Economic Modelling 26 (2009) 560–570

Contents lists available at ScienceDirect

Economic Modelling

journal homepage: www.elsevier.com/locate/econbase

On linking microsimulation and computable general equilibrium models using exact aggregation of heterogeneous discrete-choice making agents ☆

Riccardo Magnani [a,*], Jean Mercenier [b]

[a] CEPII, 9 rue Georges Pitard, 75015 Paris, France
[b] ERMES, Université Panthéon-Assas (Paris 2), 12 place du Panthéon, 75230 Paris Cedex 05, France

ARTICLE INFO

Article history:
Accepted 8 December 2008

JEL classification:
C63
C68
C81
D31
D58
E17
J10
J22

Keywords:
Microsimulation
CGE models
Exact aggregation
Discrete choice
Nested multinomial logit
Population ageing
Income inequality

ABSTRACT

Our paper contributes by bridging the gap between the (partial equilibrium) microsimulation and the computable general equilibrium (CGE) approaches, by making use of exact aggregation results from the discrete choice literature: heterogeneous individuals choosing within a set of discrete alternatives may be aggregated into a representative agent with (possibly multiple-level) constant elasticity-of-substitution/transformation preferences/technologies. These results therefore provide a natural link between the two policy evaluation approaches. We illustrate the usefulness of these results by evaluating potential effects of population ageing on the dynamics of income distribution and inequalities, using a simple overlapping generations model where individuals make leisure/work decisions, and choose a profession among a discrete set of alternatives.

1. Introduction

During the last twenty years, computable general equilibrium (CGE) models have become standard tools of quantitative policy assessment. Their appeal has built on their rigorous grounding in economic theory: agents' decision-making behaviour is derived from explicit optimisation under strictly specified technological or budget constraints, given market signals that ensure global consistency. These theoretical foundations have made CGE models appear particularly useful for ex-ante evaluations of policy reforms. However, the whole apparatus relies on the concept of "representative agent" despite unclear aggregation procedures to link these aggregate optimising decision-makers to the numerous individual agents whose behaviour they are meant to capture.

During the same period, another class of models has become increasingly popular: behavioural microsimulation models. Their appeal stems from the fact that they avoid any reliance on *typical agents* by fully taking into account the heterogeneity of individual choices as they are revealed in micro-data sets.[1] Indeed, working with myriads of actual economic agents rather than with a few hypothetical ones makes it possible to precisely identify the winners and the losers of a reform – obviously a major concern to policy-makers – yet, making it possible by simple addition to accurately measure this impact on aggregate variables. The increasing availability of large and detailed data sets on individuals makes this quite appealing. The drawback of the approach is that it is partial equilibrium in essence: for instance, individual's labour supply adjustment to some new tax incentive scheme can be quite accurately captured for *given* wages and other policy parameters, but market equilibrium and government budget constraints can be expected

☆ We wish to thank Nathalie Picard for very helpful discussions. We express our gratitude to anonymous referees for comments that helped us improve the paper, and (without implicating) Human Resources and Social Development Canada for financial support. The usual disclaimer applies.

* Corresponding author.
E-mail address: Riccardo.Magnani@cepii.fr (R. Magnani).
URL: http://mercenier.u-paris2.fr (J. Mercenier).

[1] See Bourguignon and Spadaro (2006) for an excellent survey and an extensive list of references.

0264-9993/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.econmod.2008.12.002

to have a feedback influence on the same individual's choices that is typically neglected. One could of course imagine iterating between the microsimulation and the CGE models, and indeed, a few efforts have successfully been done in this direction: see for instance Savard (2003) and the elaboration of Arntz et al. (2008) on Arntz et al. (2006). Though this iterative strategy might be satisfactory for some problems – in particular when dynamics are thought unimportant – it becomes tedious for more sophisticated apparatus such as overlapping generations (OLG) models: see however Rausch and Rutherford (2007) for progress in that direction.

In this paper, we make use of simple yet powerful exact aggregation results due to Anderson, de Palma and Thisse (1992) (here after: AdPT) who show that, under reasonably mild conditions, heterogeneous individuals that have to choose (possibly continuous amounts) within a set of discrete alternatives may be aggregated into a representative agent with constant elasticity-of-substitution (CES) preferences.[2] We illustrate how these results can be useful to CGE modellers by making available to them a growing body of empirical estimates from microeconometrics that can be used to parameterise CES/CET (constant elasticity-of-transformation) preferences/technologies in the representative agent framework. Furthermore, we argue that these results provide a natural and appealing link between the standard CGE apparatus and the microsimulations approach, and suggest that they constitute a useful alternative approach to the iterative strategy between microsimulation and CGE models. There is no free lunch, unfortunately: some details captured by the microsimulation approach could be lost, a cost that one should balance against the benefits of accounting for the general equilibrium feedbacks.

We show how to make use of these results in order to link the micro and the macro simulation approaches, and illustrate the usefulness of the methodology in the context of population ageing using a calibrated overlapping generations (OLG) model. For this, we first generate *in vitro* a micro-data set where individuals, classified in different cells according to their socio-economic characteristics, face random utility maximisation problems over sets of discrete alternatives. We focus, for illustrative purposes, on labour market participation, and particularise the discrete choices as "to work or not to work, and if work is chosen, in which profession?" in a nested multinomial logit framework.[3] We then show that the aggregation of individual choices yields a labour-supply scheme that coincides with the one derived from a macro-agent's time-allocation problem subject to smooth nested CES preferences as typically used in CGE models. The representative agent is part of a dynamic GE model which we simulate to evaluate the effects of a demographic shock on the time path of wages and interest rates. These equilibrium prices are then plugged into the microsimulation model in order to determine the response of each individual micro-agent to the changes in his/her economic environment. From this individual choice response, we can compute the income distributions consistent with general equilibrium wages, and therefore apprehend the dynamics of income inequalities induced by population ageing.

The paper is organised as follows: in Section 2, we provide a refresher on probabilistic discrete choice models. Focusing on a typical labour force participation decision problem, we show in Section 3 how to link the myriads of heterogeneous micro-agents of the microsimulation approach to a macro-agent. This macro-agent is embedded in the dynamic GE model sketched in Section 4. We then submit in Section 5 the OLG economy to an ageing shock, and plug the equilibrium prices in the microsimulation model to generate the time-path of income inequality indicators. The paper closes with a brief conclusion.

2. Discrete-choice models: a refresher

Assume a population of individuals $h = 1,\ldots, N$ has to choose among a set $i, j = 0,\ldots, I$ of discrete alternatives with associated utility levels:

$$\tilde{u}_i^h = u_i + \varepsilon_i^h \quad i = 0, \ldots, I \tag{1}$$

where u_i is a deterministic component (for now, assumed common to all individuals) and ε_i^h is a random term. Each h is therefore characterised by a draw $\varepsilon = (\varepsilon_0^h,\ldots, \varepsilon_I^h)$ in a probability distribution with cumulative distribution function $F(\varepsilon)$. Assume that individuals in this population are not only statistically identical but also statistically independent. Then, the distribution of choices is multinomial with mean $\overline{X}_i = NP_i$, $i = 0,\ldots,I$, where P_i denotes the probability that alternative i be chosen by h. $\overline{X}_i$ is the mathematical expectation of demand for alternative i; for N large enough, $\overline{X}_i$ is a close approximation of aggregate demand for i in this population. In other words, aggregate demands for each alternative may be readily determined from the choice probabilities from the individual discrete decision problem.

The probability that h will choose alternative i is:

$$\begin{aligned} P_i &= \text{prob}\left[\tilde{u}_i^h \geq \tilde{u}_j^h, \quad \forall j = 0,\ldots,I\right] \\ &= \text{prob}\left[u_i + \varepsilon_i^h \geq u_j + \varepsilon_j^h, \quad \forall j = 0,\ldots,I\right] \\ &= \text{prob}\left[\varepsilon_j^h - \varepsilon_i^h \leq u_i - u_j, \quad \forall j = 0,\ldots,I\right] \end{aligned} \tag{2}$$

The determination of the choice probabilities using $F(\varepsilon)$ is in principle always possible but in general extremely difficult, in particular if ε is assumed normally distributed as would seem natural. Fortunately, a theorem due to McFadden[4] identifies a class of cumulative distribution functions $F(\varepsilon)$ – of which the double exponential is a special case that yields the multinomial logit – for which these probabilities may be easily determined indirectly. Consider the multivariate generalised extreme value (GEV) cumulative distribution function

$$F(\varepsilon_0,\ldots,\varepsilon_I) = exp\left[-H\left(e^{-\varepsilon_0},\ldots,e^{-\varepsilon_I}\right)\right] \tag{3}$$

with H a nonnegative function defined over R_+^N satisfying the following properties: (*i*) H is homogeneous of degree $1/\mu$; (*ii*) $\lim_{x_i\to\infty} H(x_0,\ldots,x_I) = \infty$ $\forall i = 0,\ldots, I$; (*iii*) the mixed partial derivatives of H with respect to k different variables exist and are continuous, non-negative if k is odd, non-positive if k is even, $k = 0,\ldots, I$. (These technical conditions are needed to ensure that $F(\varepsilon)$ is indeed a cumulative distribution function.) Then, McFadden's GEV theorem states that the choice probabilities P_i may be determined as:

$$P_i = \mu \frac{\partial ln H\left(e^{u_0},\ldots,e^{u_I}\right)}{\partial u_i} \tag{4}$$

Many particularisations of H consistent with utility maximisation are possible, and to each corresponds a different distribution for ε. One important specification for H is:

$$H\left(e^{-\varepsilon_0},\ldots,e^{-\varepsilon_I}\right) = \sum_{i=0}^{I} e^{-\frac{\varepsilon_i}{\mu}} \tag{5}$$

[2] Discrete choice models can be extended to so-called continuous/discrete models that allow individuals to demand continuous quantities (not restricted to 0 or 1) of their preferred discrete option. See, e.g., Train (1986, Chap. 5).

[3] At the risk of being overemphatic, it seems useful to insist that the aggregation methodology is quite general and can be applied to a broad set of choices other than labour supply decisions.

[4] See McFadden (1978), p.80; 1981, p.227.