Author(s): Hyytinen, Ari; Ilmakunnas, Pekka; Maliranta, Mika

Title: Olley–Pakes productivity decomposition: computation and inference

Year: 2016

Version:

Please cite the original version:

Hyytinen, A., Ilmakunnas, P., & Maliranta, M. (2016). Olley–Pakes productivity decomposition: computation and inference. Journal of the Royal Statistical Society Series A: Statistics in Society, 179(3), 749-761. https://doi.org/10.1111/rssa.12135

To appear: Journal of the Royal Statistical Society, Series A

Olley-Pakes productivity decomposition: Computation and inference

Ari Hyytinen†

University of Jyväskylä, School of Business and Economics, Jyväskylä, Finland.

Pekka Ilmakunnas

Aalto University School of Economics, Helsinki, Finland.

Mika Maliranta

University of Jyväskylä and The Research Institute of the Finnish Economy (ETLA), Helsinki, Finland.

Summary. We show how a moment-based estimation procedure can be used to compute point estimates and standard errors for the two components of the widely used Olley-Pakes decomposition of aggregate (weighted average) productivity. When applied to business-level microdata, the procedure allows for autocovariance and heteroskedasticity robust inference and hypothesis testing about, e.g. the co-evolution of the productivity components in different groups of firms. We provide an application to Finnish firm-level data and find that formal statistical inference casts doubt on the conclusions that one might draw based on a visual inspection of the components of the decomposition.

Keywords: **inference, productivity, weighted average, generalized method of moments**
JEL classification: C10, 047

1. Introduction

Productivity growth of industries is a key driver of long-term economic growth. Starting with Bailey *et al.* (1992), a number of papers have explored how various decompositions of industry productivity growth are able to capture its key microeconomic sources, such as the reallocation of resources between firms and plants. In this context, the cross-sectional efficiency of resource allocation has attracted increased attention (e.g. Eslava *et al.* 2004, 2010, 2013, Chari 2011; see also Van Biesebroeck 2008). One reason for this is that this approach seems to avoid at least some of the problems hampering dynamic productivity decompositions (Bartelsman *et al.* 2013, Eslava *et al.* 2004, 2010). The measure that the papers in this strand of the literature typically use is the static decomposition introduced by Olley and Pakes (1996). The Olley-Pakes (OP) decomposition shows how industry productivity, which equals the weighted average of firm-level (or plant-level) productivity, can be decomposed into

†*Address for correspondence*: Ari Hyytinen, University of Jyväskylä, 40014 University of Jyväskylä, Finland.
E-mail: ari.hyytinen jyu.fi

(i) the unweighted average of the productivity of firms and (ii) a covariance-like term between activity (i.e. output or input) shares and productivity.

The within-industry covariance term between firm size and productivity is of particular economic and policy interest, because the smaller this term is, the smaller the share of activity (or resources) that gets allocated to the most productive firms. Olley and Pakes, for example, argued that the deregulation of the U.S. telecommunications equipment industry may have increased the covariance term by increasing the allocation of resources to the most productive firms. Bartlesman *et al.* (2009, 2013) argued, in turn, that a low covariance term is a good indicator of misallocation of resources and policy-induced market distortions; they provide evidence that variation in the covariance term explains an important fraction of the cross-country differences in productivity. Melitz and Polanec (2015) have recently proposed a dynamic extension of the OP decomposition.

This paper builds on the observation that hypothesis testing and inference appear to be a neglected part of the decomposition literature. In particular, the estimation of the standard errors of the two components of the OP decomposition has not, despite its increasing popularity in applied literature, received attention. While there are papers that contain statistical inference on *some* components of the decompositions (e.g. on the unweighted mean component of continuing firms or the entry and exit components), there are, to the best of our knowledge, few papers that present standard errors for all components of a decomposition simultaneously or address inference on the policy-relevant covariance-like terms. An example is Foster, Haltiwanger and Krizan (2006) who regressed productivity on indicators of entry and exit, obtaining a regression analogue to a decomposition of productivity growth to entry and exit effects and growth in continuing firms.

The aim of this paper is to start filling this gap in the literature. We outline a moment-based procedure to the estimation of the OP components and their standard errors and illustrate how a simple two-step recipe can be used for inference and hypothesis testing in applications.

It seems obvious that it would be worthwhile to have a procedure for testing hypotheses about the magnitude of the OP components as they are, like any other statistical estimate, a function of the available microdata and thus subject to sampling error. In applications, it may, for example, be of interest to plot the confidence intervals for the covariance term or to test the joint hypotheses about the temporal evolution of the OP components.

Many of earlier studies on the productivity decompositions have been based on large census-based data sets. These data sets may cover (nearly) the entire population of firms of a country. We can think of such data as a sample drawn from a larger super-population or an underlying stochastic data-generating process (see, e.g. Davidson and MacKinnon 2004, Berk *et al.* 1995, and Gelman and King 1994). An emerging tendency in the literature is towards using alternative data sources and finer decompositions (by, e.g. firm cohort, geographic region or firm type) in the hope that they help us to better understand the microeconomic determinants of industry productivity. This tendency is likely to lead to denser slicing of the available microdata, strenghening thus the need for appropriate inference procedures. One can envisage that, in the near future, researchers increasingly combine survey-based measures to the large census-based data sets to analyze, for example, how certain firm attributes, such as quality of management, organizational choices, firms' innovativeness or research and development efforts, affect the productivity-enhancing reallocation within and between industries. When such combined data are used, the need for explicitly measuring sampling error becomes stronger.

We illustrate our estimation procedure by applying it to Finnish firm-level panel data from 1995 to 2007. The application is a case study, which for brevity focuses on a single industry and cross-regional differences in productivity. We find that formal statistical inference casts some doubt on the conclusions that one might draw based on a visual inspection of the two components of OP decomposition. It also turns out that the covariance term, which often is of major policy interest, cannot be measured as accurately as the unweighted average of the productivity of firms.

We believe that the method of inference proposed in this paper should be useful in addressing substantive research questions in economics as well as in other disciplines. For example, as the rapidly growing applied literature that uses the OP decomposition shows, understanding the drivers of industry growth is clearly of major policy interest. Our method can be used to formally test how certain public policies affect the reallocation of resources within industries and thereby economic growth. Our method also has implications for statistical testing of certain aggregation procedures that have been used in previous studies.

The remainder of this paper is organized as follows: The next section presents OP decomposition. In the third section, we develop the moment-based procedure using insights from the Generalized Method of Moments (GMM) estimation. The fourth section describes the illustration that uses Finnish firm-level data. Section five concludes by discussing the scope of our contribution and areas of empirical research where our method could be useful.

2. Olley-Pakes Decomposition

We start by writing a formal expression for the OP decomposition. To this end, let s_{it} denote the activity share of firm i in period t, and φ_{it} an index of productivity. How s_{it} and φ_{it} are measured depends on the application as the decomposition can be applied either to an industry-level index of total factor productivity (TFP) or to that of labour productivity, be done in levels or in log-units and computed using either input or output shares. When input shares are used, there is a direct link to the standard measures of aggregate industry productivity (e.g. van Biesebroeck 2003 and Fox 2012).

For concreteness, we frame our analysis in terms of labour productivity. We also assume that the index of firm-level productivity is measured in log-units, i.e. $\varphi_{it} = \log(Y_{it}/L_{it})$, where Y_{it} is a measure of value added and L_{it} is the number of employees in firm i at time t. The activity shares are measured by labour inputs so that $s_{it} = L_{it}/\sum_{i=1}^{N_t} L_{it}$, where N_t refers to the number of firms in period t.

Taking a single cross-section of the microdata for period t, the OP decomposition of the aggregate productivity index of an industry is

$$\Phi_t = \overline{\varphi}_t + \sum\nolimits_{i=1}^{N_t} (s_{it} - \overline{s}_t)(\varphi_{it} - \overline{\varphi}_t), \tag{1}$$

where $\Phi_t = \sum_{i=1}^{N_t} s_{it}\varphi_{it}$ is the weighted mean of firm-level productivity, $\overline{\varphi}_t$ refers to the unweighted mean and the last term is the covariance term. The covariance term consists of the deviations of input shares around their unweighted cross-sectional mean $(s_{it} - \overline{s}_t)$ where $\overline{s}_t = 1/N_t$, and the deviations of firm productivity around their unweighted, cross-sectional mean, $(\varphi_{it} - \overline{\varphi}_t)$. To distinguish the second term of (1) from the standard sample (cross-sectional) covariance $\widehat{cov}_t(s_{it}, \varphi_{it})$, we denote $\widehat{c}_t \equiv \sum_{i=1}^{N_t}(s_{it} - \overline{s}_t)(\varphi_{it} - \overline{\varphi}_t)$ in what follows. This means that $\widehat{c}_t = \widehat{cov}_t(s_{it}, \varphi_{it}) \times N_t$.

A large part of the subsequent development in this paper is motivated by the simple observation that the two terms of (1) can be estimated jointly by regressing φ_{it} on a constant and an appropriately scaled s_{it} using Ordinary Least Squares (OLS). This insight relies on the anatomy of the regression, expressed in terms of population moments $E\left[\varphi_{it}\,|s_{it}\right] = E\left[\varphi_{it}\right] + cov\left(\varphi_{it}, s_{it}\right)\left[var\left(s_{it}\right)\right]^{-1}\left(s_{it} - E\left[s_{it}\right]\right)$. This expression immediately suggests that the two components can be captured by a single moment condition using a standard GMM procedure that we outline below.

3. Computation and Inference

3.1. Moment-Based Approach

We assume for simplicity that the microdata used to compute the OP decomposition is a balanced panel and that the sample period has length T. The aim is to compute the $2T$ decomposition terms and to estimate the associated standard errors. The approach can be generalized to many kinds of unbalanced panels (see, e.g. Wooldridge 2002, Ch. 17 and the discussion below).

We let $(\overline{\varphi}', \widehat{\mathbf{c}}')' = (\overline{\varphi}_1, ..., \overline{\varphi}_T,\ \widehat{c}_1, ..., \widehat{c}_T)'$ be the vector of decomposition terms and define scaled s_{it} as $s^*_{it} = (s_{it} - \overline{s}_t)/\widehat{\sigma}^2_t N$, where $\widehat{\sigma}^2_t$ is the cross-sectional sample variance of s_{it}, i.e. $\widehat{\sigma}^2_t = N^{-1}\sum_{i=1}^{N}(s_{it} - \overline{s}_t)^2$ in period t (where for simplicity we use N instead of $N-1$ when computing sample variances). We also let $\mathbf{D}_i$ be a $(T \times T)$ period dummy matrix with typical element $d_{it,k}$, which is equal to one if $t = k$ and equal to zero otherwise. If we use the period dummy matrix and if we collect the scaled input share data s^*_{it} for firm i into a $(T \times T)$ diagonal matrix $\mathbf{S}^*_i = diag\left[s^*_{i1}, ..., s^*_{iT}\right]$, we can define a $(T \times 2T)$ data matrix $\mathbf{X}_i = \left[\mathbf{D}_i\ \mathbf{S}^*_i\right]$. We can now write a population moment condition

$$E\left[\mathbf{X}_i{}'\left(\varphi_i - \mathbf{X}_i\beta\right)\right] = \mathbf{0}_{(2T\times 1)} \tag{2}$$

where $\varphi_i = (\varphi_{i1}, ..., \varphi_{iT})'$ and $\beta = (\theta', \gamma')' = (\theta_1, ..., \theta_T, \gamma_1, ..., \gamma_T)'$ is a $(2T \times 1)$ parameter vector.

The analogy principle says that a suitable estimator for unknown population parameters can be found by considering the sample counterpart of a population moment (see, Manski 1988). Applying this principle to (2) results in a single-equation panel Generalized Method of Moments (GMM) estimator for $\beta = (\theta', \gamma')'$ (Cameron and Trivedi 2005, pp. 744-745).

In our case, the model is just-identified, as the number of instruments is equal to the number of parameters. The moment condition therefore results in the familiar pooled OLS estimator of a linear panel model. To derive it, note that the estimator solves $N^{-1}\sum_{i=1}^{N}\left[\mathbf{X}_i{}'\left(\varphi_i-\mathbf{X}_i\widehat{\beta}\right)\right] = \mathbf{0}$. Stacking all firms $\varphi' = (\varphi_1' \cdots \varphi_N')$ and $\mathbf{X}' = (\mathbf{X}_1' \cdots \mathbf{X}_N')$, the pooled OLS estimator is $\widehat{\beta} = (\widehat{\theta}', \widehat{\gamma}')' = \left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\varphi$.

The OLS estimator is numerically equivalent to the two components on the right hand side of (1), i.e. $(\overline{\varphi}', \widehat{\mathbf{c}}')'$. The result follows from the standard results on partitioned regression and from the fact that the model is completely saturated in terms of (orthogonal) period indicators. In particular, picking any $\widehat{\theta}_t$, one can show that $\widehat{\theta}_t = \overline{\varphi}_t - \widehat{\gamma}_t\overline{s}_t^* = \overline{\varphi}_t$. The last step follows from $\overline{s}_t^* = 0$. Similarly, picking any $\widehat{\gamma}_t$, one can establish that $\widehat{\gamma}_t = \left(\sum_{i=1}^{N}(s_{it}^*)(\varphi_{it}-\overline{\varphi}_t)\right)\left(\sum_{i=1}^{N}(s_{it}^*)^2\right)^{-1}$. This expression simplifies to $\widehat{cov}_t \times N = \sum_{i=1}^{N}(s_{it}-\overline{s}_t)(\varphi_{it}-\overline{\varphi}_t)$. It is thus equal to $\widehat{c}_t$ as desired.

To summarize, obtaining the point estimates for the two OP components from microdata consists of two steps: First, s_{it} is demeaned and scaled by N times its cross-sectional variance. Second, one regresses φ_{it} on a constant and the scaled s_{it}^* using OLS. The estimator for the constant gives $\overline{\varphi}_t$, and the estimated coefficient of the slope is $\widehat{c}_t$.

3.2. *Autocovariance and Heteroskedasticity Robust Statistical Inference*

The benefit of casting the estimation of the OP decomposition in terms of a population moment condition and GMM is twofold. First, the GMM framework can be used to establish the asymptotic properties of the estimators, their asymptotic normality in particular. Second, it provides a simple way to compute autocovariance and heteroskedasticity robust standard errors (see, e.g. Cameron and Trivedi 2005, Ch. 6 and 22 and Wooldridge 2002, Ch. 14).

To describe the autocovariance and heteroskedasticity robust estimator of the standard errors, we note that moment (2) implicitly defines a $(T\times 1)$ vector of regression errors $\mathbf{u}_i \equiv \varphi_i - \mathbf{X}_i\beta$ for each firm i. If it were the case that these errors were uncorrelated over time for a given firm and homoscedastic, using the classical OLS variance-covariance estimator would lead to an estimator for the standard errors of $\overline{\varphi}_t$ and $\widehat{\gamma}_t$ that are similar (but not identical) to the conventional standard error estimators for the sample mean and covariance. However, there are strong reasons to suspect that the errors are both correlated over time for a given firm and heteroscedastic. The former should be allowed for because shocks to the productivity of firm i are likely to be persistent over time. Heteroscedasticity is also expected in most microdata and should therefore be allowed for. For example, the cross-sectional variance of productivity shocks may vary over time, leading to a form of heteroscedasticity.

Assuming independence over i and $N \to \infty$, the autocovariance and heteroskedasticity robust estimate of the asymptotic variance matrix of the estimator is

$$\widehat{V}\left[\widehat{\beta}\right] = \left[\sum_{i=1}^{N}\mathbf{X}_i'\mathbf{X}_i\right]^{-1}\sum_{i=1}^{N}\mathbf{X}_i'\widehat{\mathbf{u}}_i\widehat{\mathbf{u}}_i'\mathbf{X}_i\left[\sum_{i=1}^{N}\mathbf{X}_i'\mathbf{X}_i\right]^{-1} \tag{3}$$

where $\widehat{\mathbf{u}}_i = \varphi_i - \mathbf{X}_i\widehat{\beta}$.

Estimator (3) can be used to obtain a consistent estimate of the asymptotic covariance matrix for $\widehat{\beta}$ that is robust to within-firm autocorrelation and heteroscedasticity of unknown form. This familiar variance estimator is a suitable choice when one is unwilling to make assumptions about the within-firm autocorrelation structure or the type of heteroscedasticity in the microdata. It is also suitable when the data are a short panel and thus have relatively few observations per each firm (small T) but include many firms (large N); see Cameron and Trivedi (2005, Ch. 22 and 24) and Wooldridge (2003) for further discussion.

The estimator is easy to implement because it can be computed using a standard OLS command with an option for cluster-robust standard errors. Some of the standard econometric software (such as *Stata*) make by default a small-sample correction when computing cluster-robust standard errors, but this correction can often be undone and, in any case, does not matter in large samples. Confidence intervals and joint tests follow from standard argumentation.

In applications, it may be of interest to plot the confidence intervals of the covariance term or to test joint hypotheses about the components of the productivity decomposition. For example, testing whether the covariance term has remained stable over the sample period is equivalent to testing H_0: $\gamma_1 = ... = \gamma_T$ using standard joint testing procedures, such as the Wald test. Similarly, testing for H_0: $\gamma_{T-s} = ... = \gamma_T = 0$ corresponds to analysing the null hypothesis that the industry index of productivity during the last s years of the sample period is no higher than it would have been if the input shares were randomly allocated within the industry. Finally, the null hypothesis of constant growth (rate) of the average firm productivity can be examined by testing H_0: $\theta_2 - \theta_1 = ... = \theta_T - \theta_{T-1}$.

3.3. Discussion and Extensions

3.3.1. Mutually Exclusive Subgroups

The first, perhaps most obvious, extension to the basic procedure builds on the observation that the aggregate productivity index for a group of firms can be computed as a weighted mean of the aggregate productivities of the sub-groups of firms. If one is willing to assume that the industry affiliation of firms is time invariant (which in reality may not be the case), this observation suggests that one can assign all the firms of an industry to mutually exclusive subgroups and estimate the productivity decompositions and the associated standard errors separately for each sub-group.

To illustrate how that could be done, we assume that there are J subgroups ($j = 1, ..., J$) and take the following four steps: First, we define sub-group indicator $q_{it,j}$, which is equal to one if firm i belongs to group j in period t and is zero otherwise. This implies that in each period, the number of members in sub-group j is $N_{t,j} = \sum_{i=1}^{N} q_{it,j}$. Second, we scale the input shares by period and within each sub-group to obtain $s^*_{it,j} = (s_{it} - \overline{s}_{t,j})(\widehat{\sigma}^2_{t,j} N_{t,j})^{-1}$, where $\overline{s}_{t,j}$ is the mean and $\widehat{\sigma}^2_{t,j}$ is the cross-sectional sample variance of the input share in sub-group j in period t. By definition, $s^*_{it,j}$ is zero for firm i

in period t if it does not belong to group j during the period. Third, we let $\otimes$ denote the Kronecker product and define $\widetilde{\mathbf{d}}'_{it} = (q_{it,1}, ..., q_{it,J}) \otimes (d_{it,1}, ..., d_{it,T})$ and $\widetilde{\mathbf{s}}^{*\prime}_{it} = \left(s^*_{it,1}, ..., s^*_{it,J}\right) \otimes (d_{it,1}, ..., d_{it,T})$, which are row vectors of length JT. Finally, we use population moment condition (2) and the GMM approach to estimate the two components of the OP decomposition for each sub-group by redefining matrix $\mathbf{X}_i$ so that its t^{th} row is now $\mathbf{x}'_{it} = \left[\widetilde{\mathbf{d}}'_{it}\ \widetilde{\mathbf{s}}^{*\prime}_{it}\right]$. Of course, $\beta = (\theta', \gamma')'$ has to be redefined accordingly, i.e. to be a (column) vector of length $2JT$.

This extension is of potential interest in applications. For example, to study whether the (relative) importance of the covariance terms in an industry is similar in J geographic regions in a given period, we could test H_0: $\gamma_{t,1} = ... = \gamma_{t,J}$ using a joint test. Implementing such a test is straightforward because in each row of $\mathbf{X}_i$ the first JT terms are group-specific period indicators (i.e. the complete set of period indicators interacted with the complete set of sub-group indicators) and the next JT terms are the period and sub-group-specific input shares $s^*_{it,j}$. We illustrate a variant of this test in our application.

It is important to recognize that $\widehat{\theta}_{t,1} + \widehat{\gamma}_{t,1} .... + \widehat{\theta}_{t,J} + \widehat{\gamma}_{t,J}$ is *not* equal to Φ_t, i.e. the weighted mean of firm-level productivity in period t. However, if these estimates are weighted by the employment share of each sub-group in period t, $S_{t,j} = \sum_{i=1}^{N} q_{it,j} L_{it} / \sum_{i=1}^{N} L_{it}$, they total to Φ_t. That is, $\Phi_t = S_{t,1}\left(\widehat{\theta}_{t,1} + \widehat{\gamma}_{t,1}\right) + ... + S_{t,J}\left(\widehat{\theta}_{t,J} + \widehat{\gamma}_{t,J}\right)$, where $\sum_{j=1}^{J} S_{t,j} = 1$.

3.3.2. Industry-Level Productivity

The terms corresponding to the left hand side of (1), $\mathbf{\Phi} = (\Phi_1, ..., \Phi_T)'$, are periodic weighted averages and their standard errors can be obtained directly from an alternative moment-motivated regression. To show how, let $\varphi^*_{it} = \varphi_{it}\sqrt{L_{it}}$, collect these weighted productivity indices for firm i into $\varphi^*_i = (\varphi^*_{i1}, ..., \varphi^*_{iT})'$, and define a $(T \times T)$ matrix $\mathbf{X}^*_i$ with t^{th} row $\mathbf{x}^{*\prime}_{it} = \left(d_{1,it}\sqrt{L_{it}}, ..., d_{T,it}\sqrt{L_{it}}\right)$, where $d_{it,s}$ is a period indicator. Thus, the rows of $\mathbf{X}^*_i$ consist of weighted period dummies.

Using this notation, we obtain the following moment condition for firm i:

$$E\left[\mathbf{X}^{*\prime}_i\left(\varphi^*_i - \mathbf{X}^*_i\alpha\right)\right] = \mathbf{0}_{(T\times 1)} \tag{4}$$

where $\alpha = (\alpha_1, ..., \alpha_T)'$ is a $(T \times 1)$ parameter vector. By the analogy principle, this moment condition results in the standard pooled OLS estimator of a linear panel model that regresses φ^*_{it} on the complete set of period indicators interacted with $\sqrt{L_{it}}$. Stacking all firms $\varphi^{*\prime} = (\varphi^{*\prime}_1 \cdots \varphi^{*\prime}_N)$ and similarly for $\mathbf{X}^*$, the OLS estimator is $\widehat{\alpha} = (\mathbf{X}^{*\prime}\mathbf{X}^*)^{-1}\mathbf{X}^{*\prime}\varphi^*$. This is an estimator of $\mathbf{\Phi}$. Equivalently, if we let $\mathbf{W}_i = diag\,(L_{i1}, ..., L_{iT}))$, (4) can be rewritten as $E\left[\mathbf{D}_i{}'\mathbf{W}_i\left(\varphi_i - \mathbf{D}_i\alpha\right)\right] = \mathbf{0}$. Stacking all firms $\mathbf{D} = (\mathbf{D}'_1 \cdots \mathbf{D}'_N)$ and using the stacked $\mathbf{D}$ and $\mathbf{W} = diag\left[\mathbf{W}_1, ..., \mathbf{W}_N\right]$, the resulting estimator for $\mathbf{\Phi}$ is $\widehat{\alpha} = (\mathbf{D}'\mathbf{W}\mathbf{D})^{-1}\mathbf{D}'\mathbf{W}\varphi$.

It follows from the definition of this alternative estimator that $\widehat{\alpha} = \widehat{\theta} + \widehat{\gamma}$ and that $Var\left[\widehat{\alpha}\right] = Var\left[\widehat{\theta} + \widehat{\gamma}\right]$. This suggests two things: First, that there is an indirect way to estimate the covariance term using the difference between the weighted and unweighted means, i.e. $\widehat{\gamma} = \widehat{\alpha} - \widehat{\theta}$. The standard error of the covariance term is the square root of the diagonals of $Var\left[\widehat{\gamma}\right] = Var\left[\widehat{\alpha} - \widehat{\theta}\right]$. Second, statistical inference about the left hand side of (1) can be based on (3), which provides a consistent estimator of the asymptotic covariance matrix for $\widehat{\alpha}$ that is robust to within-firm autocorrelation and heteroscedasticity. This relation implies that it is easy to test hypotheses about, e.g. how aggregate industry productivity has developed over the sample period. For example, testing H_0: $(\theta_T - \theta_1) + (\gamma_T - \gamma_1) = 0$ would be a test of the hypothesis that there has been no (aggregate) productivity growth over the sample period (i.e. $\Phi_1 = \Phi_T$).

3.3.3. *Entry and Exit*

In some applications it may be desirable to be able to allow for unbalanced panel data that are due to the entry and exit of firms. One way to do so is to assume that there are three mutually exclusive subgroups *in each period*, i.e. the subgroups for surviving incumbents ($j = 1$), entrant firms ($j = 2$) and those firms that exit before the end of the next period ($j = 3$). One can then follow the steps outlined above for the estimation of the OP productivity components in mutually exclusive subgroups. The subgroup indicator $q_{it,j}$ is defined as follows: If firm i neither enters at t nor exits at $t+1$, $q_{it,1} = 1$ and $= 0$ otherwise. If firm i enters the data during period t, $q_{it,2} = 1$ and $= 0$ otherwise. For those firms that exit at $t+1$, $q_{it,3} = 1$ and $= 0$ otherwise. Using this notation, $\sum_{j=1}^{3} q_{it,j} = 1$.

There is, however, a remaining piece of ambiquity in how one should classify new plants that enter the data by the end of period t and exit by the end of period $t+1$. For them, $q_{it,2} = 1$ and $q_{it,3} = 1$. When the time period becomes shorter, the share of such observations gets smaller. In applications that use annual data, the number of observations of this type may be non-negligible. A practical solution to this problem is to introduce a fourth plant category for such 'experimental' short-lived entrants. An alternative way to think about this is to realize that in a dynamic context entry is backward looking and exit is forward looking. Therefore, one could - as an anonymous referee reminded us - treat a group of firms that exists in a given period both as a group of survivors and entrants relative to the previous period and as a (different) group of survivors and exitors relative to the next period. This approach would be in line with various (dynamic) versions of OP decomposition discussed and proposed by, e.g. Böckerman and Maliranta (2007), Diewert and Fox (2009), Hyytinen and Maliranta (2013) and more recently by Melitz and Polanec (2015).

Slicing the data as described above, one can, for example, compare productivity levels between entrants and surviving incumbents in a given period (e.g. test H_0: $\gamma_{t,2} + \theta_{t,2} - \gamma_{t,1} - \theta_{t,1} = 0$), study whether the (relative) productivity levels of entry vintages change over time (e.g. test H_0: $(\gamma_{t,2} + \theta_{t,2}) - (\gamma_{t-s,2} + \theta_{t-s,2}) = 0$ or H_0: $\gamma_{t,2} + \theta_{t,2} - \gamma_{t,1} - \theta_{t,1} = \gamma_{t-s,2} + \theta_{t-s,2} - \gamma_{t-s,1} - \theta_{t-s,1}$) and examine whether it is the change in the covariance component (e.g. test H_0: $(\gamma_{t,2} - \gamma_{t-s,2}) = 0$ or H_0: $\gamma_{t,2} - \gamma_{t,1} = \gamma_{t-s,2} - \gamma_{t-s,1}$), or changing average productivity of entrants (e.g. test H_0:

$(\theta_{t,2} - \theta_{t-s,2}) = 0$ or H_0: $\theta_{t,2} - \theta_{t,1} = \theta_{t-s,2} - \theta_{t-s,1})$ that drives the change.

4. Application

In this section, we apply our estimation procedure to Finnish firm-level panel data from 1995 to 2007. What follows is presented as an illustrative case study rather than a full-blown and self-contained empirical analysis. We therefore focus on the development of labour productivity in a single industry, 'Computer and related activities' (NACE 72). The industry is an example of a dynamic service industry in Finland, with high net employment growth and intensive hiring and separation rates of the employees.

4.1. Data

Our firm-level data come from the Structural Business Statistics (SBS) data of Statistics Finland. The SBS includes all firms in the Finnish business sector. For larger firms, the SBS data are primarily obtained from the Financial Statements inquiry. For those firms not covered by the inquiry, typically employing less than 20 persons, data come from the Finnish Tax Administration's corporate taxation records and Statistics Finland's Register of Enterprises and Establishments. We have used the implicit price index of the industry obtained from the Finnish National Accounts for deflation.

We measure labour productivity φ_{it} by (the logarithm of) value added per person in year 2000 prices, and activity shares s_{it} by the employment share of firms. The number of employees refers to the average number of persons engaged in the activities of a firm during the accounting period. This convention means that a person who has been employed in the firm for six months corresponds to half an employee. Employees comprise wage and salary earners and self-employed persons. Employees are converted to annual full-time employees so that, for example, an employee working half-time represents one half of a person and two employees working half-time for one year represent one annual full-time employee. We exclude from our empirical analysis the firm observations that have less than one (employed) person or that have negative value added.

Inspired by Böckerman and Maliranta (2007), Ottaviano *et al.* (2009), and Bartelsman *et al.* (2013), we focus on cross-regional differences in productivity. We estimate, in particular, OP productivity decompositions for two Finnish regions that are in many ways dissimilar. The first is Uusimaa, which is a region in Southern Finland province. Uusimaa consists of Helsinki, the capital of Finland, and 20 surrounding municipalities. The population of the region is 1.4 million, which is a quarter of the total population of Finland. The second region is Itä-Suomi (Eastern Finland). It is a sparsely populated region whose area is 7.6 times larger than that of Uusimaa but its population is only 40% of that in Uusimaa. Uusimaa is much richer than Itä-Suomi; according to the statistics of Eurostat, in 2006 the GDP per inhabitant was 56.9% above the EU average in Uusimaa but 14.7% below in Itä-Suomi.

Table 1: Descriptive statistics

	Region: Uusimaa			Region: Itä-Suomi		
Year	Number of firms	Total number of employees	Aggregate productivity	Number of firms	Total number of employees	Aggregate productivity
1995	667	10541.1	3.88	61	482.4	3.70
2000	855	21604.6	4.02	109	886.8	3.68
2005	892	24931.0	3.98	94	835.0	3.53
2007	1183	29614.7	4.09	108	807.7	3.69

Notes: Data source is Structural Business Statistics (SBS) data of Statistics Finland.

Fig. 1. Descriptive statistics

We define two mutually exclusive subgroup indicators based on the primary location of firms. Some of the multi-unit firms have activities in several regions. In these cases, the location of the firm refers to the region that has the highest within-firm employment share. The distribution of a firm's employment by region is computed by using Statistics Finland's Register of Enterprises and Establishments. The first subgroup indicator $q_{it,U}$ is equal to one if firm i is located in Uusimaa ($j = U$) in period t and zero otherwise. The second indicator $q_{it,I}$ is defined similarly for firms located in Itä-Suomi ($j = I$).

Descriptive statistics of the data can be found in Table 1. It displays for selected years the number of firms ($N_{t,j} = \sum_{i=1}^{N} q_{it,j}$), total employment ($\sum_{i=1}^{N_t} q_{it,j} L_{it}$) and the weighted average of labour productivity ($\sum_{i=1}^{N_t} q_{it,j} s_{it} \varphi_{it}$), separately for $j \in \{U, I\}$.

4.2. Results

The results that we obtained by the moment-based approach developed above are displayed in Table 2. The two panels of the table display point estimates and the associated 95% confidence intervals, based on the autocovariance and heteroskedasticity robust standard errors, for the average of labour productivity ($\widehat{\theta}$, Panel A) and the covariance term ($\widehat{\gamma}$, Panel B) separately for the two regions over the sample period from 1995 to 2007.

Looking first at the point estimates only, Table 2 suggests the following. On the one hand, there is a positive trend in average firm productivity in both the Uusimaa and Itä-Suomi regions. The point estimates also show that average productivity is higher in Uusimaa. On the other hand, it seems that the covariance component has made a negative contribution in both regions (its size diminishes over time) and that it has on average been higher in Uusimaa. The confidence intervals of the average firm productivity are much narrower in Uusimaa than in Itä-Suomi, whereas the confidence intervals of the covariance term are about the same order of magnitude.

We illustrate the benefits of our inference procedure with the following examples:

First, the null hypothesis that the level of (aggregate) industry productivity has not changed from 1995 to 2007 (i.e. H_0: $\theta_{2007,j} + \gamma_{2007,j} - \theta_{1995,j} - \gamma_{1995,j} = 0$) is rejected for Uusimaa ($j = U$) but not for Itä-Suomi ($j = I$). The p-value for the (Wald) test of the former hypothesis is 0.001, whereas it is

Panel A: Average of labour productivity by region

	Region: Uusimaa			Region: Itä-Suomi		
	Point	95% confidence interval		Point	95% confidence interval	
Year	estimate	Lower bound	Upper bound	estimate	Lower bound	Upper bound
1995	3.51	3.46	3.57	3.33	3.14	3.52
1996	3.40	3.34	3.45	3.31	3.16	3.46
1997	3.44	3.39	3.49	3.40	3.27	3.53
1998	3.66	3.60	3.71	3.35	3.15	3.55
1999	3.73	3.67	3.78	3.51	3.35	3.66
2000	3.81	3.76	3.86	3.51	3.36	3.67
2001	3.74	3.68	3.79	3.39	3.21	3.57
2002	3.73	3.68	3.78	3.45	3.34	3.57
2003	3.72	3.67	3.77	3.51	3.42	3.61
2004	3.75	3.70	3.80	3.46	3.34	3.59
2005	3.79	3.74	3.84	3.56	3.45	3.66
2006	3.81	3.77	3.85	3.62	3.51	3.73
2007	3.85	3.80	3.89	3.57	3.44	3.71

Panel B: Covariance term by region

	Region: Uusimaa			Region: Itä-Suomi		
	Point	95% confidence interval		Point	95% confidence interval	
Year	estimate	Lower bound	Upper bound	estimate	Lower bound	Upper bound
1995	0.37	0.17	0.57	0.37	0.11	0.63
1996	0.51	0.27	0.76	0.32	0.20	0.45
1997	0.49	-0.05	1.02	0.30	0.11	0.48
1998	0.32	0.02	0.62	0.29	0.11	0.48
1999	0.25	0.05	0.46	0.13	0.01	0.26
2000	0.21	0.01	0.41	0.17	0.04	0.30
2001	0.25	0.09	0.41	0.23	0.09	0.38
2002	0.22	0.06	0.38	0.00	-0.30	0.30
2003	0.25	0.04	0.47	-0.02	-0.21	0.18
2004	0.26	0.00	0.53	0.02	-0.22	0.27
2005	0.19	0.05	0.33	-0.02	-0.32	0.27
2006	0.25	0.11	0.39	0.06	-0.03	0.15
2007	0.24	0.11	0.38	0.12	-0.05	0.29

Notes: Data source is Structural Business Statistics (SBS) data of Statistics Finland. The point estimates for the average productivity and covariance term as well as the 95% confidence interval (CI) have been calculated using the GMM procedure. The standard errors that have been used to calculate the CI are autocovariance and heteroscedasticity robust.

Fig. 2. Average of labour productivity and covariance term by region

0.966 for the latter.

Second, there is a statistically significant improvement from 1995 to 2007 in the average productivity of firms in Uusimaa. For Uusimaa, we reject H_0: $\theta_{2007,U} - \theta_{1995,U} = 0$ at better than the 1% significance level. However, the same null hypothesis for Itä-Suomi cannot be rejected as confidently; its p-value is 0.042.

Third, we cannot reject H_0: $\gamma_{2007,j} - \gamma_{1995,j} = 0$ for either region (with p-values 0.255 and 0.108 for $j = U, I$, respectively). This finding suggests that the negative trend in the covariance terms cannot be measured accurately in our data.

Fourth, we clearly reject the null hypothesis that the level of industry productivity has, on average, been the same in the two regions during the period from 1995 to 2007. This hypothesis is equivalent to H_0: $\frac{1}{13}\sum_{t=1995}^{2007}(\theta_{t,U} + \gamma_{t,U} - \theta_{t,I} - \gamma_{t,I}) = 0$, and the p-value of the associated test is less than 0.001.

Fifth, there is a statistically significant difference between the two regions in the average productivity of firms; the p-value of the test for H_0: $\frac{1}{13}\sum_{t=1995}^{2007}(\theta_{t,U} - \theta_{t,I}) = 0$ is less than 0.001. However, the regional difference in the covariance terms is not statistically significant; the p-value of the test for H_0: $\frac{1}{13}\sum_{t=1995}^{2007}(\gamma_{t,U} - \gamma_{t,I}) = 0$ is 0.146.

Finally, the numbers suggest that when we compare the two regions, the last six years of our sample period might be special, with non-negligible regional differences: This is indeed what we find if we test for such differences formally: H_0: $\frac{1}{6}\sum_{t=2002}^{2007}(\theta_{t,U} + \gamma_{t,U} - \theta_{t,I} - \gamma_{t,I}) = 0$ can be rejected at better than the 1% significance level (p-value < 0.001). This result appears to be due to two things: First, the average productivity of firms in Uusimaa has, on average, been higher during these years. The p-value of the test for H_0: $\frac{1}{6}\sum_{t=2002}^{2007}(\theta_{t,U} - \theta_{t,I}) = 0$ is less than 0.001. The covariance term has also been higher in Uusimaa than in Itä-Suomi. However, the difference cannot be measured as accurately; the p-value of the test for H_0: $\frac{1}{6}\sum_{t=2002}^{2007}(\gamma_{t,U} - \gamma_{t,I}) = 0$ is 0.024.

In summary, our findings show that formal statistical inference casts in some cases clear doubt on the conclusions that one might draw about the sources of productivity dynamics based on a visual inspection of the numbers.

5. Conclusions

We show how a standard moment-based GMM procedure can be used to simultaenously compute point estimates for the components of the Olley-Pakes productivity decomposition and to estimate their standard errors. The procedure provides applied researchers with a simple two-step recipe for autocovariance and heteroskedasticity robust inference and allows for hypothesis testing about, e.g. the co-evolution of the productivity components in different groups of firms. We have framed our analysis in terms of population moments and GMM because they immediately suggest a number of ways of how the estimation and inference procedure might be extended.

We illustrate the procedure by applying it to Finnish firm-level data from 1995 to 2007. We find that formal statistical inference casts some doubt on the conclusions that one might draw based on a visual inspection of the two components of the OP decomposition. It also turns out that in our case study, the covariance term, which often is of major policy interest, cannot be measured as accurately as the unweighted average of the productivity of firms.

The scope of our analysis also deserves attention. We have proposed a method of computation and inference for one of the many decomposition methods that have been used in the literature. Further work could thus consider procedures for inference for the other popular decompositions, such as those proposed by Bailey *et al.* (1992), Foster *et al.* (2001), Petrin and Levinsohn (2013) and Melitz and Polanec (2015). Some of these methods appear to suggest that reallocation of resources has a negligible or negative contribution to aggregate productivity growth in the United States (Bailey *et al.* 2001, Foster *et al.* 2001, OECD 2003 and Petrin *et al.* 2011). This finding has not gone unchallenged. Some authors argue that it is at least partly due to problems in measurement and economic interpretation (e.g. Bartelsman *et al.* 2013), whereas others see it as a sign of deeper methodological problems (Petrin *et al.* 2011; see also Fox 2012 for a critical account). In our view, it would be of interest to know how wide the (robust) confidence intervals for the debated estimates are. Our procedure may provide a starting point for such inference.

There appears to be ample room for appropriate inference procedures. For example, there is another strand in the decomposition literature that explores the sources of aggregate productivity growth. In this strand, productivity growth is decomposed into technical change and efficiency change using Malmquist indices. For example, Färe *et al.* (1994) use the technique to identify for each point in time the so-called efficient countries, which together define the current world technology frontier. Statistical inference for these techniques has been considered by, e.g. Simar and Wilson (1999) and Fuentes *et al.* (2001).

Another call for appropriate inference procedures comes from the need to analyze the workings of public policies. The OP decomposition was originally used to examine whether and how deregulation affects aggregate productivity and reallocation. Trade and price liberalization, tax reforms, labor market reforms and new forms of subsidized government funding provide similar pseudo-experiments as they are likely to have a differential impact on firms whose productivity is heterogenenous and on the allocation of resources among such firms. Our inference procedures can be used to test, e.g. whether such policy changes affect different industries or regions differentially. Moreover, in an ongoing work, one of us is exploring how well the OP covariance component and its variants mirror policy-driven allocation distortions (Maliranta and Määttänen 2015).

We conclude by pointing out that what we, as economists, call the OP decomposition is a procedure that decomposes a weighted average into a (simple) non-weighted average and a covariance-like term. Inference procedures for such a decomposition may clearly be useful in other fields of social science and statistics.

We can provide several concrete examples of areas in which our procedure could be utilized. First, there is a closely related research strand in the field of operations research that explores when and how firm-level efficiency measures can be aggregated to industry level using weighted average estimators (see Peyrache and Coelli 2009 and references therein). Our procedure for testing hypotheses about the covariance term provides a means to test the conditions when such aggregation procedures work. If the null hypothesis that the covariance term is zero cannot be rejected, industry efficiency can be proxied by the average efficiency of the firms (see Peyrache and Coelli 2009).

The approach of this paper also has connections to aggregation theory. A classic question is whether it is justified to aggregate micro equations to macro averages. The sum of linear micro equations can be expressed as the macro (average) equation plus the covariance of micro variables and parameters, where the covariance term is the aggregation bias (e.g. van Daal and Merkies 1984, and Stoker 2008, and the references cited there). This expression has the same form as the OP decomposition. Various ways of testing for aggregation bias have been suggested (e.g. Lee, Pesaran and Pierse 1990). Our method suggests a new way of testing the bias directly by testing the significance of the covariance term.

We see that the decomposition methods can also be utilized in other fields, such as political science, sociology or geography. The procedure developed for the OP decomposition of the weighted average provides a framework for statistical inference for such non-economic applications. For example, spatial dimension is often important in political science and geography. The procedure outlined in this paper can be extended to allow for dependence that results from spatial patterns.

6. Acknowledgements

We thank two anonymous referees as well as Anni Nevalainen, Timo Kuosmanen and Otto Toivanen for useful comments and discussions and Toni Hukkanen for superb technical assistance. Data work has been carried out at Statistics Finland following its terms and conditions of confidentiality. For access to the data, contact the Research Services unit of Statistics Finland, FI-00022 Statistics Finland (email: tutkijapalvelut@stat.fi). Maliranta's work has in part been supported by the Finnish Funding Agency for Technology and Innovation (TEKES, project 441/31/08).

References

Bailey, M. N., Hulten C., and Campbell, D. (1992) Productivity dynamics in manufacturing establishments. In *Brookings Papers on Economic Activity: Microeconomics* (eds. M. Baily and C. Winston), pp. 187-249, Washington: Brooking's Institution.

Balk, B. M. (2003) The residual: On monitoring and benchmarking firms, industries, and economies with respect to productivity. *Journal of Productivity Analysis*, 20(1), 5-47.

Bartelsman, E. J., Haltiwanger J. C. and Scarpetta, S. (2009) Measuring and analyzing cross-Country differences in firm dynamics. In *Producer Dynamics* (eds. T. Dunne, B. J. Jensen and M. J. Roberts), pp. 15-82, University of Chicago Press.

Bartelsman, E., Haltiwanger, J. ja Scarpetta, S. (2013) Cross-country differences in productivity: The role of allocation and selection. *American Economic Review*, 103(1), 305-334.

Berk, R. A., Western, B., and Weiss, R. E. (1995) Statistical inference for apparent populations. *Sociological Methodology*, 25, 421-458.

Böckerman, P. and Maliranta, M. (2007) The micro-level dynamics of regional productivity growth: The source of divergence in Finland. *Regional Science and Urban Economics*, 37(2), 165-182.

Chari, A. V. (2011) Identifying the aggregate productivity effects of entry and size restrictions: An empirical analysis of license reform in India. *American Economic Journal: Economic Policy*, 3, 66-96.

Cameron, A. C. and Trivedi, P. K. (2005) *Microeconometrics: Methods and Applications*, Cambridge University Press.

Davidson, R. and MacKinnon, J.G. (2004) *Econometric Theory and Methods*, Oxford University Press.

Diewert, W. E. and Fox K. A. (2009) On measuring the contribution of entering and exiting firms to aggregate productivity growth. In *Index Number Theory and the Measurement of Prices and Productivity* (eds. Diewert, W.E., Balk, B.M., Fixler, D., Fox , K.J. and Nakamura, A.), Victoria: Trafford Publishing.

Eslava, M., Haltiwanger J., Kugler, A. D. and Kugler, M. (2004) The effects of structural reforms on productivity and profitability enhancing reallocation: Evidence from Colombia. *Journal of Development Economics*, 75(2), 333-371.

Eslava, M., Haltiwanger J., Kugler, A. D. and Kugler, M. (2010) Factor adjustments after deregulation: Panel evidence from Colombian plants. *Review of Economics and Statistics*, 92(2), 378-391.

Eslava, M., Haltiwanger, J., Kugler, A. D. and Kugler, M. (2013) Trade and market selection: Evidence from manufacturing plants in Colombia. *Review of Economic Dynamics*, 16(1), 135-158.

Foster L., Haltiwanger J. and Krizan C. J. (2006) Market selection, reallocation, and restructuring in the U.S. Retail Sector in the 1990s. *Review of Economics and Statistics*, 88(4), 748-758.

Foster L., Haltiwanger J. and Syverson C. (2008) Reallocation, firm turnover, and efficiency: Selection on productivity or profitability. *American Economic Review*, 98(1), 394-425.

Fox, K. J. (2012) Problems with (dis)aggregating productivity and another productivity paradox. *Journal of Productivity Analysis*, 37(3), 249-259.

Fuentes, H. J., Grifell-Tatje, E. ja Perelman, S. (2001) A parametric distance function approach for Malmquist productivity index estimation. *Journal of Productivity Analysis*, 15(2), 79-94.

Färe, R., Grosskopf, S., Norris, M. ja Zhongyang, Z. (1994) Productivity growth, technical progress, and efficiency change in industrialized countries. *American Economic Review*, 84(1), 66-83.

Gelman A. and King G. (1995) A unified method of evaluating electoral systems and redistricting plans. *American Journal of Political Science*, 38(2), 514-554.

Hyytinen A. and Maliranta M. (2013) Firm lifecycles and evolution of industry productivity. *Research Policy*, 42(5), 1080-1098.

Levinsohn, J. and Petrin, A. (2003) Estimating production functions using inputs to control for unobservables. *Review of Economic Studies*, 70(2), 317-342.

Lee, K.C., Pesaran, M.H., and Pierse, R.G. (1990) Testing for aggregation bias in linear models. *Economic Journal*, 100, 137-150.

Maliranta, M. and Määttänen, N. (2015) An augmented static Olley-Pakes productivity decomposition with entry and exit: measurement and interpretation, manuscript, ETLA.

Manski, C. F. (1988) *Analog Estimation Methods in Econometrics*, Chapman and Hall.

Melitz, M. J. and Polanec, S. (2015) Dynamic Olley-Pakes decomposition with entry and exit. *Rand Journal of Economics.*

OECD (2003) *The sources of economic growth in OECD countries.* OECD, Paris.

Olley, G. S. and Pakes, A. (1996) The dynamics of productivity in the telecommunications equipment industry. *Econometrica*, 64(6), 1263-1297.

Peyrache, A. and Coelli, T. (2011) Testing procedures for detection of linear dependencies in efficiency models. *European Journal of Operational Research*, 198, 647-654.

Petrin, A., White, T. K. and Reiter, J. P. (2011) The impact of plant-level resource reallocations and technical progress on U.S. macroeconomic growth. *Review of Economic Dynamics*, 14, 3-26.

Simar, L. and Wilson, P. A. (1999) Estimating and bootstrapping Malmquist indices. *European Journal of Operations Research*, 115(3), 459-471.

Stoker, T.M. (2008) Aggregation (econometrics). In *The New Palgrave Dictionary of Economics* (eds. S.N. Durlauf and L.E. Blume), The New Palgrave Dictionary of Economics Online, Palgrave Macmillan. DOI:10.1057/9780230226203.0018

To appear: Journal of the Royal Statistical Society, Series A

Van Biesebroeck, J. (2003) Productivity dynamics with technology choice: An application to automobile assembly. *Review of Economic Studies*, 70(1), 167-198.

Van Biesebroeck, J. (2008) The sensitivity of productivity estimates: Revisiting three important productivity debates. *Journal of Business and Economic Statistics*, 26(3), 321-238.

van Daal, J. and Merkies, A.H.Q.M. (1984), *Aggregation in Economic Research. From Individual to Macro Relations*, Dordrecht: D. Reidel Publishing Company.

Wooldridge, J.M. (2002) *Econometric Analysis of Cross Section and Panel Data*, MIT Press.

Wooldridge, J.M. (2003) Cluster-Sample Methods in Applied Econometrics. *American Economic Review*, 93, 133-138.