Productivity and the Age Composition of Work Teams: Evidence from the Assembly Line*

Axel Börsch-Supan†, Ismail Düzgün‡, and Matthias Weiss§

```
- - - - - - - - - - - - - - - - - - - - - - - - - - - -
- - - V E R Y   P R E L I M I N A R Y   D R A F T - - -
- - - - - in many parts still rather sketchy - - - - -
- - - - - - - - - - please do not quote - - - - - - - - - -
- - - - - - - - - - - - - - - - - - - - - - - - - - - -
```

April 17, 2007

Abstract

This papers studies the relation between the age composition of work teams and their productivity. We explore a unique data set that combines data on errors occuring in the production process of a large car manufacturer with detailed information on the personal characteristics of workers responsible for the errors. We are able to test the prevalent hypothesis regarding the effect of teams' age structure on their performance: Good performance requires a mix of young (fit, flexible) and old (experienced) workers. We find strong evidence against this hypothesis. In addition, we do not find evidence that productivity declines in older ages.

JEL codes: J24, J14, D24

Keywords: Age-Productivity Profiles, Teamwork, Age-Diverse Teams

---

*We are grateful to Melanie Lührmann and Jürgen Maurer for valuable comments and suggestions. Verena Arendt, Melanie Gräsle, and Simon Hilpert provided excellent research assistance. Funding by the State Baden-Württemberg, the Gesamtverband der Deutschen Versicherungswirtschaft, and the Hans-Böckler-Stiftung is gratefully acknowledged. We are especially grateful to the DaimlerChrysler truck assembly plant in Wörth, Germany. Without the great support from our contact persons in the different departments of the plant and from the side of the management and the works council, this project would not have been possible.

†MEA, Universität Mannheim, Germany and NBER.

‡MEA, Universität Mannheim, Germany.

§Corresponding author. MEA, Universität Mannheim, Germany, E-Mail: *weiss@mea.uni-mannheim.de.*

1 Introduction

The age composition in most of the developed world has been shifting towards older age groups for more than 100 years now. The steady population aging has intensified dramatically in consequence of the direct succession of the post war "baby boom" and the "baby bust" of the late 1960s. This accelerated aging process will have far-reaching economic consequences. Most prominent in the public discussion are the consequences for the pay-as-you-go financed social security systems. But before the babyboomers are going to retire (with their pensions having to be financed by the babybusters), extensive changes are to be expected on labor markets and in production: In Germany, e.g., the share of workers aged 55 years and older will more than double from 12% in 2005 to almost 25% in 2035. In view of these looming evolutions, it is important to better understand the relation between workers' age and their labor productivity.

Estimating age-productivity profiles has been on the agenda of labor economists for a long time. The main problem with estimating age-productivity profiles is that it requires a valid measure for productivity. There are many studies in occupational medicine, cognitive psychology, and gerontology that look at how different abilities and skills of humans evolve over their life-cycle. They look at muscle strength, sight, retentiveness, the functioning of lungs, kidney, and the heart, and many other measurable indicators. More or less concordantly, they find that from the age of 25 onwards, physical and mental fitness are deteriorating.[1] But there is certainly more to labor productivity than muscle strength, sight, and cognitive ability. Experience plays a role and is increasing with age. Hence, there is a need for more direct measures of productivity. Regarding the measurement of productivity, the existing literature can be broadly divided into four branches: (i) studies relating plant level productivity to the age of the plants' employees,[2] (ii) studies using individual wages as a productivity measure,[3] (iii) studies using interviews of managers on their employees' performance,[4] and (iv) studies using direct measures of individual productivity like, e.g., the number and quality of publications in academic research,[5] the quality of artists' paintings (in terms of auction proceeds),[6] or performance in sports and chess.[7]

These different approaches all have their vices and virtues. Plant level productivity

[1] This literature is surveyed in Skirbekk (2004) and Börsch-Supan, Düzgün, and Weiss (2005).

[2] E.g., Hellerstein and Neumark (1995) and (2004), Hellerstein, Neumark, and Troske (1999), Haltiwanger, Lane, and Spletzer (1999) and (2007), Crépon, Deniau, and Pérez-Duarte (2002), Aubert and Crépon (2004), Grund and Westergård-Nielsen (2005), Ilmakunnas and Maliranta (2005), Malmberg, Lindh, and Halvarsson (2005), and Prskawetz, Mahlberg, Skirbekk, Freund, Dworak, Lindh, Malmberg, Jans, Nordström, and Andersson (2005).

[3] E.g., Kotlikoff and Wise (1989), Kotlikoff and Gokhale (1992) and Laitner and Stolyarov (2005).

[4] E.g., Medoff and Abraham (1980), Hunter and Hunter (1984), McEvoy and Cascio (1989), Salthouse and Maurer (1996), and Schneider and Stein (2006).

[5] Jones (2005) and Weinberg and Galenson (2005).

[6] Galenson and Weinberg (2000) and (2001), and Galenson (2005).

[7] Fair (1994), (2005a), and (2005b).

can be measured easily and reliably but the level of aggregation is quite high when the goal is to study the relation between productivity and age. Furthermore, the age structure of firms is probably not exogenous.

Wages are the obvious productivity measure in many applications (returns to schooling, inter-personal comparisons, etc.) but when it comes to age profiles, the problem is that in many occupations, wages increase with age and/or seniority independently of productivity. Wage *de*creases are extremely rare.[8] Therefore, Kotlikoff and Wise (1989) look at earnings of insurance salesmen whose wages are proportional to the number of insurance contracts they sell.

Supervisors' assessments are problematic as they might reflect prejudices about age productivity profiles.

The studies subsumed as approach (iv) are able to measure productivity relatively exactly. Therefore, they can estimate age-productivity profiles quite precisely. But the occupations where this approach is feasible are rare and particular so that the results can hardly be generalized.

In addition, approaches (ii) through (iv) cannot take into account the fact that workers often work in teams and thereby affect one another's productivity. More specifically, if, e.g, older workers devote some of their working time to helping younger workers, the individual approach will underestimate older workers' productivity. Related aspects are workers' contributions to their team's work climate or how they deal with hectic situations (which again affects the productivity of the entire team).

In this paper, we follow a new approach: We look at productivity at the level of work teams. This takes into account the individual worker's contribution to her co-workers' productivity. In addition, the approach allows us to study the role of age *diversity* in teams for productivity.

Age diversity is preached as a means to boost productivity. From a purely economic view, the argument is very convincing: Older workers have the advantage of being more experienced while younger workers have the advantage of being physically and mentally fitter and more flexible. Hence, older and younger workers are complements and mixing them in teams improves productivity. But we know from psychology, that worker heterogeneity in teams can also have adverse effects. Team heterogeneity impairs identification with the team and its objectives, complicates communication, and impedes group cohesion, which are all essential for fruitful teamwork. To the best of our knowledge, the effect of age diversity on team productivity has not been studied yet.

[8] Lazear (1979) and (1981) explains the increasing age-earning profiles with incentive effects. Loewenstein and Sicherman (1991) and Frank and Hutchens (1993) show in experiments that workers have a preference for increasing wage profiles and explain this with loss aversion and problems of self-control.

2 The Data

2.1 Our Productivity Measure and Main Explanatory Variables

We exploit a unique data set that we have compiled from a truck assembly plant of the German car manufacturer DaimlerChrysler AG. At this plant, trucks are assembled by work teams on an assembly line. The quantity of output is determined by the speed of the assembly line. If work teams differ in productivity, this is not going to show up in differences in the quantity of output because the assembly line has the same speed for all teams. But production quality differs across work teams as they can make errors. Variation in productivity thus becomes manifest *only* in variation in production errors.

These errors are recorded by a "quality inspector" at the end of the assembly line. The quality inspector is able to assign every error to the work place where it happened. At any time, there is exactly one work team at any work place. In addition, every error is given a weight (between 5 and 95) that specifies the severity of the error. From this record of errors, we know which team has made how many errors of which severity on any day in 2003 through 2005. We observe 100 work teams at 50 work places on 821 days. The number of teams is double the number of work places because on every day, there is an early and a late shift. Our productivity measure is the sum of errors per team per day where the errors are rated with their respective weights. As larger teams work more and therefore make potentially more errors, we divide this sum by the number of workers in the team. E.g., if a team with 7 workers makes two errors on a day with weights 5 and 30, the sum of error weights is 35 in that team on that day. Our (inverse) productivity measure for this team for this day takes the value $\frac{35 \text{ error points}}{7 \text{ workers}} = 5$ error points per worker.

The information on errors is matched with personnel data that inform us about the daily composition of the work teams, personal characteristics of the workers such as age, sex, education, nationality, job tenure, and whether or not a worker is in her regular team.

In addition, we have data on the daily production plan which gives us information on the work load.

2.2 Some Descriptives

As the data set we use is quite unique, this section gives a brief description of the main variables we use. Table 2 in Appendix A reports descriptive statistics of all variables used in the paper.

Errors The number of errors occurring in the plant is low. We observe 14659 errors in 100 teams on 821 days. The probability that a random team on a random day makes an error is 21%. The distribution of error weights (only for those days and teams for which we observe errors) is given in Figure 1.

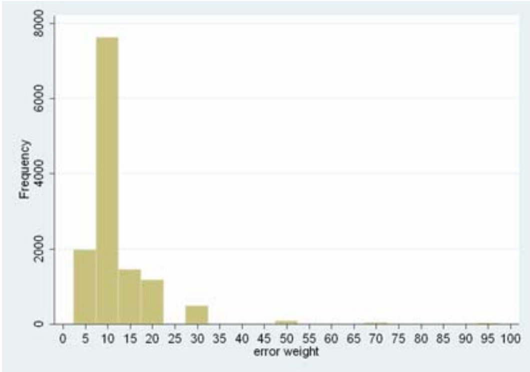

Figure 1: Distribution of error weights conditional on the observation of errors.

Age The age composition in the plant is fairly representative for the German workforce in that workers older than 55 are rare. Figure 2 shows the age distribution in the plant (red) in comparison to the age distribution of the German population. People younger

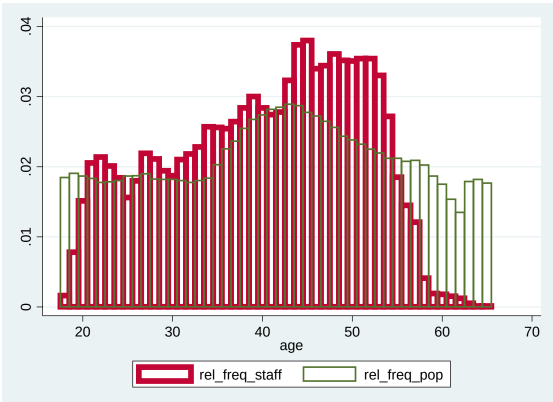

Figure 2: Age distribution in the plant (red) and in Germany (green)

then twenty are underrepresented because they are still in education or training. The share of workers aged 55 and over is low at the assembly line because many are already retired or have moved to better jobs . Figure 3 shows the distribution of average age of work teams which constitute the observation unit in our regression analysis.

Job Tenure In addition to age, we have information on workers' job tenure. Job tenure increases with age but the two variables are not perfectly correlated as workers are hired at different ages. The distribution of job tenure in the plant is shown in Figure 4. The spikes show hiring waves roughly every 5 to 10 years, the most recent having been just within the observation period (at job tenure=0). The distribution of average job tenure in work teams in Figure 5 shows that at hiring waves, the newly hired workers have been spread evenly over existing work teams as the histogram of average job tenure does not exhibit any comparable spikes. Figure 6 shows the relation between age and job tenure

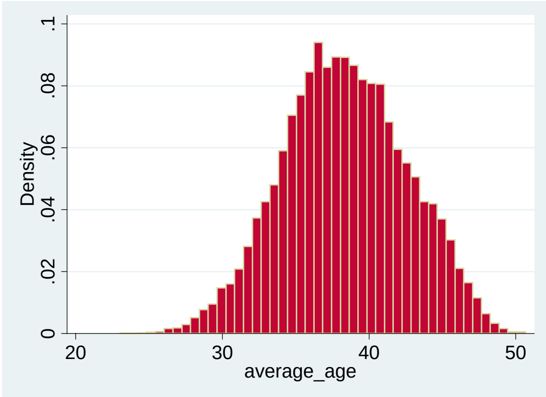

Figure 3: Distribution of average age of work teams.

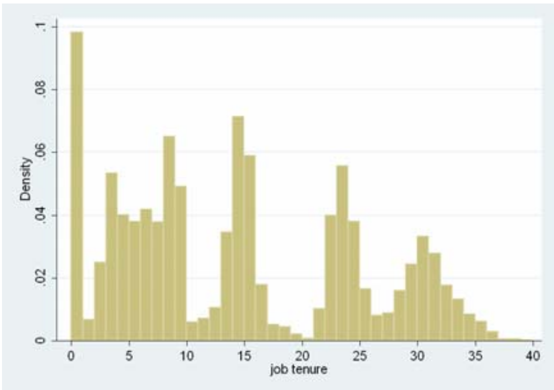

Figure 4: Distribution of job tenure in the plant.

in the plant. For any individual worker, age and job tenure are perfectly correlated over time, but as workers are hired at different ages, the overall correlation (over time *and* across workers) is “only” 0.79. The relation is tighter at the team level (see Figure 7).

Team Size The size of work teams varies between 4 and 36 workers. 90% of work teams have between 8 and 18 members (see Figure 8).

Sex The share of women in the plant is 3.7%. In 71% of all work teams, there are only men. In the other 29% of teams, women’s share is 13% on average. Within the sample period, the female share has increased from 2.9% in 2003 to 4.4% in 2005.

Nationality The composition of the personnel with respect to nationality is given in the following table:

nationality	German	French	Turkish	other
share	61.0%	31.2%	3.8%	4.0%

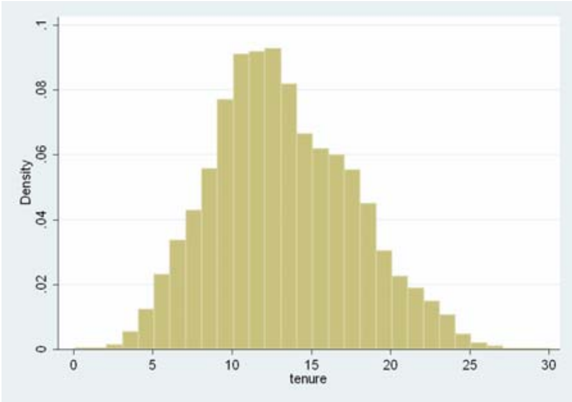

Figure 5: Distribution of average job tenure of work teams.

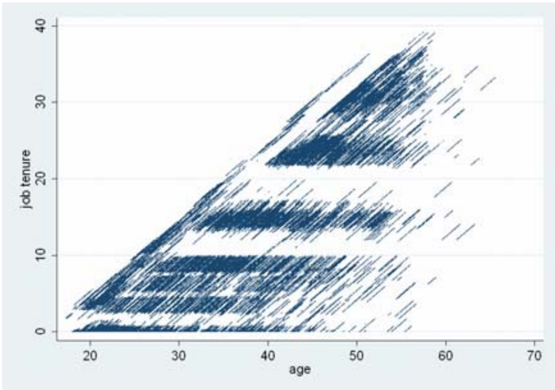

Figure 6: Scatter plot of job tenure (vertical axis) vs. age (horizontal axis).

Workload The production program and thereby the daily volume of work for every team varies over time. The required number of workers does not always exactly match the actual manning. We have daily information on the actual volume of work (measured in the number of required workers) and on actual manning for every day and every team. We use the percentage deviation of actual volume of work from actual manning as a measure of excess workload per worker. Figure 9 shows that the variation in excess workload is substantial.

Fluctuation The composition of work teams varies considerably over time. These fluctuations are due to vacation, sickness and—most importantly—due to compensatory time off for extra hours worked. Workers' contracts involve 7.5 hours per day while the assembly line runs 8 hours per work shift, so that workers accumulate one half hour overtime per day. Consequently, they can take every 16^{th} day off. This means that in a team of 16 workers, on an average day, one workers is absent due to compensatory time off. Another source of variation is given by the variation in workload. Figure 10 displays the distribution of the number of consecutive days without change in the team composition.

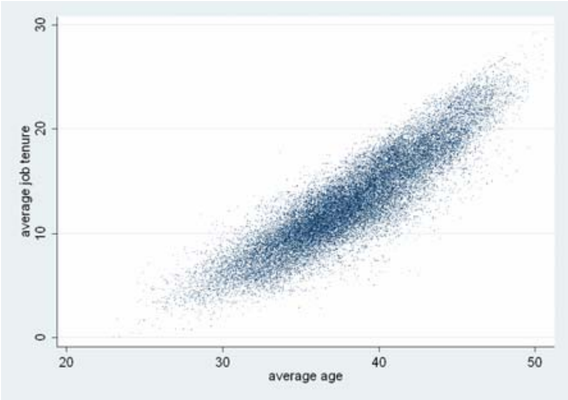

Figure 7: Scatter plot of average job tenure (vertical axis) vs. average age (horizontal axis) of work teams

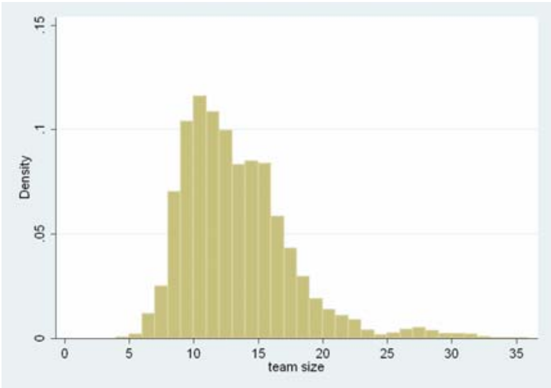

Figure 8: Distribution of team size in the plant.

It shows that team composition usually changes from day to day.

External Workers Each worker is assigned to one team as her "regular" team. But—due to fluctuations in team composition workers work outside their regular team 6% of the time on average. Some workers are explicitly designed for "team hopping" in order to stand in wherever needed. Figure 11 displays the distribution over time and across work teams of the share of workers external to the work team.

2.3 Age Diversity

Results from medicine, psychology and gerontology teach us that younger workers are physically and mentally fitter while we expect older workers to be more experienced. These different qualities of workers with different ages lead us to expect that—in addition to the average age of a work team—the age composition of teams is important for their performance. The prevalent hypothesis in this regard is ...

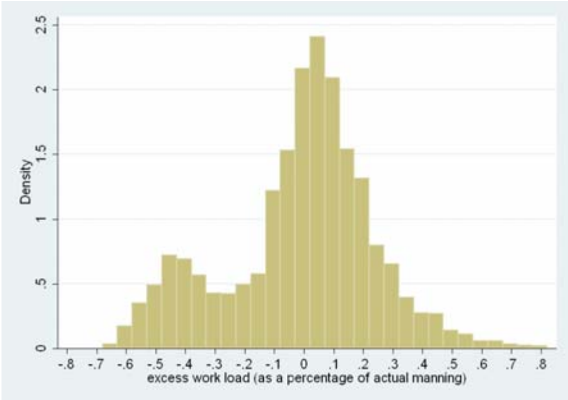

Figure 9: Distribution of excess work load (as a share of actual manning).

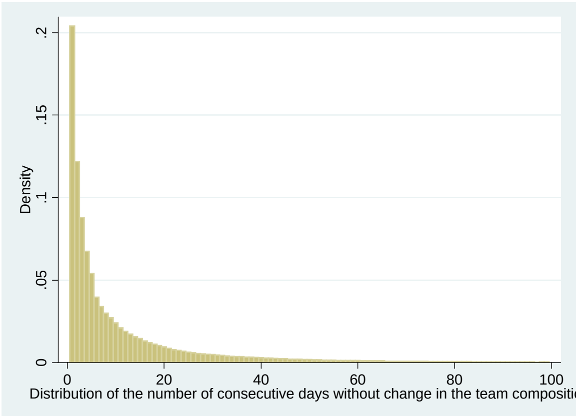

Figure 10: Distribution of the number of consecutive days without change in the team composition

Hypothesis 1 *... that good performance requires a well-balanced mix of younger (supposedly fitter, more flexible) workers and older (supposedly more experienced) workers. More age diversity is better.*

In order to test this hypothesis, we have to construct a measure of age diversity. Our interpretation of Hypothesis 1 is that having workers of each age in similar shares is good. In other words: the closer the age distribution is to uniformity, the better. Our diversity variables thus measure how close the actual age distributions in work teams are to uniformity. There are several ways to do this.

2.3.1 Closeness to Uniformity

Our preferred approach calculates the surface area overlaps of the actual age histograms of work teams from the "ideal" uniformity histogram. Figure 12 illustrates this for the showcase team with an age distribution given by the orange histogram (1 worker aged 25

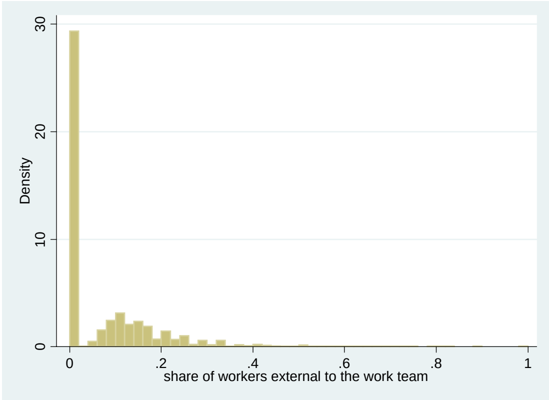

Figure 11: Distribution of the share of workers external to the work team

-35 years, 6 workers age 35 - 45 years, 3 workers aged 45 - 55 years). In terms of Hypothesis 1, the ideal team would have an age distribution as given by the blue histogram (2 workers in each age category). Our measure of how close the orange showcase team is to the blue "ideal" team is given by the surface area overlap (the hatched area in the figure). The larger this area is, the more age diverse is the work team.[9] More specifically, the hatched

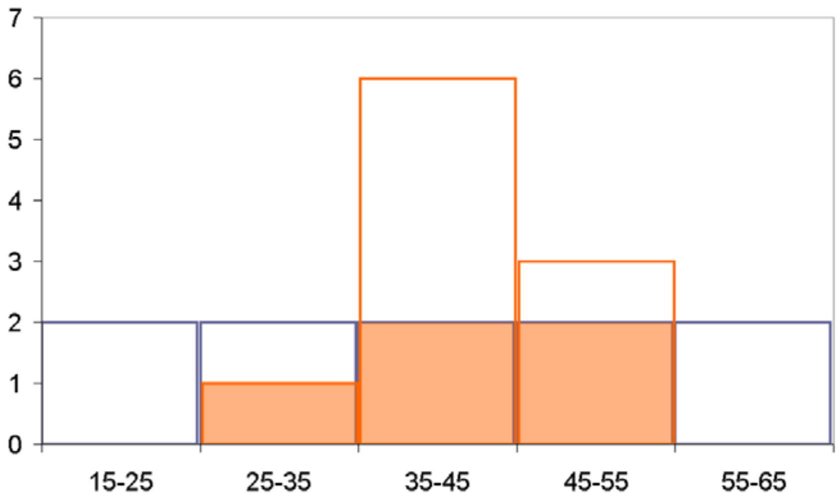

Figure 12: Illustration of our preferred measure of diversity.

area for team j is given by:

$$area_j = \sum_{i=1}^{k} \min\left(\frac{x_{ij}}{N_j}, \frac{1}{k}\right)$$

[9]Note that the concept of diversity is different from the concept of variance. While a small variance coincides with low diverstiy, a very large variance would mean that in the team there are some very young workers and some "very old" workers and no workers aged 25 - 55. Such a team would have rather *low* diversity according to our measure as only two out of five age groups would be represented in the team.

where k is the number of age categories, x_{ij} is the number of workers in team j who belong to age category i, $N_j = \sum_{i=1}^{k} x_{ij}$ is the team size, x_{ij}/N_j is the relative frequency of age i workers in team j, i.e., the height of the age i bar in the histogram. $area_j$ is between $1/k$ (if all workers in team j belong to the same age category) and 1 (if team j has uniform age distribution). In order for this diversity measure to lie within the unit interval independently of k, we normalize it to get:

$$diversity_j^{normalized} = \frac{\sum_{i=1}^{k} \min\left(\frac{x_{ij}}{N_j}, \frac{1}{k}\right) - \frac{1}{k}}{1 - \frac{1}{k}} = \frac{\sum_{i=1}^{k} \min\left(\frac{k \cdot x_{ij}}{N_j}, 1\right) - 1}{k - 1} \tag{1}$$

Figure 13 shows the distribution of this measure in our sample. A variant of it sums the

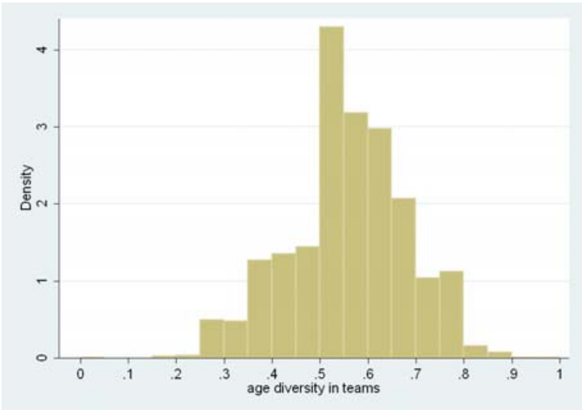

Figure 13: Distribution of our diversity measure.

squared differences in the histogram bars:

$$\sum_{i=1}^{k} \left(\frac{x_{ij}}{N_j} - \frac{1}{k}\right)^2 \in \left[0, \frac{k-1}{k}\right]$$

Normalization yields:

$$diversity_j^{squared} = 1 - \frac{k}{k-1} \cdot \sum_{i=1}^{k} \left(\frac{x_{ij}}{N_j} - \frac{1}{k}\right)^2 \tag{2}$$

The two variables given by (1) and (2) yield very similar results.

2.3.2 Probability that 2 Workers are in Different Age Categories

We also tried other measures of diversity. One is calculated as the probability that within a work team, two randomly chosen workers are in different age categories:

$$prob_j = \sum_{i=1}^{k} \frac{x_{ij}}{N_j} \cdot \left(1 - \frac{x_{ij} - 1}{N_j - 1}\right)$$

This probability is between 0 (if all workers belong to the same age category) and $\frac{k-1}{k} \cdot \frac{N_j}{N_j-1}$ (in case of uniformity). In order for this diversity measure to lie within the unit interval independently of k, we normalize it to get:

$$prob_j^{normalized} = \frac{\sum_{i=1}^{k} \frac{x_{ij}}{N_j} \cdot \left(1 - \frac{x_{ij}-1}{N_j-1}\right)}{\frac{k-1}{k} \cdot \frac{N_j}{N_j-1}} = \frac{k}{k-1} \cdot \left(1 - \sum_{i=1}^{k} \left(\frac{x_{ij}}{N_j}\right)^2\right) \tag{3}$$

If in a work team, this variable is large, it means that age diversity is high: Workers are equally spread over many age categories. If the variable is zero, it means that all workers belong to the same age category. Age diversity is minimal.

The drawback of $diversity_j^{normalized}$, $diversity_j^{squared}$, and $prob_j^{normalized}$ is that they do not take into account the "distance" between age categories. If in a team, only two out of five age categories are occupied, these diversity measures do not distinguish, whether the two age categories are close to each other (e.g., 25-35 and 35-45) or distant (15-25 and 55-65). Therefore, we considered in addition the variation coefficient of age within work teams, which explicitly takes into account the difference between each workers age and the team average.

2.3.3 Variation Coefficient of Age within Teams

The variation coefficient of age is given by

$$varcoeff_j = \frac{\sqrt{\frac{1}{N_j} \cdot \sum_{\iota=1}^{N_j} \left(age_{\iota j} - \frac{1}{N_j} \cdot \sum_{\iota=1}^{N_j} age_{\iota j}\right)^2}}{\frac{1}{N_j} \cdot \sum_{\iota=1}^{N_j} age_{\iota j}} \tag{4}$$

where age_ι is age of worker ι in team j.

While the variables in (1), (2), and (3) measure how uniformly *workers* within a team are *spread over age groups*, the variation coefficient of age measures how equally or unequally *age* is *distributed among workers*. The major drawback of the variation coefficient is that it takes the highest value if only two age categories are occupied: the youngest and the oldest. This does not exactly comply with our notion of diversity that all age groups are represented in a team.

For lack of an established theory on the subject, we try all the diversity variables. They lead to very similar results. $diversity_j^{normalized}$ from equation (1) comes closest to our conception of diversity and is used in the specifications we report in the next section.

3 Results

This section reports results on the relation between the age compostition of work teams and the number and severity of errors made in the truck assemble process. As explained in

Section 2.1, our (inverse) productivity measure is the weighted sum of errors per team per day divided by team size where each error is given a weight according to its severity. This variable is the dependent variable in all regressions. Our observation unit is a team-day. We observe 100 work teams on 821 days. As—along the assembly line—work places differ quite substantially and the allocation of workers to these work places and to early vs. late shift may be endogenous, we control for work team fixed effects. Only the day-to-day variation is used to identify our estimates.[10] This variation results from fluctuations in the work team composition due to vacation, sickness and—most importantly—due to compensatory time off for extra hours worked (see Section 2). From discussions with managers at the plant on how they replace missing workers we are confident that this variation is truly exogenous. There is no scope for optimization here.

The obvious measures of the age composition of work teams are the average age and the variation coefficient. We do not use them for the following reasons. As can be seen in Figure 3, average team age varies only between 25 and 50 years. (95% of the teams have an average age between 29 and 46 years.) Therefore, we use age shares for the categories 15-25 years, 25-35 years, 35-45 years, 45-55 years, and 55-65 years. As a measure of age diversity, we prefer the measure given in equation (1) for reasons set forth in Subsection 2.3. Results are reported in Table 1.

Results Baseline Specification The left column of Table 1 shows the results of our baseline regression. As the specification includes a number of interactions with average team age, the coefficients on the age shares cannot be interpreted as marginal effects. The estimated marginal effects including the interaction effects are reported in Table 3 in Appendix B. An increase in the share of workers aged 15-25 years (accompanied by a respective decrease in the share of 35-to-45-year-olds) is associated with an increase in errors by 0.035 (which corresponds to one tenth of a standard deviation). The youngest workers are significantly less productive than the workers in the reference age category, the prime-aged 35-to-45-year-olds. The 25-to-35-year-olds make significantly fewer errors. The productivity (in terms of errors) of workers older than 45 years does not significantly differ from that of prime-age workers.

[10] We also tried *work place* fixed effects, thereby using the variation over work shifts in addition the variation over time. The results are very similar.

Table 1: Regression Results

dependent variable: error intensity (measured as the sum of error weights per person)						
	baseline specification		correcting for sample selection		correcting for selection controlling for job tenure	
share of workers aged ...						
15 - 25 years	0.0559	(0.125)	0.0760	(0.037)	0.0932	(0.013)
25 - 35 years	-0.0564	(0.016)	-0.487	(0.039)	-0.494	(0.039)
35 - 45 years	-	-	-	-	-	-
45 - 55 years	-0.0302	(0.193)	-0.0283	(0.221)	-0.0203	(0.401)
55 - 65 years	-0.0668	(0.115)	-0.0657	(0.120)	-0.0657	(0.134)
interactions of average team age with ...						
schooling years	-0.000497	(0.037)	-0.000398	(0.094)	-0.0007602	(0.056)
share of women	-0.01224	(0.011)	-0.0133	(0.006)	-0.124	(0.200)
share of externals	0.00580	(0.044)	0.00486	(0.093)	-0.00933	(0.090)
team size	0.000843	(0.000)	0.000500	(0.004)	0.000713	(0.015)
early shift	0.00216	(0.000)	0.00195	(0.001)	-0.00135	(0.271)
days without change	0.0000429	(0.000)	0.0000464	(0.000)	0.0000317	(0.155)
excess workload	-0.000408	(0.530)	-0.000450	(0.488)	0.00276	(0.024)
div. w.r.t. nationality	-0.00315	(0.167)	-0.00284	(0.231)	0.00277	(0.546)
diversity with respect to ...						
age	0.0614	(0.002)	0.0636	(0.001)	0.0637	(0.001)
schooling years	0.0408	(0.009)	0.0412	(0.009)	0.0396	(0.012)
nationality	0.154	(0.091)	0.133	(0.146)	0.0160	(0.901)
control variables						
schooling years	0.0389	(0.000)	0.0254	(0.008)	0.0353	(0.004)
share of women	0.429	(0.016)	0.475	(0.008)	0.435	(0.092)
share of French	0.0199	(0.534)	-0.0301	(0.364)	-0.0347	(0.305)
share of Germans	-0.0288	(0.463)	-0.0656	(0.099)	-0.0801	(0.045)
share of Turkish	-0.0190	(0.639)	-0.0718	(0.084)	-0.0757	(0.069)
share of externals	-0.235	(0.035)	-0.185	(0.099)	0.131	(0.396)
team size	-0.0426	(0.000)	-0.0247	(0.000)	-0.0299	(0.001)
early shift	-0.0927	(0.000)	-0.0835	(0.000)	-0.00691	(0.838)
days without change	-0.00178	(0.000)	-0.00183	(0.000)	-0.00145	(0.014)
excess workload	0.0646	(0.019)	0.0678	(0.014)	-0.00342	(0.924)
workload · age div.	-0.0602	(0.001)	-0.0638	(0.001)	-0.0680	(0.000)
Inverse Mills Ratio	-	-	-0.0105	(0.000)	-0.00992	(0.000)
R^2 within	0.0089		0.0098		0.0103	
R^2 between	0.1040		0.0237		0.0274	
R^2 overall	0.0122		0.0082		0.0089	

continued on the next page...

Sample Selection Bias One recurrent criticism of studies on age-productivity profiles is that—due to sample selection bias—they overestimate the productivity of older workers. In fact, workers older than 55 years are underrepresented in our sample. The obvious suspicion is that the remaining workers are a positive selection. The less motivated, less healthy workers probably retire earlier. While this is probably true for samples that cover the whole economy or entire production plants, the selection into (and out of) our sample is more complex. Our sample consists only of workers on the assembly line. Even the foremen are not included. Workers aged 60 in our sample are sufficiently healthy and motivated to still be in the workforce but they were not good enough to be promoted to jobs off the assembly line.

The common problem with the correction for sample selection is that—by definition—we usually do not have information on those subjects who are not in the sample. As our sample covers several years, we do have information on those workers who enter or exit the sample within these three years. This enables us to estimate a Heckman-style selection correction model. As our observation unit in the regression is a work team while selection into the sample is an individual phenomenon, we have to aggregate individual Mills ratios to team Mills ratios (see Appendix C).

Results with Selection Correction The central columns of Table 1 and Table 3 reports the results from a regression that corrects for sample selection bias. The results show that promotion seems to be more important than retirement as a factor driving sample selection. The coefficient on the share of 15-25-year-olds increases implying that the young workers in the sample are a positive selection. The marginal effects of the shares of workers older than 45 years become significantly negative. This implies that the older workers in our sample are actually an adverse selection.[11]

Controlling for Job Tenure The obvious explanation for the good performance of older workers is experience. Older workers have more experience and therefore make fewer mistakes. We have information on workers' date of entry in the plant and can thus calculate workers' job tenure. The results of a regression that controls for job tenure are reported in the right columns of Tables 1 and 3. If job tenure is held constant, the effect of an increase in the shares of workers older than 45 years becomes insignificant. The second part of Table 1 and Table 3 show the coefficients respectively marginal effects of job tenure. Too long job tenure is found to be counterproductive, just as too short job tenure. There seems to be an optimal length of job tenure, namely 16-24 years. The non-monotoneous relation reveals that job tenure does not only measure experience. Job security may also increase with seniority and motivation may degrade with job tenure.

[11] Our strategy for correction of sample selection is not perfect yet. Promotion drives selection for younger workers while retirement drives selection for older workers. This is not taken into account in our present procedure. We will improve on this in the next version of the paper.

Table 1: Regression Results (continued)

dependent variable: error intensity (measured as the sum of error weights per person)						
	baseline specification		correcting for sample selection		correcting for selection controlling for job tenure	
share of workers with job tenure of ...						
0 - 8 years	-	-	-	-	0.000560	(0.986)
8 - 16 years	-	-	-	-	0.0766	(0.001)
16 - 24 years	-	-	-	-	-	-
24 - 32 years	-	-	-	-	0.0571	(0.014)
32 - 40 years	-	-	-	-	0.113	(0.007)
interactions of average job tenure with ...						
schooling years	-	-	-	-	0.000326	(0.445)
share of women	-	-	-	-	0.000127	(0.991)
share of externals	-	-	-	-	0.0187	(0.002)
team size	-	-	-	-	-0.000268	(0.403)
early shift	-	-	-	-	0.00408	(0.838)
days without change	-	-	-	-	0.0000128	(0.600)
excess workload	-	-	-	-	-0.00342	(0.002)
div. w.r.t. nationality	-	-	-	-	-0.00787	(0.118)

observations: 70476 (unbalanced panel of 100 work teams on 821 work days). Reference category for age shares: 35-45 years, reference category for job tenure shares: 16-24 years. Diversity measures are calculated using equation (1). $p-$values are in brackets. All specifications control for work team fixed effects.

Team Diversity We also look at the effect of age diversity on work teams' performance. According to Hypothesis 1, age diversity is good for performance because younger and older workers complement one another. This hypothesis is clearly rejected by our data. The more age diverse work teams are, the more errors they make. This finding is highly significant and robust with respect to all diversity measures we discuss in equations (1) through (4) in Section 2.3 and with respect to all specifications we tried. It seems thus, that the detrimental effect on communication and team cohesion dominates advantageous effects of complementarity. We find similar effects for diversity with respect to education and nationality.

Control Variables The coefficients on the share of women and on the interaction between the female share and average age have opposite effects. A higher share of women increases the number of errors in young teams and decreases the number of errors in older teams. It has to be noted here, that we do not observe who in the team makes the error. It might be the women but it might just as well be the men who are distracted from the presence of a woman. (The share of women on the assembly line is 3.7%.) In the latter case, older workers are less distracted by women than young workers...

The longer a team works together in the same composition, the fewer errors they make. This results accords with the result on team diversity. Communication and cohesion are important for performance and are impeded by frequent changes in team composition. The positive coefficient on the interaction with average team age suggests that older workers find it easier to deal with frequent changes in team composition.

As could be expected, excess workload is detrimental for production quality. Holding constant job tenure, older workers have more difficulties coping with excess workload. On the other hand, holding constant age, longer job tenure helps in situation with excess workload. As job tenure increases with age, the two effects offset each other.[12] We also interacted age diversity with excess workload: More age diverse teams are better in dealing with excess workload. But this positive effect of age diversity would dominate only in situations when excess workload were more than plus 100%.

Fewer errors per worker occur in larger teams. The absolute number of errors increases with team size but less than proportionately so that per worker, errors decrease with team size. Probably, in a larger team, there is more scope to correct errors right away.

Fewer errors also occur in the early shift. The morning seems to be better suited for flawless working.

Teams with a higher average education make more errors. Education is measured as years of schooling plus years of formal vocational training. As the tasks on the assembly line do not require higher education, to much of it seems detrimental.

Another significant variable is the share of workers that do not belong to the team regularly. There are workers who are especially designed for "team hopping". These are skilled and experienced workers who are able to do tasks at many different workplaces. Their presence seems to help avoid errors.

4 Conclusion

We study the relation between the age structure of work teams and their performance in an assembly plant of a car manufacturer. To our knowledge, this is the first empirical study to look at the productivity of work *teams* in relation to age—in particular age *diversity*. The analysis at the team level enables us to study the effects of team composition with respect to several dimension of heterogeneity like age, education, and nationality. In addition—and maybe as importantly—within-team externalities like a worker's contribution to her co-workers' productivity are internalized.

We use data on errors made in the production process to construct our (inverse) pro-

[12] See the insignificant coefficients on the interaction of excess workload with age in the specifications where job tenure is not controlled for.

ductivity measure. As the production quantity is given by the speed of the assembly line, which is equal across all work places, work teams that are differently productive only differ in the errors they make. From these quality data, we know for all 100 work teams the number and severity of errors they made on any day in 2003 through 2005. We combine these data with data from the personnel department that gives us the daily composition of work teams with personal characteristics of the workers. In addition, we have information on the daily work load. We correct for sample selection by aggregating individual Mills ratios to the team level. This is not perfect yet, as there are different factors driving the sample selection of younger and older workers. We will improve on our selection correction strategy in the next version of the paper.

Our findings suggest that productivity is lowest among workers aged 15 to 25 and highest among workers age 25 to 35 years. After the age of 45 years, the age-productivity profile is slightly increasing. If we control for job tenure, the age-productivity profile becomes essentially flat from age 35 onwards.

Age diversity is often preached as a means to boost productivity. Age diversity is important to benefit from complementarities between fit, flexible younger workers and experienced older workers. Yet, empirical evidence on this issue is still lacking. Our results provide strong evidence against the hypothesis that age diversity enhances productivity. Notwithstanding its virtues, diversity comes at a cost: It impairs communication and group cohesion. Both are essential for fruitful teamwork. This counteracting effect of diversity on productivity seems to dominate the beneficial effect that works through the exploitation of potential complementarities. The result is corroborated by our findings regarding diversity with respect to other characteristics such as nationality and education and by the negative effect of fluctuations in team composition.

The huge data set and the truly exogenous variation in team composition enable us to estimate the effects of age shares and age diversity quite precisely. In addition, we are able to correct potential sample selection bias. On the other hand, our results refer to a single plant only. However, we believe that our results are of general interest. Regarding our estimates of the age-productivity profile, we find it interesting that even at the assembly line—where labor is physically demanding and experience should be comparatively unimportant as tasks are rather simple and do not require substantial traing—productivity does not decline at older ages. Our results concerning age diversity may result from the fact that the tasks at the assembly line are rather homogeneous while team heterogeneity can show its strengths best when tasks are heterogeneous. Yet, other forms of production in teamwork may also involve homogeneous team tasks. In any case, our results show that age diversity should not be pursued for its own end. Diversity complicates teamwork and this counteracting effect may well sweep off the benefits from exploiting complementarities.

References

Aubert, P., and B. Crépon (2004): "La productivité des salariés âgés: une tentative d'estimation," *Économie et Statistique*, 368, 95 – 119.

Börsch-Supan, A., I. Düzgün, and M. Weiss (2005): "Altern und Produktivität: Zum Stand der Forschung," *MEA Discussion Paper*, 73-05, http://www.mea.uni–mannheim.de.

Crépon, B., N. Deniau, and S. Pérez-Duarte (2002): "Wages, Productivity, and Workers Characteristics: A French Perspective," *mimeo*, Centre de Recherche en Économie et Statistique (CREST) and Institut National de la Statistique et des Études Économiques (INSEE), Paris, France.

Fair, R. (1994): "How Fast Do Old Men Slow Down?," *The Review of Economics and Statistics*, 76(1), 103 – 118.

——— (2005a): "Estimated Age Effects in Athletic Events and Chess," *mimeo*, Cowles Foundation, Yale University, http://fairmodel.econ.yale.edu/.

——— (2005b): "Estimated Age Effects in Baseball," *mimeo*, Cowles Foundation, Yale University, http://fairmodel.econ.yale.edu/.

Frank, R. H., and R. M. Hutchens (1993): "Wage, Seniority, and the Demand for Rising Consumption Profiles," *Journal of Economic Behavior and Organization*, 21(3), 251 – 276.

Galenson, D. W. (2005): "The Greatest Artists of the Twentieth Century," *NBER Working Papers*, 11899.

Galenson, D. W., and B. A. Weinberg (2000): "Age and the Quality of Work: The Case of Modern American Painters," *Journal of Political Economy*, 108(4), 671 – 777.

——— (2001): "Creating Modern Art: The Changing Careers of Painters in France from Impressionism to Cubism," *The American Economic Review*, 91(4), 1063 – 1071.

Grund, C., and N. Westergård-Nielsen (2005): "Age Structure of the Workforce and Firm Performance," *IZA Discussion Paper*, 1816.

Haltiwanger, J. C., J. I. Lane, and J. R. Spletzer (1999): "Productivity Differences Across Employers: The Roles of Employer Size, Age, and Human Capital," *The American Economic Review*, 89(2), 94 – 98.

——— (2007): "Wages, Productivity, and the Dynamic Interaction of Business and Workers," *Labour Economics*, forthcoming.

Hellerstein, J. K., and D. Neumark (1995): "Are Earnings Profiles Steeper Than Productivity Profiles? Evidence from Israeli Firm-Level Data," *Journal of Human Resources*, 30(1), 89 – 112.

——— (2004): "Production Function and Wage Equation Estimation with Heterogeneous Labor: Evidence Fom a New Matched Employer-Employee Data Set," *NBER Working Paper*, 10325.

Hellerstein, J. K., D. Neumark, and K. R. Troske (1999): "Wages, Productivity, and Worker Characteristics: Evidence from Plant Level Production Function and Wage Equations," *Journal of Labor Economics*, 17(3), 409 – 446.

Hunter, J. E., and R. F. Hunter (1984): "Validity and Utility of Alternative Predictors of Job Performance," *Psychological Bulletin*, 96, 72 – 98.

Ilmakunnas, P., and M. Maliranta (2005): "Technology, Labour Characteristics and Wage-Productivity Gaps," *Oxford Bulletin of Economcis and Statistics*, 67(5), 623 – 645.

Jones, B. F. (2005): "Age and Great Invention," *NBER Working Paper*, 11359.

Kotlikoff, L., and D. Wise (1989): "Employee Retirement and a Firm's Pension Plan," in *The Economics of Aging*, ed. by D. Wise, pp. 279 – 334. University of Chicago Press, Chicago.

Kotlikoff, L. J., and J. Gokhale (1992): "Estimating a Firm's Age-Productivity Profile Using the Present Value of Workers' Earnings," *Quarterly Journal of Economics*, 107(4), 1215 – 1242.

Laitner, J., and D. Stolyarov (2005): "Technological Progress and Worker Productivity at Different Ages," *mimeo*, University of Michigan.

Lazear, E. (1979): "Why is There Mandatory Retirement?," *Journal of Political Economy*, 87(6), 1261 – 1284.

——— (1981): "Agency, Earnings Profiles, Productivity, and Hours Restrictions," *American Economic Review*, 71(4), 606 – 620.

Loewenstein, G., and N. Sicherman (1991): "Do Workers Prefer Increasing Wage Profiles?," *Journal of Labor Economics*, 9(1), 67 – 84.

Malmberg, B., T. Lindh, and M. Halvarsson (2005): "Productivity Consequences at the Plant Level of Work-Force Ageing: Stagnation or a Horndal Effect?," *mimeo*, Institute for Future Studies, Stockholm, Sweden.

McEvoy, G. M., and W. F. Cascio (1989): "Cumulative Evidence of the Relationship Between Employee Age and Job Performance," *Journal of Applied Psychology*, 74(1), 11 – 17.

Medoff, J. L., and K. G. Abraham (1980): "Experience, Performance, and Earnings," *Quarterly Journal of Economics*, 95(4), 703 – 736.

Prskawetz, A., B. Mahlberg, V. Skirbekk, I. Freund, M. W. Dworak, T. Lindh, B. Malmberg, A.-C. Jans, O. S. Nordström, and F. Andersson (2005): "The Impact of Population Ageing on Innovation and Productivity Growth in Europe," Vienna Institute of Demography, Austrian Academy of Sciences, Wien, Austria and Institute for Futures Studies, Stockholm, Sweden.

Salthouse, T. A., and T. J. Maurer (1996): "Aging, Job Performance, and Career Development," in *Handbook of the Psychology of Aging, 4th Edition*, ed. by J. E. Birren, and K. W. Schaie, pp. 353 – 364. Kluwer Academic Publishers, New York.

Schneider, H., and D. Stein (2006): "Personalpolitische Strategien Deutscher Unternehmen zur Bewältigung demografisch bedingter Rekrutierungsengpässe bei Führungskräften," Research Report 6, IZA, Bonn, Germany.

Skirbekk, V. (2004): "Age and Individual Productivity: A Literature Survey," in *Vienna Yearbook of Population Research*, ed. by G. Feichtinger, Vienna. Verlag der Österreichischen Akademie der Wissenschaften, Vienna.

Weinberg, B. A., and D. W. Galenson (2005): "Creative Careers: The Life Cycles of Nobel Laureates in Economics," *NBER Working Paper*, 11799.

Appendix

A Descriptive Statistics

Table 2 gives descriptive statistics of all variables used in the regressions.

Table 2: Descriptive Statistics

variable	mean	median	minimum	maximum	std. dev.
weighted sum of errors	0.0846	0	0	95	0.383
age	37.7	37.5	17.3	65.6	10.9
mean team age	37.7	37.8	22.0	56.1	4.66
age diversity	0.563	0.563	0	1	0.133
tenure	12.0	9.05	0	41.7	9.74
mean team tenure in team	12.0	11.6	0.242	32.0	4.54
schooling years	11.4	11	9	20	1.93
mean team schooling	11.4	11.2	9	14.0	0.612
schooling diversity	0.516	0.524	0	1	0.157
female (dummy)	0.0367	0	0	1	0.188
female share	0.0367	0	0	0.667	0.059
external (dummy)	0.0623	0	0	1	0.242
share of externals	0.0623	0	0	1	0.089
divers. w.r.t. nationality	0.640	0.667	0	1	0.146
early shift (dummy)	0.326	0	0	1	0.469
excess workload	0.0190	-0.015	-0.702	0.810	1.93
team size	10.2	9	3	28	3.69
days without change	14.6	5	1	270	27.2

B Significance Tests for Marginal Effects

In this appendix, we report tests on the significance of the gradient of the age and job tenure variables as well as some control variables. Table 2 reports the test results for the regressions presented in Table 1. These regressions include a set of interactions with age and job tenure. The gradient of variable x is thus the linear combination of all the coefficients that refer to variable x where coefficients on interaction terms are multiplied by the sample means of the respective variables.

The gradients of the age shares are calculated supposing that a rise in the share of a-to-a'-year-olds is accompanied by a respective decrease in the share of 35-to-45-year-olds. This change in the age composition implies a respective change in the average age of the work team. This implied change in average age is estimated on the sample. The estimated implied change in average age is used to calculate the gradients of age shares. The following equations illustrate our approach. A simplified version of the estimated equation looks as follows. Errors are a function of the age shares, excess workload and an interaction between average team age and workload.

$$errors = \beta_0 + \beta_{15-25} \cdot AgeShare_{15-25} + ... + \beta_{55-65} \cdot AgeShare_{15-25}$$

$$+\beta_{wl} \cdot workload + \beta_{wla} \cdot workload \cdot AverageAge + ... + \varepsilon$$

The effect of an increase in the share of 15-to-25-year-olds (accompanied by a respective decrease in the share of 35-45-year-olds) is given by:

$$\frac{\partial errors}{\partial AgeShare_{15-25}} = \beta_{15-25} + \beta_{wla} \cdot workload \cdot \frac{\partial AverageAge}{\partial AgeShare_{15-25}} \tag{5}$$

$\frac{\partial AverageAge}{\partial AgeShare_{15-25}}$ is estimated on our sample using the simple equation:

$$AverageAge = \alpha_{15-25} \cdot AgeShare_{15-25} + ... + \alpha_{55-65} \cdot AgeShare_{15-25} + u$$

The estimate of α_{15-25} is plugged in equation (5):

$$\overline{\frac{\partial errors}{\partial AgeShare_{15-25}}} = \beta_{15-25} + \beta_{wla} \cdot \overline{workload} \cdot \alpha_{15-25}$$

where $\overline{x}$ denotes the sample mean of variable x.

Table 3: Marginal Effects

	baseline specification		correcting for sample selection		correcting for selection controlling for job tenure	
			gradients of age shares			
15 - 25 years	0.0350	(0.048)	0.0880	(0.000)	0.0940	(0.000)
25 - 35 years	-0.701	(0.000)	-0.420	(0.013)	-0.0490	(0.007)
35 - 45 years	0	-	0	-	0	-
45 - 55 years	-0.0148	(0.388)	-0.0350	(0.046)	-0.0207	(0.325)
55 - 65 years	-0.0392	(0.213)	-0.0777	(0.016)	-0.0665	(0.086)
			gradients of *job tenure shares*			
0 - 8 years	-	-	-	-	0.0509	(0.055)
8 - 16 years	-	-	-	-	0.105	(0.000)
16 - 24 years	-	-	-	-	0	-
24 - 32 years	-	-	-	-	0.0408	(0.063)
32 - 40 years	-	-	-	-	0.0759	(0.053)
			gradient of the *share of women* at an average team age of ...			
25 years	0.108	(0.079)	0.142	(0.021)	0.126	(0.044)
35 years	-0.212	(0.470)	0.00906	(0.761)	0.00360	(0.907)
45 years	-0.150	(0.004)	-0.124	(0.018)	-0.119	(0.026)
			gradients of *other controls*			
days w/o change	-0.000189	(0.003)	-0.000205	(0.001)	-0.000214	(0.001)
excess workload	0.0191	(0.000)	0.0189	(0.000)	0.0154	(0.000)
share of externals	-0.0259	(0.107)	-0.000364	(0.983)	0.0203	(0.251)
team size	-0.0110	(0.000)	-0.00571	(0.000)	-0.00629	(0.000)
early shift	-0.0111	(0.000)	-0.00936	(0.002)	-0.00520	(0.104)
schooling years	0.0195	(0.000)	0.0103	(0.013)	0.0106	(0.013)

Gradients and are calculated from coefficients in Table 1. Unless otherwise indicated, gradients are calculated at sample means. $p-$values in brackets.

C Sample Selection

To be written.