Applied Mathematical Sciences, Vol. 9, 2015, no. 134, 6669 - 6689
HIKARI Ltd, www.m-hikari.com
http://dx.doi.org/10.12988/ams.2015.59583

Statistical Approach to Aggregation of Production Functions

D.N. Shults

Federal State Budgetary Educational Institution of Higher Professional Education
"Perm State National Research University", Perm, Russia
25th October str., 38-92, Perm, 614007, Russia

M.I. Grebnev

Federal State Budgetary Educational Institution of Higher Professional Education
"Perm State National Research University", Perm, Russia
Marshala Tolbuhina Street, 12-12, Perm, 614030, Russia

Abstract

This article explores production function (PF) aggregation problem. It is shown that a consistent description of PF on micro- and macroeconomic levels is possible only in special case of linear PF. A statistical approach to aggregation is offered. It is shown how the proposed statistical method allows to derive new types of PF. It is proved that PF's are not invariant on the micro- and macrolevels. It is also shown that the proposed statistical approach allows to obtain more accurate estimates of macroeconomic PF than traditional Cobb-Douglas function.

Keywords: production function, consistent aggregation, statistical approach

Introduction

Suppose we have a two-level economy. At the micro level, it is comprised of n economic agents described by a micro production function which relates output of i-th firm y_i to input x_i

$$y_i = f_i(x_i). \tag{1}$$

Then vector function $f: R^n \rightarrow R^n$ defines vector $y = (y_1 \quad \cdots \quad y_n)$ of output of all production units at inputs $x = (x_1 \quad \cdots \quad x_n)$.

Further, we will make a simplifying assumption that there exist no relations between production units.

For the economy as a whole, a macro production function (MAPF) $F: R \rightarrow R$ is used that maps aggregate inputs *X* into aggregate outputs *Y*.

$$Y = F(X). \tag{2}$$

Note that we need to consider relations that also exist between variables *x* and *X*, *y* and *Y*.

$$X = \chi(x), \tag{3}$$

$$Y = \varphi(y), \tag{4}$$

where $\chi: R^n \rightarrow R$ and $\varphi: R^n \rightarrow R$ are aggregation operators. In case of outputs and inputs, summation operators can be used as aggregation operators.

Background

As L. Klein stated [[6], P.94], a solution can be found in two ways. First, we can derive a macro function *F* that satisfies a set of constraints (1)-(4). This means that we must derive macroeconomic regularities and construct the complete theory of macroeconomics based on the microeconomics and statistical aggregation rules. This approach seems obvious and we will consider it in depth below.

One of first attempts to derive an aggregate production function from micro production functions was made by F. Dresch [[2]]. If output of *i*-th microeconomic agent increases by dy_i, the total output increases by $dY = \sum_i p_i dy_i$, where p_i is product price of *i*-th firm.

Then from the solution of the differential equations

$$\frac{dY}{Y} = \frac{\sum_i p_i dy_i}{\sum_i p_i y_i} \tag{5}$$

in general economic equilibrium conditions we can write the aggregate production function as:

$$Y(t) = Y(\tau) exp\left[\int_{\tau}^{t} \left(\frac{\sum_i p_i dy_i}{\sum_i p_i y_i}\right)\right],$$

where t and τ are time points for the Divisia index. First of all, note that Klein's approach to deriving aggregates fails to satisfy the Divisia index properties that are generally accepted in statistics. That fact was indicated by Klein [[6], P.96] and it caused serious contradictions between macroeconomic indicators [[3], P.134].

Shou Shan Pu showed that Klein's assumption of profit maximization by firms and perfect competition was unjustified and redundant [[13], P.300]. Shou Shan Pu formulated his own criterion for the aggregation of production functions

– the existence of a structure (pattern) for distribution of aggregate values among economic agents [[13], P.302].

K. May showed that the scheme for deriving aggregates changes completely when the condition of perfect competition is removed [[10], P.297]. This means that the aggregation of production functions may not be abstract and must be based on a specific economic situation. In his response, Klein argued that production functions refer to equations that are independent of the microeconomic system [[7], P.303], and that they should be a purely technological relationship, and should not depend on decision-making process [[7], P.305]. K. May, in his following article, argued that the productive possibilities of an economy are dependent not only upon the productive possibilities of the individual firms (reflected in production functions) but on the manner in which these technological possibilities are utilized, as determined by the socio-economic framework [[11], P.63].

Wassily Leontief [[8]] formulated a necessary and sufficient condition for a twice-differentiable production function, whose arguments are all nonnegative, to be expressible as an aggregate – the marginal rates of substitution among variables in the aggregate are independent of the variables left out of it. For the given production function $Y = f(k_1, \ldots, k_n, L)$, the aggregate production function can be written as $Y = F(K, L)$, where $K = \varphi(k_1, \ldots, k_n)$, if and only if $\frac{\partial}{\partial L}\left(\frac{\partial f/\partial k_i}{\partial f/\partial k_j}\right) = 0,\ \forall\, i \neq j$. In other words, changes in labor, the noncapital input, do not affect the substitution possibilities between the capital inputs and they can be reduced to one capital input variable [[3], P.133].

Wassily Leontief's conclusions were summarized in Nataf's conditions [[3], P.133]: the aggregate production function exists, if and only if the aggregate production function is additively separable, i.e. if it can be written in the form:

$$f(k, l) = \phi(k) + \psi(l). \tag{6}$$

F. Fisher [[3]] observed that Nataf's conditions are extremely restrictive. Relying upon May's arguments and interpreting the production function as the maximum output achievable with the given inputs, he laid down his own existence conditions:

– an output aggregate will exist if and only if a given set of relative output prices induces all firms to produce all outputs in the same proportion;
– a capital aggregate exists if and only if firms differ by at most a capital-augmenting technical difference;
– a labor aggregate will exist if and only if a given set of relative wages induces all firms to employ different labor in the same proportions.

Statistical Method for Aggregation

We have proposed a statistical method for aggregation. Below we will outline its basic concepts.

Suppose that the economic system consists of *n* elements. Production functions of each *i*-th element has the same functional form and differs in parameter a_i:

$$f_i(x) = f(a_i, x). \tag{7}$$

Next, we assume that parameters a_i are independent (across system elements) and are identically distributed random variables. That is, we can consider a_i as a realization of a random variable *A*, which has distribution function $G(a)$ and density function $g(a)$.

Provided that random variables y_i have first and second moments of the outcome, the law of large numbers holds:

$$\frac{1}{n}\sum_{i=1}^{n} y_i \xrightarrow{p} M[y_1]. \tag{8}$$

Then total output *Y* can be written as:

$$Y \xrightarrow{p} nM[y_1] = n\int f(a,\cdot)g(a)da = n\tilde{F}(\cdot). \tag{9}$$

Here, we will not specifically scrutinize the argument variables on the left and right sides. Because we are faced with extensive factors ($X = \sum_i x_i$), on the left and right sides of the equation (9) different variables are used (X and x_i respectively). In our early studies, we paid no due attention to this question. As a result, there were significant errors in our estimations. Below we will scrutinize this problem.

For now, we will assume that input *X* distribution is relatively uniform across the economy. Correspondingly, as an estimation of the macro production function we can use the following equation (10)

$$Y \approx n\int f(a,x)g(a)da \approx n\int f\left(a,\frac{X}{n}\right)g(a)da = n\tilde{F}\left(\frac{X}{n}\right), \tag{10}$$

provided that "average" production function $\tilde{F}$ is homogeneous of degree 1 (constant returns to scale), i.e. if condition $\tilde{F}(\lambda x) = \lambda\tilde{F}(x)$ is satisfied, then $F(X) = \tilde{F}(X)$.

Let us consider a case of two-factor Cobb-Douglas production function:

$$Y = AL^{\alpha}K^{\beta}, \tag{11}$$

where *Y* is total production, *K* is capital input, *L* is labor input, *A* is scale parameter, α is output elasticity of labor, β is output elasticity of capital.

Suppose that the production function parameters are correlated with each other. In this case, we can get the macro production function using the following statistical method:

$$F(K,L) = n\iiint Ag_1(A)dA\left(\frac{L}{n}\right)^{\alpha} g_2(\alpha)d\alpha\left(\frac{K}{n}\right)^{\beta} g_3(\beta)d\beta\ , \tag{12}$$

where functions $g_1(A), g_2(\alpha), g_3(\beta)$ are density functions for TFP parameters, namely *A*, α, β.

Given that the first multiplier in the macro production function equation is an expectation of random variable A, we get:

$$F(K,L) = nM[A] \iint \left(\frac{L}{n}\right)^{\alpha} g_2(\alpha)d\alpha \left(\frac{K}{n}\right)^{\beta} g_3(\beta)d\beta \,. \qquad (13)$$

In case of normal distribution of TFP parameters, the macro production function equation can be written as follows:

$$F(K,L) = nM[A] \iint \left(\frac{L}{n}\right)^{\alpha} \frac{1}{\sqrt{2\pi}\sigma[\alpha]} e^{-\frac{(\alpha-M[\alpha])^2}{2\sigma^2[\alpha]}} d\alpha \left(\frac{K}{n}\right)^{\beta} \frac{1}{\sqrt{2\pi}\sigma[\beta]} e^{-\frac{(\beta-M[\beta])^2}{2\sigma^2[\beta]}} d\beta \,, \qquad (14)$$

where $M[\alpha]$ and $M[\beta]$ are expectations for parameters α and β, while $\sigma^2[\alpha]$ and $\sigma^2[\beta]$ are their variances.

After calculations and transformations we get:

$$F(K,L) = nM[A] \left(\frac{L}{n}\right)^{M[\alpha]+\frac{1}{2}\sigma^2[\alpha]\ln\left(\frac{L}{n}\right)} \left(\frac{K}{n}\right)^{M[\beta]+\frac{1}{2}\sigma^2[\beta]\ln\left(\frac{K}{n}\right)}. \qquad (15)$$

or in logarithmic form:

$$\ln Y = \ln(nM[A]) + M[\alpha]\ln\left(\frac{L}{n}\right) + \frac{1}{2}\sigma^2[\alpha]\ln^2\left(\frac{L}{n}\right) + \\ + M[\beta]\ln\left(\frac{K}{n}\right) + \frac{1}{2}\sigma^2[\beta]\ln^2\left(\frac{K}{n}\right). \qquad (16)$$

In this way, we have constructed the macro production function based on assumptions on the functional form of TFP and normal distribution of parameters α and β.

Other Production Functions

Using the statistical method for aggregation one can construct various macro production functions for the Cobb-Douglas production function,

$$f(l,k) = Al^{\alpha}k^{\beta}; \qquad (17)$$

CES and

$$f(l,k) = A(\delta l^{-p} + (1-\delta)k^{-p})^{-m/p}; \qquad (18)$$

Leontief

$$f(l,k) = Amin\{al, bk\}. \qquad (19)$$

So, the general aggregate production function for the Cobb-Douglas production function can be written as equation (13). For various parameter distribution functions the macro production function can be written as follows:

– in case of uniform distribution ($\alpha \sim R[a,b], \beta \sim R[c,d]$):

$$F(L,K) = nM[A]\left(\frac{\left(\frac{L}{n}\right)^a - \left(\frac{L}{n}\right)^b}{ln\left(\frac{L}{n}\right)(a-b)}\right)\left(\frac{\left(\frac{K}{n}\right)^c - \left(\frac{K}{n}\right)^d}{ln\left(\frac{K}{n}\right)(c-d)}\right). \qquad (20)$$

where a and b are minimum and maximum values of parameter α; c and d are minimum and maximum values of parameter β.

– in case of triangular distribution ($\alpha \sim Tr[a,b,m_\alpha], \beta \sim Tr[c,d,m_\beta]$):

$$F(L,K) = nM[A]\left(\frac{2(m_\alpha - a)\left(\frac{L}{n}\right)^b + 2(b-m_\alpha)\left(\frac{L}{n}\right)^a - 2(b-a)\left(\frac{L}{n}\right)^{m_\alpha}}{ln^2\left(\frac{L}{n}\right)(b-a)(m_\alpha - a)(b-m_\alpha)}\right) * \left(\frac{2(m_\beta - c)\left(\frac{K}{n}\right)^d + 2(d-m_\beta)\left(\frac{K}{n}\right)^c - 2(d-c)\left(\frac{K}{n}\right)^{m_\beta}}{ln^2\left(\frac{K}{n}\right)(d-c)(m_\beta - c)(d-m_\beta)}\right) \qquad (21)$$

– for normal distribution $(\alpha \sim N(M[\alpha], \sigma^2[\alpha]), \beta \sim N(M[\beta], \sigma^2[\beta]))$ – equation (15);

– for lognormal distribution $(ln(\alpha) \sim N(\mu_\alpha, \sigma_\alpha^2), ln(\beta) \sim N(\mu_\alpha, \sigma_\beta^2))$:

$$F(L,K) = nM[A]\int\left(\frac{L}{n}\right)^\alpha \frac{1}{\alpha\sqrt{2\pi}\sigma[\alpha]} e^{-\frac{(ln\alpha - \mu_\alpha)^2}{2\sigma_\alpha^2}} d\alpha * * \int\left(\frac{K}{n}\right)^\beta \frac{1}{\beta\sqrt{2\pi}\sigma[\beta]} e^{-\frac{(ln\beta - \mu_\beta)^2}{2\sigma_\beta^2}} d\beta. \qquad (22)$$

The latter equation cannot be calculated in analytic form.

For the Leontief production function, the general macro production function can be written as follows:

$$F(L,K) = n\iiint Amin\left\{a\left(\frac{L}{n}\right), b\left(\frac{K}{n}\right)\right\} g_1(A)g_2(a)g_3(b)dAdadb, \qquad (23)$$

where $g_1(A), g_2(a), g_3(b)$ are density functions for TFP parameters respectively A, a, b.

For all distribution laws in question the macro production function takes the following unified form:

$$F(L,K) = nM[A]min\left\{M[a]\left(\frac{L}{n}\right), M[b]\left(\frac{K}{n}\right)\right\}. \qquad (24)$$

Application of Statistical Method for Aggregation

For the U.S. economy, sectoral Cobb–Douglas production functions were estimated over the period 1980-1998:

$$\ln Y = \ln A + \alpha \ln L + \beta \ln K, \qquad (25)$$

where all variables are in billion dollars.

Normal distribution was takes as a distribution law: $M[\alpha]$=0.726, $M[\beta]$=0.235, $\sigma^2[\alpha]$=0.175, $\sigma^2[\beta]$=0.050; a $M[A]$ = 4.056.

Subsequently, parameters of the obtained macro production function were estimated. Obviously, one can suggest the following methods for estimating parameters of the obtained function:

– substitution of estimated mean and estimated variance of TFP parameters;

– estimation of logarithmic equation parameters (16):

$$\ln Y = a + a_1 \ln\frac{L}{n} + a_2 \ln^2\frac{L}{n} + b_1 \ln\frac{K}{n} + b_2 \ln^2\frac{K}{n}. \quad (26)$$

Let us reduce estimated parameters of equation (26) (Table 1):

Table 1. MAPF parameter estimates

	Estimate
a	4.345
a_1	0.726
a_2	0.088
b_1	0.235
b_2	0.025

Technical Progress Accounting

Recent researches show that inputs of labor and capital explain only 4 to 50 percent of economic growth across countries [0, P.50]. So, changes in total factor productivity account for the remaining unexplained part of growth.

One way to consider technical progress in production functions is to drop an assumption that parameter A is constant in time. If written as an exponential function $A(t) = Ae^{\lambda t}$, technical progress can be exogenously included into the production function:

$$Y = Ae^{\lambda t} L^{\alpha} K^{\beta}, \quad (27)$$

where $\lambda > 0$ is a parameter describing growth rate of total factor productivity.

Such technical progress is called Hicks-neutral technical progress. Harrod neutral technical progress (labor-saving) and Solow-neutral technical progress (capital-saving) are known as well [[12], P.9].

So, the equation for calculating the macro production function can be written as follows:

$$F(K, L, t) = nM[A] \int e^{\lambda t} g_4(\lambda) d\lambda \int \left(\frac{L}{n}\right)^{\alpha} g_2(\alpha) d\alpha \int \left(\frac{K}{n}\right)^{\beta} g_3(\beta) d\beta, \quad (28)$$

where $g_4(\lambda)$ is density function for parameter λ.

This equation differs from the one that was used previously (13) by factor $\int e^{\lambda t} g_4(\lambda) d\lambda$. In this way, factor of time and factor of technical progress are included into the macro production function multiplicatively:

$$F(K, L, t) = F(K, L) \int e^{\lambda t} g_4(\lambda) d\lambda, \quad (29)$$

where $F(K, L)$ is equation for macro production function (13).

Let us try to define density function $g_4(\lambda)$ for parameter λ. To do so, let us estimate macro production function parameters across sectors over the period 1976-1998 (Table 2).

Table 2. MAPF parameter estimates by sectors taking into account technical progress

Sectors	**ln*A***	$\boldsymbol{\alpha}$	$\boldsymbol{\beta}$	λ
Food & Kindred Prod.	2.108	0.806	0.014	0.008
Textile Mills Prod.	0.719	1.002	0.039	0.006
Apparel &Related Prod.	1.751	0.387	0.127	0.022
Paper & Allied Prod.	1.207	0.774	0.211	0.004
Printing & Publishing	1.205	0.840	0.174	-0.004
Chem. & Allied Prod.	1.736	0.588	0.274	0.003
Petroleum Refining	1.855	0.865	0.642	-0.019
Rubber & Plastic Prod.	1.184	0.800	0.088	0.011
Lumber & Wood Prod.	1.005	0.934	0.154	-0.003
Furniture & Fixtures	1.210	0.705	0.122	0.012
Stone, Clay & Glass	0.839	1.112	0.092	-0.009
Primary Metal Ind.	1.401	0.757	0.117	0.008
Fabricated Metal Prod.	1.293	0.827	0.235	-0.005
Ind. Machinery, Comp.Eq.	0.961	0.804	0.171	0.007
Electric & Electr. Eq.	1.235	0.830	0.239	-0.008
Transportation Equip.	1.850	0.502	0.047	0.032
Instruments	1.106	0.839	0.020	0.007
Misc. Manufacturing	1.271	0.822	0.068	0.004
Elect. & Gas Utilities	2.061	0.794	-0.010	0.012

A histogram for parameter λ shows normal distribution $N(0{,}00471; 0{,}0001322)$. For reference, let us provide estimates by Barro and his coauthor [0, P.439] – the rate of growth of total labor productivity is estimated at the level of 0.0135 for the period 1947-1973 and 0.0076 for the period 1965-1995.

For the remaining production function parameters we used the following characteristics: $\alpha \sim N(0{,}789; 0{,}027)$, $\beta \sim N(0{,}149; 0{,}021)$, $M[A] = 3{,}928$.

Then the macro production function with technical progress can be written as follows:

$$F(K, L, t) = F(K, L) \int e^{\lambda t} \frac{1}{\sqrt{2\pi}\sigma[\lambda]} e^{-\frac{(\lambda - M[\lambda])^2}{2\sigma^2[\lambda]}} d\lambda. \tag{30}$$

Eventually, we get the following macro production function with technical progress:

$$F(K, L, t) = nM[A] e^{M[\lambda]t + \frac{1}{2}\sigma^2[\lambda]t^2} \left(\frac{L}{n}\right)^{M[\alpha] + \frac{1}{2}\sigma^2[\alpha]\ln\left(\frac{L}{n}\right)} \left(\frac{K}{n}\right)^{M[\beta] + \frac{1}{2}\sigma^2[\beta]\ln\left(\frac{K}{n}\right)}. \tag{31}$$

or in a linearized form:

$$\ln Y = a + a_1 \ln\frac{L}{n} + a_2 \ln^2\frac{L}{n} + b_1 \ln\frac{K}{n} + b_2 \ln^2\frac{K}{n} + c_1 t + c_2 t^2. \quad (32)$$

Let us compare estimated parameters of equation (32) with those obtained earlier from Table 1:

Table 3. Estimated parameters of MAPF with or without technical progress

	Estimates with technical progress	Estimates without technical progress
a	74.63	4.345
a_1	0.789	0.726
a_2	0.013	0.088
b_1	0.149	0.235
b_2	0.010	0.025
c_1	0.0047	
c_2	0.0001	

By analogy with the U.S. model described above, M. Grebnev constructed a macro production function for the Russian economy. In this case, capital is replaced with fixed investments:

$$Y = AL^{\alpha} I^{\beta}, \quad (33)$$

where Y is gross value added by types of economic activity in billion 2004 rubles; I is fixed investments in billion 2004 rubles; L is average annual number of employed in thousand people.

We got estimates for the Cobb-Douglas production function for 15 sectors. Under the assumption of their normal distribution the aggregate production function for Russia is written as follows:

$$F(I,L) = 1258{,}91 \left(\frac{L}{15}\right)^{0{,}441+0{,}135ln\left(\frac{L}{15}\right)} \left(\frac{I}{15}\right)^{0{,}216+0{,}019ln\left(\frac{I}{15}\right)}. \quad (34)$$

Note that sectoral statistics for Russia is available for 8 periods only (from 2004 through 2011). Thus, one can get most reliable estimates based on regional statistics that is available for 15 periods (from 1998 through 2012). Moreover, when the regional dimension is used, the panel expands from 15 sectors to 83 regions.

Then, in equation (33) GRP in million rubles is taken as Y; while payroll in million rubles is taken as L. All variables are reduced to 1998 prices.

Estimates were produced by the Prognoz Platform 7.2 software solution (www.prognoz.ru) for the period from 1998 to 2012 based on data from dataportal.prognoz.ru.

In contrast to estimates by sectors, estimates by regions have more clear-cut distribution: logonormal distribution for parameter α and normal distribution for parameter β. Thus, to calculate the aggregate production function, one needs to calculate the following equation:

$$F(I,L) = nM[A]\left(\frac{I}{n}\right)^{M[\beta]+\frac{1}{2}\sigma^2[\beta]\ln\left(\frac{I}{n}\right)} \int_0^{\infty} \left(\frac{L}{n}\right)^{\alpha} \frac{1}{\alpha\sqrt{2\pi\sigma_\alpha^2}} e^{-\frac{(\ln(\alpha)-\mu_\alpha)^2}{2\sigma_\alpha^2}} d\alpha. \quad (35)$$

In the integral in equation (35), all parameters except $M[A]$ are calculated based on parameters values of regional production functions. Parameter $M[A]$ as a certain balancing scale factor was estimated using the least square method for the macro production function equation. A plus of this estimation approach is that we avoid the problem of factor multicollinearity.

As a result, we got the following estimates (Table 4):

Table 4. MAPF parameter estimates for Russia

Parameter	**Value**
n	83
$M[A]$	211.3
μ_α	-1.23
σ_α	0.46
$M[\beta]$	0.17
$\sigma[\beta]$	0.10

The analysis of the obtained production function shows that it is a convex function, marginal productivity of labor and investment are positive and decreasing. Correspondingly, the second variables are negative. That is the constructed macro production function satisfies classical axioms.

Structure Analysis

Suppose that all elements of the system are identical and are characterized by a linear production function $y_i = ax_i$, where y_i is total production of i-th element, x_i is total input by i-th element, while a is return to input.

In this case, the distribution of input across elements of the system does not affect the cumulative result:

$$Y = \sum_i y_i = \sum_i ax_i = \sum_i aw_iX = aX\underbrace{\sum_i w_i}_{=1} = aX, \quad (36)$$

where Y is aggregate production across the economy, $X = \sum_i x_i$ is aggregate input, $w_i = x_i/X$ is share of input available to i-th element.

First, we can draw a certain parallel with the Coase theorem that states that under certain conditions such as zero transaction costs, the allocation of property does not affect the system efficiency. Second, in this case, as we have mentioned earlier, there exist the macro production function in an explicit form.

However, this conclusion is wrong, if we save the precondition of absolutely identical elements of the system, and remove the precondition of linearity of their production functions. Let us take the simplest production function:

$$y_i = \sqrt{x_i}. \tag{37}$$

Then

$$Y = \sum_i y_i = \sum_i \sqrt{x_i} = \sum_i \sqrt{w_i X} = \sqrt{X} \sum_i \sqrt{w_i}. \tag{38}$$

First, this means that one cannot construct a single-value macro production function even in case of identical but nonlinear production functions. More exactly, in this case the macro production function is a function of not only aggregate inputs X, but also of internal structure w_i (the way inputs are allocated).

Moreover, since $w_i < 1$, then $\sum_i \sqrt{w_i} > 1$. This means that if we had applied the representative agent approach and used $Y = \sqrt{X}$as a macro production function, the macro production function would have been underestimated by several times, since:

$$Y = \sqrt{X} \sum_i \sqrt{w_i} > \sqrt{X}. \tag{39}$$

For example, if the number of elements in the system is *n* and inputs are allocated uniformly $w_i = \frac{1}{n}$, then the actual output would exceed its estimate by $n\sqrt{\frac{1}{n}} = \sqrt{n}$ times. In real economies, distribution of inputs is less uniform, while the number of elements is measured in millions. One can envision the magnitude of errors caused by the use of the representative agent approach in macroeconomic modeling and forecasting.

To illustrate the above, we have run 1,000 simulations. During each simulation, by uniform law we generated magnitude X and weighting factors w_i for a system comprised of 20 elements (*n*=20). Then, we calculated aggregate output $Y = \sum_i \sqrt{w_i X}$ and compared it versus output values obtained using the representative agent approach $\mathrm{Y} = \sqrt{\mathrm{X}}$. The results are illustrated in the chart below (Figure 1).

We can draw the following conclusions. On the one hand, as we envisaged, the results obtained using the representative agent approach turned to be underestimated by several times. On the other hand, the scattering of points $(Y; X)$ resembles a nonlinear production function and can be approximated using regression $Y = A\sqrt{X} + \varepsilon$. So, when building econometric models we speak that random errors ε include the impact of unaccounted factors, among those factors we also imply the microeconomic structure of distribution of inputs.

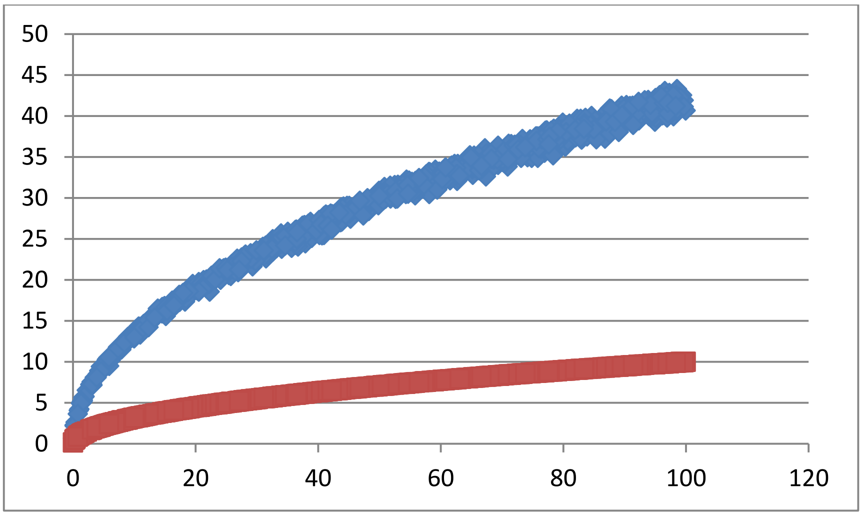

Figure 1. Total production based on precise aggregation and representative agent approach in case of nonlinear production function and identical elements

In the context of our experiment, we have estimated the regression and obtained the following results:

$$Y = 4.221\sqrt{X} + e. \tag{40}$$

The obtained estimates have high statistical significance level (99.99%).

However, assume that output elasticity of factors of production differs from 0.5. That is, the production function is written as the Cobb-Douglas production function:

$$y_i = x_i^{\alpha} = (x_i)^{\alpha}. \tag{41}$$

Then the aggregate production function can be written as:

$$Y = X^{\alpha} \sum_i w_i^{\alpha}. \tag{42}$$

If $\alpha \approx 1$, then the production function is close to linear function. For example, we have made the same calculations for $\alpha = 0{,}9$. As you can see (

Figure 2), the dependence between Y and X is close to linear dependence. The linear regression estimate is as follows:

$$Y = 1.325X^{0,9} + e, \tag{43}$$

It seems that even linear approximation for the macro production function can give good results.

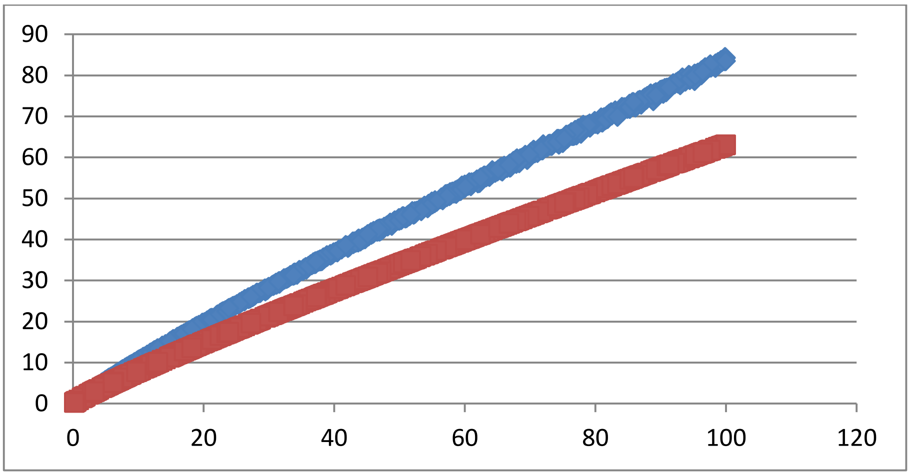

Figure 2. Total production based on precise aggregation and representative agent approach in case of $\boldsymbol{\alpha = 0.9}$.

However, if $\alpha \approx 0$, then errors can be much larger. For example, we ran the same simulations, but took $\alpha = 0.1$. As you can see (Figure 3), we have significant errors, since in this case $\sum_i w_i^\alpha$ differs from 1 by multiple times.

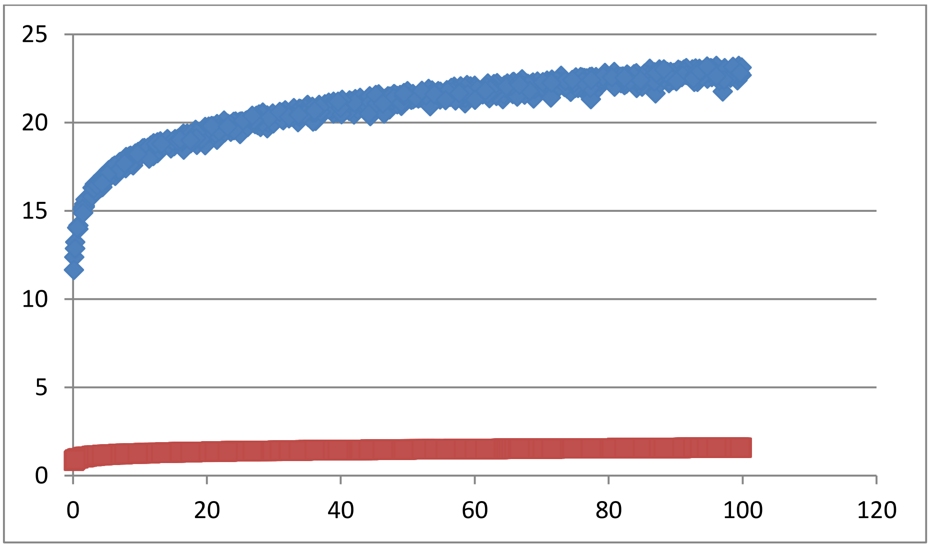

Figure 3. Total production based on precise aggregation and representative agent approach in case of $\boldsymbol{\alpha = 0.1}$.

The regression estimate is as follows:

$$Y = 14.450X^{0,1} + e, \tag{44}$$

Finally, let us consider a case of heterogeneous elements of the system:

$$Y = \sum_i (w_i X)^{\alpha_i}. \tag{45}$$

In this case, we have less grounds to approximate the macro production function in the form $Y = kX^{\alpha}$. It is practically impossible to treat this case analytically. Therefore, let us conduct a computational experiment. At first, suppose that the macro production function parameters are distributed uniformly $\alpha_i \sim U(0;1)$.

In this case, the variance should obviously increase, since scattering due to nonlinearity is added by scattering due to heterogeneity. In figure below (Figure 4), you can see scatter charts $(Y;X)$ during various experiments. For each experiment, we generated a new set of random variables α_i and ran 1,000 simulations on it.

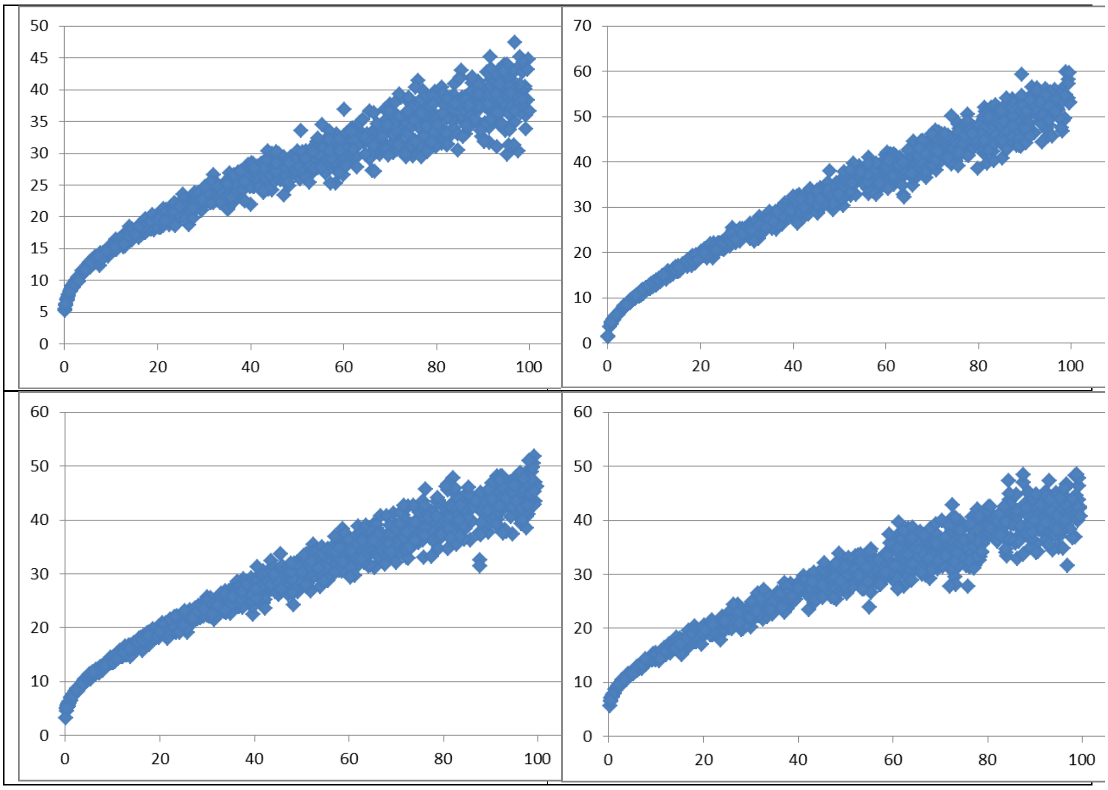

Figure 4. Different dependencies between Y and X under heterogeneous elements of the system and various generations. $\boldsymbol{\alpha_i \sim U(0;1)}$

Chart below (Figure 4) shows that in any case there exist is a nonlinear dependence between Y and X, while difference are explained by random factors α_i. If we try, as we did before, to measure the dependence between Y and X, we will have to estimate regression in the form $Y = AX^a + \varepsilon$ or even $Y = \mathrm{c} + AX^b + \varepsilon$. In the first case, parameters can be estimated using ordinary least squares based on linear transformation $\ln Y = \ln A + a \ln X + \varepsilon$. In the latter case, we need to use nonlinear least squares.

For the last experiment out of those we conducted, we got the following estimates:

$$Y = 6.01X^{0,41} + e, \tag{46}$$

and

$$Y = 6.564 + 1.820X^{0,647} + e, \tag{47}$$

Let us try to define a macro production function based on the statistical method.

Let us use formula (10) with implicit uniform distribution of weighting factors $w = \frac{1}{n}$.

Then the macro production function can be written as follows:

$$Y = n \int_0^1 \left(\frac{X}{n}\right)^a da = \frac{X - n}{\ln X - \ln n}. \tag{48}$$

The chart (Figure 5) shows aggregate outputs *Y* generated as a result of experiment and approximated by means of the statistical method using formula (48). We will not scrutinize the properties of obtained macro production function. Suffice it to point out that it has a neoclassical form that is convex in both arguments (Figure 6) and provides an “upper estimate” for our production set.

Let us remind that there exist various approaches to the production function interpretation. Regression models (46) and (47) represent an average estimate of relation between inputs *X* and outputs *Y*, since regression is, in substance, a conditional expectation. The macro production function estimated by means of the statistical method using formula (48) to a greater extent refer to majorant treatment of the production function.

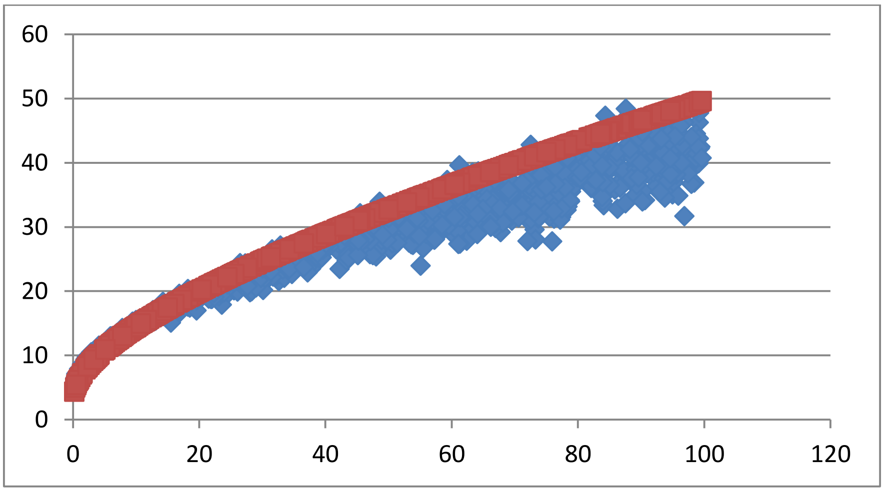

Figure 5. Generated outputs *Y* and their approximations by means of the statistical method for aggregation

Moreover, some researchers argue that structural and institutional factors must be taken into account in the aggregation of the production function. Here, it may be noted that in the context of the normative economics uniform distribution of inputs yields greater outputs (the macro production function places an upper limit on our production set). In other words, the uniform distribution of inputs in the economy in the most efficient.

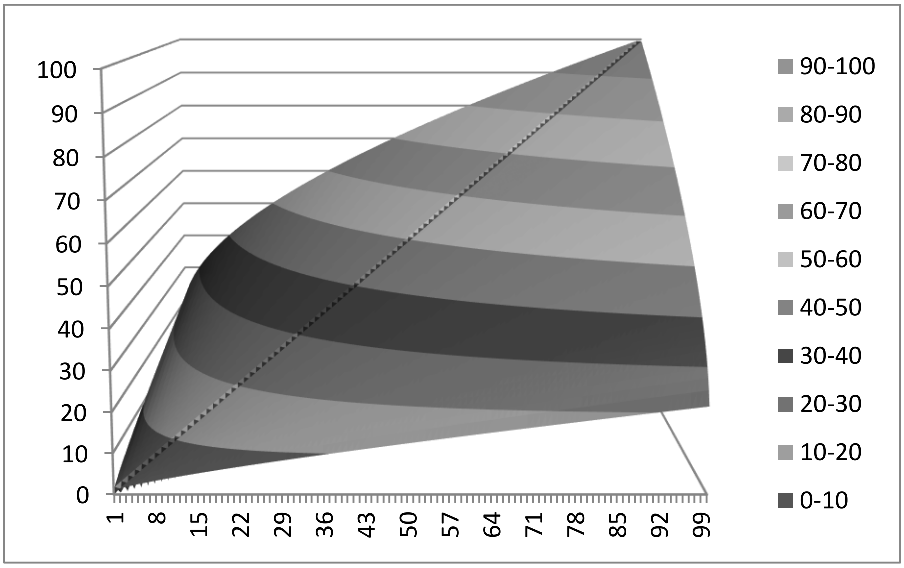

Figure 6. MAPF using formula (**48**) – dependence between *X* and *n*

From statistical estimation standpoint, formula (48) is notable for the fact that it contains no parameters. This is the only nonparametric production function that we know of.

Let us compare goodness-of-fit using R^2. Note that in some regression equations there are no constants. In such context, for the purpose of comparing model performances it is more correctly to use the following formula:

$$R^2 = \left(cor\left(Y, \tilde{Y}\right)\right)^2, \tag{49}$$

where $\tilde{Y}$ are values approximated by regression equations.

As a result, we get the following estimates of R^2.

(46)	(47)	(48)
93.47%	94.54%	94.55%

Higher R^2 in equation (47) versus (46) can be explained by a greater number of degrees of freedom (estimated parameters). In this context, it would be more correct to compare model performances based on adjusted coefficients of determination. Conspicuous is the fact that the macro production function obtained based on the statistical method (48) and having no estimated parameters whatsoever, has higher R^2.

Nonuniform Weights

In the course of finding equation (10), when we needed to get from micro variables of factor inputs x to macro inputs X, we had to assume that distribution of inputs across sectors or regions was uniform. That is we relied on the assumption that $x = \frac{X}{n}$.

In practice, this strong restriction led to significant errors in estimates of macro production function parameters. For this reason, equation (10) could be used only to construct a functional form of the macro production function, while its parameters needed to be estimated based on macroeconomic statistics.

To take into account nonuniformity of weight factors, let us treat w as an independent random variable with density function $g(w)$. Then:

$$Y(X) \approx \iint f(a, wX) g_a(a) g_w(w) da dw, \tag{50}$$

where a are production function parameters with density function $g(a)$.

This formula enables us to take into account not only nonuniformity of sectors, but it also provides a more correct transition from arguments of the micro production function to arguments of the macro production function.

The above experiments handled large amounts of economic agents. M. Grebnev […] conducted computational experiments on a system comprised of 10,000 elements ($n = 10000$), functioning of each of which is described by the Cobb-Douglas production function (similarly to the experiments conducted above):

$$y_i = x_i^{\alpha_i}, i = 1 \dots n, \tag{51}$$

where y_i is output of i-th firm, x_i is amount of factor of production used by -th firm, α_i is output elasticity of factor of production x of i-th firm.

Like in previous experiments, we are trying to test whether the statistical method for aggregation (50) is capable to produce more accurate descriptions for the system as a whole as compared to traditional econometric models. In addition to macro data, the first approach employs data about micro structure, and that is why we expect it to be more accurate.

Moreover, traditional methods that are based on the representative agent approach simply transfers functional form from micro level to macro level, while the statistical method for aggregation assumes that functional forms of micro and macro relationships may differ. So, the statistical method for aggregation is not restricted by functional form of the micro level and has greater number of degrees of freedom. This is the second reason why we expect higher accuracy from it.

To compare two approaches, we will conduct 4 experiments (each comprised of 1,000 simulations), in which we will analyze the impact of homogeneity for α and for w. In other words, we will consider various combinations:

	w homogeneous	w nonhomogeneous
α homogeneous	$\sigma[\alpha] = 0{,}01$ $\sigma[w] = 0{,}00001$	$\sigma[\alpha] = 0{,}01$ $\sigma[w] = 0{,}00003$
α nonhomogeneous	$\sigma[\alpha] = 0{,}1$ $\sigma[w] = 0{,}00001$	$\sigma[\alpha] = 0{,}1$ $\sigma[w] = 0{,}00003$

The remaining parameters will be the same in the experiments. Each experiment will include the following steps:

1. Generate normally distributed parameters α and w with parameters $M[\alpha] = 0.5$ and $M[w] = 0.0001$, as well as variances from the table;

2. Generate input of factor of production X to the economy by normal law with parameters $M[X] = 300000$ and $\sigma[X] = 50000$;

3. Calculate output y_i (51) and for the economy as a whole $Y = \sum_{i=1}^{n} y_i$;

For calculations using the aggregate production function (50), we need to define the domain of parameters: $\min_i\{w_i\} < w < \max_i\{w_i\}$, $\min_i\{\alpha_i\} < \alpha < \max_i\{\alpha_i\}$.

In our case, function (50) has no explicit form, therefore, to compare accuracy of the two approaches we measured Akaike informative criterion (AIC) for regressions between Y/n and $\tilde{Y}/n$. Here, $\tilde{Y}$ are estimated values of aggregate output obtained based on from equations (50) and modeled values obtained from the Cobb-Douglas model.

The graphs show (Figure 7) that the quality of estimations based on the statistical method for aggregation (50) in almost all cases was higher as compared to traditional macroeconomic Cobb-Douglas model despite the fact that micro data were generated by the Cobb-Douglas model (51).

The computational experiment proved that the proposed statistical method for aggregation of production functions has greater explanatory power versus traditional mathematical economic models for macro production functions.

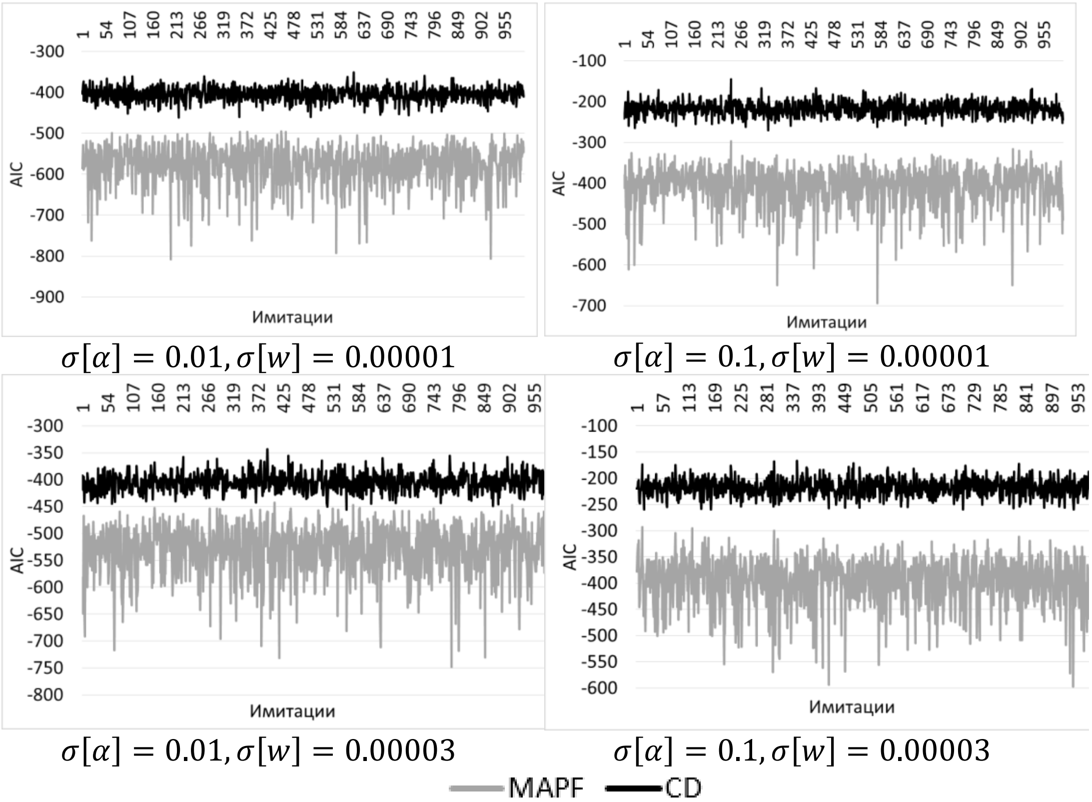

Figure 7. Akaike informative criterion

Of particular note is that everywhere above we assumed that production function parameters are independent. However, some economists prove that there exist positive economies of scale, and thus a positive relationship between a and w. Others believe that larger systems are characterized by X-inefficiency, correspondingly there must exist negative relationship between a and w.

In this paper, we will not fully cover this question and will leave it for future researches.

Main conclusions:

1. We have investigated the problem of aggregation in production functions, proposed the statistical method for aggregation of production functions, and constructed aggregate production functions for various sectoral production functions and various statistical distribution functions;
2. We have explored properties of aggregate production functions and found that in the general case production functions are not invariant across various levels of the economic hierarchy;
3. The difference between micro production function and macro production function comes from nonhomogenity of elements in the economy and the difference increases as the degree of nonhomogenity grows;

4. We have derived new aggregate production functions for the economies of the U.S.A. and Russia.

Acknowledgements. The article is implemented under complex project on high-tech production creation N 02.G25.31.0039 (RF Government Resolution No. 218 of April 9, 2010 “On Measures of State Support for the Development of Cooperation of Russian higher Education Institutions and Organizations Implementing Complex Projects for High-Tech Production” with financial support from the Ministry of Education and Science of the Russian Federation)"

References

[1] R. J. Barro, X. Sala-i-Martin, *Economic growth*, 2nd edition, The MIT Press, 2004.

[2] F.W. Dresch, Index numbers and the general economic equilibrium, *Bulletin of the American Mathematical Society*, **44** (1938), 134-141. http://dx.doi.org/10.1090/s0002-9904-1938-06708-0

[3] J. Felipe, F.M. Fisher, Aggregation in Production Functions: What Applied Economists should Know, *Metroeconomica*, **54** (2003), no. 3, 208-262. http://dx.doi.org/10.1111/1467-999x.00166

[4] F.M. Fisher, *Aggregation: Aggregate Production Functions and Related Topics: Collected Papers of Franklin M. Fisher*, The MIT Press, 1991, 280.

[5] C.I. Jones, The facts of economic growth. http://dx.doi.org/10.3386/w21142

[6] Lawrence R. Klein, Macroeconomics and the theory of rational behavior, *Econometrica*, **14** (1946), no. 2, 93-108. http://dx.doi.org/10.2307/1905362

[7] Lawrence R. Klein, Remarks on the theory of aggregation, *Econometrica*, **14** (1946), no. 4, 303-312. http://dx.doi.org/10.2307/1906912

[8] W.W. Leontief, Introduction to a theory of the internal structure of functional relationships, *Econometrica*, **15** (1947), no. 4, 361-373.

[9] J. Lintunen, O. Ropponen, Y. Vartia, Micro meets Macro via Aggregation, HECER Discussion Paper № 259, 2009. http://dx.doi.org/10.2139/ssrn.1370998

[10] K. May, The aggregation problem for a one-industry model, *Econometrica*, **14** (1946), no. 4, 285-298. http://dx.doi.org/10.2307/1906910

[11] K. May, Technological change and aggregation, *Econometrica*, **15** (1947), no. 1, 51-63. http://dx.doi.org/10.2307/1905816

[12] D. Romer, *Advanced Macroeconomics*, Second edition, McGraw-Hill, Boston, 2001, 631.

[13] Shou Shan Pu, A Note on Macroeconomics, *Econometrica*, **14** (1946), no. 4, 299-302. http://dx.doi.org/10.2307/1906911

Received: September 25, 2015; Published: November 9, 2015