Int. J. Production Economics 137 (2012) 29–35

Contents lists available at SciVerse ScienceDirect

Int. J. Production Economics

journal homepage: www.elsevier.com/locate/ijpe

Multiple-method analysis of logistics costs

Janne Engblom [a], Tomi Solakivi [b], Juuso Töyli [b,*], Lauri Ojala [b]

[a] Department of Accounting and Finance, Turku School of Economics at the University of Turku, 20014 Turun Yliopisto, Finland
[b] Department of Marketing and International Business, Turku School of Economics at the University of Turku, 20014 Turun Yliopisto, Finland

ARTICLE INFO

Article history:
Received 24 February 2011
Accepted 12 January 2012
Available online 25 January 2012

Keywords:
Logistics costs
Panel study
Mixed model
Random effect

ABSTRACT

Logistics costs comprise a significant and relevant proportion of business costs, often exceeding 10 per cent of company turnover. This article examines the differences and interdependencies in the self-reported logistics costs of manufacturing and trading companies operating in Finland. Total logistics costs are taken to consist of six individual components: transport, warehousing, inventory carrying, logistics administration, transport packaging, and indirect costs of logistics. The analysed panel data covers 241 companies identified from two surveys for the years 2005 and 2008. Logistics costs were explored through multiple methods including descriptive analysis, generalised linear mixed models (GLMM), and principal component analysis. The distributions of logistics costs measured as percentages of turnover were skewed and best described by the beta distribution. Time, the number of employees, turnover, industry, and level of internationalisation were shown to be statistically significant explanatory variables of logistics costs. Logistics costs tended to be lower in larger companies, although diseconomies of scale eventually prevail. The analysis also covers changes in costs between 2005 and 2008. In general, the results indicate the need for caution in interpreting changes in logistics costs, and for simultaneously controlling the effects of background variables.

1. Introduction

Logistics costs comprise a significant and relevant proportion of business costs: depending on the method applied and the industry in question, their share of company turnover in developed economies tends to be at least 10 per cent. However, definitions of logistics costs are many, and vary considerably. For the purposes of this article total logistics costs are measured as a percentage of turnover and comprise six individual components: transport, warehousing, inventory carrying, (logistics) administration, (transport) packaging, and indirect costs of logistics. This classification resembles that used by Naula et al. (2006) and Töyli et al. (2008).

Heskett et al. (1973) identify transportation, warehousing, inventory carrying and administration as components of logistics costs, a classification that has been widely used. Gunasekaran et al. (2001) include the opportunity cost of capital and storage, risk costs, and the possible costs of lost sales under inventory costs, thus combining the elements of inventory carrying and warehousing. Beamon (1999), on the other hand, distinguishes between operating costs and inventory costs, whereas Lambert and LaLonde (1976) separate warehousing and inventory carrying costs, but include some of the components of warehousing costs, such as the inventory service costs and storage space costs in inventory carrying. For more discussion about the definition of inventory-related costs, see also Stewart (1995a, b), Lee and Billington (1992), and Levy (1997). In sum, these four cost components seem to be in general use on the one hand, but on the other hand authors tend to define the limits of the categories in unique, inconsistent ways.

The literature identifies a wide range of logistics cost components in addition to transport, warehousing, inventory carrying and logistics administration. For example, Zeng and Rossetti (2003) add customs, risk and damage as well as handling and packaging to the list, and Ojala et al. (2007) include "other logistics costs" to reflect the fact that logistics costs can rarely be divided and measured accurately. Klaus and Kille (2007) and Klaus et al. (2010) single out order-entry costs from other administration costs as a separate component.

The ambiguities are not limited to defining and understanding logistics cost components. They are also rife in the measurement approaches and the scope of analysis, which may vary from a single function or project within a firm to the entire company, a fraction of a supply chain, or even a demand–supply network.

There appear to be three main measurement approaches. The first is statistics-based and uses various types of national accounts or industrial statistics to create an estimate for logistics costs

* Corresponding author. Tel.: +358 2 333 9248; fax: +358 2 481 4280.
E-mail addresses: janne.engblom@tse.fi (J. Engblom), tomi.solakivi@tse.fi (T. Solakivi), juuso.toyli@gmail.com (J. Töyli), lauri.ojala@tse.fi (L. Ojala).

0925-5273/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.ijpe.2012.01.007

as a share of GDP. Examples of its implementation include Elger et al, (2008), Wilson (2009) and Havenga (2010). Bowersox et al. (2003) and Rodrigues et al. (2005), applied econometric modelling based on this kind of data in order to estimate the GDP share of logistics costs worldwide. The so-called "Top 100" reports on European logistics markets compiled by Klaus et al. (2010) estimate the size of logistics expenditure, and ultimately logistics costs, in Europe based on statistical data on transport volumes, employment in logistics, and demand for logistics services.

Second, van Damme and van der Zorn (1999) and Baykasoglu and Kaplanoglu (2008), for example, take an accounting and activity-based-costing approach. In a similar vein, Pirttilä and Huiskonen (1996) apply so-called mission costing. This technique also comprises the use of specific logistics-performance indicators in various resources-planning software products for enterprises, such as the SCOR model developed by the Supply Chain Council. These have enabled accounting-based comparisons across industries and countries, especially among larger firms using elaborate ERP systems (see also Pohlen et al., 2009).

The third main measurement approach is survey-based, and logistics costs are assessed from self-reported company data. The surveys are usually questionnaire-based and result in estimates of logistics costs as a percentage of sales. The European Logistics Association and Kearney (2004, 2009) has used this approach, as have Naula et al. (2006), and Ojala et al. (2007).

This article examines the differences and the interdependencies in self-reported logistics costs of manufacturing and trading companies operating in Finland. In order to assess temporal differences, panel data on 241 companies was collected through two surveys in the years 2005 and 2008. This survey-based approach was the only feasible measurement alternative that would yield sufficiently detailed information about the logistics cost components in a broad range of companies.

Traditional assessments of logistics costs are based mainly on calculations and reports of their mean values among the target population or sample, related to a mixed set of background variables. The limitations of this method are acknowledged here (for a more detailed discussion about these limitations see, for example, Dodd et al., 2006), and a different method based on generalised linear mixed models (GLMMs) for panel data is applied. The differences of cost estimates through descriptive analysis and GLMM model based means are discussed. GLMMs take into account multivariate, longitudinal, and non-normality aspects of data, for example, and differ from descriptive analysis in which dependent variables are analysed alone in one- or two-dimensional space. GLMMs also allow the simultaneous analysis of statistical dependencies between logistics costs and several explanatory variables. Descriptive analysis may give misleading results because differences between the direct and indirect relationships among the variables remain unrecognised.

Generalised Linear Models (GLMs) represent a class of fixed-effects regression models for several types of dependent variables (i.e., continuous, dichotomous, counts; McCullagh and Nelder, 1989). In GLMs all independent variables (effects) are assumed fixed which means that there are observations from every level of that effect. If there is only a sample of levels in the data, the effect is called random (McCulloch et al., 2001). For the purpose of this article, logistics cost components are treated as continuous variables with two observations nested within randomly selected companies. This necessitates the use of random effects in the model. GLMs, which assume that all observations are independent of each other, are not appropriate for the analysis of these correlated data structures, longitudinal panel data in particular (see de Leeuw and Meijer, 2008). Analysis of this kind of multilevel data requires the addition of random cluster and/or subject effects into the regression model in order to account for the correlation. The resulting model, with both fixed and random effects, is called a linear mixed model (LMM). Given that the preliminary descriptive analysis indicated non-normality of the logistics cost components, and the beta distribution turned out to be more suitable than gamma distribution, the beta distribution was used in the reported analysis. For this reason, the word 'generalised' was added to the method description (GLMM).

2. Research design

2.1. Data and construct operationalisation

The panel data analysed in this article was a sub-sample of all 241 manufacturing and trading companies answering the identical questions about logistics cost components both in 2006 (data for 2005) (Naula et al., 2006) and 2009 (data for 2008) (Solakivi et al., 2009), which covered 13.8 and 13.6 per cent, respectively, of all the responses from both industries (variable INDUST in the statistical analysis). The logistics cost component data was measured at one-percentage-point intervals (variable COST in the statistical analysis), and the respondents were asked to assess the different logistics cost components as a percentage of the turnover, which according to Stewart (1995a, b) is a robust base for analysis. The respondents were free to choose anything between zero and 100 per cent. Given that the option "no response" was included and was the default value, zero values were conscious responses implying that the true cost was somewhere between zero and half a percentage point.

Determination of the company size was based on self-reported company turnover (TURNOV) and the number of employees (EMP). The European Commission defines the following limits: large companies have a turnover of more than 50 million euro, medium-sized companies between 10 and 50 million euro, small companies between 2 and 10 million, and micro companies a maximum of two million euro. A similar classification was used for categorising the number of employees: 1–9, 10–49, 50–249 and at least 250 employees, respectively.

The companies were also classified in terms of their level of internationalisation (INT) as domestic, export, and international based on the percentage share of sales and production from outside domestic markets. A domestic company was defined as one generating less than 10 per cent of its sales outside domestic markets, an export company as one generating over 10 per cent of its sales from outside, and an international company as also having production facilities abroad.

2.2. The GLMM model selection

Several generalised linear mixed models were used for the different logistics cost components. A general matrix form of a linear mixed model is defined as

$$\mathbf{Y} = \boldsymbol{\mu} + \boldsymbol{\varepsilon} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{b} + \boldsymbol{\varepsilon} \tag{1}$$

where $\mathbf{Y}$ is a matrix for dependent variables, $\boldsymbol{\mu}$ the mean vector and $\mathbf{X}$ and $\boldsymbol{\beta}$ are for fixed effects, similar to regression analysis, for example. Linear dependency is assumed through a link function, which in this case is identity, in other words the value of the dependent variable itself is explained. $\mathbf{Z}$ is a known matrix, the design, or incidence matrix for the random effects, and $\mathbf{b}$ is a vector that contains the coefficients of all the random variables. Matrix $\boldsymbol{\varepsilon}$ represents the residuals of the model. Moreover, if matrices $\mathbf{b}$ and $\boldsymbol{\varepsilon}$ are normally distributed

$$\begin{pmatrix} \mathbf{b} \\ \boldsymbol{\varepsilon} \end{pmatrix} \sim N\left(\begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma} \end{pmatrix} \right) \tag{2}$$

where $\mathbf{D}$ and $\boldsymbol{\Sigma}$ are covariance matrices of $\mathbf{b}$ and $\boldsymbol{\varepsilon}$. This also