Aggregate Productivity and Productivity of the Aggregate: Connecting the Bottom-Up and Top-Down Approaches

Bert M. Balk (Rotterdam School of Management, Erasmus University)

Paper prepared for the 34th IARIW General Conference

Dresden, Germany, August 21-27, 2016

Session 8C: Globalization, Growth, and Jobs III

Time: Friday, August 26, 2016 [Afternoon]

Aggregate Productivity and Productivity of the Aggregate: Connecting the Bottom-Up and Top-Down Approaches

Bert M. Balk*
Rotterdam School of Management
Erasmus University
Rotterdam
E-mail bbalk@rsm.nl

Draft, September 9, 2016

Abstract

Productivity analysis is carried out at various levels of aggregation. In microdata studies the emphasis is on individual firms (or plants), whereas in sectoral studies it is on (groupings of) industries. Now an industry is an ensemble of individual firms (decision making units) that may or may not interact with each other. In National Accounts terms this is symbolized by the fact that industry (aggregate) nominal value added is the simple sum of firm-specific nominal value added. From this viewpoint it is natural to expect there to be a relation between industry productivity and the firm-specific productivities. Yet,

*This paper draws from an extended version available at SSRN: http://ssrn.com/abstract=2585452. Presentations took place at the North American Productivity Workshop IX, 15-18 June 2016, Québec City, and the 34th General Conference of the International Association for Research in Income and Wealth, 21-27 August 2016, Dresden. Susanto Basu informed me that he had once written a paper with a virtually identical title. On inspection this turned out to be an embryonic version of Basu and Fernald (2002). Though related, my paper has a different focus.

microdata researchers do not care too much about the interpretation of the weighted means of firm-specific productivities they employ in their analyses. In this paper the consequences of this attitude are explored, based on a review of the literature.

However, a structurally similar phenomenon happens in sectoral studies, where the productivity change of industries is compared to each other and to the productivity change of some next-higher aggregate, which is usually the (measurable part of the) economy. Though there must be a relation between sectoral and economy-level measures, in most publications by statistical agencies and academic researchers this aspect is more or less neglected.

The point of departure of this paper is that aggregate productivity should be interpreted as productivity of the aggregate. It is shown that this implies restrictive relations between the productivity measure, the set of weights, and the type of mean employed.

Keywords: Producer; productivity; aggregation; bottom-up approach; top-down approach; index number theory.

JEL code: C43, D24, O47.

1 Introduction

In the first article of this series, Balk (2010), I considered productivity measurement for a single, consolidated production unit. In terms of levels, productivity is defined as real output divided by real input. Real output or input means nominal output or input deflated by some output- or input-specific price index, respectively. For the production unit considered, productivity change (through time) can then be measured as a difference or a ratio of productivities. In the latter case it appears that productivity change can also be defined directly as output quantity index divided by input quantity index.

The choice of the output and input concepts appears to be critical. Three main models can be distinguished: KLEMS-Y, KL-VA, and K-CF. Taking the composition of capital input cost into account, as set out in the companion paper Balk (2011), two more models can be added, namely KL-NVA and K-NCF. Assuming profit (defined as revenue minus total cost) to be equal to zero, or, what amounts to the same, replacing an exogenous interest rate by an endogenous rate, multiplies the number of models by two. And the introduction of a capital utilization rate further complicates the picture. Thus, there is a lot of choice here, with not unimportant empirical consequences, as illustrated by Vancauteren, Veldhuizen and Balk (2012).

Production units exist at various levels of aggregation. We see plants, enterprises, industries, countries, to name just some types of production units materializing in analyses of productivity change. Usually such units appear, more or less naturally, arranged into higher level aggregates. For instance, a number of plants belonging to the same enterprise; a certain type of enterprises defining an industry; a number of industries defining the 'measurable' part of a national economy; national economies making up the world economy. It is not difficult to perceive several sorts of hierarchy here.

As in any of these situations the structure is the same – there is an ensemble of production units, and the ensemble itself may or may not be considered as a higher level production unit –, it is interesting to study the relation between aggregate productivity (change) and productivity (change) of the aggregate.

There are basically two approaches here. Balk (2016) reviews and discusses the so-called *bottom-up* approach, the approach that takes an ensemble of individual production units as the fundamental frame of reference. The *top-down* approach is the subject of three other papers, namely Balk (2014) plus Dumagan and Balk (2016) on labour productivity, and Balk (2015) on total factor productivity.

The present paper investigates the connection between the two approaches, bottom-up and top-down. Characteristic of the approach taken in this paper is that aggregate productivity (change) should be interpreted as productivity (change) of the aggregate. It will be shown that this implies restrictive relations between the productivity measures involved, including the weights of the individual production units, and the type of mean employed.

The order of this paper is as follows. Section 2 reviews basic accounting relations. Section 3 defines the problem. Sections 4 and 5 consider value-added based total factor productivity and labour productivity, respectively. Section 6 considers gross-output based productivity. Section 7 concludes.

2 Accounting framework

We consider[1] an ensemble (or set) $\mathcal{K}^t$ of consolidated production units[2], operating during a certain time period t in a certain country or region. For each unit the KLEMS-Y *ex post* accounting identity in nominal values (or, in current prices) reads

$$C_{KL}^{kt} + C_{EMS}^{kt} + \Pi^{kt} = R^{kt} \; (k \in \mathcal{K}^t), \tag{1}$$

where C_{KL}^{kt} denotes the primary input cost, C_{EMS}^{kt} the intermediate inputs cost, R^{kt} the revenue, and Π^{kt} the profit (defined as remainder). Intermediate inputs cost (on energy, materials, and business services) and revenue concern generally tradeable commodities. It is presupposed that there is some agreed-on commodity classification, such that C_{EMS}^{kt} and R^{kt} can be written as sums of quantities times (unit) prices of these commodities. Of course, for any production unit most of these quantities will be zero. It is also presupposed that output prices are available from a market or else can be imputed. Taxes on production are supposed to be allocated to the K and L classes.

The commodities in the capital class K concern owned tangible and intangible assets, organized according to industry, type, and age class. Each production unit uses certain quantities of those assets, and the configuration of assets used is in general unique for the unit. Thus, again, for any production unit most of the asset cells are empty. Prices are defined as unit user costs and, hence, capital input cost C_K^{kt} is a sum of prices times quantities.

Finally, the commodities in the labour class L concern detailed types of labour. Though any production unit employs specific persons with certain capabilities, it is usually their hours of work that count. Corresponding prices are hourly wages. Like the capital assets, the persons employed by a certain production unit are unique for that unit. It is presupposed that, wherever necessary, imputations have been made for self-employed workers. Henceforth, labour input cost C_L^{kt} is a sum of prices times quantities.

Total primary input cost is the sum of capital and labour input cost, $C_{KL}^{kt} = C_K^{kt} + C_L^{kt}$. Profit Π^{kt} is the balancing item and thus may be positive,

[1] This section has been copied from Balk (2016). Though the time diemnsion does not play an explicit role in the present paper, the notation is retained for consistency.

[2] "Consolidated" means that intra-unit deliveries are netted out. At the industry level, in some parts of the literature this is called "sectoral". At the economy level, "sectoral" output reduces to GDP plus imports, and "sectoral" intermediate input to imports. In terms of variables to be defined below, consolidation means that $C_{EMS}^{kkt} = R^{kkt} = 0$.

negative, or zero.

The KL-VA accounting identity then reads

$$C_{KL}^{kt} + \Pi^{kt} = R^{kt} - C_{EMS}^{kt} \equiv VA^{kt} \; (k \in \mathcal{K}^t), \tag{2}$$

where VA^{kt} denotes value added, defined as revenue minus intermediate inputs cost. In this paper it will always be assumed that $VA^{kt} > 0$.

We now consider whether the ensemble of production units $\mathcal{K}^t$ can be considered as a consolidated production unit. Though aggregation basically is addition, adding-up the KLEMS-Y relations (1) over all the units would imply double-counting because of deliveries between units. To see this, it is useful to split intermediate input cost and revenue into two parts, respectively concerning units belonging to the ensemble $\mathcal{K}^t$ and units belonging to the rest of the world. Thus,

$$C_{EMS}^{kt} = \sum_{k' \in \mathcal{K}^t, k' \neq k} C_{EMS}^{k'kt} + C_{EMS}^{ekt}, \tag{3}$$

where $C_{EMS}^{k'kt}$ is the cost of the intermediate inputs purchased by unit k from unit k', and C_{EMS}^{ekt} is the cost of the intermediate inputs purchased by unit k from the world beyond the ensemble $\mathcal{K}^t$. Similarly,

$$R^{kt} = \sum_{k' \in \mathcal{K}^t, k' \neq k} R^{kk't} + R^{ket}, \tag{4}$$

where $R^{kk't}$ is the revenue obtained by unit k from delivering to unit k', and R^{ket} is the revenue obtained by unit k from delivering to units outside of $\mathcal{K}^t$. Adding up the KLEMS-Y relations (1) then delivers

$$\sum_{k \in \mathcal{K}^t} C_{KL}^{kt} + \sum_{k \in \mathcal{K}^t} \sum_{k' \in \mathcal{K}^t, k' \neq k} C_{EMS}^{k'kt} + \sum_{k \in \mathcal{K}^t} C_{EMS}^{ekt} + \sum_{k \in \mathcal{K}^t} \Pi^{kt} =$$
$$\sum_{k \in \mathcal{K}^t} \sum_{k' \in \mathcal{K}^t, k' \neq k} R^{kk't} + \sum_{k \in \mathcal{K}^t} R^{ket}. \tag{5}$$

If for all the tradeable commodities output prices are identical to input prices (which is ensured by National Accounting conventions), then the two intra-$\mathcal{K}^t$-trade terms cancel, and the foregoing expression reduces to[3]

[3]See Balk (2015, footnote 2) for the treatment of net taxes on intermediates.

$$\sum_{k\in\mathcal{K}^t} C_{KL}^{kt} + \sum_{k\in\mathcal{K}^t} C_{EMS}^{ekt} + \sum_{k\in\mathcal{K}^t} \Pi^{kt} = \sum_{k\in\mathcal{K}^t} R^{ket}. \tag{6}$$

Recall that capital assets and hours worked are unique for each production unit, which implies that primary input cost may simply be added over the units, without any fear for double-counting. Thus expression (6) is the KLEMS-Y accounting relation for the ensemble $\mathcal{K}^t$, considered as a consolidated production unit. The corresponding KL-VA relation is then

$$\sum_{k\in\mathcal{K}^t} C_{KL}^{kt} + \sum_{k\in\mathcal{K}^t} \Pi^{kt} = \sum_{k\in\mathcal{K}^t} R^{ket} - \sum_{k\in\mathcal{K}^t} C_{EMS}^{ekt}, \tag{7}$$

which can be written as

$$C_{KL}^{\mathcal{K}^t t} + \Pi^{\mathcal{K}^t t} = R^{\mathcal{K}^t t} - C_{EMS}^{\mathcal{K}^t t} \equiv VA^{\mathcal{K}^t t}. \tag{8}$$

where $C_{KL}^{\mathcal{K}^t t} \equiv \sum_{k\in\mathcal{K}^t} C_{KL}^{kt}$, $\Pi^{\mathcal{K}^t t} \equiv \sum_{k\in\mathcal{K}^t} \Pi^{kt}$, $R^{\mathcal{K}^t t} \equiv \sum_{k\in\mathcal{K}^t} R^{ket}$, and $C_{EMS}^{\mathcal{K}^t t} \equiv \sum_{k\in\mathcal{K}^t} C_{EMS}^{ekt}$. One verifies immediately that

$$VA^{\mathcal{K}^t t} = \sum_{k\in\mathcal{K}^t} VA^{kt}. \tag{9}$$

The structural similarity between expressions (2) and (8), together with the additive relations between all their elements, is the reason why the KL-VA production model is the natural starting point for studying the relation between individual and aggregate measures of productivity change. We will soon discover, however, that the bottom-up approach basically neglects this framework.

3 Bottom-up and top-down approaches connected

Let the productivity level[4] of unit k at period t be denoted by $PROD^{kt}$. The generic definition here employed is: real output divided by real input. Output can be measured as revenue (also called 'gross output') (R^{kt}) or as value added (VA^{kt}). Input can be measured as total cost ($C_{KLEMS}^{kt} \equiv C_{KL}^{kt} + C_{EMS}^{kt}$), as primary input cost (C_{KL}^{kt}), as labour input cost (C_{L}^{kt}), or as total labour

[4]On the relation between levels and indices, see Balk (2016, 21-28).

quantity (L^{kt}, where a common unit is used for the various types of labour). In all these cases, 'real' means nominal deflated by some price index, which may or may not be specific for each production unit. It is supposed that the reference period b, that is, the period for which the price index equals 1 by definition, is the same for all the units.

Each production unit comes with some measure of relative size (or, importance) in the form of a weight θ^{kt}. For each period these weights usually but not necessarily add up to 1.

The question which weights θ^{kt} are appropriate when a choice has been made as to the productivity levels $PROD^{kt}$ ($k \in \mathcal{K}^t$) has received some attention in the literature. Given that somehow $PROD^{kt}$ is output divided by input, should the weight θ^{kt} be output- or input-based? And how is this related to the type of mean – arithmetic, geometric, or harmonic? The literature does not provide us with definitive answers.[5] Indeed, as long as one stays in the bottom-up framework it is unlikely that a convincing answer can be obtained. We need the complementary top-down view.

A bit formally, the problem can be posed as follows. Generalizing the definitions introduced in Balk (2016), *aggregate productivity* is a weighted 'mean' of the individual productivities

$$PROD^t \equiv M(\theta^{kt}, PROD^{kt}; k \in \mathcal{K}^t), \tag{10}$$

where the 'mean' $M(.)$ can be arithmetic, geometric, or harmonic; the weights θ^{kt} may or may not add up to 1; and $PROD^{kt}$ can be value-added based total factor productivity, labour productivity or simple labour productivity, as defined in Balk (2016), or gross-output based total factor productivity or simple labour productivity, to be defined in this paper.

Microdata studies, where the production units considered are plants or enterprises, then concentrate on the distributional characteristics of the (large) set of individual productivities $PROD^{kt}$, the development over time of aggregate productivity $PROD^t$, and the decomposition of this development with respect to several types of firms.

Sectoral studies, where the production units considered are industries (according to some national or international classification), are usually inter-

[5]De Loecker and Konings (2006) noted that there is no clear consensus on the appropriate weights (shares) that should be used. In their own work they used employment based shares $L^{kt}/\sum_k L^{kt}$ to weigh value-added based total factor productivity indices $Q^k_{VA}(t,b)/Q^k_{KL}(t,b)$. We will return to this example.

ested in industry-specific productivity change and its components, such as capital deepening and labour-composition change. The number of industries distinguished is generally so small that separate attention can be devoted to each specific case.

In both situations the ensemble $\mathcal{K}^t$ itself can be considered as a (consolidated) higher level production unit. Using the same definitions, its productivity $PROD^{\mathcal{K}^t t}$ can be calculated. In general it will then turn out, explicitly or implicitly, that the *productivity of the aggregate*, $PROD^{\mathcal{K}^t t}$, is unequal to aggregate productivity, $PROD^t$, as defined above.[6]

Microdata analysis is usually not interested in the productivity of the aggregate. As a consequence the problem of the choice of weights and type of mean arises. Sectoral analysis usually does show productivity change of the aggregate (e.g., the economy) alongside productivity change of the component industries, however without an explicit discussion of their relationship. If there is some comparison of aggregate productivity change and productivity change of the aggregate at all, then their difference is classified as an "unexplained residual".

In this paper we will ask whether it is possible to find a set of weights and a type of 'mean' such that

$$PROD^t = PROD^{\mathcal{K}^t t}; \tag{11}$$

that is, such that aggregate productivity can be interpreted as productivity of the aggregate.

As we know, there are a number of options here. We start with the case where $PROD^{kt}$ and $PROD^{\mathcal{K}^t t}$ is value-added based total factor productivity. Next we consider value-added based labour productivity. Finally we turn to gross-output based labour and total factor productivity respectively.

[6] $PROD^t$ can be considered as a 2-stage aggregation procedure: first $PROD^{kt}$ aggregates over basic inputs and outputs per production unit k, and then $PROD^t$ aggregates over all the units $k \in \mathcal{K}^t$. $PROD^{\mathcal{K}^t t}$ can be considered as a 1-stage aggregate of the same basic inputs and outputs. See Diewert (1980, 495-498) for a similar discussion in terms of variable profit (or, value added) functions and technological change (assuming continuous time and differentiability), and the PPI Manual (2004, Chapter 18) for the cases of revenue, intermediate-input-cost, and value-added based price indices.

4 Value-added based total factor productivity

The top-down approach starts with the adding-up relation (9). This relation tells us that nominal value added of the ensemble $\mathcal{K}^t$ is the sum of nominal value added of the individual production units k making up this ensemble. Next it is important to recall that the KL-VA accounting identities of the individual units, given by expression (2), are structurally identical to the KL-VA accounting identity of the ensemble (8). This means that we can treat the ensemble as a higher level production unit, and that all the definitions of indices and levels can be applied to the individual units and the ensemble in the same way.

Real value added is nominal value added, VA^{kt}, divided (or, deflated) by some price index with reference period b, $P^k_{VA}(t,b)$. Rewriting this definition gives

$$VA^{kt} = P^k_{VA}(t,b) RVA^k(t,b) \ (k \in \mathcal{K}^t). \tag{12}$$

Nominal value added is here as it were decomposed into a price component and a quantity component. For the ensemble we have similarly

$$VA^{\mathcal{K}^t t} = P^{\mathcal{K}^t}_{VA}(t,b) RVA^{\mathcal{K}^t}(t,b), \tag{13}$$

where $P^{\mathcal{K}^t}_{VA}(t,b)$ is a value-added based price index for the ensemble $\mathcal{K}^t$ for period t relative to the reference period b. This index is supposed to be estimated from a sample of enterprises and products.

Substituting expressions (12) and (13) into expression (9) and dividing both sides by the price index $P^{\mathcal{K}^t}_{VA}(t,b)$ delivers a relation between real value added of the ensemble and real value added of the individual units,

$$RVA^{\mathcal{K}^t}(t,b) = \sum_{k \in \mathcal{K}^t} \frac{P^k_{VA}(t,b)}{P^{\mathcal{K}^t}_{VA}(t,b)} RVA^k(t,b). \tag{14}$$

It is important to observe that, unlike nominal value added – see expression (9) – real value added appears in general to be not additive; that is, $RVA^{\mathcal{K}^t}(t,b) \neq \sum_{k \in \mathcal{K}^t} RVA^k(t,b)$.

For any individual production unit, real primary input is defined by

$$X^k_{KL}(t,b) \equiv C^{kt}_{KL}/P^k_{KL}(t,b) \ (k \in \mathcal{K}^t) \tag{15}$$

where $P^k_{KL}(t,b)$ is a suitable deflator for the primary input cost of production unit k. For the ensemble the corresponding definition reads

$$X^{\mathcal{K}^t}_{KL}(t,b) \equiv C^{\mathcal{K}^t t}_{KL}/P^{\mathcal{K}^t}_{KL}(t,b), \tag{16}$$

where $C^{\mathcal{K}^t t}_{KL} \equiv \sum_{k\in\mathcal{K}^t} C^{kt}_{KL}$ and $P^{\mathcal{K}^t}_{KL}(t,b)$ is a suitable deflator for the primary input cost of the ensemble $\mathcal{K}^t$. Now, dividing both sides of expression (14) by $X^{\mathcal{K}^t}_{KL}(t,b)$ and inserting at the right-hand side $X^k_{KL}(t,b)/X^k_{KL}(t,b) = 1$ $(k \in \mathcal{K}^t)$, one obtains

$$\frac{RVA^{\mathcal{K}^t}(t,b)}{X^{\mathcal{K}^t}_{KL}(t,b)} = \sum_{k\in\mathcal{K}^t} \frac{P^k_{VA}(t,b)}{P^{\mathcal{K}^t}_{VA}(t,b)} \frac{X^k_{KL}(t,b)}{X^{\mathcal{K}^t}_{KL}(t,b)} \frac{RVA^k(t,b)}{X^k_{KL}(t,b)}. \tag{17}$$

At both sides of this identity we see value-added based total factor productivity, as introduced by Balk (2016), for the aggregate and the individual production units, respectively. Thus expression (17) can be written as

$$TFPROD^{\mathcal{K}^t}_{VA}(t,b) = \sum_{k\in\mathcal{K}^t} \frac{P^k_{VA}(t,b)}{P^{\mathcal{K}^t}_{VA}(t,b)} \frac{X^k_{KL}(t,b)}{X^{\mathcal{K}^t}_{KL}(t,b)} TFPROD^k_{VA}(t,b). \tag{18}$$

This is our desired result. It means that if $PROD^{kt}$ is defined as value-added based total factor productivity $TFPROD^k_{VA}(t,b)$, then the appropriate weights are given by

$$\phi^{kt} \equiv \frac{P^k_{VA}(t,b)}{P^{\mathcal{K}^t}_{VA}(t,b)} \frac{X^k_{KL}(t,b)}{X^{\mathcal{K}^t}_{KL}(t,b)} \; (k \in \mathcal{K}^t). \tag{19}$$

When these weights are used, then aggregate productivity $\sum_{k\in\mathcal{K}^t} \phi^{kt} PROD^{kt}$ can be interpreted as the value-added based total factor productivity of the ensemble, considered as a higher-level production unit. Notice that the weights ϕ^{kt} $(k \in \mathcal{K}^t)$ not necessarily add up to 1. Thus, though expression (18) is a weighted sum of individual productivities it is not a genuine mean.[7]

[7] Stated in our notation, Basu and Fernald (2002) consider

$$\sum_{k\in\mathcal{K}^t} \frac{VA^{kt}}{VA^{\mathcal{K}^t t}} TFPROD^k_{VA}(t,b);$$

that is, mean value-added based total factor productivity where the weights are nominal value-added shares. This, then, cannot be interpreted as value-added based total factor productivity of the ensemble, unless special conditions apply.

There is, however, another way of looking at expression (18). To see this, notice that $(P^k_{VA}(t,b)/P^{\mathcal{K}^t}_{VA}(t,b))\,TFPROD^k_{VA}(t,b)$ is so-called revenue total factor productivity; that is, the result of deflating VA^{kt} not by its unit-k-specific deflator $P^k_{VA}(t,b)$ but by the ensemble-specific deflator $P^{\mathcal{K}^t}_{VA}(t,b)$. Weighing these revenue total factor productivities by real input shares X^k_{KL} $(t,b)/X^{\mathcal{K}^t}_{KL}(t,b)$ then delivers aggregate total factor productivity. Notice that these real input shares also not necessarily add up to 1.

Expression (18) as a relation between aggregate and individual productivities is, however, not unique. To see this, instead of the adding-up relation for value added (9), we consider the adding-up relation for primary input cost,

$$C^{\mathcal{K}^t t}_{KL} = \sum_{k\in\mathcal{K}^t} C^{kt}_{KL}. \tag{20}$$

Employing definitions (15) and (16), expression (20) can be rewritten as

$$X^{\mathcal{K}^t}_{KL}(t,b) = \sum_{k\in\mathcal{K}^t} \frac{P^k_{KL}(t,b)}{P^{\mathcal{K}^t}_{KL}(t,b)} X^k_{KL}(t,b). \tag{21}$$

It is not unimportant to observe that, unlike nominal primary input cost, real primary input appears generally to be non-additive; that is, $X^{\mathcal{K}^t}_{KL}(t,b) \neq \sum_{k\in\mathcal{K}^t} X^k_{KL}(t,b)$.

Individual and aggregate real value added were defined by expressions (12) and (13) respectively. Now, dividing both sides of expression (21) by $RVA^{\mathcal{K}^t}(t,b)$ and inserting at the right-hand side $RVA^k(t,b)/RVA^k(t,b) = 1$ $(k \in \mathcal{K}^t)$, one obtains

$$\frac{X^{\mathcal{K}^t}_{KL}(t,b)}{RVA^{\mathcal{K}^t}(t,b)} = \sum_{k\in\mathcal{K}^t} \frac{P^k_{KL}(t,b)}{P^{\mathcal{K}^t}_{KL}(t,b)} \frac{RVA^k(t,b)}{RVA^{\mathcal{K}^t}(t,b)} \frac{X^k_{KL}(t,b)}{RVA^k(t,b)}. \tag{22}$$

Again employing the definition of value-added based total factor productivity, expression (22) can be written as

$$\left(TFPROD^{\mathcal{K}^t}_{VA}(t,b)\right)^{-1} = \sum_{k\in\mathcal{K}^t} \frac{P^k_{KL}(t,b)}{P^{\mathcal{K}^t}_{KL}(t,b)} \frac{RVA^k(t,b)}{RVA^{\mathcal{K}^t}(t,b)} \left(TFPROD^k_{VA}(t,b)\right)^{-1}, \tag{23}$$

or

$$TFPROD_{VA}^{\mathcal{K}^t}(t,b) = \left(\sum_{k\in\mathcal{K}^t} \frac{P_{KL}^k(t,b)}{P_{KL}^{\mathcal{K}^t}(t,b)} \frac{RVA^k(t,b)}{RVA^{\mathcal{K}^t}(t,b)} \left(TFPROD_{VA}^k(t,b) \right)^{-1} \right)^{-1}. \tag{24}$$

This is our alternative result. Thus, aggregate total factor productivity can also be obtained as a weighted *harmonic* sum of individual productivities, with weights

$$\psi^{kt} \equiv \frac{P_{KL}^k(t,b)}{P_{KL}^{\mathcal{K}^t}(t,b)} \frac{RVA^k(t,b)}{RVA^{\mathcal{K}^t}(t,b)} \; (k \in \mathcal{K}^t). \tag{25}$$

Notice that these weights not necessarily add up to 1. It is interesting to compare the structure of the two sets of weights ϕ^{kt} and ψ^{kt}. The former are based on real primary input shares and relative value-added price levels, whereas the latter are based on real output (value added) shares and relative primary input price levels.

Summarizing, there is no unique relation between the individual total factor productivities and the total factor productivity of the aggregate. One must either multiply the individual productivities by weights ϕ^{kt} and add up, or use weights ψ^{kt} and take the harmonic sum.

Additivity imposed

We observed that both real valued added and real primary input are generally non-additive.[8] A sufficient condition for additivity is that deflators for the ensemble, for value added as well as primary input, are Paasche-type indices. This can be seen as follows. Additivity of real value added,

$$RVA^{\mathcal{K}^t}(t,b) = \sum_{k\in\mathcal{K}^t} RVA^k(t,b), \tag{26}$$

is, by inserting the definitions of real value added, equivalent to

$$\frac{1}{P_{VA}^{\mathcal{K}^t}(t,b)} = \sum_{k\in\mathcal{K}^t} \frac{VA^{kt}}{VA^{\mathcal{K}^t t}} \frac{1}{P_{VA}^k(t,b)}. \tag{27}$$

[8] Notice that we are considering here additivity between production units, which is different from additivity between commodities as considered in Balk (2016, Section 4.2).

But this relation simply expresses that the value-added based deflator for the ensemble is a Paasche index of the deflators for the individual production units (recall that nominal value added is additive). Similarly, if the primary-input based deflator for the ensemble is a Paasche index of the unit-specific deflators, then

$$X_{KL}^{\mathcal{K}^t}(t,b) = \sum_{k\in\mathcal{K}^t} X_{KL}^{k}(t,b). \tag{28}$$

It is straightforward to check that if conditions (26) and (28) are satisfied, then instead of expression (18) we obtain the simpler expression

$$TFPROD_{VA}^{\mathcal{K}^t}(t,b) = \sum_{k\in\mathcal{K}^t} \frac{X_{KL}^{k}(t,b)}{X_{KL}^{\mathcal{K}^t}(t,b)} TFPROD_{VA}^{k}(t,b), \tag{29}$$

and instead of expression (24) we obtain the simpler expression

$$TFPROD_{VA}^{\mathcal{K}^t}(t,b) = \left(\sum_{k\in\mathcal{K}^t} \frac{RVA^{k}(t,b)}{RVA^{\mathcal{K}^t}(t,b)} \left(TFPROD_{VA}^{k}(t,b)\right)^{-1}\right)^{-1}. \tag{30}$$

In both cases the weights now add up to 1. The result is simple to summarize. If one weighs individual total factor productivities by real input shares then the arithmetic mean must be used, but if one weighs by real output shares then the harmonic mean must be used to arrive at an interpretable result.

Mixing this leads to unwanted effects. For example, combining the harmonic mean with real input shares leads to understating the productivity of the aggregate:

$$\left(\sum_{k\in\mathcal{K}^t} \frac{X_{KL}^{k}(t,b)}{X_{KL}^{\mathcal{K}^t}(t,b)} \left(TFPROD_{VA}^{k}(t,b)\right)^{-1}\right)^{-1} \leq TFPROD_{VA}^{\mathcal{K}^t}(t,b), \tag{31}$$

and combining the arithmetic mean with real output shares leads to overstating the productivity of the aggregate:

$$\sum_{k\in\mathcal{K}^t} \frac{RVA^{k}(t,b)}{RVA^{\mathcal{K}^t}(t,b)} TFPROD_{VA}^{k}(t,b) \geq TFPROD_{VA}^{\mathcal{K}^t}(t,b). \tag{32}$$

Both results rest on combining the mathematical fact that an harmonic mean is always less than or equal to an arithmetic mean with expressions (29) and

(30). Equality in expressions (31) and (32) holds only when all the individual productivities $TFPROD^k_{VA}(t,b)$ ($k \in \mathcal{K}^t$) are the same. Interestingly, the left-hand side of expression (32) is the target variable considered by Olley and Pakes (1996).

Also the type of mean matters. A geometric mean is greater than or equal to an harmonic mean, which implies that, using expression (30),

$$\prod_{k\in\mathcal{K}^t} \left(TFPROD^k_{VA}(t,b)\right)^{RVA^k(t,b)/RVA^{\mathcal{K}^t}(t,b)} \geq TFPROD^{\mathcal{K}^t}_{VA}(t,b). \tag{33}$$

Such a geometric mean was as target variable considered by Melitz and Polanec (2015). It is thus seen to overstate productivity of the aggregate.

Returning to the De Loecker and Konings (2006) case, it can be seen that instead of the right-hand side of expression (29) these authors considered

$$\sum_{k\in\mathcal{K}^t} \frac{L^{kt}}{L^{\mathcal{K}^t t}} TFPROD^k_{VA}(t,b), \tag{34}$$

which is a biased estimator of $TFPROD^{\mathcal{K}^t}_{VA}(t,b)$. The magnitude of the bias, and its sign, is of course an empirical matter.

5 Value-added based labour productivity

For value-added based labour productivity the setup of the previous section can simply be repeated. The only thing one needs to do is replacing real primary input by real labour input. Thus, for the individual production units real labour input is defined as

$$X^k_L(t,b) \equiv C^{kt}_L / P^k_L(t,b) \; (k \in \mathcal{K}^t). \tag{35}$$

Likewise, for the ensemble

$$X^{\mathcal{K}^t}_L(t,b) \equiv C^{\mathcal{K}^t t}_L / P^{\mathcal{K}^t}_L(t,b), \tag{36}$$

where $C^{\mathcal{K}^t t}_L \equiv \sum_{k\in\mathcal{K}^t} C^{kt}_L$ and $P^k_L(t,b)$ and $P^{\mathcal{K}^t}_L(t,b)$ are suitable deflators for the labour cost of the individual production units and the ensemble, respectively.

Labour productivity was defined as real value added divided by real labour input. Starting from the numerator of the labour productivity of the ensemble the decomposition appears to be

$$LPROD_{VA}^{\mathcal{K}^t}(t,b) = \sum_{k\in\mathcal{K}^t} \frac{P_{VA}^k(t,b)}{P_{VA}^{\mathcal{K}^t}(t,b)} \frac{X_L^k(t,b)}{X_L^{\mathcal{K}^t}(t,b)} LPROD_{VA}^k(t,b), \tag{37}$$

whereas starting from the denominator one obtains

$$LPROD_{VA}^{\mathcal{K}^t}(t,b) = \left(\sum_{k\in\mathcal{K}^t} \frac{P_L^k(t,b)}{P_L^{\mathcal{K}^t}(t,b)} \frac{RVA^k(t,b)}{RVA^{\mathcal{K}^t}(t,b)} \left(LPROD_{VA}^k(t,b) \right)^{-1} \right)^{-1}. \tag{38}$$

Both expressions relate value-added based labour productivity of the ensemble, considered as a higher level production unit, to the labour productivities of the constituent production units. Notice that the weights not necessarily add up to 1.

Simple labour productivity

Two special cases deserve our attention. First, when for labour the simple sum quantity index is used then for the individual production units labour productivity is given by

$$LPROD_{VA}^k(t,b) = \frac{RVA^k(t,b)}{C_L^{kt}/P_L^k(t,b)} = \frac{RVA^k(t,b)}{C_L^{kb} Q_L^k(t,b)} = \frac{RVA^k(t,b)}{(C_L^{kb}/L^{kb})L^{kt}} \; (k \in \mathcal{K}^t), \tag{39}$$

and real labour input by $X_L^k(t,b) = (C_L^{kb}/L^{kb})L^{kt}$, where $L^{k\tau}$ denotes production unit k's total labour quantity at period τ ($\tau = b,t$). For the ensemble similar expressions hold.

Substitution, for the individual production units as well as for the ensemble, into expression (37) and some simplification delivers the following expression,

$$\frac{RVA^{\mathcal{K}^t}(t,b)}{L^{\mathcal{K}^t t}} = \sum_{k\in\mathcal{K}^t} \frac{P_{VA}^k(t,b)}{P_{VA}^{\mathcal{K}^t}(t,b)} \frac{L^{kt}}{L^{\mathcal{K}^t t}} \frac{RVA^k(t,b)}{L^{kt}}. \tag{40}$$

This is an expression in terms of simple labour productivities, as defined in Balk (2016). Put otherwise, expression (40) can be written as

$$SLPROD_{VA}^{\mathcal{K}^t}(t,b) = \sum_{k\in\mathcal{K}^t} \frac{P_{VA}^k(t,b)}{P_{VA}^{\mathcal{K}^t}(t,b)} \frac{L^{kt}}{L^{\mathcal{K}^t t}} SLPROD_{VA}^k(t,b). \tag{41}$$

It is quite natural to assume that the labour input of the ensemble, considered as a higher level production unit, is a simple sum of the labour inputs of the constituent units; that is, $L^{\mathcal{K}^t t} = \sum_{k\in\mathcal{K}^t} L^{kt}$. Then the fractions $L^{kt}/L^{\mathcal{K}^t t}$ ($k \in \mathcal{K}^t$) are labour shares, adding up to 1. Notice, however, that these labour shares are premultiplied by relative price levels, so that the weights of the labour productivities themselves do not necessarily add up to 1. The relative price levels vanish when $P_{VA}^k(t,b) = P_{VA}^{\mathcal{K}^t}(t,b)$ ($k \in \mathcal{K}^t$); that is, when there is no differential output price change among the production units.

There is, however, another way of looking at expression (41). To see this, recall that $(P_{VA}^k(t,b)/P_{VA}^{\mathcal{K}^t}(t,b))(RVA^k(t,b)/L^{kt})$ is so-called revenue labour productivity; that is, the result of deflating VA^{kt} not by its unit-k-specific deflator $P_{VA}^k(t,b)$ but by the ensemble-specific deflator $P_{VA}^{\mathcal{K}^t}(t,b)$. Weighing these revenue labour productivities by labour shares $L^{kt}/L^{\mathcal{K}^t t}$ then delivers aggregate labour productivity.

Finally, we notice that expression (40) is the model underlying the Generalized Exactly Additive Decomposition (GEAD) (see Balk 2016, Section 5.2). But it now turns out that an alternative decomposition can be developed.

To see this, notice that, by substituting expression (39) into expression (38) and using the product relation $C_L^{kt}/C_L^{kt'} = P_L^k(t,t')Q_L^k(t,t')$, expression (38) reduces to

$$SLPROD_{VA}^{\mathcal{K}^t}(t,b) = \left(\sum_{k\in\mathcal{K}^t} \frac{C_L^{kt}/L^{kt}}{C_L^{\mathcal{K}^t t}/L^{\mathcal{K}^t t}} \frac{RVA^k(t,b)}{RVA^{\mathcal{K}^t}(t,b)} \left(SLPROD_{VA}^k(t,b)\right)^{-1} \right)^{-1}. \tag{42}$$

This expression can be used to develop an alternative to the GEAD.

If the unit labour prices are the same across production units, that is, $C_L^{kt}/L^{kt} = \alpha$ ($k \in \mathcal{K}^t$) and $C_L^{\mathcal{K}^t t}/L^{\mathcal{K}^t t} = \alpha$, then expression (42) further reduces to

$$SLPROD_{VA}^{\mathcal{K}^t}(t,b) = \left(\sum_{k\in\mathcal{K}^t} \frac{RVA^k(t,b)}{RVA^{\mathcal{K}^t}(t,b)} \left(SLPROD_{VA}^k(t,b)\right)^{-1} \right)^{-1}. \tag{43}$$

An alternative route to obtain this expression is the following. The assumption of equal unit labour prices across production units implies that $L^{\mathcal{K}^t t} = \sum_{k\in\mathcal{K}^t} L^{kt}$. Then, starting with this relation, dividing its left- and right-hand sides by $RVA^{\mathcal{K}^t}(t,b)$, and inserting at the right-hand side $RVA^k(t,b)/RVA^k(t,b) = 1$ ($k \in \mathcal{K}^t$) one obtains expression (43).

Notice that the weights in expression (43) do not add up to 1, unless additivity holds.

Additivity imposed

Second, let us assume that additivity holds; that is, $RVA^{\mathcal{K}^t}(t,b) = \sum_{k\in\mathcal{K}^t} RVA^k(t,b)$ and $L^{\mathcal{K}^t t} = \sum_{k\in\mathcal{K}^t} L^{kt}$. Instead of expression (37) we then obtain

$$SLPROD_{VA}^{\mathcal{K}^t}(t,b) = \sum_{k\in\mathcal{K}^t} \frac{L^{kt}}{L^{\mathcal{K}^t t}} SLPROD_{VA}^k(t,b), \tag{44}$$

and instead of expression (38) we obtain

$$SLPROD_{VA}^{\mathcal{K}^t}(t,b) = \left(\sum_{k\in\mathcal{K}^t} \frac{RVA^k(t,b)}{RVA^{\mathcal{K}^t}(t,b)} \left(SLPROD_{VA}^k(t,b) \right)^{-1} \right)^{-1}. \tag{45}$$

Now in both cases the weights add up to 1. The labour-share weighted arithmetic mean of simple labour productivities appears to be equal to the real-value-added-share weighted harmonic mean of simple labour productivities, and both are equal to the simple labour productivity of the aggregate.

Mixing means and weights leads to undesirable results. Using the general relation between harmonic and arithmetic means, we conclude that

$$\left(\sum_{k\in\mathcal{K}^t} \frac{L^{kt}}{L^{\mathcal{K}^t t}} \left(SLPROD_{VA}^k(t,b) \right)^{-1} \right)^{-1} \leq SLPROD_{VA}^{\mathcal{K}^t}(t,b) \tag{46}$$

$$\sum_{k\in\mathcal{K}^t} \frac{RVA^k(t,b)}{RVA^{\mathcal{K}^t}(t,b)} SLPROD_{VA}^k(t,b) \geq SLPROD_{VA}^{\mathcal{K}^t}(t,b). \tag{47}$$

Thus, a labour-share weighted harmonic mean of simple labour productivities understates labour productivity of the aggregate, while a real-value-added-share weighted arithmetic mean overstates this. The second inequality was also obtained by Van Biesebroeck (2008), though in a less direct way.

Also here the type of mean matters. As a geometric mean is less then or equal to an arithmetic mean, we conclude that

$$\prod_{k \in \mathcal{K}^t} \left(SLPROD_{VA}^{k}(t,b) \right)^{L^{kt}/L^{\mathcal{K}^t t}} \leq SLPROD_{VA}^{\mathcal{K}^t}(t,b). \tag{48}$$

Such a geometric mean of simple labour productivities features prominently in Melitz and Polanec (2015). In the Appendix of their paper decompositions based on the left-hand side and the right-hand side of expression (48) are empirically compared. The geometric mean was also considered as target variable for firms by Hyytinen and Maliranta (2013) and Maliranta and Määttänen (2015).

Notice that the right-hand side of expression (44) is the target variable of the TRAD and CSLS decompositions considered in Balk (2016, Section 5.2). Thus these decompositions are consistent; that is, aggregate productivity can be interpreted as productivity of the aggregate. However, underlying this result is the assumption of additivity, which is pretty restrictive.

6 Gross-output based productivity

There are not so many microdata studies dealing with the concept of gross-output based productivity. For any individual production unit gross-output based total factor productivity is defined as

$$TFPROD_{Y}^{k}(t,b) \equiv \frac{Y^{k}(t,b)}{X_{KLEMS}^{k}(t,b)} = \frac{R^{kt}/P_{R}^{k}(t,b)}{C_{KLEMS}^{kt}/P_{KLEMS}^{k}(t,b)} \; (k \in \mathcal{K}^t). \tag{49}$$

In the numerator we have real revenue $Y^{k}(t,b)$; that is, nominal revenue R^{kt} deflated by a k-specific revenue based price index with reference period b, $P_{R}^{k}(t,b)$. In the denominator we have real KLEMS input $X_{KLEMS}^{k}(t,b)$; that is, nominal KLEMS input cost C_{KLEMS}^{kt} deflated by a k-specific KLEMS input based price index with the same reference period, $P_{KLEMS}^{k}(t,b)$; so that the ratio $TFPROD_{Y}^{k}(t,b)$ is a dimensionless variable.

Similarly, gross-output based simple labour productivity is defined as

$$SLPROD_{Y}^{k}(t,b) \equiv \frac{Y^{k}(t,b)}{L^{kt}} \; (k \in \mathcal{K}^t); \tag{50}$$

that is, real revenue per unit of labour. The dimension of this variable is money of reference period b.

Suppose that we have access to production-unit specific data such that either of these measures can be compiled. Which weights would be appropriate? We review a number of typical studies.

Simple labour productivity

Let us start with the target variable considered by Baily, Bartelsman and Haltiwanger (2001). This is $SLPROD_Y^k(t,b)$, though instead of unit-specific deflators industry-level deflators were used. The labour unit was an hour worked. These simple labour productivities were weighted by *labour* shares; that is, by $L^{kt}/L^{\mathcal{K}^t t} = L^{kt}/\sum_{k\in\mathcal{K}^t} L^{kt}$. Thus, aggregate productivity was compiled as[9]

$$LPROD_{BBH}^{\mathcal{K}^t}(t,b) \equiv \sum_{k\in\mathcal{K}^t} \frac{L^{kt}}{\sum_{k\in\mathcal{K}^t} L^{kt}} SLPROD_Y^k(t,b) = \frac{\sum_{k\in\mathcal{K}^t} Y^k(t,b)}{L^{\mathcal{K}^t t}}. \quad (51)$$

But what precisely does this mean? To see this, we must return to the accounting identities discussed in Section 2 and notice that

$$\sum_{k\in\mathcal{K}^t} R^{kt} = \sum_{k\in\mathcal{K}^t} \sum_{k'\in\mathcal{K}^t, k'\neq k} R^{kk't} + R^{\mathcal{K}^t t}. \quad (52)$$

Thus, total revenue is the sum of revenue obtained by internal deliveries (recall that $R^{kk't}$ is the revenue obtained by unit k from delivering to unit k') and aggregate revenue $R^{\mathcal{K}^t t}$, which is the revenue obtained by the ensemble $\mathcal{K}^t$, when the ensemble is considered as a consolidated production unit. Now, imposing additivity, that is, defining the aggregate revenue-based price index as a Paasche index of the k-specific revenue based price indices,

$$\frac{1}{P_R^{\mathcal{K}^t}(t,b)} \equiv \sum_{k\in\mathcal{K}^t} \frac{R^{kt}}{\sum_{k\in\mathcal{K}^t} R^{kt}} \frac{1}{P_R^k(t,b)}, \quad (53)$$

implies that expression (52) can be written as

[9]This measure was also considered by Foster, Haltiwanger and Krizan (2001). Actually, two variants were considered, one where the labour unit is an hour worked and one where it is a worker. The geometric alternative was employed by Hyytinen and Maliranta (2013) for plants; labour quantity was thereby measured in full time equivalents.

$$P_R^{\mathcal{K}^t}(t,b)\sum_{k\in\mathcal{K}^t} Y^k(t,b) = \sum_{k\in\mathcal{K}^t}\sum_{k'\in\mathcal{K}^t,k'\neq k} R^{kk't} + R^{\mathcal{K}^t t}, \tag{54}$$

or

$$\sum_{k\in\mathcal{K}^t} Y^k(t,b) = \frac{\sum_{k\in\mathcal{K}^t}\sum_{k'\in\mathcal{K}^t,k'\neq k} R^{kk't}}{P_R^{\mathcal{K}^t}(t,b)} + \frac{R^{\mathcal{K}^t t}}{P_R^{\mathcal{K}^t}(t,b)}. \tag{55}$$

If we define real revenue of the ensemble $\mathcal{K}^t$, considered as a consolidated production unit, by $Y^{\mathcal{K}^t}(t,b) \equiv R^{\mathcal{K}^t t}/P_R^{\mathcal{K}^t}(t,b)$, then expression (55) can be simplified to

$$\sum_{k\in\mathcal{K}^t} Y^k(t,b) = \frac{\sum_{k\in\mathcal{K}^t}\sum_{k'\in\mathcal{K}^t,k'\neq k} R^{kk't}}{P_R^{\mathcal{K}^t}(t,b)} + Y^{\mathcal{K}^t}(t,b). \tag{56}$$

Substituting expression (56) into expression (51) and applying definition (50) to the ensemble considered as a production unit delivers the following relation:

$$LPROD_{BBH}^{\mathcal{K}^t}(t,b) = SLPROD_Y^{\mathcal{K}^t}(t,b)\left(1 + \frac{\sum_{k\in\mathcal{K}^t}\sum_{k'\in\mathcal{K}^t,k'\neq k} R^{kk't}}{R^{\mathcal{K}^t t}}\right). \tag{57}$$

Since nominal revenue is non-negative, it appears that aggregate BBH productivity overstates simple labour productivity of the aggregate, and that the magnitude of the bias depends on the relative extent of the intra-ensemble deliveries. The bias vanishes only when there are no intra-ensemble deliveries.

Foster, Haltiwanger and Krizan (2001) considered simple labour productivities weighted by *real output* shares; that is,

$$LPROD_{FHK}^{\mathcal{K}^t}(t,b) \equiv \sum_{k\in\mathcal{K}^t} \frac{Y^k(t,b)}{\sum_{k\in\mathcal{K}^t} Y^k(t,b)} SLPROD_Y^k(t,b). \tag{58}$$

Applying the arithmetic-harmonic mean inequality, and definitions (50) and (51) respectively, we obtain

$$LPROD_{FHK}^{\mathcal{K}^t}(t,b) \geq \frac{\sum_{k\in\mathcal{K}^t} Y^k(t,b)}{L^{\mathcal{K}^t t}} = LPROD_{BBH}^{\mathcal{K}^t}(t,b). \tag{59}$$

The right-hand side is familiar from the foregoing. Combining expressions (59) and (57) we may conclude that, even in the case of industries exhibiting no intra-ensemble trade, $LPROD_{FHK}^{\mathcal{K}^t}(t,b)$ overstates simple labour productivity of the aggregate, $SLPROD_Y^{\mathcal{K}^t}(t,b)$.

Total factor productivity

We now turn to $TFPROD_Y^k(t,b)$, a key variable considered by Bartelsman and Dhrymes (1998). They had industry and time effects removed econometrically, but that does not need to concern us here. The individual gross-output based total factor productivities were weighted by *real KLEMS input* shares $X_{KLEMS}^k(t,b)/\sum_{k\in\mathcal{K}^t} X_{KLEMS}^k(t,b)$, so that aggregate total factor productivity was compiled as

$$\begin{aligned} TFPROD_{BD}^{\mathcal{K}^t}(t,b) &\equiv \sum_{k\in\mathcal{K}^t} \frac{X_{KLEMS}^k(t,b)}{\sum_{k\in\mathcal{K}^t} X_{KLEMS}^k(t,b)} TFPROD_Y^k(t,b) \\ &= \frac{\sum_{k\in\mathcal{K}^t} Y^k(t,b)}{\sum_{k\in\mathcal{K}^t} X_{KLEMS}^k(t,b)}. \end{aligned} \tag{60}$$

Notice that, assuming that additivity at the output side holds, the numerator is given by expression (56). For the denominator a similar expression can be derived. To see this, we again return to the accounting identities in Section 2 and notice that

$$\sum_{k\in\mathcal{K}^t} C_{EMS}^{kt} = \sum_{k\in\mathcal{K}^t} \sum_{k'\in\mathcal{K}^t, k'\neq k} C_{EMS}^{k'kt} + C_{EMS}^{\mathcal{K}^t t}, \tag{61}$$

Adding at both sides $C_{KL}^{\mathcal{K}^t t} = \sum_{k\in\mathcal{K}^t} C_{KL}^{kt}$, we obtain the following accounting relation:

$$\sum_{k\in\mathcal{K}^t} C_{KLEMS}^{kt} = \sum_{k\in\mathcal{K}^t} \sum_{k'\in\mathcal{K}^t, k'\neq k} C_{EMS}^{k'kt} + C_{KLEMS}^{\mathcal{K}^t t}. \tag{62}$$

Thus, total cost is the sum of cost incurred by internal deliveries (recall that $C_{EMS}^{k'kt}$ is the cost incurred by unit k for purchases from unit k') and aggregate cost $C_{KLEMS}^{\mathcal{K}^t t}$, which is the KLEMS input cost of the ensemble $\mathcal{K}^t$, considered as a consolidated production unit. Now, imposing additivity at the input side, that is, defining the aggregate KLEMS input based price index as a Paasche index of the k-specific KLEMS input based price indices,

$$\frac{1}{P_{KLEMS}^{\mathcal{K}^t}(t,b)} \equiv \sum_{k\in\mathcal{K}^t} \frac{C_{KLEMS}^{kt}}{\sum_{k\in\mathcal{K}^t} C_{KLEMS}^{kt}} \frac{1}{P_{KLEMS}^{k}(t,b)}, \tag{63}$$

implies that expression (62) can be written as

$$P_{KLEMS}^{\mathcal{K}^t}(t,b) \sum_{k\in\mathcal{K}^t} X_{KLEMS}^{k}(t,b) = \sum_{k\in\mathcal{K}^t} \sum_{k'\in\mathcal{K}^t, k'\neq k} C_{EMS}^{k'kt} + C_{KLEMS}^{\mathcal{K}^t t}, \tag{64}$$

or

$$\sum_{k\in\mathcal{K}^t} X_{KLEMS}^{k}(t,b) = \frac{\sum_{k\in\mathcal{K}^t} \sum_{k'\in\mathcal{K}^t, k'\neq k} C_{EMS}^{k'kt}}{P_{KLEMS}^{\mathcal{K}^t}(t,b)} + \frac{C_{KLEMS}^{\mathcal{K}^t t}}{P_{KLEMS}^{\mathcal{K}^t}(t,b)}. \tag{65}$$

If we define real KLEMS input of the ensemble $\mathcal{K}^t$, considered as a consolidated production unit, as $X_{KLEMS}^{\mathcal{K}^t}(t,b) \equiv C_{KLEMS}^{\mathcal{K}^t t}/P_{KLEMS}^{\mathcal{K}^t}(t,b)$, then expression (65) can be simplified to

$$\sum_{k\in\mathcal{K}^t} X_{KLEMS}^{k}(t,b) = \frac{\sum_{k\in\mathcal{K}^t} \sum_{k'\in\mathcal{K}^t, k'\neq k} C_{EMS}^{k'kt}}{P_{KLEMS}^{\mathcal{K}^t}(t,b)} + X_{KLEMS}^{\mathcal{K}^t}(t,b). \tag{66}$$

Substituting expressions (56) and (66) into expression (60) and applying definition (49) to the ensemble considered as a production unit delivers the following relation:

$$\begin{aligned} &TFPROD_{BD}^{\mathcal{K}^t}(t,b) \\ &\quad = TFPROD_{Y}^{\mathcal{K}^t}(t,b) \frac{1 + \sum_{k\in\mathcal{K}^t} \sum_{k'\in\mathcal{K}^t, k'\neq k} R^{kk't}/R^{\mathcal{K}^t t}}{1 + \sum_{k\in\mathcal{K}^t} \sum_{k'\in\mathcal{K}^t, k'\neq k} C_{EMS}^{k'kt}/C_{KLEMS}^{\mathcal{K}^t t}}. \end{aligned} \tag{67}$$

As observed in Section 2, National Accounting conventions imply that revenue and cost of the intra-ensemble transactions are equal; that is,

$$\sum_{k\in\mathcal{K}^t} \sum_{k'\in\mathcal{K}^t, k'\neq k} R^{kk't} = \sum_{k\in\mathcal{K}^t} \sum_{k'\in\mathcal{K}^t, k'\neq k} C_{EMS}^{k'kt}.$$

Thus the magnitude of the bias of aggregate BD total factor productivity depends on the magnitude of aggregate revenue $R^{\mathcal{K}^t t}$ relative to aggregate

KLEMS input cost $C^{\mathcal{K}^t t}_{KLEMS}$. Put otherwise, the magnitude of the bias depends on aggregate profit $\Pi^{\mathcal{K}^t t}$. If aggregate profit is positive (negative), then aggregate BD total factor productivity understates (overstates) total factor productivity of the aggregate. If aggregate profit equals 0, then the bias vanishes. A sufficient condition for zero aggregate profit is that $\Pi^{kt} = 0$ for each individual production unit $k \in \mathcal{K}^t$. Of course, the bias also vanishes in the trivial case when there are no intra-ensemble deliveries.

Foster, Haltiwanger and Krizan (2001) considered total factor productivities weighted by *real output* shares; that is,[10]

$$TFPROD^{\mathcal{K}^t}_{FHK}(t,b) \equiv \sum_{k\in\mathcal{K}^t} \frac{Y^k(t,b)}{\sum_{k\in\mathcal{K}^t} Y^k(t,b)} TFPROD^k_Y(t,b). \tag{68}$$

Applying the arithmetic-harmonic mean inequality and using the definitions in expressions (49) and (60), we find that

$$TFPROD^{\mathcal{K}^t}_{FHK}(t,b) \geq TFPROD^{\mathcal{K}^t}_{BD}(t,b). \tag{69}$$

Now expression (67) above tells us that, under additivity at the input and the output side, $TFPROD^{\mathcal{K}^t}_{BD}(t,b)$ is an unbiased measure of total factor productivity of the aggregate if there are no intra-ensemble deliveries. Thus, we may conclude that in the cases studied by Foster, Haltiwanger and Krizan, which were four-digit level industries where intra-industry deliveries are unlikely, $TFPROD^{\mathcal{K}^t}_{FHK}(t,b)$ most likely overstates total factor productivity of the aggregate, $TFPROD^{\mathcal{K}^t}_{Y}(t,b)$.

The target variable of Eslava, Haltiwanger, Kugler and Kugler (2013) appears to be

$$TFPROD^{\mathcal{K}^t}_{EHKK}(t,b) \equiv \prod_{k\in\mathcal{K}^t} \left(TFPROD^k_Y(t,b)\right)^{Y^k(t,b)/\sum_{k\in\mathcal{K}^t} Y^k(t,b)}; \tag{70}$$

that is, the geometric variant of the FHK measure defined by expression (68). Using subsequently the geometric-harmonic mean inequality, definition (49), and expression (60), we obtain

[10] Actually, their multi-factor productivity index, discussed in the extended version of this paper, can be seen as a special case of $TFPROD^k_Y(t,b)$.

$$TFPROD^{\mathcal{K}^t}_{EHKK}(t,b) \geq TFPROD^{\mathcal{K}^t}_{BD}(t,b). \tag{71}$$

As we have seen, the right-hand side of this expression may or may not approximate $TFPROD^{\mathcal{K}^t}_{Y}(t,b)$.

It is now interesting to consider a recent paper by Collard-Wexler and De Loecker (2015). These authors also dealt with $TFPROD^{k}_{Y}(t,b)$ $(k \in \mathcal{K}^t)$, but to obtain aggregate productivity the individual total factor productivities were weighted by *nominal revenue* shares $R^{kt}/\sum_{k\in\mathcal{K}^t} R^{kt}$. Thus aggregate productivity was defined as

$$TFPROD^{\mathcal{K}^t}_{CWL}(t,b) \equiv \sum_{k\in\mathcal{K}^t} \frac{R^{kt}}{\sum_{k\in\mathcal{K}^t} R^{kt}} TFPROD^{k}_{Y}(t,b). \tag{72}$$

To obtain an interpretation for this mean, we first relate it to the alternative where real shares $Y^k(t,b)/\sum_{k\in\mathcal{K}^t} Y^k(t,b)$ are used as weights,

$$\begin{aligned} & TFPROD^{\mathcal{K}^t}_{CWL}(t,b) = \\ & \quad \sum_{k\in\mathcal{K}^t} \frac{Y^k(t,b)}{\sum_{k\in\mathcal{K}^t} Y^k(t,b)} TFPROD^{k}_{Y}(t,b) \\ & \quad + \sum_{k\in\mathcal{K}^t} \left(\frac{R^{kt}}{\sum_{k\in\mathcal{K}^t} R^{kt}} - \frac{Y^k(t,b)}{\sum_{k\in\mathcal{K}^t} Y^k(t,b)} \right) TFPROD^{k}_{Y}(t,b). \end{aligned} \tag{73}$$

The first term at the right-hand side of this equation is familiar; it is the FHK measure as defined by expression (68). The second term has the form of a covariance, but there is in general no compelling reason for this covariance to be positive or negative, large or small. Taken together, on the assumption that the covariance in equation (73) equals 0, it seems likely that aggregate CWL productivity overstates the productivity of the aggregate.

Some empirical comparisons

The primary purpose of the classic paper by Foster, Haltiwanger and Krizan (2001) was to compare decompositions of intertemporal change of the three aggregate measures $TFPROD^{\mathcal{K}^t}_{FHK}(t,b)$, $LPROD^{\mathcal{K}^t}_{FHK}(t,b)$, and $LPROD^{\mathcal{K}^t}_{BBH}(t,b)$. They specifically examined the Foster-Haltiwanger-Krizan (FHK) and the Griliches-Regev (GR) decomposition methods (see Balk 2016, expressions (2.43) and (2.50) respectively). It turned out that, though the levels were of course different, the FHK decompositions of $\Delta TFPROD^{\mathcal{K}^t}_{FHK}(t,b)$ and $\Delta LPROD^{\mathcal{K}^t}_{FHK}\ (t,b)$ were strikingly similar. The levels as well as the FHK decompositions of $\Delta LPROD^{\mathcal{K}^t}_{FHK}(t,b)$ and $\Delta LPROD^{\mathcal{K}^t}_{BBH}(t,b)$ differed, however, remarkably. Interestingly, for the three aggregate measures the GR decomposition delivered almost the same results. Overall, the 'within' term appeared dominant.

7 Conclusion

Our overall conclusion is that not every combination of micro-, or meso-level productivities, weights, and aggregator function (mean) leads to a nice interpretation of aggregate productivity as productivity of the aggregate. Specifically:

- An arithmetic 'mean' of value-added based total factor productivities requires weights based on relative real primary input times relative value-added based price levels.

- An harmonic 'mean' of value-added based total factor productivities requires weights based on relative real value added times relative primary input price levels.

- Under additivity the relative price levels disappear from the expressions.

- Similar results hold for value-added based (simple) labour productivities.

- An arithmetic mean of gross-output (revenue) based simple labour productivities weighted with (physical) labour input shares is likely to overstate its aggregate counterpart.

- An arithmetic mean of gross-output (revenue) based total factor productivities weighted with real input shares approximates gross-output (revenue) based total factor productivity of the aggregate; the magnitude of the bias depends on aggregate profit.

References

[1] Baily, M. N., E. J. Bartelsman and J. Haltiwanger, 2001, "Labor Productivity: Structural Change and Cyclical Dynamics", The Review of Economics and Statistics 83, 420-433.

[2] Balk, B. M., 2010, "An Assumption-free Framework for Measuring Productivity Change", The Review of Income and Wealth 56, Special Issue 1, S224-S256.

[3] Balk, B. M., 2011, "Measuring and Decomposing Capital Input Cost, The Review of Income and Wealth 57, 490-512.

[4] Balk, B. M., 2014, "Dissecting Aggregate Output and Labour Productivity Change", Journal of Productivity Analysis 42, 35-43.

[5] Balk, B. M., 2015, "Measuring and Relating Aggregate and Subaggregate Total Factor Productivity Change without Neoclassical Assumptions", Statistica Neerlandica 69, 21-28.

[6] Balk, B. M., 2016, "The Dynamics of Productivity Change: A Review of the Bottom-up Approach", in Productivity and Efficiency Analysis, edited by W. H. Greene, L. Khalaf, R. C. Sickles, M. Veall and M.-C. Voia, Proceedings in Business and Economics (Springer International Publishing, Switzerland).

[7] Bartelsman, E. J. and Ph. J. Dhrymes, 1998, "Productivity Dynamics: US Manufacturing Plants, 1972-1986", Journal of Productivity Analysis 9, 5-34.

[8] Basu, S. and J. G. Fernald, 2002, "Aggregate Productivity and Aggregate Technology", European Economic Review 46, 963-991.

[9] Biesebroeck, J. van, 2008, "Aggregating and Decomposing Productivity", Review of Business and Economics LIII, 122-146.

[10] Collard-Wexler, A. and J. de Loecker, 2015, "Reallocation and Technology: Evidence from the U. S. Steel Industry", American Economic Review 105, 131-171.

[11] Diewert, W. E., 1980, "Aggregation Problems in the Measurement of Capital", in The Measurement of Capital, edited by Dan Usher (National Bureau of Economic Research, Cambridge MA; University of Chicago Press).

[12] Dumagan, J. C. and B. M. Balk, 2016, "Dissecting Aggregate Output and Labour Productivity Change: A Postscript on the Role of Relative Prices", Journal of Productivity Analysis 45, 117-119.

[13] Eslava, M., J. Haltiwanger, A. Kugler and M. Kugler, 2013, "Trade and Market Selection: Evidence from Manufacturing Plants in Colombia", Review of Economic Dynamics 16, 135-158.

[14] Foster, L., J. Haltiwanger and C. J. Krizan, 2001, "Aggregate Productivity Growth: Lessons from Microeconomic Evidence", in New Developments in Productivity Analysis, edited by C. R. Hulten, E. R. Dean and M. J. Harper, Studies in Income and Wealth Volume 63 (The University of Chicago Press, Chicago and London).

[15] Hyytinen, A. and M. Maliranta, 2013, "Firm Lifecycles and Evolution of Industry Productivity", Research Policy 42, 1080-1098.

[16] de Loecker, J. and J. Konings, 2006, "Job Reallocation and Productivity Growth in a Post-Socialist Economy: Evidence from Slovenian Manufacturing", European Journal of Political Economy 22, 388-408.

[17] Olley S. and A. Pakes, 1996, "The Dynamics of Productivity in the Telecommunications Equipment Industry", Econometrica 64, 1263-1297.

[18] Maliranta, M. and N. Määttänen, 2015, "An Augmented Static Olley-Pakes Productivity Decomposition with Entry and Exit: Measurement and Interpretation", Economica 82, 1372-1416.

[19] Melitz, M. J. and S. Polanec, 2015, "Dynamic Olley-Pakes Productivity Decomposition with Entry and Exit", Rand Journal of Economics 46, 362-375.

[20] PPI Manual, 2004, Producer Price Index Manual: Theory and Practice (Published for ILO, IMF, OECD, UN, Eurostat, The World Bank by IMF, Washington, DC).

[21] Vancauteren, M., E. Veldhuizen and B. M. Balk, 2012, “Measures of Productivity Change: Which Outcome Do You Want?”, Paper presented at the 32nd General Conference of the IARIW, Boston MA, 5-11 August 2012.