

QU'EST-CE QU'UNE MICROSIMULATION DYNAMIQUE EN SCIENCES SOCIALES?

Martin Spielauer

Statistique Canada – Division de la modélisation

Immeuble R.-H.- Coats, 24-O

Ottawa, K1A 0T6

martin.spielauer@statcan.gc.ca

*« Perhaps consciousness arises when the brain's
simulation of the world becomes so complete that it must
include a model of itself. »*

Richard Dawkins - The Selfish Gene

Introduction

Alors que l'idée d'étudier et de projeter le développement socioéconomique et démographique d'une société en simulant un grand échantillon d'individus ainsi que leurs actions et interactions avait déjà été exprimée durant les années 50, la microsimulation dynamique n'a pas encore trouvé sa place dans la boîte à outils méthodologiques des spécialistes en sciences sociales traditionnels. La simulation réaliste d'une société requiert des données détaillées, des modèles compliqués, des ordinateurs rapides et des essais poussés. Plus les modèles deviennent complexes, plus il est difficile de comprendre leur fonctionnement et d'évaluer leur pouvoir prédictif. On pourrait présumer que la microsimulation est trop exigeante ou que les modèles de microsimulation sont des produits spécialisés ou des modèles à boîte noire douteux, qui ne sont applicables qu'avec prudence, quand il n'existe pas d'autres méthodes. Ici, toutefois, nous allons présenter un autre point de vue.

En premier lieu, la microsimulation est un outil puissant qui a déjà fait ses preuves dans des applications de complexité modérée pour lesquelles existent d'autres approches de modélisation n'ayant cependant pas la souplesse de l'approche par microsimulation.

En deuxième lieu, nous relevons (ou reconnaissons) de plus en plus souvent des défis socioéconomiques pour lesquels la microsimulation est le seul outil d'étude disponible. De surcroît, la microsimulation est une approche qui découle naturellement des paradigmes de recherche modernes et qui complète l'analyse détaillée des données.

En troisième lieu, la microsimulation est une approche dont l'heure est venue. Plus d'un demi-siècle après son introduction dans les sciences sociales, les grands obstacles à son utilisation ont pratiquement disparu. La puissance des ordinateurs a augmenté de manière exponentielle, la collecte de données longitudinales individuelles est devenue une activité de routine, les spécialistes des sciences sociales possèdent une formation en recherche longitudinale, tandis que la recherche proprement dite a évolué d'une macro à une microapproche et s'en va vers l'intégration multiniveaux. La perspective de la trajectoire de vie est devenue le paradigme dominant et la plupart des problèmes les plus pressants que nous devons résoudre sont d'une nature qui fait de la microsimulation dynamique la méthode d'étude la plus appropriée.

Un autre obstacle a également disparu aujourd'hui. Les langages de programmation, tels que Modgen, permettent maintenant aux chercheurs ne possédant que des compétences de programmation modestes — comparables à celles nécessaires pour se servir de logiciels statistiques — de mettre en œuvre leurs modèles.

Le présent chapitre offre une introduction à la microsimulation et expose les grandes idées sous-jacentes, ainsi que les points forts et les inconvénients de cette approche. Il comporte trois volets :

- premièrement, nous définissons la microsimulation dynamique dans le domaine des sciences sociales et en faisons brièvement l'historique;
- deuxièmement, nous explorons trois situations importantes où la microsimulation est une approche adéquate;
- troisièmement, nous soulignons les principaux avantages et inconvénients de la microsimulation. En ce qui concerne les avantages, nous décrivons ses points forts théoriques dans la perspective de la trajectoire de vie, ses points forts pratiques dans la perspective de l'élaboration des politiques et ses points forts techniques. En ce qui concerne les inconvénients et les limites, nous faisons la distinction entre les limites intrinsèques dues au caractère aléatoire des phénomènes et les limites plutôt transitoires dues la forte demande de données. Nous survolons également brièvement les questions liées aux ressources informatiques et d'autres questions techniques, quoique leurs coûts correspondants diminuent au cours du temps.

1 Définition de la microsimulation dynamique en sciences sociales

1.1 Qu'est-ce que la microsimulation?

Un moyen utile de définir la simulation dans le domaine des sciences sociales consiste à la percevoir comme l'utilisation intentionnelle d'un modèle. Par conséquent, si nous reculons d'un pas, la simulation en sciences sociales est à la fois un exercice de modélisation et l'exercice consistant à « exécuter le modèle », ou à « jouer » ou « expérimenter » avec celui-ci. La gamme des objectifs est aussi vaste que celle des raisons d'entreprendre la recherche : résoudre des problèmes, trouver des explications, bâtir une théorie, prédire l'avenir, et conscientiser. D'un point de vue plus pratique, nous pouvons également ajouter la formation à cette liste. Les pilotes s'entraînent sur des simulateurs de vol. Pourquoi les décideurs ne pourraient-ils pas s'entraîner au moyen de simulations informatiques à mieux saisir les effets des politiques? Et pourquoi les électeurs ne disposeraient-ils pas d'outils pour étudier les effets des mesures stratégiques proposées? La simulation en sciences sociales permet ce genre de vision.

La simulation dynamique fait intervenir le temps. Comment sommes-nous arrivés où nous en sommes aujourd'hui? À quels changements pouvons-nous nous attendre dans l'avenir? Qu'est-ce qui dicte ces changements et comment pouvons-nous influencer sur ces processus? La plupart des déclarations les mieux fondées au sujet de l'avenir s'appuient sur des simulations dynamiques. Certaines nécessitent des simulations informatiques complexes, tandis que d'autres sont le résultat d'expériences de la pensée. L'exploration de futurs scénarios et de la façon dont l'avenir est façonné par nos actes individuels est un accomplissement fondamental du cerveau humain, étroitement lié à la conscience proprement dite. Pouvoir prédire l'état futur d'un système nous permet de mieux planifier nos actions, aussi bien celles qui influencent le résultat du système que celles affectées par ce dernier. Par exemple, les prévisions météorologiques sont produites en s'appuyant sur des simulations informatiques complexes — et nous disposons à la fois de modèles relativement satisfaisants pour prédire le temps qu'il fera demain (sur lequel nous n'exerçons pas d'influence) et de modèles de simulation nettement plus controversés pour prédire les changements climatiques de long terme (sur lesquels nous pouvons influencer). La simulation dynamique sensibilise les membres du public aux problèmes qui pourraient survenir dans l'avenir, qu'il s'agisse d'une tempête demain ou de l'effet des émissions de CO₂ au cours du temps. La possibilité de prendre davantage conscience de certaines situations et de mieux planifier nos interventions existe aussi dans le domaine des sciences sociales pour des questions telles que le vieillissement de la population, la concentration de la richesse ou la viabilité des systèmes de sécurité sociale.

La microsimulation dynamique est un type particulier de simulation dynamique. Malheureusement, le terme microsimulation proprement dit peut prêter à confusion, parce que,

malgré le préfixe « micro », nous simulons un « macro » système. Le préfixe « micro » désigne essentiellement *comment* nous simulons ce système. Nombre de systèmes sont constitués d'unités de niveau plus faible. Les liquides sont constitués de particules dont le comportement est modifié quand on les chauffe, les systèmes de circulation comprennent des automobiles conduites sur un réseau de routes, et les sociétés et les économies sont formées de personnes, de ménages, d'entreprises, etc. Tous ces systèmes ont une caractéristique commune — les changements de niveau macro résultent des actions et interaction des unités de niveau micro. L'idée maîtresse de la microsimulation est que le meilleur moyen de simuler un tel système consiste souvent à modéliser et à simuler les actions et les interactions de ses unités de plus petite échelle et d'obtenir les macrorésultats par agrégation.

La microsimulation dynamique en sciences sociales peut se concevoir comme une expérimentation avec une société virtuelle comptant des milliers, voire des millions, d'individus que l'on crée et dont les trajectoires de vie se déroulent dans un ordinateur. Selon l'objectif du modèle, les individus (ou « acteurs ») font des choix d'études et de carrière, forment des unions, ont des enfants, travaillent, paient des impôts, reçoivent des avantages sociaux, divorcent, migrent, prennent leur retraite, reçoivent une pension et, éventuellement, meurent. La création d'un « jeu informatique scientifique » de ce genre comporte diverses étapes, la première étant la modélisation du comportement individuel. En microsimulation, les principales catégories de micromodèles sont les modèles statistiques et économétriques. Alors que la littérature est riche en analyses de microdonnées statistiques, la plupart des travaux de recherche s'arrêtent après l'estimation des modèles des processus individuels. Au moyen d'un modèle de microsimulation, nous allons une étape plus loin : la microsimulation ajoute la synthèse à l'analyse. Par conséquent, la deuxième étape de la microsimulation, après la modélisation des comportements individuels, est la programmation des divers modèles de comportement qui nous permettront d'exécuter les simulations du système complet. La microsimulation peut nous aider à comprendre l'effet de divers processus et de divers changements dans les processus sur le résultat global. Plus le nombre de processus interdépendants qu'il faut prendre en considération est élevé, plus il est difficile de cerner et de comprendre la contribution des facteurs individuels aux macrorésultats. Mais la microsimulation offre l'outil requis pour étudier ce genre de système.

La modélisation au niveau micro facilite la simulation des politiques. Les règlements en matière d'avantages fiscaux et de sécurité sociale sont définis au niveau individuel ou familial, si bien que la microsimulation une approche de modélisation naturelle permettant leur simulation à n'importe quel niveau de détail. Comme ce genre de règlements sont habituellement complexes et dépendent de manière non linéaire de diverses caractéristiques, telles que la composition de la famille ou le revenu (p.ex. impôts progressifs), la microsimulation est souvent le seul moyen d'étudier l'effet distributionnel et la viabilité de

long terme de ce genre de système. En analyse des politiques, une partie de la puissance de l'approche de microsimulation est déjà exploitée dans les modèles de microsimulation dits statiques. Il s'agit de modèles conçus pour étudier l'effet transversal des changements de politique, par exemple, en identifiant les gagnants et les perdants immédiats d'une réforme stratégique. Cependant, la microsimulation dynamique ajoute une toute nouvelle dimension à l'analyse des politiques, puisqu'elle permet de suivre les individus au cours de leur cycle de vie complet.

En sciences sociales, la microsimulation dynamique remonte à l'idée qu'a eue Guy Orcutt (1957) d'imiter les expériences naturelles en économie (Orcutt 1957). L'approche de modélisation qu'il a proposée correspond à ce que l'on peut appeler le courant empirique, ou guidé par les données, de modèles de microsimulation dynamique, c'est-à-dire des modèles conçus et utilisés au niveau opérationnel pour faire des prévisions et des recommandations stratégiques (Klevmarken 1997). Ce genre de microsimulation englobe les modèles microéconométriques et statistiques, ainsi que les routines comptables. Fait pendant à ce courant empirique, le courant ou la tradition théorique de la simulation orientée agents (SOA). Quoique qu'entrant dans la catégorie des modèles de microsimulation aux termes de notre définition générale, la SOA est souvent considérée comme une branche distincte de la simulation, différente de la microsimulation. Cette façon de voir tient principalement à l'objectif différent de la modélisation SOA (principalement l'exploration de théories) et les aux diverses approches utilisées dans la SOA pour modéliser le microcomportement (règles basées sur la théorie et l'intelligence artificielle). Sauf indication contraire, la présente discussion portera uniquement sur le groupe de modèles de microsimulation dynamique guidés par les données¹.

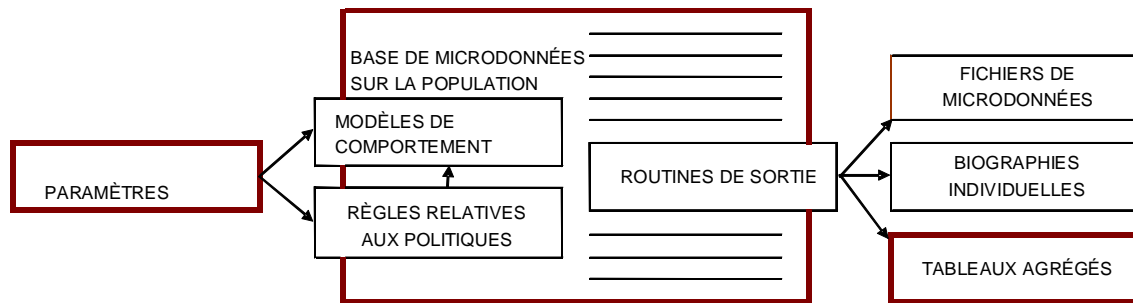
Le graphique qui suit (Figure 1) résume les principales composantes d'un modèle de microsimulation guidé par les données type. Au centre figure une base de microdonnées sur la population contenant les caractéristiques de tous les membres de cette dernière². Cette base de données est mise à jour dynamiquement durant une exécution de simulation conformément aux micromodèles de comportement et aux règles relatives aux politiques (telles que les règles relatives aux cotisations et aux prestations dans un modèle de pensions). Tous ces modèles peuvent être paramétrisés par l'utilisateur. Les résultats des simulations sont présentés sous

¹ Il convient toutefois de souligner que le langage Modgen a été utilisé avec succès pour la SOA, comme le décrit Wolfson (1999).

² Dans une perspective davantage orientée objets, la base de données sur la population peut aussi être considérée et mise en œuvre comme contenant des « cerveaux » individuels décentralisés avec des acteurs possédant des méthodes pour faire part de leurs états à un statisticien virtuel responsable de la collecte et de la présentation des données.

forme de tableaux agrégés produits par des routines de sortie. En outre, la sortie peut correspondre à des fichiers de microdonnées analysables par un logiciel statistique. Certains modèles (tels que ceux produits avec Modgen) permettent aussi de représenter graphiquement les biographies individuelles.

Figure 1 : Composantes d'un modèle de microsimulation dirigé par les données type



1.2 La vision d'Orcutt et la réalité d'aujourd'hui

La microsimulation dynamique a été introduite pour la première fois dans le domaine des sciences sociales en 1957 par Guy Orcutt dans son article fondamental intitulé « A new type of socio-economic system », dans lequel il proposait un nouveau type de modèle émanant principalement de sa frustration à l'égard des modèles de projection macroéconomique existants. Dans cet article, Orcutt examine divers défauts des modèles macroéconomiques que la microsimulation permet de surmonter. Le premier est l'« utilité prédictive limitée » des macromodèles, surtout en ce qui concerne les effets de l'intervention gouvernementale, parce qu'ils sont trop abstraits pour permettre des simulations fines des politiques. Le deuxième tient à la concentration sur les agrégats et à l'ignorance des aspects distributionnels dans les études macroéconomiques et les projections. Le troisième problème, selon Orcutt, est que les macromodèles ne tirent pas parti des connaissances existantes au sujet des unités décideuses élémentaires. La microsimulation, en revanche, n'est pas limitée par des hypothèses restrictives de « relations absurdes au sujet des unités décideuses élémentaires » afin de pouvoir procéder à l'agrégation. La modélisation au niveau où les décisions sont prises rend non seulement les modèles plus compréhensibles, mais, en cas de relations non linéaires, fait aussi que « les relations stables au niveau micro sont relativement en harmonie avec l'absence de relations stables au niveau agrégé ».

Alors que ces observations sont encore vraies un demi-siècle plus tard, certaines autres observations d'Orcutt sont de bons exemples de la façon dont les ordinateurs ont transformé la recherche. En fait, une part considérable de son article est consacrée à la justification de l'usage d'une puissance informatique coûteuse pour les simulations, fait qui était généralement considéré comme étant le domaine des mathématiciens et des solutions analytiques calculables

sur papier. Selon Orcutt, un des avantages de la microsimulation est qu'elle est « ... comprise par des personnes ne possédant qu'un modeste bagage mathématique ».

Alors que l'approche de modélisation qu'il proposait était visionnaire en 1957, à cause du manque de puissance informatique et de données à l'époque, peu après, Orcutt a été chargé de développer le premier modèle américain de microsimulation à grande échelle Dynasim. Plus tard, il a contribué au développement de son descendant CORSIM, qui a également servi de modèle de base pour le modèle canadien CANSIM et pour le modèle suédois SVERIGE. Entre-temps, des douzaines de modèles d'usage général à grande échelle et un nombre innombrable de plus petits modèles spécialisés ont vu le jour partout dans le monde (pour une liste, voir Spielauer 2007). Néanmoins, la microsimulation continue de se heurter à la résistance de l'école dominante d'économistes « imprégnés d'envie à l'égard de la physique et attachant plus d'importance à l'élégance mathématique qu'au réalisme et à l'utilité » (Wolfson 2007). De plus en plus, ce front est brisé par les demandes des décideurs préoccupés par des questions distributionnelles et des problèmes de viabilité des politiques dans le contexte de l'évolution démographique. Il en est surtout ainsi des modèles de pensions qui constituent une vitrine d'exposition des nouvelles demandes des décideurs confrontés au vieillissement de la population et à des questions de viabilité et d'équité intergénérationnelle, de même que de la puissance de la microsimulation en vue de résoudre ces problèmes. Comme les pensions de retraite individuelles dépendent des antécédents de cotisation individuels, ainsi que des caractéristiques familiales (p. ex. pension de survivant), les modèles de pensions requièrent des simulations démographiques et économiques très détaillées. D'une part, cela peut les rendre très complexes, mais d'autre part, cela leur permet de répondre à des objectifs très différents. De nombreux modèles sont conçus en tant que modèles généraux capables d'étudier divers comportements et politiques, tels que la dynamique des études, l'effet distributionnel des régimes d'avantages fiscaux, ainsi que les besoins et les dispositions en matière de soins de santé. Ce progrès a été déclenché par la demande croissante, de la part des décideurs, de projections plus détaillées pour la planification, conjuguée aux progrès réalisés dans le domaine de la collecte et du traitement des données.

2 Quand la microsimulation dynamique est-elle l'approche de simulation appropriée?

Chaque fois que nous étudions la dynamique d'un système constitué d'unités plus petites, la microsimulation est une approche possible — mais quand cela vaut-il la peine de créer des milliers ou des millions de micro-unités? À la présente section, nous donnons trois réponses à cette question, la première axée sur l'hétérogénéité de la population, la deuxième, sur la difficulté d'agréger les relations comportementales et la troisième, sur les antécédents individuels.

2.1 Hétérogénéité de la population

La microsimulation est le mode de modélisation privilégié si les individus sont différents, si les différences importent et si le nombre de combinaisons possibles de caractéristiques prises en considération est trop grand pour diviser la population en un nombre pratique de groupes.

La théorie macroéconomique classique est en majeure partie fondée sur l'hypothèse que le secteur des ménages peut être représenté par un agent représentatif. L'hypothèse est que les individus sont identiques ou, dans le cas de modèles de générations chevauchantes, qu'ils ne se distinguent que par l'âge. (Chaque cohorte est représentée par un agent représentatif.) Cependant, cette approche n'est pas applicable si les distributions à un niveau plus fin de détail ont de l'importance. Imaginons que nous souhaitons étudier la viabilité et l'effet distributionnel du régime d'avantages fiscaux. S'il n'existe qu'un seul individu représentatif et que le régime d'avantages fiscaux est équilibré, cette personne moyenne recevra en avantages et en services ce qu'elle verse en impôts et en contributions au régime d'assurance sociale (certaines de ses heures de travail étant consacrées à l'administration du système). Pour modéliser les recettes fiscales, nous devons tenir compte de l'hétérogénéité de la population — si l'impôt sur le revenu est progressif, les recettes fiscales dépendent non seulement du revenu total, mais aussi de sa distribution. Quand nous concevons une réforme fiscale, nous visons habituellement à répartir les fardeaux différemment. Nous devons représenter l'hétérogénéité de la population dans le modèle pour déterminer quels seront les gagnants et les perdants de la réforme.

La microsimulation n'est pas le seul mode de modélisation privilégié pour traiter l'hétérogénéité. L'autre option consiste à grouper les personnes par combinaisons de caractéristiques pertinentes au lieu de les représenter individuellement. Nous utilisons pour cela des modèles à cellules. Une analogie directe existe entre les deux approches en ce qui concerne le stockage des données : un ensemble d'enregistrements individuels contre un tableau croisé dans lequel chaque cellule correspond à une combinaison de caractéristiques. Nous pouvons prendre comme exemple un recensement de population. Si nous nous intéressons uniquement à une ventilation selon l'âge et le sexe, nous pourrions procéder à un recensement en comptant les personnes présentant chaque combinaison de caractéristiques. Le recensement complet pourrait être présenté dans un seul tableau sauvegardé sous forme de tableur. Cependant, si nous voulons ajouter d'autres caractéristiques que l'âge et le sexe à notre description, le nombre de cellules du tableau augmentera exponentiellement, rendant l'approche de moins en moins pratique. Par exemple, 12 variables ou caractéristiques comportant chacune 6 niveaux nous obligeraient à grouper notre population dans plus de 2 milliards de cellules ($6^{12} = 2\,176\,782\,336$). Nous nous retrouverions rapidement avec un plus grand nombre de cellules que de personnes. En présence de variables continues (p. ex. le

revenu), l'approche du groupement devient carrément impossible sans perte d'information, puisque nous devrions grouper des données (c.-à-d. définir des niveaux de revenu). La solution consiste à garder les caractéristiques de chaque personne dans un enregistrement individuel, qui correspond au questionnaire et, en dernière analyse, à une ligne dans une base de données.

Ces deux types de représentation des données (un tableau croisé par opposition à un ensemble d'enregistrements individuels) correspondent aux deux types de simulation dynamique. Dans les modèles à cellules, nous mettons à jour un tableau; dans les modèles de microsimulation, nous modifions les caractéristiques de chaque enregistrement (et créons un nouvel enregistrement à chaque événement de naissance). Dans le premier cas, nous devons trouver des formules pour représenter la façon dont l'occupation de chaque cellule évolue au cours du temps; dans le second, nous devons modéliser les changements individuels au cours du temps. Les deux approches ont pour but de modéliser les mêmes processus, mais à des niveaux différents. La modélisation au niveau macro pourrait nous épargner beaucoup de travail, mais elle n'est possible que sous des conditions contraignantes, puisque les relations comportementales individuelles proprement dites doivent être agrégées, ce qui n'est pas toujours possible. Sinon, aucune formule n'existera pour décrire comment l'occupation de chaque cellule évolue au cours du temps.

La comparaison de l'approche de microsimulation aux modèles à cellules facilite sa compréhension. Nous approfondissons ci-après cette comparaison en prenant des projections démographiques comme exemple. Dans une approche fondée sur des cellules, si nous ne nous intéressons qu'à des taux de population selon l'âge, la mise à jour d'un tableau agrégé (une pyramide de population) ne requiert que quelques éléments d'information, à savoir les taux de fécondité par âge, les taux de mortalité par âge et la répartition de la population par âge à la période précédente. En l'absence de migration, la population d'âge x à la période t est la population survivante qui était d'âge $x-1$ à la période $t-1$. Pour une hypothèse donnée de mortalité, nous pouvons calculer directement la taille future prévue de la population d'âge x . Dans une approche de microsimulation, la survie correspond à une probabilité individuelle (ou à un taux, si nous faisons la modélisation en temps continu). L'hypothèse que 95 % d'un groupe d'âge seront encore en vie dans un an résulte en un processus stochastique au niveau micro — les individus peuvent être soit en vie soit décédés. Nous tirons un nombre aléatoire compris entre 0 et 1 — s'il est inférieur au seuil de 0,95, la personne simulée survit. Ce genre d'exercice porte le nom de simulation Monte Carlo. En raison de cet élément aléatoire, chaque expérience de simulation produit un résultat agrégé légèrement différent, qui converge vers la valeur prévue à mesure que nous augmentons la taille de la population simulée. Cette différence entre les résultats agrégés, appelée variation Monte Carlo, est un attribut type de la microsimulation.

2.2 Le problème de l'agrégation

La microsimulation est le choix de modélisation adéquat si les comportements sont complexes au niveau macro, mais qu'ils sont mieux compris au niveau micro.

De nombreux comportements sont plus faciles à modéliser au niveau micro, car c'est à ce niveau que sont prises les décisions et définies les règles fiscales. Très souvent, les comportements sont aussi plus stables au niveau micro auquel il n'existe aucune interférence des effets de composition. Même une stabilité totale au niveau micro ne correspond pas automatiquement à la stabilité au niveau macro. Par exemple, si l'on examine le niveau d'études, un des meilleurs prédicteurs des décisions en matière d'éducation est le niveau d'études des parents. Donc, si nous observons une expansion éducationnelle — par exemple une hausse du taux d'obtention d'un diplôme — au niveau de la population, celle-ci n'est pas forcément due à un changement de comportement de niveau micro; elle peut tenir entièrement à l'évolution de la composition de la génération parentale.

Les règlements en matière d'impôts et de sécurité sociale relient les règles de manière non linéaire aux caractéristiques individuelles et familiales, ce qui empêche l'agrégation de leurs opérations. De nouveau, il n'existe aucune formule pour calculer directement l'effet d'une réforme ou la viabilité d'un système, même si l'on ignore les problèmes de distribution. Pour calculer les recettes fiscales totales, nous devons connaître la composition de la population selon le revenu (impôts progressifs), les caractéristiques familiales (enfants et conjoints à charge) et toutes les autres caractéristiques qui ont une incidence sur le calcul de la dette d'impôt individuelle. En recourant à la microsimulation, nous pouvons modéliser un tel système à n'importe quel niveau de détail au niveau micro, puis agréger les impôts, cotisations et avantages individuels.

2.3 Biographies individuelles

La microsimulation est le seul choix de modélisation si les biographies individuelles importent, c'est-à-dire si les processus possèdent une mémoire.

Le décrochage scolaire est influencé par les expériences antérieures de décrochage, la mortalité, par les antécédents d'usage du tabac, les pensions de vieillesse, par les antécédents de cotisation individuels, et le chômage, par les périodes antérieures de chômage et leur durée. Les processus auxquels s'intéressent les spécialistes des sciences sociales sont fréquemment de ce type, c'est-à-dire qu'ils ont une mémoire. Dans de tels processus, les événements survenus dans le passé peuvent exercer une influence directe sur ce qui se passera dans l'avenir, ce qui empêche l'utilisation de modèles à cellules, parce qu'une fois que les données d'une cellule sont saisies, toute l'information sur l'appartenance à des cellules antérieures est

perdue. Dans de tels cas, la microsimulation devient la seule option de modélisation disponible.

3 Points forts et inconvénients

Les points forts de la microsimulation se concrétisent en trois dimensions. La microsimulation est séduisante d'un point de vue théorique, car elle appuie la recherche novatrice intégrée dans les paradigmes de recherche contemporains tels que la perspective de la trajectoire de vie. (À cet égard, la microsimulation est l'étape logique qui vient après l'analyse de la trajectoire de vie.) Elle est également séduisante d'un point de vue pratique, car elle fournit des outils pour l'étude et la projection de phénomènes sociodémographiques et socioéconomiques dynamiques d'une grande importance dans le domaine de l'élaboration des politiques. Enfin, la microsimulation est séduisante du point de vue technique, puisqu'elle n'est pas restreinte en ce qui a trait au type de variable et de processus, comme cela est le cas des modèles à cellules.

3.1 Points forts de la microsimulation d'un point de vue théorique

Le changement social et démographique massif survenu au cours des dernières décennies est allé de pair avec de gigantesques progrès techniques. La capacité de produire de grandes quantités de données a stimulé la collecte de données et permis l'adoption de nouveaux plans de sondage et de nouvelles méthodes d'analyse des données. En sciences sociales, ces nouveaux développements ont été accompagnés d'un changement général de paradigme, nombre des nouvelles orientations allant dans le sens de la vision d'Orcutt. L'une d'elles est la transition générale du niveau macro au niveau micro, afin d'axer la recherche sur les individus dans leur contexte. Un autre changement a trait à l'importance accrue accordée aux processus plutôt qu'aux structures statiques, faisant ainsi intervenir les concepts de causalité et de temps. Bien que l'approche de la microsimulation soutienne ces deux nouveaux pôles de concentration, elle constitue l'outil principal d'une troisième tendance en recherche, à savoir l'évolution de l'analyse vers la synthèse (Willekens 1999). La microsimulation relie de multiples processus élémentaires afin d'engendrer une dynamique complexe et de quantifier la contribution d'un processus particulier au profil complexe d'évolution.

Ces tendances en sciences sociales se reflètent dans l'émergence du paradigme de la trajectoire de vie qui relie l'évolution sociale, la structure sociale et l'action individuelle (Giele et Elder 1998). Sa perspective pluridimensionnelle et dynamique se traduit dans la recherche longitudinale et dans la collecte de données longitudinales. Les vies individuelles sont décrites comme une multitude de trajectoires parallèles et interdépendantes, telles que poursuivre des études, travailler, former des unions et avoir des enfants. Les états de chaque trajectoire sont modifiés par des événements dont les données sur leur survenue sont recueillies dans des enquêtes et qui sont simulés respectivement dans des modèles de microsimulation. Divers

points forts de l'approche de microsimulation correspondent directement aux concepts fondamentaux de la perspective de la trajectoire de vie, ce qui en fait l'approche logique pour l'étude et la projection des phénomènes sociaux.

La microsimulation convient bien pour simuler l'interaction des trajectoires, car elle permet de modéliser à la fois les processus qui ont une mémoire (c.-à-d. que les individus se souviennent des événements passés pour divers domaines de trajectoire) et les diverses trajectoires parallèles pour lesquelles les probabilités ou risques de survenue d'un événement d'une trajectoire particulière répondent aux changements d'état dans d'autres trajectoires.

En plus de reconnaître les interactions entre les trajectoires, la perspective de la trajectoire de vie met l'accent sur l'interaction entre les individus, à savoir le concept des vies liées. La microsimulation est un outil puissant pour étudier et projeter ces interactions. Elle peut inclure des changements dans les réseaux de parenté (Wachter 1995), des transferts intergénérationnels et la transmission de caractéristiques comme l'éducation (Spielauer 2004) et la transmission de maladies comme le sida.

Dans la perspective de la trajectoire de vie, la situation et les décisions courantes d'une personne peuvent être considérées comme la conséquence d'expériences passées ou d'attentes futures, et comme une intégration de motifs individuels et de contraintes externes. De cette façon, l'agent humain et l'orientation des objectifs individuels font partie du cadre explicatif. L'un des principaux mécanismes grâce auxquels les individus relèvent les défis de la vie est la synchronisation des événements de la vie appartenant à des trajectoires parallèles, et souvent difficiles à concilier, comme travailler et élever des enfants. La microsimulation permet la modélisation des agents individuels, car toutes les décisions et tous les événements sont modélisés au niveau où ils ont lieu et les modèles peuvent tenir compte du contexte individuel. Outre ces avantages intrinsèques, la microsimulation n'impose aucune contrainte quant à la façon de modéliser les décisions; autrement dit, elle permet d'utiliser n'importe quel type de modèle de comportement qui peut être exprimé en code informatique.

3.2 Points forts de la microsimulation d'un point de vue pratique

La capacité de créer des modèles permettant de projeter les effets des politiques est l'élément central de la vision d'Orcutt. L'attrait de la microsimulation dynamique dans le domaine de l'élaboration des politiques est étroitement associé aux forces intrinsèques de cette approche. Elle permet de modéliser des politiques à n'importe quel niveau de détail, ainsi que d'aborder les questions distributionnelles et les problèmes de viabilité de long terme. Une partie de cette puissance est déjà exploitée dans les modèles de microsimulation statiques des avantages fiscaux, qui sont devenus un outil standard d'analyse des politiques dans la plupart des pays développés. Issus de l'intérêt croissant des décideurs pour l'étude des phénomènes

distributionnels, ces modèles sont toutefois limités par nature aux études transversales. Bien que des projections limitées des avantages fiscaux dans l'avenir soient possibles au moyen de modèles de microsimulation statistiques par repondération des individus d'une population initiale en vue de représenter la population dans l'avenir (et en remettant à niveau le revenu et d'autres variables), cette approche est dépourvue de la dimension longitudinale, c'est-à-dire les trajectoires de vie individuelles (et les antécédents de cotisation), qui est simulée dans les modèles dynamiques. Le rôle important de la dynamique dans les applications des politiques a été principalement reconnu dans la conception et la modélisation des régimes de pension, qui sont fortement touchés par le vieillissement démographique. Les modèles de pensions sont aussi de bons exemples d'applications dans lesquelles les biographies individuelles (cotisations) et le concept des vies liées (pension de survivant) ont de l'importance. Un autre exemple est la planification des établissements de soins dont la demande est dictée par le vieillissement démographique ainsi que par l'évolution des réseaux de parenté et la participation au marché du travail (c.-à-d. les principaux facteurs affectant la disponibilité de soins non officiels).

Compte tenu du rythme rapide d'évolution sociale et démographique, la nécessité d'adopter une perspective longitudinale a été admise rapidement dans la plupart des autres domaines stratégiques qui bénéficient des projections détaillées et du « monde virtuel », ou environnement d'essai, fourni par les modèles de microsimulation dynamique. Non seulement l'aspect longitudinal de la microsimulation dynamique est important en ce qui concerne les questions de viabilité, mais il étend aussi la portée de l'analyse des effets distributionnels des politiques. La microsimulation peut être utilisée pour analyser les distributions sur la base de la trajectoire de vie et pour résoudre les questions d'équité intergénérationnelle. La possibilité d'étudier et de comparer la distribution des taux de rendement des séries de contributions et d'avantages individuels au cours de la vie complète des individus en est un exemple.

3.3 Points forts de la microsimulation d'un point de vue technique

D'un point de vue technique, la principale force de la microsimulation tient au fait qu'elle n'est pas soumise aux contraintes typiques d'autres approches de modélisation. Contrairement aux modèles à cellules, la microsimulation permet de traiter n'importe quel nombre de variables de n'importe quel type. Comparativement aux modèles de niveau macro, il n'est pas nécessaire d'agréger les relations comportementales, ce qui, dans les macromodèles, n'est possible que sous des hypothèses contraignantes. Dans la microsimulation, la modélisation des comportements individuels n'est soumise à aucune contrainte et ce sont les résultats des comportements qui sont agrégés. Autrement dit, aucune restriction n'est appliquée au type de processus. Mais, par-dessus tout, la microsimulation permet de prendre en considération des processus non markoviens, c'est-à-dire des processus qui possèdent une mémoire. Fondée sur

des microdonnées, la microsimulation permet une agrégation flexible, car l'information peut être croisée de n'importe quelle façon, tandis que dans les approches agrégées, le schéma d'agrégation est déterminé a priori. Les résultats des simulations peuvent être présentés et expliqués simultanément de diverses façons — par des séries chronologiques agrégées, des distributions conjointes transversales, ainsi que des trajectoires de vie individuelles et familiales.

3.4 Quel est le prix? Inconvénients et limites

La microsimulation présente trois types d'inconvénients (et de préconceptions) de nature très différente, à savoir l'esthétique, les limites fondamentales inhérentes à toutes les prévisions et les coûts.

Si la beauté réside dans la simplicité et l'élégance mathématique (point de vue qui n'est pas rare chez les économistes faisant partie du courant dominant), les modèles de la microsimulation violent toutes les règles de l'esthétique. Les modèles de microsimulation à grande échelle requièrent un nombre incalculable de paramètres estimés au moyen de données provenant de diverses sources qu'il est souvent difficile de rapprocher. La simulation des politiques requiert une comptabilité fastidieuse et, étant donné leur complexité, les modèles de microsimulation posent toujours le risque de devenir des boîtes noires difficiles à utiliser et à comprendre. Même s'il est possible d'améliorer la documentation et l'interface utilisateur des modèles de microsimulation, le sacrifice de l'élégance à l'utilité sera toujours un trait de cette approche de modélisation.

Le deuxième inconvénient est plus fondamental. La limite essentielle de la microsimulation tient au fait que le niveau de détail du modèle ne va pas de pair avec la puissance globale de prédiction. Cet état de chose tient à ce que l'on appelle le caractère aléatoire de la microsimulation, causé en partie par la nature stochastique des modèles et en partie par les erreurs et biais cumulés dans les valeurs des variables. Le compromis entre le détail et le biais éventuel existe déjà à l'étape du choix des sources de données, puisque la taille de l'échantillon des enquêtes ne va pas de pair avec le niveau de détail des modèles. Il existe un compromis entre la stochasticité additionnelle introduite par les variables supplémentaires et les erreurs dues à la spécification incorrecte causées par des modèles qui sont trop simplifiés. Autrement dit, la caractéristique qui rend la microsimulation particulièrement séduisante, à avoir le grand nombre de variables que les modèles peuvent contenir, a pour prix le caractère aléatoire des modèles et l'affaiblissement résultant de la puissance prédictive à mesure que le nombre de variables augmente. Cette situation crée un compromis entre de bonnes prédictions au niveau agrégé et une bonne prédiction concernant les aspects distributionnels dans le long terme, un fait dont les modélisateurs doivent tenir compte. Ce problème de compromis n'est pas particulier à la microsimulation, mais puisque celle-ci est habituellement employée pour

produire des projections détaillées, l'importance de l'effet de la stochasticité s'accroît conséquemment. Il n'est donc pas étonnant que, dans nombre de modèles à grande échelle, certains processus sont harmonisés ou calés sur des projections agrégées obtenues par des moyens externes.

Outre sa nature fondamentale, la portée de ce caractère aléatoire dépend de la fiabilité ou de la qualité des données. À cet égard, nous pouvons observer et attendre diverses améliorations à mesure que des données de plus en plus détaillées deviennent disponibles pour la recherche, non seulement sous forme de données d'enquête, mais aussi de données administratives. Ces dernières ont propulsé la microsimulation, particulièrement en Europe, dans les pays nordiques.

Puisque la microsimulation produit non pas des valeurs prévues, mais plutôt des variables aléatoires distribuées autour des valeurs prévues, elle présente une autre forme de caractère aléatoire : la variabilité Monte Carlo, due au fait que chaque expérience de simulation produit des résultats agrégés différents. Bien que cela était fastidieux à l'époque où la puissance informatique était limitée, l'exécution d'un grand nombre d'expériences répétées et (ou) la simulation d'une grande population peut éliminer ce genre de caractère aléatoire et produire des renseignements précieux sur la distribution des résultats, en plus des estimations ponctuelles.

Le troisième type d'inconvénient est celui des coûts de développement. Les modèles de microsimulation requièrent des données de grande qualité, longitudinales et parfois d'un type très particulier, dont l'acquisition et la compilation est coûteuse. Il convient de souligner qu'il ne s'agit pas de coûts explicites associés à la microsimulation proprement dite, mais du prix à payer pour la recherche longitudinale en général et l'élaboration de politiques fondées sur des faits en particulier.

Habituellement, les modèles de microsimulation demandent aussi de gros investissements en ressources humaines et en matériel. Cependant, il faut s'attendre à ce que ces coûts continuent à diminuer, à mesure que le prix du matériel baisse et que des langages informatiques plus puissants et plus efficaces deviennent disponibles. Malgré tout, aux yeux de nombreux chercheurs, les obstacles à l'entrée sont élevés. Si nombre d'entre eux reconnaissent le potentiel de la microsimulation, ils restent sceptiques quant à la possibilité d'appliquer ses techniques dans le cadre de petits projets de recherche. Nous espérons que l'accès au langage Modgen réduit cet obstacle perçu et rend la microsimulation plus accessible au monde de la recherche. Au cours des deux dernières années, divers modèles de microsimulation à une plus petite échelle ont été développés dans le cadre de projets de doctorat ou d'études particulières. Modgen peut à la fois accélérer la programmation de petites applications et offrir une

plateforme de modélisation éprouvée et à jour pour les modèles à grande échelle, tels que les modèles LifePaths et Pohem de Statistique Canada.

Bibliographie

- Giele, J.Z. and Elder, G.H. Jr. (1998), 'Methods of Life Course Research: Qualitative and Quantitative Approaches.' SAGE Publications
- Klevmarken, N.A. (1997), 'Behavioral Modeling in Micro Simulation Models. A Survey.' Working Paper 1997:31, Department of Economics, Uppsala University.
- Orcutt, Guy (1957), 'A new type of socio-economic system', Review of Economics and Statistics, 39(2), 116-123. Reprint in International Journal of Microsimulation (2007) 1(1), 3-9 http://www.microsimulation.org/IJM/V1_1/IJM_1_1_2.pdf
- Spielauer, M. (2004) 'Intergenerational Educational Transmission within Families: Analysis and Microsimulation Projection for Austria', Vienna Yearbook of Population Research 2004, http://hw.oeaw.ac.at/3272-7inhalt/Spielauer_253_282.pdf
- Spielauer, M. (2007) 'Dynamic Microsimulation of Health Care Demand, Health Care Finance and the Economic Impact of Health Behaviours: Survey and Review.' International Journal of Microsimulation 1(1) 35-53 http://www.microsimulation.org/IJM/V1_1/IJM_1_1_5.pdf
- Wachter, K.W. (1995), '2030's Seniors: Kin and Step-Kin.' Berkeley: University of California.
- Willekens, F. (1999), 'Theoretical and Technical Orientation toward Longitudinal Research in the Social Sciences.' Population Research Centre, University of Groningen.
- Wolfson, M.C. (1999) "New Goods and the Measurement of Real Economic Growth", Canadian Journal of Economics, Volume 32, No. 2, April 1999
- Wolfson, M.C. (2007) 'Orcutt: Vision and Development', Opening address at the 1st General Conference of the International Microsimulation Association, Vienna, August 2007