



Helsinki
Center
of
Economic
Research

Discussion Papers

Micro meets Macro via Aggregation

Jussi Lintunen, Olli Ropponen and Yrjö Vartia

University of Helsinki, Department of Economics
and HECER

Discussion Paper No. 259
March 2009

ISSN 1795-0562

HECER – Helsinki Center of Economic Research, P.O. Box 17 (Arkadiankatu 7), FI-00014
University of Helsinki, FINLAND, Tel +358-9-191-28780, Fax +358-9-191-28781,
E-mail info-hecer@helsinki.fi, Internet www.hecer.fi

Micro meets Macro via Aggregation*

Abstract

This paper studies the fundamentals of the aggregation problem. The concept is most central in understanding the relations between micro and macroeconomics. Though aggregation is mentioned often it is not studied explicitly in this generality - only some special cases are governed so far. The study shows how the macro relations arise from the micro relations via aggregation. In addition, we try to give a reader an idea of under what kind of conditions the representative agent approximation is appropriate.

JEL Classification: B41, C43, C81, C82

Keywords: Aggregation, Micro Foundations, Methodology of Economics.

Jussi Lintunen

Researcher
Finnish Forest Research Institute (Metla)
P.O. Box 18 (Jokiniemenkuja 1)
FI-01301 VANTAA
FINLAND

e-mail: jussi.lintunen@metla.fi

Olli Ropponen

University Lecturer
University of Helsinki
Department of Economics
P.O. Box 17 (Arkadiankatu 7)
FI-00014
FINLAND

e-mail: olli.ropponen@helsinki.fi

Yrjö Vartia

Professor of Econometrics
University of Helsinki
Department of Economics,
P.O. Box 17 (Arkadiankatu 7)
University of Helsinki
FI-00014
FINLAND

e-mail: yrjo.vartia@helsinki.fi

* We thank Vesa Kanninen and the participants of the postgraduate seminar for supporting comments.

1 Introduction

Aggregation is mentioned in many cases and almost in every branch of economics. The importance in studying the aggregation problem lies in the frequency it is met. Who really knows what is going on in the aggregation problem? In 1946 the aggregation problem was studied by May, Pu and Klein. They all had a different approach to this problem, because they all wanted to get answers to somewhat different questions. As Klein was studying the aggregation problem, he wanted to know the conditions under which the aggregate variables were enough to explain other aggregate variable. The conditions for these are quite restrictive and exclude many of the real life phenomena out of the scope. Unlike in Pu's approach, the underlying distribution did not matter in Klein's approach. Pu wanted to know whether there exists a distribution that could be responsible for an observed macro variable. May's approach differ from the above two as he starts with the micro equations. The generality of his study is reduced as he has the solutions for the micro equations that are restricted to the case in which the theoretical equilibrium constraints are binding. Our analysis is closest to May's approach. The differences appear as we relax the assumptions that May has and we analyze aggregation explicitly, not implicitly like May. As we have different approaches when studying the aggregation, we are faced with an unclearly stated problem which also happens to appear in many places. This unclearness is pointed out multiple times by Felipe and Fisher in their paper (2003) that provides a survey on the theoretical literature on aggregation of production functions and especially what an applied economists should know about the aggregation. Independently of the approach, the problem of aggregation is about how to fit together three things: micro equations, aggregation rule and macro equations.

The aggregation problem seems to be rarely understood in all its glory and for this reason we try to clarify a reader about the concept. As we are studying this problem we are interested in how the micro and macro systems are related with each other. We concentrate here on the core of the aggregation problem: How does macro arise from a known micro system? By leaving out all the randomness in this paper we emphasize aggregation being a logical, mathematical step between micro and macro - just like calculating

an average from a given set of numbers is a logical step. One of the problems in studies on the aggregation seems to be that the aggregation problem is mixed with the other problems like the ones of modelling or estimation.¹ By leaving the randomness out, we can isolate the aggregation problem of these other arising problems.

In addition to already mentioned problems in aggregation, the third problem arising when studying aggregation seems to be the difficulty of the task. According to Lewbel (1989), representative consumer models are typically employed when one wants to ignore the complications caused by aggregation. Or as Kirman (1992) puts it: "To many macroeconomists, the aggregation problems of the sort implied by the research just described look difficult enough that the simplification of the representative individual looks more attractive, rather than less".

There has been a number of studies that are concerned with the aggregation at some level. The papers by Grunfeld and Griliches (1960) and Pesaran et al. (1989) both deal with choosing between micro and macro regression equations. They treat micro and macro as substitutes, which is very opposite to what we are doing here. We want to find a bridge between the micro system and the corresponding macro system. We want to unite these, not separate.

Also Klein (1946a, 1946b) has a different approach than we do. There is a wide technical literature on Klein's approach, where macro relations are assumed to be similar to micro relations. Leontief, Nataf and Gorman as Classic experts of Klein-aggregation have shown, that it leads to very restrictive separability and additivity assumptions (see their references). The strict conditions under which his approach works exactly and the aggregates as such are enough to describe the situation are shown by Nataf (1948) and Gorman (1953, 1959).² More recent experts are Green (1964), Fisher (1992) and Pokropp (1972, 1978). If the conditions are not met, we face aggregation bias³ by using Klein's approach. This is the bias because of

¹It seems that the aggregation problem is sometimes even confused with the estimation problem with aggregate variables.

²It is emphasized also by Ando (1971), Gupta (1971) and Browning (1993) that the aggregate variables are not enough to describe the situation on their own.

³The studies by Theil (1957), Gupta (1971) and Buse (1992) give a description for this

not aggregating, but estimating with aggregate variables. In this case the macro parameters are biased and according to Gupta, the existence of the aggregation bias can sometimes completely distort, not only the magnitude, but also the signs of the macro parameters. As Nataf and Gorman show the exact conditions, we try to give a description of when the representative consumer approximation is plausible (even if not exact) and thus parallel the Klein's approach in this sense. The general conclusion seems to be that Klein-aggregation is practically impossible (see Felipe and Fisher (2003) chapters 3-5 for discussion).

The notion of the macro relations not being similar to the micro relations is made both in linear cases⁴ (Theil (1959), Klock (1961), Ando (1971), Gupta (1971), Pesaran et al (1989)) and in nonlinear cases (Muellbauer (1975), Stoker (1986), Buse (1992), Vartia (2008a, 2008b, 2009)). Because of these differences, which can be seen as the aggregation bias, the representative agent approximation is not an appropriate approximation to describe the macro in general. Carroll (2000) thinks that in many cases the representative agent should have no future as he states: "For many purposes, the representative-consumer model should be abandoned in favor of a model that matches key microeconomic facts." In our study, we show what are the causes that separate macro, that is aggregated from micro system, from the representative consumer approximation. This way we can give for the macro the micro foundations and deduce the conditions under which the representative-agent approximation is plausible.

The basic problems preceding the logical step of aggregation are pointed out by Blundell and Stoker (2005). Because, we can define the rule for the aggregate⁵ in many ways, how to choose what the appropriate aggregate is. After the choice of the aggregation rule we still have to choose what are the suitable micro and macro levels to work with. In this paper, we take these as given as also the knowledge of the micro system. Given these choices, we are faced with the step that we call the aggregation step. In this step, we deduce what kind of macro arises from these via aggregation and for what

very central concept appearing in the context of aggregation.

⁴The linearity is a special case and this makes aggregation much simpler than in the case of non-linearities. This was noticed by De Wolff already in 1941.

⁵For example averages and index numbers constitute sets of aggregates.

reasons this deviates from the representative consumer approximation. So far the aggregation is studied explicitly only in the case of common behavior for every agent.⁶ In the study, we separate the inputs from the behaviors and like in Blundell and Stoker, we allow the heterogeneity in the inputs, but we also explicitly allow for the heterogeneity in the behaviors. So, our paper tries to give an intuition by explicitly showing what kind of macro arises from the known micro system and we also show the way to reduce our analysis to the one by Blundell and Stoker.⁷ In physics the particles with similar measurable properties do behave similarly, but this is not the case for the economical agents and in this paper we try to take that into account. Our paper explicitly studies the general case for different behaviors and tries to clarify the departures from the representative consumer approximation in this context.

The paper is organized as follows: In section 2 we introduce the framework. Section 3 shows how the aggregation of the micro system is done. Section 4 then takes a close look at the properties of the aggregate. Finally, in section 5 we show how our analysis reduced to the analysis by Blundell and Stoker.

2 Framework

Now we introduce the framework for the case that we call the standard case of aggregation (SCA).

- 1° The set of agents H^t is finite, $n^t = \text{card}(H^t)$, $H^t = \{a_1^t, \dots, a_{n^t}^t\}$ and stationary or slowly changing. Upper index refers to disjoint and consecutive time periods of equal length, typically years or quarters.
- 2° The micro level outputs $y^t(a)$, $a \in H^t$, may depend on a finite number of inputs $x_k^t(a)$, $k \in \{1, \dots, K\}$, via regular agent-wise behavioral functions $f_a \in \Phi(\mathbb{R}^K \supset \Omega \rightarrow \mathbb{R})$, $y^t(a) = f_a(x^t(a)) \in \mathbb{R}$, $x^t(a) \in \Omega$.⁸

⁶Heterogeneity in *inputs* is studied before by Ando (1971), Gupta (1971), Buse (1992) and Blundell and Stoker (2005). Blundell and Stoker give implicitly the freedom for behaviors to differ.

⁷We show how the aggregation factors relate to our notation.

⁸Also $f_a(x)$ may be allowed to be t-dependent, $f_a(x, t) = f_a^t(x)$.

Micro level outputs can be collected together into the micro system

$$y^t = f(x^t) = A(f, x^t).$$

3° Macro output Y^t is the total of $y^t(a)$, that is $Y^t = \sum_{H^t} y^t(a)$. Thus the aggregation rule, AR, in SCA is

$$AR(y^t, n^t) = n^t \langle y^t \rangle,$$

where $\langle y^t \rangle$ is the per capita output or average output per agent.

4° We treat H^t , all behaviors f_a and their inputs $x^t(a)$ as known data. This excludes all estimation problems from SCA.

5° The real problem in SCA is how to represent the dynamics of $Y^t = AR(y^t, n^t) = \sum_{H^t} f_a(x^t(a))$ in a comprehensible way. Formally $Y^t = n^t S(y^t, n^t)$, where the Synthesis Operator depends on all the functions f_a and all the inputs $x^t(a)$, $a \in H^t$.

Suppose that H^t is the set of the Finnish households, $y^t(a)$ is yearly consumption and consumption is explained by (at most) 20 input variables. Now $Y^t = n^t S(f, x^t)$ is a functional-function, which depends on 2 million behaviors (possibly heterogeneous consumption functions) f_a and of 40 million inputs. The problem in 5° is a difficult one, but not impossible, as the following special case shows. Assume all f_a 's are affine, $f_a(x) = \alpha(a) + \sum_{k=1}^K \beta_k(a)x_k$. If they are parallel with each other, then $f_a(x) = \alpha(a) + \sum_{k=1}^K \beta_k x_k$, where only the intercepts $\alpha(a)$ may differ. Now $\frac{Y^t}{n^t} = \langle y^t \rangle = \frac{1}{n^t} \sum_{H^t} f_a(x^t(a)) = \langle \alpha \rangle + \sum_{k=1}^K \beta_k \langle x_k^t \rangle$ or $Y^t = n^t \langle \alpha \rangle + \sum_{k=1}^K \beta_k X_k^t$. Here $X_k^t = n^t \langle x_k^t \rangle =$ the total of the input $x_k^t(a)$. This is standard "linear" (actually affine) and static text-book model in macro economics.⁹ Macro economists probably apply this as their prototype model, where e.g. dynamics is added by including similar lagged effects. Its coefficients are based partly on time series estimations (usually on the macro series), micro reasoning, guessing and opinions - to use Leamer's (1983) terms. The obvious simplifications and errors of the model and poor thinking are ignored at this stage or covered with endless rationalizations, details and excuses. Thus the views, estimations and opinions of

⁹Including the changing size of population n^t .

the macro behaviors differ considerably from one economist to another - or rather between different schools of thought in economics. This hides the fact that the real macro dependencies are much more stable than these opinions of it. We hope to make this strong and important conjecture at least understandable. Macro dependencies in Economics are context dependent and vary slowly in time, but once the circumstances are specified they are very stable. Macro economy moves like a train or passenger ship whenever it is going.

The mathematical cause and explanation of this is the Cancellation of Details in the mean. The Laws of Large Numbers (LLN) and the Central Limit Theorems (CLT) are based on that. Simple states, consider the mean $\langle x \rangle = \frac{1}{1000000} \sum_A x(a)$ over one million, arbitrary real values $x(a)$. Add random errors or perturbations $\varepsilon(a)$ with zero expectation and constant variance σ^2 to produce $\langle x + \varepsilon \rangle = \frac{1}{10^6} \sum (x(a) + \varepsilon(a))$. Perturbations are assumed to be independent of x 's and within themselves. Their distribution may be arbitrary and heterogenous once $\varepsilon(a) \sim IND(0, \sigma^2)$. Now $\langle x + \varepsilon \rangle = \langle x \rangle + \langle \varepsilon \rangle$, where $\langle \varepsilon \rangle \approx N(0, \frac{\sigma^2}{10^6})$. Set concretely, let $\langle x \rangle \approx 1000$ Euro and standard deviation of x is $s(x) \approx 200$ Euro. In addition, let the standard deviation of the perturbation be $\sigma = D\varepsilon = 100$ Euro, that is about 50% of the standard deviation of the original variable. Therefore the perturbed $x(a) + \varepsilon(a)$ typically differs from $x(a)$ within $\pm 2\sigma = \pm 200$ Euro, quite a lot. But $D\langle \varepsilon \rangle = \frac{\sigma}{\sqrt{10^6}} = \frac{\sigma}{1000} = \frac{100}{1000} = 0.1$. Thus $\langle x + \varepsilon \rangle$ differs from $\langle x \rangle$ typically within ± 0.2 Euro, which is totally negligible in comparison to the unperturbed x , which is being of the order of magnitude of 1000.

3 Aggregation of a micro system

3.1 Micro system

Let there be n micro units or agents a_i which define a set

$$H = \{a_1, \dots, a_n\}. \quad (1)$$

For every a_i we have a behavioral equation

$$y(i) = f_i(x(i)), \quad (2)$$

where we have output, behavior and inputs, $y_i \in \mathbb{R}$, $f_i \in \Phi(\mathbb{R}^k \rightarrow \mathbb{R})$ and $x(i) \in \mathbb{R}^k$ respectively. Outputs, behaviors and inputs can differ with *every* agent i.e. we have not enforced any of these to be the same. Now all the micro information is gathered in the set $\{y(i), f_i, x(i)\}$. The micro system can be presented in vector form by stacking the individual equations as

$$y = f(x) = \begin{pmatrix} f_1(x(1)) \\ \vdots \\ f_n(x(n)) \end{pmatrix}.$$

In this paper we take the functional forms (behaviors) as known. It is obvious, that in practice the functional forms must be modelled from the micro level data.¹⁰ However, that is not the point of this paper and could lead us astray from the main focus. When the behaviors of agents are known we can analyse the situation. It is useful to decompose the behavior to the common and deviating parts, \bar{f} and δf_i , as done in Vartia (2009). Let us define the average behavior to be ¹¹

$$\bar{f} = \frac{1}{n}(f_1 + \dots + f_n). \quad (3)$$

The deviating behavior is defined using the average behavior as $\delta f_i = f_i - \bar{f}$. With these definitions we can decompose the behaviors of the micro system into two vectors of functions:

$$f = \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix} = \bar{f}\mathbf{1} + \delta f = \begin{pmatrix} \bar{f} \\ \vdots \\ \bar{f} \end{pmatrix} + \begin{pmatrix} \delta f_1 \\ \vdots \\ \delta f_n \end{pmatrix},$$

where $\mathbf{1}$ is $(n \times 1)$ -vector of ones. With this (useful) decomposition we have the common and the deviating parts for behavior. We illustrate the decomposition of micro behaviors with three agents and equal weights in Figure 1.¹² The average behavior is the vertically calculated average of the individual behaviors.

¹⁰These may be modelled from the data for some micro level groups that are homogenous enough.

¹¹With sampling weights one has to use weighted average $\bar{f} = \sum_i w_i f_i$, where $w_i = c_i / \sum c_j$ and $c_i > 0$ for all i .

¹²Functions used here are of the form $A_i x^{b_i}$ where $A = (0.8, 1.0, 1.3)$ and $b = (0.3, 0.8, 0.8)$.

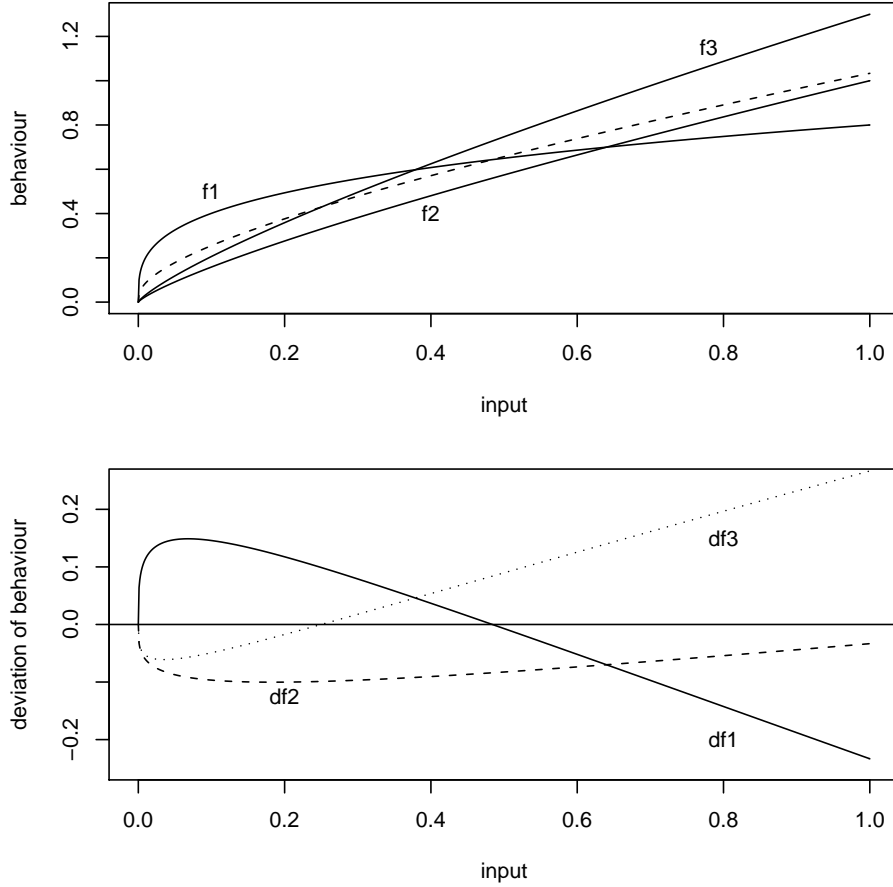


Figure 1: Illustration of the average behavior. In the upper graph three different behaviours, f_i , are drawn with solid lines and the average behaviour, \bar{f} , with the broken line. In the lower graph the deviations from the average behaviour, δf_i , are drawn for the different micro units.

Now one can present the output of a micro unit as

$$y(i) = \bar{f}(x(i)) + \delta f_i(x(i)). \quad (4)$$

Note that we do not have a standard notation to present this relation for the whole micro system in a vectorized form. This problem can be solved by introducing an Analysis operator, which is described in Appendix A. However, we can decompose the micro relation even more to bring forth the component of relation that is common to all the micro units, i.e.

$$y(i) = \bar{f}(\bar{x}) + [\bar{f}(x(i)) - \bar{f}(\bar{x})] + \delta f_i(x(i)). \quad (5)$$

The first component is the output of the representative consumer, the second term, in parenthesis, is deviation of output due to the differences in inputs and the last term the effect of deviating behavior. In the next section we formulate an aggregate of these decomposed individual equations.

3.2 Aggregation

Now let us turn to the aggregation of the micro outputs. We define aggregation as a calculation of a macro index of the micro outputs. Since the micro outputs depend on micro behaviour and inputs, it is obvious that the macro index, i.e. the aggregate, depends on those as well.

The aggregate we are studying here is the average¹³ which we denote by

$$\bar{y} = S(f, x) \doteq n^{-1} \sum_{i=1}^n f_i(x(i)). \quad (6)$$

Let us now calculate an aggregate of a set of known micro relations. Since the result is based on equation (5), we can readily state this as a theorem:

Theorem 1 (Decomposition of the aggregate) *The aggregate can be decomposed into the three parts as follows:*

$$\bar{y} = RB(\bar{x}) + NLE(x) + HE(x), \quad (7)$$

where

$$RB(\bar{x}) = \bar{f}(\bar{x}), \quad (8)$$

$$NLE(x) = n^{-1} \sum_{i=1}^n \{\bar{f}(x(i)) - \bar{f}(\bar{x})\}, \quad (9)$$

$$HE(x) = n^{-1} \sum_{i=1}^n \delta f_i(x(i)). \quad (10)$$

The first term on the right hand side, $RB(\bar{x})$, is the representative behavior with the mean inputs. The two other terms are, in a sense, unwanted. The second term, $NLE(x)$, arises from the nonlinearity of the mean behavior function and is called the nonlinearity effect. The name is based on a fact that the term is zero if there are no nonlinearities in the mean behavior, i.e.

¹³All the results can be derived also with weighted averages of outputs and behaviors (weighted by the sampling weights).

\bar{f} is affine. The third term, $HE(x)$, arises from the differences among agents' behaviors and is called the heterogeneity effect. This term vanishes if δf 's are constants that sum to 0. In that case agents' behaviours are parallel with each other. When all the inputs are identical, $x(i) = x(j)$ for all i, j , both $NLE(x)$ and $HE(x)$ vanish, but that is an uninteresting special case. In that case the distribution of the inputs would be reduced to a single point in the input space.

4 Properties of the aggregate

4.1 Level of the aggregate

The aggregation of the micro relations brought us three terms, where two of them are considered as a nuisance or at least a complication. If the terms $NLE(x)$ and $HE(x)$ are important in terms of level or dynamics of the macro output, the representative consumer assumption can be considered invalid. Therefore it is of the utmost importance to study the properties of these two terms.

It is clear, that if we have the same behavior for all the agents and the output is an affine transformation of the inputs ($x \in \mathbb{R}$)

$$f(x(i)) = a + bx(i), \quad (11)$$

then the average of the output is a function of the average of the corresponding inputs \bar{x} :

$$\bar{y} = n^{-1} \sum_{i=1}^n f(x(i)) = n^{-1} \sum_{i=1}^n a + n^{-1} \sum_{i=1}^n bx(i) = a + b\bar{x}. \quad (12)$$

So the macro output is a function of a macro input (the average \bar{x}). What is worth noticing is that in the absence of the Hamel solutions the aggregate output is a function of aggregate input only if and only if the functional form of f is affine (see Vartia (2009) pp. 14-16).

To illustrate the problem of aggregation we show that with only a slight modification to the previous restrictive model we lose the nice property, that macro output is a function of macro inputs, in a sense an exact aggregation. Functional form is still affine, but now the behaviors can vary, i.e. for agent i

we have $y_i(x(i)) = a_i + b_i x(i)$. Now the average output for the whole economy is¹⁴

$$\begin{aligned}
\bar{y} &= \frac{1}{n} \sum_{i=1}^n y_i(x(i)) = \frac{1}{n} \sum_{i=1}^n a_i + \frac{1}{n} \sum_{i=1}^n b_i x(i) \\
&= \frac{1}{n} \sum_{i=1}^n a_i + \frac{1}{n} \sum_{i=1}^n \bar{b} x(i) + \frac{1}{n} \sum_{i=1}^n \delta b_i x(i) \\
&= \bar{a} + \bar{b} \cdot \bar{x} + \frac{1}{n} \sum_{i=1}^n \delta b_i x(i). \tag{13}
\end{aligned}$$

Now the aggregate is not only a function of mean behaviour in mean input but there is also an covariance type of term. This extra term is an emergent macro property, which is not typically ready for hand for the macroeconomists. This additional term arises from the differences in behaviors among the agents.¹⁵ This is the case also in general: the macroeconomy does not follow the same functional form as the behaviors of the individuals. Neither are the averages enough to describe the macroeconomy in general. One needs higher moments of the (joint) distributions of the variables and model parameters.

From this simple example it is clear that only in a case of very strict assumptions on micro relations, one can find a macro relation of the similar form. However, from the general result of Theorem 1 it is difficult to see the nature of these emergent macro terms. Therefore we next consider the 2nd order polynomial approximation of the behaviors f_i .

It is a known result (Vartia 2008b) that when the micro relations are polynomials of order two (with two variables)

$$\psi_i = \alpha_i + \beta_{1i}x_i + \beta_{2i}y_i + \beta_{11i}x_i^2 + \beta_{22i}y_i^2 + 2\beta_{12i}x_i y_i, \tag{14}$$

then for $\bar{\psi}$ ($\bar{\psi} = \frac{1}{n} \sum_{i=1}^n \psi_i$) we get

$$\begin{aligned}
\bar{\psi} &= \underbrace{\bar{\alpha} + \bar{\beta}_1 \cdot \bar{x} + \bar{\beta}_2 \cdot \bar{y} + \bar{\beta}_{11} \cdot \bar{x}^2 + \bar{\beta}_{22} \cdot \bar{y}^2 + 2\bar{\beta}_{12} \cdot \bar{x} \cdot \bar{y}}_{RB} \\
&+ \underbrace{\bar{\beta}_{11} var(x) + \bar{\beta}_{22} var(y) + 2\bar{\beta}_{12} cov(x, y)}_{NLE} \\
&+ \underbrace{cov(\beta_1, x) + cov(\beta_2, y) + cov(\beta_{11}, x^2) + cov(\beta_{22}, y^2) + 2cov(\beta_{12}, xy)}_{HE}, \tag{15}
\end{aligned}$$

¹⁴ $\delta b_i = b_i - \bar{b}, \delta x(i) = x(i) - \bar{x}$.

¹⁵In fact this depends on the joint distribution of b and x .

where \bar{z} denotes the mean, $cov(v, z)$ is the covariance between v and z and $var(z)$ is the variance of z . It is obvious that aggregation has yielded a lot of extra structure on the macro level relation.

These results for the second order polynomials hold for more general models, when one is using the second order Taylor approximation of the micro functions. This approximation in the neighborhood of \bar{x} for the *NLE* (9) yields

$$\sum_{i=1}^n [\bar{f}(x(i)) - \bar{f}(\bar{x})] \approx \frac{1}{2} \sum_{i=1}^n (x(i) - \bar{x})' D^2 f(\bar{x}) (x(i) - \bar{x}). \quad (16)$$

Quadratic form of the equation can be calculated and the terms one gets are covariances of inputs of micro units multiplied with Taylor coefficients. One can do similar calculations for the HE term also and these calculations give sums that are related to covariances between Taylor coefficients and values of inputs and their squares. In the case of two variables, the results are identical to ones presented in the case of quadratic micro equations with two inputs (see eq. (15)).

It is worth noticing that the NLE-terms and HE-terms do not have such a clear meaning as for example the marginal rate of consumption. Maybe this is why these are not studied so much. Despite the fact that their meaning is not that straightforward, they must not be omitted.

The more elegant representation about this section with the operators is given in appendix A.

4.2 Dynamics of the aggregate

Now we have examined the levels of macro behavior derived from the heterogeneous micro behaviors with differing circumstances. Next step is to consider the effect of changes in circumstances to the macro behavior. The brief analysis presented here is done with infinitesimal changes in input vectors and some examples are given in graphs.

Changes in circumstances used here is an infinitesimal change in input vector $x(i)$ for all micro units a_i . Then the change in the mean input vector is given by equation

$$d\bar{x} = \sum_{i=1}^n w_i dx(i). \quad (17)$$

Let us decompose the change of output of a micro unit as follows:

$$\begin{aligned}
dy(i) &= df_i(x(i)) \\
&= d[f_i(\bar{x}) + f(x(i)) - f_i(\bar{x})] \\
&= f_i'(\bar{x})d\bar{x} + [f_i'(x(i))dx(i) - f_i'(\bar{x})d\bar{x}]. \tag{18}
\end{aligned}$$

Notice that derivative of behavior is $(1 \times k)$ gradient vector. To ease the aggregation phase one can decompose the change of the output a bit further by adding and subtracting $f_i'(x(i))d\bar{x}$

$$dy(i) = f_i'(\bar{x})d\bar{x} + [f_i'(x(i)) - f_i'(\bar{x})] d\bar{x} + f_i'(x(i)) [dx(i) - d\bar{x}]. \tag{19}$$

The change in macro behavior can be calculated normally as a weighted average

$$d\bar{y} = \sum_{i=1}^n w_i dy(i). \tag{20}$$

Using equation (19) one reaches directly

$$\begin{aligned}
d\bar{y} &= d\bar{x} \left\{ \bar{f}'(\bar{x}) + \sum_{i=1}^n w_i [f_i'(x(i)) - f_i'(\bar{x})] \right\} \\
&\quad + \sum_{i=1}^n w_i f_i'(x(i)) [dx(i) - d\bar{x}], \tag{21}
\end{aligned}$$

since by commutation of summation and differentiation $\sum_i w_i f_i'(\bar{x}) = \bar{f}'(\bar{x})$. By using the identity $f_i = \bar{f} + \delta f_i$ and noticing that $f_i' = \bar{f}' + \delta f_i'$ one ends up with a presentation

$$\begin{aligned}
d\bar{y} &= d\bar{x} \left[\bar{f}'(\bar{x}) + \sum_{i=1}^n w_i (\bar{f}'(x(i)) - \bar{f}'(\bar{x})) + \sum_{i=1}^n w_i (\delta f_i'(x(i)) - \delta f_i'(\bar{x})) \right] \\
&\quad + \sum_{i=1}^n w_i f_i'(x(i)) [dx(i) - d\bar{x}], \tag{22}
\end{aligned}$$

which is similar to equation (7). The last term, however, is a new one and it has a clear covariance structure. Indeed, if the input vector is one-dimensional then the extra term would be $cov(f'(x), dx)$. With multiple inputs, this term is a sum of such covariances.

By examining equation (22) one notices that everything said about equation (7) holds here, with only one exception. Here we have derivatives of

behavior, so non-linear effect appears only for third and higher order polynomials. The new term has clear implications. If changes in inputs are correlated with the behavior at the margin, then changes of the representative behavior are not adequate approximations of the macro behavior.

Let us next consider two main situations of input changes: growth and equalization of inputs. When growth happens, typically all the inputs increase. Two different growth scenarios are illustrated in Figure 2. In these

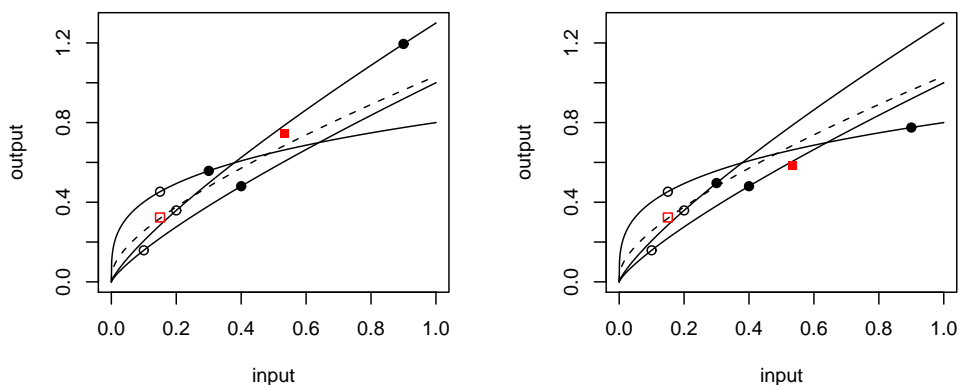


Figure 2: Two cases of growth. Initial states and total growth of the inputs is same in both figures. Despite of that, the case on the left graph has higher growth in total output. Filled symbols mark the final state.

two cases the initial states are identical and total growth of inputs is equal. Difference is in distribution of the new inputs. On the left hand side most of the growth is given to the micro unit, which has the steepest slope. On the right hand side growth is given to the micro unit, which has the most gentle slope. Growth measured as the change in the macro output is naturally higher in former case. Difference relates especially to the new covariance term in equation (22).

The other interesting situation of input change is the case of equalization of inputs. We concentrate here only to the cases where the total amount of the inputs remains constant. Let us examine two different cases, which are illustrated in Figure 3. Now we have identical final states. Difference in these two cases lies on who has to give inputs to the other micro units. On the left

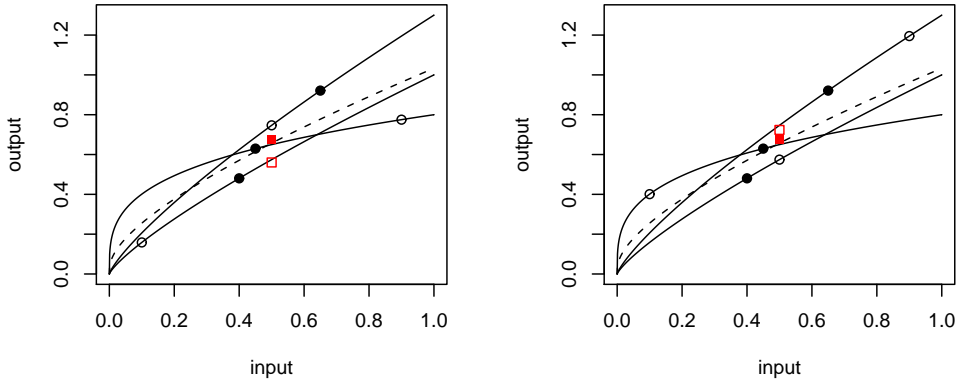


Figure 3: Two cases of equalization of inputs. Final states are identical and total amount of inputs remains the same. Still in the left hand side total output grows and on the right hand side it turns down. Filled symbols mark the final state.

hand side the micro unit with the most gentle slope gives his inputs to others. On the right hand side the agent giving inputs has the steepest slope. The consequence is that in former case the total output increases and in latter case it decreases. The situation is similar to the growth scenarios but now the sum of changes in inputs is zero. Difference between two equalization cases is completely caused by the new covariance term.¹⁶

In the above sections we have divided the macro behavior (MB) into three different components. These are representative behavior (RB), non-linearity effect (NLE) and heterogeneity effect (HE). Despite of the fact that NLE and HE are the deviations of RB from MB, NLE and HE are *not* necessarily small compared to MB. This can be seen for example from equation (15). There we have covariances and variances of variables in NLE and these are not restricted to be small compared to MB. In HE we have covariances between variables and parameters, which are neither restricted to be small. So NLE and HE can be significant determinants of MB.

Despite NLE and HE are not necessarily small, they seem to be very stable. We illustrate this in figure 4, where we have three agents. The first of the

¹⁶This is because $d\bar{x} = 0$ (see eq. (22)).

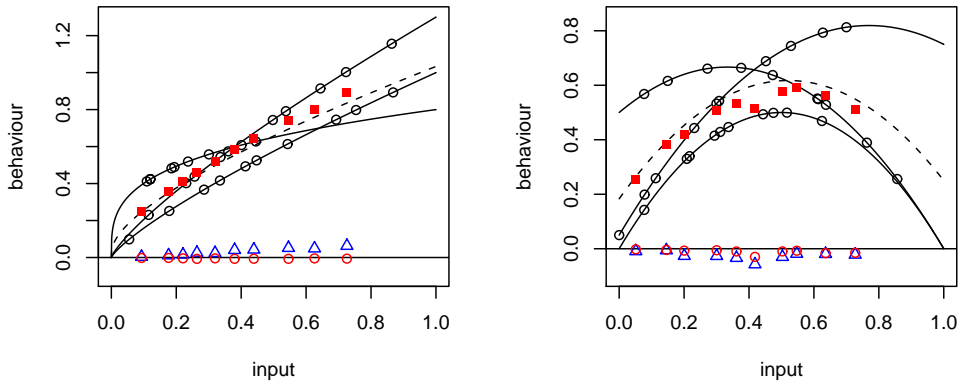


Figure 4: Results of the two Monte Carlo experiments. MB's are drawn with filled symbols, NLE's with triangles and HE's with red circles. In the left figure behaviors are exponential ($A_i x^{b_i}$) and in the right quadratic ($a_i + b_i x + c_i x^2$).

graphs has the same behaviors as figure 2. Macro behaviors are drawn with filled symbols, non-linearity effects with triangles and heterogeneity effects with circles (close to the horizontal axis). In the figure we have illustrated how the growth (in inputs) could take place. Here we have first randomly chosen inputs to all three agents and then the macro behavior is calculated. After that we have added random amounts to the original micro inputs and then again calculated macro behavior. Repeating the procedure we have 10 different macro behaviors for which we have calculated NLE's and HE's. In both figures we see that neither NLE nor HE do seem to change very much. They are like 'frozen' to the values they have.¹⁷ In the case of quadratic micro equations these are composed of covariances and variances, which are quite stable. Covariances and variances do not change a lot unless we have some kind of systematic changes in circumstances.

¹⁷Now also both these effects happen to have small values, but this doesn't have to be like this in general.

5 Comparison with Blundell and Stoker

In this section we show how does our analysis differ from the one by Blundell and Stoker in 2005 (BS). Let us rewrite the equation (6) from page 352 in their paper.

$$g_{jt}(p_t, m_{it}, z_{it}) = b_{0j}(p_t)m_{it} + b_{1j}(p_t)m_{it} \ln m_{it} + b_{2j}(p_t)m_{it}z_{it}, \quad (23)$$

where t refers to the period, i the individual, j for good, p_t refers to prices in period t , m_{it} is an income for individual i in period t , z_{it} are the individual attributes for individual i in period t , $b_{0j}(p_t)$, $b_{1j}(p_t)$ and $b_{2j}(p_t)$ are the parameters.

What if we allow (unlike BS) the coefficients b_{0j} , b_{1j} and b_{2j} to differ for different individuals i ? Then we would have

$$g_{ijt}(p_t, m_{it}, z_{it}) = b_{0ij}(p_t)m_{it} + b_{1ij}(p_t)m_{it} \ln m_{it} + b_{2ij}(p_t)m_{it}z_{it}. \quad (24)$$

An aggregate for good j in period t would then be

$$\begin{aligned} \bar{y}_{jt} &= \frac{1}{n} \sum_{i=1}^n g_{ijt}(p_t, m_{it}, z_{it}) \\ &= \frac{1}{n} \sum_{i=1}^n b_{0ij}(p_t)m_{it} + \frac{1}{n} \sum_{i=1}^n b_{1ij}(p_t)m_{it} \ln m_{it} + \frac{1}{n} \sum_{i=1}^n b_{2ij}(p_t)m_{it}z_{it} \\ &= \overline{b_{0j}(p_t)m_t} + \overline{b_{1j}(p_t)m_t \ln m_t} + \overline{b_{2j}(p_t)m_t z_t} \\ &= \overline{b_{0j}(p_t)} \cdot \overline{m_t} + \overline{b_{1j}(p_t)} \cdot \overline{m_t \ln m_t} + \overline{b_{2j}(p_t)} \cdot \overline{m_t z_t} \\ &+ \underbrace{\text{cov}(b_{0j}(p_t), m_t) + \text{cov}(b_{1j}(p_t), m_t \ln m_t) + \text{cov}(b_{2j}(p_t), m_t z_t))}_{HE(x)=HE(p_t, m_t, z_t)}. \quad (25) \end{aligned}$$

Here the last three terms correspond to the heterogeneity effect $HE(x)$ in our notation¹⁸ and emerges from the heterogeneity of agent behaviors. If coefficients b_{0ij} , b_{1ij} and b_{2ij} are same for every individual i (like in BS) then there is no differences in behavior (only in inputs) and thus $HE(x) = 0$. Also if the coefficients are the same $\forall i$, then $\bar{b}_{kj} = b_{kj}$, $k = 0, 1, 2$ and then

$$\bar{y}_{jt} = \underbrace{b_{0j}(p_t)\overline{m_t} + b_{1j}(p_t)\overline{m_t \ln m_t} + b_{2j}(p_t)\overline{m_t z_t}}_{\frac{1}{n} \sum_{i=1}^n \bar{f}(x(i))}. \quad (26)$$

¹⁸Note that x is a vector.

At this stage Blundell and Stoker define the aggregate factors which summarize the impacts of aggregation as they tell how much do the estimates for the representative consumer differ from the true parameter values. They also use the population counterparts for the sample statistics used in our notation. We want to know the causes for the departure from the representative consumer approximation. So we continue from the above:

$$\begin{aligned}
\bar{y}_{jt} &= b_{0j}(p_t)\bar{m}_t + b_{1j}(p_t)\overline{m_t \ln m_t} + b_{2j}(p_t)\overline{m_t z_t} \\
&= b_{0j}(p_t)\bar{m}_t + b_{1j}(p_t)\left(\bar{m}_t \ln \bar{m}_t - \overline{m_t \ln m_t} + \bar{m}_t \cdot \overline{\ln m_t}\right. \\
&\quad \left.+ \text{cov}(m_t, \ln m_t)\right) + b_{2j}(p_t)\left(\bar{m}_t \cdot \bar{z}_t + \text{cov}(m_t, z_t)\right) \\
&= \underbrace{b_{0j}(p_t)\bar{m}_t + b_{1j}(p_t)\overline{m_t \ln m_t} + b_{2j}(p_t)\overline{m_t \cdot z_t}}_{RB(\bar{x})=g_{jt}(p_t, \bar{m}_t, \bar{z}_t)} \\
&\quad + b_{1j}(p_t)\left(\overline{\ln m_t} - \ln \bar{m}_t\right) + b_{2j}(p_t)\text{cov}(m_t, z_t) \\
&= RB(\bar{x}) + \underbrace{b_{1j}(p_t)\left(\bar{m}_t \ln \frac{G(m_{it})}{A(m_{it})} + \text{cov}(m_t, \ln m_t)\right) + b_{2j}(p_t)\text{cov}(m_t, z_t)}_{NLE(x)} \\
&= RB(\bar{x}) + NLE(x), \tag{27}
\end{aligned}$$

where G and A are geometric and arithmetic means respectively. From here we see that the the representative consumer approximation differs from the true aggregate and this is due to non-linearity of the cross-terms.

The heterogeneity due to different behaviors, HE^{19} , has been put to zero²⁰ in the BS unlike in ours. The heterogeneity in inputs is allowed in both papers. The aggregation factors tell us how much the parameter estimates are biased because of using only the aggregate data. So these tell how much the *estimates* for representative consumer differ from the true parameter values. On contrast to BS, we are not analyzing the parameter estimates, but show what are the factors that separate representative consumer from the true dependencies. We answer to a slightly different question than BS: We want to know both why and how much whereas BS want to know how much do we make error when using representative consumer approximation. However, the slightly different questions to be asked show that BS is more practically oriented than our paper, but it is not as general as ours.

¹⁹This term includes plenty of degrees of freedom.

²⁰Parameters are the same for every individual.

Bibliography

ALBERT ANDO (1971), On a Problem of Aggregation; *International Economic Review*, Vol. 12, No.2 (June, 1971), 306-311

RICHARD BLUNDELL, THOMAS M. STOKER (2005), Heterogeneity and Aggregation; *Journal of Economic Literature*, Vol. 18, (June, 2005), 347-391

MARTIN BROWNING (1993), Estimating micro parameters from macro data alone: some pessimistic evidence; *Ricerche Economiche*, Vol. 47, 253-267

ADOLF BUSE (1992), Aggregation, Distribution and Dynamics in the Linear and Quadratic Expenditure Systems; *The Review of Economics and Statistics*, Vol. 74, No.1 (Feb., 1992), 45-53

CHRISTOPHER D. CARROLL (2000), Requirement for the Representative Consumer? Aggregate Implications of Microeconomic Consumption Behavior; *The American Economic Review*, Vol. 90, No.2 (May, 2000), 110-115

JESUS FELIPE, FRANKLIN FISHER (2003), Aggregation in Production Functions: What Applied Economists should know; *Metroeconomica*, Vol. 54, No.3, 208-262

FRANKLIN FISHER (1992), Aggregation, Aggregate Production Functions and Related Topics; *Harvester Wheatsheaf*

W.M. GORMAN (1959), Community Preference Fields; *Econometrica*, Vol. 21, No.1 (Jan., 1953), 63-80

W.M. GORMAN (1959), Separable Utility and Aggregation; *Econometrica*, Vol. 27, No.3 (July, 1959), 469-481

H. A. J. GREEN (1964), Aggregation in Economic Analysis. An Introductory Survey; *Princeton University Press, Princeton, NJ*

YEHUDA GRUNFELD, ZVI GRILICHES (1960), Is Aggregation Necessarily Bad?; *The Review of Economics and Statistics*, Vol. 42, No.1 (Feb., 1960), 1-13

KANHYA L. GUPTA (1971), Aggregation Bias in Linear Economic Models; *International Economic Review*, Vol. 12, No.2 (June, 1971), 293-305

- J.M. HEINEKE, H.M. SHEFRIN (1988), Exact Aggregation and the Finite Basis Property; *International Economic Review*, Vol. 29, No.3 (Aug., 1988), 525-538
- DALE W. JORGENSON, LAWRENCE J. LAU, THOMAS M. STOKER (1980), Welfare Comparison under Exact Aggregation; *The American Economic Review*, Vol. 70, No.2 (May, 1980), 268-272
- ALAN P. KIRMAN (1992), Whom or What does the Representative Individual Represent?; *The Journal of Economic Perspectives*, Vol. 6, No.2 (Spring, 1992), 117-136
- LAWRENCE R. KLEIN (1946a), Macroeconomics and the Theory of Rational Behavior; *Econometrica*, Vol. 14, No.2 (Apr., 1946), 93-108
- LAWRENCE R. KLEIN (1946b), Remarks on the Theory of Aggregation; *Econometrica*, Vol. 14, No.4 (Oct., 1946), 303-312
- T. KLOEK (1961), Note on Convenient Matrix Notations in Multivariate Statistical Analysis and in the Theory of Linear Aggregation; *International Economic Review*, Vol. 2, No.3 (Aug., 1961), 351-360
- EDWARD E. LEAMER (1983), Let's Take the Con out of Econometrics; *American Economic Review*, Vol. 73, No.1 (Mar., 1983), 31-43
- WASSILY LEONTIEF (1947), Introduction to a Theory of the Internal Structure of Functional Relationships; *Econometrica*, Vol. 15, No.4 (Oct., 1947), 361-373
- ARTHUR LEWBEL (1989), Exact Aggregation and a Representative Consumer; *The Quarterly Journal of Economics*, Vol. 104, No.3 (Aug., 1989), 621-633
- KENNETH MAY (1946), The Aggregation Problem for a One-Industry Model; *Econometrica*, Vol. 14, No.4 (Oct., 1946), 285-298
- KENNETH MAY (1947), Technological Change and Aggregation; *Econometrica*, Vol. 15, No.1 (Jan., 1947), 51-63
- JOHN MUELLBAUER (1975), Aggregation, Income Distribution and Consumer Demand; *The Review of Economics Studies*, Vol. 42, No.4 (Oct., 1975), 525-543

- JOHN MUELLBAUER (1976), Community Preferences and the Representative Consumer; *Econometrica*, Vol. 44, No.5 (Sep., 1976), 979-999
- ANDRE NATAF (1948), Sur la possibilite de construction de certains macro-models; *Econometrica*, Vol. 16, No.3 (July, 1948), 232-244
- M.H. PESARAN, R.G. PIERSE, M.S. KUMAR (1989), Econometric Analysis of Aggregation in the Context of Linear Prediction Models; *Econometrica*, Vol. 57, No.4 (July, 1989), 861-888
- F. POKROPP (1972), Aggregation von Produktionsfunktionen, Lecture Notes in Economics and Mathematical Systems; *Springer, Berlin, 1972*
- F. POKROPP (1978), The Functional Equations of Aggregation, Chapter 6 in Functional Equations in Economics, Eichhorn (eds.); *Addison-Wesley Publishing Company, London etc, 1978*
- SHOU SHAN PU (1946), A Note on Macroeconomics; *Econometrica*, Vol. 14, No.4 (Oct., 1946), 299-302
- THOMAS M. STOKER (1984), Completeness, Distribution Restrictions and the Form of Aggregate Functions; *Econometrica*, Vol. 52, No.4 (July, 1984), 887-907
- THOMAS M. STOKER (1986), Simple Tests of Distributional Effects on Macroeconomic Equations; *Journal of Political Economy*, Vol. 94, No.4 (Aug., 1986), 763-795
- H. THEIL (1957), Linear Aggregation in Input-Output Analysis; *Econometrica*, Vol. 25, No.1 (Jan., 1957), 111-122
- H. THEIL (1959), The Aggregation Implications of Identifiable Structural Macrorelations; *Econometrica*, Vol. 27, No.1 (Jan., 1959), 14-29
- YRJÖ VARTIA (2008a), Integration of Micro and Macro Explanations; *Discussion Paper No 239 (Nov., 2008) HECER*.
- YRJÖ VARTIA (2008b), On the Aggregation of Quadratic Micro Equations; *Discussion Paper No 248 (Nov., 2008) HECER*.
- YRJÖ VARTIA (2009), Whole and its Parts: Micro Foundations of Macro Behaviour; *Discussion Paper No 257 (Mar., 2009) HECER*.

P. DE WOLFF (1941), Income Elasticity of Demand, a Micro-Economic and a Macro-Economic Interpretation; *The Economic Journal*, Vol. 51, No.201 (Apr., 1941), 140-145

A First analysis then synthesis: $A \rightarrow S$

A.1 Analysis

Let the micro units or agents a_i define a set

$$H = \{a_1, \dots, a_n\}. \quad (28)$$

For every a_i we have a behavioral equation

$$y(i) = f_i(x(i)), \quad (29)$$

where we have output, behavior and inputs, $y(i)$, f_i and $x(i)$ respectively.²¹ Now let us define an Analysis operator A , which maps $(f, x) \in \Phi^n \times \mathbb{R}^{nk}$ to $y \in \mathbb{R}^n$, where $y = (y(1), \dots, y(n))$ is $(n \times 1)$ -vector, $f \in \Phi(\mathbb{R}^k \rightarrow \mathbb{R})^n$ and $x = (x(1), \dots, x(n)) \in \mathbb{R}^{nk}$. So we have for $A : \Phi^n \times \mathbb{R}^{nk} \rightarrow \mathbb{R}^n$

$$y = A(f, x). \quad (30)$$

In the component form this is

$$\begin{pmatrix} y(1) \\ \vdots \\ y(n) \end{pmatrix} = A \begin{pmatrix} f_1 & x(1) \\ \vdots & \vdots \\ f_n & x(n) \end{pmatrix} = \begin{pmatrix} f_1(x(1)) \\ \vdots \\ f_n(x(n)) \end{pmatrix}.$$

So the analysis operator is the one that attaches the behaviors and inputs of agents with the outputs. This is linear in its function arguments, i.e.

$$A(f + g, x) = A(f, x) + A(g, x) \quad (31)$$

$$A(\lambda f, x) = \lambda A(f, x). \quad (32)$$

We can write our behavior to be analysed as

$$A(f, x) = A(\bar{f}\mathbf{1} + \delta f, x) = A(\bar{f}\mathbf{1}, x) + A(\delta f, x), \quad (33)$$

²¹Note that $y(i) \in \mathbb{R}$, $f_i \in \Phi(\mathbb{R}^k \rightarrow \mathbb{R}) =$ such a function space and $x(i) \in \mathbb{R}^k$.

where the last equality follows from linearity of operator A . Let us write the above as

$$AB(x) = ACB(x) + AHB(x), \quad (34)$$

where on the left hand side we have the analysed behavior, which defines the whole micro system. The first term on the right hand side is the analysed common behavior, where all *behaviors* are restricted to common ones. The second term on the right hand side is the analysed heterogenous behavior, which tells us how the behaviors differ from the common behavior. From these we see how the micro system behaves (what is the typical behavior and what kind of extreme behaviors appear in the micro system).

A.2 Synthesis

Define the Synthesis operator $S : \Phi^n \times \mathbb{R}^{nk} \rightarrow \mathbb{R}$ to be:

$$S(f, x) = \frac{1}{n} \sum_{i=1}^n f_i(x(i)). \quad (35)$$

In component form this reads

$$\frac{1}{n} \sum_{i=1}^n f_i(x(i)) = S(f, x) = S \begin{pmatrix} f_1 & x(1) \\ \vdots & \vdots \\ f_n & x(n) \end{pmatrix}.$$

It is worth noticing the dependence between the analysis and the synthesis operator, which is

$$\begin{aligned} S(f, x) &= \frac{1}{n} \sum_{i=1}^n f_i(x(i)) = \frac{1}{n} \mathbf{1}' \begin{pmatrix} f_1(x(1)) \\ \vdots \\ f_n(x(n)) \end{pmatrix} \\ &= \frac{1}{n} \mathbf{1}' A \begin{pmatrix} f_1 & x(1) \\ \vdots & \vdots \\ f_n & x(n) \end{pmatrix} = \frac{1}{n} \mathbf{1}' A(f, x), \end{aligned}$$

where $\mathbf{1}$ is $(n \times 1)$ -vector and $A(f, x)$ is $(n \times 1)$ -vector. The above means that operating with the Synthesis operator gives the arithmetic mean of the components produced by the Analysis operator. The linearity of $A(f, x)$ in its f -arguments implies a similar linearity for $S(f, x)$.

Now we can write the macro behavior $MB(x)$ as follows:²²

$$\begin{aligned}
MB(x) &= S(f, x) = S(\bar{f}\mathbf{1}, x) + S(\delta f, x) \\
&= \underbrace{S(\bar{f}\mathbf{1}, \bar{x}\mathbf{1})}_{RB(\bar{x})} + \underbrace{(S(\bar{f}\mathbf{1}, x) - S(\bar{f}\mathbf{1}, \bar{x}\mathbf{1}))}_{NLE(x)} + \underbrace{S(\delta f, x)}_{HE(x)} \\
&= RB(\bar{x}) + NLE(x) + HE(x),
\end{aligned} \tag{36}$$

where the first term on the right hand side, $RB(x)$, is the representative behavior. The second term, $NLE(x)$, arises from the non-linearity of behavior functions and is called the non-linearity effect. This term is zero if there are no non-linearities in the behaviors. The third term, $HE(x)$, arises from the differences among agents and is called heterogeneity effect. This vanishes if δf 's are constants that sum to 0. In this case agents' behaviors differ from each other only by shifts (parallel family of behaviors). When all the inputs are identical, $x(i) = x(j) \forall i, j$, both $NLE(x)$ and $HE(x)$ vanish, but this is not a very realistic case.

Contents

1	Introduction	1
2	Framework	4
3	Aggregation of a micro system	6
3.1	Micro system	6
3.2	Aggregation	9
4	Properties of the aggregate	10
4.1	Level of the aggregate	10
4.2	Dynamics of the aggregate	12
5	Comparison with Blundell and Stoker	17
A	First analysis then synthesis: A→S	22
A.1	Analysis	22
A.2	Synthesis	23

²²See the similarity between the decomposition of the analysis part.