

**This is an electronic reprint of the original article.
This reprint *may differ* from the original in pagination and typographic detail.**

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

Title: How does the age structure of worker flows affect firm performance?

Year: 2016

Version:

Please cite the original version:

Ilmakunnas, P., & Maliranta, M. (2016). How does the age structure of worker flows affect firm performance?. *Journal of Productivity Analysis*, 46 (1), 43-62.
doi:10.1007/s11123-016-0471-5

All material supplied via JYX is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorised user.

How does the age structure of worker flows affect firm performance?

Pekka Ilmakunnas*

Aalto University School of Business

Mika Maliranta**

The Research Institute of the Finnish Economy (ETLA) and University of Jyväskylä

May 26, 2016

Abstract: We develop a method for decomposing firm performance to impacts coming from the inflows and outflows of workers and apply it to study whether older workers are costly to firms. Our estimation equations are derived from a variant of the decomposition methods frequently used for measuring micro-level sources of industry productivity growth. By using comprehensive linked employer-employee data, we study the productivity and wage effects, and hence the profitability effects, of the hiring and separation of younger and older workers. The evidence shows that the separations of older workers are profitable to firms, especially in the manufacturing ICT-industries. To account for the correlation of the worker flows and productivity shocks we first estimate the shocks from a production function using materials as a proxy variable. In the second step the estimated shock is used as a control variable in our productivity, wage, and profitability equations.

JEL-code: J23, J24, J63, M51

Keywords: aging, productivity, wage, profits, hiring, separation, employer-employee data

* P.O. Box 21240, FI-00076 Aalto, Helsinki, Finland, E-mail: pekka.ilmakunnas@aalto.fi

** Lönnrotinkatu 4 B, FI-00120 Helsinki, Finland, E-mail: mika.maliranta@etla.fi

*** We are thankful to many individuals at Statistics Finland, and especially to Satu Nurmi and Elias Einiö for their guidance regarding the properties of the data. The data set is publicly available for research purposes, subject to terms and conditions of confidentiality, at the Research Services unit of Statistics Finland; contact tutkijapalvelut@stat.fi, for access to these data. We have benefited from the comments of Roope Uusitalo and seminar participants at the Government Institute of Economic Research (Helsinki), Nordic Summer Institute in Labor Economics in Uppsala, EEA Congress in Vienna, CAED Conference in Chicago, Annual Meeting of the Finnish Society for Economic Research in Helsinki, World Aging and Generations Congress in StGallen, and Workshop on Labor Turnover and Firm Performance, Helsinki. We thank the referees and associate editor for useful comments. Financial support from the Yrjö Jahnsson Foundation is gratefully acknowledged. The SAS and Stata codes used in this study are available from the authors upon request.

1. Introduction

The increasing average age of the work force poses difficult challenges both to firms and the whole society. As a result of changes in the population age structure, an increasing share of firms' employees is in higher age groups. Pressures on the sustainability of pension systems have led governments to find ways of getting people to lengthen their working lives. These measures include reducing incentives of using early retirement channels or increases in the mandatory retirement age, but also extensions of the subjective right to continue working. It has also been argued that a greater awareness of the actual (low) level of pension in defined contribution pension plans will force many older employees to work longer. These developments will further increase the average age of the work force. The firms' incentives may not, however, align with the increasing supply. In contrast, firms are often reluctant to hire or even retain older employees and may prefer early retirement. Whether having older workers is profitable for the firms depends on the development of productivity and wage with age.

Ageing is most likely to affect performance in jobs that require physical strength. However, the development of cognitive abilities with age is not quite straightforward. Different kinds of abilities develop differently by age and the "productivity potential" of workers depends also on the development of the demand for different abilities (e.g. Skirbekk, 2008, Ilmakunnas et al., 2010). It is likely that the consequences of ageing therefore differ between industries.

Employees and employers tend to view these issues quite differently (e.g. van Dalen, Henkens, and Schippers, 2010). The advocates of older employees emphasize that lower turnover and higher experience may compensate age-related losses in working capacity and advise firms to hold onto their 50+ employees (e.g. Towers Perrin, 2005). On the other hand, firms often think that there is an unfavorable productivity-compensation deficit (Munnell and Sass, 2008). Declining productivity with age may be a real phenomenon or a stereotype, but it is easier to observe age-related increases in labor costs, which are related to health care costs and seniority-based pay. The discrepancy of productivity and wage has been explained by the existence of deferred payments (Lazear, 1979). According to this argument, lower pay of the employees in

their early career is repaid by the firm in the form of wage that exceeds productivity in the later career.

Economists have tried to measure age-related changes in productivity and labor costs. Individual-level productivity measures are available only in very special cases. On the other hand, the relationship between age and performance at the individual level may differ from the relationship of average age and labor productivity at the firm level, as the work environment and interaction between workers also matter. Therefore linked employer-employee data sets have been used for analyzing the impact of work force characteristics, like average age or shares of employees in different age groups, on plant- or firm-level productivity and wage (e.g. Hellerstein, Neumark, and Troske, 1999; Aubert and Crépon, 2003; Ilmakunnas, Maliranta, and Vainiomäki, 2004; Ilmakunnas and Maliranta, 2005; Daveri and Maliranta, 2007). This line of work is continuing, with emphasis on differences between firms (see e.g. Cataldi, Kampelmann, and Rycx, 2011, on ICT and non-ICT firms, Göbel and Zwick, 2012, and Mahlberg et al., 2013a, on sector differences, Göbel and Zwick, 2013, on firms with different personnel measures, and Vandenberghe, Waltenberg, and Rigo, 2013, on firms with different levels of training), differences between different types of employees (see e.g. Vandenberghe, 2013, on gender differences), regional differences (e.g. Mahlberg et al., 2013b), and on estimation methods like GMM (e.g. van Ours and Stoeldraijer, 2011, Göbel and Zwick, 2013, and Vandenberghe, 2013) or proxy variable methods (e.g. Vandenberghe, 2013).

A drawback of this line of studies is that they do not pay much attention to how the structure of the work force is determined¹. The purpose of this paper is to extend the analysis to directly examining how the age structure of the work force changes through the inflow and outflow of labor input and how the flows subsequently influence firm profitability. We disaggregate the labor flows to and from firms by age to three groups, “young” (30 or less), “middle-aged” (31 to 50), and “old” (over 50). (We emphasize that the labels “young” and “old” are used just for illustrative purposes and refer to relative age.) We decompose firm-level labor productivity change into the effects of the hiring and separation rates of the age groups and into the effect of productivity growth of those workers in different age groups who are staying in the firm. A similar decomposition can be made for firm wage growth. Combining

the two decompositions, we also obtain an equation for firm profitability change, which is the main interest of this paper. These decompositions follow Maliranta et al. (2009), who applied them to the flows of R&D workers.

The decomposition bears a resemblance to the kind of decompositions used frequently to decompose industry-level productivity change to the impacts of entry and exit of firms, and productivity growth in continuing firms (e.g Foster, Haltiwanger, and Krizan, 2001; Balk, 2015). A difference to these productivity decompositions is, however, that in our case the productivities are not observed, but are estimated. Our decomposition leads to a simple estimation equation, where the parameters have the interpretation of relative productivity levels of the different employee groups. To be able to perform the analysis we need detailed and comprehensive linked employer-employee data. We use the FLEED (Finnish Longitudinal Employer-Employee Data) data set of Statistics Finland that covers basically all firms in the country and all of their employees. The decompositions relate to performance change in the intervals 1995-97, 1997-99, 1999-2001, 2001-2003, and 2003-2005. To account for the endogeneity of the flows, we first estimate a production function using the proxy variable method suggested by Olley and Pakes (1996) and Levinsohn and Petrin (2003); see also Akerberg et al., (2007). From this estimation we obtain a measure of the productivity shocks, which is then used as a control variable in the estimated decompositions.

Our results show that there is a positive relationship between separations of older workers and firms' profitability, which is economically and statistically significant. This holds especially in the manufacturing ICT industries. The positive connection is mainly due to a lower relative productivity level of the separated older workers. Separation of younger worker groups is, on the other hand, markedly less profitable, mainly because of their higher relative productivity. As for the hiring side, our results show that the (recently) hired older workers are positively related to profitability. This is not surprising because hiring of older workers is quite rare and thus their exceptionality is expected. Hiring of younger workers is (initially) unprofitable which is consistent with the view that accumulation of firm-specific knowledge is needed for better productivity (and profitability). Indeed, our results also show that the productivity growth of the staying younger workers exceed that of the mid-aged and

even more that of the older workers. On the other hand, these productivity changes correspond to wage growth. However, we have to be careful in interpreting these associations as causal.

The structure of the paper is as follows. In section 2 we discuss related literature on turnover and firm performance. In section 3 we describe the decomposition of the growth in productivity, wage, and profitability to the impacts of the labor flows. In section 4 we describe the data set and present the estimation results. Section 5 concludes the paper with some suggestions for further research.

2. Worker turnover and firm performance

Our analysis is related to other research on labor turnover and firm performance. The traditional view both in labor economics and management literature is to emphasize the negative aspects of turnover. In the management literature (e.g. Dalton, Todor, and Krackhardt, 1982), separation is called dysfunctional, when those high-productivity workers whom the organization would like to keep, are leaving.² This involves adjustment costs in the form of rehiring and training, but also less directly in the form of disruption of informal communication structures. However, labor turnover can also be functional, i.e. in the interest of the organization. This can happen e.g. when low productivity workers quit or their separation from the firm is initiated by the employer. Replacing the leavers by new workers also brings new ideas and knowledge to the firm.³ This reasoning has generated some work where an optimal turnover rate is investigated, often using quits as the measure of turnover. There are only a few studies that have examined the separate effects of total hiring and total separation on performance (Bingley and Westergaard-Nielsen, 2004, Ilmakunnas et al., 2005, Siebert and Zhubakov, 2009), but they do not consider subgroups of the labor flows.

It is important to consider both hiring and separation flows and a detailed decomposition of them. If all employees were perfect substitutes, simultaneous separation and hiring would just cause costs without having a positive impact on productivity. The only necessary turnover would be such that is needed for expanding

or reducing the total size of the labor input. However, if there is a connection e.g. between the age structure of the work force and performance, it is the inflow and outflow of different types of employees that the firms should control to optimize the work force structure. The optimal age mix of employees is based on the relative productivities and wages of the age groups, but the choice is constrained by legal limits on layoffs, availability of different types of employees (i.e., local labor supply), and differences in the quit propensities of different employee types. Therefore, the firms may not always be at the optimum and will adjust towards it with the inflow and outflow of different age groups.

There is recent empirical research using large-scale employer-employee data sets where the inflow of new employees is explicitly seen as means of knowledge transfer and attention is paid to the decomposition of the flow. The impact of several different types of knowledge flows on productivity has been investigated: Boschma et al. (2009) decompose hiring according to experience from similar, related, or unrelated industries, Balsvik (2011) by experience in multinational enterprises, Stoyanov and Zubanov (2012) by the productivity level of the firms where the hired workers are coming from, and Parrotta and Pozzoli (2012) according to the type of knowledge carriers (highly educated or technically educated employees with above average earnings).

While the studies discussed above have decomposed the hiring flow in many different ways, only a few studies have considered also the relationship of worker outflow to productivity. Kronenberg and Carree (2010) explain productivity growth by the qualities of hired and exiting workers. The quality variables include average age of the hired and exited, average productivity of previous employer, and shares of employees hired from the same industry or exiting to the same industry. Kaiser et al. (2015) estimate a patent production function, where the explanatory variables include the shares of hired workers from patenting and non-patenting firms and exited workers disaggregated in a similar way. Maliranta et al. (2009) decompose both the hiring and separation rates by age (young/old), education (low/high), tenure (short/long), previous job (R&D/other), and current job (R&D/other) and include these rates as explanatory variables in models for productivity and profitability growth. The staying workers are included as the shares of the corresponding

subgroups. Our approach is similar to that in Maliranta et al. (2009), but we concentrate on the age effects.⁴

3. Decomposition of firm performance

Assume a production function with M worker types (age groups) $L_j, j = 1, \dots, M$, at time t as the inputs:

$$Y_t = f(L_{t1}, \dots, L_{tM}) \quad (1)$$

Take a first-order Taylor approximation of the production function about

$$X^0 = (L_1^0, \dots, L_M^0):$$

$$Y_t \approx f|_{X^0} + \sum_j \dot{a}_j f|_{X^0} (L_{tj} - L_j^0) + R = \sum_j \dot{a}_j f|_{X^0} L_{tj} + [f|_{X^0} - \sum_j \dot{a}_j f|_{X^0} L_j^0 + R] \quad (2)$$

where R is the remainder term of the approximation. Dividing by $L_t = \sum_j L_{tj}$ and denoting the term in the brackets A we obtain

$$\frac{Y_t}{L_t} \approx \sum_j \dot{a}_j g_j \frac{L_{tj}}{L_t} + \frac{A}{L_t} \quad (3)$$

where $g_j = f|_{X^0}$, i.e. the marginal product of worker type j . Assuming constant returns, we can replace the marginal productivities g_j by labor productivities Y_j / L_{tj} , where Y_j is the output accounted for by worker type j . For example in a Cobb-Douglas function the marginal products are $g_j = a_j Y_t / L_j^0 Y_j / L_{tj}$, where a_j is the power of labor input L_{tj} in the production function.⁵

We consider the development of productivity between points of time 0 and 1 (i.e. $t = 0, 1$). Given the approximation above, the firm's labor productivity at time 1 can then be expressed as the average of labor productivities, weighted by labor shares:

$$\frac{Y_1}{L_1} = \sum_j \dot{a}_j \frac{L_{1j}}{L_1} \frac{Y_{1j}}{L_{1j}} + e_{Y/L,1} \quad (4)$$

where the error term $e_{Y/L,1}$ has been included to reflect approximation errors, unobservable factors in our formulation, and the A/L term in (3). Each worker age group can further be divided into two subgroups; workers who worked in the firm at the previous point of time 0 and are still working in there, i.e., stayers (*stay*), and those who are working in the firm at time 1 but were not there at time 0, i.e., they were hired after 0 (*hire*). The firm's labor productivity level can then be expressed as follows:

$$\frac{Y_1}{L_1} = \overset{M}{\underset{j}{\overset{\circ}{\mathbf{a}}}} \frac{L_{1j,stay}}{L_1} \frac{Y_{1j,stay}}{L_{1j,stay}} + \overset{M}{\underset{j}{\overset{\circ}{\mathbf{a}}}} \frac{L_{1j,hire}}{L_1} \frac{Y_{1j,hire}}{L_{1j,hire}} + e_{Y/L,1}. \quad (5)$$

The shares of stayers and hired workers add up to one:

$$\overset{M}{\underset{j}{\overset{\circ}{\mathbf{a}}}} \frac{L_{1j,stay}}{L_1} + \overset{M}{\underset{j}{\overset{\circ}{\mathbf{a}}}} \frac{L_{1j,hire}}{L_1} = 1. \quad (6)$$

Taking this adding-up constraint into account and multiplying and dividing the first term in the right hand side of (5) by $\overset{\circ}{\mathbf{a}}_{j, L_{1j,stay}} \overset{\circ}{L_{1,stay}}$, it can be written as follows:

$$\begin{aligned} \frac{Y_1}{L_1} &= \overset{M}{\underset{j}{\overset{\circ}{\mathbf{a}}}} \frac{L_{1j,stay}}{\overset{\circ}{\mathbf{a}}_{j, L_{1j,stay}}} \frac{Y_{1j,stay}}{L_{1j,stay}} \frac{\overset{\circ}{\mathbf{a}}_{j, L_{1j,stay}}}{L_1} + \overset{M}{\underset{j}{\overset{\circ}{\mathbf{a}}}} \frac{L_{1j,hire}}{L_1} \frac{Y_{1j,hire}}{L_{1j,hire}} + e_{Y/L,1} \\ &= \overset{M}{\underset{j}{\overset{\circ}{\mathbf{a}}}} \frac{L_{1j,stay}}{L_{1,stay}} \frac{Y_{1j,stay}}{L_{1j,stay}} \frac{\overset{\circ}{\mathbf{a}}_{j, L_{1j,stay}}}{L_1} + \overset{M}{\underset{j}{\overset{\circ}{\mathbf{a}}}} \frac{L_{1j,hire}}{L_1} \frac{Y_{1j,hire}}{L_{1j,hire}} + e_{Y/L,1} \\ &= \overset{M}{\underset{j}{\overset{\circ}{\mathbf{a}}}} \frac{L_{1j,stay}}{L_{1,stay}} \frac{Y_{1j,stay}}{L_{1j,stay}} + \overset{M}{\underset{j}{\overset{\circ}{\mathbf{a}}}} \frac{L_{1j,hire}}{L_1} \frac{Y_{1j,hire}}{L_{1j,hire}} - \overset{M}{\underset{j}{\overset{\circ}{\mathbf{a}}}} \frac{L_{1j,stay}}{L_{1,stay}} \frac{Y_{1j,stay}}{L_{1j,stay}} \frac{\overset{\circ}{\mathbf{a}}_{j, L_{1j,stay}}}{L_1} + e_{Y/L,1} \\ &= \overset{M}{\underset{j}{\overset{\circ}{\mathbf{a}}}} \frac{L_{1j,stay}}{L_{1,stay}} \frac{Y_{1j,stay}}{L_{1j,stay}} + \overset{M}{\underset{j}{\overset{\circ}{\mathbf{a}}}} \frac{L_{1j,hire}}{L_1} \frac{Y_{1j,hire}}{L_{1j,hire}} - \frac{Y_{1,stay}}{L_{1,stay}} \frac{\overset{\circ}{\mathbf{a}}_{j, L_{1j,stay}}}{L_1} + e_{Y/L,1} \end{aligned} \quad (7)$$

To write the labor productivity level of the firm at time 0 we define a third subgroup, those who were in the firm at time 0, but are no longer there at time 1, i.e. those who have separated after 0 (*sepa*). We can write the time 0 productivity in an analogous way to (7), defining $\overset{\circ}{\mathbf{a}}_{j, L_{0j,stay}} \overset{\circ}{L_{0,stay}}$:

$$\frac{Y_0}{L_0} = \overset{M}{\underset{j}{\overset{\circ}{\mathbf{a}}}} \frac{L_{0j,stay}}{L_{0,stay}} \frac{Y_{0j,stay}}{L_{0j,stay}} + \overset{M}{\underset{j}{\overset{\circ}{\mathbf{a}}}} \frac{L_{0j,sepa}}{L_0} \frac{Y_{0j,sepa}}{L_{0j,sepa}} - \frac{Y_{0,stay}}{L_{0,stay}} \frac{\overset{\circ}{\mathbf{a}}_{j, L_{0j,stay}}}{L_0} + e_{Y/L,0} \quad (8)$$

We are interested in labor productivity growth, i.e., the growth of productivity level between points of times 0 and 1, i.e.

$$D \frac{Y}{L} = \frac{Y_1}{L_1} - \frac{Y_0}{L_0}. \quad (9)$$

We define the worker age groups in such a way that none of the staying workers changes her group between times 0 and 1, i.e., $L_{0j,stay} = L_{1j,stay}$ and therefore $L_{0j,stay}/L_{0,stay} = L_{1j,stay}/L_{1,stay}$ for all j . Note that people are, of course, aging over time, but the age groups should be understood as cohorts rather than absolute age groups.

We then obtain⁶

$$\begin{aligned} \frac{Y_1}{L_1} - \frac{Y_0}{L_0} = & \\ & \sum_j^M \frac{L_{0j,stay}}{L_{0,stay}} \frac{\Delta Y_{1j,stay}}{e L_{1j,stay}} - \frac{Y_{0j,stay}}{L_{0j,stay}} \frac{\Delta}{\emptyset} + \\ & \sum_j^M \frac{L_{1j,hire}}{L_1} \frac{\Delta Y_{1j,hire}}{e L_{1j,hire}} - \frac{Y_{1,stay}}{L_{1,stay}} \frac{\Delta}{\emptyset} + \\ & \sum_j^M \frac{L_{0j,sepa}}{L_0} \frac{\Delta Y_{0,stay}}{e L_{0,stay}} - \frac{Y_{0j,sepa}}{L_{0j,sepa}} \frac{\Delta}{\emptyset} + e_{Y/L,1} - e_{Y/L,0} \end{aligned} \quad (10)$$

The first set of terms on the right-hand side of equation (10) shows the productivity growth “within workers”, i.e. the productivity growth that accumulates over time for those who are staying in the firm. It can be interpreted as productivity growth due to the accumulation of human capital through experience. The within worker productivity growth may vary across the age groups, and the total effect is a labor share weighted average of productivity changes in the different groups. A firm has a rapid productivity growth when a large proportion of workers have a high productivity growth rate. These workers may have such human capital that enables them to adopt or innovate more productive techniques. In other words, these workers have dynamic long-run effects on the company’s productivity. This can be called Nelson-Phelps effect according to the seminal work by Nelson and Phelps (1966).

The second set of terms indicates the productivity effects of hiring of workers in different age groups. As can be seen from (10), hiring of type j workers has a positive impact on productivity change when these hired workers have a higher productivity level than the average staying workers. Newly hired workers may be more productive than incumbents at time 1 because they have learned more productive techniques

when they worked for the previous employer, or have more recent education, for example. Adjustment costs related to the hiring of new employees are implicitly included in our formulation. The relative productivity of the hired workers should therefore be understood as productivity net of adjustment costs.

The third set of terms indicates the productivity effects of separations of different worker age groups. Quite analogously to the hiring effect, separation of type j workers has a positive effect on productivity change when these workers have a lower productivity level than the average incumbent worker at the initial time 0. Again, the productivity impact of separations is net of adjustment costs. Finally, there is the change in the error term.

The terms of the decomposition have analogues in the firm-level productivity analyses, where productivity growth is decomposed to growth within firms, between firms and through entry and exit. Here the productivity growth of the stayers plays the role of within-firm productivity growth, and the hiring and separation effects are analogous to entry and exit. However, in (10) there is no analogue to the between effect of the firm-level analysis, since the group of stayers in an age cohort does not switch to another group (i.e., the “market shares” of the age cohorts do not change among stayers).

Besides labor productivity, we can use a similar decomposition for the average wage level in the firm, since the average wage in the firm is a share weighted average of wages in the worker groups. In this case we just replace Y in the equations above by the wage sum W . Without the error terms, equation (10) and a corresponding equation for wage growth are in principle identities. We can observe the labor flows, but we do not know the productivities, so the equations cannot be used directly for assessing productivity and wage differences between the age groups. There are some influences, however, that have not been taken into account and now included in the error term and allow us to use the equations as a basis for estimating the productivities. There are likely to be differences across firms in the productivities of different age groups. If we use (10) as a model for estimating parameters that correspond to the age-specific productivities, we will estimate average productivities. Any firm differences will

therefore be included in the error term. So far we have not taken into account other inputs, especially capital that affect productivity. Capital input could have been included already in the production function (1), in which case the capital-labor ratio would appear in the approximation (3). We will therefore include the capital-labor ratio and other control variables Z to account for other exogenous influences on firm productivity, wage, and profits. Inclusion of a constant term takes into account the productivity growth trend. After these observable influences are taken into account, the error accounts for all unobservables.

The terms of (10) can be turned into growth rates by dividing them by the average of productivity level of times 0 and 1. The growth rate is then a close approximation of a more common log-difference. We obtain the following estimation models:

$$\frac{D(Y/L)}{(Y/L)} = a_Y + \sum_j^M \mathring{a}_{Y/L,j,hire} HR_j + \sum_j^M \mathring{a}_{Y/L,j,sepa} SR_j + \sum_j^{M-1} \mathring{a}_{Y/L,j,stay} STAYSH_j + d'Z + De_{Y/L} \quad (11)$$

$$\frac{D(W/L)}{(W/L)} = a_W + \sum_j^M \mathring{a}_{W/L,j,hire} HR_j + \sum_j^M \mathring{a}_{W/L,j,sepa} SR_j + \sum_j^{M-1} \mathring{a}_{W/L,j,stay} STAYSH_j + d'Z + De_{W/L} \quad (12)$$

where $\overline{(Y/L)} = 0.5 \left((Y_0/L_0) + (Y_1/L_1) \right)$ and $\overline{(W/L)} = 0.5 \left((W_0/L_0) + (W_1/L_1) \right)$ are

averages of productivity and wage, respectively, $HR_j = L_{1j,hire} / L_1$ and

$SR_j = L_{0j,sepa} / L_0$ are the hiring and separation rates, respectively, and

$STAYSH_j = L_{0j,stay} / L_{0,stay} (= L_{1j,stay} / L_{1,stay})$ is the share of staying workers. The shares

of the M stayer groups sum up to one, so if they were all included in (11) and (12),

they would have a linear dependence with the constant term. For the hiring and exit

variables there is no such problem, as they are rates (e.g. hired of type j in relation to

the number of all workers, not to all hires) and therefore all M groups can be included.

In the estimations, we use panel data of firms over non-overlapping time periods to

calculate two-year differences within the periods. Therefore the equations to be

estimated will be indexed with i (firm) and s (period), which are not shown in (11)-

(12).

The productivity and wage gaps by age can be analyzed both on the hiring side and on the separation side. On the hiring side the coefficients of our main interest that will be estimated with equations (11)-(12) have the following interpretations:

$$b_{(Y/L),j,hire} = \frac{(Y/L)_{1,j,hire} - (Y/L)_{1,stay}}{(Y/L)} \quad (13)$$

$$b_{(W/L),j,hire} = \frac{(W/L)_{1,j,hire} - (W/L)_{1,stay}}{(W/L)}, \quad (14)$$

i.e. they measure the productivity and wage, respectively, of hired workers in age group j , compared to all staying workers. These differences are scaled by firm (over time average) productivity and average wage, respectively, which follows from the transformations of the dependent variables in (11) and (12). Equations (13)-(14) are approximately equal to log-differences, and hence the parameters measure relative differences between hired and staying workers. For the separation side, the estimable coefficients are obtained analogously as

$$b_{(Y/L),j,sepa} = \frac{(Y/L)_{0,stay} - (Y/L)_{0,j,sepa}}{(Y/L)} \quad (15)$$

$$b_{(W/L),j,sepa} = \frac{(W/L)_{0,stay} - (W/L)_{0,j,sepa}}{(W/L)}, \quad (16)$$

which measure the relative productivity and wage difference, respectively, of all staying workers and separated workers of type j . The intercepts in (11)-(12) indicate the growth rate in the reference age group of the stayers and the coefficients of the included $STAYSH_j$ variables ($M-1$ age group variables) indicate differences in the growth rate in the age groups j and in the reference group.

In this paper, we are particularly interested in profitability effects. Profitability is measured as follows:

$$P = 1 + \frac{OPM}{W(1+a)} = \frac{Y}{W(1+a)} = \frac{Y/L}{(1+a)(W/L)}, \quad (17)$$

where OPM denotes operating margin (i.e., $OPM = Y - W(1+a)$), where a , the ratio of payroll taxes to wages, is assumed to be constant over time and across the worker groups.⁷

The growth rate of profitability (17) is therefore the difference between the growth rates of productivity Y/L and wage W/L , which is approximated as

$$\frac{DP}{\bar{P}} \approx \frac{D(Y/L)}{(Y/L)} - \frac{D(W/L)}{(W/L)}, \quad (18)$$

where $\bar{P} = 0.5[P_0 + P_1]$.

By inserting (11) and (12) into (18) we obtain an equation for the profitability change

$$\frac{DP}{\bar{P}} = a + \sum_j^M \hat{a}_j b_{P,j,hire} HR_j + \sum_j^M \hat{a}_j b_{P,j,sepa} SR_j + \sum_j^{M-1} \hat{a}_j c_{P,j,stay} STAYSH_j + d'Z + De_{P/L} \quad (19)$$

where, on the basis of (18), the following approximations hold

$$b_{P,j,hire} \approx b_{(Y/L),j,hire} - b_{(W/L),j,hire} \quad (20)$$

$$b_{P,j,sepa} \approx b_{(Y/L),j,sepa} - b_{(W/L),j,sepa} \quad (21)$$

Since

$$b_{(Y/L),j,hire} = \frac{(Y/L)_{1,j,hire} - (Y/L)_{1,stay}}{(Y/L)} \approx \ln \frac{(Y/L)_{1,j,hire}}{(Y/L)_{1,stay}} \quad (22)$$

and

$$b_{(W/L),j,hire} = \frac{(W/L)_{1,j,hire} - (W/L)_{1,stay}}{(W/L)} \approx \ln \frac{(W/L)_{1,j,hire}}{(W/L)_{1,stay}} \quad (23)$$

(20) can be developed as follows

$$\begin{aligned} b_{P,j,hire} &\approx \ln \frac{(Y/L)_{1,j,hire}}{(Y/L)_{1,stay}} - \ln \frac{(W/L)_{1,j,hire}}{(W/L)_{1,stay}} = \ln \frac{(Y/W)_{1,j,hire}}{(Y/W)_{1,stay}} \\ \hat{U} b_{P,j,hire} &\approx \ln \frac{P_{1,j,hire}}{P_{1,stay}} \end{aligned} \quad (24)$$

which shows that the parameter of the hiring variable for the worker group j in the profit equation (19) can be interpreted as a measure of the relative profitability difference of the hired group j workers and all stayers at time 1.

Analogously, we obtain that

$$b_{P,j,sepa} \gg \ln \frac{P_{0,j,sepa}}{P_{0,stay}}, \quad (25)$$

which provides us a measure of the relative profitability difference of the separated group j workers and staying workers before they leave.

4. Econometric issues

When using equations (11), (12) and (19) for estimation, there are possible sources of bias. First, there can be unobservable firm heterogeneity both in firm productivity and wage levels, which is correlated with the employee characteristics (flows and shares of different age groups). In particular, new firms often start with a new work force which only slowly evolves over time (Haltiwanger, Lane, and Spletzer, 1999, 2007). Therefore, firm vintage and worker cohorts tend to be tied together, with young workers being employed in firms that have new equipment and a high productivity level. Since we are using growth rates as the dependent variables, this kind of time-invariant unobservables are eliminated. Our approach is related to the use of long differences to eliminate time-invariant unobservable effects in panel data models (e.g. Griliches and Mairesse, 1998; Haltiwanger et al., 2007). We define the growth rates and labor flows in five different time periods and pool them in estimation. We also control for some observable firm characteristics, industry and region, included in Z . These account for industry and region specific trends in the levels.

Second, there is heterogeneity across workers. Since the firms may hire the best applicants and lay off poor performers, the hiring and separation flows may be unrepresentative of the whole population. Heterogeneity can to some extent be taken into account by disaggregating the flows by age and by investigating the exit flows also by destination. For example, flows to unemployment may be a special group of all separations. Selection should affect productivity growth and wage growth in the same way under the null hypothesis of competitive labor markets (see Hellerstein and Neumark, 2007). Therefore selection should be a less serious issue when we examine their difference, i.e. the productivity-wage gaps which directly relate to our measure of firm performance.

Third, the hiring rate and the involuntary part of the separation rate (i.e. layoffs) are based on the firms' decisions and are therefore possibly correlated with productivity shocks which are part of the error terms. Time-varying shocks would remain in the equations for the growth rates.. For example, positive productivity shocks may lead to the hiring of new, young workers, which then causes an overestimate of their productivity effect (cf. Olley and Pakes, 1996; Levinsohn and Petrin, 2003). On the other hand, one of our main interests is the profitability effect of the separations of older workers. It is not quite obvious why a positive productivity or profitability shock should increase separations of older workers, which would generate a spurious positive correlation between profitability change and the separation of older workers. More probably, a negative profitability shock should increase a company's incentives to encourage its older labor force to early retirement, for instance. As a consequence, our main results are more likely biased in the direction of not showing older workers being overpaid even when they actually may be.

We use a proxy variable approach, suggested by Olley and Pakes (1996) and Levinsohn and Petrin (2003), for dealing with the shocks. Materials is the proxy variable, which is a function of capital and the productivity shock. The capital variable is the end of previous year capital stock, which is not affected by the current period shock. We first estimate a production function in level form using annual data, without disaggregating the labor input by age, including in this estimation a polynomial of capital and the proxy variable to account for the shock. The method yields an estimate of the productivity shock for each firm in each year. Then the shock is used in differenced form as an additional control variable in our productivity, wage, and profitability change models.

Another way in which we address the issue correlation of the flows and shocks is to identify those workers who have left the company for the old age pension. We consider this a more exogenous type of separation than the other separations.

Fourth, there can be productivity differences across firms in the productivity of a certain age group. This can arise from decreasing returns. For example, extensive use of younger employees in a firm lowers their marginal productivity. There may also be

genuine technological differences between firms or industries, including differences in economies of scale⁸ or efficiency differences. These factors would imply that the coefficients vary across firms. We can still obtain an unbiased estimate of the mean coefficient and account for the firm differences by correcting standard errors for clustering within firms. To account for heterogeneity in technology we also estimate the models separately for some subsets of the data. Besides the whole business sector, we present results separately for the industrial and service sectors. In addition, in each of these sectors we study ICT and non-ICT firms separately.

Finally, price differences across firms may cause biases when a common deflator is used for all firms in an industry (see e.g. Foster, Haltiwanger, and Syverson, 2008). For example, the profitability level of a low-price firm will be underestimated and that of a high-price firm overestimated. However, to the extent that there are firm differences only in price levels, the bias is eliminated when profitability changes are examined, since the log-differences of the prices would be the same. If also price growth differs, the bias is not eliminated.

5. Empirical analysis

5.1. Data and variables

The unique identification codes for persons, companies and plants used in the different registers forms the backbone of the Finnish administrative register network and the Finnish statistical system. This provides an excellent opportunity to construct cross-sectionally and dynamically representative data for various research purposes by linking different administrative data sources (see Abowd and Kramarz, 1999).

The data for this study are drawn from the *Finnish Longitudinal Employer–Employee Data* (FLEED). The data set merges comprehensive administrative records of all labor force members as well as all employers/enterprises (including information also on their establishments) subject to value added tax (VAT). It can be complemented by a range of additional information from both private and public sources. FLEED currently covers the years 1990 onwards with near-perfect traceability of employers and employees across time. The employment statistics, educational statistics, taxation

records, business register, financial statement statistics, manufacturing census as well as various surveys are among the original sources of the FLEED variables.

To define the labor flows and changes in productivity, wage, and profitability, we use 2-year windows. The flows and changes are defined for the five time periods 1995-97, 1997-99, 1999-2001, 2001-2003, and 2003-2005.⁹ The observation unit is a firm. In principle we also have data on establishments, but establishment-level information on value added and some other relevant variables, like capital intensity, are lacking beyond the manufacturing sector. Further, the links between employees and firms are more reliable than those between employees and establishments, especially in multi-unit firms. A drawback of the data is that we are not able to systematically identify the role of change in ownership, like mergers and acquisitions.

Our estimation sample covers the industry and service sectors. The industry sector is defined very broadly as consisting of branches that are not included in services, i.e., mining, manufacturing, public utilities and construction. The service sector comprises retail and wholesale trade, business services and personal services. Real estate and financial intermediation are excluded due to problems in measuring output in a reliable manner. The number of observations in the estimation sample by branches is shown in Appendix Table A.1.

The dependent variables are defined as follows. Labor productivity growth is measured as a two-year rate of change of value added per employee (defined as in (11)), average wage growth is correspondingly a two-year rate of change of wage sum per employee, and change in profitability is a two-year relative change in value added per labor costs (wages and social security payments). These variables are measured in nominal terms, and price changes are controlled by a set industry dummies that are interacted with the period dummies.¹⁰

The labor flows are based on the comparisons of employees in the firms at two points of time s and $s-2$, where s denotes the end year of a three-year period.¹¹ The flow rates are calculated separately for three age groups, “young” (< 30 years), “middle-aged” (31-50 years), and “old” (51 - years). We use fairly broad age groups to ensure that we have enough employees in the groups when hiring and separation are

disaggregated by age. With narrowly defined groups, e.g. 10-year intervals, many of the flows would be zeroes in the smaller firms. In each period age is based on the situation at the end year. For example, those who were 28 years old in $s-2$, are 30 years old in s and hence included among the “young”. Those who were 30 already in year $s-2$, are 32 in s , and hence included in the middle group. The age group classification is thus based on year s age, and not on the age at which the employees were last observed in the firm.

The hiring rate HR_{jis} for age group j is the number of new employees in firm i in the age group (those in the firm in s , but not in $s-2$) divided by the number of all employees of the firm in s . The separation rate SR_{jis} is correspondingly the number of exited employees of firm i in age group j (those in the firm in $s-2$, but no longer in s), divided by the number of all employees in the firm in $s-2$. The share of stayers, $STAYSH_{jis}$, is the number of staying employees of firm i in age group j (those in the firm both in s and $s-2$), divided by all stayers of the firm in $s-2$. The sum of these stayer shares is therefore one, so one of the groups is left out of the estimation.

As a control variable we use the change in the productivity shock, estimated using the proxy variable approach. We also control for the log of capital per employee, which is entered in two-year difference form to be consistent with the form of the dependent variables. The capital stock is measured by the book value at the end of the previous year.¹² Finally, we have a set of dummy variables. These include a set of dummies as controls for regional effects¹³ (20 regions), industry dummies (46 industries), period dummies (5 three-year periods), as well as interacted industry and period dummies to account for, besides price changes, also the effects of idiosyncratic industry shocks.

Before conducting the econometric analysis we leave out some potentially erroneous observations that may distort our results. First, we remove those observations where the number of linked employees differs more than 10% from the number of employees in the company data. This indicates that the linking of the individual and firm data is incomplete. Second, we remove some potentially influential outliers that we detected by using the method proposed by Hadi (1992, 1994). The method is useful for finding multiple outliers in multivariate data. Identification of outliers is

made on the basis of three variables: 1) the growth rate of average monthly earnings calculated from the data on individuals (Employment Statistics), 2) the growth rate of average wage calculated from the company data, and 3) the productivity growth rate. The first two variables should be highly correlated with each other because they are essentially gauging the same thing, but may sometimes differ due to possible inaccuracies in the links between employees and their employers, for instance. Wage growth is usually correlated with productivity growth, but sometimes they may be very different because of measurement errors in output and/or labor input. The identified outliers (508 out of 24 842 firm-period observations) are removed from all estimations. We also restrict the sample to firms that employ at least 10 employees and leave out the firms with over 10 000 employees; 11 observations are dropped from the sample due to the upper limit.

Table 1 gives some descriptive summary statistics of our basic sample that is used in the regression analysis below. Because some observations cannot be used in the analysis due to missing values of the explanatory variables, we are finally left with 23 738 observations. The average number of linked employees per company is 85.2, which is close to the average number of employees in these firms according to company data (85.8 persons). In other words, our regressions are based on over 2 million individual-period observations. Because we have five periods, our sample covers over 400 000 individuals per period. This figure includes those individuals who are employed by a company in our sample either in the initial ($s-2$) or end year (s) of a period or both. This is roughly one third of the total employment in the non-farm business sector.¹⁴

[INSERT TABLE 1 ABOUT HERE]

The average nominal productivity growth rate is 4.5%. Average wage growth rate, calculated from company data, is 6.7%. This is reasonably close to the average growth of monthly earnings of the linked employees, 6.9%, obtained from the register data on individuals (Employment Statistics). The average hiring rate, which is the sum of the hiring rates of the three age groups, is 28.1% and the average separation rate, sum of the separation rates of the three age groups, is 24.1%.¹⁵ Young employees account for 17.4% of the staying employees and the old workers account for 21.9%.

5.2. Basic estimation results

Table 2 reports the baseline estimates. All of the estimation results reported below are based on weighted estimation, with firm employment used as the weight.¹⁶ Since the productivity shock is a generated variable, we use robust standard errors, adjusted for clustering at the firm level. The first two columns show the results for the productivity and wage change equations with the productivity shock included as a control variable. The entries in the third column are from a separate estimation for profitability change, but they are roughly equal to the differences of the corresponding entries for productivity and wage change in columns 1 and 2. The last three columns have the corresponding results without productivity shock.

[INSERT TABLE 2 ABOUT HERE]

Consider first the results with the productivity shock included. Hiring of young employees lowers the productivity level. This implies that the young hires have lower productivity than all staying employees, presumably because of their lack of general skills. However, they also have lower wages, so that the negative effect on profitability is smaller than the effect on productivity. Hiring the mid-aged also has a negative association with productivity, but it is weaker than that of the youngest. Their connection to average wage is small and not significantly different from zero. The results on productivity and wage give a statistically significantly negative net connection to profitability. The hired old workers have a positive relationship with productivity, but it is imprecisely estimated, and the relationship with wage is negative, but not significant. The effect of hiring old employees on profitability is, however, positive and clearly significant. However, it should be noted that the number of recently hired older workers is quite small (see Table 1). These cases are thus quite exceptional and therefore it is unsurprising to find that these individuals are not a burden to the firm's profitability.

On the separation side, exiting young employees have a positive coefficient both in productivity and wage equations, implying that they have lower productivity and wage than the staying employees. Our point estimates suggest that young leavers are

underpaid, but because of a relatively large standard error the result on profitability does not differ from zero in a statistically significant way. Separation of the mid-aged has a higher coefficient in the productivity equation than the youngest. This results in a clearly significant positive connection to profitability. Separation of the oldest group has the largest coefficient both in productivity and wage equations, and in combination they yield an economically and statistically significant positive relationship between separation and profitability. These estimates indicate that the separating employees have a clearly lower productivity level than the continuing employees, but they are also paid somewhat less on average. The net result is thus an increase in profitability for all of the age groups, but the gap between pay and productivity increases with age. These results seem to support the deferred pay argument that gives rise to wage profiles that give a high pay at the end of the career.¹⁷

The results for the stayers indicate that the older employees have slower productivity growth than the young. This is consistent with the idea that the accumulation of productivity-enhancing (firm-specific) experience is greatest at the beginning of the career and then gradually slows down over time. Wages seem to develop correspondingly.

Comparison of the estimates with the productivity shock included (first three columns of Table 2) and those without the shock (last three columns) shows that the shock variable is significant in the productivity and profitability equations, but not in the wage equation. When the shock variable is included, some of the point estimates or their statistical significance differ from those obtained without the shock variable, but the main pattern of the coefficients remains the same. This shows that differencing over time has already succeeded in purging the unobservables and accounting for the time-varying shock has only a marginal additional influence on the results.

We have conducted some robustness analyses, which we briefly comment without showing the results in tables. First, we added firm size (number of employees) as a variable to account for scale economies. This had hardly any impact on the results. Second, we included worker characteristics as control variable. These included average tenure, average education years, and the share of females. Also this gave

qualitatively the same results and the changes were mostly in the third decimals of the coefficients. Third, we considered the possibility that there is a selection bias because of exit of low productivity firms. Although many firms exit, exiting firms account for a small share of employees so that we would expect that the role of the selection bias is mitigated, especially in weighted estimation. Using an inverse Mills ratio selection correction term gave indeed results that were fairly similar to those in Table 2.¹⁸ Fourth, we estimated the models without weighting. This produced in general coefficients that were lower in absolute value than those in our basic results. The pattern of the coefficients was, however, similar in the productivity and profitability equations, while the wage results had some differences. In the profitability model the only qualitative difference to the results of Table 2 was that the coefficient of the hired old employees was negative and not significantly different from zero.

5.3. Disaggregated results

To investigate whether the results hold for different sub-sectors within the whole business sector, we estimate the models separately for the industry and service sectors. Table 3 shows that there are some notable differences in the results. We find that the R^2 measures of the productivity and profitability equations are substantially higher in the industry sector than in the service sector. This may reflect a greater importance of idiosyncratic factors or measurement problems regarding productivity and profitability in service industries. In the wage equation R^2 is similar in both sectors.

[INSERT TABLE 3 ABOUT HERE]

More interestingly, the separation of older workers is particularly profitable in the industry sector. The results for the whole business sector are clearly driven by this sector. In contrast, the only statistically significant result on the hiring side in the industry sector is the negative coefficient for the mid-aged. The coefficient of the hired old workers is positive, but imprecisely estimated. These newly hired workers are paid less than their productivity. In the service sector the results are different in some respects. The labor flows of the older employees are profit neutral to firms, and hiring of young and mid-aged is negatively associated with profitability change.

The separations of the oldest age group may be driven by very different influences. Some of these employees are retiring. Some are laid off and may face periods of unemployment. Some are still looking for new jobs and quit to move to other firms. Finally, some withdraw from the labor market. To investigate the role of heterogeneity among the older employees, we have disaggregated the separation rate of the age group over 50 years into three flow rates. These are separation rate to pension (old age pension or disability pension), unemployment (including unemployment pension), and other (job-to-job moves and withdrawal from the labor market). For the sake of comparison, we have divided separations of the other age groups by destination into unemployment and other. There are very few in these age groups who end up into retirement; they have been included in the category “other”. The estimation results with this disaggregation are shown in Table 4, separately for the industry and service sectors.

[INSERT TABLE 4 ABOUT HERE]

Now the outflow of older workers into retirement and unemployment is found to have a statistically and economically significant positive relationship with productivity in the industry sector (column 1 of Table 4), indicating that these worker groups had lower than the average productivity level before they left. The results for wages (the second column) show that the wage level of these workers was below the average level, since their exit had a positive effect on average wage. Comparison of the coefficients of these exits in the productivity and wage equations indicates that these worker groups had been paid more than their productivity and their separations have been thus profitable to firms. This can also be seen in the positive coefficients of the third column, the profit equation. The productivity-wage gap is quite substantial. On the other hand, the results of Table 4 do not provide evidence that those older workers that have left the firm for some other destination, e.g. employment in another firm, had been overpaid. These workers account for roughly one third of the total separations of the older workers. So, a substantial proportion of the older workers are not found to be overpaid in our analysis. This includes besides separations to other destinations, also hired and staying older workers.

Interestingly, we do not find statistical evidence that those young or middle aged workers who separated into unemployment had been overpaid. In these age groups only exits to other destinations than unemployment have a positive impact on profitability. For the service sector the results are again different. The only significant relationships with profitability are obtained for the exit of the oldest age group to unemployment and the mid-aged to other destinations.

Our interpretation of the results is that especially the outflows to unemployment reflect the firms' choices (i.e., these separations are a selected group) whereas especially the route to old-age pension is a more exogenous event to the firm. It is worth noting that the Finnish pension and unemployment insurance systems have had an early exit route called "unemployment pension tunnel", which has allowed unemployed to withdraw from the labor market at a relatively early stage by successively transferring to unemployment compensation, unemployment pension and finally to normal pension. It has actually been relatively common for the firms to use this system for downsizing their labor force, which can be seen as an increase in the unemployment risk at an age where the "tunnel" starts (Kyyrä and Wilke, 2007).¹⁹ It can also be argued that the use of the unemployment pension has in many cases been in the mutual interest of the firms and their employees (Hakola and Uusitalo, 2005). Our results are quite consistent with the existence of this policy that makes it easy for firms to concentrate labor shedding on the older employees.

The use of temporary workers also makes it easy and cheap to downsize the labor force when necessary. Unfortunately, our data do not allow us to distinguish between temporary and permanent employees, but the temporary employees are usually young. Also Table 1 above shows that the flow rate to unemployment has been highest among the young. However, our estimation results do not give support to the view that this kind of downsizing would have been profitability-enhancing.

In order to examine more carefully what drives our findings of the negative productivity effects especially in the industry sector we have sliced our data a little bit more. First, we have classified the companies into ICT and non-ICT groups on the basis of 2- or 3-digit industries. Our starting point is the widely used classification proposed by van Ark et al. (2003). In their classification the ICT industries of the

business sector consist of a broad variety of different industries such as “Apparel” (ISIC 18), “Printing and Publishing” (ISIC 22), “Machinery” (ISIC 29), “Office and computer equipment” (ISIC 30), “Electrical machinery” (ISIC 31), “Radio, TV and communication equipment” (ISIC 32), “Instruments” (ISIC 33), “Transport equipment” (ISIC 35), “Miscellaneous manufacturing” (ISIC 36-37), “Wholesale” (ISIC 51), “Retail” (ISIC 52), “Telecommunications” (ISIC 64), “Renting of machinery” (ISIC 71), “Computer services” (ISIC 72), “R&D” (ISIC 73), and “Professional services” (ISIC 741-743). We concentrate here on a somewhat narrower definition.²⁰ More specifically, we have excluded industries 18, 35, 36-37, 51, and 52 from the group of ICT industries.

Table 5 shows the results of the profitability change equations for the ICT and non-ICT industries, with separations disaggregated by destination. The results are presented for the industry and service sectors separately. Columns (1) and (2) show the results for the ICT and non-ICT parts of the industry sector, respectively. It should be noted that all the ICT industries of the industry sector belong to manufacturing. The table shows only the estimates for the profitability equation, but they are dominated by the effects of the hiring and separation flows on productivity. The coefficients for the separations of the older workers are positive and especially in the case of workers exiting to pension the coefficient is higher for the ICT part of the industry than for the non-ICT part. This evidence supports the view that skill obsolescence lowers the relative productivity of older employees and, in addition, weakens their ability to adopt or innovate new technologies. The results of column (1) thus provide evidence that older workers may be a burden to firms particularly in the manufacturing ICT industries. However, in contrast to these results, hiring of older workers is positively related to profitability in the ICT industry. It seems that the firms are able to pick up highly productive employees. On the other hand, hiring of the mid-aged lowers profitability in the non-ICT sector. As for the young and medium aged workers, their separation to other firms has a significant positive coefficient in the ICT industries. That is, these workers in the ICT sector have profitability that is below average. In both the ICT and non-ICT parts of the service sector (columns 3 and 4) the impacts of the separations of older employees on profitability change are smaller than in the industry sector. Some of the point estimates are even negative, although statistically not significant. Interestingly, the positive association between profitability

and separations of older workers to unemployment is driven by the non-ICT part of the service sector, whereas the positive impacts of separations of young and mid-aged to other firms are driven by the ICT part of the sector.

[INSERT TABLE 5 ABOUT HERE]

One might suspect that our results are driven by some kind of downsizing effect, i.e. that firms aim to improve their profitability by reducing their work force. The results would in this case be dominated by separations in downsizing firms. It should be noted, however, that while our explanatory variables include all hiring and separation flows, we have implicitly controlled the employment growth (or decline) effect. In any case, for inspecting the downsizing hypothesis more carefully we have estimated the profitability equation by dividing the firms into two separate groups: the declining firms where employment has decreased and the expanding ones where employment has increased. Firms with stable employment are omitted. These results are shown in Table 6 for the industry and service sectors separately. Again we show only the results for the profitability change equation, but they are dominated by the effects of the flows on productivity. The main difference between growing and declining firms in the industry sector (columns 1 and 2 of the table) is that separations of older workers to unemployment has improved profitability more in the declining than in the growing firms. This supports the argument that downsizing has an important role in this labor flow. Growing industrial firms are also able to hire high-productivity older employees. Another difference is that separations of the young to other destinations than unemployment has improved profitability in the growing firms of the industry sector, but not in the declining ones. This may be an indication of more natural turnover in growing firms. In the service sector (columns 3 and 4 of the table) growing firms have hired low-productivity young and mid-aged workers, which has led to lower profitability. In contrast, hiring of these age groups to declining service sector firms has had a positive, but statistically non-significant relationship with profitability.

[INSERT TABLE 6 ABOUT HERE]

6. Conclusions

We have proposed a new way of estimating the performance effects of age using flows of labor to and from firms. The results support the argument that at the end of the working career wage exceeds productivity at least for some workers. Our results are consistent with some firm- or plant-level studies which have found a hump-shaped or declining relationship between productivity and age. Newer studies, on the other hand, tend to support a relatively flat age-productivity relationship at older ages (e.g. Göbel and Zwick, 2012, Mahlberg et al., 2013a).

Older workers appear to be particularly costly to firms in the manufacturing ICT industries. Further, the results indicate that in these cases older workers are negatively related to profitability (and productivity) growth, on top of lowering the current profitability (and productivity) level of the firm. These findings may be explained by rigidities in the wage formation which drive a wedge between the wage and productivity levels of older workers when rapid technological change makes skills obsolete at a rate that exceeds the rate of learning-by-doing. Such rigidities may derive from deferred compensation, employment protection that is focused more heavily on older workers, or insider power of the older workers in wage bargaining. On the other hand, we have found that a sizeable share of the older workers is not overpaid.

Our approach has features that help to account for many kinds of estimation biases. In particular, we have used a proxy variable method to estimate productivity shocks and used the shock as a control variable in our estimations. Still, we have to be careful in giving the results a causal interpretation, since the estimated shocks can be correlated with some other unaccounted time-varying unobservables. Therefore we have made several robustness checks of the results. In particular, separation to pension can be considered to be more exogenous than separations to other destinations and therefore more immune to estimation biases. Our finding of positive profitability effects from these separations of the older employees gives particularly strong support for our findings. We also emphasize that the intention is not to measure the performance effects of hiring or separating a randomly chosen employee, but the effects of the hires and separations actually done.

It should be noted that our estimates gauge the total association of outflows of the oldest workers with productivity. Besides a direct productivity interpretation (i.e. a worker's efficiency in her own task), the estimates of this kind of analysis may arguably also capture various indirect channels that come into being through the diffusion of knowledge between different worker groups within a firm. Important as the diffusion of the tacit knowledge of older worker to the younger ones may be in many circumstances to the employer, our results, however, suggest that generally this effect does not outweigh possible shortages in productivity or ability to adopt new and more productive techniques.

The institutional setting has obviously contributed to the results. Firms have incentives to downsize by laying off the oldest employees. On the other hand, the pension system has given incentives for wage profiles that peak at the end of the career. Our results support the view that firms have followed these incentives created by the institutions for improving their performance.

There are changes in the labor market, however, that will in the future increase incentives for keeping the aging labor force at work. Pension reforms increase the mandatory retirement ages or extend employees' rights to stay at work. There is also reduced availability of labor because of smaller age cohorts. These developments create pressures for firms to keep their older employees and to use new means for improving their performance, like changes in work organization and rotation of tasks. In addition, there is a shift away from final salary pensions to systems where the benefits are more closely tied to contributions over the working career. The new environment with longer, less fixed retirement age and fewer incentives for bargaining a back-loaded wage may give rise to flatter age-wage profiles in the future.

The approach that we have used could be extended in various ways, which are left for future work. We have disaggregated the exit by destination, but obviously an interesting question would also be where the hired employees are coming from, for example, whether they are coming from low or high productivity firms. The employees could also be examined in terms of other characteristics, e.g. education or gender. A potential problem with this kind of disaggregation is, however, that the number of worker groups would grow fast. If we had for example three educational

groups in addition to the three age groups, we would have nine age-education combinations and the models would have nine hiring rate variables, nine exit rate variables and eight stayer shares. In the smaller firms many of the age-education cells would be zeros and most likely the precision of the estimation would suffer as the coefficients would be identified from a smaller number of non-zero observations.

An interesting extension would also be to use actual wages of individual employees to calculate the wage decomposition directly instead of estimating it using average wages. However, for data reasons the results would most likely be different. The wage sum in the company data, which we have used for calculating average wages, is based on wages actually paid during the year, as reported by the firm. On the other hand, the data on individual wages in FLEED are based on tax registers and measure the total earnings of the individuals during the year. For workers who switch jobs the earnings therefore come from different firms. For example, if a person who has a low-pay job for the first half of the year switches to a high-pay job in another firm for the second half of the year, he would appear as a medium-pay hire in the new firm on the basis of his annual earnings although he is actually a high-pay hire. When average wages are used, this would not be an issue. Similar problems appear for separations. An employee who switches to another firm in the middle of the year, would contribute to the decomposition with his earnings in the previous year. The wage paid during the first half of the year of separation would not be included, unlike in the average wage. Further work is needed to investigate potential biases in both approaches.²¹

References

- Abowd JM, Kramarz F (1999) The analysis of labor markets using matched employer-employee data. In: Ashenfelter O and Card D (eds) *Handbook of Labor Economics*, volume 3b. Elsevier, Amsterdam, pp. 2629-2710
- Akerberg D, Benkard CL, Berry S, Pakes A (2007) Econometric tools for analyzing market outcomes. In: Heckman JJ, Leamer EE (eds) *Handbook of Econometrics*, Volume 6. Elsevier, Amsterdam, pp. 4171-4276
- Aghion P, Howitt, P (1996) The observational implications of Schumpeterian growth theory. *Empirical Economics* 21: 13-25
- Althin R, Färe R, Grosskopf, S (1996) Profitability and productivity changes: An application to Swedish pharmacies. *Annals of Operations Research* 66: 219-230
- Aubert P, Caroli E, Roger, M (2006) New technologies, organisations and age: Firm-level evidence. *The Economic Journal* 116: 73-93
- Aubert P, Crépon B (2003) La productivité des salariés âgés: une tentative d'estimation. *Économie et Statistique* n°368: 95-119
- Balk B M (2010) Measuring productivity change without neoclassical assumptions: A conceptual analysis. In: Diewert WE, Balk BM, Fixler D, Fox KJ, Nakamura A (eds) *Price and Productivity Measurement, Vol. 6 - Index Number Theory*. Trafford Press, Victoria, pp. 133-154
- Balk, BM (2015) The dynamics of productivity change: A review of the bottom-up approach. Available at SSRN 2585452
- Balsvik R (2011) Is labor mobility a channel for spillovers from multinationals? Evidence from Norwegian manufacturing. *Review of Economics and Statistics* 93: 285-297
- Bingley P, Westergaard-Nielsen N (2004) Personnel policy and profit. *Journal of Business Research* 57: 557-563
- Boschma R, Eriksson R, Lindgren U (2009) How does labour mobility affect the performance of plants? The importance of relatedness and geographical proximity. *Journal of Economic Geography* 9: 169-190
- Cataldi A, Kampelmann S, Rycx F (2011) Productivity-wage gaps among age groups: Does the ICT environment matter? *De Economist* 159: 193-221
- Dalton DR, Todor WD, Krackhardt DM (1982) Turnover overstated: The functional taxonomy. *Academy of Management Review* 7: 117-123
- Daniel K, Heywood JS (2007) The determinants of hiring older workers: UK evidence. *Labour Economics* 14: 35-51
- Daveri F (2004) Delayed IT usage: Is it really the drag on Europe's productivity. *CESifo Economic Studies* 50: 397-421
- Daveri F, Maliranta M (2007) Age, seniority and labour costs: Lessons from Finnish IT revolution. *Economic Policy* 22: 117-175
- Diewert WE, Fox KA (2009) On measuring the contribution of entering and exiting firms to aggregate productivity growth. In: Diewert WE, Balk BM, Fixler D, Fox KJ, Nakamura A (eds) *Price and Productivity Measurement, Vol. 6 - Index Number Theory*. Trafford Press, Victoria, pp.41-66
- Foster L, Haltiwanger J, Krizan CJ (2001) Aggregate productivity growth: Lessons from microeconomic evidence. In: Hulten CR, Dean ER, Harper MJ (eds) *New Developments in Productivity Analysis*. University of Chicago Press, Chicago, pp. 303-363

- Foster L, Haltiwanger J, Syverson C (2008) Reallocation, firm turnover, and efficiency: Selection on productivity or profitability? *American Economic Review* 98: 394-425
- Göbel C, Zwick T (2012) Age and productivity: Sector differences. *De Economist* 160: 35-57
- Göbel C, Zwick T (2013) Are personnel measures effective in increasing productivity of old workers? *Labour Economics* 22: 80-93
- Griliches Z, Mairesse J (1998) Production functions: The search for identification. In: Strøm R (ed) *Econometrics and Economic Theory in the Twentieth Century: The Ragnar Frisch Centennial Symposium*. Cambridge University Press, Cambridge, pp. 169-203
- Hadi A S (1992) Identifying multiple outliers in multivariate data. *Journal of the Royal Statistical Society, Series B* 54: 761-771
- Hadi AS (1994) A modification of a method for the detection of outliers in multivariate samples. *Journal of the Royal Statistical Society, Series B* 56: 393-396
- Hakola T, Uusitalo R (2005) Not so voluntary retirement decisions? Evidence from a pension reform. *Journal of Public Economics* 89: 2121-2136
- Haltiwanger J (1997) Measuring and analyzing aggregate fluctuations: The importance of building from microeconomic evidence. *Federal Reserve Bank of St. Louis Economic Review* 79: 55-77
- Haltiwanger JC, Lane JI, Spletzer JR (1999) Productivity differences across employers: The roles of employer size, age, and human capital. *American Economic Review* 89: 94-98
- Haltiwanger JC, Lane JI, Spletzer JR (2007). Wages, productivity, and the dynamic interaction of businesses and workers. *Labour Economics* 14: 575-602
- Hellerstein JK, Neumark D (2007) Production function and wage equation estimation with heterogeneous labor: Evidence from a new matched employer-employee data set. In: Berndt ER, Hulten CM (eds) *Hard to Measure Goods and Services: Essays in Honor of Zvi Griliches*. University of Chicago Press, Chicago, pp. 31-71
- Hellerstein JK, Neumark D, Troske KR (1999) Wages, productivity, and worker characteristics: Evidence from plant-level production functions and wage equations. *Journal of Labor Economics* 17: 409-446
- Hutchens R (1986) Delayed payment contracts and a firm's propensity to hire older workers. *Journal of Labor Economics* 4: 439-457
- Hutchens R (1988) Do job opportunities decline with age? *Industrial and Labor Relations Review* 42: 89-99
- Ilmakunnas P, Ilmakunnas S (2015) Hiring older employees: Do the age limits of early retirement and the contribution rates of firms matter? *Scandinavian Journal of Economics* 117: 164-194
- Ilmakunnas P, Maliranta, M (2005) Technology, labour characteristics and wage-productivity gaps. *Oxford Bulletin of Economics and Statistics* 67: 623-644
- Ilmakunnas P, Maliranta M, Vainiomäki J (2004) The roles of employer and employee characteristics for plant productivity. *Journal of Productivity Analysis* 21: 249-276
- Ilmakunnas P, Maliranta, M, Vainiomäki J (2005) Worker turnover and productivity growth. *Applied Economics Letters* 12: 395-398
- Ilmakunnas P, van Ours J, Skirbekk V, Weiss, M (2010) Age and productivity. In: Garibaldi P, Martins JO, van Ours J (eds) *Ageing, Health, and Productivity:*

- The Economics of Increased Life Expectancy. Oxford University Press, Oxford, pp. 133-240
- Jovanovic B (1979) Job matching and the theory of turnover. *Journal of Political Economy* 87: 972-990
- Kaiser U, Kongsted HC, Rønde T (2015) Does the mobility of R&D labor increase innovation? *Journal of Economic Behavior and Organization* 110: 91-105
- Kronenberg K, Carree M (2010) The effects of workforce composition, labor turnover, and the qualities of entering and exiting workers on productivity growth. Maastricht University, mimeo
- Kyyrä T, Wilke RA (2007) Reduction in the long-term unemployment of the elderly: A success story from Finland. *Journal of the European Economic Association* 5: 154-182
- Lazear EP (1979) Why is there mandatory retirement? *Journal of Political Economy* 87: 1261-1284
- Levinsohn J, Petrin A (2003) Estimating production functions using inputs to control for unobservables. *Review of Economic Studies* 70: 317-341
- Mahlberg B, Freund I, Crespo Cuaresma J, Prskawetz A (2013a) Ageing, productivity and wages in Austria. *Labour Economics* 22: 5-15
- Mahlberg B, Freund I, Crespo Cuaresma J, Prskawetz A (2013b) The age-productivity pattern: Do location and sector affiliation matter? *The Journal of the Economics of Ageing* 1-2: 72-82
- Maliranta M (1997) Plant-level explanations for the catch-up process in Finnish manufacturing: A decomposition of aggregate labour productivity growth. In: Laaksonen S (ed) *The Evolution of Firms and Industries. International Perspectives*. Statistics Finland, Helsinki, pp. 352-369
- Maliranta M, Ilmakunnas P (2005) Decomposing productivity and wage effects of intra-establishment labor restructuring. *ETLA Discussion Papers No. 993*, Helsinki
- Maliranta M, Mohnen P, Rouvinen P (2009) Is inter-firm labor mobility a channel of knowledge spillovers? Evidence from a linked employer-employee panel. *Industrial and Corporate Change* 18: 1161-1191
- Munnell AH, Sass SA (2008) *Working Longer. The Solution to the Retirement Income Challenge*. Brookings Institution Press, Washington, D.C.
- Nelson RR, Phelps ES (1966) Investment in humans, technological diffusion, and economic growth. *American Economic Review, Papers and Proceedings* 56: 69-75
- Olley GS, Pakes A (1996) The dynamics of productivity in the telecommunications equipment industry. *Econometrica* 64: 1263-1297
- Parrotta P, Pozzoli D (2012) The effect of learning by hiring on productivity. *RAND Journal of Economics* 43: 167-185
- Siebert WS, Zubanov N (2009) Searching for the optimal level of employee turnover: A study of a large U.K. retail organization. *Academy of Management Journal* 52: 294-313
- Skirbekk V (2008) Age and productivity potential: A new approach based on ability levels and industry-wide task demand. *Population and Development Review* 34 (Supplement): 191-207
- Stoyanov A, Zubanov N (2012) Productivity spillovers across firms through worker mobility. *American Economic Journal: Applied Economics* 4: 168-198
- Towers Perrin (2005) *The Business Case for Workers Age 50+*. AARP, Washington DC

- Vainiomäki J (1999) Technology and skill upgrading: Results from linked worker-plant data for Finnish manufacturing. In: Haltiwanger J, Lane J, Spletzer JR, Theuvs JM, Troske KR (eds) *The Creation and Analysis of Employer-Employee Matched Data*. Elsevier, Amsterdam, pp. 115-145
- van Ark B, Inklaar R, McGuckin R (2003) ICT and productivity in Europe and the United States. Where do the differences come from? *CESifo Economic Studies* 49: 295-318
- van Dalen HP, Henkens K, Schippers J (2010) Productivity of older workers: Perceptions of employers and employees. *Population and Development Review* 36: 309-330
- van Ours JC, Stoeldraijer L (2011) Age, wage and productivity in Dutch manufacturing. *De Economist* 159: 113-137
- Vandenberghe V (2010) Shed the oldies ... it's profitable. Universite Catholique de Louvain, mimeo
- Vandenberghe V (2013) Are firms willing to employ a greying and feminizing workforce? *Labour Economics* 22: 30-46
- Vandenberghe V, Waltenberg F, Rigo M (2013) Ageing and employability. Evidence from Belgian firm-level data. *Journal of Productivity Analysis* 40: 111-136

Variable	Average	p1	Median	p99
Average employment (CD)	85.8	11	26	1010
Average employment (ES)	85.2	11	26	1018
Labor productivity growth rate	0.045	-0.714	0.047	0.766
Wage growth rate (CD)	0.069	-0.413	0.069	0.540
Wage growth rate (ES)	0.067	-0.150	0.067	0.284
Profitability growth rate	-0.023	-0.698	-0.020	0.636
Capital/labor growth rate	0.064	-1.836	0.014	2.229
Change in productivity shock	-0.001	-0.080	0	0.063
Hiring of -30 years	0.141	0	0.115	0.544
Hiring of 31-50 years	0.120	0	0.095	0.438
Hiring of 51- years	0.021	0	0	0.171
Separation of -28 years	0.088	0	0.063	0.455
Separation of -28 years, unemployment.	0.010	0	0	0.100
Separation of -28 years, other	0.078	0	0.056	0.424
Separation of 29-48 years	0.113	0	0.091	0.429
Separation of 29-48 years, unemployment	0.020	0	0	0.167
Separation of 29-48 years, other	0.092	0	0.074	0.381
Separation of 49- years	0.041	0	0.028	0.214
Separation of 49- years, retirement	0.013	0	0	0.111
Separation of 49- years, unemployment	0.014	0	0	0.127
Separation of 49- years, other	0.015	0	0	0.136
Share of staying -30 years	0.174	0	0.138	0.700
Share of staying 31-50 years	0.607	0.176	0.615	1
Share of staying 51- years	0.219	0	0.206	0.615

Note: CD and ES refer to Company Data and Employment Statistics information, respectively. N = 23738.

Table 1. Descriptive statistics of the estimation sample

	Productivity	Wage	Profitability	Productivity	Wage	Profitability
D(Productivity shock)	0.875*** (0.099)	0.060 (0.044)	0.824*** (0.120)			
Dln(Capital/Labor)	0.027*** (0.007)	0.032*** (0.003)	-0.004 (0.007)	0.015** (0.007)	0.031*** (0.003)	-0.016** (0.007)
Hired -30 yrs.	-0.154*** (0.040)	-0.086*** (0.021)	-0.070* (0.036)	-0.116*** (0.041)	-0.083*** (0.021)	-0.034 (0.037)
Hired 31-50 yrs.	-0.140*** (0.040)	-0.005 (0.019)	-0.137*** (0.036)	-0.116*** (0.041)	-0.003 (0.019)	-0.115*** (0.038)
Hired 51- yrs.	0.170 (0.139)	-0.084* (0.048)	0.252* (0.134)	0.246* (0.147)	-0.078 (0.048)	0.324** (0.141)
Separated -28 yrs.	0.103** (0.049)	0.054** (0.025)	0.052 (0.052)	0.053 (0.051)	0.050** (0.025)	0.005 (0.054)
Separated 29-48 yrs.	0.185*** (0.047)	0.046** (0.020)	0.145*** (0.046)	0.143*** (0.049)	0.043** (0.020)	0.105** (0.048)
Separated 49- yrs.	0.521*** (0.084)	0.134*** (0.043)	0.392*** (0.070)	0.523*** (0.085)	0.134*** (0.043)	0.394*** (0.071)
Stayers 31-50 yrs.	-0.007 (0.029)	-0.037** (0.015)	0.030 (0.029)	-0.026 (0.030)	-0.039** (0.015)	0.012 (0.030)
Stayers 51- yrs.	-0.065* (0.034)	-0.075*** (0.016)	0.008 (0.032)	-0.082** (0.035)	-0.076*** (0.016)	-0.008 (0.033)
Observations	23738	23738	23738	23738	23738	23738
R-squared	0.217	0.122	0.219	0.201	0.121	0.201

Note: Non-reported variables include regional dummies and interactions of industry and period dummies. Employment weighted estimation. Robust standard errors in parentheses, clustered at the firm level. Firms with at least 10 and at most 10 000 employees included. * p<0.1, ** p<0.05, *** p<0.01

Table 2. Productivity, wage and profit equations in the business sector

	Industry			Services		
	Productivity	Wage	Profitability	Productivity	Wage	Profitability
D(Productivity shock)	0.951*** (0.095)	-0.006 (0.038)	0.968*** (0.113)	0.541*** (0.169)	0.247** (0.110)	0.299** (0.145)
Dln(Capital/Labor)	0.006 (0.010)	0.031*** (0.004)	-0.025*** (0.009)	0.047*** (0.007)	0.035*** (0.004)	0.013* (0.007)
Hired -30 yrs.	-0.005 (0.058)	0.021 (0.028)	-0.026 (0.055)	-0.282*** (0.053)	-0.191*** (0.031)	-0.095** (0.047)
Hired 31-50 yrs.	-0.123** (0.060)	-0.015 (0.025)	-0.112** (0.052)	-0.214*** (0.051)	-0.009 (0.029)	-0.205*** (0.047)
Hired 51- yrs.	0.220 (0.188)	-0.072 (0.060)	0.288 (0.183)	-0.056 (0.131)	-0.101 (0.083)	0.047 (0.106)
Separated -28 yrs.	0.102 (0.075)	-0.051 (0.037)	0.152** (0.068)	0.172** (0.067)	0.161*** (0.034)	0.016 (0.071)
Separated 29-48 yrs.	0.211*** (0.070)	0.046* (0.024)	0.172*** (0.066)	0.200*** (0.062)	0.071** (0.032)	0.133** (0.062)
Separated 49- yrs.	0.756*** (0.107)	0.178*** (0.054)	0.585*** (0.088)	0.158 (0.128)	0.084 (0.071)	0.077 (0.112)
Stayers 31-50 yrs.	0.007 (0.045)	-0.029* (0.017)	0.035 (0.041)	-0.010 (0.038)	-0.045* (0.023)	0.033 (0.040)
Stayers 51- yrs.	-0.028 (0.049)	-0.078*** (0.019)	0.048 (0.044)	-0.081* (0.046)	-0.075*** (0.025)	-0.009 (0.045)
Observations	12206	12206	12206	11532	11532	11532
R-squared	0.280	0.115	0.300	0.122	0.138	0.085

Note: Non-reported variables include regional dummies and interactions of industry and period dummies. Employment weighted estimation. Robust standard errors in parentheses, clustered at the firm level. Firms with at least 10 and at most 10 000 employees included. * p<0.1, ** p<0.05, *** p<0.01

Table 3. Productivity, wage and profit equations by sectors

	Industry			Services		
	Productivity	Wage	Profitability	Productivity	Wage	Profitability
D(Productivity shock)	0.951*** (0.096)	-0.004 (0.037)	0.966*** (0.113)	0.553*** (0.169)	0.256** (0.109)	0.303** (0.144)
Dln(Capital/Labor)	0.006 (0.010)	0.032*** (0.004)	-0.025*** (0.010)	0.047*** (0.007)	0.035*** (0.004)	0.013** (0.007)
Hired -30 yrs.	-0.014 (0.058)	0.011 (0.028)	-0.025 (0.056)	-0.277*** (0.053)	-0.186*** (0.031)	-0.094** (0.047)
Hired 31-50 yrs.	-0.125** (0.060)	-0.018 (0.024)	-0.111** (0.053)	-0.213*** (0.052)	-0.010 (0.029)	-0.203*** (0.048)
Hired 51- yrs.	0.232 (0.187)	-0.068 (0.061)	0.297 (0.184)	-0.068 (0.130)	-0.106 (0.083)	0.040 (0.106)
Separated -28 yrs., other destination	0.162** (0.075)	0.006 (0.035)	0.155** (0.068)	0.143* (0.075)	0.154*** (0.037)	-0.006 (0.082)
Separated -28 yrs., unemployment	-0.154 (0.203)	-0.324** (0.134)	0.168 (0.204)	0.422** (0.177)	0.163 (0.120)	0.257 (0.167)
Separated 29-48 yrs., other destination	0.351*** (0.084)	0.092*** (0.030)	0.267*** (0.082)	0.185*** (0.066)	0.059* (0.033)	0.131** (0.063)
Separated 29-48 yrs., unemployment	-0.110 (0.147)	-0.111* (0.067)	0.005 (0.140)	0.107 (0.130)	0.009 (0.071)	0.098 (0.119)
Separated 49- yrs., other destination	0.348* (0.191)	0.173* (0.091)	0.186 (0.175)	0.230 (0.198)	0.268*** (0.102)	-0.030 (0.198)
Separated 49- yrs., unemployment	1.020*** (0.172)	0.187** (0.078)	0.839*** (0.155)	0.392** (0.191)	-0.105 (0.125)	0.498*** (0.148)
Separated 49- yrs., pension	0.757*** (0.190)	0.174** (0.070)	0.586*** (0.178)	-0.251 (0.279)	0.012 (0.143)	-0.267 (0.231)
Stayers 31-50 yrs.	0.003 (0.044)	-0.029* (0.017)	0.032 (0.040)	-0.011 (0.037)	-0.043* (0.023)	0.031 (0.040)
Stayers 51- yrs.	-0.027 (0.049)	-0.078*** (0.018)	0.050 (0.044)	-0.071 (0.047)	-0.072*** (0.024)	-0.001 (0.045)
Observations	12206	12206	12206	11532	11532	11532
R-squared	0.282	0.117	0.301	0.123	0.139	0.087

Note: Non-reported variables include regional dummies and interactions of industry and period dummies. Employment weighted estimation. Robust standard errors in parentheses, clustered at the firm level. Firms with at least 10 and at most 10 000 employees included. * p<0.1, ** p<0.05, *** p<0.01

Table 4. Productivity, wage and profit equations by sector, with disaggregation of exits by destination

	Industry		Services	
	ICT	Non-ICT	ICT	Non-ICT
D(Productivity shock)	1.138*** (0.107)	0.100 (0.304)	0.470 (0.329)	0.264* (0.154)
Dln(Capital/Labor)	-0.047*** (0.012)	-0.006 (0.013)	0.009 (0.011)	0.013* (0.008)
Hired -30 yrs.	-0.144 (0.095)	0.018 (0.067)	-0.197** (0.082)	-0.060 (0.054)
Hired 31-50 yrs.	-0.041 (0.099)	-0.130** (0.060)	-0.270*** (0.079)	-0.173*** (0.058)
Hired 51- yrs.	0.418* (0.222)	0.123 (0.293)	0.003 (0.216)	0.046 (0.127)
Separated -28 yrs., other destination	0.296** (0.140)	0.085 (0.076)	0.315** (0.127)	-0.061 (0.096)
Separated -28 yrs., unemployment	0.525 (0.529)	0.073 (0.206)	0.103 (0.485)	0.278 (0.175)
Separated 29-48 yrs., other destination	0.285** (0.113)	0.254** (0.115)	0.258** (0.109)	0.082 (0.075)
Separated 29-48 yrs., unemployment	0.305 (0.304)	-0.039 (0.155)	-0.546 (0.391)	0.199* (0.119)
Separated 49- yrs., other destination	0.013 (0.290)	0.221 (0.215)	0.229 (0.422)	-0.107 (0.210)
Separated 49- yrs., unemployment	0.809*** (0.259)	0.845*** (0.179)	0.205 (0.346)	0.548*** (0.166)
Separated 49- yrs., pension	1.137*** (0.324)	0.453** (0.209)	-0.598 (0.401)	-0.209 (0.262)
Stayers 31-50 yrs.	0.120 (0.089)	0.019 (0.044)	0.056 (0.064)	0.030 (0.046)
Stayers 51- yrs.	0.064 (0.093)	0.057 (0.051)	0.147* (0.078)	-0.033 (0.051)
Observations	2990	9216	2223	9309
R-squared	0.429	0.253	0.116	0.086

Note: Non-reported variables include regional dummies and interactions of industry and period dummies. Employment weighted estimation. Robust standard errors in parentheses, clustered at the firm level. Firms with at least 10 and at most 10 000 employees included. * p<0.1, ** p<0.05, *** p<0.01

Table 5. Profitability equations by ICT/non-ICT division and sector, with disaggregation of exits by destination

	Industry		Services	
	Declining	Growing	Declining	Growing
D(Productivity shock)	0.338 (0.430)	1.015*** (0.114)	0.424* (0.244)	0.326** (0.158)
Dln(Capital/Labor)	-0.005 (0.016)	-0.031*** (0.010)	0.030** (0.014)	0.006 (0.006)
Hired -30 yrs.	-0.179 (0.118)	-0.074 (0.076)	0.124 (0.128)	-0.138** (0.057)
Hired 31-50 yrs.	-0.213* (0.124)	-0.146** (0.068)	0.100 (0.157)	-0.189*** (0.059)
Hired 51- yrs.	-0.399* (0.231)	0.503** (0.206)	0.373 (0.256)	0.020 (0.129)
Separated -28 yrs., other destination	0.021 (0.126)	0.325*** (0.088)	-0.381** (0.162)	0.142** (0.071)
Separated -28 yrs., unemployment	0.211 (0.335)	0.191 (0.208)	-0.015 (0.273)	0.246 (0.176)
Separated 29-48 yrs., other destination	0.280** (0.118)	0.385*** (0.108)	0.090 (0.143)	0.036 (0.068)
Separated 29-48 yrs., unemployment	0.016 (0.194)	-0.009 (0.175)	-0.138 (0.216)	-0.042 (0.162)
Separated 49- yrs., other destination	0.229 (0.227)	0.364* (0.211)	-0.211 (0.326)	0.133 (0.164)
Separated 49- yrs., unemployment	0.956*** (0.198)	0.464* (0.263)	0.357 (0.246)	0.317* (0.188)
Separated 49- yrs., pension	0.656*** (0.243)	0.603** (0.253)	-0.593 (0.398)	-0.488* (0.292)
Stayers 31-50 yrs.	0.072 (0.074)	-0.021 (0.046)	0.017 (0.057)	0.025 (0.046)
Stayers 51- yrs.	-0.003 (0.078)	0.076 (0.048)	-0.027 (0.076)	0.005 (0.058)
Observations	4596	6589	3352	7089
R-squared	0.334	0.374	0.163	0.114

Note: Non-reported variables include regional dummies and interactions of industry and period dummies. Employment weighted estimation. Robust standard errors in parentheses, clustered at the firm level. Firms with at least 10 and at most 10 000 employees included. * p<0.1, ** p<0.05, *** p<0.01

Table 6. Profitability equations by firm growth and sector, with disaggregation of exits by destination

Industry	N	Percent
(10-14) Mining and quarrying	97	0.41
(15-16) Food, beverages & tobacco	770	3.24
(17-18) Textiles & products	466	1.96
(19) Leather & products	108	0.45
(20) Wood & products	649	2.73
(21-22) Paper & printing	1107	4.66
(23) Coke and petroleum	6	0.03
(24) Chemicals	313	1.32
(25) Rubber & plastic	471	1.98
(26) Non-metallic mineral	397	1.67
(27-28) Basic metals	1858	7.83
(29) Machinery and equipment	1382	5.82
(30-33) Electrical equipment	738	3.11
(34-35) Transport equipment	407	1.71
(36-37) Manufacturing n.e.c.	622	2.62
(40-41) Public utilities	304	1.28
(45) Construction	2511	10.58
(50-52) Wholesale and retail trade	5654	23.82
(55) Hotels & restaurants	728	3.07
(60-64) Transportation	1739	7.33
(65-74) Other business	2665	11.23
(85&92-93) Miscellaneous services	746	3.14
Total	23738	100

Table A.1. The number of observations in the estimation sample by branches

Endnotes

¹ There is a field of literature where the determinants of hiring and employing older workers are analyzed (e.g. Hutchens, 1986, 1988; Aubert et al., 2006; Daniel and Heywood, 2007; Ilmakunnas and Ilmakunnas, 2015).

² One dictionary definition of the adjective dysfunctional is “characterized by a breakdown of normal or beneficial relationships between members of the group” (<http://www.collinsdictionary.com>).

³ The positive influences of turnover have been emphasized more formally in models where the search and matching process allocates workers to their best uses in firms (e.g. Jovanovic, 1979). Worker flows and the matching process may be particularly important for productivity when technological change is rapid (see e.g. Aghion and Howitt, 1996).

⁴ Vandenberghe (2010) has adopted our decomposition approach.

⁵ Often the labor input is disaggregated by using parameters that measure the relative productivities of the worker types (e.g. Hellerstein, Neumark and Troske, 1999). In this case, the production function is $Y = f(q_1 L_1 + \dots + q_M L_M)$ and the productivity terms are $g_j = q_j f_j \frac{\partial}{\partial X^0}$.

⁶ The decomposition is related to those used commonly in firm or plant-level productivity analysis (e.g. J. Haltiwanger, 1997), but is closer to those used by Maliranta (1997), Vainiomäki (1999), Maliranta and Ilmakunnas (2005) as well as Diewert and Fox (2009) in different contexts.

⁷ Our profit variable, the ratio of revenues and costs, is related to profitability measures used in productivity analysis (e.g. Balk, 2010, Althin et al., 1996), where it is sometimes called “return to the dollar”. If we consider only value added and assume capital fixed in the short run, our measure is the ratio of net revenue (value added Y) and variable costs. It can also be interpreted as a mark-up. Assume that firms set price equal to unit labor cost multiplied by a mark-up $1+m$, so that we have revenue $Y = (1+m)W(1+a)$. The operating margin can be written as $OPM = Y - W(1+a) = (1+m)W(1+a) - W(1+a) = mW(1+a)$, and the profit variable equals $\Pi = 1+m$.

⁸ The decomposition model is based on the assumption of constant returns to scale, but deviations from constant returns would show up as heteroscedasticity.

⁹ We leave out the period 1990-1994, because our company data are substantially less comprehensive before the year 1995.

¹⁰ It is assumed that the price development is the same for all firms within an industry. If there are firm differences in price changes, some bias may be caused, as explained in section 4 of this article. These dummies also account for changes in the payroll taxes a , which were assumed fixed in the derivation of the formulas in Section 3.

¹¹ There is some measurement error because temporary agency workers or those with an atypical employment relationship, like service contract, are not included in the employment of the firms. According to the Finnish Private Employment Agencies Association, their employees accounted for only 1% of total employment in 2010, which is below the European average. This share has been increasing in recent years, but it was still low in the period of our analysis. This kind of employment is more common in services. This is one reason why we estimate the models also by sector (industry and services). In principle, owner-managers can also cause measurement error. For small firms where the owner is working, he is included in the number of employees only if he takes at least half of his income as salary (as opposed to capital income). In any case, the restrictions we use (i.e., leaving out firms with less than ten employees and firms with missing employer-firm link) mean that this measurement error should not be large. We thank an anonymous referee for pointing out these issues to us.

¹² The proxy variable approach is based on the assumption that capital is inherited from the previous period, and the proxy variable (materials) and labor choices happen after the shock is realized. Therefore, the latter variables are endogenous, but capital stock itself is not.

¹³ Some of the firms operate in more than one region. The region of a firm refers to the one where the employment share is the highest.

¹⁴ The number of observations drops because of the reasons discussed in the text.

¹⁵ Note that these figures underestimate actual turnover among the employees, since e.g. hiring of an employee after the start of a period and subsequent separation of the same employee before the end of the period is not included in the turnover rates.

¹⁶ There are two main reasons why we prefer using employment weighted estimation. Firstly, we are ultimately concerned of productivity differences between different employment groups. Unfortunately we are unable to measure productivity at the level individuals but only at the level of firms so that in a sense we are using aggregated data. In order to give an equal weight for each individual in our analysis, we should give a larger weight to large firms than to smaller firms. Second, weighted estimation provides us with a more efficient procedure in the presence of heteroscedasticity.

¹⁷ In the Finnish pension system the pension was until 1996 based on the last four years' pay and until 2004 on the last ten year's pay in each employment relationship, which gave incentives for obtaining a high pay at the end of the career. This combination of backloaded wage and a fixed retirement age is consistent with the deferred payment model of Lazear (1979), although it is a result of a quite different institutional setting. The system has been based on a mix of centralized negotiations between labor unions, employer organizations and the government, and firm-level wage setting. Lazear's argument is somewhat difficult to use in the connection of labor flows, as it is an equilibrium model where the raising wage profile and unprofitability of the older employees is part of the "package". However, even in this case unexpected increases in the costs of older employees or an increase in the share of older employees through aging may make the system unsustainable and give raise to incentives for separations.

¹⁸ The probability of exit was modelled as a function of firm size, industry, employee characteristics (average tenure and education years, share of women), capital-labor ratio, productivity shock, and an indicator for growing firms. All of these variables were measured in the year prior to exit.

¹⁹ In addition to the unemployment pension system, also disability pension gives incentives for laying off older employees. The larger firms are responsible for paying (part of) the disability pension until the normal retirement age. This can be avoided, if the worker goes to the unemployment pension tunnel before possible disability.

²⁰ See discussion in Daveri (2004). The results obtained by using the broad definition of ICT are available on request. It should be noted that "Financial intermediation" (ISIC 65-67) industries are excluded from our estimation sample.

²¹ Also the econometric framework could be extended. We have used the proxy variable approach indirectly by estimating the shock from a level form and then using it in our difference form model. It remains an open issue how the approach could be directly applied to our model in one step. Also extensions to for example stochastic frontier models would be interesting.