e}_{it}$ basically is a "standard" error term.

We also assume that the probability density function of $\mathbf{e}_{it}$ is known and parameterized by $\boldsymbol{\psi}$, a parameter vector to be estimated. The probability density function of $\mathbf{e}_{it}$ and equation (1) allows computing $f(\mathbf{w}_{it} \mid \mathbf{z}_{it}, \boldsymbol{\rho}_i; \boldsymbol{\theta}_w)$, the probability density function of $\mathbf{w}_{it}$ conditional on $(\mathbf{z}_{it}, \boldsymbol{\rho}_i)$ parameterized by $\boldsymbol{\theta}_w \equiv (\boldsymbol{\delta}, \boldsymbol{\psi})$. The term $f(\mathbf{w}_{it} \mid \mathbf{z}_{it}, \boldsymbol{\rho}_i; \boldsymbol{\theta}_w)$ defines the "kernel"[36] likelihood function at $\boldsymbol{\theta}_w$ of the parametric RP model of $\mathbf{w}_{it}$.

Equation (1) and the independence assumptions described above define a "behavioral model" which can be used with cross-section data. With panel data additional assumptions are required in order to describe the potential dynamic features of the considered choices. In the simplest case, $\mathbf{z}_{it}$ and $\mathbf{e}_{is}$ are independent conditionally on $\boldsymbol{\rho}_i$ for any pair of years (s,t) and the $\mathbf{e}_{it}$ terms are independent across *t*. These assumptions implies that equation (1) describes a static process repeated for $t = 1,...,T$.

These assumptions are assumed to hold in our empirical application which deals with cash crop short run production choices and outcomes. The dynamic features of cash crop

[36] According to Train's (2007) terminology

production are mostly due to crop rotations. Such dynamic effects can be suitably approximated by farm specific parameters such as $\boldsymbol{\rho}_i$ when farmers base their production choices on stable rotation schemes. Short run production choices are repeated each year and follow the same process as long as the production technology, crop rotation schemes included, and the quasi-fixed factor endowment do not change. Under the assumptions given above the joint density of the vector $\mathbf{w}_i$ conditional on $(\boldsymbol{\rho}_i, \mathbf{z}_i)$, where $\mathbf{z}_i \equiv (\mathbf{z}_{i1}, ..., \mathbf{z}_{iT})$, is given by:

(2) $$f(\mathbf{w}_i \mid \mathbf{z}_i, \boldsymbol{\rho}_i; \boldsymbol{\theta}_w) = \prod_{t=1}^{T} f(\mathbf{w}_{it} \mid \mathbf{z}_{it}, \boldsymbol{\rho}_i; \boldsymbol{\theta}_w)$$

Of course, farmers' choices and outcomes are statistically linked across time due to their relying on the same parameter vector $\boldsymbol{\rho}_i$. But these choices and outcomes are assumed to be independent across time conditionally on $\boldsymbol{\rho}_i$.

5.1.2. Mixing probability distribution

The second part of a parametric RP model describes the probability distribution of the farmers' specific parameters $\boldsymbol{\rho}_i$ conditionally on the observed variables $\mathbf{z}_i$. It is assumed here that $\mathbf{z}_i$ and $\boldsymbol{\rho}_i$ are independent. The exogenous variable vector $\mathbf{z}_{it}$ contains exogenous determinants of the production choices and outcomes, *i.e.* prices and climatic conditions, which mostly vary across time. Let $g(\boldsymbol{\rho}_i; \boldsymbol{\theta}_{\boldsymbol{\rho}})$ the probability density function of $\boldsymbol{\rho}_i$. This function is defined up to the parameter vector $\boldsymbol{\theta}_{\boldsymbol{\rho}}$ to be estimated. The probability

distribution of $\boldsymbol{\rho}_i$ describes the distribution of $\boldsymbol{\rho}_i$ across the considered famers' population. Statistical estimates of $\boldsymbol{\theta}_{\boldsymbol{\rho}}$ allow the investigation of the distribution of the random parameter $\boldsymbol{\rho}_i$. These estimates can be used to test the empirical relevance of the RP specification by checking whether $\boldsymbol{\rho}_i$ actually exhibits statistically significant variability or not. Estimates of $\boldsymbol{\theta}_{\boldsymbol{\rho}}$ can also be used to interpret the empirical content of the $\boldsymbol{\rho}_i$ terms. *E.g.*, investigation of the statistical relations among the elements of $\boldsymbol{\rho}_i$ may point out some sources of unobserved heterogeneity.

Of course the choice of the "mixing" probability distribution function, *i.e.* the parametric family of the probability distribution of $\boldsymbol{\rho}_i$, is crucial to suitably capture the unobserved heterogeneity effects in the considered model. Being related to unobserved variables, this choice basically is an empirical issue. It is usually based on trials with different parametric models.[37]

5.1.3. Likelihood functions and "statistical calibration" of individual parameters

[37] Using flexible parametric models, *e.g.* finite discrete mixtures of Gaussian models, or non parametric models appears to be difficult in practice. Such models can only be used when the dimension of $\mathbf{q}_i$ is very small and with very large samples.

The probability density function of the dependent variable $\mathbf{w}_i$ conditional on its observed determinants $\mathbf{z}_i$ defines the likelihood function on which the ML estimation framework is based. For RP models, this probability density function is defined as the mean of $f(\mathbf{w}_i \mid \mathbf{z}_i, \boldsymbol{\rho}_i; \boldsymbol{\theta}_w)$, the assumed "kernel" probability distribution function, over the distribution of the random parameter $\boldsymbol{\rho}_i$, the assumed "mixing" distribution:

$$f(\mathbf{w}_i \mid \mathbf{z}_i; \boldsymbol{\theta}) = \int f(\mathbf{w}_i \mid \mathbf{z}_i, \boldsymbol{\rho}_i; \boldsymbol{\theta}_w) g(\boldsymbol{\rho}; \boldsymbol{\theta}_\rho) d\boldsymbol{\rho} \tag{3}$$

The term $\boldsymbol{\theta} \equiv (\boldsymbol{\theta}_w, \boldsymbol{\theta}_\rho)$ is the parameter vector of the considered parametric RP model. The probability density functions $f(\mathbf{w}_i \mid \mathbf{z}_i, \boldsymbol{\rho}_i; \boldsymbol{\theta}_w)$ and $g(\boldsymbol{\rho}; \boldsymbol{\theta}_\rho)$ are defined by the model and generally have simple analytical forms. But the integral in equation (3) can rarely be solved neither analytically nor numerically when $g(\boldsymbol{\rho}_i; \boldsymbol{\theta}_\rho)$ is continuous, as it is assumed here. Simulation methods need to be used for approximating the RP model individual likelihood functions $f(\mathbf{w}_i \mid \mathbf{z}_i; \boldsymbol{\theta})$ and this explains why the estimation of mixed model is often difficult.

An estimated RP model can also be used for constructing a simulation models defined as a sample of heterogeneous farm models, by using the response model (1) and by computing an estimate $\boldsymbol{\rho}_i$ for each farm of the sample used for estimating the considered RP model. These estimates of the $\boldsymbol{\rho}_i$ terms can be based on a simple "statistical calibration" procedures based on a well defined and coherent statistical background: the RP model of $\mathbf{w}_i$ and the statistical estimate of $\boldsymbol{\theta}$ obtained from this model and the data. Such procedures require either to

generate random draws from the probability distribution of $\boldsymbol{\rho}_i$ conditional on $(\mathbf{w}_i, \mathbf{z}_i)$ or to estimate its density function.

The marginal probability distribution of $\boldsymbol{\rho}_i$, characterized by the density function $g(\boldsymbol{\rho}_i; \boldsymbol{\theta}_\rho)$, is the *ex ante* or *prior* distribution of the random parameter. It describes the distribution of $\boldsymbol{\rho}_i$ in the considered farmer population. The probability distribution of $\boldsymbol{\rho}_i$ conditional on $(\mathbf{w}_i, \mathbf{z}_i)$ is designated as an *ex post* or *a posteriori* density probability distribution. Its probability density function, denoted here as $g(\boldsymbol{\rho}_i \mid \mathbf{z}_i, \mathbf{w}_i; \boldsymbol{\theta})$, sums up what is known about $\boldsymbol{\rho}_i$ for farmer *i*, according to a simple "Tell me what you do, I'll know you who you are" logic. Variables $(\mathbf{w}_i, \mathbf{z}_i)$ is the information directly brought by the data on farmer *i* and the functional form of $g(\boldsymbol{\rho}_i \mid \mathbf{z}_i, \mathbf{w}_i; \boldsymbol{\theta})$ sums up the information brought by the considered RP model. To estimate $g(\boldsymbol{\rho}_i \mid \mathbf{z}_i, \mathbf{w}_i; \boldsymbol{\theta})$ just requires to consistently estimate $\boldsymbol{\theta}$ and standard simulation methods. *E.g.*, application of Bayes' rule yields:

(4) $g(\boldsymbol{\rho}_i \mid \mathbf{z}_i, \mathbf{w}_i; \boldsymbol{\theta}) = \omega(\mathbf{w}_i, \mathbf{z}_i, \boldsymbol{\rho}_i; \boldsymbol{\theta}) g(\boldsymbol{\rho}_i; \boldsymbol{\theta}_\rho)$, where $\omega(\mathbf{w}_i, \mathbf{z}_i, \boldsymbol{\rho}_i; \boldsymbol{\theta}) \equiv \dfrac{f(\mathbf{w}_i \mid \mathbf{z}_i, \boldsymbol{\rho}_i; \boldsymbol{\theta}_w)}{f(\mathbf{w}_i \mid \mathbf{z}_i; \boldsymbol{\theta})}$.

As discussed above the $f(\mathbf{w}_i \mid \mathbf{z}_i; \boldsymbol{\theta})$ term can be estimated, with simulation methods if needed, when a consistent estimate of $\boldsymbol{\theta}$ is available.

The term $E[\boldsymbol{\rho}_i \mid \mathbf{c}_i, \mathbf{w}_i]$ is the best predictor of $\boldsymbol{\rho}_i$ conditional on $(\mathbf{w}_i, \mathbf{z}_i)$ according to the minimum squared prediction error criterion.[38] Provided that

(5) $$E[\boldsymbol{\rho}_i \mid \mathbf{c}_i, \mathbf{w}_i] \equiv \int \boldsymbol{\rho} g(\boldsymbol{\rho} \mid \mathbf{z}_i, \mathbf{w}_i; \boldsymbol{\theta}) d\boldsymbol{\rho}$$

$E[\boldsymbol{\rho}_i \mid \mathbf{c}_i, \mathbf{w}_i]$ can be estimated when a consistent estimate of $\boldsymbol{\theta}$ is available. *E.g.*, draws from $g(\boldsymbol{\rho}_i \mid \mathbf{z}_i, \mathbf{w}_i; \boldsymbol{\theta})$ can be obtained from a Metropolis-Hastings simulator (see, *e.g.*, Train 2009). But this approach is time consuming in our case. Equation (4) suggests estimating $E[\boldsymbol{\rho}_i \mid \mathbf{c}_i, \mathbf{w}_i]$ by using an importance sampling approach with $g(\boldsymbol{\rho}_i; \boldsymbol{\theta}_\rho)$ as the proposal density function. *E.g.*, the integral in equation (5) and the terms $\omega(\mathbf{w}_i, \mathbf{z}_i, \boldsymbol{\rho}_i; \boldsymbol{\theta})$ involved in equations (4) can be computed with simulation methods if needed.

An estimate of $E[\boldsymbol{\rho}_i \mid \mathbf{c}_i, \mathbf{w}_i]$ provides a prediction of $\boldsymbol{\rho}_i$ with a consistent statistical background. One can thus build a simulation model with heterogeneous farms and farmers from the considered sample and the estimated model. It suffices to "statistically calibrate" $\boldsymbol{\rho}_i$ for each sampled farm as shown above.[39]

5.2. RP acreage choice models: theoretical interest and empirical issues

---

[38] The mode of $g(\boldsymbol{\rho}_i \mid \mathbf{z}_i, \mathbf{w}_i; \boldsymbol{\theta})$ may also be used for estimating $\boldsymbol{\rho}_i$, according to the maximum likelihood criterion.

[39] We keep on using the term "calibration" here because the considered estimate of $\boldsymbol{\rho}_i$ rely on a limited amount information specific to farm *i*, *i.e.* it only relies on $(\mathbf{z}_i, \mathbf{w}_i)$

The aim of this section is to present the appealing features of RP acreage choice models and the empirical issues raised by the estimation of such models. We consider the standard MNL (SMNL) acreage share model of Carpentier and Letort (2014) because it is a simplified version of the acreage choice model that is part of the RP multi-crop model to be considered in the next sections.[40]

The SMNL acreage share model is derived as the solution to an expected profit maximization problem defined at the farm level. Let us consider a risk neutral farmer i who can allocate his fixed cropland area to K crops. Let $\mathcal{K} \equiv \{1,...,K\}$ denote the set of crops available to this farmer. This farmer is assumed to maximize his expected profit level while seeking to minimize the management costs implied by his acreage choice. In year *t*, her/his expected profit maximization problem is formally defined as:

(6) $$\max_{\mathbf{s}\geq\mathbf{0}} \left\{ \sum_{k\in\mathcal{K}} s_k \pi_{k,it} - C_{it}(\mathbf{s}) \text{ s.t. } \sum_{k\in\mathcal{K}} s_k = 1 \right\}.$$

The term $\pi_{k,it}$ denotes the return of crop *k* as it is expected by farmer *i* in year *t*. The term s_k denotes the acreage share of crop *k,* and the vector **s** denotes the related crop acreage share

[40] Interestingly, the functional form of the SMNL acreage share model is closely related to that of the MNL probability function used in discrete choice models for analyzing crop choices observed at the plot level (see, *e.g.*, Wu et al 2004; Langpap and Wu 2011; Fezzi et al 2014; Classen et al 2017). Also, the discussion related to the introduction of random parameters in the SMNL acreage share model also holds for these probabilistic discrete crop choice models.

vector with $\mathbf{s} \equiv (s_k : k \in \mathcal{K})$. The term $C_{it}(\mathbf{s})$ is defined as the implicit management cost of acreage **s**. This cost function accounts for the variable costs not included in the expected crop returns $\pi_{k,it}$. It also accounts for the implicit costs related to the constraints on the acreage choices due to limiting quantities of quasi-fixed factors or to agronomic factors. These constraints providing motives for diversifying crop acreages, the function $C_{it}(\mathbf{s})$ is assumed to be convex in **s**.[41] Carpentier and Letort (2014) show that the solution in **s** to problem (1) is the SMNL acreage share model given by

$$s_{k,it} = \frac{\exp(\alpha_0^s \pi_{k,it} - b_{k,it}^s)}{\sum_{\ell \in \mathcal{K}} \exp(\alpha_0^s \pi_{\ell,it} - b_{\ell,it}^s)} \text{ for } k \in \mathcal{K} \tag{7}$$

if the cost function is defined as $C_{it}(\mathbf{s}) = (\alpha_0^s)^{-1}\left(\sum_{k \in \mathcal{K}} s_k b_{k,it}^s + \sum_{k \in \mathcal{K}} s_k \ln s_k\right)$. The curvature parameter α_0^s needs to be strictly positive for this entropic cost function to be strictly convex in **s**. This parameter plays a crucial role in the SMNL acreage choice model, as it determines the flexibility of the acreage choices with respect to changes in the expected crop returns; the larger this parameter, the more the acreage choices respond to changes in the expected crop returns. The linear parameters of the entropic cost function $C_{it}(\mathbf{s})$, the $b_{k,it}^s$ terms, play the role of shifters of the optimal acreage share choices.

Using crop K as the reference crop and enforcing the normalization constraint $b_{K,it}^s = 0$,

---

[41] Analogous functions are used in the positive mathematical programming literature (Heckeleï and Wolff 2012) and in the multi-crop econometric literature (Carpentier and Letort 2012).

the acreage share model (7) can be log-linearized as $\ln s_{k,it} - \ln s_{K,it} = \alpha_0^s(\pi_{k,it} - \pi_{K,it}) - b_{k,it}^s$. The acreage choice shifters $b_{k,it}^s$ can then be defined as linear functions of control variables with additively separable error terms for obtaining empirically tractable equations. Let us define these functions as $b_{k,it}^s = -b_{k,0}^s - (\mathbf{d}_{k,0}^s)'\mathbf{c}_{k,it}^s - u_{k,it}^s$, where $\mathbf{c}_{k,it}^s$ is a control variable vector, and $u_{k,it}^s$ is a standard error term. Assuming that measures of the expected crop returns $\pi_{k,it}$ are available, the equation system

(8) $$\ln s_{k,it} - \ln s_{K,it} = b_{k,0}^s + \alpha_0^s(\pi_{k,it} - \pi_{K,it}) + (\mathbf{d}_{k,0}^s)'\mathbf{c}_{k,it}^s + u_{k,it}^s \text{ for } k \in \mathcal{K}$$

defines a system of linear regression equations under suitable exogeneity assumptions.

Equation systems such as the one described in equation (8) are usually extended to account for farms' and farmers' unobserved heterogeneity by specifying additively separable random farm specific effects. It suffices to replace the fixed intercept parameters $b_{k,0}^s$ by random farm specific parameters $b_{k,i}^s$ for obtaining the following system of standard panel data linear models:

(9) $$\ln s_{k,it} - \ln s_{K,it} = b_{k,i}^s + \alpha_0^s(\pi_{k,it} - \pi_{K,it}) + (\mathbf{d}_{k,0}^s)'\mathbf{c}_{k,it}^s + u_{k,it}^s \text{ for } k \in \mathcal{K}.$$

The resulting model features slope parameters that are all fixed. Standard panel data estimators allow for consistently estimating the interest parameters of this model, whether the expected crop returns $\pi_{k,it}$ are correlated with the farm specific random effects $b_{k,i}^s$ or not (Wooldridge 2010). The ability to account for such correlations is important: these may be significant, and neglecting them can bias the obtained estimates. Indeed, farm specific parameters such as the $b_{k,i}^s$ terms aim to capture the effects of unobserved factors such as

natural and material capital endowments, skills or production practices. Since such factors can also impact the expected crop returns $\pi_{k,it}$, the terms $b_{k,i}^{s}$ and $\pi_{\ell,it} - \pi_{K,it}$ may be correlated in equation (9).

Panel data acreage choice models found in the literature tend to show that additively separable farm specific effects, such as the farm specific shifters $b_{k,i}^{s}$, significantly vary across farms (see, *e.g.*, Lacroix and Thomas 2011; Platoni et al 2012; Bayramoglu and Chakir, 2016). This implies that farm specific unobserved factors significantly impact *farmers' acreage choice levels*. The main purpose of this article is to investigate a complementary question: do we also need to account for unobserved heterogeneity effects in *how farmers' acreage choices respond to economic incentives*? Extending the model specification accordingly is simple from a theoretical viewpoint. It suffices to replace the fixed acreage choice flexibility parameter α_0^s by random farm specific slope coefficients α_i^s for obtaining the following equation system:

(10) $$\ln s_{k,it} - \ln s_{K,it} = b_{k,i}^{s} + \alpha_i^s(\pi_{k,it} - \pi_{K,it}) + (\mathbf{d}_{k,0}^{s})'\mathbf{c}_{k,it}^{s} + u_{k,it}^{s} \quad \text{for} \quad k \in \mathcal{K}.$$

However, such panel data random coefficient linear models raise significant estimation issues. Even if random coefficient models have a long history in the econometrics and statistics literatures, most available methods for estimating the joint probability distribution of the random intercept and slope parameters of model (10) consider the so-called "classical random coefficient linear model" in the single equation case (see, *e.g.*, Swamy 1970; Beran et al 1996; Hsiao and Pesaran 2008). Two assumptions characterizing the classical random coefficient model are difficult to relax: its explanatory variables are assumed to be neither

endogenous with respect to its error term nor correlated with its random slope coefficients.

As will be shown below, the first assumption of the classical random coefficient model does not hold in the RP (nested MNL) acreage share model considered in our empirical application. In its log-linearized version, this model contains explanatory variables that are endogenous by construction. The second assumption may not hold, either in the acreage share model considered in the empirical application or in the simplified version of this model described in equation (10). As argued above, common unobserved farm specific factors may impact the expected crop returns $\pi_{\ell,it}$, and the farm specific shifters $b_{k,i}^{s}$, implying that these terms can be correlated. The same factors may also induce correlations between these returns and the farm specific acreage choice flexibility parameters α_i^s. Such correlations would imply that model (10) would be a system of panel data "correlated random coefficient models."

These correlation and endogeneity issues have recently received considerable attention in the econometric literature (see, *e.g.*, Wooldridge 2005; Keane 2009; Angrist and Pischke 2009; Murtazashvili and Wooldridge 2008, 2016; Arellano and Bonhomme 2011, 2012). However, most available results do not fulfill our needs. Acreage choice models are relatively large systems of interrelated equations, and we are interested in the joint probability distribution of the model random parameters. Most available results focus on the single equation case, on specific models and/or on the estimation of specific characteristics of the

random parameter distribution.[42]

As shown by Keane (2009), the main estimation issues discussed here about the estimation of the RP acreage share model (10) can be overcome by adopting a parametric full information approach. This approach consists (*a*) of explicitly modeling the $\pi_{\ell,it}$ terms and their links with the random parameters of the acreage share model and (*b*) of imposing distributional assumptions on the error terms and on the random parameters of the considered model. Our RP multi-crop model explicitly defines how the expected crop gross margins $\pi_{\ell,it}$ are constructed from the data and potentially linked with the random parameters of the acreage share model. Parametric assumptions on the random parts of the model then make explicit the structure of the endogeneity and correlation problems arising in our RP multi-crop model. Finally, ML estimators "automatically" take on the abovementioned estimation issues, as the likelihood function of any parametric version of the RP multi-crop model structurally accounts for the considered endogeneity and correlation problems.

Econometricians are often reluctant to rely on parametric models because distributional assumptions are difficult to test and can significantly impact estimation results. Moreover,

---

[42] The results surveyed by Angrist and Pischke (2009) mostly consider models with dichotomous binary explanatory variables. Murtazashvili and Wooldridge (2008, 2016) and Wooldridge (2005) focus on the estimation of the means of the random slope coefficients. Arellano and Bonhomme (2012) propose non-parametric estimators of the joint distribution of the random coefficients of linear models. However, their approach requires a relatively large time dimension and cannot handle any kind of endogeneity problem. In particular, their assumptions rule out the endogeneity problem arising in the RP acreage share model considered in our empirical application.

even if our empirical results appear to be robust to various distributions for the farm specific parameters, the scope of such robustness checks is necessarily limited in practice. However, following Keane (2009), we regard these distributional assumptions as a reasonable price to pay for the opportunity to account for rich unobserved heterogeneity effects in multi-crop models.[43]

5.3. The RP multi-crop model

This section presents the RP micro-econometric multi-crop model considered in the empirical application presented in the next section. This model is composed of three equation sub-systems describing the yield supply functions, variable input demand functions and acreage share choice models of each produced crop.

5.3.1. Acreage share choice models

The acreage share model included in the RP multi-crop model extends the SMNL acreage

[43] Distributional assumptions are widely accepted for many popular econometric models displaying specific features precisely because such assumptions allow for dealing with these features in applied work. For instance, most RP discrete choice models (Train 2009) or stochastic frontier models (Greene 2005) used by applied econometricians are parametric. Also, the multi-crop model of Lacroix and Thomas (2011), which is specified as a censored regression system for dealing with corner solutions, is also parametric.

share model (7) in two ways. First, it is defined as a RP model. Second, it is defined as a nested MNL (NMNL) acreage share model. NMNL acreage share models were proposed by Carpentier and Letort (2014) for overcoming a major limitation of SMNL models: their symmetrical consideration of the different crops of the considered crop set. In applications considering more than two arable crops, the considered crops can generally be grouped according to their competing use of quasi-fixed factors and/or according to their agronomic characteristics. In our setting, the crop set $\mathcal{K}$ is partitioned into G mutually exclusive crop groups, $\mathcal{K}(g)$ crop sets, such that $g \in \mathcal{G} \equiv \{1,...,G\}$. Crops belonging to a given group share similar agronomic characteristics and/or compete more for farmers' limited quantities of quasi-fixed factors than they compete with crops of other groups. Carpentier and Letort (2014) proposed the following (nested entropic) functional form of acreage management cost function:

$$(11) \qquad C_{it}(\mathbf{s}) = \sum_{k\in\mathcal{K}} s_k \beta^s_{k,it} + (\alpha^s_i)^{-1} \sum_{g=1}^{G} s_{(g)} \left(\ln s_{(g)} + \alpha^s_i (\alpha^s_{(g),i})^{-1} \sum_{\ell\in\mathcal{K}(g)} s_{\ell|(g)} \ln s_{\ell|(g)} \right)$$

to account for these features of the crop set. The term $s_{(g)}$ denotes the acreage share of the crops belonging to group g, *i.e.,* $s_{(g)} \equiv \sum_{\ell\in\mathcal{K}(g)} s_\ell$, while the term $s_{\ell|(g)}$ denotes the acreage share of crop ℓ in the acreage share of group g, *i.e.,* $s_{\ell|(g)} \equiv s_\ell / s_{(g)}$. The α^s_i and $\alpha^s_{(g),i}$ random farm specific parameters determine the curvature properties of $C_{it}(\mathbf{s})$.[44] The condition $\alpha_{(g),i} \geq \alpha_i > 0$ for $g \in \mathcal{G}$ is sufficient for the cost function $C_{it}(\mathbf{s})$ to be strictly convex in $\mathbf{s}$.

[44] The equality $\alpha_{(g),i} = \alpha_i$ is imposed if group g contains a single crop.

According to this cost function, the acreage management costs primarily depend on how the total land area is allocated to crop groups and then on how the land areas allocated to crop groups are allocated to crops within groups. The NMNL acreage choice model reflects this pattern. Carpentier and Letort (2014) show that the solution in **s** to the expected profit maximization problem (6) with the functional form of the cost function $C_{it}(\mathbf{s})$ described in equation (11) is given by:

$$(12) \qquad s_{k,it} = \frac{\exp\left(\alpha_{(g),i}(\pi_{k,it} - \beta^s_{k,it})\right)}{\sum_{\ell \in \mathcal{K}(g)} \exp\left(\alpha_{(g),i}(\pi_{\ell,it} - \beta^s_{\ell,it})\right)} \frac{\left(\sum_{\ell \in \mathcal{K}(g)} \exp\left(\alpha_{(g),i}(\pi_{\ell,it} - \beta^s_{\ell,it})\right)\right)^{\alpha_i \alpha^{-1}_{(g),i} - 1}}{\sum_{h \in \mathcal{G}} \left(\sum_{\ell \in \mathcal{K}(h)} \exp\left(\alpha_{(h),i}(\pi_{\ell,it} - \beta^s_{\ell,it})\right)\right)^{\alpha_i \alpha^{-1}_{(h),i}}}$$

provided that crop *k* belongs to group *g*. The first right-hand-side term of equation (12) yields the acreage share of crop *k* in the acreage share of its group, $s_{k|(g),it}$, while the second right-hand-side term yields the acreage share of group *g*, $s_{(g),it}$.

The α^s_i and $\alpha^s_{(g),i}$ farm specific parameters determine the flexibility of farmers' acreage choices. Heterogeneity in these "acreage choice flexibility parameters" may reflect differences in physical capital endowments, in labor endowments and/or in acreage choice strategies across farmers. For example, the acreage choices of farmers with limiting labor and machinery endowments may be less responsive to short term economic incentives because their acreage choices are constrained by the need to avoid unmanageable machinery and labor peak loads. Farmers may also prefer to control weeds, pests and diseases by diversifying their crop acreages rather than by using pesticides due to pro-environmental attitudes. The acreage choices of such farmers may be less responsive to crop prices. This also holds in the case of farmers committed to delivery contracts.

This NMNL acreage share model can be log-linearized as:

$$\ln s_{k,it} - \ln s_{K,it} = \alpha_i(\pi_{k,it} - \pi_{K,it} - \beta^s_{k,it}) + \left(1 - \alpha^s_i(\alpha^s_{(g),i})^{-1}\right)\ln s_{k|(g),it} - \left(1 - \alpha^s_i(\alpha^s_{(G),i})^{-1}\right)\ln s_{K|(G),it} \tag{13}$$

by using Berry's (1994) device and assuming that the benchmark crop K belongs to group G (with $\beta^s_{K,it} \equiv 0$ as a normalization constraint). In the empirical specification of the RP multi-crop model, the linear terms $\beta^s_{k,it}$ of the cost function $C_{it}(\mathbf{s})$ are decomposed as $\beta^s_{k,it} = \beta^s_{k,i} + (\boldsymbol{\delta}^s_{k,0})'\mathbf{c}^s_{k,it} + \varepsilon^s_{k,it}$ with $E[\varepsilon^s_{k,it}] = 0$. The $\mathbf{c}^s_{k,it}$ vectors contain variables used to control for observed heterogeneous factors. As the other "**c**" control variable vectors to be introduced later, these vectors contain farm characteristics, such as total land area and capital, and climatic conditions, such as rainfall and temperature. The $\beta^s_{k,i}$ farm-specific factors account for unobserved heterogeneity effects. The error terms $\varepsilon^s_{k,it}$ capture the effects of events, such as climatic events, impacting the soil state at planting. These terms are assumed to be known to farmers when they choose their acreages.

5.3.2. Yield supply and variable input demand models

The RP multi-crop model also includes crop yield supply and variable input demand equations. As in Carpentier and Letort (2014), we assume that farmers produce crop k from a variable input aggregate under a quadratic technological constraint. However, we generalize their modeling framework by assuming that the technological constraint is farm specific. In other words, we assume that the yield of crop k obtained by farmer i in year t is

given by the quadratic yield function $y_{k,it} = \beta^{y}_{k,it} - (1/2)(\alpha^{x}_{k,i})^{-1}(\beta^{x}_{k,it} - x_{k,it})^{2}$, where $x_{k,it}$ denotes the variable input use level. The term $\beta^{y}_{k,it}$ is the yield level that can be potentially achieved by farmer *i* in year *t*, while $\beta^{x}_{k,it}$ is the input quantity required to achieve this potential yield level. Assuming that farmers maximize the expected return to variable input uses and observe $\beta^{x}_{k,it}$ when determining their variable inputs, the crop variable input demand functions are given by:

(14) $$x_{k,it} = \beta^{x}_{k,it} - \alpha^{x}_{k,i} v_{k,it} p^{-1}_{k,it}$$

while the corresponding crop yield supply functions are given by:

(15) $$y_{k,it} = \beta^{y}_{k,it} - (1/2)\alpha^{x}_{k,i} v^{2}_{k,it} p^{-2}_{k,it}$$

The term $p_{k,it}$ denotes the output price of crop *k* expected by farmer *i* in year *t*. The term $v_{k,it}$ denotes the crop *k* variable input price (index) paid by this farmer. These functions show that the curvature parameter $\alpha^{x}_{k,i}$ can be interpreted as the flexibility parameter of crop *k* input use.[45] It determines the extent to which the yield supply and the input demand of crop *k* respond to input and crop prices. Heterogeneity in the input use flexibility parameters may reflect differences in the cropping practices used by farmers and/or in farmers' attitudes. For example, assume that a farmer targets high yield levels at medium crop prices. This farmer cannot increase his fertilizer uses at high crop prices very much because he already uses

[45] The parameter $\alpha^{x}_{k,i}$ is required to be positive for the considered yield function to be concave in $x_{k,it}$.

almost maximum levels of fertilizers at medium prices. In contrast, the fertilizer uses of farmers targeting medium yield levels at medium crop prices can be more responsive to crop price increases. Additionally, farmers used to high crop protection levels might be reluctant to decrease their pesticide uses because they do not know the impact of such changes on their crop yields.

In the empirical specification of the RP multi-crop model, the potential yield level and input requirement terms, $\beta^y_{k,it}$ and $\beta^x_{k,it}$, are decomposed as $\beta^y_{k,it} \equiv \beta^y_{k,i} + (\boldsymbol{\delta}^y_{k,0})'\mathbf{c}^y_{k,it} + \varepsilon^y_{k,it}$ with $E[\varepsilon^y_{k,it}] = 0$ and $\beta^x_{k,it} \equiv \beta^x_{k,i} + (\boldsymbol{\delta}^x_{k,0})'\mathbf{c}^x_{k,it} + \varepsilon^x_{k,it}$ with $E[\varepsilon^x_{k,it}] = 0$. The terms $\mathbf{c}^y_{k,it}$ and $\mathbf{c}^x_{k,it}$ are observed variable vectors used to control for farm heterogeneity and for specific climatic conditions. The term $\beta^y_{k,i}$ is the farm specific potential yield parameter of crop *k*, while the $\beta^x_{k,i}$ term is the farm specific input requirement parameter of crop *k.* These random parameters and the "input use flexibility parameters" $\alpha^x_{k,i}$ are expected to significantly vary across farms because of unobserved heterogeneity in farms' capital endowments and farmers' skills. The $\varepsilon^y_{k,it}$ and $\varepsilon^x_{k,it}$ terms are standard error terms aimed at capturing the effects on production of stochastic events such as climatic conditions or pest and weed problems.

5.3.3. Expected crop returns

The expected crop returns $\pi_{k,it}$ used in the acreage share models given in equation (13) are

constructed by considering the input demand and yield supply models described above. They are conditioned on what farmers know when choosing their acreages. Assuming that farmers' expectations of the error terms $\varepsilon_{k,it}^{y}$ and $\varepsilon_{k,it}^{x}$ are null at the time of the acreage choice, the expected crop returns $\pi_{k,it}$ can be defined as:

$$\pi_{k,it} = p_{k,it}\left(\beta_{k,i}^{y} + (\boldsymbol{\delta}_{k,0}^{y})'\mathbf{c}_{k,it}^{y,s}\right) - v_{k,it}\left(\beta_{k,i}^{x} + (\boldsymbol{\delta}_{k,0}^{x})'\mathbf{c}_{k,it}^{x,s}\right) + (1/2)\alpha_{k,i}^{x} v_{k,it}^{2} p_{k,it}^{-1} \tag{16}$$

The control variable vectors $(\mathbf{c}_{k,it}^{y,s}, \mathbf{c}_{k,it}^{x,s})$ are obtained by replacing the climatic variables contained in $(\mathbf{c}_{k,it}^{y}, \mathbf{c}_{k,it}^{x})$ by their expectations at the time of the acreage choice (we used their means over the five years preceding year *t* in our empirical application).

The RP multi-crop model is obtained by collecting the yield supply and variable input demand models and the NMNL acreage choice models (in their log-linearized versions) described above:

$$\begin{cases} y_{k,it} = \beta_{k,i}^{y} + (\boldsymbol{\delta}_{k,0}^{y})'\mathbf{c}_{k,it}^{y} - (1/2)\alpha_{k,i}^{x} v_{k,it}^{2} p_{k,it}^{-2} + \varepsilon_{k,it}^{y}, \quad k \in \mathcal{K} \\ x_{k,it} = \beta_{k,i}^{x} + (\boldsymbol{\delta}_{k,0}^{x})'\mathbf{c}_{k,it}^{x} - \alpha_{k,i}^{x} v_{k,it} p_{k,it}^{-1} + \varepsilon_{k,it}^{x}, \quad k \in \mathcal{K} \\ \ln s_{k,it} - \ln s_{K,it} = \alpha_{i}^{s}(\pi_{k,it} - \pi_{K,it}) - \alpha_{i}^{s}\left(\beta_{k,i}^{s} + (\boldsymbol{\delta}_{k,0}^{s})'\mathbf{c}_{k,it}^{s} + \varepsilon_{k,it}^{s}\right) \\ \qquad\qquad + \left(1 - \alpha_{i}^{s}(\alpha_{(g),i}^{s})^{-1}\right)\ln s_{k|(g),it} - \left(1 - \alpha_{i}^{s}(\alpha_{(G),i}^{s})^{-1}\right)\ln s_{K|(G),it}, \quad k \in \mathcal{K} \end{cases} \tag{17}$$

where the expected crop returns $\pi_{k,it}$ are given in equation (16).[46]

---

[46] Farmers' expectations of the error terms of the yield supply and input demand models, *i.e.* $\varepsilon_{k,it}^{y}$ and $\varepsilon_{k,it}^{x}$, are null at the time of the acreage choices if these terms are uncorrelated with the error terms of acreage choice models, *i.e.* with the vector $\boldsymbol{\varepsilon}_{it}^{s} \equiv (\varepsilon_{\ell,it}^{s} : \ell = 1,...,K-1)$. However, the effects of stochastic events occurring

Any equation of the RP multi-crop model contributes to the identification of its fixed parameter vector and of its random parameter joint probability distribution. However, the structure of this equation system implies specific identification patterns. In particular, the expected crop return variables $\pi_{k,it}$ are constructed with elements obtained from the yield supply and variable input demand models and then "supplied" to the acreage share models. According to this recursive structure of the RP multi-crop model, the joint probability distribution of the potential yield shifters $\beta_{k,i}^{y}$, the input requirement shifters $\beta_{k,i}^{x}$ and the input use flexibility parameters $\alpha_{k,i}^{x}$ is expected to be mostly identified by the yield supply and variable input demand equation sub-systems. The joint probability distribution of the acreage share shifters $\beta_{k,i}^{s}$ and the acreage choice flexibility parameters α_{i}^{s} and $\alpha_{g,i}^{s}$ is expected to be mostly identified by the acreage share equation sub-system.

Importantly, the acreage share equations are not classical random coefficient models. First, the expected crop *k* expected gross margin $\pi_{k,it}$ directly depends on the crop *k* random

---

before the acreage choice and captured by $\boldsymbol{\varepsilon}_{it}^{s}$, *e.g.*, soil infections or weed infestations impacting the soil state at planting, can persist during the cropping season and impact crop yields and input uses. This might induce correlations between elements of $\boldsymbol{\varepsilon}_{it}^{s}$ and $\varepsilon_{k,it}^{y}$ or $\varepsilon_{k,it}^{x}$. In the empirical application, the term $p_{k,it}(\boldsymbol{\delta}_{k,0}^{y,\varepsilon})'\boldsymbol{\varepsilon}_{k,it}^{s} - v_{k,it}(\boldsymbol{\delta}_{k,0}^{x,\varepsilon})'\boldsymbol{\varepsilon}_{k,it}^{s}$ is added in the definition of expected crop return $\pi_{k,it}$ given in equation (16). The terms $(\boldsymbol{\delta}_{k,0}^{y,\varepsilon})'\boldsymbol{\varepsilon}_{k,it}^{s}$ and $(\boldsymbol{\delta}_{k,0}^{x,\varepsilon})'\boldsymbol{\varepsilon}_{k,it}^{s}$ stand for the expectations of the error terms $\varepsilon_{k,it}^{y}$ and $\varepsilon_{k,it}^{x}$ conditional on $\boldsymbol{\varepsilon}_{it}^{s}$. They significantly increase the computation burden, but they have negligible impacts in our results.

parameters $\beta_{k,i}^{y}$, $\beta_{k,i}^{x}$ and $\alpha_{k,i}^{x}$. Hence, the crop gross margin difference $\pi_{k,it} - \pi_{K,it}$ and the random parameters of the acreage share equations can be correlated if these random parameters and those of crop *k* and/or crop *K* are correlated. Second, assuming that crop *k* belongs to crop group *g*, the acreage share of crop *k* in the acreage of group *g*, with

$$(18) \qquad s_{k|(g),it} = \frac{\exp\left[\alpha_{(g),i}\left(\pi_{k,it} - \beta_{k,i}^{s} - (\boldsymbol{\delta}_{k,0}^{s})'\mathbf{c}_{k,it}^{s} - \varepsilon_{k,it}^{s}\right)\right]}{\sum_{\ell \in \mathcal{K}(g)} \exp\left[\alpha_{(g),i}\left(\pi_{\ell,it} - \beta_{\ell,i}^{s} - (\boldsymbol{\delta}_{\ell,0}^{s})'\mathbf{c}_{\ell,it}^{s} - \varepsilon_{\ell,it}^{s}\right)\right]},$$

is a function of the slope parameters $\alpha_{(g),i}^{s}$ and the error terms $\varepsilon_{k,it}^{s}$. As a result, the $\ln s_{k|(g),it}$ variables are, by construction, endogenous in their equation and statistically linked with their slope coefficient.[47]

The likelihood function of any parametric specification of the RP multi-crop model (18) explicitly accounts for the functional forms of the $\ln s_{k|(g),it}$ and $\pi_{k,it}$ variables and for the statistical links of the random elements this model, implying that the corresponding ML estimator "automatically" takes on the endogeneity and correlation issues described above.[48] Of course, the distributional assumptions related to the random parameters (and to the errors terms) of the model also contribute to the identification of the model parameters. In

[47] Indeed, this issue arising in the log-linearized version of the NMNL acreage share model is one side of a coin, the other side being the non-linear functional form of the original version of this model. Equation (7) describes the solution in **s** to the considered expected maximization problem, while equation (8) describes (first order) conditions characterizing these solutions.

[48] The model likelihood function is provided in the first chapter of this PhD thesis.

particular, the probability distribution of α_i^s is mostly identified by the variation in the gross margin terms $\pi_{k,it} - \pi_{K,it}$ given the correlation structure linking these terms. In our model, the correlation structure between the gross margin terms $\pi_{k,it} - \pi_{K,it}$ and the random parameter α_i^s is provided by the assumed joint probability distribution of the model random parameters. This correlation structure also contributes to the identification of the probability distribution of α_i^s.

5.4. The F/RP multi-crop model

In the RP multi-crop model, farmers are allowed to respond heterogeneously to economic incentives through the random slope coefficients $\alpha_{k,i}^x$ of the price ratio variables $v_{k,it} p_{k,it}^{-1}$ in the yield supply and input demand equations and through the random slope coefficients α_i^s and $\alpha_{(g),i}^s$ of the crop return variables $\pi_{k,it}$ and of the within group crop acreage share variables $\ln s_{k|(g),it}$ in the acreage shares equations.

The fixed and random parameter (F/RP) multi-crop model is defined for investigating the impacts of ignoring the variability in the production choice flexibility parameters $\alpha_{k,i}^x$, α_i^s and $\alpha_{(g),i}^s$. This model is obtained from the RP multi-crop model by imposing homogeneity restrictions on the production choice flexibility parameters. The restrictions $\alpha_{k,i}^x = \alpha_{k,0}^x$ for $k \in \mathcal{K}$ are imposed on the variable input choice flexibility parameters, and the restrictions

$\alpha_i^s = \alpha_0^s$ and $\alpha_{(g),i}^s = \alpha_{(g),0}^s$ for $g \in \mathcal{G}$ are imposed on the acreage choice flexibility parameters. The F/RP multi-crop model is related to the standard random effect models. It contains additively separable farm specific random parameters: the farm-specific potential yield levels $\beta_{k,i}^y$, the input requirement levels $\beta_{k,i}^x$ and the acreage choice shifters $\beta_{k,i}^s$. However, the expected crop returns $\pi_{k,it}$ depending on the yield supply and input demand shifters, $\beta_{k,i}^y$ and $\beta_{k,i}^x$, distinguishes this model from the usual panel data multi-crop models.

5.5. Estimation of the RP model and calibration of the Sim-RP model

This section presents the parametric specification of the RP multi-crop model and briefly describes the estimation procedure used for its estimation. We also provide a sketch of the procedure that we implement for estimating its parameter and for calibrating the resulting simulation model, hereafter designated the Sim-RP model.[49] Analogous estimation and calibration procedures are used for estimating the F/RP model and calibrating the simulation model obtained from this model, the Sim-F/RP model.

5.5.1. Parametric specification of the RP multi-crop model and exogeneity assumptions

[49] Further details on these topics are available in the second chapter of this PhD thesis.

The RP multi-crop model described in equation (17) comprises five main components. (*a*) Farmers' production choice variables are collected in the dependent variable vector $\mathbf{w}_{it}$. (*b*) Price and control variables are collected in the explanatory variable vector $\mathbf{z}_{it}$. (*c*) The coefficients of the control variable are collected in the vector $\boldsymbol{\delta}_0$. (*d*) The potential yield, input requirement and acreage choice shifter parameters are collected in the additively separable random parameter vector $\boldsymbol{\beta}_i$. The input use flexibility and acreage choice flexibility parameters are collected in the production choice flexibility parameter vector $\boldsymbol{\alpha}_i$. The vector $\boldsymbol{\rho}_i = (\boldsymbol{\beta}_i, \ln \boldsymbol{\alpha}_i)$ collects the random parameters of the model.[50] (*e*) The model error terms are collected in the vector $\boldsymbol{\varepsilon}_{it}$.

The random parameter vector of the model is assumed to be normally distributed with $\boldsymbol{\rho}_i \sim \mathcal{N}(\boldsymbol{\mu}_0, \boldsymbol{\Omega}_0)$.[51] This probability distribution describes the distribution of the random parameters across the farmer population represented by the considered sample. The error term vector is assumed to be normally distributed with $\boldsymbol{\varepsilon}_{it} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Psi}_0)$. We assume that the

---

[50] With $\boldsymbol{\beta}_i = (\boldsymbol{\beta}_i^y, \boldsymbol{\beta}_i^x, \boldsymbol{\beta}_i^s)$ and $\boldsymbol{\alpha}_i = (\boldsymbol{\alpha}_i^x, \boldsymbol{\alpha}_i^s)$ according to obvious notations.

[51] Although standard in RP models, our distribution choice for the random parameter vector is admittedly convenient. We considered, as robustness checks, different Johnson's transformations of the multivariate normal distribution (Stanfield *et al* 1996) for the random parameter vector. While allowing for a wide range of distribution functional forms, these transformations have little impact on our estimation results and on their interpretation. Note that $\boldsymbol{\alpha}_i$ is log-normally distributed.

random parameter vector $\boldsymbol{\rho}_i$, the error term vector $\boldsymbol{\varepsilon}_{it}$ and the price and the control variable vector $\mathbf{z}_{it}$ are mutually independent. Finally, we assume that the error terms $\boldsymbol{\varepsilon}_{it}$ are independent across time and that the explanatory variable vector $\mathbf{z}_{it}$ is strictly exogenous with respect to these error terms.

The exogeneity assumptions characterizing this assumption set are standard in short panel data multi-crop econometric models (see, *e.g.*, Lacroix and Thomas 2011; Platoni et al 2012; Bayramoglu and Chakir 2016). In particular, climatic and anticipated price variables are generally considered as exogenous with respect to the model's random elements. These variables are mostly determined by factors that are external to individual farms. Additionally, as they mostly vary across time in our sample, they can be considered as uncorrelated with the time invariant random parameters. Note also that restrictions on the serial correlation of the $\boldsymbol{\varepsilon}_{it}$ error terms are required for identifying the probability distribution of the model's random parameters (Arellano and Bonhomme, 2012).[52]

5.5.2. Sketch of the estimation procedure

[52] These restrictions on the serial correlations of the model error terms can be relaxed to account for dynamic features of the multi-crop technology. Our estimation framework can be adapted accordingly. Yet, this would require a significant extension of our modeling framework that is out of the scope of this article. In particular, rational farmers are expected to consider these serial correlations for improving their yield expectations.

The aim of the estimation procedure is to estimate the fixed parameter vector of the model $\boldsymbol{\delta}_0$, the variance matrix of the error term vector $\boldsymbol{\Psi}_0$ and the parameters of the random parameter joint probability distribution, $\boldsymbol{\mu}_0$ and $\boldsymbol{\Omega}_0$. Let the term $\boldsymbol{\theta}_{w,0}$ denote the parameter sub-vector collecting the elements of $\boldsymbol{\delta}_0$ and $\boldsymbol{\Psi}_0$ and let $\boldsymbol{\theta}_{\rho,0}$ denote the parameter sub-vector collecting the elements of $\boldsymbol{\mu}_0$ and of $\boldsymbol{\Omega}_0$. Of course, the term $\boldsymbol{\theta}_0 \equiv (\boldsymbol{\theta}_{w,0}, \boldsymbol{\theta}_{\rho,0})$ defines the parameter vector to be estimated. The ML estimation framework is well suited for estimating $\boldsymbol{\theta}_0$ in our parametric RP multi-crop model. Let $\mathbf{w}_i$ denote the observed choice sequence of farmer i and $\mathbf{z}_i$ denote the corresponding exogenous variable vector. The structural and distributional assumptions defining the RP multi-crop model allow for determining the probability density function of $\mathbf{w}_i$ conditional on $\mathbf{z}_i$ and $\boldsymbol{\rho}_i$ in analytical closed form. Let $f(\mathbf{w}_i \mid \mathbf{z}_i; \boldsymbol{\rho}_i; \boldsymbol{\theta}_{w,0})$ denote this density function.

Let $\varphi(\mathbf{u}; \mathbf{B})$ denote the probability density function of $\mathcal{N}(\mathbf{0}, \mathbf{B})$ at $\mathbf{u}$. Due to the distribution of $\boldsymbol{\varepsilon}_{it} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Psi}_0)$ and the functional form of the model (equations 11 and 12), the probability density function of $\mathbf{w}_{it}$ conditional on $\mathbf{z}_{it}$ and $\boldsymbol{\rho}_i$ is given by:

(19a) $$f(\mathbf{w}_{it} \mid \mathbf{z}_{it}; \boldsymbol{\rho}_i; \boldsymbol{\theta}_{w,0}) = (\alpha_i^s)^{1-G} \left(\prod_{g \in \mathcal{G}} (\alpha_{(g),i}^s)^{1-K_g} \right) \left(\prod_{k \in \mathcal{K}} s_{k,it}^{-1} \right) \times \varphi\left(\mathbf{e}_{it}(\mathbf{w}_{it}, \mathbf{z}_{it}, \boldsymbol{\rho}_i; \boldsymbol{\delta}_0); \boldsymbol{\Psi}_{it} \right)$$

with

(19b) $$\mathbf{e}_{it}(\mathbf{w}_{it}, \mathbf{z}_{it}, \boldsymbol{\rho}_i; \boldsymbol{\delta}_0) \equiv \left[\boldsymbol{\varepsilon}_{it}^y, \boldsymbol{\varepsilon}_{it}^x, \mathbf{e}_{it}^s \right],$$

where

(19c) $e^s_{k,it} = p_{k,it}(\boldsymbol{\delta}^{y,\varepsilon}_{k,0})'\varepsilon^s_{k,it} - v_{k,it}(\boldsymbol{\delta}^{x,\varepsilon}_{k,0})'\varepsilon^s_{k,it} - \varepsilon^s_{k,it}$.

The derivation of $f(\mathbf{w}_{it} \mid \mathbf{z}_{it};\boldsymbol{\rho}_i;\boldsymbol{\theta}_{w,0})$ makes use of Berry's (1994) device for inverting Nested Logit functions and the fact that $\varepsilon^s_{k,it}$ is assumed to be correlated to $\varepsilon^y_{k,it}$ or $\varepsilon^x_{k,it}$.[53] The term $p_{k,it}(\boldsymbol{\delta}^{y,\varepsilon}_{k,0})'\boldsymbol{\varepsilon}^s_{k,it} - v_{k,it}(\boldsymbol{\delta}^{x,\varepsilon}_{k,0})'\boldsymbol{\varepsilon}^s_{k,it}$ is added to the $\varepsilon^s_{k,it}$ in equation (17):

$$\begin{cases} y_{k,it} = \beta^y_{k,i} + (\boldsymbol{\delta}^y_{k,0})'\mathbf{c}^y_{k,it} - (1/2)\alpha^x_{k,i}v^2_{k,it}p^{-2}_{k,it} + \varepsilon^y_{k,it}, \quad k \in \mathcal{K} \\ x_{k,it} = \beta^x_{k,i} + (\boldsymbol{\delta}^x_{k,0})'\mathbf{c}^x_{k,it} - \alpha^x_{k,i}v_{k,it}p^{-1}_{k,it} + \varepsilon^x_{k,it}, \quad k \in \mathcal{K} \\ \ln s_{k,it} - \ln s_{K,it} = \alpha^s_i(\pi_{k,it} - \pi_{K,it}) - \alpha^s_i\left(\beta^s_{k,i} + (\boldsymbol{\delta}^s_{k,0})'\mathbf{c}^s_{k,it} + \varepsilon^s_{k,it}\right) \\ \qquad\qquad + \left(1 - \alpha^s_i(\alpha^s_{(g),i})^{-1}\right)\ln s_{k|(g),it} - \left(1 - \alpha^s_i(\alpha^s_{(G),i})^{-1}\right)\ln s_{K|(G),it}, \quad k \in \mathcal{K} \end{cases}$$

through the linear terms $\beta^s_{k,it} = \beta^s_{k,i} + (\boldsymbol{\delta}^s_{k,0})'\mathbf{c}^s_{k,it} + \varepsilon^s_{k,it}$. The terms $(\boldsymbol{\delta}^{y,\varepsilon}_{k,0})'\boldsymbol{\varepsilon}^s_{k,it}$ and $(\boldsymbol{\delta}^{x,\varepsilon}_{k,0})'\boldsymbol{\varepsilon}^s_{k,it}$ stand for the expectations of the error terms $\varepsilon^y_{k,it}$ and $\varepsilon^x_{k,it}$ conditional on $\boldsymbol{\varepsilon}^s_{it}$. The term $\boldsymbol{\Psi}_{it}$ is a function of $\boldsymbol{\Psi}_0$, $p_{k,it}$, $(\boldsymbol{\delta}^{y,\varepsilon}_{k,0})$, $v_{k,it}$ and $(\boldsymbol{\delta}^{x,\varepsilon}_{k,0})$.

The likelihood function of observation *i* at $\boldsymbol{\theta}$ is then given by

(20a) $\ell_i(\boldsymbol{\theta}) = \int f(\mathbf{w}_i \mid \mathbf{z}_i;\boldsymbol{\rho};\boldsymbol{\theta}_w)g(\boldsymbol{\rho};\boldsymbol{\theta}_\rho)d\boldsymbol{\rho}$

where the term $g(\boldsymbol{\rho}_i;\boldsymbol{\theta}_{\rho,0})$ denotes the probability density function of the random parameter vector $\boldsymbol{\rho}_i$. Due to $\boldsymbol{\rho}_i \sim \mathcal{N}(\boldsymbol{\mu}_0,\boldsymbol{\Omega}_0)$, then

---

[53] Further details regarding the derivation of $f(\mathbf{w}_{it} \mid \mathbf{z}_{it};\boldsymbol{\rho}_i;\boldsymbol{\theta}_{w,0})$ can be found in the first Chapter of this PhD thesis.

(20b) $g(\boldsymbol{\rho}_i;\boldsymbol{\theta}_{\rho,0}) = \varphi(\boldsymbol{\rho}_i - \boldsymbol{\mu}_0;\boldsymbol{\Omega}_0)$.

The likelihood function $\ell_i(\boldsymbol{\theta})$ can be obtained neither in analytical closed form nor numerically because of the multidimensional integral over the random parameter distribution. The individual likelihood functions, *i.e.* the $\ell_i(\boldsymbol{\theta})$ terms, must be integrated with simulation methods, implying that the estimators of $\boldsymbol{\theta}$ must be simulated counterparts of the standard ML estimator. The $\ell_i(\boldsymbol{\theta})$ term can be estimated by generating S independent (pseudo-)random draws from $g(\boldsymbol{\rho}_i;\boldsymbol{\theta}_\rho)$, the $\tilde{\boldsymbol{\rho}}_{i,s}(\boldsymbol{\theta}_\rho)$ terms for $s = 1,...,S$. The empirical mean $S^{-1}\sum_{s=1}^{S} f(\mathbf{w}_i \mid \mathbf{z}_i;\tilde{\boldsymbol{\rho}}_{i,s}(\boldsymbol{\theta}_\rho);\boldsymbol{\theta}_w)$ is an unbiased simulator of $f(\mathbf{w}_i \mid \mathbf{z}_i;\boldsymbol{\theta}_w)$.

While applied econometricians mostly use Simulated ML (SML) estimators in similar contexts, applied statisticians mostly rely on on Stochastic Expectation-Maximization (SEM) algorithms for computing asymptotically equivalent estimators (see, *e.g.*, Jank and Booth, 2003). The asymptotic properties, as S and N grows to infinity with S rising faster than $N^{1/2}$, of the SML and SEM estimators are those of the "infeasible" ML estimator of $\boldsymbol{\theta}$ (Jank and Booth, 2003). Computing a SML estimator of our RP multi-crop model would involve solving a single but intimidating simulated likelihood maximization problem.[54]

---

[54] First, the dimension of the parameter vector to be estimated is relatively large because of the dimension of the considered RP equation system. Second, the functional form of the simulated likelihood function of our panel data RP multi-crop model is particularly awkward. It is integrated by simulation methods over the joint

SEM algorithms are extensions of the expectation-maximization (EM) algorithms proposed by Dempster et al (1977) and, as such, they are particularly well-suited for maximizing likelihood functions involving missing variables such as random parameters (see, *e.g.*, McLachlan and Krishnan 2008; Lavielle 2014). EM algorithms consist of iterating two steps, the expectation (E) step and the maximization (M) step, until numerical convergence.

The EM algorithm is particularly well suited for computing ML estimators in cases where the model of interest involves hidden variables such as random parameters. It consists in iterating two steps, the Expectation step (E step) and the Maximization step (M step), until numerical convergence. It basically replaces a large ML problem by a sequence of simpler maximization problems. In our case, the E step consists of computing expectations of the log of the individual "complete data" likelihood functions, $f(\mathbf{w}_i \mid \mathbf{z}_i;\boldsymbol{\rho}_i;\boldsymbol{\theta}_w)g(\boldsymbol{\rho}_i;\boldsymbol{\theta}_\rho)$, conditionally on the observed data.[55] At the jth iteration, this expectation is integrated over the probability distribution of $\boldsymbol{\rho}_i$ conditional on $(\mathbf{w}_i,\mathbf{z}_i)$ as this distribution is characterized by the estimate of $\boldsymbol{\theta}_0$ that is obtained from the last iteration, $\boldsymbol{\theta}^{(j-1)}$. These conditional expectations are denoted $L_i^E(\boldsymbol{\theta} \mid \boldsymbol{\theta}^{(j-1)})$ with

---

probability distribution of the model random parameter vector and it considers a relatively large dependent variable vector with a panel data structure.

[55] The "complete data" contains the observed dependent variables $\mathbf{w}_i$ and the random parameters $\boldsymbol{\rho}_i$ that are considered as unobserved variables in EM algorithms. Accordingly, "complete data" likelihood function corresponds to the density function of $(\mathbf{w}_i,\boldsymbol{\rho}_i)$ conditional on $\mathbf{z}_i$.

(21)
$$L_i^E(\boldsymbol{\theta}\,|\,\boldsymbol{\theta}^{(j-1)}) = \int \ln f(\mathbf{w}_i\,|\,\mathbf{z}_i;\boldsymbol{\rho};\boldsymbol{\theta}_w)h(\boldsymbol{\rho}\,|\,\mathbf{w}_i,\mathbf{z}_i;\boldsymbol{\theta}^{(j-1)})d\boldsymbol{\rho} + \int \ln g(\boldsymbol{\rho};\boldsymbol{\theta}_\rho)h(\boldsymbol{\rho}\,|\,\mathbf{w}_i,\mathbf{z}_i;\boldsymbol{\theta}^{(j-1)})d\boldsymbol{\rho}.$$

where the term $h(\boldsymbol{\rho}_i\,|\,\mathbf{w}_i,\mathbf{z}_i;\boldsymbol{\theta}_0)$ denotes the probability density function of the random parameter vector $\boldsymbol{\rho}_i$ conditional on $(\mathbf{w}_i,\mathbf{z}_i)$. The M step of the jth iteration then consists of maximizing in $\boldsymbol{\theta}$ the congruent expectation of the sample "complete data" log-likelihood function for obtaining an update of the estimates of $\boldsymbol{\theta}_0$. This implies that $\boldsymbol{\theta}^{(j)}$ is defined as the solution in $\boldsymbol{\theta}$ to the maximization problem:

(22a)
$$\boldsymbol{\theta}^{(j)} = \arg\max_{\boldsymbol{\theta}} \sum_{i=1}^{N} L_i^E(\boldsymbol{\theta}\,|\,\boldsymbol{\theta}^{(j-1)})$$

or by finding

(22b) $\boldsymbol{\theta}^{(j)}$ such that $\sum_{i=1}^{N} L_i^E(\boldsymbol{\theta}^{(j)}\,|\,\boldsymbol{\theta}^{(j-1)}) > \sum_{i=1}^{N} L_i^E(\boldsymbol{\theta}^{(j-1)}\,|\,\boldsymbol{\theta}^{(j-1)})$, if possible.

In the case of RP models the maximization problem (21) is much simpler to solve than the corresponding SML maximization problem. Indeed, EM algorithms aim at replacing large and involved ML maximization problems by sequences of simple maximization problems. Moreover, they are shown to converge to a maximum of the considered likelihood function.

Equation (22a) defines a standard EM algorithm while condition (22b) defines a Generalized EM algorithm (Dempster et al 1977). In some cases to define $\boldsymbol{\theta}^{(j)}$ by condition (22b) is much less computationally demanding than to maximize $\sum_{i=1}^{N} L_i^E(\boldsymbol{\theta}\,|\,\boldsymbol{\theta}^{(j-1)})$ in $\boldsymbol{\theta}$.

EM algorithms allow taking advantage of the mixed structure of RP models. Equation (22) is specific to models involving hidden variables such as random parameters. In our case it is used to split the M step into two simpler problems, *i.e.* to find

(23a) $$\boldsymbol{\theta}_w^{(j)} \text{ such that } Q_N^c(\boldsymbol{\theta}_w^{(j)} \mid \boldsymbol{\theta}_{n-1}) > Q_N^c(\boldsymbol{\theta}_w^{(j-1)} \mid \boldsymbol{\theta}_{n-1})$$

where

(23b) $$Q_N^c(\boldsymbol{\theta}_w \mid \boldsymbol{\theta}^{(j-1)}) \equiv \sum_{i=1}^N \int \ln f(\mathbf{w}_i \mid \mathbf{z}_i; \boldsymbol{\rho}; \boldsymbol{\theta}_w) h(\boldsymbol{\rho} \mid \mathbf{w}_i, \mathbf{z}_i; \boldsymbol{\theta}^{(j-1)}) d\boldsymbol{\rho}) ,$$

and to obtain

(23c) $$\boldsymbol{\theta}_\rho^{(j)} \equiv \arg\max_{\boldsymbol{\theta}_\rho} Q_N^q(\boldsymbol{\theta}_\rho \mid \boldsymbol{\theta}^{(j-1)})$$

where

(23d) $$Q_N^c(\boldsymbol{\theta}_\rho \mid \boldsymbol{\theta}^{(j-1)}) \equiv \sum_{i=1}^N \int \ln g(\boldsymbol{\rho}; \boldsymbol{\theta}_\rho) h(\boldsymbol{\rho} \mid \mathbf{w}_i, \mathbf{z}_i; \boldsymbol{\theta}^{(j-1)}) d\boldsymbol{\rho}$$

and

(23e) $$\boldsymbol{\theta}^{(j)} \equiv (\boldsymbol{\theta}_w^{(j)}, \boldsymbol{\theta}_\rho^{(j)}) .$$

The parameters of the "behavioral model" on the one hand, and those of the "mixing" probability distribution model on the other hand are separately updated in this M step.

Stochastic EM algorithms

The EM algorithm described above would lead to the ML estimator of $\boldsymbol{\theta}$. The SEM algorithms were proposed to extend the use of the EM algorithms in cases where the E step requires integration by simulation methods, as is the case for our RP multi-crop model.

Numerous SEM algorithms have been proposed in the computational statistics literature, thereby offering a rich toolbox for estimating a wide range of RP models.

Wei, Tanner (1990), proposed the Monte Carlo EM algorithm (MCEM). This algorithm allows approximating the E step using Monte Carlo approximation. In our application, the expectations in equation (21) were integrated with an (self-normalized) Importance Sampling (IS) simulator employed in a similar context by Train (2008, 2009). This IS simulator uses $g(\boldsymbol{\rho};\boldsymbol{\theta}_{\rho}^{(j-1)})$ as the proposal density function for estimating expectations of functions of $\boldsymbol{\rho}_i$ conditional on $(\mathbf{w}_i,\mathbf{z}_i)$ parameterized by $\boldsymbol{\theta}^{(j)}$.[56,57] It was used for approximating the expectations of the E step, *i.e.* $Q_N^c(\boldsymbol{\theta}_w \mid \boldsymbol{\theta}_{n-1})$ was approximated by:

(24a) $$\tilde{Q}_{N,S_n}^c(\boldsymbol{\theta}_w \mid \boldsymbol{\theta}^{(j-1)}) \equiv \sum_{i=1}^{N} S_n^{-1} \sum_{s=1}^{S_n} \sum_{t=1}^{T} \tilde{\omega}_{i,s}(\boldsymbol{\theta}^{(j-1)}) \ln f(\mathbf{w}_{it} \mid \mathbf{z}_{it};\tilde{\boldsymbol{\rho}}_{i,s}(\boldsymbol{\theta}_{\rho}^{(j-1)});\boldsymbol{\theta}_w),$$

while $Q_N^c(\boldsymbol{\theta}_\rho \mid \boldsymbol{\theta}^{(j-1)})$ was approximated by:

(24b) $$\tilde{Q}_{N,S_n}^q(\boldsymbol{\theta}_\rho \mid \boldsymbol{\theta}^{(j-1)}) \equiv \sum_{i=1}^{N} S_n^{-1} \sum_{s=1}^{S_n} \tilde{\omega}_{i,s}(\boldsymbol{\theta}^{(j-1)}) \ln g\left(\tilde{\boldsymbol{\rho}}_{i,s}(\boldsymbol{\theta}_{\rho}^{(j-1)});\boldsymbol{\theta}_\rho\right),$$

---

[56] It can be seen is a direct application of equation (4). It was used to estimate the conditional expectations in equations (14)–(15) as well as to calibrate the farmer's specific parameters in our empirical application.

[57] This proposal density is unlikely to be very efficient when the random parameters strongly impact the modeled choice and outcomes. Moreover the use of normal density functions as proposal density is usually unwarranted owing to their light tails. But this proposal density has two main advantages. First, its use is simple. Second, it allows interpreting the obtained estimator of $\boldsymbol{\theta}$ as a Method of Simulated Score estimator which is closely linked to the usual SML estimator (Train 2007). Statisticians rarely refer to SML estimators. They often approximate the $E[\ln \kappa(\mathbf{w}_i,\boldsymbol{\rho}_i \mid \mathbf{z}_i;\boldsymbol{\theta}) \mid \mathbf{z}_i,\mathbf{w}_i;\boldsymbol{\theta}^{(j-1)}]$ terms by directly drawing from the *ex post* density $h(\boldsymbol{\rho} \mid \mathbf{z}_i,\mathbf{w}_i;\boldsymbol{\theta}^{(j-1)})$, *e.g.* by using Metropolis-Hasting simulators.

where

(24c) $$\tilde{\omega}_{i,s}(\boldsymbol{\theta}) \equiv \frac{\prod_{t=1}^{T} f(\mathbf{w}_{it} \mid \mathbf{z}_{it}; \tilde{\boldsymbol{\rho}}_{i,s}(\boldsymbol{\theta}_{\rho}); \boldsymbol{\theta}_{w})}{S_n^{-1} \sum_{s=1}^{S_n} \prod_{t=1}^{T} f(\mathbf{w}_{it} \mid \mathbf{z}_{it}; \tilde{\boldsymbol{\rho}}_{i,s}(\boldsymbol{\theta}_{\rho}); \boldsymbol{\theta}_{w})} \ .$$

The objective functions of problems (22a) and (22b) takes advantage of the panel structure of the data. The objective functions $\tilde{Q}^{c}_{N,S_n}(\boldsymbol{\theta}_w \mid \boldsymbol{\theta}^{(j-1)})$ and $\tilde{Q}^{q}_{N,S_n}(\boldsymbol{\theta}_\rho \mid \boldsymbol{\theta}^{(j-1)})$ can be interpreted as the log-likelihood functions of standard models. They are simply weighted by the $\tilde{\omega}_{i,s}(\boldsymbol{\theta}^{(j-1)})$ terms and involve simulated pseudo-observations, the $(\mathbf{w}_i, \mathbf{z}_i, \tilde{\boldsymbol{\rho}}_{i,s}(\boldsymbol{\theta}_{\rho}^{(j-1)}))$ vectors for $i = 1,...,N$ and $s = 1,...,S_n$. To solve $\boldsymbol{\theta}_\rho$ and $\boldsymbol{\theta}_w$, the simulated counterparts of problems (23a) and (23c) appeared to be much easier than to directly maximize in $\boldsymbol{\theta}$ the corresponding sample simulated log-likelihood function

(25) $$\tilde{L}_{N,S}(\boldsymbol{\theta}) \equiv \sum_{i=1}^{N} \ln\left(S^{-1} \sum_{s=1}^{S} \left(\prod_{t=1}^{T} f(\mathbf{w}_{it} \mid \mathbf{z}_{it}; \tilde{\boldsymbol{\rho}}_{i,s}(\boldsymbol{\theta}_{\rho}); \boldsymbol{\theta}_{w}) \right) \right) \ .$$

The dimension of $\boldsymbol{\theta}$ is quite large in our case and the functional form of $\tilde{L}_{N,S}(\boldsymbol{\theta})$ makes it difficult to split its maximization problem into smaller optimization problems. The Newton-type algorithms usually employed for maximizing the simulated likelihood function in problem (25) only have local convergence properties. This makes the definition of "good" starting values a crucial and difficult issue.

Deterministic EM algorithms increase the sample log-likelihood at each iteration, implying that it generally leads to a (local) maximum of the considered likelihood function. SEM algorithms do not necessarily monotonically increase the simulated sample log-

likelihood due to the simulation noise, but they are expected to do so when S is "large enough".[58]

SAEM algorithms and ECM algorithms

The MCEM algorithm need to have S very "large enough" for convergence, that is not practical feasible in our case. The SEM algorithm used for estimating the RP multi-crop model is a SAEM algorithm (Delyon et al 1999; Lavielle 2014).[59] This algorithm is relatively easy to code and monitor. The SAEM algorithms consisting in a class of SEM with two main advantages. First, SAEM algorithms are numerically stable despite their requiring integrations relying on simulation methods at each of their iterations. Second, they allow using simplified versions of the M step. Contrary to the MCEM algorithm, the SAEM algorithm uses, at each iteration of the algorithm, all the simulations before the considered iteration. In addition to the simulation (S) step at the E step, the SAEM algorithm (SA) Stochastic Approximation step. Its (S) step uses an importance sampling simulator (Caffo et

[58] The main drawback of EM algorithms is their linear convergence rate. They are slower than Newton-type algorithms which enjoy a quadratic convergence rate. However, this comparison only holds locally, *i.e.* within a neighborhood of the likelihood function. Indeed, EM algorithms are known to quickly converge to a neighborhood of the likelihood function but to slowly converge within this neighborhood. This led Ruud (1991) to suggest designing hybrid algorithms with an EM algorithm starting the maximization process and a Newton-type algorithm terminating it.

[59] The SAEM algorithm is presented in the second chapter of this PhD thesis.

al 2005) and it was used for approximating the expectations of the (SA) step, *i.e.* $Q_N^c(\boldsymbol{\theta}_w \mid \boldsymbol{\theta}_{n-1})$ was approximated by:

(26a) $$\tilde{Q}_{N,S_n}^c(\boldsymbol{\theta}_w \mid \boldsymbol{\theta}^{(j-1)}) \equiv \tilde{Q}_{N,S_n}^c(\boldsymbol{\theta}_w \mid \boldsymbol{\theta}^{(j-2)}) + \gamma_j \left[\mathcal{H}_w(\tilde{\boldsymbol{\rho}}_{i,s}(\boldsymbol{\theta}_\rho^{(j-1)})) - \tilde{Q}_{N,S_n}^c(\boldsymbol{\theta}_w \mid \boldsymbol{\theta}^{(j-2)}) \right]$$

with

(26b) $$\mathcal{H}_w(\tilde{\boldsymbol{\rho}}_{i,s}(\boldsymbol{\theta}_\rho^{(j-1)})) \equiv \sum_{i=1}^N S_n^{-1} \sum_{s=1}^{S_n} \sum_{t=1}^T \tilde{\omega}_{i,s}(\boldsymbol{\theta}^{(j-1)}) \ln f(\mathbf{w}_{it} \mid \mathbf{z}_{it}; \tilde{\boldsymbol{\rho}}_{i,s}(\boldsymbol{\theta}_\rho^{(j-1)}); \boldsymbol{\theta}_w),$$

while $Q_N^q(\boldsymbol{\theta}_\rho \mid \boldsymbol{\theta}^{(j-1)})$ was approximated by:

(26c) $$\tilde{Q}_{N,S_n}^q(\boldsymbol{\theta}_\rho \mid \boldsymbol{\theta}^{(j-1)}) \equiv \tilde{Q}_{N,S_n}^q(\boldsymbol{\theta}_w \mid \boldsymbol{\theta}^{(j-2)}) + \gamma_j \left[\tilde{\mathcal{H}}_\rho(\tilde{\boldsymbol{\rho}}_{i,s}(\boldsymbol{\theta}_\rho^{(j-1)})) - \tilde{Q}_{N,S_n}^q(\boldsymbol{\theta}_\rho \mid \boldsymbol{\theta}^{(j-2)}) \right]$$

with

(26d) $$\tilde{\mathcal{H}}_\rho(\tilde{\boldsymbol{\rho}}_{i,s}(\boldsymbol{\theta}_\rho^{(j-1)})) \equiv \sum_{i=1}^N S_n^{-1} \sum_{s=1}^{S_n} \tilde{\omega}_{i,s}(\boldsymbol{\theta}^{(j-1)}) \ln g\left(\tilde{\boldsymbol{\rho}}_{i,s}(\boldsymbol{\theta}_\rho^{(j-1)}); \boldsymbol{\theta}_\rho\right),$$

where

(26e) $$\tilde{\omega}_{i,s}(\boldsymbol{\theta}) \equiv \frac{\prod_{t=1}^T f(\mathbf{w}_{it} \mid \mathbf{z}_{it}; \tilde{\boldsymbol{\rho}}_{i,s}(\boldsymbol{\theta}_\rho); \boldsymbol{\theta}_w)}{S_n^{-1} \sum_{s=1}^{S_n} \prod_{t=1}^T f(\mathbf{w}_{it} \mid \mathbf{z}_{it}; \tilde{\boldsymbol{\rho}}_{i,s}(\boldsymbol{\theta}_\rho); \boldsymbol{\theta}_w)} .$$

The term γ_j is a decreasing sequence such that $\sum \gamma_j = +\infty$ and $\sum \gamma_j^2 < +\infty$.

Importantly, its M step takes advantage of the structure of the considered panel data RP multi-crop model. It is decomposed into three simple (weighted) likelihood maximization problems: the first updates the fixed parameters of the input demand and yield supply sub-systems, the second updates the fixed parameters of the acreage share model sub-system and

the last updates the parameters of the random parameter probability distribution. Indeed, equations (26a) and (26b) shows that the maximization problem (22) can be decomposed into two simpler maximization problems. The first aims at updating the estimates of $\boldsymbol{\theta}_{\rho,0}$ while the second aims at updating the estimates of $\boldsymbol{\theta}_{w,0}$. This decomposition of problem (22) is specific to RP models. The equation system structure of our RP multi-crop model can also be exploited for further decomposing the maximization problem aimed at updating the estimates of $\boldsymbol{\theta}_{w,0}$. We consider the input demand and yield supply sub-systems on the one hand, and the acreage share model sub-system on the other hand. SML estimators do not allow such decompositions.

We finally end up with a SAEM algorithm for which the elements of $\boldsymbol{\theta}_{\rho}^{(j)}$ are computed as weighted empirical means and covariances whereas the elements of $\boldsymbol{\theta}_{w}^{(j)}$ (without the covariance matrice) are defined as weighted Feasible Generalized Least Squares estimators.[60]

5.5.3. The Sim-RP model and its "statistical" calibration

The estimation of the RP multi-crop model delivers an estimate, $\hat{\boldsymbol{\theta}}$, of the model "deep" parameter vector $\boldsymbol{\theta}_0$. This estimate includes an estimate of the parameter vector that are

[60] A detailed presentation of this algorithm is provided in the second chapter of this PhD thesis.

common to all farms, $\hat{\boldsymbol{\delta}}$, and an estimate of the probability distribution of $\boldsymbol{\rho}_i$ across the farm population. However, estimating the RP multi-crop model does not directly yields estimates of the farm specific parameters $\boldsymbol{\rho}_i$ for each sampled farm. As will be shown below, obtaining such estimates is relatively easy and useful for constructing simulation models for analyzing the effects agri-environmental policies such as EU biofuel support considered in our empirical application.

The functional form of considered RP multi-crop model describes how farmers' production choices respond to their drivers according to the multivariate equation $\mathbf{w}_{it} = \mathbf{w}(\mathbf{r}_{it}, \mathbf{c}_{it}, \boldsymbol{\varepsilon}_{it}; \boldsymbol{\delta}_0, \boldsymbol{\rho}_i)$. This equation defines the production choice vector $\mathbf{w}_{it}$ as a function of its economic drivers $\mathbf{r}_{it}$ and of the production conditions $\mathbf{c}_{it}$ (control variables) and $\boldsymbol{\varepsilon}_{it}$ (error terms). This response function depends on parameters that are common to all farms, $\boldsymbol{\delta}_0$, and on parameters that are farm specific, $\boldsymbol{\rho}_i$. It can be used for constructing a simulation model based on the sampled farms used in the estimation process. The Sim-RP model considered in our empirical application is constructed by computing an estimate of $\boldsymbol{\rho}_i$ for each farm, as well as an estimate of $\boldsymbol{\varepsilon}_{it}$ for each observation, of our sample. This simulation model is thus defined as a set of heterogeneous farm models. The construction of these farm specific models is described in what follows.

Assuming that $\hat{\boldsymbol{\rho}}_i$ is a reliable estimate of $\boldsymbol{\rho}_i$, the residual term $\hat{\boldsymbol{\varepsilon}}_{it}$ is defined as the solution in $\boldsymbol{\varepsilon}$ to the equation system $\mathbf{w}_{it} = \mathbf{w}(\mathbf{r}_{it}, \mathbf{c}_{it}, \boldsymbol{\varepsilon}; \hat{\boldsymbol{\delta}}, \hat{\boldsymbol{\rho}}_i)$ for ensuring that the considered model exactly reproduces the observed choice vector $\mathbf{w}_{it}$. The simulated response of farmer *i* to the

economic drivers described by $\mathbf{r}$ under the production conditions characterized by $\mathbf{c}_{it}$ and $\hat{\boldsymbol{\varepsilon}}_{it}$ can then be defined as $\mathbf{w}(\mathbf{r}, \mathbf{c}_{it}, \hat{\boldsymbol{\varepsilon}}_{it}; \hat{\boldsymbol{\delta}}, \hat{\boldsymbol{\rho}}_i)$.

Reliable estimates of $\boldsymbol{\rho}_i$ can be obtained for each sampled farm based on simple "statistical calibration" procedures relying on a well-defined statistical background: the estimated RP multi-crop model and what is known in the data about each farmer. These calibration procedures exploit the information content of sequence of production choices $\mathbf{w}_i$ and the conditions of these choices $\mathbf{z}_i$ for estimating $\boldsymbol{\rho}_i$ by relying on the links among these terms that are summarized in the estimated RP multi-crop model. The calibration procedure used for constructing our Sim-RP model aims at computing the expectation of $\boldsymbol{\rho}_i$ conditional on $(\mathbf{w}_i, \mathbf{z}_i)$. This defines $\hat{\boldsymbol{\rho}}_i$ as the best prediction of $\boldsymbol{\rho}_i$ given $(\mathbf{w}_i, \mathbf{z}_i)$ according to the minimum mean squared error criterion.[61] In our empirical application, we compute the terms $\hat{\boldsymbol{\rho}}_i$ for each farm of our sample by using an importance sampling simulator based on the equation $\hat{\boldsymbol{\rho}}_i = \int \boldsymbol{\rho} h(\boldsymbol{\rho} \mid \mathbf{w}_i, \mathbf{z}_i; \hat{\boldsymbol{\theta}}) d\boldsymbol{\rho}$.

[61] The term $\hat{\boldsymbol{\rho}}_i$ could have been be defined as the value of $\boldsymbol{\rho}$ that maximizes the conditional probability function $h(\boldsymbol{\rho} \mid \mathbf{w}_i, \mathbf{z}_i; \boldsymbol{\theta}_0)$, according to a ML estimation logic.

5.6. Empirical application: impact of the UE support to biofuels on acreage choices

As an illustrative application of the modeling framework proposed in this paper, we use a set of French data (*i*) to estimate the multi-crop RP model presented above, (*ii*) to estimate the effects of EU biofuel support on farmers' choices and (*iii*) to compare the statistical and simulation performances of this model with those of its F/RP counterpart.

5.6.1. Data

Our data set is a balanced panel data sample with 816 observations describing the production choices of 136 French farmers located in the Meuse *département*[62] over the years 2006 to 2011. It was obtained from the Meuse Management Centre (*Centre de Gestion de la Meuse*) and provides detailed information on crop production for each farm: acreages, yields, crop prices at the farm gate and cost accounting. The aggregated variable input (pesticides and fertilizers) price index was constructed with the indices made available at the regional level by the French Department of Agriculture. The climatic variables are provided at the municipality level by Meteo France, the French national meteorological service. All price variables are deflated by the agricultural production cost index provided by the French

[62] A *département* is a relatively small French territorial division (6211 km^2 for the Meuse *département*).

Department of Agriculture.

Our main objective in this empirical application being to investigate the importance of unobserved heterogeneity effects, the dataset has been selected to contain farms without trivial heterogeneous characteristics.[63] We focus on farms specialized in grain production over the years 2006 to 2011. The sampled farms only and always produce three crops – soft wheat, barley and rapeseed – and always use their entire arable land.[64] Their land area mostly ranges from 90 ha to 300 ha, with a mean of 150 ha. Standard summary statistics reveal that the sampled farmers' production choices are relatively rigid: their annual means evolve moderately despite large output price variations across years. As shown in Table 1, approximately 51% of the crop yield, 65% of the input use and 65% of the acreage share total variances are due to variations across farms. This suggests that farms' and farmers' heterogeneity accounts for a large part of the observed differences in farmers' production choices, even in small areas such as the Meuse *département* and for samples of farms with identical production sets.

---

[63] This partly explains why the size of our sample is admittedly limited. Nevertheless, the estimated precision level of our estimators shows that this sample size is sufficient for the purpose of this study. Indeed, the panel data parametric RP multi-crop model considered in this article describes 8 continuous production choices. This implies that its 172 parameters are estimated with 6528 observations corresponding to the farm/year/choice triplets observed in our dataset. Specifically, the probability distribution of the model random parameter vector is characterized by 104 parameters that are identified by 1088 farm/choice pairs repeated 6 times each.

[64] These selection criteria are not as stringent as they may appear. During the considered period, the acreages of the considered crops cover more than 90% of the arable land devoted to cash crops in the Meuse *département,* and more than 85% of the farmers produce these three crops in the considered area.

Table 1. Sample Summary Statistics

		Crop acreage shares (%)			Crop yields (ton/ha)			Variable input uses (€2005/ha)		
	Year	**Wheat**	**Barley**	**Rapeseed**	**Wheat**	**Barley**	**Rapeseed**	**Wheat**	**Barley**	**Rapeseed**
Annual means	**2006**	38.0	35.0	27.1	7.06	6.75	3.07	265	228	312
	2007	37.3	34.3	28.4	7.03	5.86	3.33	263	226	310
	2008	39.0	34.5	26.5	7.01	6.73	3.33	285	243	336
	2009	38.1	33.8	28.0	7.74	7.12	3.75	268	227	327
	2010	40.8	30.9	28.4	7.21	6.67	3.25	258	218	335
	2011	39.3	32.4	28.2	6.79	5.91	3.37	254	222	357
Standard deviation		0.08	0.10	0.06	0.83	0.94	1.18	0.55	0.53	0.72
Between farm variance in total variance (%)		65.3	69.6	61.9	50.7	51.1	50.6	65.8	65.4	64.8

		Crop gross margins (€/ha)			Crop prices (€/ton)			Variable input price indices (1 in 2005)		
	Year	**Wheat**	**Barley**	**Rapeseed**	**Wheat**	**Barley**	**Rapeseed**	**Wheat**	**Barley**	**Rapeseed**
Annual means	**2006**	520	413	357	115	98	233	1.11	1.09	1.14
	2007	1046	710	630	190	164	295	1.11	1.10	1.12
	2008	630	612	665	143	137	329	1.29	1.29	1.28
	2009	410	283	470	107	90	259	1.55	1.59	1.57
	2010	960	683	815	175	141	373	1.19	1.18	1.21
	2011	890	700	981	182	171	433	1.36	1.38	1.33

5.6.2. Modeling details

The estimated models are defined as equation systems with 8 equations each: 3 variable input demand equations, 3 yield supply equations and 2 acreage share equations (the share of rapeseed being obtained from those of wheat and barley). The RP model contains 13 random parameters, while the F/RP contains 8 random parameters. The empirical acreage share choice models were designed by grouping wheat and barley into a "cereals" nest with rapeseed being the only crop in its "oilseeds" nest. This implies that the (between nests) α_i^s parameter drives the land allocation to cereals and to rapeseed, while the (within cereals) $\alpha_{(\text{cereals}),i}^s$ parameter drives the allocation of the cereal acreage to wheat and to barley. Rapeseed is used as the benchmark crop in the acreage share models. We have tested several crop price anticipation schemes, including cases where price anticipations differ based on whether they are considered in the acreage choice models or in the input demand and yield supply models, various adaptive anticipation schemes and futures prices. Our results led us to adopt the simplest naïve anticipation scheme.

5.6.3. Statistical performances of the RP model

The *ex post* and *ex ante* distributions of the random parameters corresponding to the farm specific effects ($\boldsymbol{\beta}_i$) and choice flexibility parameters ($\boldsymbol{\alpha}_i$) are represented on Figures 1a and 1b respectively. We can notice here that the two distributions are almost superimposed for all

parameters, which reflects a good specification of our model.

Figure 1a. Ex post and ex ante probability distribution of the farm specific effects β_i

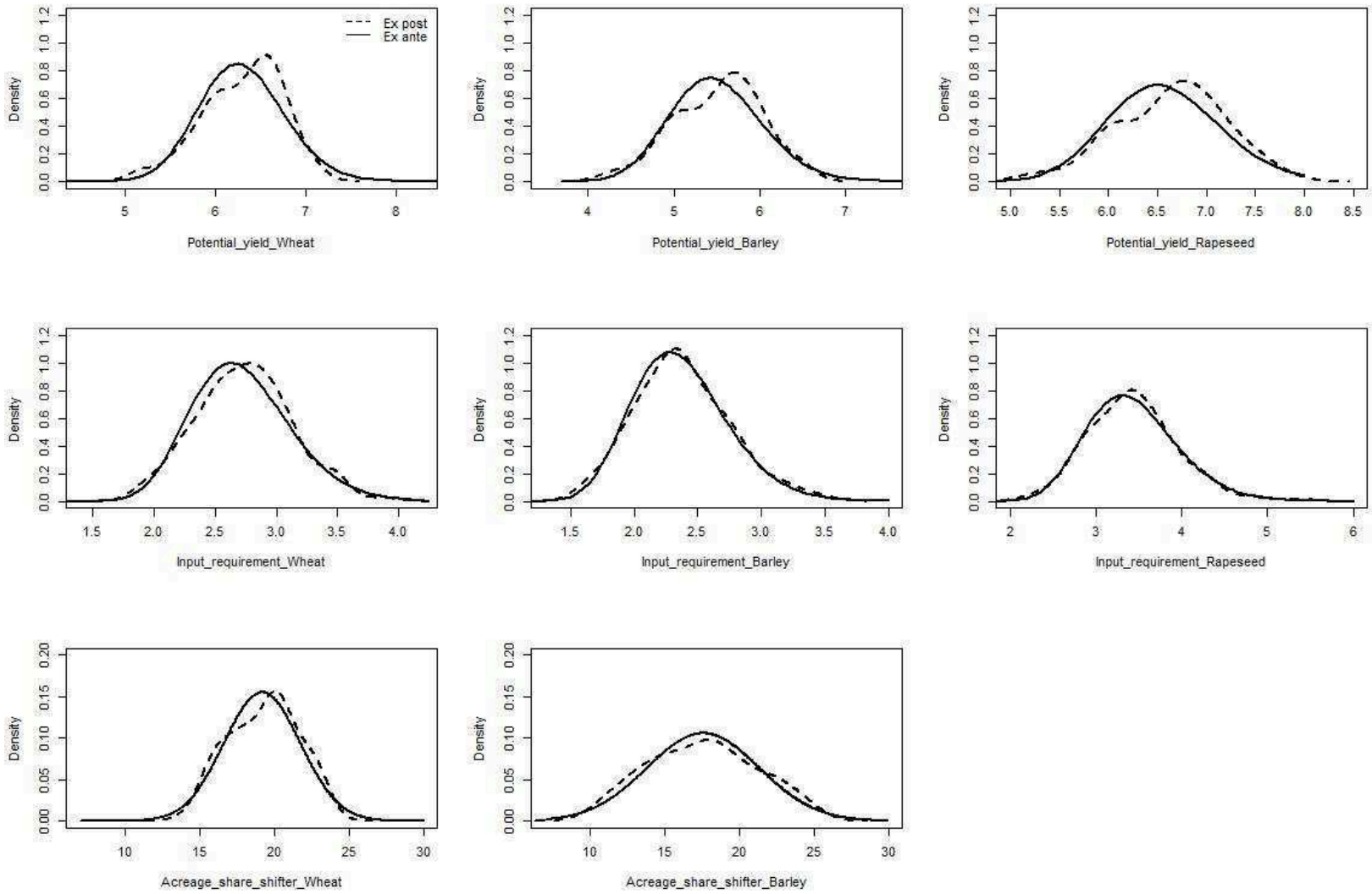
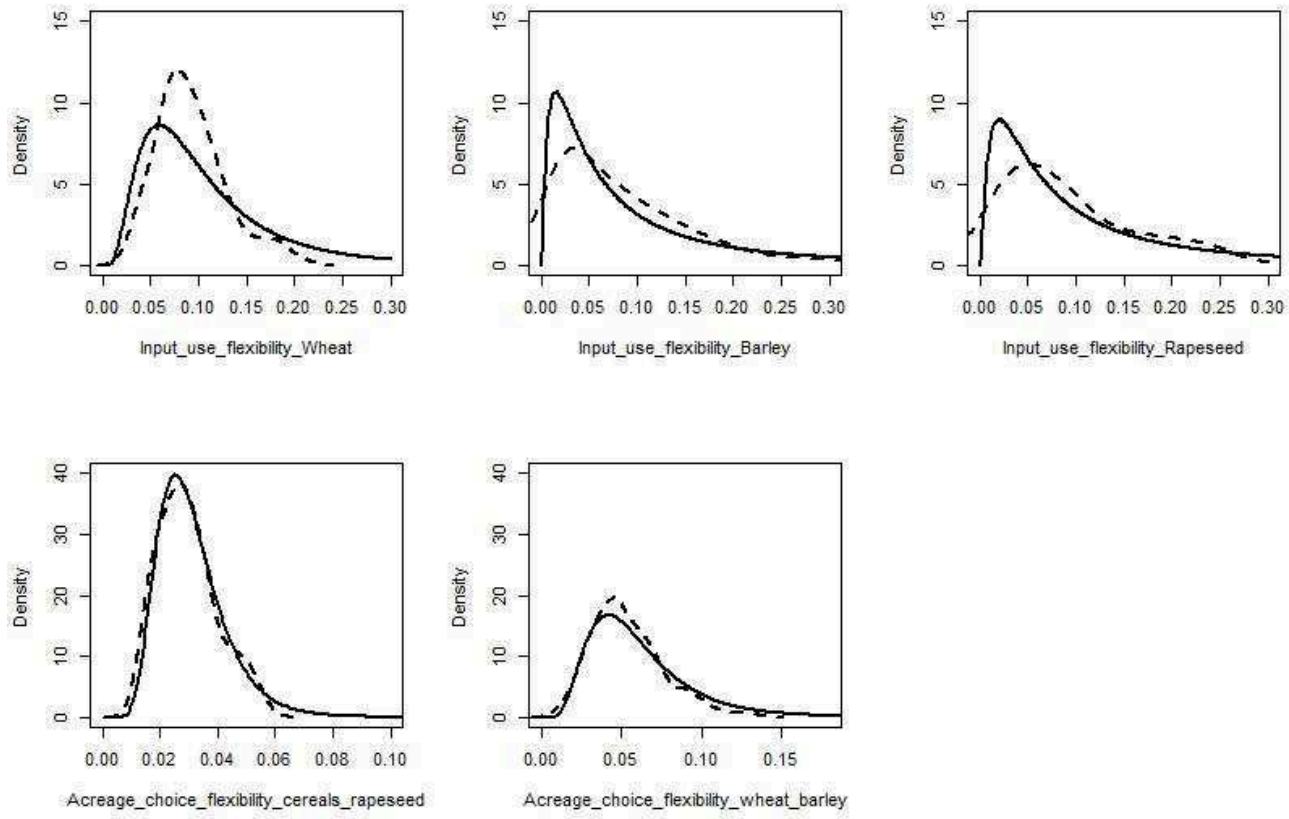


Figure 1b. Ex post and ex ante probability distribution of the farm specific effects α_i



As explained in section 5.3, the estimated parametric model allows a statistical calibration of the (random) parameters, $\boldsymbol{\rho}_i$, for each farm/farmer of the sample. Once the *ex ante* distribution of $\boldsymbol{\rho}_i$ in the population has been estimated we "statistically calibrate" the specific parameters for each individual i based on the *ex post* density of $\boldsymbol{\rho}_i$. More precisely, the "statistically calibrated" parameters $\hat{\boldsymbol{\rho}}_i$ are set to the means of $\boldsymbol{\rho}_i$ conditional on observed variables $(\mathbf{w}_i, \mathbf{z}_i)$ for each sampled farm/farmer. Having calibrated the individual parameters for each farm/farmer, we are able to compute the individual yields, input use variables and acreages predicted by the NMNL model. These predictions are then used to compute the "Sim-R^2" criteria, reported in Table 2, which measure the quality of the predictions of the observed choices by the Sim-RP and Sim-F/RP models. The Sim-R^2 criterion is analogous to the R^2 criterion of the standard linear regression model; for a given choice variable and a given model, it is defined as the ratio of the empirical variance of the predictions of this variable to the empirical variance of the observed variable. The results reported in Table 2 show that the Sim-RP model slightly outperforms the Sim-F/RP model: the Sim-R^2 values of both simulation models are close to each other, with slightly higher values for the Sim-RP model. The obtained Sim-R^2 values are close to the corresponding shares of the between farm variance in the total variance of the observed choice variables, demonstrating that the additively separable farm specific effects $\boldsymbol{\beta}_i$ "anchor" the predicted choices at the farm averages of these choices.

Table 2 also reports test results and fit criteria used for comparing the statistical performances

of the considered models, *i.e.,* the RP and F/RP models.[65] A standard likelihood ratio (LR) tests can be used for testing the "homogeneity" restriction $\boldsymbol{\alpha}_i = \boldsymbol{\alpha}_0$; it suffices to test the joint nullity of the corresponding parameters in $\boldsymbol{\Omega}_0$. These "homogeneity" restrictions are strongly rejected by the data, indicating that the production choice flexibility parameters $\boldsymbol{\alpha}_i$ significantly vary across farms.

[65] Our estimations were conducted by using the SAS® IML software. The MCEM algorithms were implemented using 3000 (scrambled Halton) draws to obtain the results presented herein. Such a large number of draws were used to neglect the simulation noise when computing the asymptotic distribution of the estimators. Similar results were obtained with 500 draws or more.

Table 2. Fit Criteria of the RP and F/RP Models

		RP Model	F/RP Model
Input demand equation Sim-R^2	Wheat	0.65	0.64
	Barley	0.64	0.63
	Rapeseed	0.65	0.64
Yield supply equation Sim-R^2	Wheat	0.45	0.40
	Barley	0.52	0.46
	Rapeseed	0.41	0.40
Acreage share equation Sim-R^2	Wheat	0.58	0.56
	Barley	0.66	0.64
	Rapeseed	0.57	0.56
Sample (simulated) log-likelihood		-1670.3	-1832.9
		Results of the "homogeneity" restriction tests	
Benchmark model		RP model: $\boldsymbol{\beta}_i$ and $\boldsymbol{\alpha}_i$	
Tested hypothesis		F/RP model: $\boldsymbol{\alpha}_i = \boldsymbol{\alpha}_0$	
Number of tested nullity constraints		19	
Estimated LR test statistic		325.2	
Critical value of test statistic (1%)		33.7	

5.7. Estimated distribution of the random parameter vector of the RP model

Selected parameter estimates of the RP model parameters are reported in Tables 3–5. These results show that most parameters of this model are precisely estimated. In particular, the estimates of the standard deviations of the random parameters $\boldsymbol{\alpha}_i$ are all significantly non

null,[66] thereby confirming farmers' heterogeneous responses to economic incentives. The acreage choice flexibility parameter α_i^s displays significant heterogeneity across farms, with a mean estimated at 0.031 and a standard deviation estimated at 0.012. It is worth noting that the estimate of $\alpha_{(\text{cereals}),i}^s$ exceeds that of α_i^s, as is sufficient for the considered acreage share choice model to be well behaved, for 133 of the 136 sampled farms. The parameters $\alpha_{(\text{cereals}),i}^s$ and α_i^s are significantly positively correlated, with a correlation coefficient exceeding 0.76. The estimates of the input use flexibility parameters $\boldsymbol{\alpha}_i^x$ are small, with means approximately 0.10, indicating that price changes have limited impact on the variable input uses and on the yield levels of the considered farmers. Unsurprisingly, our results indicate that farms are generally productive for all crops or none and that the yield levels and input use intensities are consistent. The potential yield parameters $\beta_{k,i}^y$ are strongly positively correlated with each other, with estimated correlation coefficients ranging from 0.92 to 0.89, and they are significantly correlated with the corresponding input requirement parameters $\beta_{k,i}^x$, with estimated correlation coefficients ranging from 0.68 to 0.33.

The acreage choice flexibility parameters α_i^s and $\alpha_{(\text{cereals}),i}^s$ are weakly (and not significantly) correlated to the random parameters of the crop yield supply and input demand equations. As a consequence, the acreage choice flexibility parameters α_i^s and $\alpha_{(\text{cereals}),i}^s$ are

[66] With the notable exception of the wheat input use flexibility parameter.

only weakly (and not significantly) correlated to the expected crop gross margins $\pi_{k,it}$, with estimated correlation coefficients not exceeding 0.10 in absolute value. These correlations only slightly impact our simulation results. Hence, the correlation between these economic drivers and their random slope coefficients in the acreage share model does not appear to be a serious issue in our application.

Table 3. Selected Parameter Estimates of the Input Demand Models[a]

	RP input demand models			F/RP input demand models		
	Wheat	***Barley***	***Rapeseed***	***Wheat***	***Barley***	***Rapeseed***
Error term ($\varepsilon_{k,it}^{x}$) std dev	0.354*	0.342*	0.458*	0.344*	0.341*	0.445*
	(0.013)	(0.014)	(0.018)	(0.013)	(0.013)	(0.017)
Input requirement ($\beta_{k,i}^{x}$)						
Mean	2.744*	2.386*	3.429*	2.886*	2.533*	3.738*
	(0.077)	(0.067)	(0.111)	(0.049)	(0.046)	(0.065)
Std dev	0.416*	0.393*	0.542*	0.421*	0.396*	0.525*
	(0.076)	(0.079)	(0.090)	(0.047)	(0.057)	(0.072)
Input use flexibility ($\alpha_{k,i}^{x}$)						
Mean	0.109*	0.130*	0.144*	0.263*	0.263*	0.466*
	(0.027)	(0.044)	(0.048)	(0.034)	(0.029)	(0.046)
Std dev [b]	0.643	1.196*	1.141*	0	0	0
	(0.347)	(0.586)	(0.530)			

a. Estimated standard errors of the ML estimator are in parentheses. b. Variance of the log of the random parameter.
Note: Asterisk (*) denotes a statistically significant parameter at the 5% level.

Table 4. Selected Parameter Estimates of Yield Supply Models[a]

	RP yield models			F/RP yield models		
	Wheat	***Barley***	***Rapeseed***	***Wheat***	***Barley***	***Rapeseed***
Error term ($\varepsilon_{k,it}^{y}$) std dev	0.643*	0.694*	0.938*	0.649*	0.694*	0.938*
	(0.023)	(0.025)	(0.030)	(0.023)	(0.026)	(0.036)
Potential yield ($\beta_{k,i}^{y}$)						
Mean	6.296*	5.520*	6.638*	6.432*	5.587*	6.696*
	(0.092)	(0.101)	(0.115)	(0.052)	(0.062)	(0.074)
Std dev	0.476*	0.545*	0.592*	0.444*	0.533*	0.633*
	(0.102)	(0.098)	(0.109)	(0.047)	(0.057)	(0.072)
Input use flexibility ($\alpha_{k,i}^{x}$)						
Mean	0.109*	0.130*	0.144*	0.263*	0.263*	0.466*
	(0.027)	(0.044)	(0.048)	(0.034)	(0.029)	(0.046)
Std dev [b]	0.643	1.196*	1.141*	0	0	0
	(0.347)	(0.586)	(0.530)			

a. Estimated standard errors of the ML estimator are in parentheses. b. Variance of the log of the random parameter.
Note: Asterisk (*) denotes a statistically significant parameter at the 5% level.

Table 5. Selected Parameter Estimates of the Acreage Share Models[a]

	RP acreage share models			F/RP acreage share models		
	Wheat	***Barley***	***Rapeseed***	***Wheat***	***Barley***	***Rapeseed***
Error term ($\varepsilon^{s}_{k,it}$) std dev	7.672*	8.393*	0	7.750*	8.965*	0
	(0.262)	(0.305)		(0.382)	(0.436)	
Acreage share shifters ($\beta^{s}_{k,i}$)						
Mean	19.202*	17.568*	0	15.275*	13.361*	0
	(0.505)	(0.718)		(1.011)	(1.346)	
Std dev	2.579*	3.775*	0	6.874*	9.879*	0
	(0.686)	(0.693)		(0.984)	(0.853)	
Acreage choice flexibility	***Cereals vs rapeseed*** **(α^{s}_{i})**	***Wheat vs barley*** **($\alpha^{s}_{(\text{cereals}),i}$)**		***Cereals vs rapeseed*** **(α^{s}_{i})**	***Wheat vs barley*** **($\alpha^{s}_{(\text{cereals}),i}$)**	
Mean	0.031*	0.061*		0.032*	0.047*	
	(0.002)	(0.005)		(0.001)	(0.002)	
Std dev [b]	0.383*	0.502*		0	0	
	(0.162)	(0.219)				

a. Estimated standard errors of the ML estimator are in parentheses. b. Standard deviation of the log of the random parameter.
Note: Asterisk (*) denotes a statistically significant parameter at the 5% level.

5.7.1. Estimation of the impact of EU biofuel support: simulations with the Sim-RP model

From the French farmers' perspective, EU biofuel support mostly led to an increase in the rapeseed price. We consider a 33% increase in the price of rapeseed. This price increase is roughly the median of the estimated increases in rapeseed price due to EU biofuel support found in the literature (Louhichi and Valin, 2012). Given that our observations cover a period during which EU biofuel support was already in place, its effects on farmers' production choices are estimated by simulating a 25% decrease in the observed rapeseed prices. Due to

space limitations, we focus on the comparison of the simulated effects obtained with the Sim-F/RP and Sim-RP models on the production choices related to rapeseed. Because they outperform the F/RP counterparts from a statistical viewpoint, the RP and Sim-RP models are now considered as the benchmark models. In what follows, we analyze the effects of the effects of EU biofuel support on average from 2006 to 2011 for attenuating the effects of the production choice drivers varying across time.

An increase in the price of rapeseed is expected to increase farmers' rapeseed variable input uses and yield levels. The estimated increases in rapeseed variable input uses due to the EU support to the price of rapeseed range from 0.1% to 8.6% with a mean of 1.3%. The corresponding estimated increases in rapeseed yields range from 0.1% to 2.8% with a mean of 0.6%. The corresponding own-price elasticities of the rapeseed yield supplies are positive with a maximum at 0.11 and a mean of 0.02. Similarly, the rapeseed price elasticities of the rapeseed variable input demands are positive with a maximum at 0.34 and a mean of 0.05. These limited effects at the intensive margin are due to the low estimates of input use flexibility parameters $\alpha^{x}_{\text{rape},i}$ obtained from the RP model.

An increase in the price of rapeseed is expected to increase farmers' rapeseed acreages according to a simple two-step mechanism: the increase in the rapeseed price increases the rapeseed expected return, and this increases the optimal rapeseed acreage. According to the simulation results obtained with the Sim-RP model, EU biofuel support led to an average

increase of 4.7% in farmers' rapeseed acreages over the 2006–2011 period. The estimated increases in the rapeseed acreages significantly depend on the considered farms. They range from 1.9% to 9.8% with a 1.7% standard deviation across farms. The corresponding own-price elasticities of the rapeseed acreage range from 0.03 to 0.30 with a mean of 0.14. They are larger than those obtained with the usual panel data multi-crop models and French data (see, *e.g.*, Lacroix and Thomas 2011; Bayramoglu and Chakir 2016), but they are comparable to the short run own-price elasticities of soybean acreages estimated by Hendricks et al (2014) with US data.

5.7.2. Comparison of the simulations obtained from the RP and F/RP models

The estimated impacts of EU biofuel support obtained from the Sim-F/RP model significantly differ from those obtained from the Sim-RP model. These differences are mostly due to how farmers' responses to economic incentives are modeled in the considered models: with random parameters in the RP model and with fixed parameters in the F/RP model. As illustrated by Figure 2, the estimated effects on rapeseed acreage shares obtained with the Sim-F/RP and Sim-RP models differ significantly. The distribution of the effects estimated with the Sim-F/RP model is slightly upward biased, with a mean of 5.0% and is much less spread out than that of the effects estimated with the Sim-RP model. The standard deviation of the effects estimated by the Sim-F/RP model equals 0.7%, while that of the effects estimated with the Sim-RP simulation model equals 1.7%. Indeed, the Sim-F/RP model systematically overestimates the small increases in rapeseed acreages and underestimates the

large ones, as shown in Figure 3.

Figure 2. Distribution of the effects of EU biofuel support on Rapeseed Acreage Choices

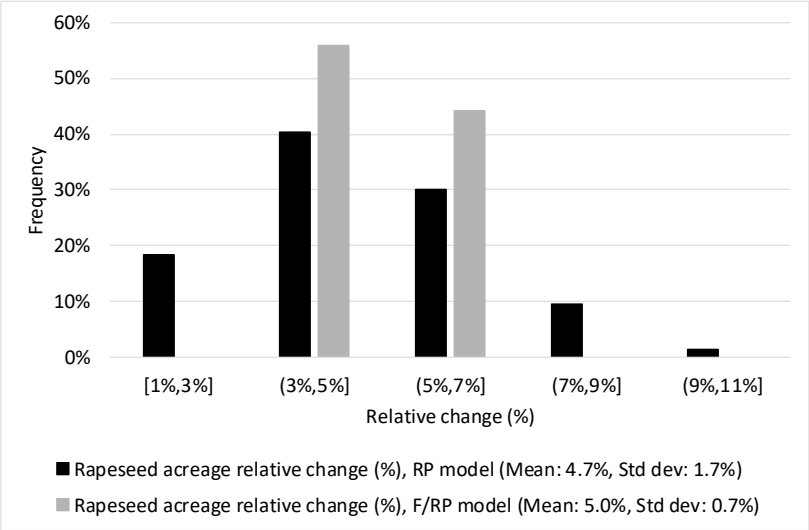
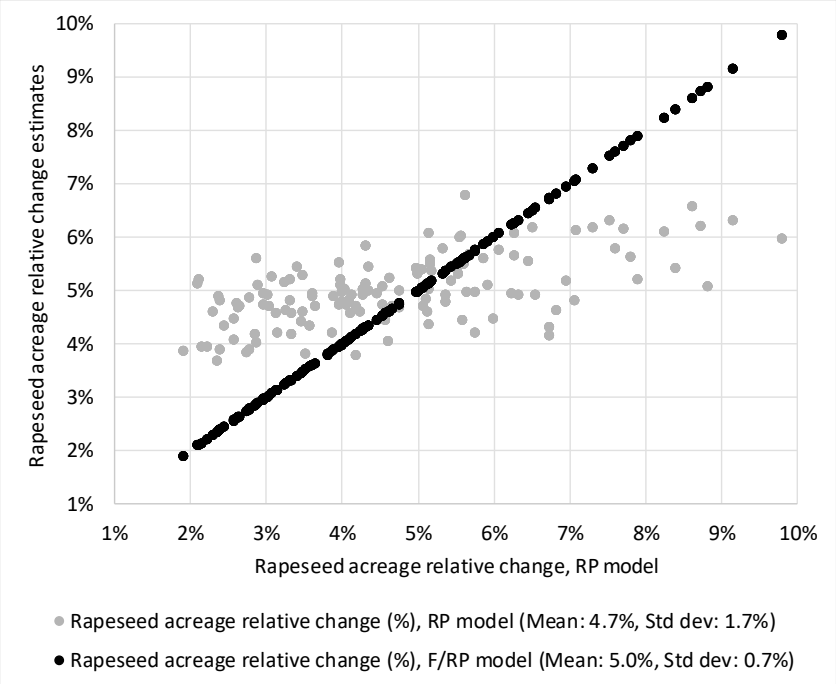


Figure 3. Effects of EU biofuel support on Rapeseed Acreage Choices



The differences in the estimates of the Sim-RP and Sim-F/RP model have a main origin: the variability in the acreage choice flexibility random parameter α_i^s around its mean in the Sim-RP model. According to our estimation results, the between farm heterogeneity in the simulated rapeseed acreage changes mostly depends on the product $\beta_{\text{rape},i}^y \alpha_i^s$: the rapeseed potential yield parameter $\beta_{\text{rape},i}^y$ largely determines the expected rapeseed returns of farmer *i*, while the acreage choice flexibility parameter α_i^s is the main driver of her/his cropland allocation to cereals and to rapeseed.[67] As shown in Table 5, the estimated value of the fixed parameter α_0^s of the F/RP model, 0.032, is almost identical to the estimated value of the mean of the corresponding random parameter α_i^s of the RP model, 0.031. The estimates of the rapeseed potential yield parameter $\beta_{\text{rape},i}^y$ obtained from the RP and F/RP models are very close to each other: their correlation coefficient equals 0.96. Hence, the increases in the rapeseed acreages estimated with the Sim-RP and Sim-F/RP mostly differ because the Sim-RP model allows heterogeneous responses of farmers' acreages to changes in expected returns, while the Sim-F/RP model imposes homogeneous responses.

Figures 4 and 5 display the estimated distributions of the effects of EU biofuel support on rapeseed yields and variable input uses. The Sim-F/RP model tends to significantly

[67] The α_i^s and $\beta_{\text{rape},i}^y$ parameters are negatively correlated. However, this correlation only slightly attenuates the variability of the simulated effects.

overestimate the effects of the considered rapeseed price increase. The effects on the rapeseed yields obtained from the Sim-RP model are mostly concentrated around their mean, *i.e.,* approximately 0.6%, while those obtained from the Sim-F/RP model mostly range from 1% to 4% with a mean of 2.6%. Similarly, the effects on the rapeseed input use levels obtained from the Sim-RP model are mostly concentrated around their mean, *i.e.,* approximately 1.2%, while those obtained from the Sim-F/RP model mostly range from 3% to 6% with a mean of 4.9%.

Figure 4. Effects of EU biofuel support on Rapeseed Variable Input Uses

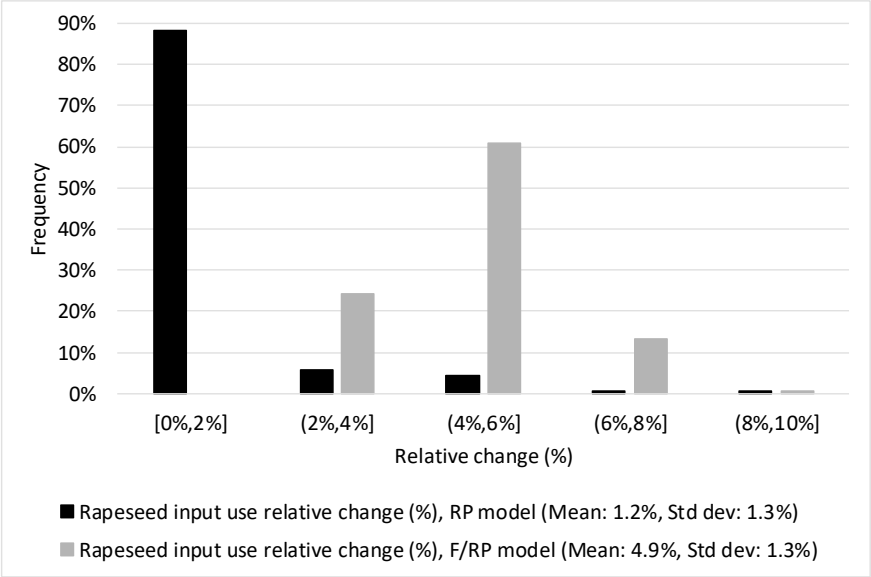
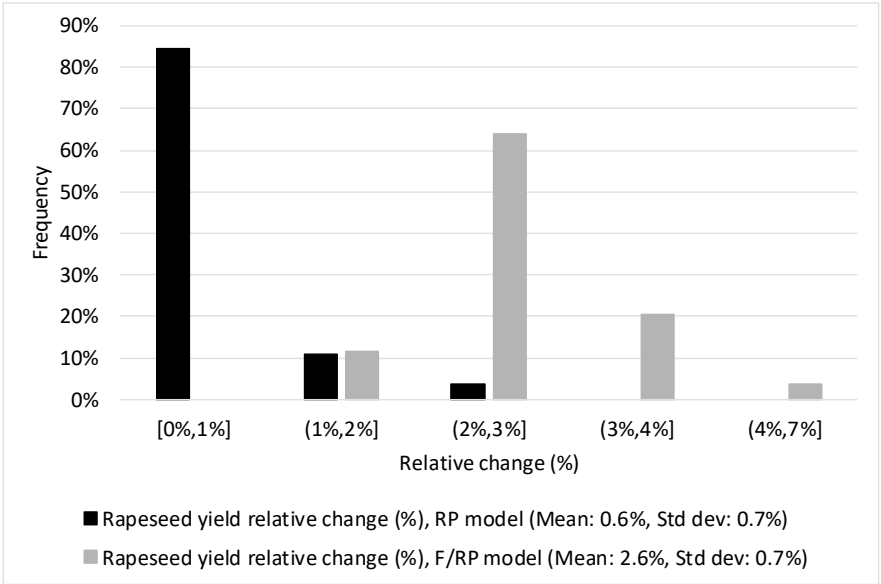


Figure 5. Effects of EU biofuel support on Rapeseed Yields



These significant differences in the estimated effects of EU biofuel support at the intensive margin mostly come from the estimates of the rapeseed input use flexibility parameters considered in the Sim-RP and Sim-F/RP models. The estimate of $\alpha^{x}_{\text{rape},0}$ used in the F/RP model equals 0.47, while the estimates of $\alpha^{x}_{\text{rape},i}$ used in the Sim-RP model mostly range from 0.03 (1st decile) to 0.18 (8th decile). Although they significantly vary across farms according to our estimation results, the estimates of the $\alpha^{x}_{\text{rape},i}$ parameters are small and induce rather homogeneous input use and yield responses to price changes. In this case random slope parameters induce less heterogeneous responses than fixed slope parameters.

5.7.3. Lagged acreage shares as additional control variables

Taken together, our empirical results tend to demonstrate that farmers' production choices strongly depend on unobserved and heterogeneous drivers. Nevertheless, this conclusion must be nuanced because our models are essentially static. We estimated the F/RP and RP models with the (log of the) lagged acreage shares of wheat and of rapeseed as additional control variables in each equation of these models to investigate the potential effects of dynamic features of acreage choices on our estimation results. Lagged acreage shares are exogenous with respect to the RP model error terms $\boldsymbol{\varepsilon}_{it}$ under the considered assumptions. However, these variables depend, by construction, on the random parameters $\boldsymbol{\rho}_i$, implying that these parameters can capture the most persistent features of the modeled process dynamics.

As for the versions "without lagged acreage shares", the homogeneity restriction test obtained with the versions "with lagged acreage shares" strongly rejects the F/RP model against the RP model. This confirms that unobserved heterogeneity matters even with lagged acreage shares as control variables. The lagged acreage shares have statistically significant effects in the RP model, especially in its acreage share equations. The lagged rapeseed acreage share tends to increase the current cereal acreage share, while the lagged wheat acreage share tends to increase the current rapeseed acreage share. These results are consistent with the rapeseed/wheat/barley rotation that is standard in the studied area. However, introducing lagged acreage shares has very limited impacts on the estimates of the random parameter distribution of the RP multi-crop model. This suggests that the dynamic

features of the multi-crop technology and of the production choices of the sampled farmers are highly persistent and largely captured by the random parameters of the RP multi-crop model. Taken together these results suggest that the sampled farmers account for crop rotation effects in their acreage choices, mostly by relying on stable crop rotation schemes but also through small adjustments in the short run. This interpretation is consistent with the findings of Hendricks et al (2014), but it requires further empirical work based on a suitable dynamic setting to be validated.

Concluding remarks

Taking for granted that many unobserved heterogeneous factors can impact farmers' production decisions, our empirical results tend to show that unobserved heterogeneity really matters in micro-econometric multi-crop models. Most key parameters of the RP model describing the production choices of French grain crop producers in the Meuse *département*, *i.e.,* the usual additively separable farm effects and the production choice flexibility parameters, exhibit significant variability across farmers. Our simulation results also show that these farmers have responded heterogeneously to the increase in the rapeseed price induced by EU biofuel support. Omitting this variability in farmers' responses leads (*i*) to significantly overestimating the effects of EU biofuel support on rapeseed input uses and yield levels and (*ii*) to significantly overestimating (underestimate) the small (large) increases in farmers' rapeseed acreages induced by this support.

Importantly, our results show that farmers' economic performances and responsiveness to

economic incentives are significantly heterogeneous even in small geographical areas such as the Meuse *département* and when considering farms that are observationally homogenous as in our sample. Of course, neither our empirical results nor our interpretations can be generalized without further empirical investigations. Nevertheless, they suggest that unobserved heterogeneity in farmers' responses to economic incentives is likely to be a relevant issue for other micro-econometric analyses of agricultural production choices, particularly when considering large geographical areas displaying heterogeneous production conditions that cannot be controlled by observed variables.

From a methodological viewpoint, our study also demonstrates that suitably specified RP multi-crop micro-econometric models allow accounting for unobserved heterogeneity in a fairly flexible way. Importantly, such micro-econometrics models can be used for "statistically calibrating" simulation models composed of sets of heterogeneous farm models such as our Sim-RP multi-crop model. Additionally, according to our experience, ML estimators computed with SAEM algorithms appear to be competitive alternatives to SML estimators, the implementation of which is challenging for large multivariate RP models such as RP multi-crop models.

Of course, this empirical modeling framework can be adapted to micro-econometric multi-crop models other than the one considered here, although it would lead to estimation issues that depend on the considered model. This framework could also be considered, albeit with substantial modifications, for defining estimation procedures for the parameters of micro-economic (positive) mathematical programming models that are often used for agricultural policy analysis.

Limitations of our empirical study and modeling framework call for improvement and further research. Its reliance on distributional assumptions admittedly is a limitation of our RP multi-crop modeling framework. Because avoiding distributional assumptions would either raise challenging estimation issues or imply eliminating unobserved heterogeneity effects from the model, an interesting research topic would be to consider the use of flexible parametric probability distributions, such as multivariate distributions constructed with copulas. Additionally, specifying a full covariance matrix for the random parameters of the RP multi-crop model significantly increases the number of parameters to be estimated: the joint probability distribution of the 13 random parameters of our model is characterized by 78 covariance parameters. The use of latent variables – such as latent productivity indices – for structuring the random parameter distribution can significantly reduce the dimension of the estimation problem.

Accounting for dynamic features of the multi-crop technology and of farmer's choices would obviously significantly improve our multi-crop modeling framework. In particular, the results on the effects of the lagged acreage shares in the RP multi-crop model suggest that crop rotation effects and how they are managed by farmers are important research topics. Preliminary results extending in this direction the theoretical modeling framework considered here suggest that the estimation of dynamic RP multi-crop model is possible, though computationally challenging.

Similarly, corner solutions in acreage choices is a pervasive issue when considering the estimation of multi-crop models. Extending the multi-crop models and the estimation procedures presented in this article for their accounting for corner solutions is possible.

Ongoing research projects on this topic have already yielded promising empirical results (Koutchadé et al, 2017).

Chapitre 6

Dealing with corner solutions in multi-crop micro-econometric models: an endogenous regime switching approach with regime fixed costs.

Introduction

Corner solution problems are pervasive in micro-econometric acreage choice models because farmers rarely produce the same crop set in a considered sample, even in samples considering specialized farms. These problems raise significant modelling issues. The numerous null crop acreages observed in farms' level datasets must be modelled as farmers' choice. Yield and input use level are not observed for the crops not produced by the farmers. As these unobserved netput levels might explain why farmers do not produced some crops, endogenous selection issues arise. Agricultural economists usually use two approaches for coping with null crop acreages. First, crops can be aggregated for eliminating or, at least, attenuating the occurrence of null crop acreages. Of course, this approach can lead to substantial information loss. Second, corner solutions can be dealt with by specifying acreage choices as a system of censored regressions (see, e.g., Platoni et al., 2012). However, if censored regression systems explicitly account for null crop acreages from a statistical viewpoint, they cannot consistently represent acreage choices with corner solutions. Arndt et al. (1999) made this point, for consumer demand systems. More generally, acreage choice models suitably accounting for corner solutions need to be specified as endogenous regime switching (ERS) models, in which production regimes are defined as the subsets of crops with non-null acreages – *i.e*. by the subsets of actually produced crops.

Regime and acreage choice decisions are closely linked since these decisions are taken simultaneously and depend on common drivers. For instance, the choice of the set of crops to be produced depends on the optimal acreages of these crops. Importantly, responses to

crop price changes of crop acreage decisions depend on the regime in which these crops are produced. For instance, winter wheat crop acreages cannot respond to corn price changes in regimes where winter wheat is produced whereas corn is not produced. Censored regression systems cannot account for such effects since in these modelling framework farmers' acreage choices are described by a model that is common to all production regimes. In our ERS modelling framework farmers' acreage choices are described by models that are specific to each production regime.

According to our knowledge, micro-econometric ERS models involving multiple corner solutions were defined only for modelling consumer demand systems (see, e.g., Kao et al. 2001) or firm input demand systems (see, e.g., Chakir et al. 2004), following the pioneering works of Wales and Woodland (1983) and of Lee and Pitt (1986). However, these models have rarely been used in practice, probably because their estimation is challenging, and despite the development of estimation procedures with simulation methods (see, e.g., Kao et al. 2001).

The main aim of our paper is to propose an ERS model specifically designed for empirically modeling acreage choices with corner solutions. This model defines farmers' production choice models as resulting from a profit maximization problem. It is fully coherent from a micro-economic point of view and includes regime fixed costs, which is to our knowledge a unique feature compared to other ERS models with multiple corner solutions found in the economic literature. These regime fixed costs allow accounting for unobserved costs, such as marketing or management costs, which depend on the set of crops grown simultaneously.

The ERS model we propose defines a Nested MultiNomial Logit (NMNL) acreage choice model (Carpentier and Letort, 2014; Koutchade et al., 2015) for each potential production regime. The regime choice is based on a discrete choice model in which farmers choose the subset of crops they produce by comparing the profit levels associated to each regime. The econometric model derived from this framework is theoretically consistent – in its deterministic and in its random parts – and can be combined with yield supply and variable input demand functions. Furthermore, following Koutchade et al. (2015), this model accounts for the unobserved heterogeneity in farmers' behaviors through the specification of random parameters. *I.e.* we assume that most model parameters, specifically those respond to economic incentive, are farmer specific and estimate their distribution across the farmers' population described by our sample. Given that our model is fully parametric, it can be efficiently estimated within a Maximum Likelihood (ML) estimation framework. The structure of the model and the functional form of its likelihood function actually make the version of the Simulated Expectation-Maximisation (SEM) algorithm developed by Delyon et al. (1999) especially suitable for maximizing the sample likelihood function. Importantly, once their probability distribution has been estimated, each farmer specific parameter can be 'statistically calibrated' for simulation purpose.

We illustrate the empirical tractability of our approach by estimating seven crops – and eight production regimes – production choice model for a sample of French arable crop producers. The estimated model is then used to simulate the impacts of a crop price change on acreages and illustrate how accounting for endogenous production regime choices and production regime fixed costs can affect the simulation results.

Our results tend to show that our ERS multi-crop model with regime fixed costs perform well according to the standard fit criteria. They also tend to show that the regime fixed costs significantly matter for explaining the production regime choices and that the decision to produce a crop or not plays a major role in the acreage choice responses to economic incentives. In particular, our simulation results show that the acreage choice responses to price changes exhibit threshold effects due to the regime fixed costs.

The rest of the chapter is organized as follows. The approach proposed to account for endogenous regime switching and regime fixed costs in the modelling of acreage decisions is presented in the first section. The structure of the econometric model of acreage and production choices is described in the second section. Identification and estimation issues are discussed in the third section. The illustrative estimation results are provided in the fourth section. Finally, we conclude.

6.1. Endogenous regime switching acreage choices with regime fixed costs

This section presents the theoretical modelling framework we propose for dealing with corner solutions in micro-econometric acreage choice models. We adopt an ERS approach for the resulting models to be fully consistent from a micro-economic viewpoint. We also allow for regime fixed costs for improving the ability of the resulting models to capture the effects of potentially important drivers of farmers' acreage choices.

Let consider a risk neutral farmer i, who can allocate his fixed cropland to K crops. Let

$\mathcal{K} \equiv \{1,...,K\}$ denote the set of crops available to this farmer and let $\mathcal{R} \equiv \{1,...,R\}$ denote the set of feasible production regimes, *i.e.* the set of crop subsets considered by the farmer when choosing the crops he/she will produce.[68] The term $\mathcal{K}^{+}(r)$ denotes the subset of crops produced in regime *r*, with $\mathcal{K}^{+}(1) = \mathcal{K}$ by convention, while the term $\mathcal{K}^{0}(r)$ denotes subset of crops not produced in regime *r*. Finally, let the term $\mathbf{s} \equiv (s_k : k \in \mathcal{K})$ denote an acreage share vector, with $\mathbf{s} \geq \mathbf{0}$ and $\mathbf{s}'\boldsymbol{\iota} = 1$ where $\boldsymbol{\iota}$ is the unitary column vector with dimension *K*, and let the function $\rho(\mathbf{s})$ define the regime of the acreage share vector **s**.

In period *t* the considered farmer is assumed to solve the following expected profit maximization problem:

$$\max_{\mathbf{s}} \left\{ \mathbf{s}'\boldsymbol{\pi}_{it}^{e} - C_{it}(\mathbf{s}) - D_{it}\left(\rho(\mathbf{s})\right) \text{ s.t. } \mathbf{s} \geq \mathbf{0} \text{ and } \mathbf{s}'\boldsymbol{\iota} = 1 \right\} \tag{1}$$

where $\boldsymbol{\pi}_{it}^{e} \equiv (\pi_{k,it}^{e} : k \in \mathcal{K})$ is the vector of crop returns expected when choosing **s**, $C_{it}(\mathbf{s})$ is the implicit management cost of acreage **s** and $D_{it}(r)$ is the fixed cost of production regime *r*.

The acreage management cost function $C_{it}(\mathbf{s})$ accounts for the crop variable costs not included in the crop gross margins and for the implicit costs related to the constraints on the acreage choices due the limiting quantities of quasi-fixed factors or to agronomic factors.

---

[68] Some combinations of crops are not possible because of agronomic constraints. $\mathcal{R}$ thus does not contain all the possible subsets of $\mathcal{K}$ but the subsets that can potentially be chosen by farmers.

The quasi-fixed factor and agronomic constraints providing motives for diversifying crop acreages, the function $C_{it}(\mathbf{s})$ is assumed to be convex in **s**. In order to ensure that the solution in **s** to problem (1) is unique we strengthen this assumption by assuming that $C_{it}(\mathbf{s})$ is strictly convex in **s**. Such cost functions are used in the Positive Mathematical Programming literature (see, e.g., Heckelei et al., 2012; Howitt, 1995) and in the multi-crop econometric literature (see, e.g., Carpentier and Letort, 2011, 2014). Ignoring the regime fixed costs, the optimal acreage choice is determined by maximizing the sum of the crop expected gross margins $\pi^{e}_{k,it}$ weighted by their acreage shares s_k minus the costs associated to the crop acreage **s**, $C_{it}(\mathbf{s})$. In this model the management costs of the crop acreage prevent farmers to solely produce the most profitable crop.

The regime fixed cost terms $D_{it}(r)$ introduce discrete elements in farmers' acreage choices with $D_{it} \in \{D_{it}(r) : r \in \mathcal{R}\}$. These terms account for the hidden fixed costs incurred by the farmer for any acreage choice in the regime. They include fixed costs related to the marketing process of the crop products or those incurred when purchasing specific variable inputs, when renting specific machines, when seeking crop specific advises, *etc*. These costs do not depend on the chosen acreage in a given regime, they only depend on the crop set defining this regime.

The smooth acreage management cost function $C_{it}(\mathbf{s})$ and the discontinuous regime fixed cost function $D_{it}\left(\rho(\mathbf{s})\right)$ are expected to impact farmers' crop diversification in opposite directions. While limiting quantities of quasi-fixed factors impose constraints fostering crop diversification, regime fixed costs are expected to foster crop specialization. In particular, the

regime fixed costs are expected to be non-decreasing in the number of produced crops.[69]

It appears convenient, as well as intuitively appealing, to decompose farmers' acreage choice problem into two steps, as commonly done in acreage choice models based on censored regression systems (see, e.g., Fezzi and Bateman, 2011; Lacroix and Thomas, 2011; Sckokai and Moro, 2006, 2009). Namely, we distinguish the production regime choice from the acreage choice while assuming that both choices are linked by the effects of common observed or unobserved drivers.

Our modelling framework is based on a standard backward induction approach according to which farmers choose their production regime after examining their expected profit in each possible production regime. First, the acreage choice problem is solved for each potential regime:

(2) $$\max_{\mathbf{s}} \left\{ \mathbf{s}'\boldsymbol{\pi}_{it}^{e} - C_{it}(\mathbf{s}) \text{ s.t. } \mathbf{s} \geq \mathbf{0},\ \mathbf{s}'\boldsymbol{\iota} = 1 \text{ and } s_k = 0 \text{ if } k \in \mathcal{K}^0(r) \right\},$$

yielding regime specific optimal acreage shares:

(3a) $$\mathbf{s}_{it}(r) \equiv \arg\max_{\mathbf{s}} \left\{ \mathbf{s}'\boldsymbol{\pi}_{it}^{e} - C_{it}(\mathbf{s}) \text{ s.t. } \mathbf{s} \geq \mathbf{0},\ \mathbf{s}'\boldsymbol{\iota} = 1 \text{ and } s_k = 0 \text{ if } k \in \mathcal{K}^0(r) \right\}$$

[69] Note however that in specific empirical settings the $D_{it}(r)$ terms may also capture the effects of exogenous factors preventing farmer *i* to produce specific crops, e.g. due to unsuitable soils or to lacking outlets. In the empirical application presented in section 4, such features are unlikely to occur. Our sample covers a limited geographical area and we only consider crops which can be profitably produced in this area.

and regime specific optimal expected profit – regime fixed excluded – levels:

(3b) $$\Pi_{it}(r) \equiv \max_{\mathbf{s}} \left\{ \mathbf{s}'\boldsymbol{\pi}_{it}^{e} - C_{it}(\mathbf{s}) \text{ s.t. } \mathbf{s} \geq \mathbf{0}, \mathbf{s}'\boldsymbol{\iota} = 1 \text{ and } s_k = 0 \text{ if } k \in \mathcal{K}^0(r) \right\} .$$

for $r \in \mathcal{R}$. Second, the production regime r_{it} is determined by comparing the regime specific expected profit levels while accounting for the production regime fixed costs, i.e. r_{it} is defined as the solution in *r* to a simple maximization problem with:

(4) $$r_{it} \equiv \max_{r} \left\{ \Pi_{it}(r) - D_{it}\left(\rho\left(\mathbf{s}_{it}(r)\right)\right) \text{ s.t. } r \in \mathcal{R} \right\} .$$

The obtained optimal regime r_{it} is assumed to be unique as multiple solutions can only occur in knife-edge cases. Of course, the optimal acreage choice $\mathbf{s}_{it}$ and the expected profit level Π_{it} are obtained by combining equation (4) and equations (3), with $\mathbf{s}_{it} = \mathbf{s}_{it}(r_{it})$ and $\Pi_{it} = \Pi_{it}(r_{it})$.

The regime specific acreage choices $\mathbf{s}_{it}(r)$ are derived from optimization problems that only differ from one regime to the other due to nullity constraints. These constraints are sufficient for these acreage choices to respond significantly differently to economic changes. For instance, the regime *r* acreage choice, $\mathbf{s}_{it}(r)$, doesn't respond to changes in the expected returns of the crops not produced in regime *r*. Similarly, the wheat acreage is expected to be more responsive to the price of wheat in farms producing barley than in farms not producing other straw cereals. Acreage choice models based on censored regression systems cannot reproduce such patterns.

Note that the regime fixed cost considered in the maximization problem (4) determining the optimal regime r_{it} is not simply $D_{it}(r)$ but $D_{it}\left(\rho\left(\mathbf{s}_{it}(r)\right)\right)$. The reason is that the acreage $\mathbf{s}_{it}(r)$ may not belong to regime *r*, depending on the functional form chosen for the cost function $C_{it}(\mathbf{s})$. This acreage is only guaranteed to belong to a regime 'included' in regime *r* in the sense that some crops of $\mathcal{K}^{+}(r)$ may not be produced in the regime $\rho\left(\mathbf{s}_{it}(r)\right)$.

The Multinomial Logit (MNL) modelling framework proposed by Carpentier and Letort, (2014) is especially convenient in this context. It is based on functional forms of the acreage management cost function ensuring that the regime specific acreage share $\mathbf{s}_{it}(r)$ and expected profit $\Pi_{it}(r)$ are obtained in analytical closed forms and that are smooth in their parameters. For instance, if we assume that the functional form of the acreage management cost function is given by the 'entropic' function:

(5a) $$C_{it}(\mathbf{s}) = \sum\nolimits_{k\in\mathcal{K}^{+}(r)} s_k \beta^{s}_{k,it} + \alpha_i^{-1} \sum\nolimits_{k\in\mathcal{K}^{+}(r)} s_k \ln s_k \quad \text{with} \quad \alpha_i > 0$$

then the regime specific acreage share vectors $\mathbf{s}_{it}(r)$ are given by:

(5b) $$s_{k,it}(r) = j_k(r) \frac{\exp\left(\alpha_i(\pi^{e}_{k,it} - \beta^{s}_{k,it})\right)}{\sum\nolimits_{\ell\in\mathcal{K}^{+}(r)} \exp\left(\alpha_i(\pi^{e}_{\ell,it} - \beta^{s}_{\ell,it})\right)} \quad \text{with} \begin{cases} j_k(r) = 1 \ \text{ if } \ k \in \mathcal{K}^{+}(r) \\ j_k(r) = 0 \ \text{ if } \ k \in \mathcal{K}^{0}(r) \end{cases}$$

while the corresponding expected profit levels $\Pi_{it}(r)$ are given by:

(5c) $$\Pi_{it}(r) = \alpha_i^{-1} \ln \sum\nolimits_{\ell\in\mathcal{K}^{+}(r)} \exp\left(\alpha_i(\pi^{e}_{\ell,it} - \beta^{s}_{\ell,it})\right) .$$

These specific properties of the MNL modelling framework significantly simplifies the

specification of the acreage choice models featuring corner solutions. They basically imply that the production regime choice can be defined as a standard discrete choice, *i.e.* that of the most profitable production regime among a predetermined regime set.

Indeed, the optimal acreage share of crop *k* in regime *r*, *i.e.* $s_{k,it}(r)$, is ensured to be strictly positive if crop *k* belongs to regime *r*, *i.e.* if $k \in \mathcal{K}^{+}(r)$, as shown by equation (5b) in the case standard MNL acreage share choice models (ensuring that $\mathbf{s}_{it}(r)$ necessarily belong to regime *r*). Of course, $s_{k,it}(r)$ is almost null when crop *k* is much less profitable than the other crops of the considered regime.[70] This implies that corner solutions are handled in a specific way in the MNL modelling framework: their characterization doesn't rely on the qualification conditions related to the acreage non-negativity constraints which would be involved in the case where the cost function $C_{it}(\mathbf{s})$ was chosen to be quadratic in **s**. The acreage share non-negativity constraints never bind in the MNL framework, they just imply that the optimal acreage shares of the least profitable crops of a given crop set are very small.[71] The acreage shares of the least profitable crops only become null at the production regime choice stage, when these crops are excluded from the produced crop set characterizing the chosen

---

[70] The term $s_{k,it}(r)$ decreases in $\pi^{e}_{k,it} - \beta^{s}_{k,it}$ and tends to 0 as the term $\pi^{e}_{k,it} - \beta^{s}_{k,it}$ goes to $-\infty$. The larger α_i is, the more rapidly $s_{k,it}(r)$ decreases in $\pi^{e}_{k,it} - \beta^{s}_{k,it}$.

[71] From a technical viewpoint, this property comes from properties of the entropy function: the term $-s_k \ln s_k$ tends to 0 as s_k decreases to 0 (leading to the continuity extension at 0 of these terms and the related convention stating that $s_k \ln s_k = 0$ if $s_k = 0$) while its derivative in s_k tends to $+\infty$ as s_k decreases to 0.

production regime.[72]

6.2. An ERS micro-econometric multi-crop model with regime fixed costs

This section presents the structure of the ERS micro-econometric multi-crop model considered in the empirical application presented in the next section. This model is composed of three equation subsystems describing the yield supply functions, variable input demand functions and the acreage share choice models of each produced crop on the one hand, and of a probabilistic production regime choice model on the other hand. This micro-econometric multi-crop model can be interpreted as an extension, to an ERS framework with regime fixed costs, of the model proposed by Carpentier and Letort (2014). As in Koutchade et al. (2015) we adopt a random parameter approach for accounting for farmers' unobserved

[72] Indeed, if the acreage management cost function $C_{it}(\mathbf{s})$ were chosen to be quadratic in $\mathbf{s}$ – as in the usual PMP framework or as in the econometric acreage choice model of Guyomard et al. (1996), of Moore and Negri (1992) or of Carpentier and Letort (2011) – farmers' acreage choice problem would be defined as quadratic programming problem. It would then be possible to adapt the approaches developed by Wales and Woodland (1980) or by Lee and Pitt (1986) for modelling acreage choices with corner solutions. Following the primal approach of Wales and Woodland (1980), one would define empirically tractable estimating equations – for recovering the parameters of the cost function – based on the first order conditions of quadratic acreage choice problem, including the qualification conditions related to the acreage non-negativity constraints. However, the resulting modelling framework would ignore production regime fixed costs. To account for regime costs would raise significant difficulties as the per regime optimal expected profit functions could only be obtained numerically and would be characterized by salient discontinuities in the parameters to be estimated.

heterogeneity.

The considered ERS micro-econometric multi-crop model is presented in three steps. First, we present the production choice models defined at the crop level, *i.e.* the yield supply and variable input demand models. Second, we present the acreage share choice models. Finally, we describe the production regime choice model. This presentation is organized according to the structure of the model: the crop level production choice models are used for defining the acreage share choice models, which are themselves, used for defining the production regime choice model.

6.2.1. Yield supply and variable input demand models

We assume that farmers produce crop *k* from a variable input aggregate under a quadratic technological constraint. *I.e.*, we assume that the yield of crop *k* obtained by farmer *i* in year *t* is given by:

(6a) $$y_{k,it} = \beta^{y}_{k,it} - 1/2 \times \gamma^{-1}_{k,i}(\beta^{x}_{k,it} - x_{k,it})^2$$

where $x_{k,it}$ denotes the variable input use level. The $\gamma_{k,i}$ parameter is required to be (strictly) positive for the production function to be (strictly) concave in $x_{k,it}$. It determines the extent to which the yield supply and the input demand of crop *k* respond to the input and crop prices. The terms $\beta^{y}_{k,it}$ and $\beta^{x}_{k,it}$ have direct interpretations in the considered yield function. The term $\beta^{y}_{k,it}$ is the yield level that can be potentially achieved by farmer *i* in year *t* while $\beta^{x}_{k,it}$ is the

input quantity required to achieve this potential yield level.

These parameters are decomposed as:

(6b) $$\beta_{k,it}^{y} \equiv \beta_{k,i}^{y} + (\mathbf{a}_{k,0}^{y})'\mathbf{z}_{k,it}^{y} + \varepsilon_{k,it}^{y} \text{ and } \beta_{k,it}^{x} \equiv \beta_{k,i}^{x} + (\mathbf{a}_{k,0}^{x})'\mathbf{z}_{k,it}^{x} + \varepsilon_{k,it}^{x}$$

where the terms $\mathbf{z}_{k,it}^{y}$ and $\mathbf{z}_{k,it}^{x}$ are observed variable vector used to control for farm heterogeneity. The $\beta_{k,i}^{y}$ and $\beta_{k,i}^{x}$ terms are farmer specific parameters aimed at capturing unobserved heterogeneity across farms and farmers. These terms, as well as the $\gamma_{k,i}$ random parameter, mainly capture two kinds of effects: those of the natural and material factor endowment of farms (e.g. soil quality and machinery quality) and those of the skills of farmers. The $\varepsilon_{k,it}^{y}$ and $\varepsilon_{k,it}^{x}$ terms are standard error terms aimed at capturing the effects on production of stochastic events (e.g. climatic conditions, and pest and weed problems). We assume that farmer *i* is aware of the content of $\varepsilon_{k,it}^{x}$ when deciding his variable input uses.

Assuming that farmer *i* maximizes the expected return to variable input uses of each crop, we can easily derive the demand of the variable input for crop *k*:

(7a) $$y_{k,it} = \beta_{k,i}^{y} + (\mathbf{a}_{k,0}^{y})'\mathbf{z}_{k,it}^{y} - 1/2 \times \gamma_{k,i} w_{k,it}^{2} p_{k,it}^{-2} + \varepsilon_{k,it}^{y}$$

and the corresponding yield supply:

(7b) $$x_{k,it} = \beta_{k,i}^{x} + (\mathbf{a}_{k,0}^{x})'\mathbf{z}_{k,it}^{x} - \gamma_{k,i} w_{k,it} p_{k,it}^{-1} + \varepsilon_{k,it}^{x} \ .$$

The terms $p_{k,it}$ and $w_{k,it}$ respectively denote the expected output and input prices of crop

k. Assuming that the expectations of the terms $\varepsilon_{k,it}^{y}$ and $\varepsilon_{k,it}^{x}$ of farmer i are null at the beginning of the cropping season, this farmer expects the following return to the variable input:

(8) $$\pi_{k,it}^{e} = p_{k,it}\left(\beta_{k,i}^{y} + (\mathbf{a}_{k,0}^{y})'\mathbf{z}_{k,it}^{y}\right) - w_{k,it}\left(\beta_{k,i}^{x} + (\mathbf{a}_{k,0}^{x})'\mathbf{z}_{k,it}^{x}\right) + 1/2 \times \gamma_{k,i} w_{k,it}^{2} p_{k,it}^{-1}$$

for crop k when she/he chooses her/his acreage shares.

6.2.2. Acreage share choice models

As discussed in Carpentier and Letort (2014), the (Standard MNL) acreage share choice model given in equation (5b) appears to be rather rigid because it treats the different crops symmetrically. Indeed, arable crops can often be grouped according to their competing for the use of quasi-fixed factors or according to their agronomic characteristics. The ERS micro-econometric multi-crop model considered here contains a '3 level-Nested Multinomial Logit" (NMNL) acreage share choice model, which derives from an entropic acreage management cost function as proposed by Carpentier and Letort (2014). In our setting, the crop set $\mathcal{K}$ is partitioned into G mutually exclusive groups of crops, each group $g \in \{1,...,G\}$ being itself partitioned into $M(g)$ subgroups of crops. The mth subgroup of the gth group is defined as the crop subset $\mathcal{K}(m,g)$. Crops (resp. subgroups) belonging to a same subgroup (resp. group) are assumed to have similar agronomic characteristics and to compete more for farmers' limiting quantities of quasi-fixed factors than they compete with crops (resp. subgroups) of other subgroups (resp. of other groups). The three level nested structure of the

crop set used in the empirical application is depicted in Figure 4.1. The corresponding acreage management cost function is given by:

$$C_{it}(\mathbf{s}) = \sum_{k\in\mathcal{K}} s_k \beta^s_{k,it} + \sum_{g=1}^{G} \alpha_i^{-1} s_{(g)} \ln s_{(g)} + \sum_{g=1}^{G} s_{(g)} \alpha^{-1}_{(g),i} \sum_{m=1}^{M(g)} s_{m|(g)} \ln s_{m|(g)} + \sum_{g=1}^{G} s_{(g)} \sum_{m=1}^{M(g)} s_{m|(g)} \alpha^{-1}_{m|(g),i} \sum_{\ell\in\mathcal{K}(m,g)} s_{\ell|(m,g)} \ln s_{\ell|(m,g)} \tag{9}$$

where $s_{(g)}$ denotes the acreage share of group *g*, $s_{m,(g)}$ that of subgroup *m* in group *g*, and $s_{k|(m,g)}$ that of crop *k* in the subgroup *m* of group *g*. The α_i, $\alpha_{(g),i}$ and $\alpha_{m|(g),i}$ are farm specific parameters determining the flexibility of farmers' acreage choices.[73] The larger they are, the more the acreage share choice respond to economic incentives (because the less management costs matter). The condition $\alpha_{m|(g),i} \geq \alpha_{(g),i} \geq \alpha_i > 0$ is sufficient for the cost function $C_{it}(\mathbf{s})$ to be strictly convex in **s**.

The linear terms of the cost function $C_{it}(\mathbf{s})$ are decomposed as:

$$\beta^s_{k,it} \equiv \beta^s_{k,i} + (\mathbf{a}^s_{k,0})'\mathbf{z}^s_{k,it} + \varepsilon^s_{k,it} \tag{10}$$

where $\mathbf{z}^s_{k,it}$ are explanatory variable used to control for observed heterogeneous factors. The $\beta^s_{k,i}$ farm specific factors account for heterogeneity effects unobserved in the data. The error terms $\varepsilon^s_{k,it}$ capture the effects of stochastic variation of the cost due to random events such

[73] We have $\alpha_{g,m,i} = \alpha_{g,i}$ if subgroup *m* contains a single crop. Similarly, we have $\alpha_{g,i} = \alpha_i$ if group *g* contains a single subgroup.

as climatic events impacting the soil state at planting. The content of these terms are assumed to be known to farmers when they choose their acreages.

These error terms are assumed to be independent from the error terms of the yield and input demand functions, $\varepsilon^{y}_{k,it}$ and $\varepsilon^{x}_{k,it}$. *I.e.*, we assume that the potential links between the error terms of the acreage choice model on the one hand and those of the yield supply and input demand functions on the other hand are negligible. To relax this assumption is possible but significantly increases the estimation burden. In a similar context, Koutchade et al (2015) found that the $\varepsilon^{s}_{k,it}$ error terms were not significantly correlated with the error terms $\varepsilon^{y}_{\ell,it}$ and $\varepsilon^{x}_{\ell,it}$.

The optimal acreage share choices of farmers as given by equation (2a) can be derived for any production regime. It suffices to solve the maximization problem given in equations (3). For instance, ten acreage share subsystems are considered in our empirical application, one for each production regime represented in the data. Of course, the functional form of the derived acreage choice function depends on the subset of crops included in the considered regime. Assuming that crop k belongs to the mth subgroup of the gth group, we obtain:

$$\text{(11a)}\quad s_{k,it}(r) = j_k(r)\frac{\exp(\alpha_{g,m,i}\pi_{k,it})\left(\theta_{(m,g),it}(r)\right)^{\alpha_{g,i}\alpha^{-1}_{g,m,i}-1}\left\{\sum_{n=1}^{M(g)}\left(\theta_{(n,g),it}(r)\right)^{\alpha_{(g),i}\alpha^{-1}_{n(g),i}}\right\}^{\alpha_i\alpha^{-1}_{g,i}-1}}{\sum_{h=1}^{G}\left\{\sum_{n=1}^{M(h)}\left(\theta_{(n,h),it}(r)\right)^{\alpha_{(h),i}\alpha^{-1}_{n(h),i}}\right\}^{\alpha_i\alpha^{-1}_{h,i}}}$$

and:

$$\text{(11b)}\qquad \Pi_{it}(r) = \alpha_i^{-1}\ln\sum_{h=1}^{G}\left\{\sum_{n=1}^{M(h)}\left(\theta_{(n,h),it}(r)\right)^{\alpha_{(h),i}\alpha^{-1}_{n(h),i}}\right\}^{\alpha_i\alpha^{-1}_{(h),i}}$$

where:

(11c) $$\theta_{(n,h),it}(r) = \sum_{\ell \in \mathcal{K}(n,h)} j_\ell(r) \exp\left(\alpha_{n|(h),i}(\pi_{\ell,it} - \beta^s_{\ell,it})\right) .$$

6.2.3. Production regime choice model

Observing that the regime *r* optimal acreage choice $\mathbf{s}_{it}(r)$ necessarily belongs to regime *r* in the MNL case considered here, the regime specific expected profit levels can easily be used for defining a regime choice model according to the choice problem described in equation (4). Let define the regime fixed costs as $D_{it}(r) = d_i(r) - \sigma_i^{-1}\varepsilon^r_{it}$. The farm specific parameters $d_i(r)$ aim at capturing the effects of unobserved factors affecting the regime fixed costs. The error terms ε^r_{it} aim at capturing the effects of stochastic factors. They are assumed to be independent from the other elements of the model and to be independently distributed according to a type I extreme value distribution. This implies that the production regime choice is specified as a standard Multinomial Logit discrete choice model.

The probability of the choice of regime *r* – conditionally on σ_i, and on the $d_i(r)$ and $\Pi_{it}(r)$ terms for $r \in \mathcal{R}$ – is thus given by:

(12) $$P[r_{it}=r]=\frac{\exp\left(\sigma_i(\Pi_{it}(r)-d_i(r))\right)}{\sum_{q\in\mathcal{R}}\exp\left(\sigma_i(\Pi_{it}(q)-d_i(q))\right)}\ .$$ [74]

The farm specific parameter σ_i is assumed to be positive. This scale parameter allows determining the extent to which the regime specific expected profit levels minus the corresponding fixed cost, *i.e.* the $\Pi_{it}(r)-d_i(r)$ terms, explain the production regime choice as regards to the effects of the ε_{it}^r idiosyncratic terms. The higher σ_i, the more the 'deterministic' terms $\Pi_{it}(r)-d_i(r)$ impact the observed regime choices.

The regime fixed costs $d_i(r)$ can be specified in different ways. These costs are expected to increase with the number of crops, since transaction costs and labor requirements related to a production regime increase with the number of crops produced in that regime. It should however be noted that the costs associated to a regime do not necessarily correspond to a sum of fixed costs associated to each crop produced in the regime. The regime fixed costs are actually likely to be sub-additive in the individual crop fixed costs. *E.g.* farmers may purchase

---

[74] Note however that the error terms $\varepsilon_{k,it}^s$ of the acreage choice model contained in the expected profit levels $\Pi_{it}(r)$ can only be directly recovered from the data for the crops produced by farmer *i* in year *t*. Indeed, we can recover the vector $\boldsymbol{\varepsilon}_{r,it}^{s,+}=(\varepsilon_{k,it}^s : k\in\mathcal{K}^+(r))$ while the vector $\boldsymbol{\varepsilon}_{r,it}^{s,0}=(\varepsilon_{k,it}^s : k\in\mathcal{K}^0(r))$ cannot. We used Laplace approximations for integrating the expectation of the probability function $P[r_{it}=r]$ over the probability distribution of $\boldsymbol{\varepsilon}_{k,it}^{s,0}$ conditional on $\boldsymbol{\varepsilon}_{k,it}^{s,+}$ (see, e.g., Harding and Hausman, 2007). This expectation is part of the likelihood function of the model.

inputs specific to different crops from the same supplier, implying savings in the related transaction costs[75]. One way to specify $d_i(r)$ is thus to consider as a constant term by regime, *i.e.* $d_i(r) = d_i^r$. This specification allows the identification of regime fixed costs associated to each regime observed in the data, but does not permit to recover the costs corresponding to non-produced regimes. Yet this can be an issue for simulations based on the estimated model, if one seeks, for instance, to study the impacts of changes in market conditions on the adoption of new production regimes. To overcome this issue, another way to specify $d_i(r)$ is to consider a sum of fixed costs associated to each crop produced in the production regime. More precisely, if there exist a regime $r(0)$ that is contained in all regimes, *i.e.* all farmers in the sample produce at least the crops included in $r(0)$, this regime is defined as a reference regime with null cost. The fixed costs corresponding to the other production regimes are then specified as the sum of fixed costs associated to each crop produced in the considered regime but not in the reference regime: $d(r) \equiv \sum_{k \in \mathcal{K}^+(r) \setminus \mathcal{K}^+(0)} d_{k,i}^c$. The term $d_{k,i}^c$ represents the fixed costs associated to each crop not produced in the reference regime. This last specification offers the possibility to recover the regime fixed costs of non-produced regimes that contain at least the reference regime and has been used in our empirical application, section 4.

[75] Similarly, different crops may generate work peak loads at the same dates. These peak loads lead to increase the acreage management costs represented by $C(\mathbf{s})$. However, theyonly generate one single fixed cost, since the farmer must be on his farm at these dates, whether this is due to a single crop or to several crops.

6.3. Parametric specification and estimation procedure

The ERS multi-crop micro-econometric model presented in section 2 is composed of three main parts: a subsystem of yield supply and input demand equations (equations 7), subsystems of acreage share equations (equation 11a) and a probabilistic production regime choice model (equation 12). In this section, we briefly present the econometric procedure used to estimate this model.[76]

6.3.1. Distributional assumptions

Most parameters of the model are farmer specific, which allows accounting for the heterogeneity in the performance levels as well as in the responses to economic incentives of the sampled farmers. Yet, standard data set, even panel data sets, do not permit a direct estimation of each individual parameter: the objective of the estimation here is to characterize the distribution of these parameters in the population described by our sample. To do so, we rely on a random parameter approach, as proposed in Koutchade et al. (2015).

Given the rather complex structure and the size of our model, we adopt a fully parametric

[76] Specification and estimation details are available in supplementary data.

framework. Apart from the modelled variables, *i.e.* the crop yield levels $y_{k,it}$, the crop input use levels $x_{k,it}$, the crop acreage shares $s_{k,it}$ for $k \in \mathcal{K}$, and the production regimes the production regimes r_{it} for $r \in \mathcal{R}$ collected in the vector $\mathbf{c}_{it}$, and the fixed parameters, *i.e.* the terms $\mathbf{a}_{k,0}^{y}$, $\mathbf{a}_{k,0}^{y}$ and $\mathbf{a}_{k,0}^{s}$ for $k \in \mathcal{K}$ collected in the vector $\mathbf{a}_0$, the considered model contains five subsets of random elements:

The farm specific parameter vectors $\boldsymbol{\delta}_i$ collecting the potential yield parameters $\beta_{k,i}^{y}$, the input requirement parameters $\beta_{k,i}^{x}$, the input use flexibility parameters $\ln \gamma_{k,i}$ and the cost function linear parameters $\beta_{k,i}^{s}$ for $k \in \mathcal{K}$; the acreage choice flexibility parameters $\ln \alpha_i$, $\ln \alpha_{(g),i}$ and $\ln \alpha_{m|(g),i}$ for $m = 1,...,M(g)$ and $g = 1,...,G$; the regime fixed cost parameters d_i^r (or $d_{k,i}^c$ depending on the specification) for $r \in \mathcal{R}$ and the regime choice scale parameter $\ln \sigma_i$. The vectors $\boldsymbol{\delta}_i$ are assumed normally and independently distributed across farms.

The explanatory variable vectors $\mathbf{z}_{it}$ containing the crop prices $p_{k,it}$, the variable input prices $w_{k,it}$ and the control variable vector $\mathbf{z}_{k,it}^{y}$, $\mathbf{z}_{k,it}^{x}$ and $\mathbf{z}_{k,it}^{s}$ for $k \in \mathcal{K}$.

The yield supply and input demand error term vectors $\boldsymbol{\varepsilon}_{it}^{yx}$ containing the error terms $\varepsilon_{k,it}^{y}$ and $\varepsilon_{k,it}^{x}$ for $k \in \mathcal{K}$ and that are assumed normally and independently distributed across farms.

The acreage share error term vectors $\boldsymbol{\varepsilon}_{it}^{s}$ containing the error terms $\varepsilon_{k,it}^{s}$ for $k \in \mathcal{K}$ and that

are assumed normally and independently distributed across farms.

The production regime error term vectors $\boldsymbol{\varepsilon}_{it}^{\rho}$ containing the error terms ε_{it}^{r} for $r \in \mathcal{R}$. These terms ε_{it}^{r} are assumed independent across regimes and farms, well as distributed according to a type I extreme value distribution.

We further assume that the error term vectors $\boldsymbol{\varepsilon}_{it}^{yx}$, $\boldsymbol{\varepsilon}_{it}^{s}$ and $\boldsymbol{\varepsilon}_{it}^{\rho}$ are mutually independent, and that the explanatory variables $\mathbf{z}_{it}$ are (*i*) strictly exogenous with respect to these error term vectors and (*ii*) independent of the random parameters $\boldsymbol{\delta}_i$. The vector $\mathbf{z}_{it}$ contains prices, climatic variables and characteristics of the farms' fixed factor endowments. We finally assume that that the error term vectors $\boldsymbol{\varepsilon}_{it}^{yx}$, $\boldsymbol{\varepsilon}_{it}^{s}$ and $\boldsymbol{\varepsilon}_{it}^{\rho}$ are independent across years.

As the explanatory variable vector $\mathbf{z}_{it}$ doesn't contain any lagged endogenous variable, the considered model can be interpreted either as an essentially static model or as a reduced form model as regards the dynamic features of the modelled choices. It is notably difficult to empirically disentangle the effects of farmers' unobserved heterogeneity from those of unobserved persistent dynamic features of the modelled processes (see, e.g., Angrist and Pischke, 2009 or Arellano and Bonhomme, 2011). For instance, the random parameters $\boldsymbol{\delta}_i$ are likely to capture the effects on farmers' production choices and performances of the stable crop rotation schemes that these farmers' seem to rely on.[77] Our assuming that the error term

[77] See, *e.g.*, Koutchadé et al (2015) for an empirical analysis providing arguments confirming this hypothesis.

vectors $\boldsymbol{\varepsilon}_{it}^{yx}$, $\boldsymbol{\varepsilon}_{it}^{s}$ and $\boldsymbol{\varepsilon}_{it}^{\rho}$ are serially independent across years is mostly based on this hypothesis. We do not assume that farmers' choices and performances are not significantly impacted by unobserved dynamic features. But we assume that these dynamic features are sufficiently persistent for being mostly captured by the random parameters of our model. Of course, dynamic features of crop production and of farmers' choice are important topics. But these are also difficult ones. Their empirical investigation with models involving corner solutions in acreage choices and regime fixed costs is left for further research.

6.3.2. Estimation

The aim of the estimation procedure is to obtain statistical estimates of two parameter sets: the fixed parameters and the elements of variance matrices of the error term vectors of the model – collected in the vector $\boldsymbol{\theta}_0$ – on the one hand, and the parameters of the probability distribution of its random parameter vector – collected in the vector $\boldsymbol{\eta}_0$. We use a Maximum Likelihood (ML) estimator computed via an Expectation Maximization (EM) algorithm (Dempster et al., 1977) for estimating $(\boldsymbol{\theta}_0, \boldsymbol{\eta}_0)$.

Let the function $g(\mathbf{u} \mid \mathbf{v}; \boldsymbol{\lambda})$ generically define the probability distribution function of the random vector $\mathbf{u}$ conditional on the random vector $\mathbf{v}$ parameterized by $\boldsymbol{\lambda}$. The likelihood function at $(\boldsymbol{\theta}, \boldsymbol{\eta})$ of $\mathbf{c}_i$ conditional on $\mathbf{z}_i$ is given by $\ell_i(\boldsymbol{\theta}) = \int g(\mathbf{c}_i \mid \mathbf{z}_i, \boldsymbol{\delta}_i; \boldsymbol{\theta}) g(\boldsymbol{\delta}_i; \boldsymbol{\eta}) d\boldsymbol{\delta}_i$. The assumed serial independence of the model error terms imply that:

(13) $$g(\mathbf{c}_i \mid \mathbf{z}_i, \boldsymbol{\delta}_i; \boldsymbol{\theta}) = \prod_{t \in \mathcal{H}(i)} g(\mathbf{c}_{it} \mid \mathbf{z}_{it}, \boldsymbol{\delta}_i; \boldsymbol{\theta})$$

where $\mathcal{H}(i)$ is the observation history of farm i. Let denote $\mathbf{y}_{it}^{+} \equiv (y_{k,it} : k \in \mathcal{K}_{it}^{+})$, $\mathbf{x}_{it}^{+} \equiv (x_{k,it} : k \in \mathcal{K}_{it}^{+})$, $\mathbf{s}_{it}^{+} \equiv (s_{k,it} : k \in \mathcal{K}_{it}^{+})$ and r_{it} the observed choices of farmer i at time t. The probability density function of the observed data $\mathbf{c}_{it} \equiv (\mathbf{y}_{it}^{+}, \mathbf{x}_{it}^{+}, \mathbf{s}_{it}^{+}, r_{it})$ for farmer i at time t conditional on $\mathbf{z}_{it}$ and $\boldsymbol{\delta}_i$ is :[78]

(14) $$\begin{aligned} g(\mathbf{c}_{it} \mid \mathbf{z}_{it}, \boldsymbol{\delta}_i; \boldsymbol{\theta}) &= g(\mathbf{y}_{it}^{+}, \mathbf{x}_{it}^{+}, \mathbf{s}_{it}^{+}, r_{it} \mid \mathbf{z}_{it}, \boldsymbol{\delta}_i; \boldsymbol{\theta}) \\ &= g(\mathbf{y}_{it}^{+}, \mathbf{x}_{it}^{+} \mid \mathbf{z}_{it}, \boldsymbol{\delta}_i, \mathbf{s}_{it}^{+}, r_{it}; \boldsymbol{\theta}) \times g(r_{it} \mid \mathbf{z}_{it}, \boldsymbol{\delta}_i, \mathbf{s}_{it}^{+}; \boldsymbol{\theta}) \times g(\mathbf{s}_{it}^{+} \mid \mathbf{z}_{it}, \boldsymbol{\delta}_i; \boldsymbol{\theta}) \\ &= g(\mathbf{y}_{it}^{+}, \mathbf{x}_{it}^{+} \mid \mathbf{z}_{it}, \boldsymbol{\delta}_i; \boldsymbol{\theta}) \times P(r_{it} \mid \mathbf{z}_{it}, \boldsymbol{\delta}_i, \mathbf{s}_{it}^{+}; \boldsymbol{\theta}) \times g(\mathbf{s}_{it}^{+} \mid \mathbf{z}_{it}, \boldsymbol{\delta}_i; \boldsymbol{\theta}) \\ &= g(\mathbf{y}_{it}^{+}, \mathbf{x}_{it}^{+} \mid \mathbf{z}_{it}, \boldsymbol{\delta}_i; \boldsymbol{\theta}_{r_{it}}) \times P(r_{it} \mid \mathbf{z}_{it}, \boldsymbol{\delta}_i, \boldsymbol{\varepsilon}_{it}^{s,+}; \boldsymbol{\theta}) \times g(\mathbf{s}_{it}^{+} \mid \mathbf{z}_{it}, \boldsymbol{\delta}_i; \boldsymbol{\theta}_{r_{it}}). \end{aligned}$$

The likelihood function $\ell_i(\boldsymbol{\theta})$ can be obtained neither analytically nor numerically. But it can be estimated via simulation methods for computing Simulated ML estimators of $(\boldsymbol{\theta}_0, \boldsymbol{\eta}_0)$. Albeit possible, the empirical implementation of this approach is difficult due to the dimension of our parameter of interest and to the rather complex functional form of the likelihood function $\ell_i(\boldsymbol{\theta})$.[79]We choose to compute our ML estimator via an EM algorithm

---

[78] This decomposition is used to rely on the functional forms of $\mathbf{y}_{it}^{+}$, $\mathbf{x}_{it}^{+}$, $\mathbf{s}_{it}^{+}$ and r_{it}. It allows accounting for the dependence between all random variables, which is not always the case in likelihood functions of regime switching model.

[79] In particular, the probability distribution functions $g(\mathbf{c}_{it} \mid \mathbf{z}_{it}, \boldsymbol{\delta}_i; \boldsymbol{\theta})$ cannot be obtained in analytical closed form. These functions contain the probability functions at $\boldsymbol{\theta}$, denoted by, of the regime choice of farmer i in year t (r_{it}), conditional on the random parameters ($\boldsymbol{\delta}_i$), explanatory variables ($\mathbf{z}_{it}$) and on the acreage shares

because EM algorithms are particularly well suited for estimating models with missing variables such as random parameters. The aim of these algorithms replace a large ML maximization problem by a sequence of simpler problems. EM algorithms iterate an Expectation (E) step and a Maximization (M) step until numerical convergence. In our case, we consider $\boldsymbol{\delta}_i$ as missing data. The E step consisting of computing the expectation of the probability distribution functions $g(\mathbf{c}_i \mid \mathbf{z}_i, \boldsymbol{\delta}_i; \boldsymbol{\theta})$ and $g(\boldsymbol{\delta}_i; \boldsymbol{\eta})$ conditional on the observed choices $\mathbf{c}_i$ and explanatory variables $\mathbf{z}_i$ according to the probability distribution functions obtained from the preceding iteration. Denoting the last available estimates of $(\boldsymbol{\theta}_0, \boldsymbol{\eta}_0)$ by $(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\eta}})$, this step aims at obtaining the 'modified' log-likelihood functions given by

$$\hat{\ell}_i(\boldsymbol{\theta}) = \int \ln\left[g(\mathbf{c}_i \mid \mathbf{z}_i, \boldsymbol{\delta}; \boldsymbol{\theta})\right] g(\boldsymbol{\delta} \mid \mathbf{c}_i, \mathbf{z}_i; \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\eta}}) d\boldsymbol{\delta} \tag{15}$$

---

of the produced crops ($\mathbf{s}_{it}^+$), i.e. the terms $g(r_{it} \mid \mathbf{z}_{it}, \boldsymbol{\delta}_i, \mathbf{s}_{it}^+; \boldsymbol{\theta})$. These probability functions are computed as expectations over the joint probability distributions function of the terms $\varepsilon_{k,it}^s$ for $k \in \mathcal{K}^0(r_{it})$. These error terms are arguments of the regime expected profit levels $\Pi_{it}(r)$ that are themselves arguments of the production regime choice. But they must be integrated out in the model likelihood function because they cannot be recovered from the data. We use Laplace approximations for computing the expectations yielding the terms $g(r_{it} \mid \mathbf{z}_{it}, \boldsymbol{\delta}_i, \mathbf{s}_{it}^+; \boldsymbol{\theta})$ (see, *e.g.*, Harding and Hausman, 2007, for an application of this integration approach in a related context). Also, the parameters in $g(\mathbf{y}_{it}^+, \mathbf{x}_{it}^+ \mid \mathbf{z}_{it}, \boldsymbol{\delta}_i; \boldsymbol{\theta}_{r_{it}})$ and $g(\mathbf{s}_{it}^+ \mid \mathbf{z}_{it}, \boldsymbol{\delta}_i; \boldsymbol{\theta}_{r_{it}})$ depends on regime in belong the farmer i at time t. To have a function of $\boldsymbol{\theta}$, we need to treat $\varepsilon_{k,it}^{yx,0}$ and $\varepsilon_{k,it}^{s,0}$ as latent variables and to consider $g(\mathbf{y}_{it}^+, \mathbf{x}_{it}^+, \varepsilon_{k,it}^{yx,0} \mid \mathbf{z}_{it}, \boldsymbol{\delta}_i; \boldsymbol{\theta})$ and $g(\mathbf{s}_{it}^+, \varepsilon_{k,it}^{s,0} \mid \mathbf{z}_{it}, \boldsymbol{\delta}_i; \boldsymbol{\theta})$ -, rather than $g(\mathbf{y}_{it}^+, \mathbf{x}_{it}^+ \mid \mathbf{z}_{it}, \boldsymbol{\delta}_i; \boldsymbol{\theta}_{r_{it}})$ and $g(\mathbf{s}_{it}^+ \mid \mathbf{z}_{it}, \boldsymbol{\delta}_i; \boldsymbol{\theta}_{r_{it}})$ respectively.

and

$$\hat{\ell}_i(\boldsymbol{\eta}) = \int \ln\left[g(\boldsymbol{\delta} \mid \boldsymbol{\eta})\right] g(\boldsymbol{\delta} \mid \mathbf{c}_i, \mathbf{z}_i; \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\eta}}) d\boldsymbol{\delta} \,. \tag{16}$$

The M step then consists of maximizing in the interest parameters $\boldsymbol{\theta}$ and $\boldsymbol{\eta}$ the sample modified log-likelihood functions obtained from the E step. *I.e.* it consists of solving the problems $\max_{\boldsymbol{\theta}} N^{-1} \sum_{i=1}^{N} \hat{\ell}_i(\boldsymbol{\theta})$ and $\max_{\boldsymbol{\eta}} N^{-1} \sum_{i=1}^{N} \hat{\ell}_i(\boldsymbol{\eta})$. This is where EM algorithms take advantage of the likelihood functions of random parameter models in general. Our model being composed of four equation sub-systems, the structure of the terms $\hat{\ell}_i(\boldsymbol{\theta})$ allows for further splitting the maximization problem $\max_{\boldsymbol{\theta}} N^{-1} \sum_{i=1}^{N} \ln \hat{\ell}_i(\boldsymbol{\theta})$ into a few simpler maximization problems. Also, given equation (13), $\hat{\ell}_i(\boldsymbol{\theta})$ can be decomposed as follows:

$$\begin{aligned} \hat{\ell}_i(\boldsymbol{\theta}) = & \int \sum_{t \in \mathcal{H}(i)} \ln\left[g(\mathbf{y}_{it}^{+}, \mathbf{x}_{it}^{+} \mid \mathbf{z}_{it}, \boldsymbol{\delta}_i; \boldsymbol{\theta}_{r_{it}})\right] g(\boldsymbol{\delta} \mid \mathbf{c}_i, \mathbf{z}_i; \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\eta}}) d\boldsymbol{\delta} \\ & + \int \sum_{t \in \mathcal{H}(i)} \ln\left[P(r_{it} \mid \mathbf{z}_{it}, \boldsymbol{\delta}_i, \boldsymbol{\varepsilon}_{it}^{s,+}; \boldsymbol{\theta})\right] g(\boldsymbol{\delta} \mid \mathbf{c}_i, \mathbf{z}_i; \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\eta}}) d\boldsymbol{\delta} \\ & + \int \sum_{t \in \mathcal{H}(i)} \ln\left[g(\mathbf{s}_{it}^{+} \mid \mathbf{z}_{it}, \boldsymbol{\delta}_i; \boldsymbol{\theta}_{r_{it}})\right] g(\boldsymbol{\delta} \mid \mathbf{c}_i, \mathbf{z}_i; \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\eta}}) d\boldsymbol{\delta} \end{aligned} \tag{17}$$

However, the 'modified' likelihood functions $\hat{\ell}_i(\boldsymbol{\theta})$ and $\hat{\ell}_i(\boldsymbol{\eta})$ can be integrated neither analytically nor numerically. Stochastic versions of the EM algorithm – *i.e.* the so-called SEM or Monte Carlo EM (MCEM) algorithms – have been proposed in the computational statistics literature for combining the advantages of the EM algorithm and simulation methods (see, e.g., McLachlan and Krishnan, 2008). In this study, we rely on the Stochastic

Approximation EM algorithm proposed by Delyon et al. (1999) combined with Importance Sampling methods via Quasi-Monte Carlo simulation, proposed in the same context by Jank (2005) for estimating the 'modified' likelihood functions $\hat{\ell}_i(\boldsymbol{\theta})$ and $\hat{\ell}_i(\boldsymbol{\eta})$. Thus, at iteration j, S values $\left\{\tilde{\boldsymbol{\rho}}_{i,s}(\boldsymbol{\eta}^{(j-1)}) : s = 1,...,S\right\}$ are quasi-randomly drawn from the distribution $g(\boldsymbol{\rho}_i;\boldsymbol{\eta}^{(j-1)})$, and corresponding weights $\tilde{\omega}_{i,s}(\boldsymbol{\theta},\boldsymbol{\eta})$ are computed as:

$$(18) \qquad \tilde{\omega}_{i,s}(\boldsymbol{\theta},\boldsymbol{\eta}) \equiv \frac{\prod_{t=1}^{T}\begin{pmatrix} g(\mathbf{y}_{it}^{+},\mathbf{x}_{it}^{+} \mid \mathbf{z}_{it},\tilde{\boldsymbol{\delta}}_{i,s}(\boldsymbol{\eta});\boldsymbol{\theta}_{r_{it}}) \\ P(r_{it} \mid \mathbf{z}_{it},\tilde{\boldsymbol{\delta}}_{i,s}(\boldsymbol{\eta}),\boldsymbol{\varepsilon}_{it}^{s,+};\boldsymbol{\theta})g(\mathbf{s}_{it}^{+} \mid \mathbf{z}_{it},\tilde{\boldsymbol{\delta}}_{i,s}(\boldsymbol{\eta});\boldsymbol{\theta}_{r_{it}}) \end{pmatrix}}{S_n^{-1}\sum_{s=1}^{S_n}\prod_{t=1}^{T}\begin{pmatrix} g(\mathbf{y}_{it}^{+},\mathbf{x}_{it}^{+} \mid \mathbf{z}_{it},\tilde{\boldsymbol{\delta}}_{i,s}(\boldsymbol{\eta});\boldsymbol{\theta}_{r_{it}}) \\ P(r_{it} \mid \mathbf{z}_{it},\tilde{\boldsymbol{\delta}}_{i,s}(\boldsymbol{\eta}),\boldsymbol{\varepsilon}_{it}^{s,+};\boldsymbol{\theta})g(\mathbf{s}_{it}^{+} \mid \mathbf{z}_{it},\tilde{\boldsymbol{\delta}}_{i,s}(\boldsymbol{\eta});\boldsymbol{\theta}_{r_{it}}) \end{pmatrix}} .$$

These simulated values are then used to approximate the expectations in the E step, *i.e.* $\hat{\ell}_i(\boldsymbol{\theta})$ and $\hat{\ell}_i(\boldsymbol{\eta})$. The term $\tilde{\omega}_{i,s}(\boldsymbol{\theta},\boldsymbol{\eta})$ depends on $P(r_{it} \mid \mathbf{z}_{it},\tilde{\boldsymbol{\delta}}_{i,s}(\boldsymbol{\eta}),\boldsymbol{\varepsilon}_{it}^{s,+};\boldsymbol{\theta})$, which cannot be computed directly since it depends on the $\boldsymbol{\varepsilon}_{it}^{s,0}$ that are unobserved: $P(r_{it} \mid \mathbf{z}_{it},\tilde{\boldsymbol{\delta}}_{i,s}(\boldsymbol{\eta}),\boldsymbol{\varepsilon}_{it}^{s,+};\boldsymbol{\theta})$ needs to be integrated on the distribution of $\boldsymbol{\varepsilon}_{it}^{s,0}$ conditional on $\boldsymbol{\varepsilon}_{it}^{s,+}$:

$$(19) \qquad P(r_{it} \mid \mathbf{z}_{it},\boldsymbol{\delta}_i,\boldsymbol{\varepsilon}_{it}^{s,+};\boldsymbol{\theta}) = \int P(r_{it} \mid \mathbf{z}_{it},\boldsymbol{\delta}_i,\boldsymbol{\varepsilon}_{it}^{s,+},\boldsymbol{\varepsilon}_{it}^{s,0};\boldsymbol{\theta})g(\boldsymbol{\varepsilon}_{it}^{s,0} \mid \boldsymbol{\varepsilon}_{it}^{s,+};\hat{\boldsymbol{\theta}})d\boldsymbol{\varepsilon}_{it}^{s,0} .$$

This way of computing the probability may be time consuming. We use a Laplace approximation (Harding and Hausman, 2007) to reduce the computation time.

The M step of the algorithm can be simplified by using in addition an inner ECM algorithm in the considered SAEM algorithm at each iteration. This procedure consists in estimating the $\boldsymbol{\theta}$ parameter as:

$$(20)\ \boldsymbol{\theta}^{(j)} = \arg\max_{\boldsymbol{\theta}} N^{-1}\sum_{i=1}^{N}\left[\begin{array}{c}\int\left(\sum_{t\in\mathcal{H}(i)}\left[\int \ln g(\mathbf{y}_{it}^{+},\mathbf{x}_{it}^{+},\boldsymbol{\varepsilon}_{it}^{yx,0}\mid\mathbf{z}_{it},\boldsymbol{\delta}_i;\boldsymbol{\theta})g(\boldsymbol{\varepsilon}_{it}^{yx,0}\mid\boldsymbol{\varepsilon}_{it}^{yx,+})d\boldsymbol{\varepsilon}_{it}^{yx,0}\right]\right.\\ \left. g(\boldsymbol{\delta}\mid\mathbf{c}_i,\mathbf{z}_i;\boldsymbol{\theta}^{(j-1)},\boldsymbol{\eta}^{(j-1)})\right)d\boldsymbol{\delta}\\ +\int\left(\sum_{t\in\mathcal{H}(i)}\left[\int \ln g(\mathbf{s}_{it}^{+},\boldsymbol{\varepsilon}_{it}^{s,0}\mid\mathbf{z}_{it},\boldsymbol{\delta}_i;\boldsymbol{\theta})g(\boldsymbol{\varepsilon}_{it}^{yx,0}\mid\boldsymbol{\varepsilon}_{it}^{yx,+})d\boldsymbol{\varepsilon}_{it}^{s,0}\right]\right.\\ \left. g(\boldsymbol{\delta}\mid\mathbf{c}_i,\mathbf{z}_i;\boldsymbol{\theta}^{(j-1)},\boldsymbol{\eta}^{(j-1)})\right)d\boldsymbol{\delta}\end{array}\right]$$

such that $N^{-1}\sum_{i=1}^{N}\hat{\ell}_i(\boldsymbol{\theta}^{(j)}) \geq N^{-1}\sum_{i=1}^{N}\hat{\ell}_i(\boldsymbol{\theta}^{(j-1)})$ at each iteration, rather than:

$$(21)\ \boldsymbol{\theta}^{(j)} = \arg\max_{\boldsymbol{\theta}} N^{-1}\sum_{i=1}^{N}\left[\begin{array}{l}\int\sum_{t\in\mathcal{H}(i)}\ln\left[g(\mathbf{y}_{it}^{+},\mathbf{x}_{it}^{+}\mid\mathbf{z}_{it},\boldsymbol{\delta}_i;\boldsymbol{\theta}_{r_{it}})\right]g(\boldsymbol{\delta}\mid\mathbf{c}_i,\mathbf{z}_i;\boldsymbol{\theta}^{(j-1)},\boldsymbol{\eta}^{(j-1)})d\boldsymbol{\delta}\\ +\int\sum_{t\in\mathcal{H}(i)}\ln\left[g(\mathbf{s}_{it}^{+}\mid\mathbf{z}_{it},\boldsymbol{\delta}_i;\boldsymbol{\theta}_{r_{it}})\right]g(\boldsymbol{\delta}\mid\mathbf{c}_i,\mathbf{z}_i;\boldsymbol{\theta}^{(j-1)},\boldsymbol{\eta}^{(j-1)})d\boldsymbol{\delta}\\ +\int\sum_{t\in\mathcal{H}(i)}\ln\left[P(r_{it}\mid\mathbf{z}_{it},\boldsymbol{\delta}_i,\boldsymbol{\varepsilon}_{it}^{s,+};\boldsymbol{\theta})\right]g(\boldsymbol{\delta}\mid\mathbf{c}_i,\mathbf{z}_i;\boldsymbol{\theta}^{(j-1)},\boldsymbol{\eta}^{(j-1)})d\boldsymbol{\delta}\end{array}\right]$$

6.3.3. Calibration of the simulation model

The estimated ERS multi-crop micro-econometric model can be used for "statistically calibrating" its random parameters for each farm of our sample and thus for obtaining a simulation model consisting of a sample of farm specific "calibrated" models (see, e.g., Koutchade et al., 2015). The underlying idea of this procedure is to use the estimated distribution of the random parameters and farmers' observed choices compute estimates of the farm specific parameters according to a "Tell me what you did, I will tell you who you are" logic.

Interestingly, the Expectation step of the SEM algorithm we use relies on computations closely related to this calibration procedure since both rely on the probability distributions of the random parameters $\boldsymbol{\delta}_i$ conditional on the observed choices $\mathbf{c}_i$ and explanatory variables

$\mathbf{z}_i$. In this study, the specific parameter $\boldsymbol{\delta}_i$ of farm *i* is calibrated as the mode – *i.e.* according to a ML 'calibration' criterion – of its simulated probability distribution conditional on $(\mathbf{c}_i, \mathbf{z}_i)$, *i.e.* on what is known about farm *i* in the data. Also, this 'statistical calibration' procedure and its counterpart in the SAEM algorithm allow for calibrating the random parameters corresponding to crops that have not been grown by the considered farmer or corresponding to regime fixed costs for regimes that have not been chosen by the considered farmer.

6.4. Empirical application

6.4.1. Data and empirical specification of the model

The model is estimated on an unbalanced panel data set containing 2276 observations of 415 French grain crop producers in the North and North-East of France, over the years 2006 to 2011. This sample has been extracted from data provided by an accounting agency located in the French territorial division *La Marne*. It contains detailed information about crop production for each farm (acreages, yields, input uses and crop prices at the farm gate). We consider seven crops (or crop aggregates): sugar beet, alfalfa, peas, rapeseed, winter crops (wheat and barley), corn and spring barley, which represent more than 80% of the total

acreage in the considered area.[80]

The variable input aggregate account for the use of fertilizers, pesticides and seeds. The corresponding price index is computed as a standard Tornqvist index. When a farmer doesn't produce a crop the corresponding output and input prices are unobserved. These missing prices were estimated by the yearly average of the corresponding observed prices. All prices are deflated by the hired production services price index (base 1 in 2006) obtained from the French department of Agriculture. This aggregated price index mainly depends on the price indices of machinery, fuel and hired labor, the main inputs involved in the implicit acreage management cost function. The climatic variables are provided at the municipality level by Météo France, the French national meteorological service.

Figure 1 depicts the three levels nesting structure that we adopt for the seven crops. In a first level we distinguish a cereal group composed of winter crops, corn and spring barley, and a group of crops that are generally planted at the head of rotation: corn, alfalfa, peas and rapeseed.

[80] The EU sugar beet subsidy scheme requires limited adjustments in our application. This scheme is based on sugar beet production quotas – held by farmers on a historical basis – with subsidized prices. Yet, the actual sugar beet production largely exceeds the subsidized quota for all sugar beet producers of our sample. This suggests that these farmers choose their sugar beet acreages, yield levels and input use levels according to the off-quota sugar beet prices.

Figure 1. Nesting structure of the acreage choice model

Groups	Cereals			Rotation heads			
Subgroups	Winter cereals	Spring cereals		Oil and protein crops			Root crops
Crops	Winter cereals*	Spring barley	Corn	Rapeseed	Protein peas	Alfalfa	Sugar beet

* Winter wheat (mostly) and winter barley

This structure is intended to reflect the basic rotation scheme of grain producers in France. In a second level, the cereal group is split into two subgroups: winter cereals on the one hand and other cereals on the other hand, in order to account for the differences in planting seasons between those cereals. The 'head of rotation' group is split into an 'oilseeds and protein crops' subgroup and a subgroup including only sugar beet (the only root crop considered here).

This implies that, in the first level, the (between nests) α_i^s parameter drives the land allocation to cereals and to rotation heads. In the second level, the (within cereals) $\alpha^s_{(\text{cereals}),i}$ parameter drives the allocation of the cereal acreage to winter cereals and to spring barley, and the (within rotation heads) $\alpha^s_{(rotation\ heads),i}$ parameter drives the allocation of the rotation heads to oil and protein crops and to root crops. In the third level, the (within spring cereals) The $\alpha^s_{(spring\ cereals),i}$ parameter drives the allocation of spring cereals to spring barley and to corn while the (within oil and protein crops) $\alpha^s_{(oil\ and\ protein\ crops),i}$ parameter drive the allocation of oil and protein crops to rapeseed, to protein peas and to alfalfa. Winter cereals are used as the benchmark crop in the acreage share models.

Table 1 provides descriptive statistics concerning the production regimes observed in the

data.

Table 1. Descriptive statistics of the production regimes represented in the data

Regime number	Crops produced in the regime							Regime frequency	Average gross margin (€/ha)[d]
	Winter cereals	*Corn*	*Spring Barley*	*Sugarbeet*	*Alfalfa*	*Peas*	*Rapeseed*		
1								06.6%	953
2								21.5%	1014
3								11.8%	930
4								48.6%	1007
5								02.8%	989
6								02.5%	825
7								04.9%	970
8								01.3%	738
Average acreage share[a]	38.3% (0.09)	02% (0.05)	18.2% (0.07)	14.5% (0.06)	09.7% (0.05)	1.7% (0.03)	15.6% (0.06)		
Average acreage share[a] if produced[a]	38.3% (0.09)	08.4% (0.07)	18.2% (0.07)	15.1% (0.06)	10.9% (0.04)	5.9% (0.03)	15.6% (0.06)		
Production frequency	100%	23.7%	100%	96.14%	88.4%	28.%	100%		
Average gross margin (€/ha)[a]	842.6 (327)	872.38 (448.6)	755.5 (286.8)	1789.3 (378.8)	561.9 (286.3)	662.6 (269)	842.9 (311.1)		
Average yield (t/ha)[a]	8.58[b] (0.88)	9.23 (1.73)	6.82 (1.21)	95.19[c] (1.30)	12.62 (1.92)	4.72 (1.28)	3.89 (0.64)		
Average price (€/t)[a]	148.8[b] (31.05)	131.29 (33.85)	155.53 (35.30)	24.657[c] (2.99)	72.19 (15.49)	197.7 (25.5)	322.58 (64)		
Average production cost (fertilizers and pesticides)[a]	431.3 (90.8)	307.9 (74.2)	294.1 (69.9)	547.3 (126.4)	350.1 (124.8)	245.8 (66)	414.8 (83)		

a Empirical standard deviation in parentheses. b Winter wheat. c. Off-quota price. d. Sugar beet subsidies excluded.

Based on these seven crops, 127 regimes, could theoretically be grown by farmers. Only eight of them are actually observed in our sample.[81] All farmers grow winter crops and at least two additional crops. The most frequent regimes in the sample (regimes 2, 3 and 4) actually include five or six crops. It is interesting to note that most farmers adopt different production regimes over the 5 years of our sample: only 8 out of 415 farmers have not changed their production regime during the period. The average gross margins associated to each regime are reported in the last column of Table 1. An interesting feature appears here: the most frequently chosen regimes are not the ones that lead on average to the highest gross margin. For instance, regime 2 – which excludes corn – is characterized by the highest observed gross margin, but has been adopted in only 21.5% of the observations. This comes to illustrate the fact that farmers' choices of production regime are driven by other factors than gross returns, such as the acreage management and regime fixed costs represented in our model.

Regarding regime fixed costs, we rely on the second type of specification presented in section 6.2.3 and assume that regime fixed costs are equal to a sum of fixed costs associated to each crop produced in the regime. We consider regime 8, which contains all other regimes represented in our data, as the reference regime (r_0) and specify regime fixed costs as $d_i(r) \equiv \sum_{k \in \mathcal{K}^+(r) \setminus \mathcal{K}^+(r_0)} d^c_{k,i}$. As winter crop, spring Barley and rapeseed are always produced in the considered sample, we assume that the corresponding costs, $d^c_{k,i}$, for these crops are zero. The regime fixed cost $d_i(r)$ associated to regime 8 is thus equal to zero and only the fixed costs $d^c_{k,i}$ corresponding to corn, sugar beet, alfalfa and peas have to be estimated. We have four

[81] This are the 8 regimes the most represented, in way statistical inference could be done.

random parameter rather than seven in the case of regime fixed costs as a constant. Given that and after estimation, it is possible to recover the (relative) regime fixed costs of regimes that contain at least regime 8, and at most regime 1.

6.4.2. Estimation results

The parameter estimates of the yield, input demand, acreage shares and regime choice equations are reported in Tables 2 to 5.

As shown in Table 2, the expectations of random parameters representing potential yields, $\beta_{k,i}^{y}$, are significantly estimated for all crops and their values lie in reasonable ranges regarding the average yields observed in the sample (Table 1). More importantly, the variances of their distributions are also significantly different from zero for all crops. These parameters thus significantly vary across farms, despite the fact that we control for observed factors characterizing farm heterogeneity (land and capital endowments and climate). This comes to illustrate the importance of unobserved farm heterogeneity and the relevance of the use of a random parameter approach that allows dealing with this heterogeneity. The parameter estimates of the input demand equation, reported in Table 3, confirm this: not only are the expectations of random parameter significantly estimated but also their variance. This is true for the "constant" of the model $\beta_{k,i}^{x}$ (the input use requirement) but also for its "slope", $\alpha_{k,i}^{x}$, which represents the response of farmers to change in netput prices. We can also notice that these price parameters are all positive and thus satisfy the restrictions imposed by the concavity of production functions.

Table 2. Selected Parameter Estimates of Yield Supply Models[a]

		RP yield models						
		Winter cereals	*Corn*	*Spring Barley*	*Sugar beet*	*Alfalfa*	*Peas*	*Rapeseed*
Error term $\varepsilon_{k,it}^{y}$								
	Std dev	0.66*	1.83*	0.95*	9.7*	2.96*	1.72*	0.49*
		(0.016)	(0.066)	(0.02)	(0.02)	(0.02)	(0.03)	(0.016)
Potential yield $\beta_{k,i}^{y}$								
	Mean	8.71*	9.06*	6.81*	95.6*	12.23*	4.15*	4.04*
		(0.02)	(0.04)	(0.02)	(0.32)	(0.04)	(0.03)	(0.01)
	Std dev	0.26*	0.65*	0.33*	5.7*	0.69*	0.51*	0.24*
		(0.01)	(0.03)	(0.01)	(0.17)	(0.02)	(0.02)	(0.01)
Input use flexibility $\alpha_{k,i}^{x}$								
	Mean	0.43*	0.08*	0.30*	0.49*	0.25*	0.33*	0.79*
		(0.01)	(0.004)	(0.003)	(0.03)	(0.001)	(0.01)	(0.02)
	Std dev	0.13*	0.09*	0.05*	0.58*	0.02	0.18*	0.31*
		(0.01)	(0.04)	(0.002)	(0.06)	(0.03)	(0.01)	(0.01)

a. Estimated standard errors of the ML estimator are in parentheses. Note: Asterisk (*) denotes a statistically significant parameter at the 5% level.

Table 3. Selected Parameter Estimates of the Input Demand Models[a]

		RP input demand models						
		Winter cereals	*Corn*	*Spring Barley*	*Sugar beet*	*Alfalfa*	*Peas*	*Rapeseed*
Error term $\varepsilon_{k,it}^{x}$								
	Std dev	0.52*	0.59*	0.41*	0.84*	0.88*	0.6*	0.58*
		(0.01)	(0.02)	(0.009)	(0.017)	(0.019)	(0.024)	(0.013)
Input requirement $\beta_{k,i}^{x}$								
	Mean	4.36*	2.57*	2.92*	5.44*	3.15*	2.29*	4.44*
		(0.02)	(0.02)	(0.01)	(0.03)	(0.03)	(0.02)	(0.02)
	Std dev	0.37*	0.33*	0.24*	0.54*	0.44*	0.37*	0.41*
		(0.02)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.02)
Input use flexibility $\alpha_{k,i}^{x}$								
	Mean	0.43*	0.08*	0.30*	0.49*	0.25*	0.33*	0.79*
		(0.01)	(0.004)	(0.003)	(0.03)	(0.001)	(0.01)	(0.02)
	Std dev	0.13*	0.09*	0.05*	0.58*	0.02	0.18*	0.31*
		(0.01)	(0.04)	(0.002)	(0.06)	(0.03)	(0.01)	(0.01)

a. Estimated standard errors of the ML estimator are in parentheses. Note: Asterisk (*) denotes a statistically significant parameter at the 5% level.

Turning to the parameter estimates of the acreage share equations in Table 4, again, the expectations and variance of random parameters are significantly estimated. The range of expectations of the acreage flexibility parameters are consistent with the theory: namely,

$\alpha^s_{(spring\ cereals),i} \geq \alpha^s_{(\text{cereals}),i} \geq \alpha^s_i > 0$ and $\alpha^s_{(oil\ and\ protein\ crops),i} \geq \alpha^s_{(rotation\ heads),i} \geq \alpha^s_i > 0$, which are sufficient condition for the acreage model to be well-behaved.

Table 4. Selected Parameter Estimates of the Acreage Share Models[a]

		RP acreage share models						
		Winter cereals	***Corn***	***Spring Barley***	***Sugar beet***	***Alfalfa***	***Peas***	***Rapeseed***
Error term $\varepsilon^s_{k,it}$								
	Std dev	-	11.12*	9.91*	6.25*	6.77*	8.56*	7.09*
			(0.38)	(0.19)	(0.13)	(0.15)	(0.28)	(0.16)
Acreage share shifters $\beta^s_{k,i}$								
	Mean	-	17.41*	13.88*	24.51*	11.15*	18.78*	11.07*
			(0.73)	(0.37)	(0.23)	(0.24)	(0.37)	(0.24)
	Std dev	-	3.92*	4.19*	3.96*	2.70*	2.62*	2.20*
			(0.015)	(0.017)	(0.025)	(0.057)	(0.012)	(0.014)
Acreage choice flexibility		***Cereals vs rotation heads***	***Within Cereals***		***Within Rotation heads***	***Within spring cereals***		***Within oil and protein crops***
		α^s_i	$\alpha^s_{(\text{cereals}),i}$		$\alpha^s_{(rotation\ heads),i}$	$\alpha^s_{(spring\ cereals),i}$		$\alpha^s_{(oil\ and\ protein\ crops),i}$
	Mean	0.046*	0.053*		0.073*	0.53*		0.11*
		(0.001)	(0.0008)		(0.001)	(0.029)		(0.0016)
	Std dev	0.015*	0.013*		0.025*	0.64*		0.02*
		(0.0009)	(0.0008)		(0.0013)	(0.029)		(0.002)

a. Estimated standard errors of the ML estimator are in parentheses. Note: Asterisk (*) denotes a statistically significant parameter at the 5% level.

By comparing, in Tables 2 to 4, the estimated standard deviations of the "constant" terms in our model ($\beta^y_{k,i}$, $\beta^x_{k,i}$ and $\beta^s_{k,i}$) to the estimated standard deviations of error terms, we can notice that, although the variations in these random parameters account for a significant part of the observed variations in farmers' choices, the variations in the model error terms account for a comparable part. Even if crop production is known to be significantly affected by numerous random events, this indicates that there is still room for further improving the explanatory power of our model. Further research should thus be conducted in that direction, by looking for additional control variables in particular.

Finally, as shown in Table 5, the regime costs associated to crops, $d_{k,i}^{c}$, and the scale parameter, σ_i, of the regime choice equation are significantly estimated and heterogeneous across the sample. The mean value of the scale parameter, 1.80, is relatively large which reflects a relative importance of the deterministic part of the model: the profit and regime fixed costs in the regime choice appear to be significant drivers of the regime choices. As mentioned in section 2.3, we expected regime fixed costs to increase with the addition of a crop to a regime. If this is on average true for some regimes without alfalfa (for instance, moving from regime 8 to regime 5 by adding corn and sugarbeet increases fixed costs on average), it is not always the case. In particular, the fixed cost associated to alfalfa being negative (-4.7), introducing this crop in production regimes appears to decrease regime fixed costs (for instance moving from regime 4 to regime 1 by adding alfalfa and peas decreases fixed costs on average). Two reasons might explain this result. First, in the *Marne* region, the downstream (dehydration) industry contributes to part of the cropping and harvesting of alfalfa, which comes to decrease the workload for farmers growing this crop. Second, like other protein crops, alfalfa exhibits good agronomic properties, especially when used as a previous crop for cereals, and does not require a lot of work during the cropping season contrary to other protein crops like peas. The decrease in regime costs induced by the introduction of alfalfa might thus be explained to the expected benefits of growing this crop. This result should however be considered cautiously given the high and significant estimated variability of regime fixed costs across farmers: those parameters are highly heterogeneous in the population and the ranking of regime costs prevailing for one farmer might not be the same for another one.

Table 5. Parameter Estimates of Regime Choice Models

	Crops produced in the regime							
	Winter cereals	***Corn***	***Spring Barley***	***Sugar beet***	***Alfalfa***	***Peas***	***Rapeseed***	
Fixed costs $d_{k,i}^{c}$								
Mean	-	3.8* (0.24)	-	0.3* (0.12)	-4.7* (0.28)	1.3* (0.04)	-	
Std dev	-	4.16* (0.10)	-	2.22* (0.05)	4.4* (0.11)	0.67* (0.01)	-	
Regime number								**Mean regime fixed costs**
1								0.7
2								-3.1
3								-0.6
4								-4.4
5								4.1
6								3.8
7								0.3
8								-
Scale parameter σ_i								
Mean	1.80 (0.073)							
Std dev	1.4 (0.07)							

a Standard deviation in parentheses. b Winter wheat. c. Off-quota price. d. Sugar beet subsidies excluded.

Once we have estimated the parameters characterizing the distribution of the random parameters $\boldsymbol{\delta}_i$, we can "statistically calibrate" those parameters for each farmer in our sample and thus obtain a set of farmer specific "calibrated" models (Koutchadé et al., 2017a). We thus statistically calibrate the value of $\boldsymbol{\delta}_i$ for each farmer in our sample. One interesting feature is that this procedure also allows us to calibrate the parameters of the yield, input demand and acreage equations corresponding to crops that have not been grown by the considered farmer as well as farmer specific regime fixed costs for regime that have never been chosen by the considered farmer.

The estimated farmer specific models allow us to compute some fitting criteria, Sim-R^2, which are reported in Tables 7 to 9. The Sim-R^2 criterion measures the quality of the prediction of the observed choices of farmers by the estimated models. Its construction is analogous to

that of the R^2 criterion of the standard linear regression model: for a given choice variable and a given model, the Sim-R^2 criterion is defined as the ratio of the empirical variance of the prediction of this variable to the empirical variance of the observed variable. These estimated criteria tend to show that the proposed model offers a satisfactory fit to our data, with a better fit to the major crops than to the ones less frequently produced or with smaller acreages (peas, corn and, to a lesser extent, alfalfa).

Furthermore, using the estimated farmer specific models to predict the regimes choices observed in our data, we find our model to exhibit a relatively good predictive power with 72.24% of regime choices accurately predicted.

Table 7. Fitting criteria of Yield Supply Models (*Sim-R²*)

Regime number	Crops produced in the regime						
	Winter cereals	*Corn*	*Spring Barley*	*Sugar beet*	*Alfalfa*	*Peas*	*Rapeseed*
1	0.265	0.202	0.368	0.426	0.218	0.344	0.385
2	0.295		0.318	0.487	0.305	0.404	0.464
3	0.452	0.253	0.411	0.398	0.254		0.470
4	0.407		0.350	0.414	0.292		0.450
5	0.263	0.416	0.270	0.529			0.411
6	0.368	0.126	0.315				0.378
7	0.312		0.429	0.294			0.396
8	0.201		0.336				0.473

Table 8. Fitting criteria of Input Demand Models (*Sim-R²*)

Regime number	Crops produced in the regime						
	Winter cereals	*Corn*	*Spring Barley*	*Sugar beet*	*Alfalfa*	*Peas*	*Rapeseed*
1	0.376	0.283	0.319	0.254	0.205	0.597	0.319
2	0.387		0.354	0.286	0.237	0.381	0.345
3	0.439	0.330	0.378	0.334	0.270		0.374
4	0.479		0.442	0.394	0.347		0.448
5	0.372	0.256	0.257	0.205			0.190
6	0.331	0.243	0.319				0.315
7	0.486		0.520	0.318			0.521
8	0.164		0.152				0.251

Table 9. Fitting criteria of Acreage Share Models (*Sim-R²*)

Regime number	Crops produced in the regime						
	Winter cereals	*Corn*	*Spring Barley*	*Sugar beet*	*Alfalfa*	*Peas*	*Rapeseed*
1		0.642	0.443	0.948	0.808	0.501	0.415
2			0.352	0.970	0.715	0.541	0.496
3		0.600	0.365	0.782	0.722		0.407
4			0.324	0.762	0.676		0.385
5		0.412	0.372	1.000			0.155
6		0.396	0.361				0.222
7			0.390	0.781			0.284
8			0.197				0.766

Importantly, our investigations on this issue tend to demonstrate that our results are robust to various distributional assumptions related to the model random parameters.

6.4.3. Simulation results

The structure of the proposed ERS multi-crop micro-econometric model allows for investigating the relative importance of the main drivers of production regime choices. For that purpose, we consider the simulation model obtained from the estimated one by calibrating the farm specific parameters for each farm of our sample. Then we use this simulation model for investigating the prediction power of three elements of the regime choice models: the weighted sum of the expected crop gross returns $\mathbf{s}_{it}(r)'\boldsymbol{\pi}_{it}^{e}$, the acreage management costs $C_{it}\left(\mathbf{s}_{it}(r)\right)$ and the regime fixed costs d_i^r for $r \in \mathcal{R}$. We simulate the regime choices according to each of these elements as well as combinations of these elements, and then confront them, on average, with the observed regime choices. Taken together these simulation results confirm that regime fixed costs matter, but mainly in combination with the other drivers of the regime choice model. The maximization of gross margins, or the minimization of acreage management costs or regime fixed cost alone leads to very poor predictions of regime choices. Considering pairs of these choice criteria only slightly improve the predictions, while considering together these three

criteria unsurprisingly provides predicted choices very close, on average, to the observed ones.

Price elasticity. To illustrate the relevance of the approach we propose to deal with corner solutions in acreage choices, we simulate the impacts of changes in expected crop prices on acreage choices. Acreage price elasticities play a crucial role in this type of exercise. Yet these elasticities account both for the impact of crop prices on acreages within any given regime and for the switch in production regimes induced by crop price changes. These two effects can be distinguished by generalizing to multiple regimes the decomposition originally proposed by McDonald and Moffit (1980) in the case of a Tobit model.

Indeed, expected acreage shares depend on conditional expected acreage shares given the chosen production regime and on the probability of regime choice. Formally, the expected acreage share of crop k for farmer i in year t is given by:

(22) $$E\left(s_{k,it}\right)=\sum_{r=1}^{8}\Pr\left(r_{it}=r\right)\times E\left(s_{k,it}\middle|r_{it}=r\right) \text{ for } k\in\mathcal{K} \text{ and } r\in\mathcal{R},$$

and its variation with crop price can be decomposed into two parts: changes in the probabilities of regime choice and changes in acreage shares within each regime:

(23) $$\frac{\partial}{\partial p_{l,it}}E\left(s_{k,it}\right)=\sum_{r=1}^{8}\left[\frac{\partial}{\partial p_{l,it}}\Pr\left(r_{it}=r\right)\times E\left(s_{k,it}\middle|r_{it}=r\right)+\frac{\partial}{\partial p_{l,it}}E\left(s_{k,it}\middle|r_{it}=r\right)\times\Pr\left(r_{it}=r\right)\right].$$

The (unconditional) price elasticities of acreage are given by:

(24) $$\begin{aligned}\varepsilon_{s_{k,it}|p_{j,it}}&=\frac{\partial}{\partial p_{j,it}}E\left(s_{k,it}\right)\times\frac{p_{j,it}}{E\left(s_{k,it}\right)}\\&=\frac{\hat{y}_{j,it}p_{j,it}}{E\left(s_{k,it}\right)}\left[\begin{array}{l}\sum_{r=1}^{8}\left[\frac{\partial}{\partial\pi_{j,it}}\Pr\left(r_{it}=r\right)\times E\left(s_{k,it}\middle|r_{it}=r\right)\right]\\+\ \sum_{r=1}^{8}\left[\frac{\partial}{\partial\pi_{j,it}}E\left(s_{k,it}\middle|r_{it}=r\right)\times\Pr\left(r_{it}=r\right)\right]\end{array}\right]\end{aligned}$$

where

$$(25)\ \frac{\partial}{\partial \pi_{j,it}} \Pr\left(r_{it}=r\right) = \sum_l \sigma_i \times \begin{Bmatrix} \Pr\left(r_{it}=r\right) - \left[\Pr\left(r_{it}=r\right)\right]^2 & \text{if } l=r \\ -\Pr\left(r_{it}=r\right) \times \Pr\left(r_{it}=l\right) & \text{otherwise} \end{Bmatrix} \times \begin{Bmatrix} s_{j,it} & \text{if } j \in l \\ 0 & \text{otherwise} \end{Bmatrix}.$$

The functional form of $\frac{\partial}{\partial \pi_{j,it}} \partial E\left(s_{k,it} \middle| r_{it}=r\right)$ depends on the considered regime. For the complete regime, notably, we have:

$$(26)\quad \begin{aligned} &\frac{\partial}{\partial \pi_{j,it}} E\left(s_{k,it} \middle| r_{it}=r\right) \\ &= s_{k,it} \begin{cases} \alpha_{g,n,i} + (\alpha_{g,i} - \alpha_{g,n,i}) s_{k|\mathcal{K}_{gn,it}} + (\alpha_i - \alpha_{g,i}) s_{k|\mathcal{K}_{gn,it}} s_{\mathcal{K}_{gn}|\mathcal{N}_{g,it}} - \alpha_i s_{k,it} & \text{if } j=k \\ (\alpha_{g,i} - \alpha_{g,n,i}) s_{j|\mathcal{K}_{gn,it}} + (\alpha_i - \alpha_{g,i}) s_{j|\mathcal{K}_{gn,it}} s_{\mathcal{K}_{gn}|\mathcal{N}_{g,it}} - \alpha_i s_{j,it} & \text{if } \begin{cases} h=g\ , \\ m=n \\ \text{and } j \neq k \end{cases} \\ (\alpha_i - \alpha_{g,i}) s_{j|\mathcal{K}_{gm,it}} s_{\mathcal{K}_{gm}|\mathcal{N}_{g,it}} - \alpha_i s_{j,it} & \text{if } \begin{cases} h=g\ , \\ m \neq n \end{cases} \\ -\alpha_i s_{j,it} & \text{if } h \neq g \end{cases} \end{aligned}$$

The corresponding formula for each regime can easily be deduced from equation (26).

The average acreage own price elasticities of our farm sample are reported in Table 10. They have expected signs and, because of the crop disaggregation level of our data, are larger than those commonly found in the literature. The decomposition of these elasticities shows that a large part of the price effects on acreages can be due to the inclusion or not of these crops in the production regimes chosen by farmers. For crops like corn or peas with small overall acreage shares, changes in the production regimes account for about one third of the estimated price elasticities. However, changes in the production regimes can also be substantial for frequently produced crops. For instance, they account for a quarter of the sugar beet acreage own price elasticities.

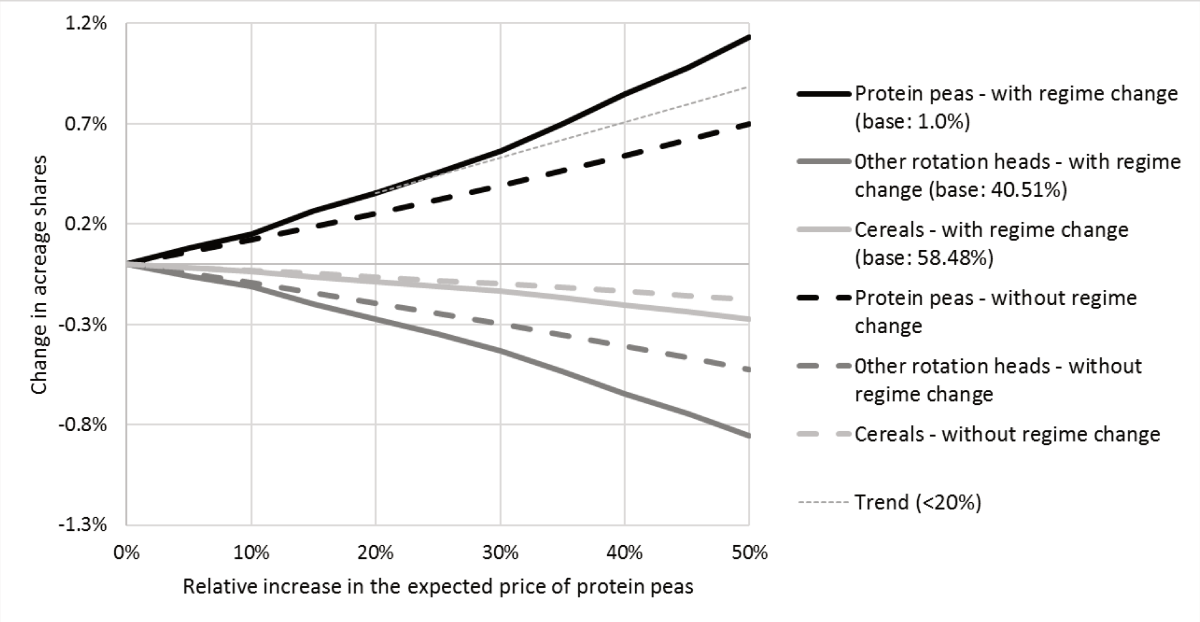
Table 10. Own price elasticities of acreages computed at sample average

	Winter cereals	*Corn*	*Spring Barley*	*Sugarbeet*	*Alfalfa*	*Protein peas*	*Rape Seed*
Overall elasticity	0.334	4.26	0.44	1.39	0.74	1.22	0.76
Part of the elasticity due to:							
Change in acreage within regime	0.332	2.33	0.43	1.236	0.598	0.714	0.753
Change in production regime	0.002	1.932	0.009	0.157	0.146	0.508	0.014

The impact of the production regime choice is further highlighted by simulating the effects of increases in the price of peas on the acreages of the crop. Owing to their fixing atmospheric nitrogen for themselves as well as for their following crops in the rotation this crop is often considered as 'diversification crop' of significant interest. Yet, protein peas acreages have declined over the last decade in the considered area mostly because of lacking profitability, as regards to that of the other rotation heads in particular.[82] The simulated impacts of increases in the price of peas on acreages are reported on Figure 2.

[82] In other parts of France the extension of a soil infection (*aphanomyces*) severely impacts peas yields and explains the decrease peas acreages. The diversified cropping systems used in *La Marne* seems to limit this extension.

Figure 2. Simulated impacts of changes in the price of peas on acreage shares

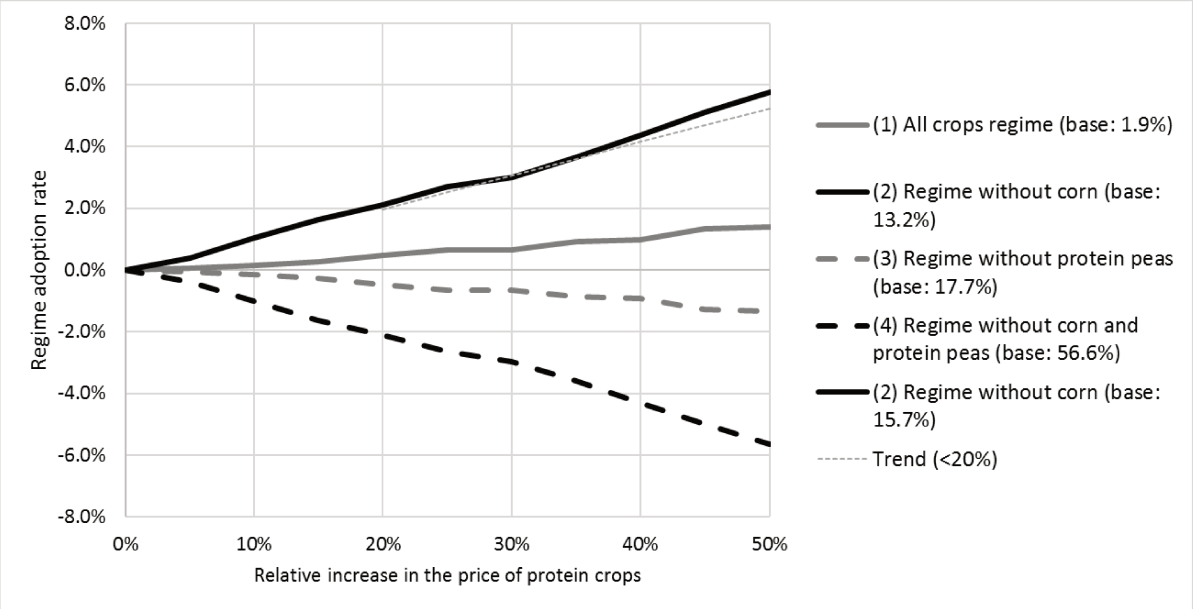


According to our results, a 50% increase in the price of peas would increase the average peas acreage share by 1.2%, from 1.0% to 2.2%. These additional peas acreages would mainly replace those of other rotation heads: the average combined acreage share of rapeseed, alfalfa and sugar beet would decrease by around 0.85% while that of cereals would only decrease by around 0.3%. This illustrates the interests in considering the crop – agronomic and management – characteristics when specifying the acreage management cost function. This also suggests that the increase in the rapeseed price due to the EU support to bio-fuels has played significant role in the decrease in the peas acreages in the considered area.

Interestingly, about two thirds of the increase in the peas acreage would be due to new producers. This also explains another feature of our simulation results. The simulated increases in the peas acreage in not linear in the price of peas: in particular, the increase in the peas acreages is more pronounced above the 20% price increase than below. Figure 3 shows that the

adoption rates of the production regimes including protein peas, regime 2 in particular, have similar patterns. This is partly explained by the threshold effects generated by the production regime fixed costs.

Figure 3. Simulated impacts of changes in the price of peas on regime adoption rates



Concluding remarks

The main aims of this article are twofold. First, it presents an original modelling framework for dealing with corner solutions in multi-crop micro-econometric models. This framework is based on the ERS approach, implying that it is fully consistent from an economic viewpoint. It also explicitly considers regime fixed costs. These features make the proposed ERS multi-crop micro-econometric models suitable for analyzing, and to some extent for disentangling, the effects of the main drivers of farmers' acreage choices at disaggregation levels at which the corner solution issue is pervasive. For instance, our estimation and simulation results and the

structure of the considered model tend to demonstrate the expected crop returns are not the sole significant drivers of farmers' acreage choices, at least in the short run.

Second, the application presented in this article illustrates the empirical tractability of ERS models, of random parameter ERS models in particular, for investigating farmers' production choices. Of course, to estimate such models raises challenging issues. But, this is also necessary for estimating structured micro-econometric models suitably accounting for important features characterizing micro-economic agricultural production data, among which significant unobserved heterogeneity. In particular, to estimate such models enables analysts to calibrate simulation models consisting of samples of farm specific models. According to our experience, ML estimators computed with SEM or MCEM algorithms combining simulation methods appear to be interesting alternatives to Simulated ML estimators for relatively large systems of interrelated equations such as the random parameter ERS models considered in our empirical application.

Of course, significant specification and estimation issues remain to be addressed for ERS multi-crop micro-econometric models such as ours to meet the needs of the agricultural production economist community. However, as fostering crop diversification tend to become an important agri-environmental objective in many countries, including those of the European Union, the modelling framework proposed in this article can be seen as a first step in the right direction.

Conclusion Générale

Les travaux présentés dans cette thèse proposent différentes améliorations des modèles économétriques de choix de production agricole en se basant sur le modèle multiculture de forme logit multinomial proposé par Carpentier et Letort (2014). Ce modèle, qui intègre des éléments agro-environnementaux, permet de représenter les comportements de court terme des agriculteurs dans un cadre statique. Les décisions modélisées sont les choix d'assolement, d'allocation d'intrants variables et de rendement. Les améliorations que nous proposons concernent spécifiquement la prise en compte de l'hétérogénéité inobservée des agriculteurs et de leurs conditions de production, et le traitement des problèmes de solution en coin dans les choix d'assolement, en cohérence avec la théorie micro-économique de la production agricole. Nous mobilisons pour cela des approches de spécification issues d'autres champs de la micro-économétrie appliquée que l'économie de la production agricole et des approches d'estimation issues de la statistique computationnelle. Nous contribuons aussi, de façon plus large, à la littérature en adaptant ces techniques à des modèles beaucoup plus complexes que ceux généralement étudiés. Deux applications de ces techniques sont présentées dans la thèse, et les résultats obtenus montrent que ces améliorations sont nécessaires pour avoir des modèles plus fins et ainsi apporter des solutions cohérentes aux nouvelles questions posées aux économistes agricoles

Nos travaux sur la prise en compte de l'hétérogénéité inobservée des agriculteurs et de leurs conditions de production sont, à notre connaissance, les premiers à intégrer les effets de cette hétérogénéité dans un modèle multicultures, aussi bien au niveau des termes additifs des équations (les constantes du modèle) que sur les paramètres de flexibilité de choix de production (les coefficients des variables explicatives), via la spécification de modèles à paramètres aléatoires.

L'approche d'estimation, basée sur l'algorithme EM, que nous proposons est utilisée pour

estimer une version à paramètre aléatoire d'un modèle de choix de production multiculture de Carpentier et Letort (2014) sur un échantillon de producteurs de la Meuse spécialisés en grandes cultures. Cette application illustre l'intérêt de notre approche pour prendre en compte de façon flexible les effets de l'hétérogénéité inobservée des agriculteurs, et pour fournir les valeurs des paramètres de chaque exploitation à des fins de simulation. Nos résultats soulignent l'importance de la prise en compte de l'hétérogénéité inobservée dans les modèles micro-économétriques de choix de production agricole. Nos résultats d'estimation et de simulation montrent en effet que les agriculteurs réagissent de façon hétérogène aux incitations économiques, même dans une région géographique relativement petite telle que le département de la Meuse. Nos résultats de simulation sur les effets du soutien de l'Union Européenne au bio-carburant montrent en particulier que l'utilisation des modèles usuels, c'est-à-dire les modèles à paramètres fixes et semi-fixes (effets fixes individuels additivement séparables) peuvent sous-estimer ou surestimer les effets simulés.

Nos résultats empiriques et nos interprétations ne peuvent être généralisés sans plus d'investigation. Ils suggèrent toutefois que l'hétérogénéité inobservée est un problème important à prendre en compte lorsqu'on cherche à analyser les décisions de production des agriculteurs, particulièrement lorsqu'on considère des échantillons couvrant des espaces géographiques relativement larges, donc caractérisés par des qualités de sol et/ou des conditions climatiques hétérogènes ne pouvant être prises en compte par des variables de contrôle disponibles dans les données.

Le modèle de choix d'assolement logit multinomial emboîté que nous considérons est similaire aux modèles de choix d'assolement fondés sur des choix discrets de culture à la parcelle. Nos résultats suggèrent donc qu'adopter une approche à paramètres aléatoires dans les modèles de choix discret à la parcelle pourrait être utile, ce qui est confirmé par les travaux très récents de Claassen *et al* (2017).

Nos travaux sur la prise en compte de solutions en coin dans les modèles de choix d'assolement utilisent une approche originale et cohérente d'un point de vue micro-économique. Cette approche est basée sur un modèle micro-économétrique à changement de régimes endogènes. A chaque régime on associe un coût fixe permettant de tenir compte notamment des coûts de commercialisation et/ou des contraintes spécifiques à la production de chaque régime. Le modèle micro-économétrique multiculture à changement de régime endogène que nous proposons permet d'étudier les effets des différents déterminants de choix d'assolement et de cultures à un niveau désagrégé. Dans l'application que nous proposons, une version à paramètres aléatoires de ce modèle est estimée sur un échantillon de producteurs de grandes cultures de la Marne. Nos résultats d'estimation et de simulation montrent que les coûts fixes associés aux régimes jouent un rôle important dans le choix des agriculteurs de produire ou non certaines cultures et que les marges brutes espérés ne sont pas les seuls déterminants (importants) des choix d'assolement. Les coûts fixes associés aux régimes constituent aussi un déterminant très important, du moins à court terme.

L'estimation dans un cadre paramétrique des modèles micro-économétriques considérés dans ces travaux engendre des problèmes d'estimation spécifiques, à savoir des problèmes d'intégration de la vraisemblance du modèle, de simulation de la loi conditionnelle de la vraisemblance complète, d'intégration dans le calcul des probabilités de choix de régime et du traitement des performances non observées des agriculteurs pour les cultures qu'ils ont choisi de ne pas produire. Nous proposons des solutions pour tenir compte de ces problèmes majeurs. Ces solutions sont basées sur les techniques de calcul d'estimateurs du maximum de vraisemblance *via* les versions stochastiques de l'algorithme EM, les algorithmes SAEM de Delyon *et al* (1999) en particulier. Combinées avec les méthodes d'échantillonnage préférentiel

ces techniques permettent d'estimer des modèles dont la vraisemblance est très compliquée, des modèles multivariés à paramètres aléatoires et à changement de régime endogène par exemple.

De manière plus générale, ces techniques offrent des alternatives très intéressantes aux estimateurs du maximum de vraisemblance simulée que les économètres emploient habituellement.

Tout exercice de modélisation micro-économétrique repose sur arbitrage entre d'une part la richesse et le « réalisme » des modèles spécifiés, et d'autre part la difficulté liée à l'estimation concrète de ces modèles. Nous avons choisi l'option de privilégier la spécification du modèle, quitte à devoir mettre en œuvre des méthodes d'estimation complexes et caractérisées par des temps de calculs importants. En effet, imposer des restrictions ou accepter des approximations bien choisies – *e.g.* ne pas introduire de variables de contrôle dans les modèles, approximer certains aspects de la spécification du modèle probabiliste – aurait énormément simplifié nos procédures d'estimation.

Il est cependant important de mentionner que le codage et la mise en œuvre de ces procédures est très chronophage. Certains problèmes purement numériques doivent être gérés. Par exemple, les problèmes d'*overflow* liés au calcul de fonctions exponentielles nécessitent la mise en place de procédures spécifiques. Ensuite, les estimations reportées dans les chapitres 5 et 6 nécessitent des temps de calcul variant de quelques heures à quelques jours selon le nombre de simulations employées. Ces temps de calculs « réduits » ont été obtenus en mettant en œuvre les estimations sur une machine multi-processeurs dédiée et en parallélisant les calculs. Ces temps de codage et de calcul réduisent les possibilités de recherche de spécification sur le modèle, ce qui en retour nuit à l'interprétation des estimations obtenues.

Les constats décrits ci-dessus suggèrent qu'il serait très utile de comparer nos estimations et

nos temps de calcul avec ceux relatifs à des spécifications simplifiées des modèles considérés ici. Ceci permettrait d'évaluer l'impact de certaines hypothèses de modélisation. Ceci permettrait également de rendre certaines de nos procédures d'estimation plus opérationnelles pour d'autres travaux de micro-économétrie appliquée. Néanmoins, et comme l'illustrent les points discutés dans la suite, enrichir la spécification d'un modèle est toujours tentant.

Bien sûr, nos études empiriques et nos cadres de modélisation présentent certaines limites qui ouvrent de nouvelles perspectives de recherche.

Tout d'abord, l'approche à paramètres aléatoires dans un cadre entièrement paramétrique a un défaut majeur : elle repose sur l'utilisation d'un grand volume de paramètres à estimer. C'est d'autant plus vrai dans le cadre nos travaux que nous avons choisi de spécifier nos modèles en imposant le moins de restrictions possibles sur leurs paramètres de façon à avoir des modèles les plus généraux possibles de manière à prendre en compte de manière flexible l'hétérogénéité inobservée de nos données. Ayant choisi des paramètres aléatoires et des termes d'erreurs suivant des lois normales multivariées, une grande partie des paramètres que nous estimons sont des paramètres de covariance. Une solution pour remédier à ce problème serait de combiner l'approche à paramètres aléatoires à une approche à indices latents. Cela revient à introduire, non plus des effets additifs aléatoires dans le modèle, mais des fonctions linéaires de variables latentes propres aux sous-groupes de culture, ces sous-groupes de cultures étant définis a priori sur la base de considérations agronomiques. Le nombre de sous-groupes étant inférieur au nombre de cultures considérées, cette approche permet de réduire la taille des matrices de variance-covariance à estimer. Il apparait relativement aisé de définir des variables latentes ayant du sens d'un point de vue agronomique dans le cas des équations de rendement et d'utilisation intrants. Dans le cas des rendements, la variable latente pourrait par exemple correspondre à « un indice latent de productivité propre à chaque exploitation et dans le cas des

équations d'intrants variables à un indice d'intensité des pratiques culturales. Ceci n'est toutefois pas aussi simple dans le cas des équations de choix d'assolements. En effet les paramètres de ces équations sont sensés capter les effets des (très) nombreux facteurs inobservés impactant les choix d'assolement des agriculteurs. Ces facteurs étant très variés, il semble difficile de leur donner une structure *a priori*.

Ensuite, l'approche entièrement paramétrique que nous utilisons est une limite de nos modèles à paramètres aléatoires. Toutefois, ne pas spécifier complètement la distribution des paramètres aléatoires et des termes d'erreur dans nos modèles engendrerait d'autres problèmes d'estimation beaucoup plus compliqués à résoudre. Une solution « à mi-chemin » entre l'approche paramétrique employée ici et l'approche non paramétrique serait d'utiliser des distributions de probabilité flexibles, telles que les distributions multivariées construites à partir des copules.

Enfin la prise en compte des effets de rotation et de la dynamique des choix de production constitue une voie de recherche potentielle très utile, mais également très difficile. Les effets des rotations sont les clés de voûte des systèmes de production agro-écologique dont de nombreux pays cherchent à promouvoir l'utilisation. Les approches empiriques actuellement proposées dans la littérature tiennent en effet compte de ces effets de façon *ad hoc*, et souvent peu cohérente d'un point de vue agronomique (voir, *e.g.*, Carpentier *et a*l, 2015). D'un point de vue théorique, tenir compte des effets des rotations à partir des modèles proposés dans le cadre de cette thèse est possible mais ce nécessite de résoudre des problèmes d'estimation spécifiques et difficiles, et certainement à partir d'approches très lourdes en matière en calcul (Carpentier et Gohin, 2014 ; 2015).

Bibliographie

Bibliographie

Ackerberg, D., Benkard, C.L., Berry, S., and Pakes, A. (2007). Econometric tools for analyzing market outcomes. Handbook of Econometrics 6, 4171–4276.

Allassonnière, S., Kuhn, E., and Trouvé, A. (2010). Construction of Bayesian deformable models via a stochastic approximation algorithm: a convergence study. Bernoulli 16, 641–678.

Allenby, G.M., Arora, N., and Ginter, J.L. (1998). On the heterogeneity of demand. Journal of Marketing Research 384–389.

Angrist, J.D., and Pischke, J.-S. (2008). Mostly harmless econometrics: An empiricist's companion (Princeton university press).

Arellano, M., and Bonhomme, S. (2011a). Identifying distributional characteristics in random coefficients panel data models. The Review of Economic Studies 79, 987–1020.

Arellano, M., and Bonhomme, S. (2011b). Nonlinear panel data analysis. Annu. Rev. Econ. 3, 395–424.

Arndt, C. (1999). Demand for herbicide in corn: An entropy approach using micro-level data. Journal of Agricultural and Resource Economics 204–221.

Arndt, C., Liu, S., and Preckel, P.V. (1999). On dual approaches to demand systems estimation in the presence of binding quantity constraints. Applied Economics 31, 999–1008.

Arrow, K.J., and Enthoven, A.C. (1961). Quasi-concave programming. Econometrica: Journal of the Econometric Society 779–800.

Atanassov, E.I. (2004). On the discrepancy of the Halton sequences. Math. Balkanica (NS) 18, 15–32.

Atanassov, E.I., and Durchova, M.K. (2002). Generating and testing the modified Halton sequences. (Springer), pp. 91–98.

Bayramoglu, B., and Chakir, R. (2016). The impact of high crop prices on the use of agro-chemical inputs in France: A structural econometric analysis. Land Use Policy 55, 204–211.

Beran, R., Feuerverger, A., and Hall, P. (1996). On nonparametric estimation of intercept and slope distributions in random coefficient regression. The Annals of Statistics 24, 2569–2592.

Berry, S.T. (1994). Estimating discrete-choice models of product differentiation. The RAND Journal of Economics 242–262.

Bhat, C.R. (2001). Quasi-random maximum simulated likelihood estimation of the mixed multinomial logit model. Transportation Research Part B: Methodological 35, 677–693.

Bhat, C.R. (2003). Simulation estimation of mixed discrete choice models using randomized and scrambled Halton sequences. Transportation Research Part B: Methodological 37, 837–855.

Booth, J.G., and Hobert, J.P. (1998). Standard errors of prediction in generalized linear mixed models. Journal of the American Statistical Association 93, 262–272.

Booth, J.G., and Hobert, J.P. (1999). Maximizing generalized linear mixed model likelihoods with an automated Monte Carlo EM algorithm. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 61, 265–285.

Boyd, S., and Vandenberghe, L. (2004). Convex optimization (Cambridge university press).

Breitung, J., and Pesaran, M.H. (2008). Unit roots and cointegration in panels. In The Econometrics of Panel Data, (Springer), pp. 279–322.

Buchecker, M.M. Pipeline Parallelism Performance Practicalities.

Caffo, B.S., Jank, W., and Jones, G.L. (2005). Ascent-based Monte Carlo expectation–maximization. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 67, 235–251.

Cappé, O., Douc, R., Guillin, A., Marin, J.-M., and Robert, C.P. (2008). Adaptive importance sampling in general mixture classes. Statistics and Computing 18, 447–459.

Carpentier, A., and Gohin, A. (2014). Accounting for crop rotations in acreage choice modeling: a tractable modeling framework. (European Association of Agricultural Economists), p.

Carpentier, A., and Gohin, A. (2015). On the economic theory of crop rotations: value of the crop rotation effects and implications on acreage choice modeling. On the Economic Theory of Crop Rotations: Value of the Crop Rotation Effects and Implications on Acreage Choice Modeling (2015).

Carpentier, A., and Letort, E. (2011). Accounting for heterogeneity in multicrop micro-econometric models: implications for variable input demand modeling. American Journal of Agricultural Economics 94, 209–224.

Carpentier, A., and Letort, E. (2014). Multicrop production models with Multinomial Logit acreage shares. Environmental and Resource Economics 59, 537–559.

Carpentier, A., Gohin, A., Sckokai, P., and Thomas, A. (2015). Economic modelling of agricultural production: past advances and new challenges. Revue d'Études En Agriculture et Environnement 96, 131–165.

Chakir, R., Bousquet, A., and Ladoux, N. (2004). Modeling corner solutions with panel data: Application to the industrial energy demand in France. Empirical Economics 29, 193–208.

Chambers, R.G., and Just, R.E. (1989). Estimating multioutput technologies. American Journal of Agricultural Economics 71, 980–995.

Chang, H.-H., and Stout, W. (1993). The asymptotic posterior normality of the latent trait in an

IRT model. Psychometrika 58, 37–52.

Chavas, J.-P., and Holt, M.T. (1990). Acreage decisions under risk: the case of corn and soybeans. American Journal of Agricultural Economics 72, 529–538.

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

Cherchi, E., and Guevara, C.A. (2012). A Monte Carlo experiment to analyze the curse of dimensionality in estimating random coefficients models with a full variance–covariance matrix. Transportation Research Part B: Methodological 46, 321–332.

Claassen, R., Langpap, C., and Wu, J. (2017). Impacts of Federal Crop Insurance on Land Use and Environmental Quality. American Journal of Agricultural Economics 99, 592–613.

Costinot, A., Donaldson, D., and Smith, C. (2016). Evolving comparative advantage and the impact of climate change in agricultural markets: Evidence from 1.7 million fields around the world. Journal of Political Economy 124, 205–248.

Delyon, B., Lavielle, M., and Moulines, E. (1999). Convergence of a stochastic approximation version of the EM algorithm. Annals of Statistics 94–128.

Dempster, A.P., Laird, N.M., and Rubin, D.B. (1977). Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society. Series B (Methodological) 1–38.

Douc, R., Guillin, A., Marin, J.-M., and Robert, C.P. (2007). Minimum variance importance sampling via population monte carlo. ESAIM: Probability and Statistics 11, 427–447.

Eaton, J., Kortum, S., and Kramarz, F. (2011). An anatomy of international trade: Evidence from French firms. Econometrica 79, 1453–1498.

Fezzi, C., and Bateman, I.J. (2011). Structural agricultural land use modeling for spatial agro-environmental policy analysis. American Journal of Agricultural Economics 93, 1168–1188.

Fezzi, C., Bateman, I., Askew, T., Munday, P., Pascual, U., Sen, A., and Harwood, A. (2014). Valuing provisioning ecosystem services in agriculture: the impact of climate change on food production in the United Kingdom. Environmental and Resource Economics 57, 197–214.

Gelman, A., Carlin, J.B., Stern, H.S., and Rubin, D.B. (2003). Bayesian Data Analysis, (Chapman & Hall/CRC Texts in Statistical Science).

Geweke, J. (1991). Efficient simulation from the multivariate normal and student-t distributions subject to linear constraints and the evaluation of constraint probabilities. (Citeseer), pp. 571–578.

Gohin, A., Carpentier, A., Koutchade, O.P., and Bareille, F. (2015). Amélioration de la représentation de l'offre agricole dans les modèles macroéconomiques (Rapport, Agence de L'Environnement et de la Maîtrose de L'Energie).

Gourieroux, C., and Monfort, A. (1990). Simulation based inference in models with heterogeneity. Annales d'Economie et de Statistique 69–107.

Gourieroux, C., and Monfort, A. (1993). Simulation-based inference: A survey with special reference to panel data models. Journal of Econometrics 59, 5–33.

Greene, W. (2005). Reconsidering heterogeneity in panel data estimators of the stochastic frontier model. Journal of Econometrics 126, 269–303.

Greene, W.H. (2004). Interpreting estimated parameters and measuring individual heterogeneity in random coefficient models.

Greene, W.H. (2008). The econometric approach to efficiency analysis. The Measurement of Productive Efficiency and Productivity Growth 1, 92–250.

Greene, W.H., and Hensher, D.A. (2003). A latent class model for discrete choice analysis: contrasts with mixed logit. Transportation Research Part B: Methodological 37, 681–698.

Guyomard, H., Baudry, M., and Carpentier, A. (1996). Estimating crop supply response in the presence of farm programmes: application to the CAP. European Review of Agricultural Economics 23, 401–420.

Hajivassiliou, V. (2000). Some practical issues in maximum simulated likelihood (Cambridge University Press).

Hajivassiliou, V.A., and Ruud, P.A. (1994). Classical estimation methods for LDV models using simulation. Handbook of Econometrics 4, 2383–2441.

Halton, J.H. (1960). On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals. Numerische Mathematik 2, 84–90.

Harding, M.C., and Hausman, J. (2007). Using a Laplace approximation to estimate the random coefficients logit model by nonlinear least squares. International Economic Review 48, 1311–1328.

Heckelei, T., and Wolff, H. (2003). Estimation of constrained optimisation models for agricultural supply analysis based on generalised maximum entropy. European Review of Agricultural Economics 30, 27–50.

Heckelei, T., Britz, W., and Zhang, Y. (2012). Positive mathematical programming approaches–recent developments in literature and applied modelling. Bio-Based and Applied Economics 1, 109–124.

Heckman, J.J. (1976). The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. In Annals of Economic and Social Measurement, Volume 5, Number 4, (NBER), pp. 475–492.

Heckman, J.J. (1977). Sample selection bias as a specification error (with an application to the estimation of labor supply functions).

Heckman, J.J. (1981). The incidental parameters problem and the problem of initial conditions

in estimating a discrete time-discrete data stochastic process.

Heckman, J.J. (2001). Micro data, heterogeneity, and the evaluation of public policy: Nobel lecture. Journal of Political Economy 109, 673–748.

Hendricks, N.P., Smith, A., and Sumner, D.A. (2014). Crop supply dynamics and the illusion of partial adjustment. American Journal of Agricultural Economics aau024.

Holt, M.T. (1999). A linear approximate acreage allocation model. Journal of Agricultural and Resource Economics 383–397.

Hoogerheide, L., Opschoor, A., and Van Dijk, H.K. (2012). A class of adaptive importance sampling weighted EM algorithms for efficient and robust posterior and predictive simulation. Journal of Econometrics 171, 101–120.

Jank, W. (2005). Quasi-Monte Carlo sampling to improve the efficiency of Monte Carlo EM. Computational Statistics & Data Analysis 48, 685–701.

Jank, W., and Booth, J. (2003). Efficiency of Monte Carlo EM and simulated maximum likelihood in two-stage hierarchical models. Journal of Computational and Graphical Statistics 12, 214–229.

Ji, Y., Rabotyagov, S., and Kling, C.L. (2014). Crop Choice and Rotational Effects: A Dynamic Model of Land Use in Iowa in Recent Years. (Agricultural and Applied Economics Association), p.

Just, R.E., and Pope, R.D. (2001). The agricultural producer: theory and statistical measurement. Handbook of Agricultural Economics 1, 629–741.

Just, R.E., and Pope, R.D. (2003). Agricultural risk analysis: adequacy of models, data, and issues. American Journal of Agricultural Economics 85, 1249–1256.

Just, R.E., Zilberman, D., and Hochman, E. (1983). Estimation of multicrop production functions. American Journal of Agricultural Economics 65, 770–780.

Kao, C., Lee, L., and Pitt, M.M. (2001). Simulated maximum likelihood estimation of the linear expenditure system with binding non-negativity constraints. Annals of Economics and Finance 2, 203–223.

Keane, M.P. (2009). Simulated maximum likelihood estimation besed on first-order conditions. International Economic Review 50, 627–675.

Kiefer, J., and Wolfowitz, J. (1952). Stochastic estimation of the maximum of a regression function. The Annals of Mathematical Statistics 23, 462–466.

Koutchadé, O. P., Carpentier, A., and Femenia, F. (2017a)Modeling heterogeneous farm responses to EU biofuel support with a random parameter multicrop model. American Journal of Agricultural Economics (forthcoming)

Koutchade, P., Carpentier, A., and Femenia, F. (2017b). Dealing with corner solutions in multi-crop micro-econometric models: an endogenous regime switching approach with regime fixed costs. XV EAAE Congress 2017

Koutchadé, P., Carpentier, A. and Femenia, F. (2015). Empirical modeling of production decisions of heterogeneous farmers with random parameter models. SMART-LERECO Working Paper WP15-10

Lacroix, A., and Thomas, A. (2011). Estimating the environmental impact of land and production decisions with multivariate selection rules and panel data. American Journal of Agricultural Economics 93, 784–802.

Lange, K. (1995). A quasi-Newton acceleration of the EM algorithm. Statistica Sinica 1–18.

Langpap, C., and Wu, J. (2011). Potential environmental impacts of increased reliance on corn-based bioenergy. Environmental and Resource Economics 49, 147–171.

Lavielle, M. (2014). Mixed effects models for the population approach: models, tasks, methods and tools (CRC press).

Lee, L.-F. (1978). Unionism and wage rates: A simultaneous equations model with qualitative and limited dependent variables. International Economic Review 415–433.

Lee, L.-F., and Pitt, M.M. (1986). Microeconometric demand system with binding nonnegativity constraints: the dual approach. Econometrica: Journal of the Econometric Society 1237–1242.

Letort, E. (2009). Modélisation micro-économétrique des choix de production des agriculteurs.

Levine, R.A., and Casella, G. (2001). Implementations of the Monte Carlo EM algorithm. Journal of Computational and Graphical Statistics 10, 422–439.

Lichtenberg, E., and Smith-Ramirez, R. (2011). Slippage in conservation cost sharing. American Journal of Agricultural Economics 93, 113–129.

Lindsay, B.G., and Lesperance, M.L. (1995). A review of semiparametric mixture models. Journal of Statistical Planning and Inference 47, 29–39.

Livingston, M., Roberts, M.J., and Zhang, Y. (2015). Optimal sequential plantings of corn and soybeans under price uncertainty. American Journal of Agricultural Economics 97, 855–878.

Louhichi, K., and Valin, H. (2012). Impact of EU biofuel policies on the French arable sector: A micro-level analysis using global market and farm-based supply models. Revue d'Etudes En Agriculture et Environnement-Review of Agricultural and Environmental Studies 93, 233–272.

Maddala, G.S., and Nelson, F.D. (1975). Switching regression models with exogenous and endogenous switching. (Business and Economics Section), pp. 423–426.

Mannering, F.L., Shankar, V., and Bhat, C.R. (2016). Unobserved heterogeneity and the statistical analysis of highway accident data. Analytic Methods in Accident Research 11, 1–16.

McCulloch, C.E. (1994). Maximum likelihood variance components estimation for binary data. Journal of the American Statistical Association 89, 330–335.

McCulloch, C.E. (1997). Maximum likelihood algorithms for generalized linear mixed models. Journal of the American Statistical Association 92, 162–170.

McDonald, J.F., and Moffitt, R.A. (1980). The uses of Tobit analysis. The Review of Economics and Statistics 318–321.

McFadden, D. (1989). A method of simulated moments for estimation of discrete response models without numerical integration. Econometrica: Journal of the Econometric Society 995–1026.

McLachlan, G., and Krishnan, T. (2007). The EM algorithm and extensions (John Wiley & Sons).

McLachlan, G., and Peel, D. (2004). Finite mixture models (John Wiley & Sons).

Meng, X.-L., and Rubin, D.B. (1993). Maximum likelihood estimation via the ECM algorithm: A general framework. Biometrika 80, 267–278.

Moore, M.R., and Negri, D.H. (1992). A multicrop production model of irrigated agriculture, applied to water allocation policy of the Bureau of Reclamation. Journal of Agricultural and Resource Economics 29–43.

Murtazashvili, I., and Wooldridge, J.M. (2008). Fixed effects instrumental variables estimation in correlated random coefficient panel data models. Journal of Econometrics 142, 539–552.

Murtazashvili, I., and Wooldridge, J.M. (2016). A control function approach to estimating switching regression models with endogenous explanatory variables and endogenous switching. Journal of Econometrics 190, 252–266.

Perali, F., and Chavas, J.-P. (2000). Estimation of censored demand equations from large cross-section data. American Journal of Agricultural Economics 82, 1022–1037.

Platoni, S., Sckokai, P., and Moro, D. (2012). Panel data estimation techniques and farm-level data models. American Journal of Agricultural Economics 94, 1202–1217.

Robbins, H., and Monro, S. (1951). A stochastic approximation method. The Annals of Mathematical Statistics 400–407.

Ruud, P.A. (1991). Extensions of estimation methods using the EM algorithm. Journal of Econometrics 49, 305–341.

Savic, R.M., and Karlsson, M.O. (2009). Importance of shrinkage in empirical bayes estimates for diagnostics: problems and solutions. The AAPS Journal 11, 558–569.

Sckokai, P., and Moro, D. (2006). Modeling the reforms of the common agricultural policy for arable crops under uncertainty. American Journal of Agricultural Economics 88, 43–56.

Sckokai, P., and Moro, D. (2009). Modelling the impact of the CAP Single Farm Payment on

farm investment and output. European Review of Agricultural Economics 36, 395–423.

Shamlin, D. (2004). Threads unraveled: a parallel processing primer. In SUGI, 29, pp. 217–228.

Shonkwiler, J.S., and Yen, S.T. (1999). Two-step estimation of a censored system of equations. American Journal of Agricultural Economics 81, 972–982.

Stanfield, P.M., Wilson, J.R., Mirka, G.A., Glasscock, N.F., Psihogios, J.P., and Davis, J.R. (1996). Multivariate input modeling with Johnson distributions. (IEEE Computer Society), pp. 1457–1464.

Swamy, P.A. (1970). Efficient inference in a random coefficient regression model. Econometrica: Journal of the Econometric Society 311–323.

Thomas, N. (1993). Asymptotic corrections for multivariate posterior moments with factored likelihood functions. Journal of Computational and Graphical Statistics 2, 309–322.

Thomas, N., and Gan, N. (1997). Generating multiple imputations for matrix sampling data analyzed with item response models. Journal of Educational and Behavioral Statistics 22, 425–445.

Train, K. (2007). A recursive estimator for random coefficient models. University of California, Berkeley.

Train, K.E. (2009). Discrete choice methods with simulation (Cambridge university press).

Von Davier, M., and Sinharay, S. (2007). An importance sampling EM algorithm for latent regression models. Journal of Educational and Behavioral Statistics 32, 233–251.

Wales, T.J., and Woodland, A.D. (1983). Estimation of consumer demand systems with binding non-negativity constraints. Journal of Econometrics 21, 263–285.

Wei, G.C., and Tanner, M.A. (1990). A Monte Carlo implementation of the EM algorithm and the poor man's data augmentation algorithms. Journal of the American Statistical Association 85, 699–704.

Wooldridge, J.M. (2005a). Fixed-effects and related estimators for correlated random-coefficient and treatment-effect panel data models. The Review of Economics and Statistics 87, 385–390.

Wooldridge, J.M. (2005b). Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity. Journal of Applied Econometrics 20, 39–54.

Wooldridge, J.M. (2010). Econometric analysis of cross section and panel data (MIT press).

Wu, C.J. (1983). On the convergence properties of the EM algorithm. The Annals of Statistics 95–103.

Wu, J. (2000). Slippage effects of the conservation reserve program. American Journal of Agricultural Economics 82, 979–992.

Wu, J., and Brorsen, B.W. (1995). The impact of government programs and land characteristics on cropping patterns. Canadian Journal of Agricultural Economics/Revue Canadienne d'agroeconomie 43, 87–104.

Wu, J., Adams, R.M., Kling, C.L., and Tanaka, K. (2004). From microlevel decisions to landscape changes: an assessment of agricultural conservation policies. American Journal of Agricultural Economics 86, 26–41.

Yen, S.T. (2005). Zero observations and gender differences in cigarette consumption. Applied Economics 37, 1839–1849.

Yen, S.T., and Lin, B.-H. (2006). A sample selection approach to censored demand systems. American Journal of Agricultural Economics 88, 742–749.

Zellner, A. (1962). An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. Journal of the American Statistical Association 57, 348–368.