

Peters, Jan Cornelius

Working Paper

Do age complementarities affect labor productivity? Evidence from German firm level data

Economics Working Paper, No. 2016-10

Provided in Cooperation with:

Christian-Albrechts-University of Kiel, Department of Economics

Suggested Citation: Peters, Jan Cornelius (2017) : Do age complementarities affect labor productivity? Evidence from German firm level data, Economics Working Paper, No. 2016-10, Kiel University, Department of Economics, Kiel

This Version is available at:

<https://hdl.handle.net/10419/168033>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Do age complementarities affect labor productivity? Evidence from German firm level data

by Jan Cornelius Peters



Do age complementarities affect labor productivity?

Evidence from German firm level data*

Jan Cornelius Peters^{a,b}

^aJohann Heinrich von Thünen Institute, Institute of Rural Studies,
Bundesallee 50, 38116 Braunschweig, Germany, Cornelius.Peters@thuenen.de, +49/531/596-5503

^bIAB Northern Germany, Regional Research Network of the Institute for Employment Research,
Projensdorfer Str. 82, 24106 Kiel, Germany

July 11, 2017
(slightly revised version)

Abstract

In Germany, as in many other European countries, there will be a shift in the workforce age structure in the next decades. The number of older workers will increase, and the number of younger and middle aged workers will decline. This paper provides evidence how the shift in the relative labor supply affects labor productivity, taking into account that differently aged workers are suggested to be imperfect substitutes. Using a cross sectional linked employer-employee data set from 2012, translog cost functions are estimated. To control for the skill level of the workers a nested production structure is applied. This allows to analyze age complementarities within groups of workers that have a comparable skill level. Based on the estimated parameters, pairwise elasticities of complementarity and factor price elasticities are computed. The results indicate that workers that belong to different age groups are complementary factors. But the degree of complementarity differs, depending on the age and the skill level of the workers. The complementarities especially arise between younger and middle aged workers. The highest degree of complementarity is observed between younger and middle aged high skilled labor. Simulating how the expected shift in the age structure affects labor productivity indicates that the productivity of younger and middle aged workers will increase. In contrast, the productivity of older workers will significantly decline caused by their increasing share in the workforce.

JEL classification: C31, D24, J11, J31

Keywords: Age complementarities, demographic change, labor-labor substitution, linked employer-employee, translog

*I gratefully thank Antje Buche, Katerina Homolkova, Monika Jungbauer-Gans and especially Johannes Bröcker and Annekatrin Niebuhr for their very helpful remarks and suggestions. Furthermore, I have benefited from fruitful discussions with seminar participants at the Institute for Employment Research, Ifo Dresden, the IfR/SDU doctoral seminar 2014 and the Annual Meeting of the German Economic Association (VfS) 2015. Financial support from the German Research Foundation (DFG) is gratefully acknowledged as part of the project "Diversity and individual careers" (NI 596/5-1). The usual disclaimer applies. An earlier version of this paper has been published as conference paper in the series 'Beiträge zur Jahrestagung des Vereins für Socialpolitik 2015: Ökonomische Entwicklung - Theorie und Politik - Session: Empirical Labor Economics', URL: <https://www.econstor.eu/handle/10419/112941>.

1 Introduction

In the next decades, Germany, as well as many other European countries, will face a shift in the age structure of the workforce due to demographic change. According to a labor market projection commissioned by the Federal Ministry of Labor and Social Affairs, the labor force aged 15 to 24 years will decline from 2010 to 2030 by almost one million persons, and the labor force aged 25 to 54 years by almost five million. On the contrary, the labor supply of workers older than 55 years will increase by almost three million individuals (BMAS, 2013). Considering a certain degree of complementarity between the workers of each group, it is to be expected that such a shift in age structure leads to downward pressure on the productivity, and thereby on the wage rate, of older workers because their share in total employment increases. In contrast, the productivity and the wage rate of younger and middle aged workers presumably will increase since they become relatively rare. Quantitatively, these effects depend on the degree of complementarity between workers that are of a different age. The larger the degree of complementarity between them, the larger is the effect of demographic change on the (relative) productivity of age groups.

Complementary relationships between differently aged workers are supposed to exist as the sets of characteristics relevant in the production process differ across younger and older workers (Garloff/Roth, 2016). Younger workers are stronger in generating and recombining new knowledge, as found by Horn and Cattell (1967) who refer to this as ‘fluid intelligence’. Furthermore, they typically have more physical strength and agility. In contrast, older workers tend to have more experience and abilities to understand difficult situations as well as to use and apply existing skills, i.e., ‘crystallized intelligence’ (Horn/Cattell, 1967; Börsch-Supan/Weiss, 2007; Börsch-Supan, 2013). Empirical results by, e.g., Göbel and Zwick (2013) support that differences in abilities cause complementarities between younger and older workers. Based on German linked employer-employee data, they find that the productivity of younger and older workers in establishments with mixed-age working teams is significantly higher than in other establishments. They relate the productivity gains to the knowledge exchange between different age groups and to benefits of sharing tasks according to comparative advantages (Göbel/Zwick, 2013).

In this paper, I quantify the degree of complementarity of workers that belong to different age groups by estimating cross factor price elasticities. These indicate how the wage rate of age group i changes due to a change in the quantity of workers that belong to age group j . This is of special interest, as it provides information on the impact of demographic change, as well as of any policy that impacts on relative labor supply, such as changes in retirement age, on relative wages. Based on the estimated elasticities, I thereby apply model calculations and show how the expected demographic change might impact on relative wages of younger, middle aged, and older workers in Germany.

In contrast to most previous papers, my analysis is based on a linked employer-employee data set. This allows the investigation of age complementarities on establishment level, i.e., directly at the workplace. It is likely that this yields better estimates of the parameters characterizing establishments’ labor demand than aggregate data (Hamermesh, 1996). Most other studies analyze complementarities between different groups of workers using data referring to the regional or national level. These studies might suffer from ecological fallacy as aggregate results are interpreted as if they pertain to behavior of firms (see discussion by, e.g., Naz et al., 2015).

Taking into account the interaction of skill and age, I analyze age complementarities on the establishment level within groups of low, medium, and high skilled workers by estimating translog cost functions. Thereby, I impose less restrictions on the elasticities as compared with using CES-functions, which are often used in this context. One basic assumption of a CES-function is that the elasticity between any two groups of workers that

differ in age is the same, independently of how old the workers within the two groups are. In contrast, I provide empirical evidence that the degree of complementarity varies between age groups.

The remainder of this paper is structured as follows. In section 2 I review the related literature. In section 3 I describe the theoretical framework of my empirical analysis, in section 4 the empirical specification and in section 5 the data set. In section 6 I discuss the results of my empirical analysis, and in section 7 I set out my conclusions.

2 Related literature

Several studies assess the imperfect substitutability of differently aged workers by analyzing cohort effects (e.g., Freeman, 1979; Welch, 1979; Berger, 1985; Sapozhnikov/Triest, 2007; Morin, 2015). Under the hypothesis that workers of different age are imperfect substitutes, an increase in the relative size of an age group has a negative effect on the wage rate of the members of that group. For Germany, the impact of cohort size on cohort specific wages was recently analyzed by Garloff and Roth (2016). They estimate the effect of changes in the share of the youth in total regional employment on the wages of youth workers within the regional labor market. As expected, their findings indicate that wages of younger workers decrease if their share in total regional employment increases.

Similarly, Card and Lemieux (2001) relate the raise in the cohort specific college wage premium to cohort specific shifts in the relative supply of highly educated workers. They observe for the United States, the United Kingdom, and Canada that the college wage premium for younger workers significantly increased relative to the college wage premium for older workers. The empirical analysis indicates that this can be explained by the observation that the age-group specific relative supply of college-educated labor grew faster in the group of older workers than in the group of younger workers.

In their analysis Card and Lemieux (2001) employ a nested CES-framework. This strategy has been adopted by several scholars. Recently, there are several studies focusing on the impact of immigration on wages of natives (e.g., Borjas, 2003; Ottaviano/Peri, 2012; Haas et al., 2013). They apply a nested CES-framework and aim at estimating the degree of complementarity between natives and immigrants within groups of workers that are similar with respect to education and work experience.

Fitzenberger and Kohn (2006) use the CES-framework described by Card and Lemieux (2001) to analyze complementarities between differently skilled workers for Germany. They estimate the parameters of a production function which refers to the national level using information on a random sample of employees based on the IAB employment subsample (IABS). When analyzing complementarities between differently skilled workers, Fitzenberger and Kohn (2006) take into account the imperfect substitutability of age groups, as they model the amount of differently skilled labor as CES-subaggregates of the skill-specific employment quantities of individuals in different age groups. Their findings confirm that differently aged employees are imperfect substitutes.

However, one drawback of the nested CES-framework applied by Card and Lemieux (2001), Fitzenberger and Kohn (2006) and others is that it is based on “strong functional form assumptions” (Fitzenberger/Kohn, 2006: 24). By modelling skill-specific employment quantities as CES-subaggregates of different age groups, it is assumed that the elasticity of substitution between any pair of differently aged workers is the same, i.e., the degree of complementarity between younger and middle aged workers is the same as between younger and older workers as well as between middle aged and older workers (Hamermesh/Grant, 1979; Ottaviano/Peri,

2012). Suppose, for example the abilities and the experience of a worker change monotonically with her age, then this assumption likely is too restrictive.

As an alternative to a CES-function, a more flexible translog-function may be used to describe the imperfect substitutability of differently aged workers in the production process. It allows the degree of complementarity between age groups to vary. So far, there are only very few studies estimating age complementarities based on translog-functions. Grant and Hamermesh (1981) distinguish between younger and different types of older labor. Their findings for the United States indicate that the degree of complementarity between the different groups of workers indeed differs. Similarly, Arntz and Gregory (2014) find that differently aged workers are complements in the knowledge production process using regional data for Germany. However, this is only true for younger and older workers. Younger and middle aged workers as well as middle aged and older workers seem to be perfect substitutes.

One drawback of the studies by Grant and Hamermesh (1981) and Arntz and Gregory (2014) is that they disregard the interaction of skill and age. As stressed by Fitzenberger and Kohn (2006) it is important to analyze age complementarities within groups of workers that have a comparable skill level.

Taking this into account, I distinguish workers according to their skill level and approximate skill specific labor input by a flexible translog-function rather than by a more restrictive CES-function allowing the degree of complementarity to vary between age groups. In order to quantify the parameters that describe firms' demand for workers of different age, I estimate an establishment level cost function based on a linked employer-employee data set. Estimating a cost function is based on the weaker assumption that firms minimize costs rather than assuming profit maximization (Hamermesh/Grant, 1979).¹

3 Theoretical framework

To identify complementarities between different age groups, I apply a nested production framework which is in line with neoclassical production theory. This well established approach (see, e.g., Card/Lemieux, 2001; Borjas, 2003; Fitzenberger/Kohn, 2006; Ottaviano/Peri, 2012; Haas et al., 2013) is based on the assumption that the demand for different types of labor is separable from the demand for other inputs (Hamermesh, 1996). As illustrated by figure 1, the top level describes the output of establishment f , Y_f , as a function of technology A_f , intermediate products X_f , capital κ_f , and labor L_f .

The labor aggregate L_f is distinguished according to the qualification of the workers (index: q). On the next level the skill specific labor input L_{qf} itself is partitioned into groups that differ according to the age of the workers (index: $i \in \{y, m, o\}$): *younger workers* (y), *middle aged workers* (m), *older workers* (o). It is important to note that L_{qf} is not the pure sum of, respectively, low, medium, and high skilled workers employed by the establishment, but compound labor input, that also depends on complementarities between differently aged workers with skill level q . This paper's aim is to quantify the degree of complementarity between differently aged workers. Therefore, the empirical analysis focuses on the analysis of the third level of the nested production framework, i.e., on the age composition of skill-specific labor inputs.

Assuming establishment f choose L_{iqf} , the amount of labor with age i and skill level q , such that the costs to accumulate L_{qf} units of compound labor with skill level q at wage rate vector \mathbf{w}_{qf} are minimized, then, the total

¹ Hamermesh and Grant (1979) also note that one drawback of estimating a cost function is that it requires to invert the matrix of estimated coefficients to obtain the elasticities of complementarity. A problem arises if one parameter has a very high standard error as this reflects in the standard error of the computed elasticities. However, my analysis does not suffer from this.

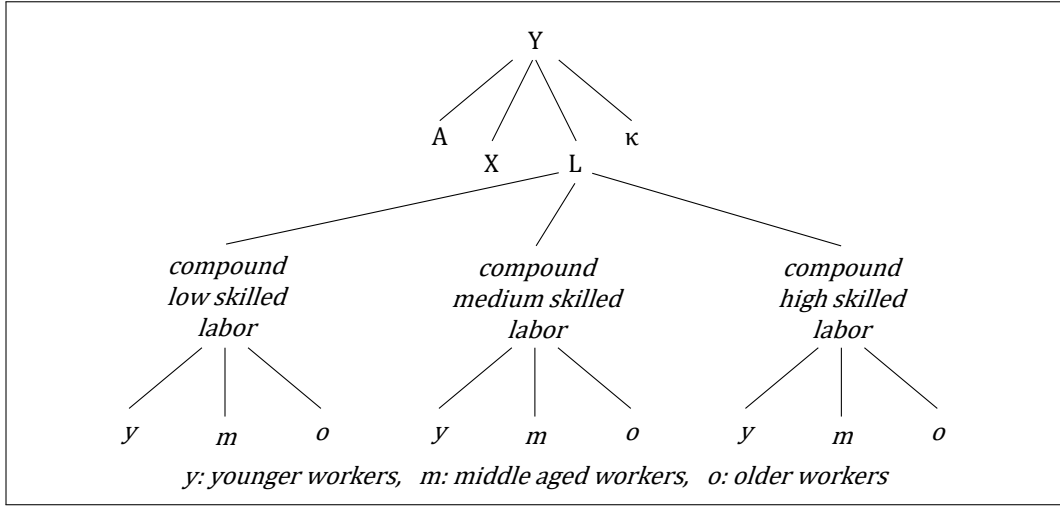


Figure 1: Nested production framework

costs of L_{qf} are given by $C_q(\mathbf{w}_{qf}, L_{qf})$. As suggested by Christensen et al. (1973) and others, $\ln(C_q(\mathbf{w}_{qf}, L_{qf}))$ can be approximated by a function that is quadratic in the logarithms, see also Greene (2012):

$$\ln(C_q(\mathbf{w}_{qf}, L_{qf})) = \alpha_q + \sum_i \alpha_{iq} v_{iqf} + \frac{1}{2} \sum_i \sum_j \beta_{ijq} v_{iqf} v_{jqf} + l_{qf} \quad (1)$$

with $v_{iqf} := \ln(w_{iqf})$ and $l_{qf} := \ln(L_{qf})$

I assume that $C_q(\mathbf{w}_{qf}, L_{qf})$ is linear homogeneous in L_{qf} . This assumption implies that the cost minimizing, relative composition of L_{qf} with respect to the different age groups is independent of the level of compound labor input with skill level q . Its cost minimizing composition is given by (2) which is a system of share equations that depend on the same parameters as (1).² S_{iqf} is the value share of age group i within skill group q and establishment f and the parameters β_{ijq} can be interpreted as share elasticities which indicate the skill level specific response of i 's value share to an increase of j 's wage rate (Jorgenson, 1986).

$$S_{iqf}(\mathbf{w}_{qf}) \equiv \frac{L_{iqf}^* w_{iqf}}{\sum_j L_{jqf}^* w_{jqf}} = \alpha_{iq} + \sum_j \beta_{ijq} v_{jqf}, \quad (2)$$

Estimates of (2) enable computation of the Hicks partial elasticities of complementarity, c_{ijq} , and factor price elasticities, θ_{ijq} (Hicks, 1970; Sato/Koizumi, 1973). In the context of my analysis, the elasticities provide information on age complementarities by skill level. The derivation of the elasticities is given in appendix B.³

In terms of a translog cost function c_{ijq} is given by:

$$c_{ijq} = \frac{|\mathbf{K}_q^{ij}|}{|\tilde{\mathbf{K}}_q|}, \quad (3)$$

where $|\tilde{\mathbf{K}}_q|$ is the determinant of matrix $\tilde{\mathbf{K}}_q$ and $|\mathbf{K}_q^{ij}|$ is the cofactor of K_{ijq} . Matrix $\tilde{\mathbf{K}}_q$ is a transformation of

² (2) has been derived by taking the partial derivative of (1) with respect to v_{iqf} and by applying Shephard's Lemma. See appendix A for a detailed description.

³ As the elasticities depend on the cost shares, the elasticities vary across observations. I omit the index f in the description of the elasticities for simplicity. In the empirical analysis the elasticities are computed at mean predicted shares.

the bordered Hessian matrix⁴ of $C_q(\mathbf{w}_{qf}, L_{qf})$:

$$\tilde{\mathbf{K}}_q = \left[\begin{array}{c|ccc} 0 & k_{1q} & \dots & k_{nq} \\ \hline k_{1q} & & & \\ \vdots & & k_{ijq} & \\ k_{nq} & & & \end{array} \right]$$

with

$$\begin{aligned} k_{iq} &= S_{iq}, \\ k_{ijq} &= \beta_{ijq} + S_{iq}S_{jq}, \text{ if } i \neq j \\ k_{iiq} &= \beta_{iiq} + S_{iq}S_{iq} - S_{iq}. \end{aligned} \quad (4)$$

Based on c_{ijq} the factor price elasticity is given by:

$$\theta_{ijq} = \frac{\partial \ln(w_{iq})}{\partial \ln(L_{jq})} = S_{jq}c_{ijq}. \quad (5)$$

If $c_{ijq} > 0 (< 0)$ the age groups i and j within skill group q are called quantity complements (substitutes) since an increase in L_{jq} leads to an increase (decline) in w_{iq} . θ_{ijq} measures the percentage change in w_{iq} in response to a one percent increase in L_{jq} given wage rates are flexible (Sato/Koizumi, 1973).

4 Empirical specification and econometric issues

Since estimates of translog cost functions like (2) do not, per se, fulfill the properties of a microeconomic cost function, parameter constraints have to be imposed. To maintain consistency with Young's theorem of integrable functions, symmetry constraints have to be fulfilled, i.e., $\beta_{ijq} = \beta_{jiq}$, $i \neq j$. In addition, $C_q(\mathbf{w}_{qf}, L_{qf})$ has to be monotone as well as linear homogeneous and concave in w_{iq} (Jorgenson, 1986).

Monotonicity implies that the (predicted) cost shares have to be non-negative. However, since $C_q(\mathbf{w}_{qf}, L_{qf})$ is quadratic in the logarithms of the wage rates, it is not possible to impose parameter constraints that ensure monotonicity (Jorgenson, 1986). Thus, whether this property is fulfilled has to be checked after the estimation.

The other two properties, linear homogeneity and concavity, however, can be imposed. Linear homogeneity requires that the parameters α_{iq} add up to unity and the parameters β_{ijq} add up to zero:

$$\sum_i \alpha_{iq} = 1, \quad \sum_i \beta_{ijq} = 0. \quad (6)$$

Concavity of $C_q(\mathbf{w}_{qf}, L_{qf})$ implies that its Hessian matrix \mathbf{H}_q is negative semi-definite. Diewert and Wales (1987) show that this is fulfilled provided that $C(\mathbf{w}_{qf}, L_{qf}) > 0$, if and only if the matrix \mathbf{K}_q (as defined in (4)) is negative semi-definite.

However, since \mathbf{H}_q and \mathbf{K}_q depend on the data, i.e., on the cost shares S_{iqf} , a different Hessian matrix and thus a different matrix \mathbf{K}_q is obtained for each observation. As shown by Diewert and Wales (1987), imposing concavity globally destroys the flexibility of the translog function and may result in biased estimates as restrictions

⁴ For details see appendix B. The bordered Hessian matrix and its derivation are given in appendix C.

on own and cross price elasticities are imposed. Therefore, Ryan and Wales (2000) suggest to impose concavity locally at a normalization point. If the sample mean is chosen as normalization point, i.e., all wage rates are centered around their respective mean, the logarithmic wage rates are zero at this point and S_{iqf} equals α_{iq} , see (2), such that \mathbf{H}_q and \mathbf{K}_q in this point only depend on parameters.⁵ The ij th element of \mathbf{K}_q is then given by:

$$k_{ijq} = \beta_{ijq} - \alpha_{iq}\delta_{ij} + \alpha_{iq}\alpha_{jq} \quad (7)$$

with $\delta_{ij} = 1$ if $i = j$ and 0 otherwise, compare (4). Hence, in this point parameter restrictions can be imposed.

Concavity of the cost function in the neighborhood of the normalization point can be attained by setting $-\mathbf{K}_q = \mathbf{D}\mathbf{D}'_q$ (Ryan/Wales, 2000), where \mathbf{D} is a lower triangular matrix.⁶ Solving (7) for β_{ijq} and replacing $-k_{ijq}$ by the ij th element of $\mathbf{D}\mathbf{D}'_q$ ⁷ yields non linear constraints that impose concavity of $C_q(\mathbf{w}_{qf}, L_{qf})$ in the neighborhood of the normalization point (Ryan/Wales, 2000: 255):

$$\beta_{ijq} = -(DD')_{ijq} + \alpha_{iq}\delta_{ijq} - \alpha_{iq}\alpha_{jq}. \quad (8)$$

If the parameter constraint given by (8) as well as the symmetry constraint, i.e., $\beta_{ijq} = \beta_{jiq}$, are inserted into the system of share equations given by (2) and random disturbances added, (2) changes to (9).⁸ Since the parameters β_{ijq} are replaced by (8), the parameters d_{ijq} are estimated instead of β_{ijq} . It is worth noting that the number of parameters d_{ijq} is equal to the number of parameters β_{ijq} . The non-linear constraint given by (8) ensures that the parameters β_{ijq} are in a certain relationship to each other such that the Hessian matrix of $C_q(\mathbf{w}_{qf}, L_{qf})$ is negative semi-definite, at least in the reference point.

$$\begin{aligned} S_{yqf} &= \alpha_{yq} + (-d_{yyq}d_{yyq} + \alpha_{yq} - \alpha_{yq}\alpha_{yq})v_{yqf} + (-d_{ymq}d_{yyq} - \alpha_{yq}\alpha_{mq})v_{mqf} + (-d_{yoq}d_{yyq} - \alpha_{yq}\alpha_{oq})v_{oqf} + \varepsilon_{yqf} \\ S_{mqf} &= \alpha_{mq} + (-d_{ymq}d_{yyq} - \alpha_{yq}\alpha_{mq})v_{yqf} + (-d_{ymq}d_{ymq} - d_{mmq}d_{mmq} + \alpha_{mq} - \alpha_{mq}\alpha_{mq})v_{mqf} \\ &\quad + (-d_{yoq}d_{ymq} - d_{moq}d_{mmq} - \alpha_{mq}\alpha_{oq})v_{oqf} + \varepsilon_{mqf} \\ S_{oqf} &= \alpha_{oq} + (-d_{yyq}d_{yoq} - \alpha_{yq}\alpha_{oq})v_{yqf} + (-d_{yoq}d_{ymq} - d_{moq}d_{mmq} - \alpha_{mq}\alpha_{oq})v_{mqf} \\ &\quad + (-d_{yoq}d_{yoq} - d_{moq}d_{moq} - d_{ooq}d_{ooq} + \alpha_{oq} - \alpha_{oq}\alpha_{oq})v_{oqf} + \varepsilon_{oqf}. \end{aligned} \quad (9)$$

The indices y , m , and o refer to younger, middle aged, and older workers respectively, that belong to skill group q within establishment f . ε_{iqf} are unobservable random disturbances for the cost shares.

Since the cost shares S_{yqf} , S_{mqf} , and S_{oqf} sum up to unity for each observation qf , the sum of the error terms $\sum_i \varepsilon_{iqf}$ is equal to zero for all observations (Jorgenson, 1986). Hence, the error covariance matrix is singular. Greene (2012) therefore suggests to estimate just $n - 1$ share equations. Furthermore, he notes that computing maximum likelihood estimates of the parameters ensures invariance with respect to the choice of which share equation is dropped.⁹

As stressed for example by Grant and Hamermesh (1981), one only derives unbiased estimates of a cost func-

⁵ Please note, centering the wage rates does not affect the parameter estimates, except the estimates of α_{iq} .

⁶ The idea of employing the Cholesky factorization to impose concavity goes back to Lau (1978).

⁷ Here: $(\mathbf{D}\mathbf{D}')_{\mathbf{q}} = \begin{bmatrix} d_{yyq}^2 & d_{yyq}d_{myq} & d_{yyq}d_{oyq} \\ d_{yyq}d_{myq} & d_{myq}^2 + d_{mmq}^2 & d_{myq}d_{oyq} + d_{mmq}d_{omq} \\ d_{yyq}d_{oyq} & d_{myq}d_{oyq} + d_{mmq}d_{omq} & d_{oyq}^2 + d_{omq}^2 + d_{ooq}^2 \end{bmatrix}$.

⁸ The adding-up constraints $\sum_i \alpha_{iq} = 1$ and $\sum_i \beta_{ijq} = 0$ are directly imposed when estimating the system of equations.

⁹ I estimate (9) using Stata's iterative feasible generalized least squares estimator which is equivalent to ML estimation with multivariate normal disturbances and also invariant to the choice of the estimated share equations (see Stata's manual *nlsur - Estimation of nonlinear systems of equations*, page 9, <http://www.stata.com/manuals13/rnlsur.pdf>).

tion, or a system of share equations that is based on factor prices, if the latter are exogenous to the firm. Hamermesh (1996) therefore argues that the cost function would ideally be embedded in a system of equations that also describes the labor supply reaction or workers' choices. As an alternative, strictly exogenous variables might be used to generate instruments for the factor prices in the cost function. However, to find such variables is already a severe problem for studies that consider homogeneous labor. In the case of heterogeneous labor it is "insolvable" (Hamermesh, 1996: 70). Therefore, one has to decide whether to treat quantities as exogenous and to estimate a production function, or to treat wage rates as exogenous and to estimate a cost function. Hamermesh (1996) argues that in the case of microeconomic data it is reasonable to assume the latter since small units like establishments face perfect elastic labor supply in the long run. As a robustness check, I exclude all establishments with 50 employees or more to address that large establishments may have local monopsony power to set wages.

5 Data

To estimate the parameters of (9) that are needed to calculate c_{ij} and θ_{ij} , I use detailed information on a 20 percent random sample of all establishments located in Germany that employed at least one worker subject to social security contributions on June 30, 2012. The public sector, temporary employment agencies, and crop and animal production, hunting and related service activities are excluded. The considered number of establishments amounts to 508,514. For these establishments my raw data set contains all spells of employment that were reported by the establishments in 2012, i.e., 9,124,309 individual employment spells. The information on the spells of employment are taken from the IAB employment history (BeH).

The BeH is based on the integrated notification procedure for health, pension, and unemployment insurance. Employers are required to submit notifications to the responsible social security agencies concerning all of their employees covered by social security as well as marginally employed at least once a year. The Federal Employment Agency collects and edits these notifications, which the IAB subsequently integrates into the BeH. Thus, the BeH contains very reliable information on all spells of employment as long as the workers are not exempt from social security contributions, like civil servants and self-employed persons (about 12 percent of total employment in Germany).

Among other things, each observation in the BeH provides information on individual and job characteristics as well as on the length of the specific spell of employment and the paid gross wage sum during this time period. Due to the social security threshold, the wage information is right-censored and in 2012 only reported up to € 67,200 per year in West Germany and € 57,600 per year in East Germany. To avoid biased estimates, I partly impute the wages of employees subject to social security contributions with estimated wages.¹⁰

Aggregating the information provided by the 9 million individual spells of employment from the BeH, a unique establishment level data set is generated.¹¹ The units of observation are skill groups within the considered establishments. As I am interested in composition effects of the skill groups at the establishment level with regard to the age of the workers, I exclude very small skill groups, which I define as groups with less than five members, i.e., $\sum_i L_{iqf} < 5$. The final number of observations that are used in the empirical analysis amounts to

¹⁰ I follow Reichelt (2015) and apply an interval regression to estimate the (daily) wages above the thresholds, see appendix D.

¹¹ Six percent of the employment spells are not considered. I exclude apprentices from the analysis. Furthermore, employment spells of interns, and student trainees are only considered as long as the employees have at least a bachelors degree and have an occupation that requires a vocational training of at least two years or higher education.

290,334 skill groups within 208,868 establishments.¹²

To distinguish the employees of an establishment according to their skill level, i.e., low skilled, medium skilled, and high skilled, the requirement level of the individual job is used.¹³ I define workers with an occupation that does not require a vocational training of at least two years as low skilled. Medium skilled workers are those with an occupation that requires a vocational training of at least two years, and workers with an occupation that requires at least a bachelor's degree, a qualification as master craftsman or technician, or an equivalent technical school or college graduation are considered as high skilled.

I use the requirement level of the job instead of the highest educational degree of a worker which is also provided in the BeH in order to identify low skilled, medium skilled, and high skilled labor for two reasons. First, the requirement level enables assignment of workers to the different skill groups according to the complexity of their actual task and not according to their formal level of qualification. Haas et al. (2013) show that workers with the same formal educational degree, but with different job requirements, are imperfect substitutes. Second, 27 percent of the nine million spells of employment do not provide information on the educational degree of the worker, whereas the information on the requirement level is missing only in two percent of the employment spells.¹⁴

For each skill group within an establishment, I compute its individual composition with regard to the age of the workers. To be more specific, I compute the establishment-specific cost share of age group i within skill group q and establishment f , S_{iqf} . Unfortunately, the BeH does not provide information on non-wage labor costs. Therefore, the computed cost shares refer to wage costs only. Summary statistics are provided by table 1. In addition to the cost shares, the wages of the different age-skill groups, w_{iqf} , that enter (9) on the right hand side are calculated. w_{iqf} indicates how much establishment f paid in 2012 on average per day to an employee with age i and skill level q . For the calculation of w_{iqf} only spells employment are used that refer to full-time employment as more detailed information on working hours is not available. Moreover, when calculating w_{iqf} only spells of employment with a length of at least eight days and a gross-daily wage above the limit of marginally employed are considered. In 40 percent of the cases it is not possible to calculate w_{iqf} since establishment f does not employ workers with age i and skill level q whose spells of employment fulfill the criteria above. In these cases $w_{iqr(f)}$ is used as an approximation of w_{iqf} , where $w_{iqr(f)}$ is the average gross daily wage of group iq in region $r(f)$, i.e., the local labor market as defined by Kosfeld and Werner (2012) in which establishment f is located. $w_{iqr(f)}$ is calculated based on information from the Employment Statistics of the Federal Employment Agency. It can be interpreted as the daily wage that establishment f would have to pay to employ one unit of type iq labor in the local labor market where it is located.

The age group to which an employee belongs is defined by his or her day of birth that is also given in the BeH. In accordance with Mahlberg et al. (2013), employees younger than 30 years are considered as *younger workers*, workers aged 30 to 49 years as *middle aged workers* and workers that are 50 years old or older as *older workers*.¹⁵

¹² Without excluding skill groups with less than five members, the number of observations would be 885,377.

¹³ The requirement level of a job is the second dimension of the German Classification of Occupations 2010 (KldB 2010). A detailed description can be found in Paulus and Matthes (2013).

¹⁴ If the requirement level is missing but the highest educational degree is not, then the latter is used to assign the workers to one of the skill groups. This reduces the share of employment spells that cannot be assigned to one of the three groups to 0.07%. In 65.8% of the cases in which information on the requirement level as well as on the highest educational degree is available the requirement level accords with the reported highest educational degree. In 16.6% the reported highest educational degree is higher and in 17.6% lower than the requirement level.

¹⁵ 0.4 percent of the nine million spells of employment are not assigned to one of the age groups since the worker is implausible old (above 75 years).

Table 1: Summary statistics

	Mean	SD	Min	Max
Groups of low skilled labor (N = 64405)				
Cost shares				
Younger labor	0.215	0.244	0.000	1.000
Middle aged labor	0.459	0.259	0.000	1.000
Older labor	0.326	0.266	0.000	1.000
ln(Average gross daily wages)				
Younger labor	3.976	0.227	2.558	5.621
Middle aged labor	4.155	0.295	2.565	6.268
Older labor	4.229	0.273	2.570	6.127
Groups of medium skilled labor (N = 173313)				
Cost shares				
Younger labor	0.205	0.213	0.000	1.000
Middle aged labor	0.501	0.232	0.000	1.000
Older labor	0.294	0.238	0.000	1.000
ln(Average gross daily wages)				
Younger labor	4.203	0.273	2.560	6.710
Middle aged labor	4.336	0.365	2.558	6.214
Older labor	4.424	0.351	2.559	6.460
Groups of high skilled labor (N = 52617)				
Cost shares				
Younger labor	0.121	0.178	0.000	1.000
Middle aged labor	0.561	0.237	0.000	1.000
Older labor	0.318	0.252	0.000	1.000
ln(Average gross daily wages)				
Younger labor	4.506	0.286	2.570	6.368
Middle aged labor	4.823	0.417	2.568	7.441
Older labor	4.960	0.408	2.567	6.959

6 Results

I start with a regression where I pool the observations that belong to different skill groups. Thus, by now I assume that the parameters and elasticities do not differ across skill levels. The first and the second column of table 2 contain the estimated parameters of (9), i.e., the system of share equations including the parameter constraints ensuring that the cost function is concave. As the adding-up constraints, see (6), are imposed when estimating the model, not all parameters are explicitly reported.¹⁶ Inserting the reported parameters into (8) and (6), i.e., the concavity and the adding-up constraints, the parameters β_{ij} are obtained. They are reported in the third column of table 2. Based on these parameters, the Hicks partial elasticities of complementarity, c_{ij} , and factor price elasticities, θ_{ij} and θ_{ji} , are calculated using (3) and (5), respectively. As θ_{ij} depends on the cost share of age group j and θ_{ji} on the cost share of age group i , θ_{ij} is not symmetric, i.e., $\theta_{ij} = \theta_{ji}$ only holds if $S_{iq} = S_{jq}$.

Since I center the wage rate of each age-skill group around their mean, the results for α_i indicate that, at the reference point where the wage rates correspond to their respective mean, the cost share of younger workers amounts to about 20 percent, the cost share of middle aged workers to about 50 percent, and, hence, the cost share of older workers to about 30 percent. Furthermore, the results in column three indicate that the cost share of each age group declines, given an increase in the wage rate of the respective age group. This shows that firms react on increasing relative wages by substituting one age group by another. However, there is also a

¹⁶Please note, the adding-up constraints given by (6) require that $\sum_i \beta_{ij}$ is zero. The corresponding constraints with regard to d_{ij} are obtained by combining the adding-up constraints with the constraints that are imposed to ensure that the cost function is concave which is given by (8).

Table 2: Estimated parameters and calculated elasticities,
average effects assuming $\alpha_{iq} = \alpha_i$, $\beta_{ijq} = \beta_{ij}$

i	j	α_i	d_{ij}	β_{ij}	c_{ij}	θ_{ij}	θ_{ji}
y	y	0.194*** (0.000)	0.455*** (0.002)	-0.050*** (0.001)	-3.273*** (0.022)	-0.628*** (0.004)	-0.628*** (0.004)
m	m	0.503*** (0.000)	0.528*** (0.001)	-0.066*** (0.001)	-0.801*** (0.004)	-0.403*** (0.002)	-0.403*** (0.002)
o	o			-0.137*** (0.001)	-1.394*** (0.006)	-0.426*** (0.002)	-0.426*** (0.002)
y	m		-0.192*** (0.002)	-0.010*** (0.001)	0.999*** (0.008)	0.502*** (0.004)	0.192*** (0.002)
y	o			0.061*** (0.001)	0.413*** (0.008)	0.126*** (0.002)	0.079*** (0.002)
m	o			0.076*** (0.001)	0.690*** (0.004)	0.211*** (0.001)	0.347*** (0.002)
N		290335					

Standard errors in parentheses, calculation based on Delta-method, significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The data is normalized at the sample mean. Monotonicity and concavity of the cost function are satisfied at all data points. The concavity constraint is locally imposed at the sample mean. The elasticities are calculated at mean predicted shares.

negative relationship between the cost share of younger workers and the wage rate of middle aged workers (and vice versa due to the symmetry constraint). This negative cross effect points to the imperfect substitutability of younger and middle aged workers. However, positive cross price elasticities, per se, do not necessarily imply that those groups of workers are perfect substitutes.¹⁷

The results in column four to six indicate how the productivity of age groups change in response to changes in the amount of age specific labor input. As expected, the marginal productivity of an age group declines if its size increases. The effect is largest for younger workers which is in line with findings by Welch (1979). He observes that the negative effect of cohort size on wages of the same cohort declines over the career.

The cross effects indicate that all three age groups are complementary input factors, i.e., all age groups benefit from working together with differently aged workers. But the degree of complementarity differs. This points out that flexible functional forms, like a translog function, are needed to quantify the degree of complementarity between different age groups. A CES framework, for example, is too restrictive as the elasticities of substitution between any two pairs of age groups are assumed to be the same.

My results show that the highest degree of complementarity is observed for younger and middle aged workers, the lowest for younger and older workers. With regard to younger workers this indicates that they especially benefit of working together with middle aged workers. One reasonable explanation is that younger workers benefit from the working experience of middle aged workers. At the same time it might be beneficial that the knowledge base and the abilities of younger and middle aged workers are not too different which might enhance cooperation. An alternative interpretation of the cross effects is that it is more difficult to substitute middle aged workers by either younger or older workers, than younger by older workers (and vice versa) as middle aged workers have a combination of cognitive abilities of younger workers and experience of older workers. This combination make them especially valuable, what in turn decreases substitutability.

I use the calculated own and cross factor price elasticities, θ_{ij} , to simulate how the demographic change might affect the productivity of the three age groups due to the expected shift in relative labor supply (compare Grant/-

¹⁷ It follows from the adding-up constraint given by (6) that the parameters β_{ym} as well as β_{mo} need to be positive if β_{ym} and β_{ii} are negative.

Hamermesh, 1981). It is assumed that the shift in relative labor supply reflects in a corresponding shift in the age structure of (skill-specific) employment. Furthermore, the model calculation is based on the assumption that the productivity of compound skill-specific labor in terms of the value of its marginal product, $P \frac{\partial Y}{\partial L} \frac{\partial L}{\partial L_q}$ where P denotes the price of one unit of output Y , remains constant for all three skill groups.¹⁸ In the model calculation, I examine a scenario where the number of younger workers decreases by 21 percent, and the number of middle aged workers by 16.5 percent. The number of older workers is assumed to increase by 3.9 percent.¹⁹

The results of the model calculation are summarized in table 3. They suggest that the productivity of younger workers will increase most (by five percent) due to the assumed changes in the age structure of the workforce. The own factor price elasticity of younger workers (θ_{yy}) would predict an productivity increase of the wage rate by more than 13 percent in response to the assumed decline of younger labor by 21 percent. Further upward pressure on the wage of younger workers comes from the increasing number of older workers as it lets the relative amount of young workers decline. In contrast, the decreasing number of middle aged workers puts downward pressure on the wages of younger workers due to the complementary relationship. In total, the productivity of younger workers is therefore expected to rise ‘only’ by about five percent.

Similarly, the decline in the number of middle aged workers and the increase in the number of older workers let the productivity of middle aged workers increase. The decreasing number of younger workers, however, puts downward pressure on the wage of middle aged workers. The total effect is expected to be a productivity increase of about three percent. For older workers the model calculation predicts a strong decrease in productivity, i.e., a decline by nine percent. Predominantly this is caused by the decreasing number of younger and middle aged workers, wherefore the relative amount of older workers increases. The ceteris paribus effect of the increasing number of older workers suggests a reduction in productivity by less than two percent.

Table 3: Results of a model calculation

Expected percentage change in productivity of ...		
younger workers	middle aged workers	older workers
5.4 (0.06)	3.4 (0.02)	-9.0 (0.04)

Effects based on factor price elasticities (θ_{ij}, θ_{ji} in table 2) and the scenario: number of younger workers -21%, middle aged workers -16.5%, older workers: +3.9%. Standard errors in parentheses, calculation based on Delta-method.

By relaxing the assumption that the elasticities do not vary across skill levels, heterogeneous effects across skill levels become visible, see table 4. Own price elasticities are highest in absolute terms within groups of high-skilled labor, especially for younger high-skilled labor. It indicates that a relatively small decline in the number of younger high-skilled workers leads to a relatively large productivity increase of the remaining younger workers.

The cross effects are heterogeneous across skill groups as well. The degree of complementarity between younger and middle aged workers is highest within groups of high-skilled labor. The same result is obtained for the degree of complementarity between middle aged and older workers. This supports the findings of Welch (1979) who observes that the degree of complementarity between cohorts increases by skill level.

¹⁸ This assumption implies that the marginal costs of compound skill-specific labor are assumed to remain unchanged since at the cost minimum the marginal value product equals marginal costs (compare appendix B: for the model calculation $d\phi$ is assumed to be zero).

¹⁹ This scenario is based on the relative change of the population aged 15-29 years, 30-49 years, and 50-70 years that is expected by the German Federal Statistical Office until 2030 (DESTATIS, 2009). Changes in labor participation rates are not considered.

Surprisingly and in contrast to the expectation that especially within groups of high skilled labor cognitive abilities of younger workers and experience of older workers complement each other, I find that the degree of complementarity between younger and older workers decreases in the skill level. The presence of younger workers seems to have a higher importance for the productivity of older workers (and vice versa) in groups of low skilled workers than in other skill groups. One reasonable explanation for the high degree of complementarity between younger and older workers with jobs that require only little specific training is that some of their tasks are often connected with physical burden. Hence, older low skilled workers might benefit from working together with younger workers as tasks could be shared according to the individual physical capacity. In addition, especially within groups of low skilled workers with tasks that do not require specific training, general labor market experience might still enhance productivity. Therefore, younger low skilled workers might benefit from older low skilled workers who presumably have at least some general working experience.

The high degree of substitutability between younger and older high-skilled labor supports the view of Fitzenberger and Kohn (2006). They argue that university graduates have in general (independent of their age) high competence to solve problems which reduces complementarities. Nevertheless, it remains surprising that, within groups of high skilled labor, younger and older workers are rather substitutes than complements, but younger and middle aged workers as well as middle aged and older workers complement each other. But the finding is robust.

As stressed earlier, my estimates can only be interpreted as causal effects if wage rates are exogenous to the firms. As a robustness check, I repeat the estimations with a reduced sample that only covers rather small establishments, defined as establishments with less than 50 employees. The additional results are summarized in table E.1 and table E.2 in appendix E. Qualitatively the results are similar to those discussed in this section. Quantitatively, the parameter estimates and calculated elasticities that are based on the reduced sample are ten percent to 20 percent smaller than those that are based on the full sample. As a further robustness check, I restrict my sample to skill groups with, on the establishment level, at least ten members. Again I observe qualitatively similar results, even though the number of observations reduces significantly.²⁰

²⁰ The additional results are available from the author upon request.

Table 4: Elasticities by skill level

i	j	low skilled labor			medium skilled labor			high skilled labor		
		c_{ij}	θ_{ij}	θ_{ji}	c_{ij}	θ_{ij}	θ_{ji}	c_{ij}	θ_{ij}	θ_{ji}
y	y	-2.936*** (0.053)	-0.630*** (0.011)	-0.630*** (0.011)	-2.257*** (0.014)	-0.463*** (0.003)	-0.463*** (0.003)	-7.870*** (0.183)	-0.954*** (0.022)	-0.954*** (0.022)
m	m	-0.818*** (0.010)	-0.376*** (0.005)	-0.376*** (0.005)	-0.707*** (0.004)	-0.354*** (0.002)	-0.354*** (0.002)	-1.042*** (0.019)	-0.585*** (0.010)	-0.585*** (0.010)
o	o	-1.242*** (0.014)	-0.405*** (0.005)	-0.405*** (0.005)	-1.324*** (0.006)	-0.390*** (0.002)	-0.390*** (0.002)	-1.847*** (0.021)	-0.587*** (0.007)	-0.587*** (0.007)
y	m	0.890*** (0.019)	0.409*** (0.009)	0.191*** (0.004)	0.768*** (0.006)	0.385*** (0.003)	0.157*** (0.001)	1.892*** (0.053)	1.061*** (0.030)	0.229*** (0.006)
y	o	0.678*** (0.022)	0.221*** (0.007)	0.146*** (0.005)	0.265*** (0.006)	0.078*** (0.002)	0.054*** (0.001)	-0.337*** (0.043)	-0.107*** (0.014)	-0.041*** (0.005)
m	o	0.566*** (0.009)	0.185*** (0.003)	0.260*** (0.004)	0.669*** (0.004)	0.197*** (0.001)	0.335*** (0.002)	1.119*** (0.018)	0.355*** (0.006)	0.628*** (0.010)
N		64405			17313			52617		

Standard errors in parentheses, calculation based on Delta-method, significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The data is normalized at the respective sample mean. Monotonicity and concavity of the cost function are satisfied at more than 99 percent of the respective data points. The concavity constraint is locally imposed at the respective sample mean. The elasticities are calculated at mean predicted shares.

7 Conclusion

In this study, I analyze complementarities between differently aged workers. This is of special interest since demographic change leads to a shift in the workforce age structure in Germany as well as in many other European countries which likely affects (relative) productivity of workers. Taking into account the interaction of skill and age, I analyze age complementarities within groups of workers with a comparable skill level. Contrary to previous papers, my analysis is based on a linked employer-employee data set. This allows the investigation of age complementarities on establishment level, i.e., directly at the workplace. Studies operating at the aggregate level might suffer from ecological fallacy because their results are interpreted as if they pertain to behavior of firms. Moreover, in contrast to most other studies, I provide empirical evidence on the way the degree of complementarity varies between age groups. Often it is assumed that the substitutability of different age groups is constant, meaning that it does not depend on the age groups that are considered. My findings suggest that this assumption does not hold.

The results of this study indicate that workers who belong to different age groups are complements and that the degree of complementarity differs across the considered age groups and by skill level. Even though, for each skill level the largest degree of complementarity is identified between younger and middle aged workers which especially applies to the group of high-skilled labor. In contrast, younger and older high-skilled workers seem to be substitutes rather than complements. This surprises, as Arntz and Gregory (2014) find that cognitive abilities of younger workers and experience of older workers complement each other in generating innovations, which typically involves high-skilled labor.

Based on my results, I simulate how the expected demographic change impacts on the productivity of the different age groups. Due to the complementary relationship between older workers and the two other groups, the expected decline in younger and middle aged workers puts pressure on the wages of older workers. Under the assumption that wages are completely flexible and that the overall productivity of compound labor remains constant, one could for this age group expect a decline in wages of about nine percent until 2030. On the contrary, my results show that the wage of younger and middle aged workers is expected to rise by about five percent and three percent, respectively, since their relative shares decline.

Appendix

A Deriving the system of share equations

The partial derivative of $\ln(C_q(\mathbf{w}_{qf}, L_{qf}))$, as given by (1), with respect to $\ln(w_{iqf})$ corresponds to

$$\frac{\partial \ln(C_q(\mathbf{w}_{qf}, L_{qf}))}{\partial \ln(w_{iqf})} = \alpha_{iq} + \sum_j \beta_{ijq} \ln(w_{jqf}). \quad (10)$$

$\frac{\partial \ln(C_q(\mathbf{w}_{qf}, L_{qf}))}{\partial \ln(w_{iqf})}$ is the elasticity of $C_q(\mathbf{w}_{qf}, L_{qf})$ with regard to w_{iqf} . Thus:

$$\frac{\partial \ln(C_q(\mathbf{w}_{qf}, L_{qf}))}{\partial \ln(w_{iqf})} = \frac{\frac{dC_q(\mathbf{w}_{qf}, L_{qf})}{C_q(\mathbf{w}_{qf}, L_{qf})}}{\frac{dw_{iqf}}{w_{iqf}}} = \frac{dC_q(\mathbf{w}_{qf}, L_{qf})}{dw_{iqf}} \frac{w_{iqf}}{C_q(\mathbf{w}_{qf}, L_{qf})}. \quad (11)$$

According to Shephard's Lemma, see appendix C: $\frac{dC_q(\mathbf{w}_{qf}, L_{qf})}{dw_{iqf}} = L_{iqf}^*$, where L_{iqf}^* is the cost minimizing number of type iq workers given the price vector \mathbf{w}_{qf} and L_{qf} . Hence, the cost minimizing share S_{iqf} of type iq workers in type q workers amounts to:

$$S_{iqf}(\mathbf{w}_{qf}) = \frac{L_{iqf}^* w_{iq}}{\sum_j L_{jqf}^* w_{jqf}} = \alpha_{iq} + \sum_j \beta_{ijq} \ln(w_{jqf}). \quad (12)$$

B Deriving the Hicks partial elasticity of complementarity and the factor price elasticity

The derivation of the Hicks partial elasticity of complementarity and of factor price elasticity are provided by, e.g., Sato and Koizumi (1973). As the elasticities are of central importance for my empirical analysis, I sketch it here in more detail. For simplicity, I omit the indices q and f .

Sato and Koizumi (1973) start with Shephard's Lemma. Accordingly, the partial derivative of a cost function with respect to factor price of input i corresponds to the conditional factor demand of i :

$$\underbrace{\frac{\partial C(\mathbf{w}, L)}{\partial w_i}}_{C_i} = L_i(\mathbf{w}, L). \quad (13)$$

Assuming the cost function is linear homogeneous in L , the corresponding total differential is given by, compare appendix C:

$$C_i \frac{dL}{L} + \sum_j C_{ij} dw_j = dL_i. \quad (14)$$

This equation can be transferred into a system of total differentials and expressed in matrix form:

$$\underbrace{\begin{bmatrix} 0 & C_1 & \dots & C_n \\ C_1 & C_{11} & \dots & C_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ C_n & C_{n1} & \dots & C_{nn} \end{bmatrix}}_{\tilde{\mathbf{H}}} \underbrace{\begin{bmatrix} \frac{dL}{L} \\ dw_1 \\ \vdots \\ dw_n \end{bmatrix}}_{d\mathbf{w}} = \underbrace{\begin{bmatrix} Ld\phi \\ dL_1 \\ \vdots \\ dL_n \end{bmatrix}}_{d\mathbf{L}}$$

where $\tilde{\mathbf{H}}$ denotes the bordered Hessian matrix of the cost function, see appendix C The first row refers to the change of minimum total costs due to changes in factor prices. ϕ denotes the Lagrange multiplier of cost minimization which in equilibrium equals marginal costs (Sato/Koizumi, 1973).

It follows that $d\mathbf{w} = \tilde{\mathbf{H}}^{-1} d\mathbf{L}$. Hence, the partial derivative of w_i with respect to L_j is given by:

$$\frac{\partial w_i}{\partial L_j} = (\tilde{\mathbf{H}}^{-1})_{ij} \equiv \frac{|\mathbf{H}^{ij}|}{|\tilde{\mathbf{H}}|}, \quad (15)$$

where $|\mathbf{H}^{ij}|$ denotes the cofactor of C_{ij} and $|\tilde{\mathbf{H}}|$ the determinant of $\tilde{\mathbf{H}}$.

Therefore, the factor price elasticity is given by

$$\frac{\partial \ln(w_i)}{\partial \ln(L_j)} = \frac{\partial w_i}{\partial L_j} \frac{L_j}{w_i} = \frac{|\mathbf{H}^{ij}|}{|\tilde{\mathbf{H}}|} \frac{L_j}{w_i}. \quad (16)$$

By multiplying the right hand side by $\frac{w_j C}{w_j C}$, the factor price elasticity can be expressed as a function of j 's cost share, S_j , and the Hicks partial elasticity of complementarity, c_{ij} :

$$\frac{\partial \ln(w_i)}{\partial \ln(L_j)} = \underbrace{\frac{L_j w_j}{C}}_{S_j} \underbrace{\frac{C}{w_i w_j} \frac{|\mathbf{H}^{ij}|}{|\tilde{\mathbf{H}}|}}_{c_{ij}}. \quad (17)$$

As I consider a translog specification of a cost function, the Hicks partial elasticity of complementarity simplifies to:

$$c_{ij} = \frac{|\mathbf{K}^{ij}|}{|\tilde{\mathbf{K}}|}, \quad (18)$$

where $\tilde{\mathbf{K}}$ is a transformation of $\tilde{\mathbf{H}}$:

$$\tilde{\mathbf{K}} = \begin{bmatrix} 0 & K_1 & \dots & K_n \\ K_1 & K_{11} & \dots & K_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ K_n & K_{n1} & \dots & K_{nn} \end{bmatrix}$$

with, compare appendix C,

$$\begin{aligned} K_i &= C_i \frac{w_i}{C(\mathbf{w}, L)} \\ &= S_i \end{aligned} \quad (19)$$

$$\begin{aligned} K_{ii} &= C_{ii} \frac{w_i^2}{C(\mathbf{w}, L)} \\ &= S_i^2 - S_i + \beta_{ii} \end{aligned} \quad (20)$$

$$\begin{aligned} K_{ij} &= C_{ij} \frac{w_i w_j}{C(\mathbf{w}, L)} \\ &= S_i S_j + \beta_{ij}. \end{aligned} \quad (21)$$

That in the case of a translog cost function it holds

$$\frac{|\mathbf{K}^{ij}|}{|\tilde{\mathbf{K}}|} = \frac{C}{w_i w_j} \frac{|\mathbf{H}^{ij}|}{|\tilde{\mathbf{H}}|}, \quad (22)$$

can be seen if $\tilde{\mathbf{K}}$ is expressed in matrix notation:

$$\tilde{\mathbf{K}} = \frac{\widehat{\mathbf{W}} \tilde{\mathbf{H}} \widehat{\mathbf{W}}}{C(\mathbf{w}, L)},$$

where $\widehat{\mathbf{W}}$ is a diagonal matrix:

$$\widehat{\mathbf{W}} \equiv \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & w_1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & w_n \end{bmatrix}.$$

Hence:

$$\tilde{\mathbf{K}}^{-1} = C(\mathbf{w}, L) \widehat{\mathbf{W}}^{-1} \tilde{\mathbf{H}}^{-1} \widehat{\mathbf{W}}^{-1} \quad (23)$$

and

$$\frac{|\mathbf{K}^{ij}|}{|\tilde{\mathbf{K}}|} \equiv (\mathbf{K}^{-1})_{ij} = C(\mathbf{w}, L) \frac{1}{w_i} (\mathbf{H}^{-1})_{ij} \frac{1}{w_j} \equiv \frac{C}{w_i w_j} \frac{|\mathbf{H}^{ij}|}{|\tilde{\mathbf{H}}|}. \quad (24)$$

C Deriving the bordered Hessian matrix $\tilde{\mathbf{H}}$ of the cost function $C(\mathbf{w}, L)$

In the following, I derive the (bordered) Hessian matrix of the translog cost function given by equation (1). For simplicity, I omit the indices q and f . The result can also be found in, e.g., Hamermesh and Grant (538ff. 1979).

$$\tilde{\mathbf{H}} = \left[\begin{array}{c|ccc} 0 & C_1 & \dots & C_n \\ \hline C_1 & & & \\ \vdots & & \mathbf{H} & \\ C_n & & & \end{array} \right]$$

$$\mathbf{H} = \begin{bmatrix} C_{11} & \dots & C_{1n} \\ \vdots & \ddots & \vdots \\ C_{n1} & \dots & C_{nn} \end{bmatrix}$$

$$\begin{aligned} C(\mathbf{w}, L) &= \exp \left[a_0 + \sum_i \alpha_i \ln(w_i) + \frac{1}{2} \sum_i \sum_j \beta_{ij} \ln(w_i) \ln(w_j) + l \right] \\ C_i = \frac{\partial C(\mathbf{w}, L)}{\partial w_i} &= C(\mathbf{w}, L) \left(\alpha_i \frac{1}{w_i} + \sum_j \beta_{ij} \frac{1}{w_i} \ln(w_j) \right), \text{ with } \beta_{ij} = \beta_{ji} \\ &= \frac{C(\mathbf{w}, L)}{w_i} \left[\alpha_i + \sum_j \beta_{ij} \ln(w_j) \right] \\ &= \frac{C(\mathbf{w}, L)}{w_i} S_i (= L_i^*, \text{ Shephard's Lemma}) \end{aligned} \quad (25)$$

$$\begin{aligned} C_{ii} = \frac{\partial^2 C(\mathbf{w}, L)}{\partial w_i \partial w_i} &= C(\mathbf{w}, L) \left(\alpha_i \frac{1}{w_i} + \sum_j \beta_{ij} \frac{1}{w_i} \ln(w_j) \right) \left(\alpha_i \frac{1}{w_i} + \sum_j \beta_{ij} \frac{1}{w_i} \ln(w_j) \right) \\ &\quad + C(\mathbf{w}, L) \left(-\alpha_i \frac{1}{w_i} \frac{1}{w_i} + \left(-\sum_j \beta_{ij} \frac{1}{w_i} \frac{1}{w_i} \ln(w_j) + \beta_{ii} \frac{1}{w_i} \frac{1}{w_i} \right) \right) \\ &= C(\mathbf{w}, L) \left[\frac{1}{w_i} \left(\alpha_i + \sum_j \beta_{ij} \ln(w_j) \right) \frac{1}{w_i} \left(\alpha_i + \sum_j \beta_{ij} \ln(w_j) \right) + \frac{1}{w_i^2} \left(-\alpha_i - \sum_j \beta_{ij} \ln(w_j) + \beta_{ii} \right) \right] \\ &= \frac{C(\mathbf{w}, L)}{w_i^2} [S_i^2 - S_i + \beta_{ii}] \end{aligned} \quad (26)$$

$$\begin{aligned} C_{ij} = \frac{\partial^2 C(\mathbf{w}, L)}{\partial w_i \partial w_j} &= C(\mathbf{w}, L) \left(\alpha_i \frac{1}{w_i} + \sum_j \beta_{ij} \frac{1}{w_i} \ln(w_j) \right) \left(\alpha_j \frac{1}{w_j} + \sum_i \beta_{ij} \frac{1}{w_j} \ln(w_i) \right) + C(\mathbf{w}, L) \beta_{ij} \frac{1}{w_i w_j}, \quad i \neq j \\ &= C(\mathbf{w}, L) \frac{1}{w_i} \left(\alpha_i + \sum_j \beta_{ij} \ln(w_j) \right) \frac{1}{w_j} \left(\alpha_j + \sum_i \beta_{ij} \ln(w_i) \right) + C(\mathbf{w}, L) \beta_{ij} \frac{1}{w_i w_j} \\ &= \frac{C(\mathbf{w}, L)}{w_i w_j} \left[\left(\alpha_i + \sum_j \beta_{ij} \ln(w_j) \right) \left(\alpha_j + \sum_i \beta_{ij} \ln(w_i) \right) + \beta_{ij} \right] \\ &= \frac{C(\mathbf{w}, L)}{w_i w_j} [S_i S_j + \beta_{ij}] \end{aligned} \quad (27)$$

D Imputation of censored wages

Since the wages in the BeH are only reported up to the social security threshold (2012: €67,200 per year in West Germany, €57,600 in East Germany), I have to estimate those wages which are above the censoring limit. I use additional information from the BeH, i.e., the age, sex, and nationality of the workers as well as requirement level of the job, the employment status, the establishment size, the industry (2-digit level), the occupation (3-digit level), and the type of region the employees are working in.

Following Reichelt (2015), I apply interval regression to estimate the wages above the threshold and add an error term to the estimates, i.e., a random value from a normal distribution. As the wage information is right censored, the interval regression is equivalent to tobit regression. Since the estimated wage has to be larger than the censoring limit, the random variable is drawn from a truncated distribution as suggested by Gartner (2005). Wages of certain groups like marginal employed, interns, and student trainees should not be affected by the social security threshold. They are therefore excluded from the interval regression, whose results are summarized in table D.1. Figure D.1 shows the wage distribution before and after imputation.

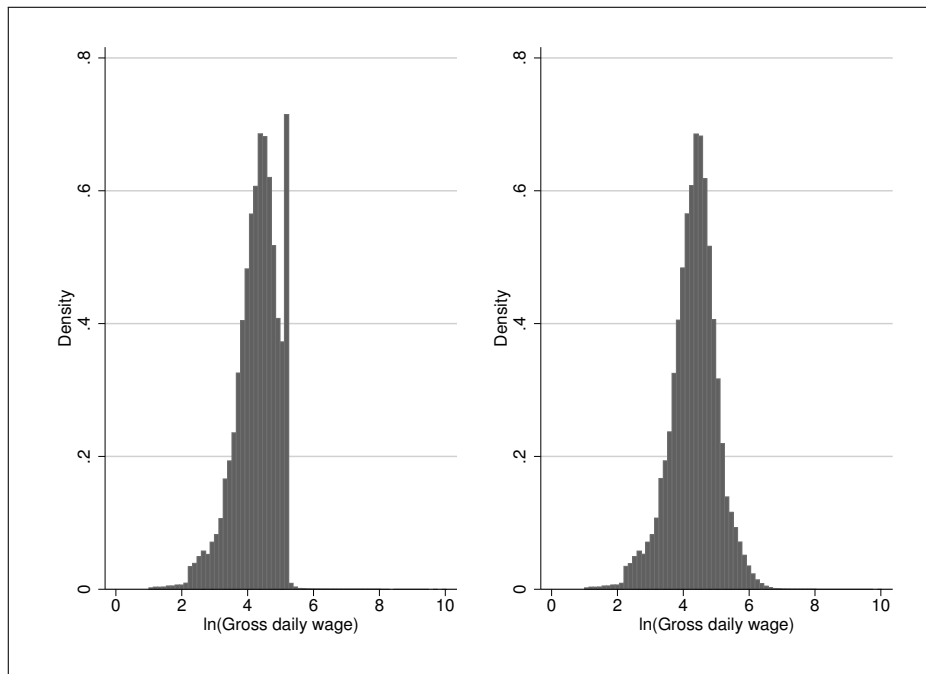


Figure D.1: Wage distribution before (left figure) and after (right figure) imputation

Table D.1: Results of an interval regression to partly impute the censored wages of employees liable to social security

Dependent variable: ln(Censored gross daily wage)		
Age	0.040***	(311.32)
Age ²	-0.000***	(-254.03)
Female	-0.189***	(-345.38)
German	0.074***	(96.50)
Requirement level according to KldB 2010 (see Paulus/Matthes, 2013), reference: Unskilled or semi-skilled activities		
Specialist activities	0.204***	(283.20)
Complex specialist activities	0.408***	(409.92)
Highly complex activities	0.609***	(553.25)
Parttime	-0.527***	(-967.73)
Employment status, reference: Employees liable to social security without special characteristics		
Employees in partial retirement	0.118***	(73.02)
Causal worker	0.466***	(94.22)
Employed pensioner (subject to social security contributions)	-0.349***	(-104.63)
Seaman	0.313***	(28.86)
Seaman in partial retirement	0.706***	(12.59)
Establishment size: ln(Number of employees)	0.076***	(609.69)
Type of region, reference: Agglomerations in West Germany		
Urbanized region in West Germany	-0.038***	(-75.40)
Rural region in West Germany	-0.049***	(-60.32)
Agglomeration in East Germany	-0.127***	(-166.33)
Urbanized region in East Germany	-0.211***	(-240.32)
Rural region in East Germany	-0.227***	(-202.68)
Constant	2.423***	(74.32)
<hr/>		
Total number of observations	6,173,784	
Censored observations	525,152	
Uncensored observations	5,648,632	
<hr/>		
Log-likelihood full model	-4.523e+06	
Log-likelihood intercept only	-6.870e+06	

t-statistics in parentheses, significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Beside the explanatory variables reported in this table industry fixed effects (2-digit level: 82 industries) and occupation fixed effects (3-digit level: 144 occupations) are considered. Marginal employed, interns, and student trainees are not considered as their wage should not be affected by the social security threshold.

E Further results

Table E.1: Results based on a reduced sample: only establishments with less than 50 employees, average effect assuming $\alpha_{iq} = \alpha_i$, $\beta_{ijq} = \beta_{ij}$

i	j	α_i	d_{ij}	β_{ij}	c_{ij}	θ_{ij}	θ_{ji}
y	y	0.205*** (0.000)	0.505*** (0.002)	-0.092*** (0.002)	-2.599*** (0.042)	-0.524*** (0.016)	-0.524*** (0.003)
m	m	0.502*** (0.001)	0.543*** (0.001)	-0.093*** (0.002)	-0.730*** (0.007)	-0.367*** (0.002)	-0.367*** (0.002)
o	o			-0.169*** (0.002)	-1.339*** (0.006)	-0.396*** (0.002)	-0.396*** (0.002)
y	m		-0.220*** (0.002)	0.008*** (0.001)	0.854*** (0.007)	0.429*** (0.008)	0.172*** (0.001)
y	o			0.084*** (0.001)	0.322*** (0.007)	0.095*** (0.002)	0.065*** (0.001)
m	o			0.085*** (0.001)	0.658*** (0.004)	0.194*** (0.001)	0.331*** (0.002)
N		226290					

Standard errors in parentheses, calculation based on Delta-method, significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The data is normalized at the sample mean. Monotonicity and concavity of the cost function are satisfied at all data points. The concavity constraint is locally imposed at the sample mean. The elasticities are calculated at mean predicted shares.

Table E.2: Elasticities by skill level based on reduced sample: only establishments with less than 50 employees

i	j	low skilled labor			medium skilled labor			high skilled labor		
		c_{ij}	θ_{ij}	θ_{ji}	c_{ij}	θ_{ij}	θ_{ji}	c_{ij}	θ_{ij}	θ_{ji}
y	y	-2.320*** (0.044)	-0.524*** (0.010)	-0.524*** (0.010)	-2.081*** (0.013)	-0.434*** (0.003)	-0.434*** (0.003)	-5.203*** (0.113)	-0.730*** (0.016)	-0.730*** (0.016)
m	m	-0.706*** (0.009)	-0.329*** (0.004)	-0.329*** (0.004)	-0.679*** (0.004)	-0.340*** (0.002)	-0.340*** (0.002)	-0.869*** (0.016)	-0.488*** (0.009)	-0.488*** (0.009)
o	o	-1.151*** (0.014)	-0.356*** (0.004)	-0.356*** (0.004)	-1.298*** (0.006)	-0.377*** (0.002)	-0.377*** (0.002)	-1.711*** (0.021)	-0.509*** (0.006)	-0.509*** (0.006)
y	m	0.767*** (0.017)	0.357*** (0.008)	0.173*** (0.004)	0.723*** (0.006)	0.362*** (0.003)	0.151*** (0.001)	1.427*** (0.038)	0.802*** (0.021)	0.200*** (0.005)
y	o	0.540*** (0.019)	0.167*** (0.006)	0.122*** (0.004)	0.248*** (0.006)	0.071*** (0.002)	0.052*** (0.001)	-0.243*** (0.034)	-0.072*** (0.010)	-0.034*** (0.005)
m	o	0.502*** (0.009)	0.155*** (0.003)	0.234*** (0.004)	0.651*** (0.004)	0.189*** (0.001)	0.326*** (0.002)	0.967*** (0.016)	0.288*** (0.005)	0.543*** (0.009)
N		45302			147292			33696		

Standard errors in parentheses, calculation based on Delta-method, significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The data is normalized at the respective sample mean. Monotonicity and concavity of the cost function are satisfied at more than 99 percent of the respective data points. The concavity constraint is locally imposed at the respective sample mean. The elasticities are calculated at mean predicted shares.

References

- Arntz, M., Gregory, T. (2014). What old stagers could teach us: Examining age complementarities in regional innovation systems. ZEW Discussion Papers 14-050, ZEW - Zentrum für Europäische Wirtschaftsforschung / Center for European Economic Research.
- Berger, M. C. (1985). The effect of cohort size on earnings growth: A reexamination of the evidence. *Journal of Political Economy*, 93(3), 561–573.
- BMAS (2013). Arbeitsmarktprognose 2030 Eine strategische Vorausschau auf die Entwicklung von Angebot und Nachfrage in Deutschland. Bundesministerium für Arbeit und Soziales Referat Information, Publikation, Redaktion (Ed.), Bonn.
- Borjas, G. J. (2003). The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market. *The Quarterly Journal of Economics*, 118(4), 1335–1374.
- Börsch-Supan, A. (2013). Myths, scientific evidence and economic policy in an aging world. *The Journal of the Economics of Ageing*, 1–2, 3–15.
- Börsch-Supan, A., Weiss, M. (2007). Productivity and age: Evidence from work teams at the assembly line. me Discussion Paper 148-2007, me - Munich Center for the Economics of Aging.
- Card, D., Lemieux, T. (2001). Can falling supply explain the rising return to college for younger men? A cohort-based analysis. *The Quarterly Journal of Economics*, 116(2), 705–746.
- Christensen, L. R., Jorgenson, D. W., Lau, L. J. (1973). Transcendental logarithmic production frontiers. *The Review of Economics and Statistics*, 55(1), 28–45.
- DESTATIS (2009). Bevölkerung Deutschlands bis 2060 - 12. koordinierte Bevölkerungsvorausberechnung. Statistisches Bundesamt (Ed.), Wiesbaden.
- Diewert, W. E., Wales, T. J. (1987). Flexible functional forms and global curvature conditions. *Econometrica*, 55(1), 43–68.
- Fitzenberger, B., Kohn, K. (2006). Skill wage premia, employment, and cohort effects: Are workers in Germany all of the same type? IZA Discussion Papers 2185, IZA - Institute for the Study of Labor.
- Freeman, R. B. (1979). The effect of demographic factors on age-earnings profiles. *The Journal of Human Resources*, 14(3), 289–318.
- Garloff, A., Roth, D. (2016). Regional age structure and young workers' wages. IAB-Discussion Paper 06/2016, Institute for Employment Research (IAB) - The Research Institute of the Federal Employment Agency, Nuremberg.
- Gartner, H. (2005). The imputation of wages above the contribution limit with the German IAB employment sample. FDZ Methodenreport 02/2005, The Research Data Centre (FDZ) of the Federal Employment Agency in the Institute for Employment Research, Nuremberg.
- Göbel, C., Zwick, T. (2013). Are personnel measures effective in increasing productivity of old workers? *Labour Economics*, 22(C), 80–93.

- Grant, J. H., Hamermesh, D. S. (1981). Labor market competition among youths, white women and others. *The Review of Economics and Statistics*, 63(3), pp. 354–360.
- Greene, W. H. (2012). *Econometric Analysis*. Upper Saddle River, NJ: Pearson, 7 ed. International Edition.
- Haas, A., Lucht, M., Schanne, N. (2013). Why to employ both migrants and natives? A study on task-specific substitutability. *Journal for Labour Market Research*, 46(3), 201–214.
- Hamermesh, D. S. (1996). *Labor Demand*. Princeton, New Jersey: Princeton University Press.
- Hamermesh, D. S., Grant, J. (1979). Econometric studies of labor-labor substitution and their implications for policy. *Journal of Human Resources*, 14(4), 543–562.
- Hicks, J. (1970). Elasticity of substitution again: Substitutes and complements. *Oxford Economic Papers*, 22(3), pp. 289–296.
- Horn, J. L., Cattell, R. B. (1967). Age differences in fluid and crystallized intelligence. *Acta Psychologica*, 26, 107–129.
- Jorgenson, D. W. (1986). Econometric methods for modeling producer behavior. In Z. Griliches, M. D. Intriligator (Eds.) *Handbook of Econometrics*, vol. 3, chap. 31, (pp. 1841–1915). Elsevier.
- Kosfeld, R., Werner, A. (2012). Deutsche Arbeitsmarkregionen - Neuabgrenzung nach den Kreisgebietsreformen 2007-2011. *Raumforschung und Raumordnung*, 70, 49–64.
- Lau, L. J. (1978). Testing and imposing monotonicity, convexity and quasiconvexity constraints. In M. Fuss, D. McFadden (Eds.) *Production Economics: A Dual Approach to Theory and Applications*, vol. 1 of *History of Economic Thought Chapters*, chap. A4. McMaster University Archive for the History of Economic Thought.
- Mahlberg, B., Freund, I., Cuaresma, J. C., Prskawetz, A. (2013). Ageing, productivity and wages in Austria. *Labour Economics*, 22, 5 – 15.
- Morin, L.-P. (2015). Cohort size and youth earnings: Evidence from a quasi-experiment. *Labour Economics*, 32, 99 – 111.
- Naz, A., Niebuhr, A., Peters, J. C. (2015). What's behind the disparities in firm innovation rates across regions? Evidence on composition and context effects. *The Annals of Regional Science*, 55(1), 131–156.
- Ottaviano, G. I. P., Peri, G. (2012). Rethinking the effect of immigration on wages. *Journal of the European Economic Association*, 10(1), 152–197.
- Paulus, W., Matthes, B. (2013). The German classification of occupations 2010 - Structure, coding and conversion table. FDZ-Methodenreport 08/2013, The Research Data Centre (FDZ) of the Federal Employment Agency in the Institute for Employment Research, Nuremberg.
- Reichelt, M. (2015). Using longitudinal wage information in linked data sets - The example of ALWA-ADIAB. FDZ-Methodenreport 01/2015, Research Data Centre of the Federal Employment Agency (BA) at the Institute for Employment Research (IAB), Nuremberg.
- Ryan, D. L., Wales, T. J. (2000). Imposing local concavity in the translog and generalized Leontief cost functions. *Economics Letters*, 67(3), 253–260.

Sapozhnikov, M., Triest, R. K. (2007). Population aging, labor demand, and the structure of wages. Working Papers 07-8, Federal Reserve Bank of Boston.

Sato, R., Koizumi, T. (1973). On the elasticities of substitution and complementarity. *Oxford Economic Papers*, 25(1), pp. 44–56.

Welch, F. (1979). Effects of cohort size on earnings: The baby boom babies' financial bust. *Journal of Political Economy*, 87(5), pp. S65–S97.