On Linking Microsimulation and Computable General Equilibrium Models Using Exact Aggregation of Heterogeneous Discrete-choice Making Agents

Riccardo Magnani (Cepii, Paris)

Jean Mercenier (Université de Paris 2)

Workshop on Linking Microsimulation and Macro Models
Institute for Employment Research

Nuremberg, December 15-16, 2008

Objective

The objective of the paper is to develop a new methodology allowing to link the General Equilibrium approach and the Microsimulation approach by using an exact aggregation result of individual discrete choices

Our methodology

Integration of the two approaches by using the exact aggregation result due to Anderson, de Palma and Thisse (1992):

Heterogeneous individuals, who have to choose among a set of discrete alternatives, may be aggregated into a representative agent with CES preferences

Our methodology

This result implies:

- **that it is not necessary to work with a big number of individuals in order to take into account the individual heterogeneity**
- **that it is not necessary to iterate because the equilibrium prices obtained in the GE model are already computed by taking into account the individual heterogeneity**

Individuals can be classified in different groups according to their socio-economic characteristics (age, sex, education...)

Our methodology

We can apply our methodology to discrete choices concerning:

- **Labor market decisions**
  - **whether to work or not**
  - **in which profession**
  - **retirement age**
- **Education decisions**
- **Migration**

How to implement

- **We need:**
  - **A Micro data-set to estimate the individual preferences**
  - **A General Equilibrium model to evaluate the effects on the equilibrium prices**
  - **A Microsimulation model to evaluate the effects at the individual level**

How to implement

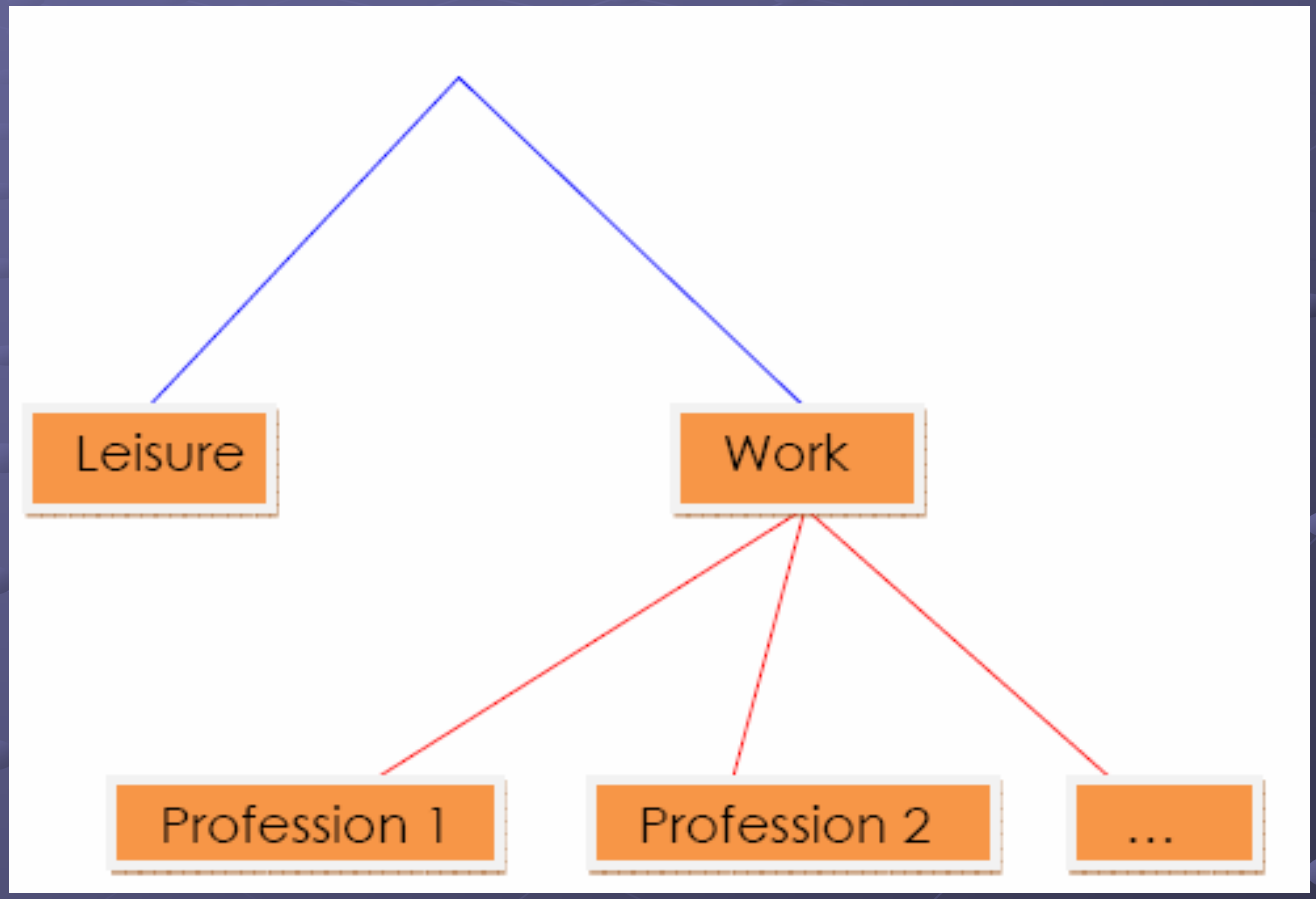
- **Step 1 – For each socio-economic group: Estimation and Aggregation of individual preferences**
- **Step 2 – Introduction of the functions that aggregate the individual preferences into a GE model**
- **Step 3 – Simulation of macroeconomic shock and computation of the equilibrium prices**
- **Step 4 – Evaluation of the impacts at the individual level (choices) and on income distribution, inequality, poverty...**

Estimation and Aggregation of the individual preferences

- **Individuals belonging to a specific socio-economic group (by age, sex, education…) have to choose:**
  - **Whether to work or not**
  - **In which profession**

- **Two-stage decision problem**

The Nested Multinomial Logit



1) Choice of the profession

Utility of choosing profession *i*:

$$\overline{u}_i^h = \ln \theta_i + \ln \overline{w}_i + \epsilon_i^h \qquad i \in A_1$$

It depends on:

- **market wage in profession *i***
- **characteristics specific to profession *i***
- **double exponential stochastic term (correlated, with dispersion parameter μ_2)**

2) Choice of whether to work or not

Utilities of not working and working are:

$$\begin{cases} \overline{u}^h_{A_0} = \ln \theta_{A_0} + \varepsilon^h_{A_0} \\ \\ \overline{u}^h_{A_1} = G_{A_1} + \varepsilon^h_{A_1} \end{cases}$$

$$\text{with} \quad G_{A_1} = \mu_2 \ln \sum_{j \in A_1} \exp\left(\frac{\ln \theta_j + \ln \overline{w}_j}{\mu_2}\right)$$

$$\text{and} \quad \varepsilon^h_{A_1} = \max_{j \in A_1} \left(u_j + \epsilon^h_j\right) - G_{A_1}$$

They depend on:

- **disutility of working**
- **market wages, consistent with the 2nd stage decision problem.**

G_{A1} is the expected maximum utility of a subset of alternatives

- **double exponential stochastic terms (dispersion parameter μ_1)**

Probabilities

The probability of choosing profession *i* is:

$$P_i^{A_1} = \frac{\theta_i^{\frac{1}{\mu_2}} \cdot \overline{w}_i^{\frac{1}{\mu_2}}}{\sum_{j \in A_1} \theta_j^{\frac{1}{\mu_2}} \cdot \overline{w}_j^{\frac{1}{\mu_2}}} \qquad i \in A_1$$

The probability of working is:

$$P_{A_1} = \frac{\left[\sum_{j \in A_1} \theta_j^{\frac{1}{\mu_2}} \cdot \overline{w}_j^{\frac{1}{\mu_2}}\right]^{\frac{\mu_2}{\mu_1}}}{\theta_{A_0}^{\frac{1}{\mu_1}} + \left[\sum_{j \in A_1} \theta_j^{\frac{1}{\mu_2}} \cdot \overline{w}_j^{\frac{1}{\mu_2}}\right]^{\frac{\mu_2}{\mu_1}}}$$

Aggregation of individual choices

The number of individuals (belonging to a specific group) who decide to work in profession *i* is given by:

$$L_i^{\text{sup}} = \frac{\left[\sum_{j\in A_1} \theta_j^{\frac{1}{\mu_2}} \cdot \overline{w}_j^{\frac{1}{\mu_2}}\right]^{\frac{\mu_2}{\mu_1}}}{\theta_{A_0}^{\frac{1}{\mu_1}} + \left[\sum_{j\in A_1} \theta_j^{\frac{1}{\mu_2}} \cdot \overline{w}_j^{\frac{1}{\mu_2}}\right]^{\frac{\mu_2}{\mu_1}}} \cdot \frac{\theta_i^{\frac{1}{\mu_2}} \cdot \overline{w}_i^{\frac{1}{\mu_2}}}{\sum_{j\in A_1} \theta_j^{\frac{1}{\mu_2}} \cdot \overline{w}_j^{\frac{1}{\mu_2}}} \cdot N \qquad i \in A_1$$

This is a labor supply function that perfectly aggregates the individual preferences

Aggregation of individual choices

$$\frac{L_i^{sup}}{L_j^{sup}} = \left(\frac{\theta_i}{\theta_j}\right)^{\frac{1}{\mu_2}} \left(\frac{\overline{w}_i}{\overline{w}_j}\right)^{\frac{1}{\mu_2}}$$

$$\frac{L^{sup}}{1-L^{sup}} = \left[\frac{\left(\sum_{j\in A_1} \theta_j^{\frac{1}{\mu_2}} \cdot \overline{w}_j^{\frac{1}{\mu_2}}\right)^{\mu_2}}{\theta_{A_0}}\right]^{\frac{1}{\mu_1}} = \left[\frac{e^{G_{A_1}}}{\theta_{A_0}}\right]^{\frac{1}{\mu_1}}$$

μ_2 is the inverse of the transformation elasticity between professions

μ_1 is the inverse of the transformation elasticity between work and leisure

Representative agent

As in AdPT, we can write an optimization problem for a representative agent (one for each cell) who decides the optimal allocation of his time into leisure and professional activities

This optimal time allocation coincides with the one generated from the aggregation of the individual discrete choices

An illustration

We apply our methodology to labor choices in the context of population ageing

Objective: Evaluate the effects of population ageing on the dynamics of the income distribution and inequalities

Micro data-set

- **We generate *in vitro* a micro data-set of 51,850 individuals (39,525 individuals aged 15-64)**
- **Individuals are classified on the basis of their age and sex**
- **Individuals have to choose whether to work or not and, if yes, in which profession (*Prof0* and *Prof1*)**

Micro data-set

Number of individuals by age and sex

	Males	*Females*
15-24	4 000	4 500
25-34	4 000	4 500
35-44	3 800	4 275
45-54	3 600	4 050
55-64	3 200	3 600
65-74	2 800	3 150
75-84	2 000	2 250
85-94	1 000	1 125
Total	24 400	27 450

Micro data-set

Generation of individual wages

$$\ln w_{i,g,s}^{h} = \varphi_{0i} + \varphi_{1i} \cdot g + \varphi_{2i} \cdot g^2 + \varphi_{3i} \cdot s + \psi_i^h$$

	Prof-0	Prof-1
constant	5.000	5.500
g	0.400	0.450
g^2	-0.030	-0.035
s	-0.300	-0.320

$$\psi_i^h = N(0, 0.5)$$

Micro data-set

General statistics on individual wages by age and sex for the two professions

	Males					*Females*				
	Obs	Mean	Std. Dev.	Min	Max	Obs	Mean	Std. Dev.	Min	Max
15-24	2335	222.668	114.057	37.206	994.496	2416	165.755	83.429	35.147	725.862
25-34	2393	298.008	148.925	54.594	1586.799	2582	225.667	118.394	41.730	972.842
35-44	2192	386.735	196.312	47.203	1767.226	2453	287.372	143.336	51.962	1258.786
45-54	1976	475.490	240.257	69.690	2424.641	2233	346.310	175.671	55.017	1863.712
55-64	1374	527.044	253.920	94.021	1934.993	1586	401.647	213.611	66.931	2025.210
	Males					*Females*				
	Obs	Mean	Std. Dev.	Min	Max	Obs	Mean	Std. Dev.	Min	Max
15-24	872	429.900	213.679	107.592	1872.320	775	325.198	157.753	80.321	1345.630
25-34	982	620.612	311.012	87.031	2378.202	969	455.526	213.806	83.345	1765.826
35-44	1158	767.865	384.479	182.279	3267.305	1017	591.654	297.682	118.489	2442.261
45-54	1217	928.794	504.224	186.391	7067.726	1057	704.562	342.523	140.431	2455.936
55-64	1059	1043.794	567.688	196.841	5463.498	791	783.668	384.234	164.134	2882.709

Micro data-set

- The aggregation result is “exact” if all the individuals belonging to the same socio-economic group earn the same wage

$$L_i^{\sup} = \frac{\left[\sum_{j\in A_1} \theta_j^{\frac{1}{\mu_2}} \cdot \overline{w}_j^{\frac{1}{\mu_2}}\right]^{\frac{\mu_2}{\mu_1}}}{\theta_{A_0}^{\frac{1}{\mu_1}} + \left[\sum_{j\in A_1} \theta_j^{\frac{1}{\mu_2}} \cdot \overline{w}_j^{\frac{1}{\mu_2}}\right]^{\frac{\mu_2}{\mu_1}}} \cdot \frac{\theta_i^{\frac{1}{\mu_2}} \cdot \overline{w}_i^{\frac{1}{\mu_2}}}{\sum_{j\in A_1} \theta_j^{\frac{1}{\mu_2}} \cdot \overline{w}_j^{\frac{1}{\mu_2}}} \cdot N \qquad i \in A_1$$

- In the labor supply functions used the GE model we consider, for each cell, the average wage
- In the microsimulation model, individuals choose by considering their individual wage (different from the average level)
- It is important to check that discrepancy between the macro and the micro labor supply is sufficiently small. We find that the error is lower than 0.01%

Micro data-set

We generate the preference parameters, so we can determine the choices for each individual

Transformation elasticities of the labor supply functions (inverse of μ_1 and μ_2)

	Leisure / Work	Prof-0 / Prof-1
Males		
15-24	0.900	1.665
25-34	0.800	1.590
35-44	0.700	1.490
45-54	0.600	1.450
55-64	0.500	1.375
Females		
15-24	0.850	1.700
25-34	0.750	1.625
35-44	0.650	1.575
45-54	0.550	1.540
55-64	0.450	1.475

Micro data-set

Work-Leisure choice and choice of the profession
General statistics by age and sex

	Leisure / Total	Work / Total	Prof-0 / Work	Prof-1 / Work
Males				
15-24	19.8%	80.2%	72.8%	27.2%
25-34	15.6%	84.4%	70.9%	29.1%
35-44	11.8%	88.2%	65.4%	34.6%
45-54	11.3%	88.7%	61.9%	38.1%
55-64	24.0%	76.0%	56.5%	43.5%
Females				
15-24	29.1%	70.9%	75.7%	24.3%
25-34	21.1%	78.9%	72.7%	27.3%
35-44	18.8%	81.2%	70.7%	29.3%
45-54	18.8%	81.2%	67.9%	32.1%
55-64	34.0%	66.0%	66.7%	33.3%

The OLG-GE model

- **Standard OLG model to evaluate the macroeconomic effects of population ageing**
- **8 generations (15-24, 25-34,…, 85-94) that coexist at each time**
- **Heterogeneity: individuals differ in age and sex (as in the micro-data set)**
- **One representative firm uses capital and labor (two professions)**
- **A PAYG pension system: the replacement ratio is endogenously determined to guarantee the equilibrium**

The OLG-GE model

Representative agents decide:

- **The intertemporal profile of consumption**

- **The labor supply at each period (the labor supply functions come from the aggregation of the individual preferences)**

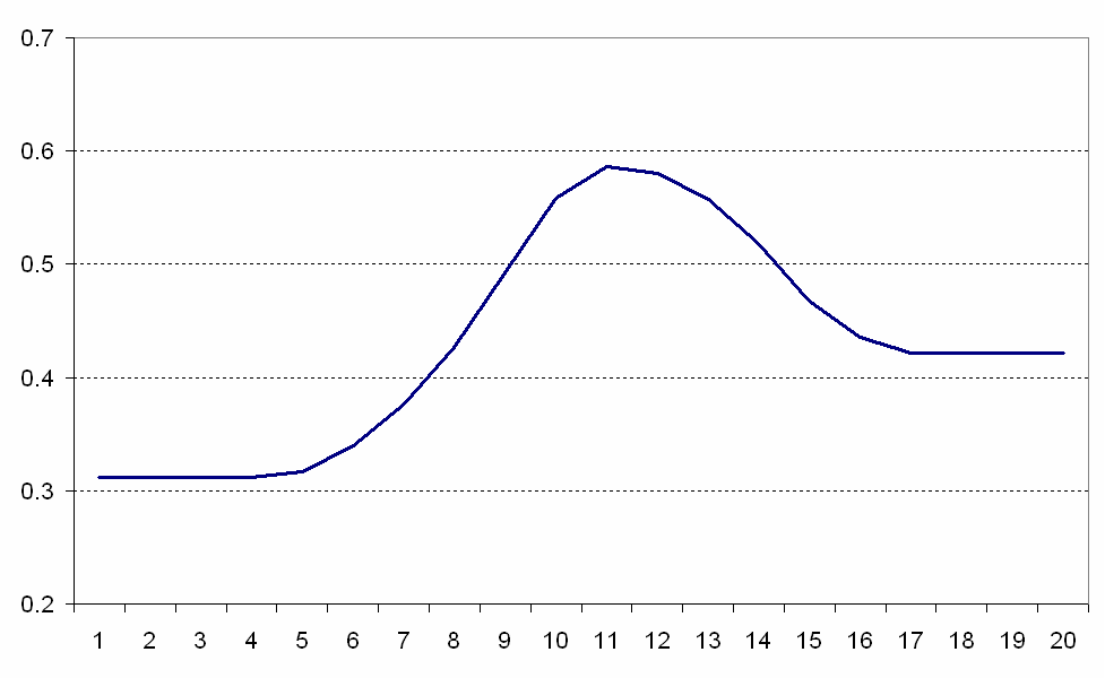
The Macroeconomic shock

Reduction in fertility rates and increase in survival rates

	1	2	3	4	5	6	7	8	9	> 9
η	1.000	1.000	1.000	0.930	0.925	0.919	0.911	0.903	0.892	1.000
$\Gamma_{15\text{-}24}$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$\Gamma_{25\text{-}34}$	0.950	0.950	0.950	0.950	0.950	0.950	0.950	0.950	0.950	0.950
$\Gamma_{35\text{-}44}$	0.947	0.947	0.947	0.947	0.947	0.947	0.947	0.947	0.947	0.947
$\Gamma_{45\text{-}54}$	0.889	0.889	0.889	0.889	0.895	0.901	0.907	0.912	0.919	0.918
$\Gamma_{55\text{-}64}$	0.875	0.875	0.875	0.875	0.882	0.890	0.897	0.904	0.912	0.911
$\Gamma_{65\text{-}74}$	0.714	0.714	0.714	0.714	0.750	0.786	0.821	0.857	0.893	0.893
$\Gamma_{75\text{-}84}$	0.500	0.500	0.500	0.500	0.563	0.625	0.688	0.750	0.812	0.813

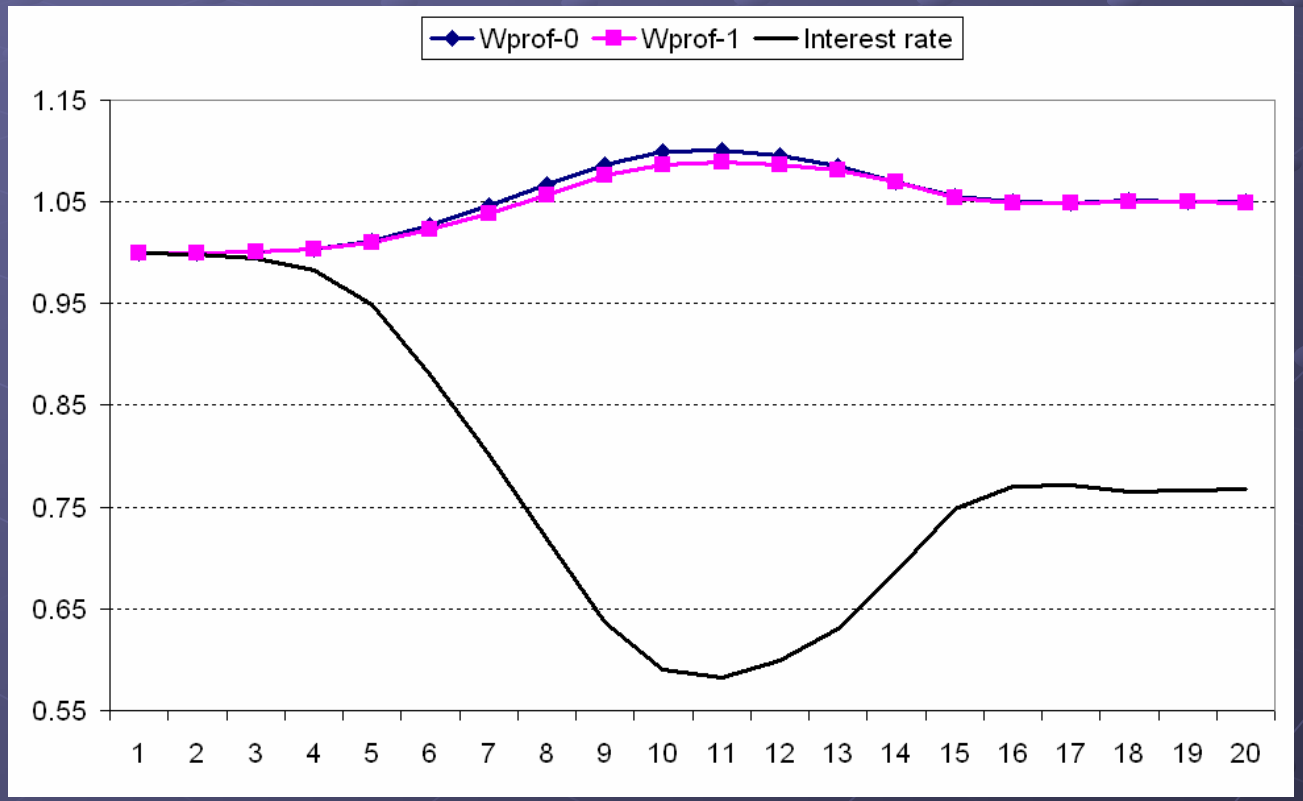
The Macroeconomic shock

Evolution of the old-age dependency ratio



Macroeconomic Results

Impact on factor prices



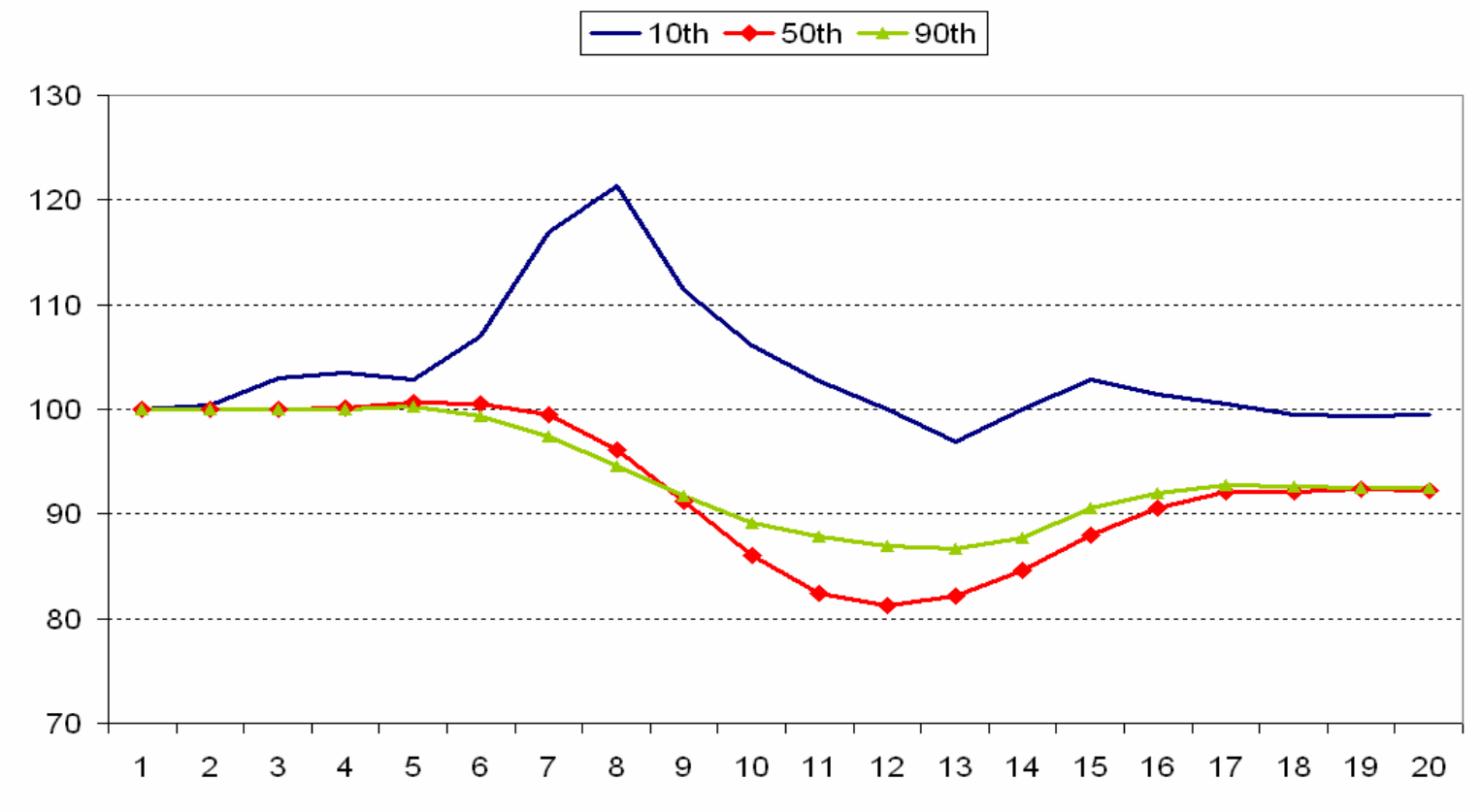
Microsimulation model

- **We introduce the GE time-path of factor prices into the Microsimulation model**

- **We determine the effects on the dynamic of:**
  - **labor choices at the individual level**
  - **individual income (labor income + capital income)**
  - **income distribution**
  - **inequalities**

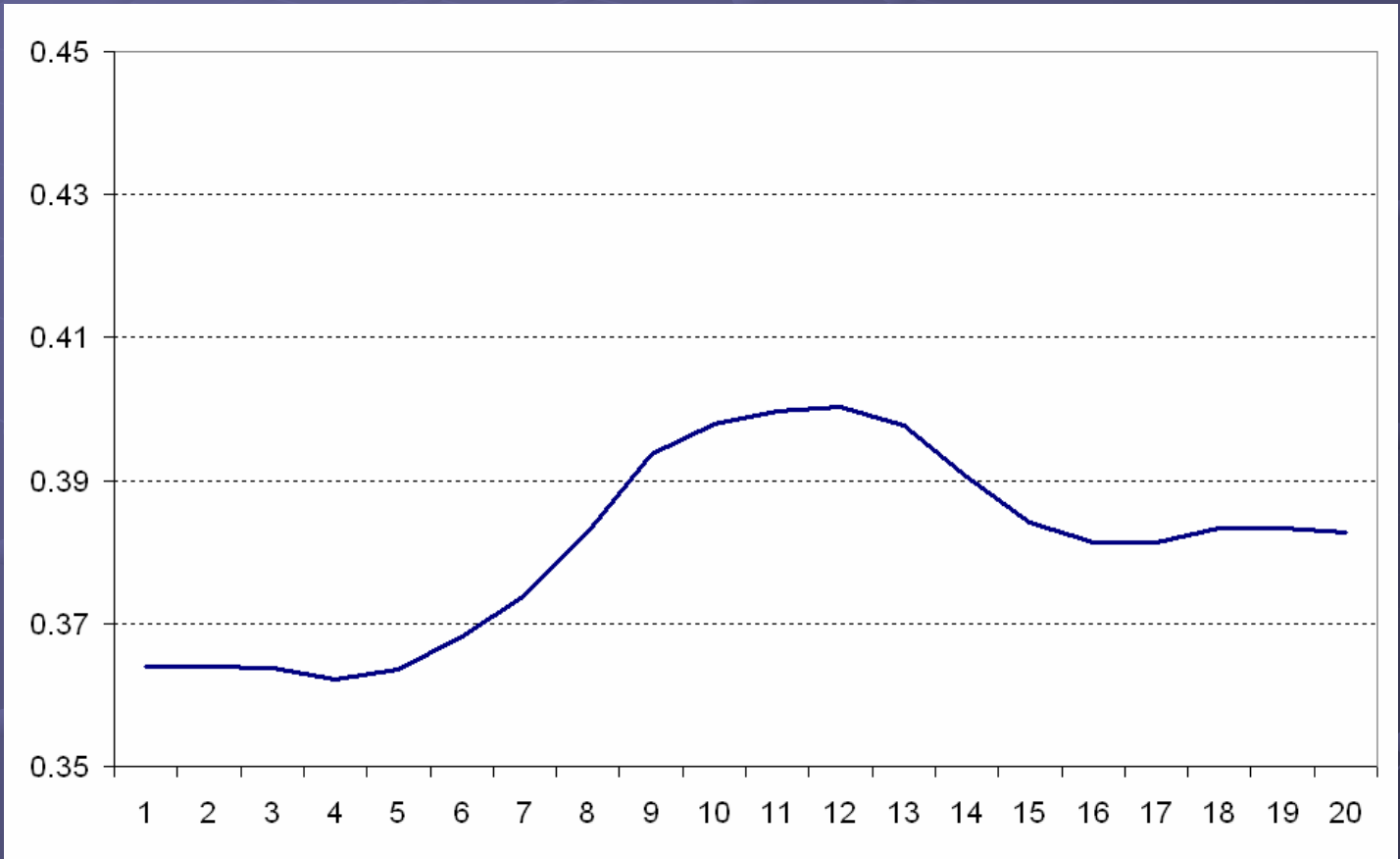
Microeconomic Results

Evolution of the 10th, 50th and 90th percentiles



Microeconomic Results

Evolution of the Gini Index
age group 45-54



Conclusions

- **We developed a methodology that allows to integrate GE and Microsimulation approaches by the aggregation of individual discrete choices**
- **It is not necessary to work with a big number of individuals in the CGE model**
- **It is not necessary to iterate between the two models, thanks to the exact aggregation property**

Restrictions in order to guarantee the exact aggregation property

- This aggregation property holds if the labor supply computed in the CGE model coincides with the labor supply computed in the microsimulation model
  - Given a shock, the average variation in the net wage in the general equilibrium model must coincide with the average variation in the net wage in the microsimulation model
  - Fiscal rules must be simple: we need a simple system of taxes and benefits that allows to link in a simple way the gross wage to the net wage

Future research

- **Implementation of this methodology to the Canadian case in the context of population ageing**
- **FMGD - *Fichier de microdonnées à grande diffusion* - 2001**
- **Individuals choose:**
  - **Whether to work or not**
  - **In which profession (10 professions)**
  - **The type of the contract (Full-time or Part-time)**
  - **The investment in education (5 education levels)**

Technical aspects

- Estimation of a nested (3-level) multinomial logit
- Generation of a potential wage for each non-observed option
  - Estimation of a wage equation with correction of the selection bias when selection is specified as a multinomial logit (Lee (1983), Dubin and McFadden (1984), Dahl (2002), Bourguignon et al. (2007))
- Generation of Gumbel error terms
  - Correlated for the options belonging to the same nest
  - Uncorrelated for the options belonging to different nests