AGGREGATE PRODUCTIVITY GROWTH IN THE PRESENCE OF (PERSISTENTLY) HETEROGENEOUS FIRMS

Giovanni Dosi†, Marco Grazzi°, Le Li‡, Luigi Marengo¶, and Simona Settepanella◇

†Institute of Economics, Scuola Superiore Sant'Anna, Italy
°Department of Economic Policy, Università Cattolica del Sacro Cuore, Italy
‡Institute of Business Research, Chuo University, Japan
¶Department of Business and Management, LUISS University, Italy
◇Department of Mathematics, Hokkaido University, Japan

June 2019

Abstract

In this paper we introduce a new decomposition method for aggregate productivity growth. Our methodology does not require to impose a production function but uses only the empirical input output data of production. Moreover it computes productivity in the very same way both at the individual establishment level and at the aggregate level, while standard methods compute the latter as a weighted average of the former, were the choice of weights is arbitrary and the loss of information implied by aggregation is substantial. We show that our methodology is particularly appropriate when production units are more heterogeneous and we test out methodology both with artificial and with empirical data.

JEL codes: D24; C67; C81; O30
Keywords: Productivity measure; Decomposition of aggregate productivity growth; Firm heterogeneity

1 Introduction

In recent years an extremely robust evidence regarding firm- and plant- level longitudinal microdata has highlighted striking and persistent heterogeneity across firms operating in the same industry. A large body of research from different sectors in different countries (Baily et al., 1992; Baldwin and Rafiquzzaman, 1995; Bartelsman and Doms, 2000; Disney et al., 2003a; Dosi, 2007; Syverson, 2011) documents the emergence of the following "stylized facts": first, wide asymmetries in productivity across firms; second, significant heterogeneity in relative input intensities even in presence of the same relative input prices; third, high intertemporal persistence in the above properties. Fourth, such heterogeneity is maintained also when increasing the level of disaggregation, thus plausibly reducing the diversity across firms' output.

The latter property has been vividly summarized by Grilliches and Mairesse (1999) " *We [...] thought that one could reduce heterogeneity by going down from general mixtures as "total manufacturing" to something more coherent, such as "petroleum refining" or "the manufacture of cement". But something like Mandelbrot's fractal phenomenon seems to be at work here also: the observed variability-heterogeneity does not really decline as we cut our data finer and finer. There is a sense in which different bakeries are just as much different from each others as the steel industry is from the machinery industry.*"

The bottom line is that firms operating in the same industry display a large and persistent degree of technological heterogeneity while there does not seem to be any clear sign that either the diffusion of information on different technologies, or the working of the competitive mechanism bring about any substantial reduction of such a heterogeneity, even when involving massive differences in efficiencies, as most incumbent theories would predict.

This evidence poses serious challenges not only to theory of competition and market selection, but also to any theoretical or empirical analysis which relies upon some notion of industry or sector defined as a set of production units producing under rather similar input prices with equally similar technologies, and the related notion of "the technology" of an industry represented by means of a sectoral production function. Indeed, the aggregation conditions needed to yield the canonic production functions building from the technologies of micro entities are extremely demanding, basically involving the identity of the latter up to a constant multiplier (Fisher, 1965; Hulten, 2001).

Note that these problems do not only concern the neoclassical production function, whose well known properties may either not fit empirical data or fit only spuriously, but also non neoclassical representations of production at the industry level. If input-output coefficients à la Leontief (1986) are averages over distributions with high standard deviations and high skewness, average input coefficients may not provide a meaningful representation of the technology of that industry. Moreover, one cannot take for granted that changes of such coefficients can be interpreted as indicators of productivity change as they may be just caused by some changes in the distribution of production among heterogeneous units, characterized by unchanged technologies.

How does one then account for the actual technology - or, better, the different techniques - in such industry? Hildenbrand (1981) suggests a direct and agnostic approach which instead of estimating some aggregate production function, offers a representation of the empirical production possibility set of an industry in the short run based on actual microdata. Each production unit is represented as a point in the input-output space whose coordinates are input requirements and output levels at full capacity. Under the sole assumptions of divisibility and additivity of production processes, the production possibility set is represented geometrically by the space formed by the finite sum of all the line segments linking the origin and the points representing each production unit, called a zonotope (see below). Hildenbrand then derives the actual "production function" (one should more accurately say "feasible" production function) and shows that " *short-run efficient production functions do not enjoy the well-known properties which are frequently assumed in production theory. For example, constant returns to scale never prevail, the production functions are never homothetic, and the elasticities of substitution are never constant. On the other hand, the competitive factor demand and product sup-*

ply functions [...] will always have definite comparative static properties which cannot be derived from the standard theory of production" (Hildenbrand, 1981, p. 1095).

Dosi et al. (2016) move a step forward and show that by further exploiting the properties of zonotopes it is possible to obtain rigorous measures of heterogeneity and productivity change without imposing on data a model like that implied by standard production functions. In particular, they develop a measure of industry productivity change that takes into consideration the entire observed production possibility set derived from observed heterogeneous production units, instead of considering only an efficient frontier.

The contribution of this work is twofold. First, we extend the industry-level productivity measure derived from the main diagonal of the (industry) zonotope (Dosi et al., 2016) to the firm level. Our firm-level measure of productivity maintains the high degree of flexibility (both in terms of number of inputs and outputs) as the industry-level measure and does not require to impose most of the assumptions generally imposed in the standard production function framework. Second, within the vectorial framework, i.e. zonotope, where multiple inputs are taken into account, we point out that, besides the classic "within" and "between" effects (Foster et al., 2001), the change of heterogeneity of the industry also contributes to the aggregate (industry) productivity growth. Based on the consistent measure of productivity in industry- and firm- levels abovementioned, we develop a new method to decompose the APG into three parts: two items, as counterparts of "within" and "between" effects, and one item measuring contribution from heterogeneity change. Some toy examples are provided to help understand our new method and empirical evidence to support change of heterogeneity as one non-trivial component of the decomposition of APG.

The rest of the paper is organized as follows. Section **??** summarizes the main methods for decomposing APG which have been used in previous studies and investigates the potential problems of these methods. Section 3 points out the change of heterogeneity in firm-level contributes to APG and proposes one decomposition method, based on vectorial setting up, to measure this contribution. Section 4 applies our proposed decomposition method on empirical data and shows that the contribution of the change of heterogeneity is not trivial based on empirical evidence. Finally, section 5 concludes.

2 Decomposition of APG in Previous Literature

Many empirical studies[1] highlight the effect of reallocation and heterogeneous productivity on aggregate productivity growth (APG). To do so, virtually they consider the decomposition of APG. In section 2.1, we summarize two main decomposition methods popular in previous studies according to Foster et al. (2001). In section 2.2 we discuss the potential issues with these methods.

2.1 A brief review of the existing decomposition methods

Foster et al. (2001) compare the two major decomposition methods of an industry aggregate productivity. The aggregate productivity P^t at time t is defined as the weighted

[1]Among the many, Baily et al. (1992); Olley and Pakes (1996); Bartelsman and Dhrymes (1998); Dwyer (1998); Haltiwanger (1997) use data from US, Aw et al. (1997) from Taiwan, Liu and Tybout (1996) from Chile and Colombia, Grilliches and Regev (1995) from Israel, and Dosi et al. (2015) from France, Germany, and the UK.

average of the productivities of individual firms:

$$P^t := \sum_{i \in I^t} w_i^t p_i^t \tag{1}$$

where I^t is the set of all firms in the industry at time t, w_i^t is the weight/share (e.g. output share) of firm i in this industry, and p_i^t is the productivity of firm i. Aggregate productivity growth is defined as the difference of the aggregate productivities between two (consecutive) time periods.

A decomposition method for aggregate productivity growth was introduced by Baily et al. (1992):

$$\sum_{i \in I^t} w_i^t p_i^t - \sum_{i \in I^{t-1}} w_i^{t-1} p_i^{t-1} = \sum_{i \in C} w_i^{t-1} \Delta p_i^t + \sum_{i \in C} p_i^t \Delta w_i^t + \sum_{i \in N} w_i^t p_i^t - \sum_{i \in X} w_i^{t-1} p_i^{t-1} \tag{2}$$

where C denotes continuing firms, i.e. those active in both periods, N denotes the set of firm are new entries in the industry, X denotes firms that have exited the industry and the operator Δ represents change from time $t-1$ to time t, i.e. $\Delta p_i^t = p_i^t - p_i^{t-1}$ and $\Delta w_i^t = w_i^t - w_i^{t-1}$. The first term, originally called "fixed shares", represents a within-firm component given by each firm's productivity change, weighted by the initial shares in the industry. The second term, originally called "share effect", represents a between-firm component that reflects changing shares, weighted by the productivities of the final period. The last two terms represent the contribution of the firms that, respectively, entered and exited the industry.

A second decomposition method was introduced by Grilliches and Regev (1995). It only differs from the previous one in the weights used to compute the "within" and the "between" effects. Rather than using either the initial or the final weights, this method employs their averages:

$$\sum_{i \in I^t} w_i^t p_i^t - \sum_{i \in I^{t-1}} w_i^{t-1} p_i^{t-1} = \sum_{i \in C} \bar{w}_i \Delta p_i^t + \sum_{i \in C} \bar{p}_i \Delta w_i^t + \sum_{i \in N} w_i^t p_i^t - \sum_{i \in X} w_i^{t-1} p_i^{t-1} \tag{3}$$

where the bar over a variable indicates the average $\bar{w}_i = \frac{w_i^t + w_i^{t-1}}{2}$ and $\bar{p}_i = \frac{p_i^t + p_i^{t-1}}{2}$. In this decomposition, the first term can be interpreted as a within effect which is measured by the sum of productivities weighted by the average (across time) shares. The second term represents a between effect where the changes in the shares are indexed by the average firm-level productivities. The last two terms represent the contribution of the firm which entered and exited the industry.[2]

Finally, Haltiwanger (1997) refines this decomposition by weighting the between effect with the deviations of initial firm productivities from the initial industry index. Because of this deviation term a continuing firm whose output increases and a new entry will contribute positively to the index only if their productivities are higher than the aggregate, while an exiting firm will contribute positively only if its productivity was lower than the aggregate. Without such a deviation the between effect may be non zero even when all individual productivities remain constant if the share of entering and exiting firms are different. Below we will discuss how our methodology deals with this kind of problem.

[2]To be more precise, Grilliches and Regev (1995) treat all entering and exiting firms as one firm.

2.2 Some problems with the current decomposition methods

Defining aggregate productivity as a weighted average of the productivities of individual firms leaves open the question of which weights should be used. Different weights will, in general, deliver different aggregate values and therefore also different decompositions.

As an illustration, consider the extreme case of a hypothetical industry composed of the two highly heterogeneous firms described in Table 1.

Table 1: An Hypothetical Industry with Two Firms

Firm	Labour	Output	Labour Productivity
A	1	100	100
B	100	1	0.01
Aggregate (Industry)	101	101	1

To compute the aggregate productivity (AP henceforth) of this industry we can use either input or output shares as weights. With the former we obtain an AP of 1, while with the latter we obtain a value very close to 100, i.e. the productivity of firm A.

Alternatively, instead of computing AP as a weighted average, we could compute it directly, by considering the industry as one large firm producing the industry total output with the industry total input. This measure is 1 in our example, and it is equal to the value obtained with inputs as weights because we are considering a case with only one input. With multiple inputs this equality does not hold in general.

When we have multiple inputs we normally aggregate them into a synthetic measure called total factor productivity (TFP). But the calculation of TFP requires the assumption of a specific production function and the empirical estimation of its parameters.

This "direct" aggregate measure presents some important advantages over measures obtained as weighted averages. The first and most important feature is that AP is computed exactly in the same way for the individual and for the aggregate. Suppose for instance that some production units merge into a single entity (for instance two firms merge legally, if our unit of analysis are firms) keeping exactly the same input-output structure. Weighted average measures would change[3] although the input-output structure of the industry has remained exactly the same. The productivity computed directly on the aggregate would instead remain unchanged.

More in general, no matter the method adopted, when we try to summarize multiple pieces of information, i.e. quantities of each input and quantity of output, into one number, i.e. the productivity level, we inevitably lose information. One number is not enough to represent the complete production activity and it is possible to have firms with very different production activities but same productivity level. Thus, since every time we compute a productivity value we lose information, in principle it is preferable to minimize the number of such computation. Now, if we compute AP as a weighted average of the productivities of n firms we will lose information $n+1$ times, whereas if we compute productivity directly on the aggregate level, we only do it once. There is more information extracted from firm-level to industry level, by aggregating firms' production activities than firms' productivities. Moving now to average productivity growth (APG

[3]To be more precise, AP will be different if computed using outputs as weights and using TFP with multiple inputs. It would not change in the case of a single input.

henceforth), one can choose the weights of either the initial period or of the final one or some average value. Also in this case the choice will matter and different weights may produce significantly different decompositions. For example, Haltiwanger (1997) argues that, since Baily et al. (1992) use the productivities of the final period in computing the "share effect", this effect does actually capture both the between effect and the covariance term.

Finally, when we have

As a result, it seems more reasonable to define AP in the aggregate level instead of following definition (1). Then the next question is whether without defining the weighted-average form for AP, we can still perform a decomposition of its variations into a within and a between component? Contrary to what is usually believed, in the next session we show that the answer to this question is yes.

3 Productivity growth in firms and industries: a unified framework

In this section we propose our decomposition method of AP and APG. We start with an empirical representation of an industry as a set of heterogeneous firms first introduced by Hildenbrand (1981) and later developed by Dosi et al. (2016). Such a representation does not assume the existence of a production function, but nevertheless allows to compute rigorous aggregate measures of productivity and, as we show below, decompositions of their variation which preserve coherence when passing from individual firms to industry aggregates. We first introduce some notation and definitions in subsection 3.1. Then, in subsection 3.2, we propose our measure of productivity both for individual firms and for an industry and show how the latter can be decomposed into a "within" and a "between" effect. In section 3.3, we further discuss the role of heterogeneity among firms in APG. In section 3.4, we extend our proposed decomposition method by taking into account firm entry and exit and, finally, in section 3.5 we illustrate our methodology with some toy examples, before presenting, in section 4, an empirical application to real data.

3.1 Notation and definitions

Following Koopmans (1977) and Hildenbrand (1981) we represent the actual technique of a production unit i by means of the vector of its *production activity*:

$$a_i^t = \left(\alpha_{i,1}^t, \cdots, \alpha_{i,(l-1)}^t, \alpha_{i,l}^t\right) \in \mathbb{R}_+^l, \tag{4}$$

where $\alpha_{i,l}^t$ is the output in period t and $\left(\alpha_{i,1}^t, \cdots, \alpha_{i,(l-1)}^t\right)$ is the vector of inputs. In this section we will analyze the case of single-output activities. The extension to multiple outputs will be briefly discussed in appendix B. If I^t denotes the set of all production units within one industry at time t, the aggregate (industry) production activity can be defined as:

$$d^t = \left(\beta_1^t, \cdots, \beta_{l-1}, \beta_l^t\right) = \left(\sum_{i\in I^t} \alpha_{i,1}^t, \cdots, \sum_{i\in I^t} \alpha_{i,(l-1)}^t, \sum_{i\in I^t} \alpha_{i,l}^t\right) \in \mathbb{R}_+^l \quad , \tag{5}$$

i.e. the sum of all individual firm production activities in the industry.

The productivity p_i^t of the production unit i at time t can be measured as the tangent of the angle $\theta(a_i^t)$ that the vector a_i^t forms with the space of inputs (Dosi et al., 2016). To give the intuition behind this measure of productivity, let us consider the case of only one input. Clearly, the larger the angle that the vector representing a production activity forms with the input axis, and therefore the smaller the angle it forms with the output axis, the more productive is the activity. By extension to the case of multiple inputs we obtain the following productivity indicator:

$$p_i^t := tg\left(\theta(a_i^t)\right) = \frac{\alpha_{i,l}^t}{||pr\left(a_i^t\right)||} \tag{6}$$

where the map

$$\begin{aligned} pr : \mathbb{R}^l &\to \mathbb{R}^{l-1} \\ (x_1, \cdots, x_l) &\mapsto (x_1, \cdots, x_{l-1}) \end{aligned}$$

is the projection map on the space of inputs[4].

Similarly, we define the aggregate productivity AP of the industry at time t, denoted by P^t, as

$$P^t := tg\left(\theta(d^t)\right) = \frac{\beta^t}{||pr\left(d^t\right)||} \quad . \tag{7}$$

Notice that while $\beta^t = \sum_{i \in I^t} \alpha_{i,l}^t$ is the total output of the industry, in general $||pr\left(d^t\right)|| \neq \sum_{i \in I^t} ||pr\left(a_i^t\right)||$, the equality holding only either in the case of a unique input or, in the case of multiple inputs, only when all the vectors $pr(a_i^t)$ lie on the same line and therefore production activities are perfectly homogeneous and differ only in their scale. If instead techniques are heterogeneous and firms use different combinations of inputs, the inequality $||pr\left(d^t\right)|| \neq \sum_{i \in I^t} ||pr\left(a_i^t\right)||$ holds. This heterogeneity component is an important feature of our model and we will further discuss it later in the paper.

3.2 Decomposing APG into *within* and *between* Effects

AP, defined in equation (7), can be further written as a weighted average of individual productivities p_i^t, since

$$P^t = \frac{\beta^t}{||pr\left(d^t\right)||} = \frac{\sum_{i \in I^t} \alpha_{i,l}^t}{||pr\left(d^t\right)||} = \sum_{i \in I^t} \frac{||pr\left(a_i^t\right)||}{||pr\left(d^t\right)||} \frac{\alpha_{i,l}^t}{||pr\left(a_i^t\right)||}$$

from which, by (6), we get the decomposition

$$P^t = \sum_{i \in I^t} w_i^t p_i^t \tag{8}$$

where the weights

$$w_i^t := \frac{||pr\left(a_i^t\right)||}{||pr\left(d^t\right)||} \tag{9}$$

represent the input-based weights defined as the relative length of individual input vectors $||pr\left(a_i^t\right)||$ over industry input vector $||pr\left(d^t\right)||$. As already mentioned, the length of the

[4] This can be easily generalised to multi-output case simply considering a different projection map (see Dosi et al. (2016)).

industry input vector $||pr\left(d^t\right)||$ is not necessarily equal to the sum of the lengths of the individual input vectors $||pr\left(a_i^t\right)||$, thus $\sum_{i\in I^t} w_i^t$ is not necessarily equal to 1 except for the one input case or for perfectly homogeneous firms.

Equality (8) indicates that AP can be written as a weighted average of individual productivities p_i^t. However, it is important to stress a fundamental methodological difference between how we obtain this weighted average and the standard approaches that we briefly surveyed in section **??** above. In the latter, AP is **defined** and hence computed as the weighted average of individual productivities, no matter how such individual productivities are measured. In our framework instead it is defined and computed in exactly the same way as we compute individual productivities. Moreover, the measure of AP we propose is also straightforward generalization of the one-input-one-output case. When there is only one input, the industry input vector degenerates to one number and the tangent of the angle we use to measure AP, according to (7), becomes the quotient of the sum of all outputs divided by the sum of all inputs. Input based weights in (9), $w_i^t = \frac{\alpha_{i,1}^t}{\sum_{i\in I^t}\alpha_{i,1}^t}$ are nothing else than input shares with sum equal to 1.

Now we show that our measure of AP can indeed be decomposed and the standard effects outlined by the traditional literature can indeed be easily computed also in our framework. For the sake of simplicity we first introduce our decomposition method only for the set C of continuing firms, i.e. all those that are active both in period $t-1$ and in period t. Entry and exit will be introduced later in section 3.4.

Such continuing firms are described at time $t-1$ by the vector set $\{a_i^{t-1}\}_{i\in C} \in \mathbb{R}_+^l$ and at time t by the vector set $\{a_i^t\}_{i\in C} \in \mathbb{R}_+^l$. Let d^{t-1} and d^t, computed according to (5), represent the corresponding aggregates at time $t-1$ and t respectively. Given all the production activity vectors, aggregate and individual productivity at $t-1$ and t can be easily computed according to (7) and (6) respectively.

We can now decompose APG, defined as the difference of AP between two consecutive years, into within-firm and between-firm components[5]:

$$\Delta P^t = \underbrace{\sum_{i\in C} \bar{w}_i \Delta p_i^t}_{Within} + \underbrace{\sum_{i\in C} \bar{p}_i \Delta w_i^t}_{Between} \;, \tag{10}$$

where Δ represents the variation from year $t-1$ to year t and $\bar{w}_i, \bar{p}_i$ are average of weights and productivities respectively.

In the above decomposition, the *within* term represents the contribution given to APG by the variations of the individual productivities and it is therefore similar to the "within" effect in the current literature. The *between* term present instead an important difference when compared with the "between" effect in the standard literature. In the latter, the weights w_i^t are defined either as input or output shares and, in both cases, $\sum_{i\in C} w_i^t = 1$. This is not the case in our decomposition (10) where, since the sum of the lengths of individual input vectors is not necessarily equal to the length of the sum of individual input vectors, i.e. the length of the industry input vector, we have, in general, $\sum_{i\in C} w_i^t \neq 1$. We will discuss this point in details in the next subsection and show that, actually, our *between* effect can be further decomposed.

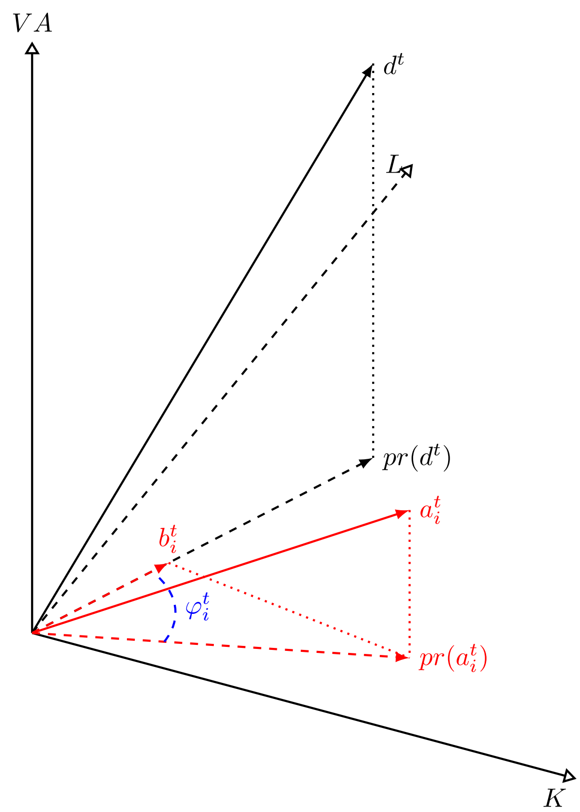
[5]All the mathematical details which lead to this decomposition can be found in appendix A.

3.3 Further Decomposing the *Between* Effect

In this section, we show that we can further decompose our *Between* term into two parts, an input weights component and a heterogeneity component.

Given the aggregate d^t and the individual a_i^t production activities, consider their projections on the input space and call them $pr\left(d^t\right)$ and $pr\left(a_i^t\right)$. Figure 1 provides a graphical representation for the special case with two inputs, capital K and labour L, and one output. Notice that if all individual production activities used the same proportion of inputs and differed only in scale and/or productivity, all the projection vectors $pr\left(d^t\right)$ and $pr\left(a_i^t\right)$ would overlap. On the other hand, the further away $pr\left(a_i^t\right)$ is from $pr\left(d^t\right)$, the more the combination of inputs used by firm i differs from the industry average combination. To measure this difference, we introduce φ_i^t which is the angle formed by the individual input vectors $pr\left(a_i^t\right)$ and the industry input vector $pr\left(d^t\right)$. Notice that given vectors d^t and a_i^t, and thus $pr\left(d^t\right)$ and $pr\left(a_i^t\right)$, it is easy to compute $\cos\varphi_i^t$ for each firm i at time t.

Figure 1: A graphical explanation for the decomposition of the between effect



In the input space we denote by b_i^t the projection of $pr\left(a_i^t\right)$ onto the industry projection vector $pr\left(d^t\right)$ (see Figure 1 for a graphical example). Notice that the length $||b_i^t||$ of the vector b_i^t can be regarded as the contribution of $pr\left(a_i^t\right)$ to the length of the industry input vector $pr\left(d^t\right)$. Hence, from now on, we will refer to the length $||pr\left(a_i^t\right)||$ of the firm input vector $pr\left(a_i^t\right)$ as the **actual input size** of firm i and to the length $||b_i^t||$ of the vector b_i^t as the **contributing input size** of firm i. It is easy to see that $||pr\left(d^t\right)|| = \sum_{i\in I^t} ||b_i^t||$.

Since

$$||pr\left(a_i^t\right)|| = \frac{||b_i^t||}{\cos\varphi_i^t} \quad ,$$

we can decompose w_i^t as

$$w_i^t = \frac{||pr\left(a_i^t\right)||}{||pr\left(d^t\right)||} = \frac{||b_i^t||}{\cos\varphi_i^t}\frac{1}{||pr\left(d^t\right)||} = \frac{||b_i^t||}{||pr\left(d^t\right)||}\frac{1}{\cos\varphi_i^t}$$

that is as the product:

$$w_i^t = s_i^t \cdot h_i^t \tag{11}$$

of what we could call the "**input weights**":

$$s_i^t = \frac{||b_i^t||}{||pr\left(d^t\right)||} \tag{12}$$

and a "**heterogeneity coefficient**":

$$h_i^t = \frac{1}{\cos\varphi_i^t} \ . \tag{13}$$

Equation (11) shows that the input based weights w_i^t in (9) mix together two different effects. The first one, s_i^t, represents the contribution of individual firms to the length of the industry input vector, i.e. our equivalent to the input weights in an multiple input case. Notice that in this case we have $\sum_{i\in C} s_i^t = 1$ and that, in the case of only one input s_i^t is the standard input share weight. The second one, that we named "heterogeneity coefficient" h_i^t, measures to which degree the individual input combinations are different from the industry average combination. The larger this difference, the bigger the angle φ_i^t and therefore the coefficient h_i^t. Thus, the sum $\sum_{i\in C} h_i^t$ can be regarded as an index of the heterogeneity of input combinations among productive units, and $\sum_{i\in C} \Delta h_i^t$ measures the variations of such heterogeneity.

Given these two effects, we can further decompose the *Between* term in (10) into two parts and thus refine our decomposition of APG as follows:[6]

$$\Delta P^t = \underbrace{\sum_{i\in C} \bar{w}_i \Delta p_i^t}_{Within} + \underbrace{\sum_{i\in C} \bar{p}_i \bar{h}_i \Delta s_i^t}_{Between^{is}} + \underbrace{\sum_{i\in C} \bar{p}_i \bar{s}_i \Delta h_i^t}_{Heterogneity} \tag{14}$$

where the *Within* term is as the same as the one in (10) and the sum of *Between*^{is}, i.e. the changes of the contributions given by individual firms to the industry total inputs, plus *Heterogeneity*, i.e. the changes in the heterogeneity among input combinations, is equal to the *Between* term in (10). It is easy to see that when there is only 1 input, h_i^t is always equal to 1 thus *Heterogeneity* goes to 0 and the two decompositions (14) and (10) coincide.

3.4 Decomposing APG with Entering and Exiting Firms

In this section we expand our decomposition of APG by accounting for the contributions given by production units which may enter or exit the industry during the period under

[6] Appendix A contains the mathematical details of the derivation of equation 14.

consideration. Let C be the set of all continuing firms, i.e. those that are active both in $t-1$ and in t, N the set of entering firms which are active in t but not in $t-1$, and X the set of exiting firms which are active in $t-1$ but not in t. Let vector sets $\{a_i^{t-1}\}_{i\in\{C\cup X\}} \in \mathbb{R}_+^l$ and $\{a_i^t\}_{i\in\{C\cup N\}} \in \mathbb{R}_+^l$ represent all firms active in the industry in $t-1$ and t respectively. According to equation (8) we have

$$\begin{aligned}\Delta P^t &= \sum_{i\in\{C\cup N\}} w_i^t p_i^t - \sum_{i\in\{C\cup X\}} w_i^{t-1} p_i^{t-1}\\ &= \sum_{i\in C}(w_i^t p_i^t - w_i^{t-1}p_i^{t-1}) + (\sum_{i\in N} w_i^t p_i^t - \sum_{i\in X} w_i^{t-1}p_i^{t-1})\end{aligned}$$

where for all the continuing firms, the term $\sum_{i\in C}(w_i^t p_i^t - w_i^{t-1}p_i^{t-1})$ can be further decomposed as in equation (14) and finally we have

$$\begin{aligned}\Delta P^t =& \underbrace{\sum_{i\in C}\bar{w}_i\Delta p_i^t}_{Within} + \underbrace{\sum_{i\in C}\bar{p}_i\bar{h}_i\Delta s_i^t}_{Between^{is}} + \underbrace{\sum_{i\in C}\bar{p}_i\bar{s}_i\Delta h_i^t}_{Heterogeneity}\\ &+ \underbrace{\sum_{i\in N} w_i^t p_i^t - \sum_{i\in X} w_i^{t-1}p_i^{t-1}}_{NetEntry}.\end{aligned} \tag{15}$$

3.5 Toy Example

Table 2 provide an illustrative toy example of an industry composed of 5 production units (firms) producing a unique output with two inputs, K and L. The first three columns report the input-output data in year 1. The fourth and fifth columns report the length $||pr\,(a_i^1)\,||$ of the input vectors and the length $||a_i^1||$ of the production activity vectors respectively. In the sixth column, we compute the productivities according to equation(6). Columns 7th to 12th report the same data referred to year 2. The only change taking place between the two years is an increase of the heterogeneity of the input combinations among firms, as visualized in Figure 2. The industry becomes more productive, as productivity increases from 0.5001 to 0.5294. Applying decomposition method in equation (14) this 0.0293 increase of productivity can be decomposed in the following way:

$$\underbrace{0.0293}_{APG} = \underbrace{0}_{Within} \underbrace{-0.0006}_{Between^{is}} + \underbrace{0.0299}_{Heterogeneity}.$$

which confirms that the within effect is null and the between effect is basically due to the increase of heterogeneity among firm, which is indeed the only phenomenon taking place between the two years. The percentages of *Between*^{is} and *Heterogeneity* over APG are -2.05% and 102.05% respectively and clearly the APG in this case is mostly driven by increasing in the inputs heterogeneity[7]. This contribution given by change in heterogeneity of inputs can be measured by our proposed decomposition method.

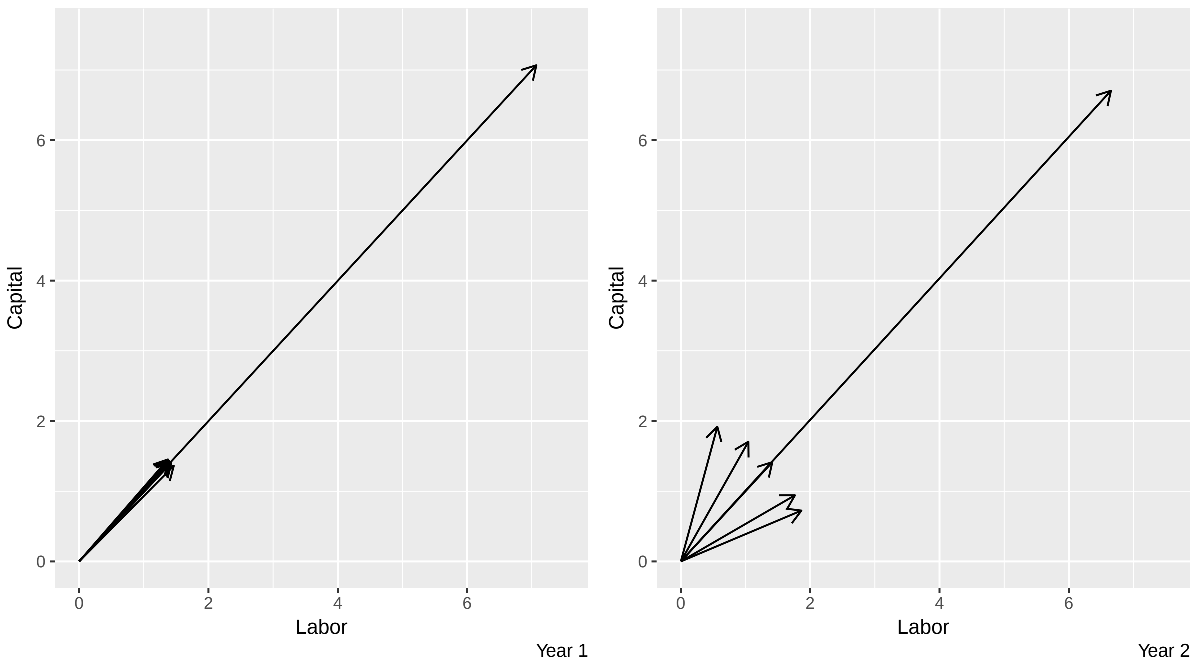
Let us now continue our toy example assuming that in year 3 everything remains unchanged from year 2 except that firm 3 doubles its output with the same inputs. The first six columns of Table 3 report these hypothetical data for year 3. Because of the

[7] Consistently, the heterogeneity measure introduced by Dosi et al. (2016) increases from 2.09025e-06 to 0.00728504.

Table 2: Toy Example with Five Firms and Increasing Heterogeneity

	Year 1						Year 2					
	$1st$ L	$2nd$ K	$3rd$ Output	$4th$ length of input vector	$5th$ length of vector	$6th$ $tg(\cdot)$	$7th$ L	$8th$ K	$9th$ Output	$10th$ length of input vector	$11th$ length of vector	$12th$ $tg(\cdot)$
Firm 1	1.414	1.414	1.000	2.000	2.236	0.500	1.414	1.414	1.000	2.000	2.236	0.500
Firm 2	1.464	1.362	1.000	2.000	2.236	0.500	1.764	0.942	1.000	2.000	2.236	0.500
Firm 3	1.424	1.404	1.000	2.000	2.236	0.500	1.864	0.724	1.000	2.000	2.236	0.500
Firm 4	1.374	1.453	1.000	2.000	2.236	0.500	1.044	1.706	1.000	2.000	2.236	0.500
Firm 5	1.394	1.434	1.000	2.000	2.236	0.500	0.564	1.919	1.000	2.000	2.236	0.500
Industry	7.071	7.068	5.000	9.998	11.178	0.5001	6.651	6.705	5.000	9.444	10.686	0.529

Figure 2: Toy Example with Five Firms and Increasing Heterogeneity



increase of firm 3's productivity, AP increases from 0.5294 to 0.6353 and this increase can be decomposed as:

$$\underbrace{0.1059}_{APG} = \underbrace{0.1059}_{Within} + \underbrace{0}_{Between^{is}} + \underbrace{0}_{Heterogeneity}$$

which indicates that APG is completely driven by the technical change operated by Firm 3, and therefore only the $Within$ term is different from zero. Finally, let us suppose that between year 3 and year 4 all firms hold their productivities and heterogeneity coefficients h_i^t constant and only the input shares weights s_i^t change (the last six columns in Table 3). APG is now totally imputed to variations of the input weights, i.e. to the $Between^{is}$ effect:

$$\underbrace{-0.0060}_{APG} = \underbrace{0}_{Within} \underbrace{-0.0060}_{Between^{is}} + \underbrace{0}_{Heterogeneity} .$$

These toy examples show that our measure correctly captures the phenomena driving

APG. The next section provides an empirical application to real data.

Table 3: Toy Example with Five Firms - Dynamic from Year 3 to Year 4

	Year 3							Year 4						
	L	K	Output	$\|\|pr(\cdot)\|\|$	$tg(\cdot)$	s_i^t	h_i^t	L	K	Output	$\|\|pr(\cdot)\|\|$	$tg(\cdot)$	s_i^t	h_i^t
Firm 1	1.414	1.414	1.000	2.000	0.500	0.212	1.000	1.494	1.494	1.057	2.113	0.500	0.224	1.000
Firm 2	1.764	0.942	1.000	2.000	0.500	0.202	1.046	1.766	0.943	1.001	2.002	0.500	0.203	1.046
Firm 3	1.864	0.724	2.000	2.000	1.000	0.193	1.095	1.777	0.690	1.906	1.906	1.000	0.184	1.095
Firm 4	1.044	1.706	1.000	2.000	0.500	0.206	1.028	1.081	1.766	1.035	2.071	0.500	0.213	1.028
Firm 5	0.564	1.919	1.000	2.000	0.500	0.186	1.137	0.533	1.811	0.944	1.888	0.500	0.176	1.137
Industry	6.651	6.705	6.000	9.444	0.635	1.000	1.000	6.651	6.705	5.943	9.444	0.629	1.000	1.000

4 An Empirical Application

4.1 Data and methodology

In order to show that our methodology is also empirically relevant, in this section we apply our APG decomposition to real firm-level data from **AMADEUS**, a commercial database provided by Bureau van Dijk. The edition at our access (October 2015) contains balance sheets and income statements for over 21 million European firms over the period 2004-2013. We focus on firms from several industries at the 4-digit NACE classification for three European countries, namely France, UK and Italy. These industries have been randomly selected from those with at least 20 firms (including continuous, exiting, and entering firms) during the time period under investigation for all the three countries. For reasons of space we report here only the results for seven selected industries, listed in Table 4. Results for other industries are available upon request. Number of employees and fixed assets are chosen as proxies for the two inputs labour and capital and turnover as a proxy for output. In some cases we add a third input, proxied by material costs. All of these values, except for the number of employees, are measured in thousands Euros and expressed in 2010 prices using the appropriate deflator for the 4-digit industry and the country under consideration[8].

Table 4: List of Selected Industries

NACE	Name of Industry
2014	Manufacture of other organic basic chemicals
2120	Manufacture of pharmaceutical preparations
2593	Manufacture of wire products, chain and springs
2630	Manufacture of communication equipment
2712	Manufacture of electricity distribution and control apparatus
2813	Manufacture of other pumps and compressors
2920	Manufacture of bodies (coachwork) for motor vehicles; manufacture of trailers and semi-trailers

We compute APG and its decomposition for two time period: between 2004 and 2007 and between 2010 and 2013. We compute four-years rather than yearly variations because the latter would mostly be very small. Moreover, we omit from consideration 2008 and 2009, the two years when the economic crisis hit more harshly the countries under consideration. Indeed an analysis of APG in these years is worthwhile, but since our exercise here is mainly for illustration purposes, we prefer not to include two years in which very abrupt variations of productivity are mainly driven by plunging output and high exit rates.

Table 5 shows APG and its decomposition for the period 2004-07 in the left panel and 2010-13 in the right panel. In the left panel, column 1 shows the value of APG, columns 2-4 the contributions to APG given by, respectively, Entering, Exiting and Continuing firms. The latter is further decomposed in a $Within$ and $Between$ effect in column 5. The $Between$ effect of column 5 is further decomposed in a $Between^{is}$ and a $Heterogenenity$ effects computed as in (14) and reported in column 6. Finally, in column 7 we report an

[8]Deflators for 4-digit industries are provided by *Eurostat* (`https://ec.europa.eu/eurostat/data/database`). When the 4-digit deflators for a specific country are not available, more aggregate deflators, e.g. 3-digit or 2-digit deflators for that country, are adopted.

heterogeneity coefficient coherent with our methodology and introduced by Dosi et al. (2016).[9]

Columns 8 to 14 report the same results for the time span 2010-13.

The table contains some interesting results. For instance, for what concerns industry "*2014*" from 2004 to 2007, we observe positive $Within$ and negative *Between* effects in all three countries. The former indicates that relatively big firms increased their productivity, while the latter signals that their contribution to the size of the industry (measured in input) has decreased. An opposite behaviour can be instead found, for instance, in industry "*2120*" in UK and Italy, where firms became less productive but increased their contribution to the total size of the industry.

4.2 Comparing Decomposition Results from Several Decomposition Methods

In this section, we compare our methodology with the standard one on empirical data. For the standard benchmark we adopt the method proposed by Grilliches and Regev (1995) and summarized in equation (3) above which is the one more directly comparable to ours. In order to compute AP defined as the weighted average (1), we need measures of individual productivities, which are usually proxied by TFP. To estimate TFP, according to Levinsohn and Petrin (2003)[10] , we need one more variable to serve as the *proxy* variable necessary in their method. We add in material cost and assume 3-input-1-output activity. Notice that by doing this, we lose all the firms from UK since in our original dataset, there is no observation for the material cost for the firms from UK. Also we have to drop the firms, from France and Italy, with observations for number of employees and fixed assets but not material cost. Using number of employees as the proxy for the weight s_i^t, we compute APGs from 2004 to 2007 for each country/industry cell and report them in column (1) of Table **??**. Decomposition results, by following (3), can be found in column 2. Similar APGs and their decomposition but using output as the proxy for the weight s_i^t can be found in column 3 and 4 respectively in the same table. In column 5, we report the APG, where AP follows (7), from 2004 to 2007 and the contribution from continuing firms, i.e. $Within$ and *Between* are reported vertically in column 6. The decomposition of *Between* can be found vertically in column 7. Similar to Table 5, industry heterogeneity change proxied by the percentage change of *Gini coefficient* is reported in column 8. From column 9 to 16, we report similar decomposition results of different methods based on the dynamics from 2010 to 2013.

[9]The coefficient is a Gini volume coefficient which is the ration between the volume of the zonotope formed by the actual production activities and the volume of the zonotope of an industry with the same size but maximum heterogeneity. For more details see Dosi et al. (2016) p. 885.

[10]We use **Stata** command **levpet** (Petrin et al., 2004) to estimate TFP with number of employees as *free* variable, fixed assets as *capital* variable, and material cost as *proxy* variable respectively. Dependent value is proxied by revenue not value added. "grid" search is set. Other options for the command follow the default value.

Table 5: APG Decompositions for Selected Industries in France, Italy and UK, 2004-07 and 2010-13

		From 2004 to 2007							From 2010 to 2013						
NACE	Ctry.	APG	Enter	-Exit	Continue	*Within* *Btw.*	*Btw.^{is}* *Heterogeneity*	*Gini* Growth (%)	APG	Enter	-Exit	Continue	*Within* *Btw.*	*Btw.^{is}* *Heterogeneity*	*Gini* Growth (%)
		1	2	3	4	5	6	7	8	9	10	11	12	13	14
2014	FR	1.60	0.31	-0.34	1.63	2.57 -0.94	-3.27 2.34	42.24	-0.61	0.67	-0.20	-1.08	-0.26 -0.81	-0.81 0	-40.92
2014	UK	0.31	0.36	-0.00	-0.05	5.46 -5.51	-10.58 5.07	37.99	0.30	0.32	-0.17	0.15	1.28 -1.12	-1.13 0.01	-10.47
2014	IT	-0.00	0.40	-0.01	-0.40	0.49 -0.89	-0.89 0	98.76	1.97	0.49	-0.14	1.63	-0.46 2.09	2.09 0	-3.70
2120	FR	0.89	0.25	-0.47	1.11	0.01 1.1	1.25 -0.14	-7.88	-0.32	0.41	-0.47	-0.25	-0.04 -0.21	-0.09 -0.12	-7.69
2120	UK	0.47	0.36	-0.01	0.12	-0.23 0.35	0.36 -0.01	2.98	-1.15	0.90	-0.19	-1.86	0.19 -2.05	-2.1 0.05	75.09
2120	IT	-0.12	0.17	-0.24	-0.04	-0.49 0.44	0.44 0	3.37	-0.27	0.81	-0.37	-0.72	-0.48 -0.24	-0.2 -0.04	-10.22
2593	FR	1.25	0.71	-1.11	1.65	0.54 1.11	1.09 0.02	-8.46	4.54	4.89	-2.74	2.39	0.17 2.22	2.23 -0.01	-6.48
2593	UK	0.45	1.13	-0.44	-0.24	0.69 -0.93	-0.96 0.02	36.36	-0.51	0.52	-0.15	-0.87	2.15 -3.02	-4.64 1.62	-17.64
2593	IT	0.41	0.86	-0.07	-0.38	0.14 -0.52	-0.52 0	21.33	-0.00	0.57	-0.18	-0.39	-0.04 -0.36	-0.36 0	10.44
2630	FR	1.96	2.64	-3.30	2.62	-1.93 4.56	4.64 -0.08	165.30	-8.18	2.90	-6.47	-4.61	-0.49 -4.12	-4.13 0.01	-56.58
2630	UK	-1.07	1.40	-0.25	-2.21	0.66 -2.87	-7.54 4.67	-15.48	-0.47	0.21	-0.26	-0.42	-0.61 0.18	0.11 0.07	-30.54
2630	IT	-2.26	0.88	-0.47	-2.67	-2.03 -0.64	1.31 -1.94	53.77	-1.71	0.17	-0.76	-1.12	-33.76 32.65	38.31 -5.66	-29.79
2712	FR	0.10	3.00	-1.09	-1.81	-0.01 -1.79	-1.78 -0.01	-23.71	-1.87	1.74	-1.70	-1.91	-0.03 -1.88	-1.88 0	102.80
2712	UK	6.44	5.63	-0.19	1.00	-3.63 4.63	5.84 -1.21	279.46	-5.40	0.20	-5.16	-0.44	-1.54 1.1	1.09 0.01	-76.85
2712	IT	0.58	0.17	-0.01	0.42	0.42 0	0 0	-11.67	-0.09	0.41	-0.04	-0.45	-1.52 1.07	1.08 -0.01	-25.45
2813	FR	-0.70	0.37	-2.00	0.93	-0.01 0.95	0.95 0	34.75	0.59	1.81	-2.28	1.05	-0.44 1.49	1.49 -0.01	-28.72
2813	UK	-2.67	0.47	-0.03	-3.12	0.09 -3.21	-3.26 0.05	109.51	0.20	0.06	-0.03	0.17	-0.15 0.31	0.32 0	-24.87
2813	IT	0.57	1.35	-0.14	-0.63	0.52 -1.16	-1.16 0	56.87	-0.45	0.76	-0.55	-0.66	-1.17 0.51	1.32 -0.8	1.08
2920	FR	1.00	2.12	-2.56	1.44	1.04 0.4	0.87 -0.46	6.02	-0.17	2.30	-4.67	2.20	-0.04 2.23	2.24 -0.01	-6.80
2920	UK	0.16	0.95	-0.26	-0.52	0.47 -0.99	-0.99 0	14.39	1.42	0.40	-0.27	1.29	1.26 0.02	0.03 -0.01	-2.67
2920	IT	1.00	1.29	-1.02	0.72	1.49 -0.77	-1 0.23	7.49	0.11	0.85	-0.10	-0.64	0.46 -1.09	-1.08 -0.01	-44.96

Table 6: The APG Decomposition Results of Different Decomposition Methods (An Alternative Structure)

		From 2004 to 2007											From 2010 to 2013										
Decomposition Method		Decomposition (3)						Proposed Method (15)					Decomposition (3)						Proposed Method (15)				
NACE	Ctry.	APG	*Enter* / *-Exit*	*Within* / *Btw*	APG	*Enter* / *-Exit*	*Within* / *Btw*	APG	*Enter* / *-Exit*	*Within* / *Btw.*	*Btw.^{is}* / *Heterogeneity*	*Gini Growth* (%)	APG	*Enter* / *-Exit*	*Within* / *Btw*	APG	*Enter* / *-Exit*	*Within* / *Btw*	APG	*Enter* / *-Exit*	*Within* / *Btw.*	*Btw.^{is}* / *Heterogeneity*	*Gini Growth* (%)
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
2014	FR	1.77	1.05	1.8	2.39	1.05	4.59	0.04	0.13	0.43	-0.4	-9.82	-3.35	4.37	-1.37	-12.05	2.56	-5.07	-0.18	0.36	-0.1	-0.32	-58.54
			-2.13	1.05		-2.73	-0.53		-0.27	-0.26	0.14			-1.5	-4.85		-0.69	-8.84		-0.1	-0.34	-0.02	
2014	IT	0.67	1.1	0.43	0.60	1.04	0.34	0.05	0.32	0.17	-0.5	291.02	-0.25	1.75	-0.49	0.02	0.91	-0.02	0.39	0.22	0.02	0.44	-25.29
			-0.03	-0.84		-0.03	-0.76		-0.01	-0.44	0.06			-1.22	-0.28		-0.7	-0.15		-0.11	0.27	-0.17	
2120	FR	0.59	3.08	0.61	3.25	1.41	0.87	0.27	0.17	-0.16	0.63	-10.34	2.01	3.9	0.53	-3.42	3.55	1.05	-0.07	0.22	-0.02	-0.03	-22.48
			-3.38	0.27		-3.29	4.27		-0.35	0.61	-0.02			-3.13	0.7		-8.53	0.52		-0.27	-0.01	0.03	
2120	IT	0.52	1.34	0.47	1.12	1.02	0.96	0.02	0.1	0.07	0.1	15.61	-0.24	7.74	-0.05	0.17	6.59	0.31	0.05	0.58	-0.08	-0.26	-10.96
			-2.84	1.54		-2.16	1.3		-0.21	0.08	-0.03			-0.87	-7.06		-2.8	-3.92		-0.21	-0.24	0.01	
2593	FR	-0.23	2.53	-0.13	-0.09	1.89	0.22	-0.07	0.21	0.02	0.17	-26.31	-1.21	8.75	-0.26	-0.73	9.46	-0.1	-0.15	0.97	-0.02	0.2	-35.31
			-3.99	1.36		-4.05	1.86		-0.4	0.11	-0.06			-12.06	2.36		-12.19	2.1		-1.28	0.19	-0.01	
2593	IT	0.05	0.36	-0.05	0.06	0.32	-0.04	-0.01	0.45	-0.08	-0.34	10.30	0.25	0.54	-0.04	0.16	0.38	-0.03	0.06	0.4	-0.02	-0.21	47.68
			-0.03	-0.23		-0.03	-0.19		-0.04	-0.33	0			-0.07	-0.18		-0.06	-0.12		-0.12	-0.2	0.01	
2630	FR	1.65	3.81	0.06	3.51	4.59	1.02	-0.36	0.53	-0.25	0.7	519.39	5.33	16.92	-0.12	5.68	21.54	-0.54	-0.23	1.9	-0.04	-0.67	-73.20
			-8.2	5.98		-7.79	5.69		-1.26	0.61	-0.09			-11.04	-0.44		-10.19	-5.14		-1.53	-0.57	0.11	
2630	IT	-1.25	9.62	-0.17	5.84	15.99	1.58	-0.00	0.5	0.04	-0.24	104.81	3.11	3.96	-44.3	1.05	5.09	-39.54	-0.38	0.17	-28.14	33.91	-14.13
			-5.66	-5.05		-5.95	-5.78		-0.3	-0.25	-0.01			-9.13	52.57		-14.08	49.57		-0.42	28.02	-5.89	
2712	FR	0.40	5.99	0.18	0.49	7.13	0.26	-0.07	1.04	-0.06	-0.67	-45.69	0.25	5.15	0.11	0.07	4.13	0.12	0.10	0.62	0.01	-0.09	22.45
			-3.02	-2.76		-2.68	-4.22		-0.4	-0.66	0.02			-3.68	-1.33		-3.15	-1.03		-0.44	-0.09	0	
2712	IT	0.48	2.41	0.43	0.53	1.36	0.37	0.35	0.14	0.2	0.05	24.67	-1.08	3.93	-1.59	-3.92	3.25	-2.43	-0.09	0.32	-0.14	-0.18	-55.92
			-0.18	-2.19		-0.14	-1.06		-0.01	0.02	-0.03			-0.62	-2.8		-0.34	-4.4		-0.03	-0.24	-0.07	
2813	FR	1.52	0.05	1.34	1.60	0.03	1.97	-0.19	0.13	-0.14	0.51	39.92	0.66	0.15	0.65	1.88	0.14	1.17	0.13	0.88	-0.04	0.47	-46.77
			-0.24	0.37		-0.23	-0.17		-0.69	0.52	0.01			-0.48	0.34		-0.3	0.87		-1.21	0.49	0.02	
2813	IT	1.38	7.68	0.3	2.17	9.13	0.68	0.02	0.51	-0.01	-0.41	92.80	0.08	7.79	-0.86	-1.89	7.97	-1.54	0.00	0.42	-0.12	-0.06	11.53
			-0.74	-5.86		-0.89	-6.76		-0.06	-0.42	-0.01			-3.18	-3.67		-5.81	-2.5		-0.27	-0.03	0.03	
2920	FR	10.90	11.93	7.29	14.08	11.2	9.64	0.02	0.34	0	0.14	24.57	3.73	12.52	1.53	-2.04	14.53	1.5	0.02	0.46	0.01	0.47	-2.34
			-12.61	4.29		-14.68	7.91		-0.44	0.13	-0.02			-24.6	14.28		-34.12	16.04		-0.92	0.46	0	
2920	IT	-4.83	14.62	-2.03	-25.87	14.82	-14.98	0.04	0.36	-0.05	0.07	19.42	6.20	14.95	1.16	-3.02	13.28	-0.64	0.08	0.41	0.04	-0.32	-45.42
			-11.4	-6.02		-22.53	-3.17		-0.33	0.07	0			-10.94	1.03		-1.27	-14.39		-0.05	-0.32	0	

5 Conclusions

Thanks to the increasing availability of longitudinal establishment- and firm- level data, a rapidly growing body of empirical literature has analyzed the relative importance between firm-level increase in productivity and the reallocation of market share to the aggregate productivity growth, i.e. so-called "within" and "between" effects, across individual producers within narrowly defined sectors. At the same time, the empirical evidence shows a highly significant and persistent degree of heterogeneity among firms and establishments in the input combinations and in their productivities even in the presence of the same relative input prices and in narrowly defined industries, thus with relatively homogeneous types of output.

Such heterogeneity poses serious challenges to the use of standard aggregate production functions and aggregate productivity one can derive from them. In this paper, building upon a geometric representation of the empirical production possibility set first suggest by Hildenbrand (1981) and developed by Dosi et al. (2016), we have introduced a new decomposition method for APG which 1) computes individual and aggregate productivity in the same way, instead of computing the latter as some arbitrary weighted average of the individual indicators; 2) reduces the loss of information implied by standard decomposition methods; 3) allows for a precise measure of the contribution given by variations in heterogeneity.

Our methodology can be applied to empirical data and the preliminary application we present in this paper, on some selected industries in France, Italy and the UK, show that indeed the contributions to APG that can be attributes to changes of firm-level heterogeneity are far from negligible.

References

Aw, B.-Y., S. Chen, and M. Roberts (1997): "Firm-level Evidence on Productivity Differentials, Turnover and Exports in Taiwanese Manufacturing," *NBER Working Paper 6235*.

Baily, M. N., C. Hulten, and D. Campbell (1992): "Productivity dynamics in manufacturing plants," *Brookings papers on economic activity (Microeconomics)*, 1992, 187–249.

Baldwin, J. R. and M. Rafiquzzaman (1995): "Selection versus evolutionary adaptation: Learning and post-entry performance," *International Journal of Industrial Organization*, 13, 501–522.

Bartelsman, E. J. and P. J. Dhrymes (1998): "Productivity dynamics: US manufacturing plants, 1972–1986," *Journal of productivity analysis*, 9, 5–34.

Bartelsman, E. J. and M. Doms (2000): "Understanding productivity: Lessons from longitudinal microdata," *Journal of Economic literature*, 38, 569–594.

Disney, R., J. Haskel, and Y. Heden (2003a): "Entry, exit and establishment survival in UK manufacturing," *The Journal of Industrial Economics*, 51, 91–112.

——— (2003b): "Restructuring and productivity growth in UK manufacturing," *The Economic Journal*, 113, 666–694.

Dosi, G. (2007): "Statistical regularities in the evolution of industries. a guide through some evidence and challenges for the theory," in *Perspectives on Innovation*, ed. by F. Malerba and S. Brusoni, Cambridge University Press, chap. 5.

Dosi, G., M. Grazzi, L. Marengo, and S. Settepanella (2016): "Production Theory: Accounting for Firm Heterogeneity and Technical Change," *The Journal of Industrial Economics*, 64, 875–907.

Dosi, G., D. Moschella, E. Pugliese, and F. Tamagni (2015): "Productivity, market selection, and corporate growth: comparative evidence across US and Europe," *Small Business Economics*, 45, 643–672.

Dwyer, D. W. (1998): "Technology locks, creative destruction, and nonconvergence in productivity levels," *Review of Economic Dynamics*, 1, 430–473.

Fisher, F. M. (1965): "Embodied technical change and the existence of an aggregate capital stock," *The Review of Economic Studies*, 32, 263–288.

Foster, L., J. C. Haltiwanger, and C. J. Krizan (2001): "Aggregate productivity growth: lessons from microeconomic evidence," in *New developments in productivity analysis*, University of Chicago Press, 303–372.

Griliches, Z. and J. Mairesse (1999): *Production functions: The search for identification*, Cambridge University Press: Cambridge.

Griliches, Z. and H. Regev (1995): "Productivity and firm turnover in Israeli industry," *Journal of Econometrics*, 65, 175–203.

Haltiwanger, J. (1997): "Measuring and analyzing aggregate fluctuations: the importance of building from microeconomic evidence," *Federal Reserve Bank of St. Louis Review*, 79, 55–78.

Hildenbrand, W. (1981): "Short-Run Production Functions Based on Microdata," *Econometrica*, 49, 1095–1125.

Hulten, C. R. (2001): "Total factor productivity: a short biography," in *New developments in productivity analysis*, University of Chicago Press, 1–54.

Koopmans, T. C. (1977): "5 Examples of Production Relations," *The microeconomic foundations of macroeconomics*, 144.

Leontief, W. W. (1986): *Input-output economics*, Oxford University Press on Demand.

Levinsohn, J. and A. Petrin (2003): "Estimating production functions using inputs to control for unobservables," *The Review of Economic Studies*, 70, 317–341.

Liu, L. and J. R. Tybout (1996): "Productivity growth in Chile and Colombia: the role of entry, exit, and learning," *Industrial Evolution in Developing Countries: Micro Patterns of Turnover, Productivity and Market Structure, New York: Oxford University Press for the World Bank*, 73–103.

Olley, G. S. and A. Pakes (1996): "The dynamics of productivity in the telecommunications equipment industry," *Econometrica*, 64, 1263–1297.

Petrin, A., B. P. Poi, and J. Levinsohn (2004): "Production function estimation in Stata using inputs to control for unobservables," *Stata Journal*, 4, 113–123.

Syverson, C. (2011): "What determines productivity?" *Journal of Economic literature*, 49, 326–365.

A Decomposition of aggregate productivity growth

In this Appendix we go through the mathematical details behind the decompositions in equations (10) and (14).

Decomposition in equation (10) By the decomposition of productivity as $P^t = \sum_{i \in C} w_i^t p_i^t$ described in equation (8) we get that

$$\Delta P^t = P^t - P^{t-1} = \sum_{i \in C} w_i^t p_i^t - \sum_{i \in C} w_i^{t-1} p_i^{t-1} \quad .$$

The above equality holds if we add and subtract the same quantity from its right side, that is

$$\Delta P^t = \sum_{i \in C} w_i^t p_i^t - \sum_{i \in C} w_i^{t-1} p_i^{t-1} + \sum_{i \in C} w_i^t p_i^{t-1} - \sum_{i \in C} w_i^t p_i^{t-1} + \sum_{i \in C} w_i^{t-1} p_i^t - \sum_{i \in C} w_i^{t-1} p_i^t$$

which we can re-write as

$$\Delta P^t = \sum_{i \in C} \frac{w_i^{t-1} + w_i^t}{2} (p_i^t - p_i^{t-1}) + \sum_{i \in C} \frac{p_i^{t-1} + p_i^t}{2} (w_i^t - w_i^{t-1})$$

which becomes equation (10)

$$\Delta P^t = \underbrace{\sum_{i \in C} \bar{w}_i \Delta p_i^t}_{Within} + \underbrace{\sum_{i \in C} \bar{p}_i \Delta w_i^t}_{Between}$$

with the notation $\bar{w}_i = \frac{w_i^{t-1} + w_i^t}{2}, \bar{p}_i = \frac{p_i^{t-1} + p_i^t}{2}$ and $\Delta p_i^t = p_i^t - p_i^{t-1}, \Delta w_i^t = w_i^t - w_i^{t-1}$.

Decomposition in equation (14) Following the decomposition:

$$\Delta P^t = \sum_{i \in C} \bar{w}_i \Delta p_i^t + \sum_{i \in C} \bar{p}_i \Delta w_i^t$$

in equation (10), we can further decompose the coefficient $\Delta w_i^t = w_i^t - w_i^{t-1}$ as follows. Since $w_i^t = s_i^t h_i^t$ (see equation (11)), we get equality

$$\Delta w_i^t = s_i^t h_i^t - s_i^{t-1} h_i^{t-1}$$

which can be modified by adding and subtracting the same quantity as follows

$$\Delta w_i^t = s_i^t h_i^t - s_i^{t-1} h_i^{t-1} + (s_i^t h_i^{t-1} - s_i^t h_i^{t-1}) + (s_i^{t-1} h_i^t - s_i^{t-1} h_i^t) \quad .$$

We can then re-write the right side of the equality as follows

$$\Delta w_i^t = \frac{h_i^{t-1} + h_i^t}{2} (s_i^t - s_i^{t-1}) + \frac{s_i^{t-1} + s_i^t}{2} (h_i^t - h_i^{t-1}) \quad .$$

If we denote by $\bar{h}_i = \frac{h_i^{t-1} + h_i^t}{2}$ and $\bar{s}_i = \frac{s_i^{t-1} + s_i^t}{2}$ the average sums and by $\Delta s_i^t = s_i^t - s_i^{t-1}$ and $\Delta h_i^t = h_i^t - h_i^{t-1}$ the variations then Δw_i^t becomes

$$\Delta w_i^t = \bar{h}_i \Delta s_i^t + \bar{s}_i \Delta h_i^t$$

which, substituted in equation (10), gives

$$\Delta P^t = \underbrace{\sum_{i \in C} \bar{w}_i \Delta p_i^t}_{Within} + \underbrace{\sum_{i \in C} \bar{p}_i \bar{h}_i \Delta s_i^t}_{Between^{is}} + \underbrace{\sum_{i \in C} \bar{p}_i \bar{s}_i \Delta h_i^t}_{Heterogeneity}$$

that is equation (14).

B Decomposition of aggregate productivity with multiple outputs

In this appendix we give a quick description of the aggregate productivity and its decomposition in the case of more than one output.

Notation and definitions During the period t the production unit i, which is described by the vector

$$a_i^t = \left(\alpha_{i,1}^t, \cdots, \alpha_{i,m}^t, \alpha_{i,m+1}^t, \cdots \alpha_{i,m+n}^t\right) \in \mathbb{R}_+^{m+n} \quad , \tag{16}$$

produces $a_{i,out}^t = \left(\alpha_{i,m+1}^t, \cdots \alpha_{i,m+n}^t\right)$ units of n outputs by means of $a_{i,in}^t = \left(\alpha_{i,1}^t, \cdots, \alpha_{i,m}^t\right)$ units of m inputs. Denote by I^t the set of all firms within the industry at time t. Then the set of vectors $\{a_i^t\}_{i \in I^t} \in \mathbb{R}_+^{m+n}$ represents the production activities of all in the industry at time t. Thus the aggregate (industry) production activity d^t can be written as the sum of individual firm production activity:

$$\begin{aligned} d^t &= \left(\beta_1^t, \cdots, \beta_m^t, \beta_{m+1}^t, \cdots, \beta_{m+n}^t\right) \\ &= \left(\sum_{i \in I^t} \alpha_{i,1}^t, \cdots, \sum_{i \in I^t} \alpha_{i,m}^t, \sum_{i \in I^t} \alpha_{i,m+1}^t \cdots, \sum_{i \in I^t} \alpha_{i,m+n}^t\right) \in \mathbb{R}_+^{m+n} . \end{aligned} \tag{17}$$

If we denote by

$$\begin{aligned} pr_{out} : \mathbb{R}^{m+n} &\to \mathbb{R}^n \\ a_i^t &\mapsto a_{i,out}^t \end{aligned}$$

and by

$$\begin{aligned} pr_{in} : \mathbb{R}^{m+n} &\to \mathbb{R}^n \\ a_i^t &\mapsto a_{i,in}^t \end{aligned}$$

the projections of the production activities a_i^t (and similarly of the industry vector d^t) on the spaces of outputs and inputs respectively, then the formula to compute the industry and firms productivities become

$$P^t := tg\left(\theta(d^t)\right) = \frac{||pr_{out}\left(d^t\right)||}{||pr_{in}\left(d^t\right)||} \tag{18}$$

and

$$p_i^t := tg\left(\theta(a_i^t)\right) = \frac{||pr_{out}\left(a_i^t\right)||}{||pr_{in}\left(a_i^t\right)||} \tag{19}$$

respectively, where $\theta(.)$ denotes the angle of vectors d^t and a_i^t with the space of inputs. Notice that, since in this case output $a_{i,out}^t$ is a multidimensional vector then, in general, $||pr_{out}\left(d^t\right)|| \neq \sum_{i \in I^t} ||pr_{out}\left(a_i^t\right)||$, unless all output vectors are proportional or there is only one output. If we denote by φ_i^t the angle formed by the vectors $pr_{in}\left(a_i^t\right)$ and $pr_{in}\left(d^t\right)$ and by σ_i^t the angle formed by the vectors $pr_{out}\left(a_i^t\right)$ and $pr_{out}\left(d^t\right)$, we get that

$$||pr_{out}\left(d^t\right)|| = \sum_{i \in I^t} \left(||pr_{out}\left(a_i^t\right)|| \cos \sigma_i^t\right) \tag{20}$$

which, substituted in equation (18), gives

$$P^t = \frac{\sum_{i\in I^t}\left(||pr_{out}\left(a_i^t\right)||\cos\sigma_i^t\right)}{||pr_{in}\left(d^t\right)||} = \sum_{i\in I^t}\left(\cos\sigma_i^t \frac{||pr_{in}\left(a_i^t\right)||}{||pr_{in}\left(d^t\right)||}\frac{||pr_{out}\left(a_i^t\right)||}{||pr_{in}\left(a_i^t\right)||}\right) ,$$

that is

$$P^t = \sum_{i\in I^t} u_i^t p_i^t$$

where the "weight" coefficient

$$u_i^t := k_i^t \cdot w_i^t$$

is defined as the product of an "output homogeneity" measure

$$k_i^t := \cos\sigma_i^t$$

and the input-based-weight

$$w_i^t := \frac{||pr_{in}\left(a_i^t\right)||}{||pr_{in}\left(d^t\right)||} \quad .$$

Notice that k_i^t is a decreasing function of σ_i^t when $\sigma_i^t \in \left[0, \frac{\pi}{2}\right]$. That is smaller σ_i^t's correspond to larger k_i^t's and indicate that the vector $pr_{out}(a_i^t)$ is closer to the vector $pr_{out}(d^t)$, i.e. less "output heterogeneity". The fact that more output-based-homogeneity coincides with bigger k_i^t explains why we name k_i^t as output homogeneity measure.

Decomposing the aggregate industry growth Let us denote by C, N, and X the sets of continuing, entering, and exiting firms respectively. The aggregate productivity growth from time $t-1$ to time t is given by

$$\Delta P^t = \sum_{i\in C}\bar{u}_i\Delta p_i^t + \sum_{i\in C}\bar{p}_i\Delta u_i^t + \sum_{i\in N}u_i^t p_i^t - \sum_{i\in X}u_i^{t-1}p_i^{t-1} . \tag{21}$$

where for any variable $x^t \in \mathbb{R}$ at time t, operator Δ represents its change from $t-1$ to t, i.e. $\Delta x^t \equiv x^t - x^{t-1}$, and $\bar{x} \equiv \frac{x^t + x^{t-1}}{2}$. We can further decompose Δu_i^t as

$$\begin{aligned}\Delta u_i^t &= \frac{k_i^{t-1}+k_i^t}{2}(w_i^t - w_i^{t-1}) + \frac{w_i^{t-1}+w_i^t}{2}(k_i^t - k_i^{t-1}) \\ &= \bar{k}_i\Delta w_i^t + \bar{w}_i\Delta k_i^t \\ &= \bar{k}_i\left(\bar{h}_i\Delta s_i^t + \bar{s}_i\Delta h_i^t\right) + \bar{w}_i\Delta k_i^t \end{aligned} \tag{22}$$

Finally, by substituting (22) into (21) we have

$$\Delta P^t = \underbrace{\sum_{i\in C}\bar{u}_i\Delta p_i^t}_{Within} + \underbrace{\sum_{i\in C}\bar{p}_i\bar{k}_i\bar{h}_i\Delta s_i^t}_{Between^{is}} + \underbrace{\sum_{i\in C}\bar{p}_i\bar{k}_i\bar{s}_i\Delta h_i^t}_{Heterogeneity} + \underbrace{\sum_{i\in C}\bar{p}_i\bar{w}_i\Delta k_i^t}_{Homo^{out}} + \underbrace{\sum_{i\in N}u_i^t p_i^t - \sum_{i\in X}u_i^{t-1}p_i^{t-1}}_{NetEntry} . \tag{23}$$

Notice that when $n=1$, i.e. there is only one output, the angle σ_i^t between the individual output vector and the aggregate output vector degenerates to 0. Thus for all firm i over all time t, we have

$$k_i^t = \cos\sigma_i^t = 1$$

and thus $\bar{k}_i = 1$, $\Delta k_i^t = 0$ and $u_i^t = w_i^t$. As a result, the decomposition (23) degenerates to (15).