# Are Jobs More Polarized in ICT Firms?

Petri Böckerman
Seppo Laaksonen
Jari Vainiomäki

DISCUSSION PAPER SERIES

I Z A

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

# Are Jobs More Polarized in ICT Firms?

**Petri Böckerman**
*Turku School of Economics,*
*Labour Institute for Economic Research and IZA*


**Seppo Laaksonen**
*University of Helsinki*


**Jari Vainiomäki**
*University of Tampere*

# ABSTRACT

# Are Jobs More Polarized in ICT Firms?[*]

We perform decompositions and regression analyses that test the routinization hypothesis and implied job polarization at the firm level. Prior studies have focused on the aggregate, industry or local levels. Our results for the abstract and routine occupation groups are consistent with the routinization hypothesis at the firm level. The observed changes are linked to ICT adoption. Thus, disappearing middle-level (routine) work can be traced to firm-level technological change.

Corresponding author:

Petri Böckerman
Labour Institute for Economic Research
Pitkänsillanranta 3A
FI-00530 Helsinki
Finland
E-mail: petri.bockerman@labour.fi

---

## 1. Introduction

According to the dominant view in the literature, the driving force behind the increase in wage differentials, education premiums and skill-upgrading observed in industrialized countries since the 1970s/1980s has been skill-biased technological change (SBTC) (see e.g., Bound and Johnson (1992); and Acemoglu (2002) for a review). However, Card and DiNardo (2002), and Acemoglu and Autor (2010) have noted several anomalies that are difficult to explain with the standard SBTC model. Acemoglu and Autor (2010) presented a task-based model of technological change to explain these anomalies. Earlier, Autor, Levy and Murnane (ALM) (2003) focused on computer-based production technologies and also argued for an approach that distinguishes between the tasks that workers do in their jobs and the skills they use to perform these tasks. Adoption of computers at the workplace changes the demand for tasks that are performed at work. Computers replace the performance of routine tasks, but complement non-routine analytical and interactive tasks. ALM (2003) showed that a decline in computer prices leads to increased demand for total routine task input, including computer capital, but a reduction in the routine task input of human labour. This process is dubbed the routinization hypothesis (or routine-biased technological change). The increase in analytical non-routine tasks increases the relative demand for educated workers, because they have a comparative advantage in performing these tasks and in computer usage. On the other hand, the demand for intermediate-skilled (educated) workers declines, because they often work in routine-task intensive jobs.

Autor, Katz and Kearney (2006) and Goos and Manning (2007) argued that the routinization process may lead to job polarization, where employment growth is concentrated in low and high skill (wage) occupations, while jobs in the middle of the skill distribution continuum are diminished.[1] Low wage (education), non-routine manual and service jobs are unlikely to be directly affected by computerization, but non-homothetic preferences in product demand may lead to increasing demand for low paid services and non-routine manual jobs.

Acemoglu and Autor (2010) present empirical evidence for their task-based model using aggregate-level regressions for the wage changes of different skill groups

---

[1] Jenkins (1995) precedes the current literature on job polarization.

(defined by sex, education and experience). They measure the relative advantage of the skill groups in performing abstract, routine or service tasks by the shares of each skill group in abstract, routine, and service occupations prior to the computer era. The increase over time (which proxies for technological change) in the coefficients for these initial abstract and service shares, compared to routine shares, in explaining wage changes is consistent with their task-based model for technological change, routinization and polarization.[2]

ALM (2003) provided industry-level evidence for their model. They used task measures derived from the Dictionary of Occupation Titles (DOT) and showed that analytical and interactive non-routine task inputs increase more and routine task inputs decrease more in industries that invest more heavily in computer capital. Michaels, Natraj and Van Reenen (2014) used country-industry panel data, and showed that wage bill shares (and relative wages) for both high and low education levels are positively related to industry ICT capital, whereas those for the middle educated are negatively related to ICT. This pattern is consistent with job polarization. Also, Goos et al. (2010) used country-industry data to show that polarization is a pervasive phenomenon in 16 European countries, and furthermore, that the routine task intensity of occupations is a more important driver of employment polarization than offshorability.

Autor et al. (2013a) applied a spatial equilibrium model to show that the falling cost of computing, combined with preference for variety in consumption, can explain both the rise in service occupations and polarization in employment and wages. Using commuting-zone data they found that local labour markets that initially specialized in routine tasks adopted more information technology, experienced growth in service occupations and experienced wage polarization between 1980 and 2005. An alternative explanation for the decline in routine (manufacturing) work is import penetration from, or offshoring to, low cost countries. Autor et al. (2013b) exploited the commuting-zone variation in rising Chinese import competition between 1990 and 2007 to examine recent labour market trends in terms of employment and wages in

---

[2] It should be noted that Acemoglu and Autor (2010) describe their empirical exercise as "…highly preliminary – indeed, it is intended as an example of an empirical approach rather than a test of the theory…".

the US local labour markets. They showed that increasing import penetration explains one-quarter of the aggregate decline in US manufacturing employment, and reduces wages in local labour markets. Autor et al. (2013b) reported that highly trade-exposed and highly technology-exposed commuting zones are dissimilar in their occupational employment structures. Automation and trade therefore affect local labour markets differentially in terms of employment structure and wages.

The novelty of this paper is that we study the routinization hypothesis and implied job polarization using firm-level data. As noted in the discussion above, most (if not all) empirical studies of this issue have been conducted using different types of aggregate data (total economy, industries, locations). Our point of departure in this paper is to examine changes in the structure of labour demand at the firm level using linked data; this allows us to study the technology explanation for routinization and job polarization at the micro level, where actual labour demand decisions are made. In this way, we account for the effects of the compositional changes in product demand on the structure of employment and/or wages. The compositional changes in production confound aggregate studies if they are driven by factors other than technology (for example by offshoring or consumer preferences). Furthermore, we are able to examine routinization and polarization at the firm level using firm-specific adoption of specific Information and Communication Technologies (ICT) as explanatory variables for employment structure in firm-level regressions. Using firm-level ICT indicators allows us to establish a more direct link between observed occupational or educational changes and computer-based technological change than aggregate studies based only on time effects or industry-level aggregates. Using micro data, we are also able to perform decompositions of changes in educational and occupational employment (wage bill) shares into within firm, between firm, and entry-exit components to gain insight into the likely sources of these changes.

The paper is structured as follows. Section 2 describes the data. The results from decompositions are provided in Section 3. Section 4 discusses the specification of firm-level regressions and reports the estimates. Section 5 concludes.

## 2. Data

We use linked data that matches wage structure statistics to firm-level technology indicators. Harmonized Wage Structure Statistics (HWSS) data of Statistics Finland combines annual wage structure statistics data into harmonized panel data, where all important wage measures and classifications, such as industry and occupation, are consistent across years and sectors. The core of the annual wage structure statistics is the set of firm and individual level wage surveys of employer federations for their member firms. Statistics Finland augments these with samples of non-member firms and sectors not covered by the employer data. The harmonized data is available for the private sector for every year from 1995 onwards. In this paper, we exclude the recession years following the financial crisis after 2008, because our interest is in the structural effects of new technology.

Harmonization over time is needed because of the differences and changes in both collective wage contracts and classifications used over time and across sectors. The annual harmonization across different sectors takes into account the differences in wage concepts and compensation components used in different collective agreements. For example, hourly and monthly pay schedules are made fully comparable.

In the panel data education, occupation and industry variables are harmonized over time to the latest versions of standard classifications of Statistics Finland. Formal education is available from a comprehensive register of completed degrees. The industry classification of firms is available at the 5-digit level but used in the analyses at the 2-digit level. Occupation codes in the primary data are converted into international ISCO 2001 codes at the 5-digit level but used in the analyses at the 3-digit level. Unfortunately, it is not possible to completely harmonize some occupations for white-collar manufacturing workers over the break point 2001-2002 due to a classification change in the primary employer survey data. Hence, we either perform all our estimations using separate data before and after this break point using the periods 1995-2001 and 2002-2008, or focus only on the latter period.

This longitudinal data for the years 1995-2008 contains some 600 000–750 000 employees per year. Approximately 28 000 firms exist in the data for at least one year

during the period 1995-2008. Using sampling weights, these data are representative of the total private sector (with the exception of the smallest firms, which are exempt from the wage surveys of employer associations and Statistics Finland).

Our wage concept, the "hourly wage for regular working time" includes basic pay as well as various supplements for working conditions and performance-pay paid on a regular basis. It does not include overtime pay or one-off items, such as holiday and annual performance bonuses. In addition to wages we observe regular working hours per month for each worker in these wage statistics. Because the employer of each person is known, we are able to calculate the total number of employed persons and their total monthly wage bill for each firm in these statistics. Finally, we observe the education level and the occupation of each person, so we are able to disaggregate these measures according to education and occupation in order to examine the employment structure at the firm level.

The wage bill is divided into three education groups (low, intermediate and high) and three occupation groups (abstract, routine and service occupations). We have also constructed similar shares for hours worked and employed persons, but the results for these are similar to the wage bill, so we report only the latter. The low education group consists of those with basic compulsory education only. The high education group consists of those with a university level bachelor's degree or higher. The intermediate group consists of all degrees in between, i.e., from vocational to non-university higher degrees that usually involve two to four years of education. Our occupational grouping is an application of the classification presented in Acemoglu and Autor (2010) to the Finnish ISCO occupations. The abstract group includes managers, professionals and technicians; the routine group includes sales, clerical, production and operator's work; and services include occupations in protection, food preparation, building and grounds, cleaning and personal care and services.

Using this data we first illustrate the aggregate pattern of employment polarization in the Finnish private sector.[3] As Figure 1 shows, changes in employment shares by initial occupational wage deciles have been U-shaped in both the 1995-2001 and

---

[3] Earlier Finnish evidence on polarization at the aggregate level is provided in Asplund et al. (2011), Mitrunen (2013) and Böckerman et al. (2013).

2002-2008 periods, similar to the polarization pattern documented for the UK in Goos and Manning (2007). On the other hand, we find no indication of wage polarization in Finland in Böckerman et al. (2013), where wage growth increases almost linearly with initial wage levels.

Figure 1 here

We augment wage and employment data with firm-level variables for the adoption of Information and Communication Technology. Our data on the use of information technology and electronic commerce in firms originates from the Statistics Finland survey "Use of Information Technology in Enterprises" (ICT survey). The survey is a stratified random sample of firms in the sampling frame of the Business Register of Statistics Finland. It covers all large firms (100 employees or more) and a random sample of smaller firms with more than 5 employees. Because the HWSS data samples also for smaller firms, matching these two data sources reduces the number of firms in the linked data substantially. Therefore, our results are representative only for the larger firms. The variables in this survey describe the usage of ICT in firms, for example, Internet, intranet, broadband, home pages, services offered via home pages, electronic commerce, and electronic data interchange (EDI). The full list and explanations of variables is provided in Appendix 1.

The variables that describe various aspects of ICT are highly correlated as they measure the underlying characteristics of firms that affect the adoption of new technologies.[4] We use factor analysis to compress this information into latent factors, which we use as explanatory variables in our regressions. This alleviates multicollinearity and variance inflation in the estimated models, because the factors are orthogonal. We use the principal factors method and based on the eigenvalues three factors are adequate to describe the common variance of the ICT indicators. The cumulative variance explained is 71%. Factor loadings are documented in Appendix 2. We call Factor 1 EDI as it loads on variables related to the usage of electronic data interchange by the firm for various purposes (sending and receiving invoices or

---

[4] For the 16 ICT indicators only two pair-wise correlations are not statistically significant at the 5% level. The correlations are usually in the interval 0.30-0.50 (max 0.64) and statistically significant at the 1% level.

orders, or sending transport documents). Factor 2 loads on a large number of variables related to broadband or mobile access to the Internet, the firm's having a web site, and that the firm orders or sells through computer networks. This factor also loads on the firm having enterprise resource planning (ERP), but we call this the Internet factor for short. The third factor loads on two variables indicating whether the firm shares supply chain management (SCM) data with suppliers or customers.

Furthermore, we match task input measures at the occupation level from Goos, Manning and Salomons (GMS) (2010) into the wage data. These authors provide measures of the abstract, routine and service intensity for each 2-digit occupation. The measures are derived from the 2006 version of the Occupational Information Network (ONET) database, which provides the occupational attributes and characteristics of workers in 812 US SOC (Standard Occupational Classification) occupations. GMS (2010) manually convert these to the International Standard Classification of Occupations (ISCO), so we can match the GMS (2010) measures to our data at the 2-digit ISCO level. GMS (2010) use 96 ONET variables related to worker characteristics, worker requirements and work activities to create their measures for the Abstract, Routine and Service task intensities of different occupations. The task information is gathered from job incumbents, occupational analysts and occupational experts, who evaluate how important these task variables are in each occupation on a scale from 1 (not important at all) to 5 (extremely important). The 96 ONET variables are divided into one of the three groups of Abstract, Routine and Service tasks. Abstract task variables measure things like critical thinking and complex problem solving. Routine task variables measure things like manual dexterity, finger dexterity and operation monitoring. Service task variables measure assisting and caring for others, service orientation, and establishing and maintaining interpersonal relationships. The actual task measures are averages of these variables for each SOC occupation. They are converted to an ISCO occupation, using US employment in SOC cells as weights. Each task measure is normalized to have zero mean and unit standard deviation and they are available at the 2-digit ISCO level from GMS (2010).

Figure 2 documents the development of the task importance variables as weighted means across 3-digit occupations over time. The decline in importance of Routine tasks shows that employment in routine intensive occupations has decreased steadily

over the years. The increasing importance of both Abstract and Service tasks correspondingly implies increasing employment in abstract and service intensive occupations. Defining Routine intensity as the ratio of Routine importance to the sum of Abstract and Service Importance as in Goos et al. (2011), these trends imply a decreasing trend in Routine intensity. The pattern is consistent with the Routinization hypothesis of ALM (2003).

Figure 2 here

## 3. Decompositions for employment and wage bill shares

The main variables of interest in our examinations are the shares of the total wage bill by educational and occupational groups at the firm level, and their changes over time. To obtain proximate information about the possible sources of changes in employment structure, we present firm-level decompositions for the changes in wage bill shares.[5] This decomposition augments the Berman, Bound and Griliches (1994) industry-level decomposition to an unbalanced panel of firms with entry and exit.[6] The aggregate change in the wage bill share of a worker group defined by education or occupation (indexed by g) can be decomposed as follows:

$$\Delta P^A = \sum_i \Delta S_i \overline{P_i} + \sum_i \Delta P_i \overline{S_i} + w_t^N \left( P_t^N - P_t^S \right) + w_{t-s}^D \left( P_{t-s}^S - P_{t-s}^D \right)$$

where $P = \dfrac{E^g}{E}$, $\quad P_i = \dfrac{E_i^g}{E_i}$, $\quad S_i = \dfrac{E_i}{E}$, $\quad w_t^N = \dfrac{E_t^N}{E_t^A}$ and $w_{t-s}^D = \dfrac{E_{t-s}^D}{E_{t-s}^A}$.

$P$ is the aggregate share of the skill group in the total wage bill (denoted by E), $P_i$ is the corresponding share in firm $i$ ($i = 1,...,N$), $S_i$ is the share of firm $i$ in the aggregate wage bill, $\Delta$ indicates change over the period (t-s, t), and bar an average over the period's initial (t-s) and final year (t) values. Superscripts indicate the sums or shares

---

[5] We have also performed the decompositions for employment shares and working hour's shares, but the results are essentially similar to those we report for wage bill shares.
[6] See Vainiomäki (1999) for a detailed derivation of this decomposition augmented to include entry and exit effects.

for all firms (A), surviving firms (S), entering firms (N) and exiting firms (D). It can be shown that the entry and exit effects can also be written as

$$ENTRY = w_t^N \left( P_t^N - P_t^S \right) = \left( P_t^A - P_t^S \right)$$

$$EXIT = w_{t-s}^D \left( P_{t-s}^S - P_{t-s}^D \right) = \left( P_{t-s}^S - P_{t-s}^A \right).$$

These effects therefore depend on the deviation of the entering and exiting plant's average skill group shares with that of continuing plants. The entry effect is positive and greater the *higher the group's share in new plants* is compared to continuing plants ($P_t^N \geq P_t^S$). Similarly, the exit effect is positive and greater the *lower the group's share is in exiting plants* compared to continuing plants ($P_{t-s}^S \geq P_{t-s}^D$). However, it is noteworthy that the entry effect is also given by the simple difference between the group's aggregate wage bill share for all firms and continuing firms in the *final* year of the period. Similarly, the exit share is given by the simple difference in the shares for continuing firms and exiting firms in the *initial* year of the period.

The other two terms are standard from industry-level decompositions. The first sum is the *between firms effect*, which captures the shifts of employment (wage bill) between firms with different average wage bill shares. It is positive if the wage bill shifts towards firms which have a high wage bill share of the skill group in question. The second sum is the *within firms effect,* which captures changes in the wage bill share within each firm, weighted by the firm's average share of the total wage bill. The within component captures technological change within firms, the between component captures product demand changes across firms, and the entry/exit components reflect the demographic changes in firm population. The education groups in our decompositions are Basic, Intermediate and High, as explained above. The occupation groups are Abstract, Routine, and Services, following the definitions in Acemoglu and Autor (2010).

Table 1 reports the decomposition of changes in wage bill shares by education groups for the period 2002-2008.[7] We find that the within and total changes for the low and intermediate education groups are negative for this period. The respective changes for the highest educated are large positive. The entry component for the basic education group and the exit component for the highest educated are also positive. The between components for all education groups are minimal. These changes imply a rapid skill upgrading at the highest educational level during the 2000s. This overwhelmingly occurs within firms. The patterns are broadly consistent with job polarization in the sense that the intermediate education group loses shares, but in general these results show that the development has been "linear" with respect to education. The largest decline in shares occurs for the lowest educated and the largest increase for the highest educated.

Table 1 here

Table 2 shows the decompositions of change in wage bill shares by occupation groups. In contrast to education, both within and between components are important for occupational changes and affect in the same direction, except for the service occupations. We find that in total the routine occupation share declines and the abstract and service occupation shares increase, so that the total change is clearly consistent both with the routinization hypothesis and with job polarization. The entry and exit effects are small in general, but the entering firms are mildly service intensive and less intensive in abstract occupations. The shifts in production between different firms (the between component) seems to be more important in explaining polarization in the occupational shares than in educational shares. The shifts in production towards service-intensive firms and away from routine-intensive firms clearly contribute to the polarized pattern of total change. This suggests that changes in product demand may have a role in explaining the increase in the service occupations. However, for the abstract and routine occupations, the overwhelming majority of change occurs within existing firms, which is consistent with technological change being important in explaining the declining shares of routine occupations.

---

[7] We present these decompositions here as background for the regressions below only for the 2002-2008 period, because the ICT variables in regressions are available only for the 2000's.

Table 2 here

As a preliminary analysis of the effect of new technology on wage shares, we depict in Figure 3 the distributions of the within contributions for each occupation and education group partitioned by the median of each ICT factor. For each factor we find that a decrease in the routine share and an increase in the abstract share are more likely when the factor score is above the median. In contrast, the service share is not affected. For education groups the pattern is distinctly different. The distributions tend to tilt towards increasing shares for the highly educated and towards declining shares for both the intermediate and low educated. In the next section we quantify these effects using regressions.

Figure 3 here

## 4. Specifications and results from firm-level regressions

In order to examine the importance of the technology explanation to the shifts in the structure of labour demand, we estimate the following equations, at the firm level, for the wage bill shares of the education groups (g=Low, Middle, High), and occupation groups (g=Abstract, Routine, Service):

$$\Delta SHR_{it}^{g} = c^{g} + \beta_1^{g} ICT_{it} + \beta_2^{g} \ln SIZE_{it} + \beta_3^{g} X_{it} + u_{it}^{g}$$

where ICT denotes the adoption of new technology in each firm, lnSIZE is the log of firm size (employment), and X denotes a vector of other control variables (two-digit industry indicators in all models, and the firm-level average education and age in some models). Berman, Bound and Griliches (1994) showed that this type of share equation can be derived from a short-run trans-log cost function to examine the relative demand for different labour groups. However, this specification omits relative wages that we control with the use of two-digit industry indicators. Berman, Bound and Griliches (1994) also account for all permanent differences between industries. Michaels, Natraj and Van Reenen (2014) derive similar equations from a three-input

CES production function, which allows for ICT capital to substitute for the medium educated, and to complement the highly educated.

The polarization hypothesis implies that following the adoption of ICT (increase in ICT capital) the wage bill share of the highly educated (skilled) workers increases ($\beta_1^{High} > 0$) and the share of the middle educated (skilled) declines ($\beta_1^{Middle} < 0$), and the change in the share of the lowest educated is ambiguous. We perform a factor analysis on a large number of ICT indicators at the firm level to create underlying ICT factors, which we use as explanatory variables in our regressions instead of ICT capital, which is not available at the firm level.

We also perform similar regressions for occupational groups, i.e., Abstract, Routine, and Service occupations. Analogous to the treatment of education, the polarization hypothesis now implies that technological change increases the demand and therefore the wage bill share of the Abstract occupations ($\beta_1^{Abstract} > 0$), reduces the share of the Routine occupations ($\beta_1^{Routine} < 0$) and has an ambiguous effect on the Service occupations ($\beta_1^{Service} = ?$). The ambiguity for service occupations is based on the assumption that service occupations are technologically independent of ICT. Furthermore, the demand effects are general equilibrium effects, which are not manifest at the firm level in those exact firms that adopt new technologies. The decompositions above on the other hand incorporate the demand effects in all other firms. They indicated that the shifts in production between firms and the entry of new service intensive firms do play a role in explaining the rise in service occupations.

We address possible endogeneity issues below, but begin our analysis with OLS results for educational shares in column (1) in Table 3. The explanatory variables for the change in the wage bill share of each education group include the three ICT factors (our main interest) and the lagged level of the dependent (to control for the regression-towards-mean phenomenon), log firm size, log capital-output ratio and two-digit industries as control variables. Firm size controls for the fact that the

adoption of ICT technologies correlates positively with the size of the firms.[8] To the extent that firm size itself or omitted factors that are correlated with size (e.g., the quality of firm management) have effects on the structure of labour demand, omitting firm size would obscure the results for ICT factors.[9] Our results are based on comparing similar sized "ICT and non-ICT firms" (different levels of ICT factors). Similarly, we include capital-output ratio as a control for possibly differential effect of capital-skill complementarity on wage growth and employment development of different education and occupation groups. Our first-differenced equations already account for all firm fixed effects in levels,[10] but here we attempt to control for the possibly differential effects of capital intensity on the wage growth (e.g., increasing skill premiums) or on employment growth by education or occupation. We use the firm's average capital-output ratio over the whole period of 2002-2008 to measure its capital intensity. The two-digit industry indicators control for all other differences among industries, including differences in wage and employment growth from overall technological progress between industries. It should be noted that any association between a firm's ICT factors and employment structure that we find prevails across firms within two-digit industries.

The pattern of OLS results in Table 3 is more supportive of SBTC than of the routinization hypothesis. All the ICT factors are positively associated with the increase in high skilled wage bills and negatively associated with the share of both medium and basic education groups, without much difference between these latter groups. The fact that the intermediate educated are not strongly negatively affected by ICT factors is contradictory to the predictions of job polarization. The regression coefficients are also numerically significant. We find that the wage bill share of the highly educated increases 1.7 %-points faster in firms that have one standard deviation (0.43) higher score for the EDI factor, 1.3 %-points faster for the Internet factor (s.d. 0.25), and 0.9 %-points faster for the SCM factor (s.d. 0.46). These effects and their sum, 3.6%, are economically large compared to the total increase of 7.4% of

---

[8] The (weighted) correlations of individual ICT indicators with log firm size are all statistically significant at the 1% level and in the interval 0.10-0.44.

[9] The effects for both ICT factors or for the original ICT indicators (not reported) are essentially similar but the coefficients are smaller and less significant if firm size is not included in the equation as a control variable.

[10] Our estimating equation can be interpreted as a first differences version of a level equation for wage bill shares. Then, firm fixed effects in levels of these shares are controlled by differencing.

the high-educated wage bill share over the 2002-2008 period as reported in Table 1. The sum of one standard deviation effects of ICT factors (and the total change) for Intermediate education is -2.0% (-2.8%) and for Basic education -2.4% (-4.6%). The ICT factors together explain about half of the increase in the high education share and the decline in the low education share.

Table 3 here

The results for occupational shares are reported in Table 4. In contrast to education groups, the OLS results in columns (1) are now supportive of the occupational routinization hypothesis that is associated with the ICT factors. The EDI and Internet factors are positively and significantly related to the change in the share of Abstract occupations, and negatively and significantly related to the share of routine occupations. The pattern is similar for the SCM factor but all SCM coefficients are insignificant. The service share is mostly independent of ICT at the firm level, with only the EDI factor gaining significance in OLS regressions. The magnitude of the coefficients is again economically significant compared to the mean of the dependent variable, or to the total change in shares in Table 2. A one standard deviation increase in the EDI factor increases the abstract occupation share by 2.4 %-points and reduces the routine occupation share by 1.8 %-points. Corresponding effects for the Internet factor are 2.1 %-points and −1.9 %-points using the OLS coefficients in columns (1). One standard deviation changes in ICT factors would explain essentially all of the observed changes in occupational wage bill shares. The sum and total changes are 5.4% (4.5%) for the abstract group, -4.5% (-5.7%) for the routine group and -1.0% (1.2) for the service occupations.

The results for education and occupational groups are notably different. This pattern is explained by the shares of education groups in different occupations. We find that the low educated are concentrated in routine occupations (72%), and the highly educated in abstract occupations (81%); see Table A2 of Appendix 3. However, the occupation structure of medium educated is more varied; 60% work in routine occupations but over a quarter (26%) is found in abstract occupations. Taking into account the fact that the number of intermediate educated workers is over three times the number of highly educated workers there are about the same absolute number of medium and

15

highly educated workers in abstract occupations. The fact that medium educated work to a substantial extent in abstract occupations provides an explanation for our apparently contradicting results for educational and occupational groups. To illustrate the importance of the occupational distribution of different education groups on our results we compute predicted educational effects based on occupational effects. Table A3 documents the predicted values on the within contributions and the effect of ICT for each education group using the within contributions in Table 2 and the estimated OLS coefficients for ICT factors for each occupational group in Table 4 (and the shares in Table A2). The structure of these predicted effects is similar to the estimated (OLS) effects for education groups. These back-of-envelope calculations show that the results for education and occupational groups are consistent with each other, as the educational results reflect the occupational effects and occupational distribution of each education group. Overall, our interpretation from this is that our results support the routinization hypothesis and that ICT contributes to routinization. Our educational results reflect the fact that a substantial share of medium educated work in abstract occupations, which mitigates the effect of routinization on the intermediate educated as a group.

Table 4 here

Instead of factor scores, we have used summary indicators of ICT variables that resemble the factors used in our analyses (not reported). We created three summary variables as means of individual ICT items that have the largest loadings on each of the three factors. The OLS results for these summary indicators are qualitatively similar (including statistical significance) to OLS results for the factor scores. We have also included all ICT indicators both individually and jointly as explanatory variables without summarizing them. Individually many ICT indicators are significant in abstract (high education) and routine (low education) share equations, but only a few items are statistically significant at least at the 10% level when they are included jointly. The strongest and most consistent effects relate to the firm using electronic data interchange in receiving invoices. This increases the shares of abstract occupations and of the highly educated, and reduces the shares of routine occupations and of intermediate and low educated statistically significantly at the 1% level. In addition, the existence of a firm home page and broadband connection increase the

abstract occupation share and reduce the routine occupation share significantly at least at 10% level.

As already noted above, our estimating equations are essentially first-differenced versions of the levels equations for wage bill shares, so any endogeneity related to the unobserved firm fixed effects in the levels equations is eliminated from our results. Other sources of endogeneity bias, however, remain in our models. First, measurement error in explanatory variables causes the standard attenuation bias. Second, the lagged level of a wage bill correlates by definition with its change. Third, there is the possibility of reverse causality in ICT, i.e., shocks to the firm's wage bill shares causing firms to change their investments in new technology (adoption of ICT).

Regarding the lagged level of the dependent variable we use the corresponding wage bill share as an instrument (i.e., year 2001 value as an instrument for year 2002 value). For the ICT factors, we use predicted factor scores as instruments for the observed factor scores. To obtain the predicted scores, we first run unweighted linear probability models for each of the ICT indicators that were included in the factor analysis. The firm-level explanatory variables in these regressions are log firm size, lagged average education years and average age of firm's employees (year 2002 values), log capital-output ratio as mean over the period 2002-2008, the period means of computer investments relative to sales and as share of fixed investments in machinery, the period mean of computer expenditures relative to sales, the indicator for foreign ownership being over 50 percent and export/import indicators. As explanatory variables in the probability models, we also include average R&D intensity, three task-importance variables, average education years and age of employees at the 3-digit industry level and one-digit industries. We expect the above variables to relate to ICT adoption in firms. For example, the industry-level values of task intensities reflect firms' *potential* to replace routine tasks with ICT or potential to benefit from ICT in performing analytical tasks. The predicted factor scores are then calculated as weighted averages of predicted values from these models; this is done using the same weights as were used in calculating the observed factor scores from observed ICT indicators. In the second stage of IV estimation we use these predicted factor scores, the routine-intensity at the firm level and 3-digit industry level averages

of observed factor scores as instruments for the ICT factors.[11] Task intensities and other industry-level variables should be exogenous with respect to idiosyncratic firm-level shocks to the wage bill, but for the other variables we rely on the over-identification tests for the instruments to check the validity of the exclusion restrictions.

The results of our instrumental variables estimations are presented in columns (2) in Tables 3 and 4 for the education groups and the occupation groups, respectively. The importance of the lagged level of the dependent variable in all models confirms the regression-towards-the-mean phenomenon, but it also implies that the educational employment structure changes quite slowly even over the six-year period. The coefficients for the lagged dependent variable estimate ($\rho^6-1$), where $\rho$ is the AR(1) parameter for the level of wage bill share. The IV estimates in column (1) of Table 5 imply that $\rho^6$ is approximately 0.8 for the highly and medium educated, and approximately 0.6 for the low education group, which imply high persistence in educational wage bill shares at the firm level.

For the educational groups the coefficients for ICT factors from IV estimations mostly lose statistical significance; see column (2) of Table 3. The coefficients increase in absolute value compared to OLS results as expected due to measurement error bias, but the coefficients obtain, in general, implausibly large values and only a few coefficients remain significant. The standard errors of IV estimates are considerably inflated compared to the OLS standard errors, although the instrument set seems to have reasonable predictive power according to the first-stage F-tests. The instruments for the lagged level of the wage bill have strong explanatory power in the first stage. The over-identifying restriction test passes for the high education group (at the 5% level) and the medium education group, but not for the low education group.

The IV results for the occupational groups in Table 4 retain the significance for the Internet factor in column (2), but there are also some changes in significance of ICT

---

[11] Michaels et. al. (2014) have used the proportion of routine manual tasks in the industry as an instrument for future ICT growth. They argue that high routine industries were most likely to be affected by falling ICT prices and therefore to adopt ICT. We use routine intensity at the firm level as an instrument for ICT adoption.

factors. The EDI factor becomes insignificant in IV results whereas the SCM factor becomes statistically significant compared to the OLS results. However, our primary results remain intact. The abstract occupation share is positively related to (some) ICT factors, and the routine occupation share is correspondingly negatively affected by the ICT factors. The absolute size of the regression coefficients increases consistently with measurement error attenuation in OLS, but this may also be a sign of weak instruments, although the first-stage F-tests for additional instruments exceed 10 for two of the factors. The instruments pass the over-identifying restriction test for the abstract and routine occupation groups, but not for the service occupations.

In summary, our IV results confirm the high persistence in educational and occupational employment structure even when the lagged level of the dependent variable is instrumented to account for the reverse correlation and measurement error. Our strategy of instrumenting ICT factors confirms our main result, namely that the routinization and polarization of employment is related to the adoption of new ICT technologies by firms.

## 5. Conclusions

Using the new Harmonized Wage Structure Statistics (HWSS) data of Statistics Finland, we first document patterns of employment polarization in the Finnish private sector labour market. We establish that there has been considerable job polarization at the aggregate level. The structure of changes in the employment shares by initial occupational wage deciles is distinctly U-shaped. Matching task measures for individuals at the occupational level, we also find that the changes in aggregate occupational structure are consistent with the routinization hypothesis, i.e., aggregate average routine intensity of employment declines in a trend-like manner over the period 1995-2008.

Our paper introduces a new feature, however, in its use of decompositions and regression analyses that test for the routinization hypothesis and job polarization at the firm level, instead of the aggregate, industry or local levels as in prior studies. Using firm-level approach, we are able to study routinization and job polarization at the micro level, where actual labour demand decisions are made.

Our firm-level decompositions for the changes in wage bill shares indicate likely reasons behind these changes. The decompositions for educational groups show that changes in education shares are towards more educated groups in a "linear" fashion with respect to education level. This pattern is consistent with the standard SBTC model. The total change occurs overwhelmingly within firms, which is suggestive of a technological cause of these changes. As for changes in occupational shares, we find that the increase in abstract and the decline in routine occupations also occur substantively within existing firms. However, our decompositions also show that for the service occupations the shifts in production towards service-intensive firms and the entry of new service-intensive firms is relatively more important than for other occupation groups. This pattern indicates that changes in product demand may have a role in explaining the increase in service occupations, which produces a polarized pattern of total changes for occupation groups, that is increasing abstract and service shares and declining routine share.

Furthermore, we examined the technology-based explanations for routinization and job polarization at the firm level by using firm-level indicators of ICT usage as explanatory variables in the firm-level regressions. We first perform a factor analysis on a large number of indicators for ICT adoption at the firm level to obtain factor scores for three ICT factors. We then use these factor scores as explanatory variables in regressions for changes in wage bill shares of different education and occupation groups. OLS regressions show that these ICT factors are associated with increases in the demand for highly educated workers and reductions in the demand for the low educated, while the intermediate education group is independent of ICT. In regressions for occupation groups, we find that ICT factors are associated with increases in abstract occupation shares and declines in routine occupation shares. These occupational patterns support the routinization hypothesis at the firm level. Since routinization is the main mechanism producing polarization, these results are also consistent with job polarization. The service occupation share is independent of ICT at the firm level (as hypothesized), and the increasing aggregate share of services relates to the demand effects, as noted above. A possible explanation for the apparent contradiction in our results with respect to educational and occupational employment structure is that the assumed substitutability of middle educated workers for ICT technologies does not hold, but, in our analysis, it does prevail for occupations

20

classified as routine. Alternatively, our educational results reflect the fact that a substantial share of medium educated work in abstract occupations, which mitigates the effect of routinization on the intermediate educated as a group.

Finally, our strategy of instrumenting ICT factors confirms our main result, namely that the adoption of new ICT technologies by firms is (causally) related to the routinization and polarization of employment. Furthermore, our IV results confirm a high level of persistence in educational and occupational employment structures even when the lagged level of the dependent variable is instrumented to account for the reverse correlation and measurement error.

**References**

Acemoglu, D. (2002). 'Technical change, inequality, and the labor market'. *Journal of Economic Literature*, 40: 7–72.

Acemoglu, D. and Autor D. (2010). 'Skills, tasks and technologies: Implications for employment and earnings'. In O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics,* Volume 4B. Amsterdam: North-Holland, pp. 1043–1171.

Asplund, R., Barth E., Lundborg, P. and Nilsen, K. M. (2011). 'Polarization of the Nordic labour markets'. *Finnish Economic Papers*, 24: 87–110.

Autor, D. and Dorn, D. (2013). 'The growth of low-skill service jobs and the polarization of the US labor market'. *The American Economic Review,* 103: 1553–1597.

Autor, D., Dorn, D. and Hanson G.H. (2013a). 'The geography of trade and technology shocks in the United States'. *The American Economic Review,* 103:220–225.

Autor, D., Dorn, D. and Hanson G.H. (2013b). 'The China Syndrome: Local labor market effects of import competition in the United States'. *The American Economic Review,* 103: 2121–2168.

Autor, D., Levy, F. and Murnane, R. J. (2003). 'The skill content of recent technological change: An empirical exploration'. *Quarterly Journal of Economics,* 118: 1279–1333.

Autor, D., Katz, L. and Kearney, M. (2006). 'The polarization of the U.S. labor market'. National Bureau of Economic Research, Working Paper No. 11986.

Autor, D., Katz, L. and Kearney, M. (2008). 'Trends in U.S. wage inequality: Re-assessing the revisionists'. *Review of Economics and Statistics,* 90: 300–323.

Berman, E., Bound, J. and Griliches, Z. (1994). 'Changes in the demand for skilled labor within US manufacturing: Evidence from the annual survey of manufactures'. *Quarterly Journal of Economics,* 109: 367–397.

Bloom, N., Draka, M. and Van Reenen, J. (2011). 'Trade induced technical change? The impact of Chinese imports on innovation, IT and productivity'. Centre for Economic Performance, Discussion Paper No. 1000.

Bound, J. and Johnson, G. (1992). 'Changes in the structure of wages in the 1980s: An evaluation of alternative explanations'. *The American Economic Review,* 83: 371–392.

Böckerman, P., Laaksonen, S, and Vainiomäki, J. (2013). 'Is there job polarization at the firm level?'. Tampere Economic Working Papers, No. 91, October 2013.

Card, D. and DiNardo J. E. (2002). 'Skill-biased technological change and rising wage inequality: Some problems and puzzles'. *Journal of Labor Economics,* 20: 733–782.

Goos, M. and Manning A. (2007). 'Lousy and lovely jobs: The rising polarization of work in Britain'. *Review of Economics and Statistics,* 89: 118–133.

Goos, M., Manning, A. and Salomons, A. (2010). 'Explaining job polarization in Europe: The roles of technology and globalization'. Centre for Economic Performance, Discussion Paper No. 1026.

Jenkins, S. P. (1995). 'Did the middle class shrink during the 1980s? UK evidence from kernel density estimates'. *Economics Letters*, 49: 407–413.

Michaels, G., Natraj, A. and Van Reenen, J. (2014). 'Has ICT polarized skill demand? Evidence from eleven countries over 25 years'. *Review of Economics and Statistics*, 96: 60–77.

Mitrunen, M. (2013). Työmarkkinoiden polarisaatio Suomessa. Government Institute for Economic Research, VATT Mimeo 33.

Vainiomäki, J. (1999). 'Technology and skill upgrading: Results from linked worker-plant data for Finnish manufacturing'. In J. Haltiwanger, J. Lane, J. Spletzer, J. Theeuwes and K. Troske (eds.), *The Creation and Analysis of Employer-Employee Matched Data.* Amsterdam: Elsevier/North-Holland, pp. 115–145.

Figure 1. Employment polarization.
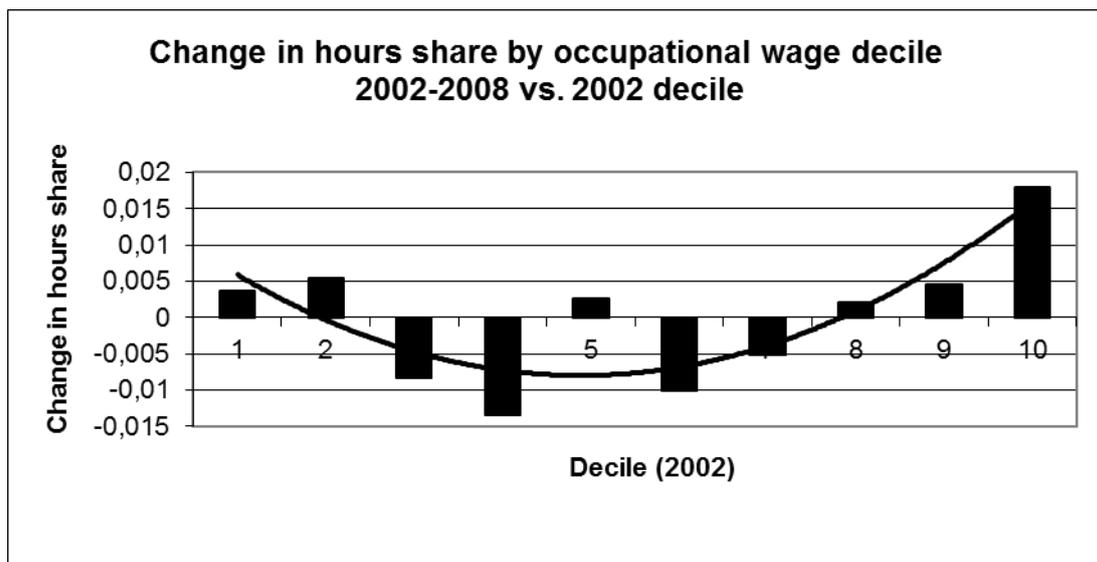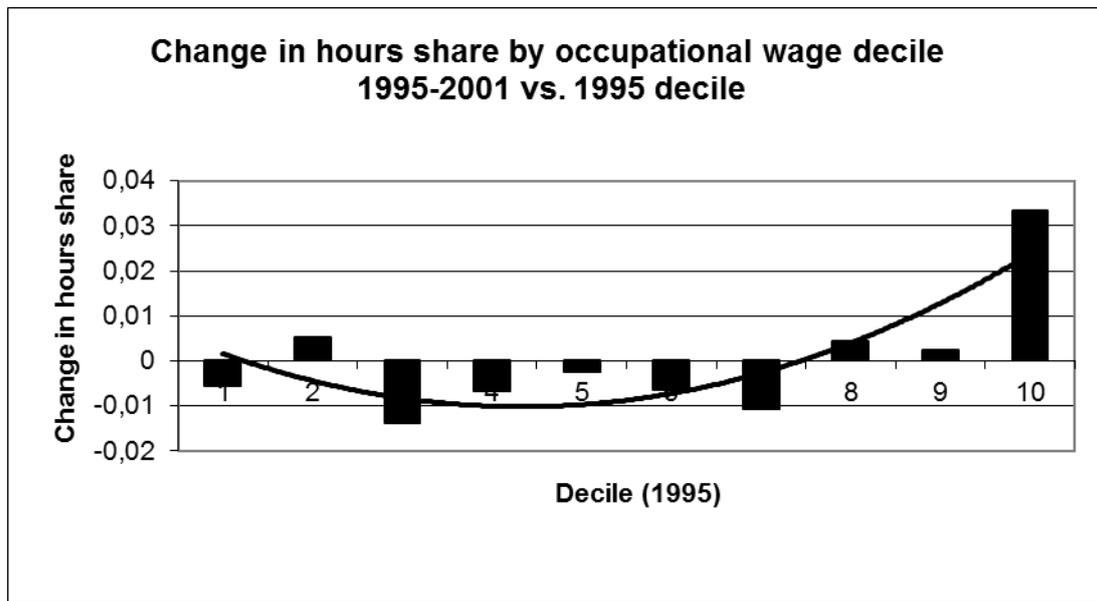
Figure 2. Development of Abstract, Routine and Service task importance.
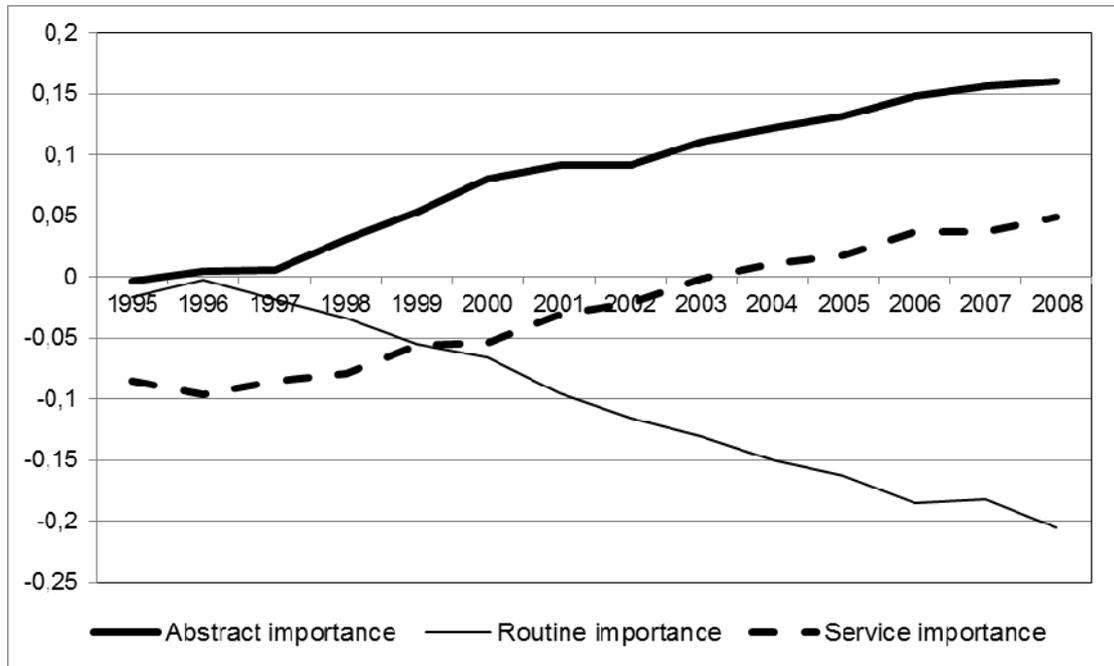


*Note:* Hours weighted means of task variables across 3-digit occupations.

Figure 3. Distributions of within contributions by ICT factors.
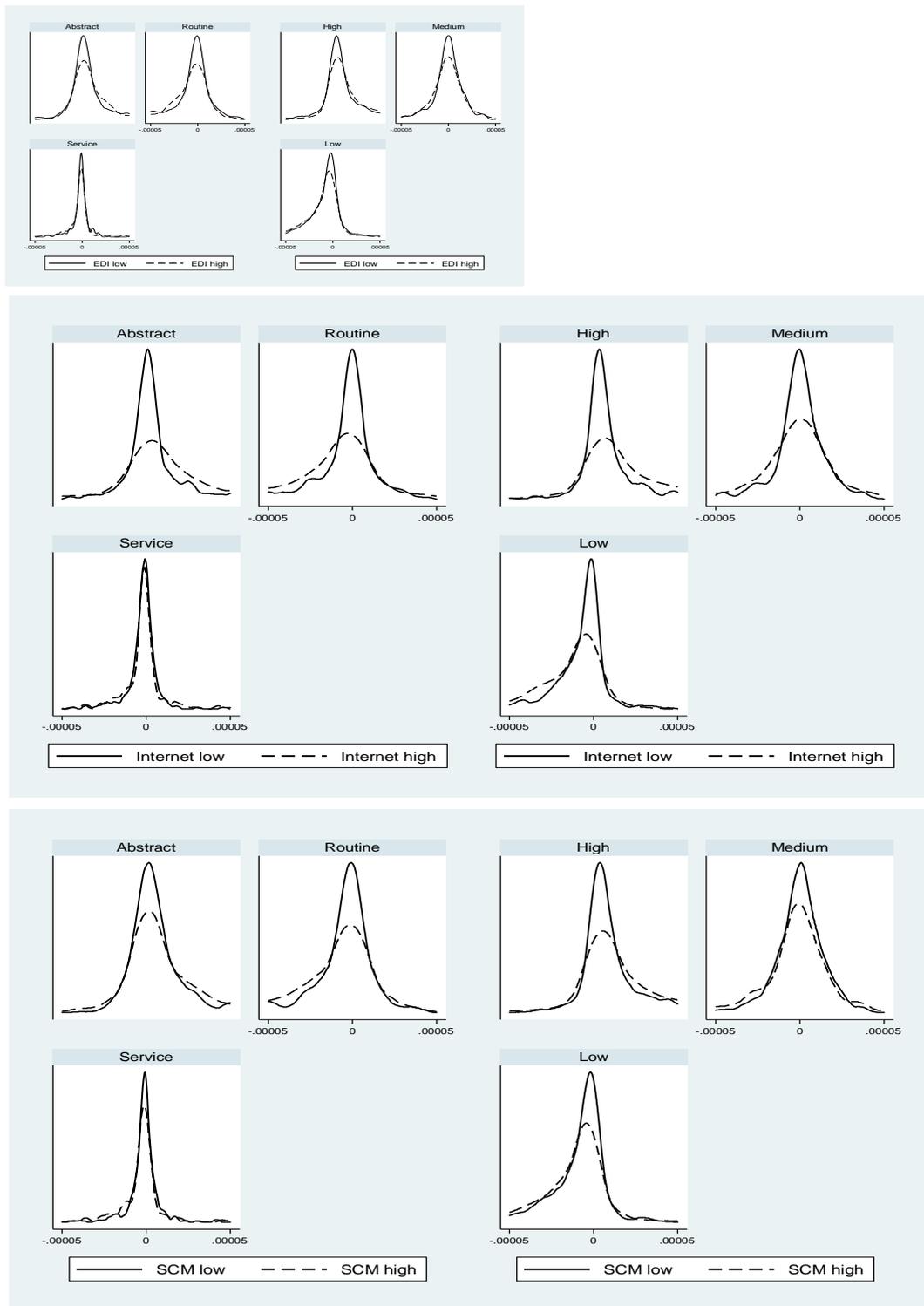
Table 1. Decompositions for wage bill share by education.

| Period | Education | Within | Between | Entry | Exit | Total | Share (2008) |
|---|---|---|---|---|---|---|---|
| 2002-2008 | Basic | -0.051 | -0.001 | 0.012 | -0.006 | -0.046 | 0.138 |
| 2002-2008 | Intermediate | -0.018 | 0.001 | -0.004 | -0.007 | -0.028 | 0.571 |
| 2002-2008 | High | 0.068 | 0.000 | -0.007 | 0.013 | 0.074 | 0.291 |

Table 2. Decompositions for wage bill share by occupation.

| Period | Occupation | Within | Between | Entry | Exit | Total | Share (2008) |
|---|---|---|---|---|---|---|---|
| 2002-2008 | Abstract | 0.045 | 0.004 | -0.004 | 0.000 | 0.045 | 0.464 |
| 2002-2008 | Routine | -0.038 | -0.019 | -0.001 | 0.001 | -0.057 | 0.450 |
| 2002-2008 | Service | -0.007 | 0.015 | 0.005 | -0.001 | 0.012 | 0.086 |

Table 3. Regressions for the change in wage bill shares by education group.

| Variable | High | | Medium | | Low | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (1) | (2) | (1) | (2) |
| ICT Factors | OLS | IV | OLS | IV | OLS | IV |
| EDI | 0.040*** | 0.023 | -0.020* | 0.017 | -0.024*** | -0.046 |
| | (0.011) | (0.064) | (0.011) | (0.060) | (0.006) | (0.041) |
| Internet | 0.050*** | 0.086 | -0.029* | -0.149** | -0.031** | -0.075 |
| | (0.018) | (0.069) | (0.016) | (0.062) | (0.012) | (0.067) |
| SCM | 0.019* | 0.070 | -0.008 | -0.096* | -0.013** | -0.085** |
| | (0.011) | (0.061) | (0.009) | (0.056) | (0.006) | (0.035) |
| | | | | | | |
| Lagged dependent (level) | -0.215*** | -0.207*** | -0.260*** | -0.243*** | -0.431*** | -0.396*** |
| | (0.045) | (0.046) | (0.052) | (0.051) | (0.068) | (0.069) |
| | | | | | | |
| ln(Size) | -0.004 | -0.008 | -0.001 | 0.008 | 0.004** | 0.014** |
| | (0.003) | (0.009) | (0.003) | (0.008) | (0.002) | (0.006) |
| | | | | | | |
| ln(Capital-output) ratio | 0.004 | 0.003 | -0.0004 | 0.001 | -0.005* | -0.004 |
| | (0.004) | (0.004) | (0.003) | (0.004) | (0.003) | (0.003) |
| | | | | | | |
| F-test/Chi$^2$ (industry) | 23.0 | 497.2 | 8.65 | 176.5 | 85.4 | 563.5 |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| | | | | | | |
| N | 1110 | 1110 | 1110 | 1110 | 1110 | 1100 |
| | | | | | | |
| First stage F-tests: Lagged dependent EDI Internet SCM | | 998.3 24.1 12.8 6.7 | | 263.6 23.5 12.4 5.9 | | 384.3 25.5 12.1 7.7 |
| Over-identifying restrictions | | 8.33 (0.08) | | 2.01 (0.73) | | 12.7 (0.01) |
| Dependent Mean (weighted) | 0.052 | 0.052 | -0.003 | -0.003 | -0.049 | -0.049 |

*Notes:* Weighted by the product of sampling weight and total working hours of the firm. Robust t-values reported. The dependent variable is a six-year difference over the period 2002-2008.
In columns (2) the level of the lagged dependent variable in 2002 is instrumented with its value in the previous year (2001) and the ICT factors are instrumented with three-digit industry values for factor scores as well as with firm level predicted values for the factor scores and firm routine intensity (see main text for more details).

Table 4. Regressions for the change in wage bill shares by occupation group.

| Variable | Abstract | | Routine | | Service | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (1) | (2) | (1) | (2) |
| ICT Factors | OLS | IV | OLS | IV | OLS | IV |
| EDI | 0.056*** | 0.176 | -0.043* | -0.117 | -0.011*** | -0.049 |
| | (0.017) | (0.114) | (0.016) | (0.098) | (0.004) | (0.035) |
| Internet | 0.083*** | 0.336** | -0.075** | -0.240** | -0.009 | -0.105 |
| | (0.029) | (0.147) | (0.030) | (0.119) | (0.010) | (0.069) |
| SCM | 0.021 | 0.277** | -0.016 | -0.194** | -0.006 | -0.093** |
| | (0.016) | (0.121) | (0.016) | (0.097) | (0.004) | (0.044) |
| | | | | | | |
| Lagged dependent (level) | -0.243*** | -0.253*** | -0.267*** | -0.253*** | -0.166 | -0.337*** |
| | (0.048) | (0.048) | (0.042) | (0.045) | (0.136) | ((0.115) |
| | | | | | | |
| ln(Size) | -0.015*** | -0.053*** | 0.010** | 0.036*** | 0.004*** | 0.018** |
| | (0.005) | (0.019) | (0.005) | (0.015) | (0.002) | (0.008) |
| | | | | | | |
| ln(Capital-output) ratio | 0.021*** | 0.016** | -0.019*** | -0.016** | -0.003* | -0.0003 |
| | (0.007) | (0.008) | (0.007) | (0.007) | (0.002) | (0.002) |
| | | | | | | |
| F-test/Chi$^2$ (industry) | 21.9 | 379.6 | 5.2 | 135.3 | 5.0 | 175.9 |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| | | | | | | |
| N | 1110 | 1110 | 1110 | 1110 | 1110 | 1110 |
| | | | | | | |
| First stage F-tests: Lagged dependent EDI Internet SCM | | 200.4 23.6 12.1 6.2 | | 281.7 23.4 12.2 6.1 | | 14.6 23.9 12.2 6.2 |
| Over-identifying restrictions | | 6.62 (0.16) | | 4.42 (0.35) | | 16.7 (0.002) |
| Dependent Mean (weighted) | 0.031 | 0.031 | -0.030 | -0.030 | -0.002 | -0.002 |

*Notes:* Weighted by the product of sampling weight and total working hours of the firm. Robust t-values reported. The dependent variable is a six-year difference over the period 2002-2008.
In columns (2) the level of the lagged dependent variable in 2002 is instrumented with its value in the previous year (2001) and the ICT factors are instrumented with three-digit industry values for factor scores as well as with firm level predicted values for the factor scores and firm routine intensity (see main text for more details).

Appendix 1. Variables and definitions for terms used in the ICT survey (Source: Statistics Finland)

Variables

| | |
|---|---|
| PC | Firm uses computers |
| INTER | Firm has Internet connection |
| WEB | Firm has a website |
| EPURCH | Firm orders through computer networks (websites or EDI) |
| ESALES | Firm sells through computer networks (websites or EDI; not email orders) |
| BROAD | Firm has broadband (ADSL, SDSL, cable modem; faster than ISDN) |
| MOB | Firm has mobile access to internet (laptop, mobile phone; 3G, 4G or slower) |
| ERP | Firm has Enterprise Resource Planning system (ERP program); 2006 onwards |
| SCMT | Sharing Supply Chain Management (SCM) data with suppliers using computer networks (inc. Internet); regular exchange of information on demand forecasts, inventories, production plans, delivery progress, product planning; available 2006-2009 |
| SCMA | Sharing SCM data with customers; see above; available 2006-2009 |
| CRMINF | Management and sharing of customer information with other business functions within firm; 2006 onwards |
| CRMANA | Firm analyses customer information for marketing purposes (price setting, promoting sales, choosing delivery channels); 2006 onwards |
| AUTTIED | Electronic data interchange used |
| AUTLASVA | Receiving e-invoices using electronic data interchange; available 2007-2009 |
| AUTLASLA | Sending e-invoices using electronic data interchange; available 2007-2009 |
| AUTKULJ | Using electronic data interchange in sending transport documents; available 2007-2009 |

Definitions of some terms

Broadband

Broadband is a telecommunications connection with a capacity of at least 256 Kbps. In the statistics on the use of information technology in enterprises, broadband has in practice been defined through the type of technology used in the connection as either DSL (e.g., ADSL) or other broadband connection (faster than a traditional telephone modem or ISDN).

E-invoice

An e-invoice is an electronic invoice constructed according to a generally used message format, whose data can be handled and interpreted automatically. E-invoices are transmitted via a telecommunications service provider or a bank. E.g, Finvoice, eInvoice, TEAPPSXML, PostiXML.

E-mail invoice

An e-mail invoice is an invoice sent as a pdf-file attached to an e-mail.

EDI

EDI (Electronic Data Interchange) is a procedure by which information located in an enterprise's data system is used to produce a specified data flow that is transmitted electronically to a receiving enterprise, where it is directly incorporated into the data system (e.g., order, payment order for invoice, price list or product catalogue).

EDI commerce

EDI commerce is electronic commerce that takes place between enterprises through the medium of EDI.

EDI invoice

An EDI invoice is an electronic invoice in machine code according to the EDI structure standards. EDI invoices are often sent via a telecommunications service provider.

Electronic invoice

An electronic invoice is an invoice transmitted in electronic form: an EDI invoice, an e-invoice, an e-mail invoice or some other electronic invoice. Payments entered by a customer into an online banking system or direct debit are not electronic invoices.

Homepage

A homepage here is defined as an enterprise's own Internet homepages or its section in the homepages of a group. Homepages do not refer, for example, to publication of an enterprise's contact details on various company and address lists.

Internet sales

Internet sales are communications between a person and a data system. Online shopping, as defined here, is an order placed by completing and sending a ready-made electronic form on the Internet and shopping in actual Internet shops. Orders placed with a standard email message are not defined as online shopping. Purchases made on an extranet subject to the same conditions are also counted as Internet sales.

Online shopping

Online shopping is the ordering of goods and services via a computer network, regardless of payment or delivery method.

Appendix 2. Factor analysis of ICT variables.

The variables in the ICT survey are (mostly) binary indicators. Factor analysis assumes that the observed variables are continuous (or at least ordinal), because they are modelled as linear combinations of continuous latent factors. One could proceed with a factor analysis for binary variables, which specifies a logistic link function between observed indicators and latent factors. Alternatively, one can continue to use ordinary factor analysis but base it on a tetrachoric correlation matrix. Tetrachoric correlations are estimates for the correlation coefficient of latent bivariate normal distribution based on observed binary variables. The results from factor analysis using tetrachoric correlations are usually similar to those from the binary factor analysis. The difference is mainly that tetrachoric correlations treat binary variables as incompletely observed underlying variables rather than observed items of latent factors.

We used tetrachoric correlations for the 16 binary indicators of the usage of different aspects of ICT in the firms in the ICT survey. Using orthogonal rotation and the principal factors method yielded the following rotated factor loadings for three factors with eigenvalues greater than one.

Table A1. Rotated factor loadings (pattern matrix) and unique variances

| Variable | Factor 1 | Factor 2 | Factor 3 | Uniqueness |
|---|---|---|---|---|
| Firm has website | 0.1322 | 0.8854 | -0.1564 | 0.1742 |
| Firm has broadband | 0.1354 | 0.7020 | 0.3243 | 0.3837 |
| Firm has mobile access to internet | 0.2011 | 0.7382 | 0.2517 | 0.3513 |
| Firm orders through computer networks | 0.2590 | 0.5490 | 0.3133 | 0.5333 |
| Firm sells through computer networks | 0.4405 | 0.5793 | 0.1155 | 0.4570 |
| Firm has Enterprise Resource Planning | 0.2516 | 0.6405 | 0.3845 | 0.3786 |
| Shares SCM data with suppliers | 0.3013 | 0.0714 | 0.8988 | 0.0963 |
| Shares SCM data with customers | 0.2640 | 0.2855 | 0.8003 | 0.2083 |
| Shares customer information (within firm) | 0.1774 | 0.6967 | 0.3841 | 0.3356 |
| Analyses customer information for marketing | 0.1907 | 0.6689 | 0.4096 | 0.3484 |
| Receives e-invoices via electronic data exchange | 0.7500 | 0.3795 | 0.0792 | 0.2872 |
| Sending e-invoices via electronic data exchange | 0.7880 | 0.4563 | 0.0414 | 0.1692 |
| Electronic data interchange used | 0.9605 | 0.1052 | 0.2008 | 0.0260 |
| Receiving orders | 0.7872 | 0.2495 | 0.2425 | 0.2593 |
| To suppliers | 0.7729 | -0.1093 | 0.3617 | 0.2598 |
| Sending transport documents | 0.7375 | 0.1138 | 0.3758 | 0.3020 |

Source: ICT panel 2001-2009 (information refers to the end of previous year, i.e., 2000-2008). Variables are indicators that the firm has or utilizes the technology indicated by the variable name.

Factor 1: Loads on using automated data exchange for sending and receiving invoices and orders and sending transport documents. We call this factor EDI (for electronic data interchange).

Factor 2: Loads on firm having a website and access to internet, selling and placing orders via computer networks (website or EDI). It also loads on the firm having a special program (CRM) for sharing and analysing customer information within the firm, and the firm having Enterprise Resource Planning (ERP). We call this a general Internet factor.

Factor 3: Loads on sharing supply chain management (SCM) data with suppliers or customers via computers (demand forecasts, inventory levels, production plans, deliveries, product planning information); we call this factor SCM.

Appendix 3. Information on education and occupational groups.

Table A2. Occupational composition of education groups.

|  | Low | Medium | High |
|---|---|---|---|
| Abstract | 0.14 | 0.26 | 0.81 |
| Routine | 0.72 | 0.60 | 0.15 |
| Service | 0.15 | 0.13 | 0.04 |
|  | 1 | 1 | 1 |

Note: Based on the average distribution for the years 2002 and 2008.

Table A3. Predicted and observed values for the within contributions and ICT effects.

|  | Low | Medium | High |
|---|---|---|---|
| Pred. within contr. | -0.022 | -0.012 | 0.030 |
| Observed | -0.051 | -0.018 | 0.068 |
|  |  |  |  |
| Pred. EDI | -0.025 | -0.013 | 0.038 |
| Obs. EDI (OLS) | -0.024 | -0.020 | 0.040 |
|  |  |  |  |
| Pred. Internet | -0.044 | -0.025 | 0.055 |
| Obs. Internet (OLS) | -0.031 | -0.029 | 0.050 |
|  |  |  |  |
| Pred. SCM | -0.010 | -0.005 | 0.014 |
| Obs. SCM (OLS) | -0.013 | -0.008 | 0.019 |

Note: Predicted values for the within contributions and the effect of ICT for each education group are obtained by weighting the within contributions in Table 2 and the estimated OLS coefficients for ICT factors for each occupational group in Table 4 by the occupational employment shares in Table A2 for each education group. The observed values for education groups are from Tables 1 and 3.