Available online at www.sciencedirect.com
ScienceDirect
Labour Economics 13 (2006) 611–638
LABOUR ECONOMICS
www.elsevier.com/locate/econbase

A gender wage gap decomposition for matched employer–employee data☆

Ossi Korkeamäki, Tomi Kyyrä *

Government Institute for Economic Research, P.O. Box 1279, FIN-00101 Helsinki, Finland

Received 27 April 2004
Available online 31 May 2005

Abstract

In this paper we evaluate the extent to which the gender wage gap in the Finnish manufacturing sector is attributable to within-job wage differentials, sex differences in individual qualifications, and disproportionate concentration of women in lower-paying firms and lower-paying jobs within firms. We use matched employer–employee data to compare wage differentials between similarly qualified female and male workers who are doing similar work for the same employer. Our modelling approach employs a correlated random effects specification to account for the hierarchical grouped structure of the underlying data.

JEL classification: J16; J31; J71; C21
Keywords: Gender wage gap; Wage discrimination; Sex segregation; Random effects model

☆ The earlier version of the paper was circulated under the title "Explaining gender wage differentials: Findings from a random effects model". We appreciate the helpful comments received at the 3rd Nordic Workshop on the Economic Analysis of Linked Employer–Employee Data in Bergen, 2nd Nordic Econometrics Meeting in Bergen, CAED Conference in London, and EALE Conference in Seville. We are grateful to the Confederation of Finnish Industry and Employers for access to their data. Antti Luukkonen kindly provided his measures of job complexity levels for our use. The suggestions of a co-editor and anonymous referees considerably improved the paper.

* Corresponding author. Tel.: +358 9 703 2942; fax: +358 9 703 2969.
E-mail address: tomi.kyyra@vatt.fi (T. Kyyrä).

0927-5371/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.labeco.2005.02.001

1. Introduction

A huge body of literature has emerged to explain why the gender wage gap persistently exists in virtually all labour markets (see Altonji and Blank, 1999; Blau and Kahn, 2000, for recent surveys). Traditional attempts to explain the wage gap focused on sex differences in individual qualifications and their rewards in the labour market. More recently, the importance of the segregation of women and men into different jobs has been recognised. This line of research emphasizes that wages are closely tied to the characteristics of jobs, not only to the individuals who hold them. If typical female jobs pay lower wages than jobs dominated by men, the mean earnings of women can fall short of men's earnings even in the absence of within-job wage differentials between sexes.

Attempts to quantify the segregation effects on the wage gap were distorted by a lack of appropriate data for a long time. Consequently, most of the early analysis focused on segregation among occupations, firms, or industries only. This is clearly unsatisfactory, as women and men are further segregated into different jobs within firms. In recent years important advances have been made by access to large matched employer–employee data sets that contain multiple observations on workers with the same employer. When information on occupations or job titles is available, such data enable wage comparisons between male and female workers who are doing similar work for the same employer. This kind of comparative analysis has been conducted by Petersen and Morgan (1995), Petersen et al. (1997), Meyersson Milgrom et al. (2001), Groshen (1991), Datta Gupta and Rothstein (2001), and Bayard et al. (2003). In the first three of these studies observed sex differentials in mean wages within jobs are simply aggregated to form various wage decompositions. This approach has the obvious drawback that variation in individual characteristics is left uncontrolled. In the other studies wages are regressed against a set of control variables and fraction female in the worker's industry, firm, occupation, and/or job.[1] The key idea is that the regression coefficients of the various fraction female variables capture the relationship between the wage rate and 'femaleness' of the underlying labour market structure.

It should be noted that a common practice in the fraction female regressions above has been to neglect the grouping in the underlying data. For example, observations on workers resulting from the same firm are interpreted as being independent.[2] However, intuition suggests that we should expect workers in the same firm to be more homogeneous than those in a sample drawn randomly from the population of all firms. Workers in the same firm share many common factors, some of which may be observable (e.g. firm size, fraction female) but many are not (e.g. market power, managerial ability). In the regression analysis the effect of such unobservables serves as a latent firm effect that will be absorbed into the error term. Moreover, since different jobs require different skills and qualifications, we can further expect that within a given firm workers who are doing the same job are more homogenous than

[1] In a related paper we apply this method to the Finnish data; see Korkeamäki and Kyyrä (2002). Groshen's (1991) specifications do not include control variables.

[2] Bayard et al. (2003) report the standard errors adjusted for intraestablishment error correlation.