Occupational Mobility: Specialization and Diversification over the Life-Cycle

Christina Gathmann
Stanford University

Uta Schoenberg
University of Rochester

April, 2005

Very Preliminary and Incomplete

Abstract

The previous literature on career choice has either assumed that worker productivity in a job is match-specific and thus independent across jobs or has restricted mobility to be along a one-dimensional job ladder. In this paper, we combine elements of the two approaches to model horizontal occupational mobility and movements up an occupational ladder. To reduce the dimensionality of career choice, we define occupations as a combination of a small number of tasks. Occupations differ in the weight they attach to each task and how sensitive productivity in the occupation is to tasks. Occupational mobility arises in our model from the accumulation of occupation-specific capital, that is partially transferable across occupations and changes in demand conditions, which we model as changes in the weights occupations attach to tasks. Our empirical analysis uses a unique administrative dataset from Germany to test the empirical implications of our model. Preliminary evidence suggests that the assumption of pure match-specific productivity is rejected in our data. While we find evidence for a job ladder (vertical occupational mobility), most occupational mobility occurs to occupations with similar tasks. Current work focuses on empirical tests of increasing specialization or diversification over the life-cycle and its implications for wage growth.

Correspondence: Christina Gathmann, Stanford Unviersity, cgathman@stanford.edu; Uta Schoenberg, University of Rochester, utas@troi.cc.rochester.edu. We are responsible for all errors.

1 Introduction

How do individuals choose their labor market career over the life-cycle? While there is large literature on educational choice and the returns to schooling, the literature on career choice is much more limited and dispersed.

Modelling the career choices of individuals is complicated by the fact that there are a large number of choices available and decisions are interdependent over time. In principle, the most general setup would be a Roy model. Individuals have heterogenous skill endowments (ex-ante heterogeneity) and there are no costs associated with switching occupations or jobs. In the static setup, workers go where their given skills are rewarded highest. Productivity in one occupation may be correlated with productivity in another in an abitrary form. As prices of skills in occupations or sectors change (or as the consequence of other exogenous shocks), individuals change sectors. Thus, occupational mobility is in response to exogenous shocks and not a basic feature of the decision problem. This model becomes quickly untractable as the number of occupations.increases (Heckman and Sedlacek, 1985; 1990; Gould, 2002; Keane and Wolpin, 1998).[1]

The previous literature has followed two approaches to further reduce the dimensionality of the problem.[2] The first approach assumes that productivity in one occupation is uncorrelated with productivity at another occupation. Productivity is match-specific or a combination of the match and accumulated capital (general or specific). Learning about the quality of the match is therefore independent across jobs and occupations. That is, tenure on the last job does only reveal that the last match was good (which reduces probability of separation and thus the probability of finding an even better match) but does not say anything about the quality of the current match. This substantially reduces the dimensionality of the problem. But it also makes it less interesting because workers do not learn anything about their productivity across similar jobs or occupations.[3] The optimal policy has a reservation wage property. Mobility is an equilibrium outcome of optimizing process as information about match productivity accumulates and occurs independently of

[1] Heckman and Sedlacek (1985) translate the Roy model (Roy, 1951) within a market setting where individuals choose between two (three) alternatives (occupations, industries). Gould (2002) is a version of Heckman and Sedlacek to three sectors within a parametric setting. The model allows for both sorting according to absolute (skills are similarly rewarded across sectors) and comparative advantage (skills are paid different prices across sectors). Dahl (2002) extends the Roy model to polychotomous choices using semiparametric estimation. This relaxes the restrictive normality assumptions. The focus of the analysis has thus far been on characteristics of the sorting equilibrium and not on the mobility decision or frequency and patterns of mobility.

[2] A third approach has been to limit the choice set to only two or three occupations. For example, Siow (1984), Rosen (1992) and Zarkin (1985) analyze the investment decision between law school (teacher certification) and entering the labor market instead. Sauer (1998) and Stinebrickner (2001) study occupational change out of market for lawyers and teachers respectively as a function of outside opportunities. These models have a structure similar to models of educational investment.

[3] McCall (1990) is an exception.

external shocks to productivity.[4] Occupations cannot be ranked and there is no job ladder.[5]

The second perspective orders occupations vertically (like products with different qualities) and individuals sort according to absolute advantage. In this perspective, occupations form a job ladder where individuals start at the bottom and then move from simple to more complex jobs (Gibbons and Katz, 1992). These models stress learning-by-doing and accumulation of specific skills, which are (partially or fully) transferable across rungs of the ladder. An early example where movement across jobs (without explicit reference to occupations) is associated with different learning intensities is Rosen (1972). Because of investment in skills, models are dynamic. These models share with the older human capital approach the focus on the accumulation of skills in the labor market and its associated wage growth.[6] Jovanovic and Nyarko (1997) combine horizontal and vertical occupational mobility where low-level jobs are 'training grounds' for complex ones.[7]

In this paper, we combine features of both approaches into a general framework of career choice. To reduce the dimensionality inherent in occupational choice, we model occupations as a combination of a (smaller) number of tasks. Occupations differ in the weights they attach to a vector of tasks.[8] Two aspects are important: first the relative importance of one task versus another task (e.g. tasks that require manual skills and tasks that require cognitive skills). Second, occupations also differ in the skill requirements for each task (e.g. consider a bookkeeper and a computer programmer. Both perform mostly cognitive tasks, but the former needs less skills than the latter). Combining the two makes the setup general enough to allow for both horizontal occupational mobility and an occupational ladder (vertical occupational mobility).

The model is dynamic to allow for occupational mobility. Workers choose how much time to spend

[4] If a job is a pure experience good, all wage growth is within jobs due to learning about the match quality. If the job is a pure inspection or search good in contrast, all wage growth occurs between jobs as workers exploit better outside alternatives.

[5] Empirical work by Neal (1995) and Parent suggest that that industry-specific human capital is important. Kombourov et al. (2002) and Pavan (2004) argue that human capital is occupation-specific. However, all studies assume that industry- and occupation-specific producitivy is match-specific - i.e. workers' productivity at one occupation is uncorrelated with his productivity at another occupation.

[6] One of the most interesting examples explicitly analyzes occupational mobility decisions from the worker's perspective is Sicherman and Galor (1990).

[7] One example that combines features of both job search and a Roy-type model is Miller (1984). He extends the job matching framework to allow for multiple (lateral) types of jobs. Jobs belong to different occupations if agent's belief about the match-specific components of productivity are different; otherwise, they belong to the same occupation. By working in a particular job, the individual learns about his job-specific (not occupation-specific) productivity only. In other words, working in a job in an occupational group does not reveal anything about the other jobs in the same occupation. This is a consequence of productivity defined purely in terms of job-specific components. His economic predictions include that labor market entrants will be found in jobs with high informational value about individual productivity (similar to Rosen's model (1972) where some jobs have high learning content). In these jobs, only few successful experienced workers will be found, who learned that their maximum productivity is in this job (for example, the arts). In contrast, jobs with fewer information exhibit less turnover, higher returns to recent entrants on average and a greater concentration of wages about the mean. The theoretical decision problem is not solvable because in general the decision rule depends on characteristics of all jobs. To derive the decision rule for this type of a multi-armed bandit decision problem, he uses the Gittins index, which reduces the choice problem to comparing one-dimensional indices.

[8] In this sense, our approach is similar to characteristics or hedonic models, in which goods are a combination of characteristics and prices for these characteristics.

on each task, which determines their occupational choice in each period. Occupational mobility arises in the model for three potential reasons: first, workers accumulate human capital in each task when working in an occupation that uses these tasks. This is similar to the stepping stone model of Jovanovic and Nyarko (1997).Occupational mobility is planned from the start; workers choose a particular occupation when young because it prepares them for a particular occupation when old; workers should move up the 'occupational ladder' within each task. Second, if weights (or equivalently prices) for tasks change because of new technology or demand conditions.[9] Finally, if workers in addition learn about their productivity in tasks while using them in an occupation. Productivities in each task would then follow a martingale (see Appendix B for an outline of a model with learning).

Two possible extensions we will pursue in the future is to include the employment margin. If employees with unemployment or nonemployment spells differ from continuously employed individuals, our results are only valid for the subset of workers without career disruptions. Incorporating the employment is important for another reason: it could be that people with intermediate nonworking spells are more likely to experience downward occupational mobility over their career and associated declines in real wages. A second extension is to incorporate firm decisions about hiring and firing into the model. This brings us closer to an equilibrium model of the labor market where heterogeneous workers are matched to heterogeneous firms.

Our empirical analysis tests empirical predictions of our model with a unique administrative dataset from the Bundesanstalt fuer Arbeit in Germany that combines detailed information about occupational positions and mobility both within and between firms. The data have a number of advantages: first, there is practically no measurement error in reported wages and occupational codes, both of which are serious problems in household surveys like the PSID or NLSY used in prior studies. In addition, we have much larger samples than what is typically available in survey data. Further, the data contains more detailed individual information including educational degree, nationality and other characteristics that are often missing in administrative datasets from other countries.

The empirical results show that there is evidence of a ladder in occupational mobility. There is however also substantial mobility between occupations with similar skill requirements (which we call horizontal mobility). Our data also rejects a pure matching model, which assumes that all occupational capital is lost

[9] Autor et al. (2003) show that there has been a shift in the demand for tasks over the past decades, with non-routine cognitive skills becoming increasingly important. See Spitz (2004) for evidence for Germany.

when workers switch occupations. Wage growth is highest for those changing employers and occupations (7 percent) while the wage growth of employer change within same occupation and of occupational changes within the same employer are similar (5 percent). In contrast, wage growth within the same occupation and remaining at the same employer is not statistically different from zero.

... to be completed ...

The structure of the paper is as follows. The next section outlines a simple model of sorting into occupations and occupational mobility over the life-cycle. The third section introduces our dataset and provides institutional background on labor market institutions and wage setting in Germany. In Section 4, we analyze occupational mobility patterns. Section 5 decomposes wage growth over the life-cycle. In Section 6 ,we provide additional results to show the robustness of the estimates. Finally, Section 7 discusses potential implications of the results and concludes.

2 Theoretical Framework

2.1 Choice between Occupations

There are two types of skill, A and B. Occupations differ with respect to the importance of each skill. There are occupations in which mostly one skill is used, and occupations in which mostly the other skill is used. Workers choose the occupation that maximizes life-time wages. The labor market is perfectly competitive. Every period, workers are paid a wage equal to their expected productivity.

We assume that the marginal return to performing a task is decreasing. Hence, if a worker has worked 5 hours on task A, and 0 hours on task B, the worker will be more productive in task B, even if he generally is equally good at both tasks. The following production function captures the idea of decreasing marginal returns.

$$y = \sqrt{pA} + \sqrt{(1-p)B}.$$

There are two periods. Let A_1 and B_1 denote workers' (expected) productivity at task A and B at the end of the first period. At the beginning of the second period, workers maximize with respect to p :

$$\max_{p_2} \quad \sqrt{p_2 A_1} + \sqrt{(1-p_2)B_1}$$

The first order condition is

$$\begin{aligned}\frac{A_1}{2\sqrt{p_2 A_1}} &= \frac{B_1}{2\sqrt{(1-p_2)B_1}}, \text{ or}\\ p_2 &= \frac{A_1}{A_1+B_1}.\end{aligned}$$

The first order condition is sufficient for a maximum. If a worker is equally good at both tasks, he spends the same amount of time at both tasks. If $A_1 > B_1$, he spends more time on task A, while he spends more time on task B if $A_1 < B_1$. Workers' utility or wage in the second period can be computed as

$$w_2 = \sqrt{\frac{A_1^2}{A_1+B_1}} + \sqrt{\frac{B_1^2}{A_1+B_1}} = (A_1+B_1)\sqrt{\frac{1}{A_1+B_1}} = \sqrt{\frac{(A_1+B_1)^2}{A_1+B_1}} = \sqrt{A_1+B_1}.$$

What about occupational choice in the first period? Of course, workers anticipate their optimal occupational choice in the second period. Let A_0 and B_0 denote workers' expected ability at task A and B at the beginning of the first period. Workers maximize

$$\max_{p_1} \sqrt{p_1 A_0} + \sqrt{(1-p_1)B_0} + E[\sqrt{A_1+B_1}|A_0, B_0].$$

We need to make assumptions on how A_1 and B_1 depend on A_0 and B_0. Suppose

$$\begin{aligned}A_1 &= A_0 + \sqrt{p_1 H} + \varepsilon_A,\\ B_1 &= B_0 + \sqrt{(1-p_1)H} + \varepsilon_B.\end{aligned}$$

This specification first implies that workers accumulate more human capital of type A or B the more time they spend performing the task. Human capital accumulation is concave, meaning that workers accumulate the most human capital when $p = 0.5$ (i.e. $\sqrt{p_1 H} + \sqrt{(1-p_1)H}$ is maximized if $p = 0.5$. This does not mean that all skills workers acquire will be useful at the occupation they choose in the second period. Note that human capital accumulation is independent of workers' skill - all workers accumulate the same amount of human capital of a particular type if they spend the same amount of time at the task. This specification further implies that occupational choice in the first period has no impact on how much workers learn about their productivity during the first period. We can think of the productivity shock as a macroeconomic shock - sometimes one task is in higher demand, sometimes the other. Later, we also consider models in which workers learn about their true productivity in each task in the first period, and they learn the more about the task the more time they spend at the task.

Suppose that ε_A and ε_B are both normally distributed with mean 0 and variance σ^2. ε_A and ε_B are independent. Hence, $\varepsilon_A + \varepsilon_B$ is normally distributed with mean 0 and variance $2\sigma^2$. Furthermore ε_A and ε_B are independent of A_0 and B_0, respectively. (Note that since the support of the normal distribution is unbounded, A_1 and B_1 may be negative. The objective function, however, is only defined for A_1 and $B_1 > 0$. We assume that A_1 and $B_1 < 0$ happens with very low probability, and in case it happens, A_1 and $B_1 = 0$). These assumptions imply that

$$\begin{aligned} E[\sqrt{A_1 + B_1}|A_0,B_0] &= E[\sqrt{A_0 + \sqrt{p_1 H} + \varepsilon_A + B_0 + \sqrt{(1-p_1)H} + \varepsilon_B}|A_0, B_0] \\ &= \int_{-\infty}^{\infty} \sqrt{\left|A_0 + \sqrt{p_1 H} + \varepsilon_A + B_0 + \sqrt{(1-p_1)H} + \varepsilon_B\right|} \cdot \\ &\quad \frac{1}{\sqrt{4\pi\sigma^2}} \exp(-0.5\frac{(\varepsilon_A + \varepsilon_B)^2}{2\sigma^2}) d(\varepsilon_A + \varepsilon_B). \end{aligned}$$

The first order condition becomes

$$\begin{aligned} &\frac{A_0}{2\sqrt{p_1 A_0}} - \frac{B_0}{2\sqrt{(1-p_1)B_0}} + \int_{-\infty}^{\infty} \frac{1}{2\sqrt{\left|A_0 + \sqrt{p_1 H} + \varepsilon_A + B_0 + \sqrt{(1-p_1)H} + \varepsilon_B\right|}} \cdot \\ &\quad \frac{1}{\sqrt{4\pi\sigma^2}} \exp(-0.5\frac{(\varepsilon_A + \varepsilon_B)^2}{2\sigma^2}) d(\varepsilon_A + \varepsilon_B)(\frac{H}{2\sqrt{p_1 H}} - \frac{H}{2\sqrt{(1-p_1)H}}) = 0. \end{aligned}$$

Let me first point out that if I had specified human capital accumulation as $p_1 H + (1-p_1)H$, instead of$\sqrt{p_1 H} + B_0 + \sqrt{(1-p_1)H}$, the rule for the optimal occupational choice in the first period would be the same as in the second period, i.e. $p_1 = \frac{A_0}{A_0+B_0}$. This is because under this specification, workers accumulate the same amount of human capital, independently of the choice of p in the first period. This rule is no longer optimal under our specification of human capital accumulation. Workers with $A_0 = B_0$ still choose $p = 0.5$, and thus fully diversify. Workers with $A_0 > B_0$ choose $0.5 < p_1 < \frac{A_0}{A_0+B_0}$, while workers with $A_0 < B_0$ choose $\frac{A_0}{A_0+B_0} < p_1 < 0.5$. Hence, compared to the productivity-maximizing occupational choice, workers choose an occupation closer to $p = 0.5$. In the first period, workers specialize less, and diversify more. To see this, note that the second term is positive if $p < 0.5$, zero if $p = 0.5$, and negative if $p > 0.5$. Suppose $A_0 > B_0$. The first term is zero if $p = \frac{A_0}{A_0+B_0}$, i.e. $p > 0.5$. But then the second term is negative. Hence, for the whole term to be zero, we require p to be smaller than $\frac{A_0}{A_0+B_0}$ and closer to 0.5. The same argument applies if $A_0 < B_0$. The first term is 0 if $p = \frac{A_0}{A_0+B_0}$, i.e. $p < 0.5$. But then the second term is positive. Hence, for the whole term to be zero, we require p to be greater than $\frac{A_0}{A_0+B_0}$ and closer to 0.5. The intuition for this result is clear. Workers accumulate most human capital if they spend the same amount of time at both tasks. This motive is absent in the second period. Consequently, workers specialize more in the second than in the first period.

Moreover, the second term is the more important the higher H and the higher σ^2, the shock to tasks A and B. This should imply that in the first period, workers diversify the more the higher human capital acquisition and the greater the productivity shock. However, the second term does not seem to depend to much on the variance of productivity shocks. Here are some simple computations:

$\int_{-\infty}^{\infty} \frac{1}{2\sqrt{|30+\sqrt{0.2}+\sqrt{0.8}+x|}} \frac{1}{\sqrt{12\pi}} \exp(-0.5\frac{x^2}{6})dx =$

3: 8.9519×10^{-2}

2: 8.9449×10^{-2}

1: 0.08938

0.5: 8.9346×10^{-2}

0.25: 8.9329×10^{-2}

The result that workers specialize more the older they get seems very intuitive. However, this result is partly driven by functional form assumptions. If we instead assume that human capital and skills are complements, we can easily get the opposite result - workers specialize more in the first period. SupposeA_1 and B_1 are almost always positive and equal

$$
\begin{aligned}
A_1 &= A_0(1+\sqrt{p_1 H}) + \varepsilon_A, \\
B_1 &= B_0(1+\sqrt{(1-p_1)H}) + \varepsilon_B.
\end{aligned}
$$

Then the first order condition for occupational choice in the first period becomes:

$$\frac{A_0}{2\sqrt{p_1 A_0}} - \frac{B_0}{2\sqrt{(1-p_1)B_0}} + \int_{-\infty}^{\infty} \frac{1}{2\sqrt{\left|A_0(1+\sqrt{p_1 H}) + \varepsilon_A + B_0(1+\sqrt{(1-p_1)H}) + \varepsilon_B\right|}} \cdot$$
$$\frac{1}{\sqrt{4\pi\sigma^2}} \exp(-0.5\frac{(\varepsilon_A+\varepsilon_B)^2}{2\sigma^2}) d(\varepsilon_A+\varepsilon_B)(\frac{HA_0}{2\sqrt{p_1 H}} - \frac{HB_0}{2\sqrt{(1-p_1)H}}) = 0.$$

Set $p_1 = \frac{A_0}{A_0+B_0}$ so that the first term is 0. Suppose $A_0 > B_0$. Then the second term (i.e. $\frac{HA_0}{2\sqrt{p_1 H}} - \frac{HB_0}{2\sqrt{(1-p_1)H}}$) is positive. Hence, workers choose an occupation with $p > \frac{A_0}{A_0+B_0}$, and thus specialize more in the first period. The same argument applies if $A_0 < B_0$. This effect is the stronger the more skilled workers are. Hence, skilled workers should choose more specialized occupations. (Derek Neal has a paper in the JHR where he makes this point too, in a very different model).

An alternative specification is $A_1 = A_0(1+pH) + \varepsilon_A$, and $B_1 = B_0(1+(1-p)H) + \varepsilon_A$. The first order condition becomes

$$\frac{A_0}{2\sqrt{p_1 A_0}} - \frac{B_0}{2\sqrt{(1-p_1)B_0}} + \int_{-\infty}^{\infty} \frac{1}{2\sqrt{|A_0(1+p_1 H) + \varepsilon_A + B_0(1+(1-p_1)H) + \varepsilon_B|}} \cdot$$
$$\frac{1}{\sqrt{4\pi\sigma^2}} \exp(-0.5\frac{(\varepsilon_A+\varepsilon_B)^2}{2\sigma^2}) d(\varepsilon_A+\varepsilon_B) H(A_0 - B_0) = 0.$$

If $p_1 = \frac{A_0}{A_0+B_0}$, the first term is positive. Suppose $A_0 > B_0$ so that $p > 0.5$. Then the second term is positive. Hence, the optimal p will be greater than $\frac{A_0}{A_0+B_0}$, and as in the previous case, workers specialize more in the first than in the second period. In fact, these results are also fairly intuitive. If skills and human capital accumulation are complements, then the return to human capital acquisition is higher at the task which the worker performs better. Hence, the worker should optimally choose an occupation that gives more weight to this task.

3 Adding a Job Ladder

Suppose occupations do not only differ with respect to how much time workers spend on each of the two tasks, but also with respect to how valuable the tasks are in each occupation:

$$
\begin{aligned}
y_o &= a_o + b_o\sqrt{p_o A} + a_o + b_o\sqrt{(1-p_o)B} \\
&= 2a_o + b_o(\sqrt{p_o A} + \sqrt{(1-p_o)B})
\end{aligned}
$$

As in Gibbons and Waldman (1999), the higher b_0, the lower a_0. Hence, occupations with a high b are occupations in which output is particularly sensitive to skills. These occupations may be thought of as on the top of the 'occupational ladder'. If we additionally assume that productivity shocks or learning do not depend on b_o, adding a job ladder model to the model outlined above does not lead to many complications. This is because a_o and b_o have no impact on the optimal choice of p in the first and second period. Hence, the optimal occupational choice follows a simple 2-step rule:

1. First, workers choose the optimal p_o.

2. Second, workers choose the optimal a_o and b_o.

More skilled workers will sort into jobs with a higher b_o, and a lower $a_{o.}$ The model potentially captures two important aspects about occupations. First, the horizontal aspect: Occupations differ with respect to the relative importance of one skill versus another. This makes skills occupation-specific. Second, the vertical aspect: Some occupations are higher up the 'occupational ladder' than others.

The model is a lot more complicated if the choice of b_o affects how much workers learn about their true productivity. Suppose that at the end of the first period, firms and workers observe $a_o + b_o\sqrt{p_o A} + \varepsilon_A$, and $a_o + b_o\sqrt{p_o B} + \varepsilon_B$. Because of the non-linearity, it is not certain that there is an easy formulae firms and workers can use in order to update their beliefs of their true productivity. Moreover, then we should have interdependencies between p and b - which will be a lot harder to analyze.

3.1 An Example with Linear Technology

Six occupations should be enough to highlight the main features of the model. The production functions are a special case of **??**. Suppose there are two occupations that specialize in task A. One occupation requires more skills than the other. Let the production functions be

$$
\begin{aligned}
y_1 &= a_1 + b_1 T_A, \text{ and} \\
y_2 &= a_2 + b_2 T_A.
\end{aligned}
$$

Let $a_1 > a_2$, and $b_2 > b_1$. If workers have to choose between these two occupations, they prefer occuapation 1 if $T_A < T^*$, and occupation 2 if $T_A \geqslant T^*$, where T^* satisfies $T^* = \frac{a_1 - a_2}{b_2 - b_1}$. Hence, occupation 2 is the skilled occupation, and occupation 1 is the unskilled occupation.

Similarly, there are two occupations that specialize in task B. One occupation requires more skills than the other. Let the production functions be

$$
\begin{aligned}
y_3 &= a_1 + b_1 T_B, \text{ and} \\
y_4 &= a_2 + b_2 T_B.
\end{aligned}
$$

If workers have to choose between these two occupations, they prefer occuapation 3 if $T_B < T^*$, and occupation 4 if $T_B \geqslant T^*$, where T^* satisfies $T^* = \frac{a_1 - a_2}{b_2 - b_1}$. Occupation 4 is the skilled occupation, and occupation 3 is the unskilled occupation.

Finally, there are two occupations where workers perform both tasks. Again, one occupation requires more skills than the other. The production functions are

$$
\begin{aligned}
y_5 &= a_1 + b_1(\lambda_o T_A + (1 - \lambda_o) T_B), \text{ and} \\
y_6 &= a_2 + b_2(\lambda_o T_A + (1 - \lambda_o) T_B).
\end{aligned}
$$

Let $\widetilde{T}$ denote $\widetilde{T} = \lambda_o T_A + (1 - \lambda_o) T_B$. If workers have to choose between these two occupations, they prefer occuapation 5 if $\widetilde{T} < \widetilde{T}^*$, and occupation 6 if $\widetilde{T} \geqslant \widetilde{T}^*$, where $\widetilde{T}^*$ satisfies $\widetilde{T}^* = \frac{a_1 - a_2}{b_2 - b_1}$.

Suppose there is only one period. Wages are equal to productivity. Workers have no preference for variety, and only care about income. Which occupation do workers choose? It is easy to see that no worker chooses occupations 5 and 6. They choose occupations 1 or 2 if $T_A > T_B$. They select occupation 1 if $T_A < T^*$, and occupation 2 if $T_A \geqslant T^*$. They choose occupations 3 or 4 if $T_A < T_B$. They select occupation 3 if $T_B < T^*$, and occupation 4 if $T_B \geqslant T^*$. Figure **??** illustrates workers' occupational choice.

First conclusion: 'General' occupations only survive in a 1-period model if

- workers have a preference for variety
- there are decreasing marginal returns in performing a task; in this case production functions are not linear in task-specific skills.

In a multi-period model, there are additional reasons for why general occupations exist:

- workers accumulate human capital that is not transferable across tasks *and* there are shocks to productivity
- workers learn about their task-specific productivity only if they have a chance to perform a task

Let's look at a simple two-period model. Suppose the evolution of task-specific skills is as follows.

$$T_A^2 = \begin{cases} T_A^1 \text{ if } o_1 = 3, 4; \\ T_A^1 + h_A + s_A \text{ if } o_1 = 1, 2; \\ T_A^1 + h_A + \lambda_o s_A \text{ if } o_1 = 5, 6. \end{cases}$$

If in period 1 a worker worked in occupation 3 or 4 - where task A is not performed-, his skills at task A in period 2 are the same as they were in period 1. If, on the other hand, in period 1 he worked in occupation 1 or 2 -where only task A is perfomed-, he accumulates task-specific human capital of amount s_A and in addition he learns how good he is at performing task A. The learning aspect is captured by the productivity shock h_A, with $E[h_A] = 0$. Workers only learn about their true productivity at a task if they experience it. Finally, if a worker started at occupations 5 or 6 where λ_o percent of tasks performed are of type A, he again learns his true procutivity at task A. He also accumulates task-specific skills, but only of amount $\lambda_o s_A$. The evolution of skills in task B is analoguous.

Some workers now have an incentive to choose occupations 5 or 6 in the first period, for two related reasons.

1. Choosing occupations 5 and 6 allows them to learn their true productivity at both tasks.

2. Choosing occupations 5 and 6 allows them to accumulate human capital at both task.

Not all workers want to choose occupations 5 and 6, since in period 2 they will work for a specialized job, and they accumulate more task-specific human capital at specialized jobs.

This is my conjecture about occupational choice in the first period. Workers who are about equally good at both tasks choose to work in one of the general occupations. Among those, the more talented ones choose occupation 6. Workers who are clearly better at one task than at the other choose to work in the specialized occupations. Among those, the more talented workers choose the more skilled occupation. Figure **??** illustrates this.

The model should yield a number of interesting implications, such as

- occupations 5 and 6 are 'training grounds'; here the proportion of young workers should be higher; workers then move to more specialized occupations

- there is also a 'job ladder' from occupation 1 to 2, and from occupation 3 to 4, i.e. more workers move from occupation 1 to2 and 3 to4 than vice versa

- There are more movements between occupations 1 and 2 and 3 and 4 than there are from occupations 1 and 2 to 3 and 4, and vice versa

- The model should also yield predictions concerning wages of movers vs. stayers as well as wage gains.

4 Data and Institutional Background

4.1 Data

The dataset is drawn from adminstrative social security records between 1975 and 1995 [2001].[10] In January of 1973, a unified health, retirement and unemployment insurance system was introduced in Germany. The Employee Sample (*Beschaeftigtenstichprobe*) of the Institute for Employement Research is a one percent sample of all employees insured in the German social security system. This covers more than 80 percent

[10] See Appendix A for more details on the data and the construction of our sample used in the empirical analysis.

of all employed individuals in Germany. Excluded are the self-employed, persons employed in the military, public servants and jobs that are exempt from social security contributions because of low working hours. The sample size is large with each year containing around 200,000 observations.

Each employer has to report the exact beginning and end of any employment relation (new hires and employees leaving the firm) subject to social security contributions. In addition, employers provide information about all employees at the end of each year. The data has several advantages over commonly used household surveys or adminstrative datasets from other countries. Unlike individual survey data, we observe the exact date of a job change as well as the wage associated with each job. Another major advantage is that gross earnings are measured without error as they form the basis for calculating social security contributions. Misreporting is subject to heavy penalties. Also, since occupations are the basis for wage bargaining between unions and employer, there is a high degree of consistency and accuracy of occupational codes across firms. In contrast, occupational codes in datasets like the NLSY are highly subjective.

In contrast to administrative datasets from other countries, the data reports detailed individual characteristics of the employee including educational level, vocational degree, nationality, marital status and number of children. Like most datasets of this nature, earnings are right censored because of an upper ceiling to social security payments. In the empirical analysis below, we use tobit models or quantile regression to account for the censoring. Another unique feature of the data is that we have information about the type of interruption if the individual is not employed. In particular, the data distinguishes between maternity leave, publicly funded training programs and unemployment spells.[11] Finally, we combine our information on individual employees with plant-level data about firm size, the educational structure of the workforce and detailed industry classification and region of operation.

We restrict our sample to West German men and women aged 45 years or younger in 1975 and add all individuals entering the sample in later years. We require that individuals are employed at least for 5 five consecutive years during the sample period. Since labor force attachment, job mobility patterns and occupational choice are very different for men and women, we run separate analyses for men and women throughout the paper. Table X contains descriptive statistics for men and women over the whole sample period.

[11] Note that only spells for individuals who are eligible for any of the benfits mentioned in the text are observed in the data. In addition, public pay measures (*Arbeitsbeschaffungsmassnahmen*) are not contained in the data.

4.2 Horizontal and Vertical Occupational Mobility: Measurement

In order to relate occupations to each other, we use the average number and types of tasks employed in each occupation. The data on tasks performed in each occupation is taken from BIBB (data on "Evolution von Qualifikationsanforderungen" of the Bundesinstitut für Berufsbildung).

To measure the degree of specialization within each occupation, we use three different definitions. According to the first definition, the specialization measure is equal to the average number of tasks performed in an occupation. No distinction is made between different types of tasks. Here is an example. Consider two workers who both perform two tasks. For the first worker, the two tasks are: *bauen* and *anbauen*, i.e. two very similar tasks. For the second worker, the two tasks are: *bauen* and *forschen*, i.e. two rather different tasks. For both workers, the specialization measure is 2. At the individual level, this measure can take values between 1 and 20, i.e. there are 20 different tasks. At the occupation level, the minimum (average) number of tasks performed is 1 (occupation 9), the maximum is 4.5 (occupation 92). For definition 2, I divided each task into three types: manual, analytic, interactive. According to this definition, the specialization measure is equal to the average number of types of tasks performed in an occupation. I therefore ignore whether a worker performs more than one task of a particular type. At the individual level, this measure can take values from 1 to 3. At the occupation level, the occupation with the lowest index is occupation 9 (1), the occupation with the highest index is 86 (2.12). Definition 3 is a mixture of these two definitions. Tasks of a different type are given twice the weight as tasks within the same type. At the occupational level, this measure varies from 1 (occupation 9) to 3.23 (occupation 75). All specialization measures are strongly postively correlated. The correlation coefficient depends on how many observations I use per occupation. It is at least 0.8, no matter which weights I use.

To measure the horizontal distance between occupations, we use two definitions. The first one takes into account all 20 tasks, and makes no distinction between different types of tasks. Consider two occupations, A and B. Let p_{At} and p_{Bt} denote the proportion of workers who perform task t in occupation A and B. The first distance measure is:

$$d_1 = \sum_{i=1}^{20} (p_{At} - p_{Bt})^2.$$

The greater this number, the more different are occupations A and B. In principle, this distance measure could be as high as 20. In practice, the maximum distance observed between two occupations is 3.20 (between occupations 36 and 69, Textilausruester and Bankfachleute).

For the second distant measure, we divide the 20 tasks into three groups, manual, analytical, and interactive. Let q_{At} and q_{Bt} denote the proportion of workers who perform a task of type t in occupation A and B. The second distance measure is:

$$d_2 = \sum_{i=1}^{3} (q_{At} - q_{Bt})^2.$$

This measure ignores that workers in different occupations may perform different tasks of the same type. In principle, this distance measure could be as high as 3. The maximum distance observed is 2.14 (between occupations 11 and 69, Betonhersteller and Bankfachleute).

Finally, we need to rank occupations vertically. For now, we use the average wage in each occupation, using data on all employed male workers from 1975 to 1995, to rank occupations. In order to deal with censoring, we estimated a censored regression. The occupation that pays the lowest wage is ranked 1, the occupation that pays the highest wage is ranked 77. When we look separately at the three education groups, we run censored regressions separetely for each education group. Results are very similar if we just use the same ranking for all education groups. Rankings are strongly positively correlated across education groups.

An alternative approach to define vertical distance between two occupations is to use the *absolute* difference between the average wage that are paid in the two occupations (i.e. if the average wage in occupation A is 10, in the average wage in occupation B is 12, the vertical distance between these two occupations is 2).

5 Empirical Results on Occupational Mobility and Sorting

5.1 Occupational Mobility

Table 1 shows the proportion of workers switching occupations at least once, by actual experience. We assume that the occupational switch occured before or at the beginning of the second spell. We also report the number of observations for each education group and for each experience level. It is evident that better educated workers are less likely to switch occupations. Occupational mobility declines with experience for all education groups.

Table 2 reports results from several models of occupational mobility. The dependent variable is equal to 1 if the worker switched occupations, and 0 otherwise. Unconditional on any other variable, better educated

Table 1: Occupational mobility by experience

experience	unskilled	apprentices	university graduates
hline 1	0.4188	0.2528	0.2264
	5721	*23994*	*5530*
2	0.3153	0.1844	0.1386
	5385	*22349*	*5065*
3	0.2759	0.1537	0.1230
	4704	*19966*	*4504*
4	0.2322	0.1351	0.1145
	4164	*17675*	*3922*
5	0.1999	0.1149	0.1067
	3682	*15395*	*3345*
6	0.1675	0.1056	0.0976
	3260	*13169*	*2776*
7	0.1546	0.0988	0.0846
	2872	*11069*	*2292*
8	0.1280	0.0897	0.0799
	2532	9136	1877
9	0.1284	0.0792	0.0735
	2228	*7206*	*1538*
10	0.1207	0.0783	0.0697
	1898	5592	1247

workers are less likely to switch occupations. Differences between unskilled workers and apprentices are stronger than between apprentices and university graduates. Controlling for other variables, such as wages, experience, and occupation tenure, the difference between apprentices and university graduates disappears. University graduates and apprentices continue to be less likely to switch occupations. As expected, workers earning higher wages are less likely to switch occupations. Unconditional on occupation tenure, experience has a negative impact on the probability that a worker switches occupations. The impact of experience on occupational mobility declines if we control for wages. If we also control for occupation tenure, the impact of experience switches signs: Conditional on occupation tenure, more experienced workers are more likely to switch occupations. Conditional (and unconditional) and experience, workers with more occupation tenure are less likely to switch occupations. These results are robust across education groups. Interestingly, wages seem to have a stronger deterrent effect on occupational mobility for better educated workers.

Table **??** shows how occupational mobility in the past affects current occupational mobility. For each experience level, workers who have switched occupations more often in the past, are more likely to switch occupations this period. Clearly, the number of previous occupational switches has a strong *positive* impact on the probability that a worker switches occupations this period. The table also reports the wage at labor market entry as well as the current wage by number of previous occupational switches, separately for each

Table 2: Determinants of occupational mobility

	1	2	3	4	5
all					
apprentice	-0.1456 (0.0019)	-0.1016 (0.0019)	-0.1610 (0.0019)	-0.1288 (0.0020)	-0.0721 (0.0020)
university	-0.1647 (0.0018)	-0.0835 (0.0023)	-0.1714 (0.0017)	-0.1206 (0.0021)	-0.0732 (0.0022_
log-wage		-0.1982 (0.0017)		-0.1279 (0.0019)	-0.0965 (0.0018)
exp			-0.0420 (0.0006)	-0.0295 (0.0007)	0.1412 (0.0010)
exp^2			0.0012 (0.0001)	0.0007 (0.0001)	-0.0097 (0.0001)
occten					-0.2824 (0.0011)
$occten^2$					0.0192 (0.0001)
unskilled					
log-wage		-0.1949 (0.0039)		-0.0709 (0.0044)	-0.0434 (0.0047)
exp			-0.0569 (0.0015)	-0.0481 (0.0016)	0.12910 (0.0024)
exp^2			0.0014 (0.0001)	0.0010 (0.0001)	-0.0085 (0.0002)
occten					-0.3403 (0.0027)
$occten^2$					0.0215 (0.0002)
apprentices					
log-wage		-0.1993 (0.0022)		-0.1364 (0.0024)	-0.0979 (0.0022)
exp			-0.0413 (0.0008)	-0.0294 (0.0009)	0.1573 (0.0013)
exp^2			0.0014 (0.0001)	0.0009 (0.0001)	-0.0117 (0.0001)
occten					-0.2922 (0.0015)
$occten^2$					0.0216 (0.0002)
university graduates					
log-wage		-0.1889 (0.0036)		-0.1589 (0.0041)	-0.1217 (0.0035)
exp			-0.0324 (0.0014)	-0.0127 (0.0015)	0.1053 (0.0020)
exp^2			0.0013 (0.0001)	0.0003 (0.0001)	-0.0062 (0.0002)
occten					-0.1936 (0.0023)
$occten^2$					0.0121 (0.0002)

experience level. Clearly, workers who end up switching occupations more often in the future earned lower wages at labor market entry. They also currently earn a lower wage.

If we break this analysis down by education, the sample size becomes quite small. Nevertheless, there are some interesting differences and similarities across education groups. We observe for all education groups that workers who have switched occupations more often in the past are more likely to switch occupations this period. For both apprentices and university graduates, the number of (future and past) job switches negatively affects both entry and current wages at all experience levels. The relationship between future mobility and entry wages is particularly strong for university graduates. Since university graduates experience a larger wage growth when switching occupations than apprentices, past occupational switches have about the same negative impact on current wages for university graduates and apprentices. The pattern is different for unskilled workers. At low experience levels, workers who switch occupations earned *higher* wages, both at labor market entry and after switching jobs. At higher experience levels, on the other hand, the number

Table 3: History of occupational mobility

			2nd year			
# switches	0	1				
# observations	23,861	8,938				
Prob. switch	14.76	33.55				
entry wage		-0.1827(0.0057)				
current wage		-0.0478 (0.0053)				
			3rd year			
# switches	0	1	2			
# observations	18,398	8,325	2,451			
Prob. switch	11.46	22.50	39.05			
entry wage		-0.1566 (0.0060)	-0.2517 (0.0094)			
current wage		-0.0267 (0.0052)	-0.1489 (0.0084)			
			4th year			
# switches	0	1	2	3		
# obsrvations	14,604	7,567	2,848	742		
Prob. switch	10.13	17.17	25.39	40.84		
entry wage		-0.1487 (0.0064)	-0.2239 (0.0090)	-0.3192 (0.0159)		
current wage		-0.0277 (0.0053)	-0.1077 (0.0076)	-0.2253 (0.0140)		
			5th year			
# switches	0	1	2	3	4	
# observations	11,572	6,739	2,929	962	220	
Prob. switch	7.81	14.30	20.83	30.04	43.18	
entry wage		-0.1466 (0.0068)	-0.2090 (0.0091)	-0.3105 (0.0145)	-0.3571 (0.0275)	
current wage		-0.0101 (0.0053)	-0.0838 (0.0072)	-0.1828 (0.0116)	-0.2881 (0.0236)	
			6th year			
# switches	0	1	2	3	4	5
# observations	9,203	5,765	2,807	1,054	312	64
Prob. switch	7.51	12.23	15.89	23.43	28.53	46.88
entry wage		-0.1348 (0.0074)	-0.1997 (0.0095)	-0.2829 (0.0140)	-0.3301 (0.0237)	-0.4323 (0.0486)
current wage		0.0033 (0.0057)	-0.0601 (0.0073)	-0.1367 (0.0111)	-0.2292 (0.0196)	-0.3514 (0.0467)

of (future and past) job switches negatively affects both entry and current wages. The impact of the number of job switches on wages is weaker for unskilled workers than for apprentices and university graduates.

5.2 Existence of a Job Ladder

Are workers more likely to move to an occupation in which wages are higher? Does wage growth of occupation switchers depend on whether a worker moved to a better or lower paying occupation? Table 4 shows that among all occupation swtichers, 56.9 % move to an occupation that pays a higher wage. The proportion is considerably higher for university graduates (65.4 %). Moreover, wage growth of workers who move up exceeds wage growth of movers who move down by at least 6 %. Allso note that wage growth of workers who switch occupations and move down is substantially larger than wage growth of workers who do not switch occupations.

Table 4: Occupational mobility and Ranking of Occupations

	all	unskilled	apprentices	university graduates
prop. moving up, conditional on moving	56.85 %	55.22 %	56.57 %	65.43 %
wage growth, moving up	0.1944	0.2251	0.1549	0.3564
	(0.0035)	(0.0076)	(0.0041)	(0.0107)
wage growth, moving down	0.1185	0.1152	0.0963	0.1708
	(0.0040)	(0.0084)	(0.0048)	(0.0120)
wage growth, stayers	0.0536	0.0609	0.0514	0.0551
	(0.0008)	(0.0020)	(0.0010)	(0.0013)

In order to further investigate the hypothesis whether workers 'move up the job ladder' over the life cycle, I computed the average ranking of occupations by experience. Table 5 reports results. For all education groups, the average ranking increases with experience. Hence, workers move to higher paying occupations over the life cycle. Also note that unskilled workers, apprentices, and university graudates tend to work at very different occupations - university graduates work at occupations that pay considerably higher wages.

Table 5: Ranking of Occupations by Experience

experience	all	unskilled	apprentices	university graduates
1	42.05	32.53	39.50	62.15
2	43.30	34.63	40.31	65.94
3	43.73	35.70	40.75	66.59
4	44.09	36.66	41.19	67.14
5	44.35	37.37	41.61	67.28
6	44.62	37.77	42.21	67.57
7	44.62	38.37	42.45	67.47
8	44.84	38.80	42.92	67.71
9	44.90	38.88	43.28	67.71
10	44.87	38.88	43.43	67.81

The next table (table 6) re-investigates the relationship between wage growth when switching from occupation A to occupation B, and the difference in the average wage between the two occupations (dbeta). (Hence, if the (log) average wage in occupation A is 4.5, and the (log) average wage in occupation B is 4.8, then dbeta is 0.3 when the worker switches from A to B, and -0.3 when the worker switches from B to A.) The coefficient on the difference between the average wage in the two occupations is positive, but less than 1. Hence, workers experience a higher wage growth if they move to an occupation which on average pays higher wages. However, the average wage growth of a worker who switches from occupation A to B is considerably smaller than the average wage difference between occupation A and B.

Next, I look at transition matrices (table 7 to 10). I divided workers into 5 categories such that aprrox-

Table 6: Wage growth and Vertical Distance Between tOccupations

	all	unskilled	apprentices	university graduates
dbeta	0.2994	0.4235	0.2632	0.5004
	(0.0118)	(0.0232)	(0.0164)	(0.0278)

imately 20 % work in each category (I used all workers for this, not only workers who switch occupations. This is not that easy to do, since some occupations employ many more workers than others). Category 1 consists of occupations that pay the lowest wages, category 5 consists of occupations that pay the highest wages. The rows of each table denote the category a worker worked in in period $t-1$, while the columns of each table denote the category a worker works in in period t. Along the diagonal, the first number corresponds to workers who do not switch occupations (and hence remain in the same category), while the second number refers to workers who switch occupations, but stay in the same category. Table 7 shows the number of workers moving from one category to another, between period $t-1$ and t. Workers who start in low wage occupations (category 1) and switch occupations, are more likely to stay in category 1 or move to category 2 than to move to a higher category. Similarly, workers who start in high wage occupations (category 5) and switch occupations, are more likely to switch to occupations in category 5 or 4 than to occupations in a lower category. We observe the same pattern if we look at columns instead of rows. Workers who work in category 1 in period t, are most likely to have worked in category 1 in period $t-1$, in least likely to have worked in category 5. Also note that the proportion of workers who switch occupations is lower in high wage occupations. (Devide the number of workers who stay in the same occupation between period $t-1$ and t by the number of all workers in that category in period $t-1$). Finally, note that workers are more likely to move up a category than to move down a category. For instance, 36537 workers worked in category 1 in period $t-1$, but only 35581 workers in period t. Similarly, only 36518 workers worked in category 5 in period $t-1$, and 38771 in period t.

Table 8 shows the wage in period $t-1$ for the different groups of workers. First note that workers who stay in their occupation earn considerably higher wages than workers who switch occupations, no matter whether the work in high or low wage occupations. Moreover, the category workers move to has a strong impact on the wage which workers earned at their previous occupation. Workers who worked in category 1 in period $t-1$ and then move to category 5 earn a higher wage in period $t-1$ than workers who started out

Table 7: Transition Matrix: Number of Workers

	1	2	3	4	5	all
1	28101 (s) 2739 (m)	2474	1417	1363	443	36537
2	2191	34755 (s) 1826 (m)	1479	1148	619	42018
3	1445	1525	37033 (s) 1170 (m)	1109	1062	43345
4	964	871	673	33424 (s) 1526 (m)	1535	38993
5	141	141	175	949	33325 (s) 1787 (m)	36518
all	35581	41593	41947	39519	38771	197411

in category 1, switched occupations, but remained in category 1. We observe the same pattern in all other starting categories.

Table 8: Transition Matrix: Wage in period t-1

	1	2	3	4	5	all
1	4.74 (s) 4.45 (m)	4.72	4.76	4.74	4.96	4.76
2	4.57	4.83 (s) 4.65 (m)	4.64	4.66	4.72	4.79
3	4.65	4.68	4.88 (s) 4.70 (m)	4.73	4.81	4.85
4	4.52	4.62	4.68	4.91 (s) 4.71 (m)	4.86	4.88
5	4.66	4.82	4.80	4.96	5.19 (s) 4.97 (m)	5.17
all	4.69	4.79	4.85	4.88	5.14	4.87

Table 9 looks at wages in period t instead. Holding the category in period $t-1$ fixed (i.e. moving along rows), wages in period $t-1$ are increasing in the category workers move to. This is not that surprising: Workers who work in high wage occupations earn higher wages, even conditional on the category the worker worked in last period. Moreover, holding the category in period t fixed, i.e. moving along columns, wages tend to be increasing in the category in period $t-1$. Hence, the starting category positively affects wages, conditional on the category the worker goes to.

Table 10 looks at wage growth. First note that workers who do not switch occupations experience the lowest wage growth - typically lower than workers who move down. Moreover, conditional on workers' category in period $t-1$, wage growth tends to be higher if the worker moves to a higher category. Notice that wage growth tends to be lower than differences between average wages in two categories.

Table 9: Transition Matrix: Wage in Period t

	1	2	3	4	5	all
1	4.78 (s) 4.58 (m)	4.72	4.76	4.74	4.96	4.76
2	4.67	4.88 (s) 4.65 (m)	4.82	4.77	5.01	4.86
3	4.71	4.78	4.94 (s) 4.85 (m)	4.86	5.07	4.92
4	4.67	4.75	4.83	4.98 (s) 4.88 (m)	5.09	4.96
5	4.72	4.86	4.93	5.10	5.25 (s) 5.25 (m)	5.24
all	4.75	4.85	4.92	4.96	5.24	4.95

Table 10: Transition Matrix: Wage Growth

	1	2	3	4	5	all
1	4.55 % (s) 13.59 % (m)	16.9 %	22.13 %	17.85 %	33.51 %	7.60 %
2	9.99 %	4.77 % (s) 12.20 % (m)	17.52 %	11.14 %	28.80 %	6.35 %
3	5.28 %	10.14 %	5.16 % (s) 15.35 % (m)	12.61 %	26.5 %	6.32 %
4	14.45 %	12.73 %	14.90 %	6.53 % (s) 17.77 % (m)	23.15 %	8.10 %
5	5.58 %	4.08 %	12.67 %	13.55 %	4.59 % (s) 27.63 % (m)	7.00 %
all	10.89 %	13.35 %	17.87 %	14.93 %	26.77 %	16.16 %

5.3 A Test of the Pure Matching Model

A *pure* matching model of occupational choice (such as Neal (1998) and Pavan (2004)) assumes that turnover rates are constant across occupations. The emphasis here is on *pure*. Of course, this does not mean that occupation-specific human capital is irrelevant if we reject this hypothesis. The same holds for the next hypothesis. This is clearly rejected by the data. Table lists the p-value of the hypothesis that occuaptions have no effect on mobility, conditional and unconditional on control variables.

Table 11: Hypothesis: Quit rates are the same across occupations

	all	unskilled	apprentices	university graduates
no controls	0.000	0.000	0.000	0.000
controls	0.000	0.000	0.000	0.000

A *pure* matching model of occupational choice also implies that condtional on moving, the occupation a worker works in this period does not affect the occupation the worker moves to. Similarly, their occupation one period before should have no effect on moving to a particular occupation. I tried to test the hypothesis for the two occupations with the most workers: 27 (Schlosser) and 78 (Bürofachkräfte). Table shows the proportion of workers who quit occupation 27, and move to a new occupation (column 3). It also the proportion of workers who move to occupation 27, and started at a different occupation (column 4). Table reports the same numbers for occupation 78. Finally, the table reports the proportion we would observe if mobility were random. This is the proportion we should observe if a worker's starting occupation has no impact on the occupation a worker goes to (and the other way round). This is computed as follows. I first compute the number of workers who work in occupations other than occupation 27. I then devide the number of workers in a particular occupation by this number. 'Random mobility' thus assumes that if one occupation employs twice as many workers as another, a worker is twice as likely to go to that occupation (or twice as likely to have come from that occupation). For both occupations, random mobility differs from observed mobility. I have performed a chi^2-test to test for the hypothesis that the two distributions (i.e. column 2 a 3, as well as column 3 and 4) are the same. This hypothesis is clearly rejected (p-value$<$0.000). For instance, 43 % of workers who worked in occupation 27 this period move to 7 out of the 77 occupation, but these 7 occupations only comprise 27 % of the total workforce. This is even more extreme for workers who worked in occupation 78 this period. Here, 70 % of workers move to 8 out of the 77 occupation, but these 8 occupations only employ 22 % of the total workforce.

Occupation 27: Not surprisingly, workers who worked in occupation 27 (Schlosser) this period are unlikely to move to typical white collar office jobs, such as Kaufleute, Bürofachkräfte (compare column 2 with column 1). Similarly, workers who worked in white collar office jobs this period are unlikely to move to occupation 27 (compare column 2 and 4). It seems that workers are more likely to move to and come from similar occupations. I investigate this further in one of the sections below. Moreover, notice that disproportionally many workers move from occupation 27 to occupation 62 (Schlosser and Techniker). Less workers switch from occupation 62 to occupation 27. Occupation 62 is higher up in the ranking than occupation 27. It could be that occupation 27 is a 'stepping stone' for occupation 62. Also notice that few workers (less than proportionally) move to occupation 60 (Ingenieure), a related occupation that is even further up the ranking. It thus seems hard to make the transition from occupation 27 to occupation 60.

Bürofachkräfte: It is evident that workers who worked in occupation 78 this period are more likely to

Table 12: Occupation 27: Schlosser

	random mobility	worker left occupation 27: Where does worker go to? N = 2894	worker joined occupation 27: Where does worker come from? N = 2364
Bürofachkräfte	6.44	2.04	0.76
Elektriker	5.77	4.91	5.20
Kaufleute	5.08	2.42	1.65
Ingenieure	4.86	1.83	0.17
Lagerarbeiter	4.54	6.50	7.36
Mechaniker	4.29	5.67	10.28
Bankkaufleute	3.66	0.38	0.16
Techniker	3.58	6.95	1.18
Schienenfahrzeugführer	3.48	5.29	4.06
Hilfsarbeiter	1.54	3.42	3.89
Montierer	3.22	7.39	8.25
Tischler	3.04	1.87	2.66
Löter	0.57	3.04	2.88

move to related occupations, such as Kaufleute, Buchhalter and Unternehmensberater. Similarly, workers who move to occupation 78 are more likely to have worked in one of these occupations before. Notice that the proportion of workers joining the occupation 75 (Unternehmensberater) is twice as high as the proportion of workers coming from that occupation. Maybe occupation 78 is a 'stepping stone' for occupation 75.

Table 13: Occupation 78: Bürofachkräfte

	random mobility	worker left occupation 78: Where does worker go to? N = 2141	worker joined occupation 78: Where does worker come from? N = 2454
Schlosser	6.70	0.84	2.40
Elektriker	5.77	1.17	2.89
Kaufleute	5.07	23.26	20.54
Ingenieure	4.86	2.20	1.75
Lagerarbeiter, -verwalter	4.54	7.29	8.96
Mechaniker	4.29	1.17	2.04
Bankkaufleute	3.66	5.46	3.87
Techniker	3.58	4.95	3.50
Schienenfahrzeugführer	3.48	3.64	4.20
Montierer	3.22	0.56	1.67
Tischler	3.04	0.47	1.22
Buchhalter	2.21	11.49	5.83
Unternehmensberater	1.62	9.20	4.60
Speditionskaufleute	0.87	4.76	4.32

5.4 Occupational mobility and similarity

Are workers more likely to move to similar occupation, conditional on moving? Table compares the average distance measure that is observed in the data with the average distance measure that we would observe under random mobility. It also compares the quartiles and median under observed and random mobility. See the previous section for a definition for 'random mobility'. Mean and quartiles are lower under observed mobility for both distance measures, indicating that workers are more likely to move to similar occupations. Differences tend to be higher for apprentices than for unskilled workers and, in particular, university graduates.

Table 14: Occupational Mobility and Similarity

	mean		1st quartile		median		3rd quartile	
	random	observed	random	observed	random	observed	random	observed
all: types	0.5357	0.2260	0.0307	0.0163	0.1915	0.0544	1.0714	0.1915
all: tasks	1.2272	0.7557	0.5410	0.3288	1.1128	0.6729	1.7977	1.0730
unskilled: types	0.2837	0.1747	0.0140	0.0123	0.0477	0.0435	0.2028	0.1241
unskilled: tasks	0.8826	0.7145	0.3961	0.3238	0.7791	0.6407	1.2033	1.0180
apprentices: types	0.4597	0.2385	0.0224	0.0153	0.1010	0.0482	0.9523	0.2015
apprentices: tasks	1.0744	0.7649	0.4558	0.3277	1.1013	0.6999	1.6577	1.0844
university: types	0.2961	0.3083	0.0554	0.0554	0.1133	0.1079	0.4157	0.3573
university: tasks	0.9596	0.8286	0.4868	0.4460	0.9165	0.6473	1.2903	1.1168

A different way to test the hypothesis that workers are more likely to move to similar occupations is to estimate mobility models. I ran 77 probit models (one for each occupation) of the following type. The dependent variable is equal to one if the worker goes to occupation A (B, C..), 0 if he moves to another occupation. I only use information on workers who have switched occupations. The explanatory variable is the distance index between occupation A (the occupation which the worker has joined), and the occupation which the worker has left. Both distance measures lead similar answers. For all occupations, the more distant two occupations, the less likely that a worker moves to this occupation. The coefficient is significant at (at least) a 5 % level for 74 out of the 77 occupations. The coefficient varies accross occupations. It is particularly high at occupations 27,62, 68,78, and particularly low at occupations 8, 9, 30, and 40. The pattern is roughly similar for all education groups.

Do these results really mean that workers are more likely to move to occupations in which similar tasks are performed? Recall that the previous section provided some evidence that workers are more likely to move to occupations that on average pay similar wages as their current occupation. Suppose that the vertical distance measure that is based on wages is positvely correlated with the similarity measure that is based on tasks. The results above could simply mean that workers are more likely to move to occupations that are vertically close, but not to occupations that are close in terms of tasks performed.

Table 15 reports the correlation between the vertical distance measure and our similarity measure. It is appearant that the vertical distance measure is positively correlated with the task distance measure. This I find quite interesting in itself. It means that occupations in which similar tasks are performed on average pay similar wages. Also note that the correlation between the two distance measures is higher for university graduates.

Table 15: Correlation between the vertical distance and task distance measure

	all	unskilled	apprentices	university graduates
types	0.5026	0.2233	0.4532	0.7967
tasks	0.4524	0.3972	0.3976	0.6440

I reran the probit regressions above, and included the vertical distance measure as an explanatory variable, instead of the task distance measure. For most occupations, the closer the occupation, the more likely is that the worker to move to that occupation. The coefficient is significant for 70 out of the 74 occupations, and roughly similar in magnitude as the distance measure based on tasks. When I control for both the vertical and the task distance measure, the task distance measure continues to have a negative and significant effect on the probability of moving to a certain occupation, for most occupations. Hence, conditional on the average wage paid in occupations, workers are more likely to move to occupations in which similar tasks are performed. The vertical distance measure also continues to have a negative impact on the probability of moving to a certain occupation, but this effect is not significant for some occupations.

Do workers who move to similar occupations in terms of tasks experience a lower or higher wage growth? Table 16 shows the wage growth of occupational movers, conditional on the task distance measure. The upper panel uses the distance measure that is based on types, the lower panel uses the distance measure that is based on tasks. It turns out that workers who switch to occupations that are more similar in terms of tasks experience a lower wage growth. In fact, this relationship is strongly concave - if a worker switches to an occupation that is *very* different to his previous one, his wage growth may be lower than if he had switched to a similar occupation. Notice that the relationship between the distance measure and wage growth of occupational switchers is particularly strong for university graduates.

Is this positive relationship between wage growth and distance due to the fact that an occupation that is very different from another is higher up the job ladder, and workers switching to an occupation higher up

Table 16: Wage growth and task distance

	all	unskilled	apprentices	university graduates
		types		
s	0.0698 (0.0050)	0.0665 (0.0114)	0.0403 (0.0060)	0.2061 (0.0151)
s	0.2997 (0.0171)	0.2671 (0.0387)	0.2379 (0.0203)	0.6332 (0.0477)
s^2	-0.1958 (0.0142)	-0.1667 (0.0320)	-0.1747 (0.0172)	-0.3529 (0.0374)
dbeta	0.2703 (0.0119)	0.4168 (0.0233)	0.2399 (0.0167)	0.3789 (0.0309)
s	0.2457 (0.0170)	0.2471 (0.0383)	0.2018 (0.0204)	0.4681 (0.0489)
s^2	-0.1715 (0.0141)	-0.1579 (0.0317)	-0.1547 (0.0172)	-0.2906 (0.0372)
		tasks		
s	0.0553 (0.0028)	0.0536 (0.0057)	0.0348 (0.0033)	0.1662 (0.0088)
s	0.1560 (0.0075)	0.1556 (0.0151)	0.1219 (0.0090)	0.4182 (0.0259)
s^2	-0.0635 (0.0044)	-0.0695 (0.0096)	-0.0541 (0.0052)	-0.14569 (0.0141)
dbeta	0.2772 (0.0118)	0.4144 (0.0232)	0.2483 (0.0165)	0.3773 (0.0295)
s	0.1648 (0.0075)	0.1505 (0.0150)	0.1170 (0.0090)	0.3959 (0.0255)
s^2	-0.0708 (0.0044)	-0.0688 (0.0095)	-0.0545 (0.0052)	-0.1604 (0.0139)

the job ladder experience a higher wage growth? In order to check this, I next included the difference in the average wage between two occupations into the wage regression (dbeta). This has little effect on the impact of the distance measure on wage growth.

Note: For the wage growth regressions, I haven't done anything about censoring. For unskilled workers and apprentices, censoring is not a problem. For univeristy graduates, 5.5 % of the observations are censored (i.e. either the starting or the joining wage, or both, are censored).

5.5 Occupational mobility and specialization

Table 17 reports the relationship between specialization of an occupation and experience. The analysis is based on definition 1, i.e. the unweighted average number of tasks performed in an occupation. The more tasks are performed on average in an occupation, the less specialized the occupation. (Hence, our specialization measure is in fact a generality measure). Clearly, the average number of tasks performed increases with experience for all education groups. This implies that workers move to more general occupations over the life-cycle. Note that university graduates are employed at occupations where more tasks are performed. We observe the same pattern if a different definition for specialization is used.

The negative relationship between experience and specialization could be due to the fact that more general occupations pay higher wages, and that workers move to occupations that pay higher wages over the life-cycle. Table 18 reports the correlation between the average wage paid in an occupation and the generality measure. (The correlation coefficient depends on which data I use to compute it. For table 18 I

Table 17: Specialization and experience

experience	all	unskilled	apprentices	university graduates
1	2.795	2.466	2.713	3.490
2	2.823	2.497	2.718	3.633
3	2.831	2.495	2.721	3.667
4	2.837	2.486	2.729	3.696
5	2.840	2.487	2.736	3.710
6	2.845	2.493	2.746	3.726
7	2.845	2.494	2.748	3.730
8	2.842	2.497	2.756	3.746
9	2.845	2.483	2.764	3.747
10	2.844	2.475	2.763	3.750

used yearly data so that occupations with more workers are given more weight. The qualitative pattern is the same if I use only one observation per occupation, or if I use different weights). The generality measure and the average wage paid in an occupation are strongly positively correlated, no matter which generality measure I use. Interestingly, this correlation is much weaker for unskilled workers than for apprentices and univeristy graduates. It is strongest for university graduates.

Table 18: Correlation beta and specialization measure

	all	unskilled	apprentices	university graduates
tasks	0.7955	0.1924	0.6786	0.7015
types	0.7316	0.0336	0.5980	0.7947
mixture	0.7977	0.1488	0.6761	0.7568

Table **??** confirms these results. Here, I regressed the log-wage on the generality measure, with and without other controls, for the three generality measures. For all education groups, workers working in more general occupations earn a higher wage. Controlling for education, experience, experience squared, occupation tenure, occupation tenure squared and year effects reduces the impact of generality on wages. Interestingly, the impact of the average number of tasks performed in an occupation on wages is particularly strong for university graduates. This holds true for all definitions of generality.

Table 19: Wages and Specialization

	all	unskilled	apprentices	university graduates
task	0.2132 (0.0011)	0.0396 (0.0036)	0.0904 (0.0015)	0.2704 (0.0031)
+ controls	0.0828 (0.0011)	0.0183 (0.0031)	0.0687 (0.0013)	0.1967 (0.0028)
type	0.3932 (0.0026)	-0.0510 (0.0076)	0.0909 (0.0032)	0.8756 (0.0096)
+ controls	0.0777 (0.0025)	-0.0787 (0.0066)	0.0476 (0.0028)	0.6465 (0.0087)
mixture	0.2962 (0.0016)	0.02881 (0.0051)	0.1104 (0.0021)	0.4492 (0.0049)
+ controls	0.1023 (0.0016)	0.0008 (0.0044)	0.0792 (0.0019)	0.3297 (0.0044)

Table 20 revisits the relationship between generality and experience. Here, I first regressed the average number of tasks performed in an occupation on experience. In line with the results reported in table 17, the coefficient on experience is positive, indicating that workers move to more general occupations over the life-cycle. I then additionally control for the vertical ranking of each occupation, in order to account for the fact that workers move to higher paying occupations over the life-cycle. The coefficient on experience now flips sign: Conditional on average wages paid in an occupation, workers move to more specialized jobs over the life-cycle. The coefficient on experience decreases if I additionally control for education, but remains statistically significant. The coefficient on experience increases if I also control for year effects. This suggests that over time, workers have moved to more general occupations. Results are qualititatively similar when other specialization measures are used. There are some interesting differences across educations groups. For unskilled workers and apprentices, the coefficient on education is negative if we control for the average wage paid in an occupation. For university graduates, on the other hand, the coefficient on experience remains positive, even if we control for averages wages and year effects.

Table 20: Specialization and experience revisited

	all	unskilled	apprentices	university graduates
		tasks		
exp	0.0090 (0.0003)	0.0017 (0.0005)	0.010 (0.0003)	0.0230 (0.0007)
+ beta	-0.0020 (0.0002)	-0.0010 (0.0005)	-0.0028 (0.0003)	0.0095 (0.0005)
+ educ	-0.0005 (0.0002)			
+ year	-0.0020 (0.0002)	-0.0058 (0.0006)	-0.0045 (0.0003)	0.0092 (0.0005)
		types		
exp	0.0032 (0.0001)	-0.0005 (0.0002)	0.0045 (0.0002)	0.0068 (0.0002)
+ beta	-0.0011 (0.0001)	-0.0005 (0.0002)	-0.0006 (0.0001)	0.0019 (0.0001)
+ educ	-0.0004 (0.0001)			
+ year	-0.0018 (0.0001)	-0.0041 (0.0003)	-0.0020 (0.0001)	0.0017 (0.0001)
		mixture		
exp	0.0061 (0.0002)	0.0006 (0.0003)	0.0073 (0.0002)	0.01491 (0.0004)
+ beta	-0.0015 (0.0001)	-0.0008 (0.0003)	-0.0017 (0.0002)	0.0057 (0.0003)
+ educ	-0.0005 (0.0001)			
+ year	-0.0019 (0.0002)	-0.0050 (0.0004)	-0.0033 (0.0002)	0.0055 (0.0003)

Table 21 repeats the analysis, but with a different dependent variable. The dependent variable now is the average wage paid in an occupation. For unskilled workers and apprentices, the coefficient on experience is positive, even if we control for the generality measure and year effects. For university graduates, on the other hand, the coefficient on experience becomes negative, once we condition on the generality measure and year effects.

Table 21: Ranking and experience revisited

	all	unskilled	apprentices	university graduates
exp	0.0039 (0.0001)	0.0043 (0.0001)	0.0044 (0.0001)	0.0050 (0.0002)
+ specialization	0.0019 (0.0001)	0.0042 (0.0001)	0.0028 (0.0001)	0.0007 (0.0001)
+ educ	0.0023 (0.0001)			
+ year	0.0018 (0.0001)	0.0043 (0.0002)	0.0027 (0.0001)	-0.0009 (0.0001)

Table 22 analyzes the relationship between mobility and specialization. For all education groups, workers who are employed in occupations where more tasks are performed are less likely to switch occupations. This holds true if we control for experience, experience squared, education, occupation tenure, occupation tenure squared, year effects, and in particular wages. Results are robust to alternative measures for specialization.

Table 22: Mobility and specialization

	all	unskilled	apprentices	university graduates
task	-0.0749 (0.0009)	-0.0671 (0.0032)	-0.0616 (0.0013)	-0.0876 (0.0018)
+ controls	-0.0481 (0.0010)	-0.0469 (0.0031)	-0.0428 (0.0012)	-0.0565 (0.0019)
type	-0.1498 (0.0021)	-0.1142 (0.0067)	-0.1034 (0.0028)	-0.2732 (0.0053)
+ controls	-0.1011 (0.0022)	-0.1029 (0.0064)	-0.0835 (0.0025)	-0.1859 (0.0054)
mixture	-0.1065 (0.0013)	-0.0949 (0.0046)	-0.0846 (0.0018)	-0.1409 (0.0028)
+ controls	-0.0709 (0.0014)	-0.0718 (0.0044)	-0.0613 (0.0017)	-0.0934 (0.0029)

6 Wage Growth and Occupational Mobility

7 Robustness Analysis

8 Conclusion

... to come ...

References

[1] Abowd, J.M. and F. Kramarz (1997): "Internal and External Labor Markets: An Analysis of Matched Longitudinal Employer-Employee Data," *NBER Working Paper* No. 6109.

[2] Antonovics, K. and L. Golan (2005): ”Experimentation and Job Choice”, mimeo, Carnegie Mellon University, Tepper School of Business.

[3] Autor, D. (2003), " ", *Quarterly Journal of Economics*,

[4] Baker, G.; M. Gibbs and B. Holmstrom (1994a): ”The Internal Economics of the Firm: Evidence from Personnel Data,” *Quarterly Journal of Economics*, 109: 881-919.

[5] Baker, G.; M. Gibbs and B. Holmstrom (1994b): ”The Wage Policy of a Firm,” *Quarterly Journal of Economics*, 109: 920 - 955.

[6] Burdett, K. (1978): ”A Theory of Employee Job Search and Quit Rates,” *American Economic Review*, 68: 212-20.

[7] Chan, W. (1996): ”External Recruitment versus Internal Promotion,” *Journal of Labor Economics*, 14: 555-70.

[8] Dustmann, C. and C. Meghir (2003): ”Wages, Experience and Seniority,” *IFS Working Paper 01/01.*

[9] Flinn, C. (1986): ”Wages and Job Mobility of Young Workers,” *Journal of Political Economy*, 94: S88-S110.

[10] Gibbons, R.; L.F. Katz; T. Lemieux and D. Parent (2002): ”Comparative Advantage, Learning and Sectoral Wage Determination,” *NBER Working Paper* No.8889.

[11] Gibbs, M.; K. Ierulli and E. Meyersson Milgrom (2003a): ”Careers in Firm and Occupational Labor Markets,” *mimeo, Graduate School of Business, University of Chicago.*

[12] Ierulli, K., E.P. Lazear and E. Meyersson Milgrom (2003): ”Wage Mobility in Sweden,” *mimeo, Stanford University.*

[13] Johnson, W. (1978): ”A Theory of Job Shopping,” *Quarterly Journal of Economics*, 261-77.

[14] Jovanovic, B. (1979a): ”Firm Specific Capital and Turnover,” *Journal of Political Economy*, 87: 1246-60.

[15] Jovanovic, B. (1979b): ”Job Matching and the Theory of Turnover,” *Journal of Political Economy*, 87: 972-90.

[16] Jovanovic, B. (1984): "Matching, Turnover, and Unemployment," *Journal of Political Economy*, 92: 108-22.

[17] Jovanovic, B. and Y. Nyarko (1997): "Stepping Stone Mobility," *Carnegie Rochester Series for Public Policy*, 46: 289-337.

[18] Kambourov, G. and I. Manovskii (2002): "Occupational Specificity of Human Capital," *mimeo, University of Toronto.*

[19] Kwon, I. and E. Meyersson Milgrom (2003): "Boundary of Internal Labor Markets: Do We Have the Empirical Facts Right?," *mimeo, University of Michigan.*

[20] Lazear, E.P. (2003): "“Firm-Specific Human Capital: A Skill-Weights Approach," *NBER Working Paper* No. w9679.

[21] Lazear, E.P. and P. Oyer (2003): "Internal and External Labor Markets: A Personnel Economics Approach," *NBER Working Paper* No. 10192.

[22] McCue, K. (1996): "Promotions and Wage Growth," *Journal of Labor Economics*, 14: 175-209.

[23] Medoff, J. and K. Abraham (1980): "Experience, Performance, and Earnings," *Quarterly Journal of Economics*, 95: 703-36.

[24] Medoff, J. and K. Abraham (1981): "Are Those Paid More Really More Productive," *Journal of Human Resources*, 16: 186-216.

[25] Moscarini, G. and F. Vella (2003): "Occupational Mobility and Employment Reallocation: Evidence from the NLSY79," *mimeo, Yale University.*

[26] Neal, D. (1995): "Industry-Specific Capital: Evidence from Displaced Workers," *Journal of Labor Economics*, 13: 653-77.

[27] Neal, D. (1999): "The Complexity of Job Mobility among Young Men," *Journal of Labor Economics*, 17: 237-61.

[28] O'Flaherty, B. and A. Siow (1995): "Up-or-Out Rules in the Market for Lawyers," *Journal of Labor Economics*, 13: 709-33.

[29] Parent, D. (2000): "Industry-Specific Capital and the Wage Profile: Evidence from the National Longitudinal Study of Income Dynamics," *Journal of Labor Economics*, 18: 306-23.

[30] Pavan, R. (2004): "Career Choice and Wage Growth", University of Chicago, mimeo.

[31] Rosen, S. (1972): "Learning and Experience in the Labor Market," *Journal of Human Resources*, 7: 325-42.

[32] Shaw, K. (1987): "Occupational Change, Employer Change, and the Transferability of Skills," *Southern Economic Journal*, 53: 702-19.

[33] Sicherman, N. and O. Galor (1990): "A Theory of Career Mobility," *Journal of Political Economy*, 98: 145-76.

[34] Spitz, (2004),

[35] Topel, R. (1991): "Specific Capital, Mobility and Wages: Wages Rise with Job Seniority," *Journal of Political Economy*, 99: 145-76.

[36] Topel, R. and M. Ward (1992): "Job Mobility and the Careers of Young Men," *Quarterly Journal of Economics*, 107: 439-79.

[37] Upward, R. and P. Wright (2003): "Snakes or Ladders? Skill Upgrading and Occupational Mobility in the US and UK during the 1990s," *mimeo, University of Nottingham.*

[38] Weiss, Y. (1971): "Learning by Doing and Occupational Specialization," *Journal of Economic Theory*, 3: 189-98.

A Data

In our data, firms are not required to notify an occupational switch to the authorities. Hence, we do not know the exact date when a worker switched occupations. Consider first a worker who switches firms on April 1st. For this worker, we observe two spells during the year he switches, the first from January 1st to March 31st, the second from April 1st to December 31st. Suppose that the occupation variable is A in the first spell, and B in the second spell. For this worker, and firm switchers in general, it is reasonable to assume that he worked on January 1st at occupation A, and April at occupation B. He may have switched occupations once more between January and April, and between April and December. Next, consider a worker who stayed with the same employer for (at least) two years. For this worker, we observe two spells, both from January 1st to December 31st. Suppose that the first spell classifies the worker as in occupation A, while this spell classifies him as in occupation B. For this worker, it is reasonable to assume that on January 1st he was working in occupation A, and on January first one year later in occupation B. He may have switched occupations more than once. We do not know when exactly the worker switched occupations. This makes it somewhat arbitrary to compute occupational mobility rates by (actual) experience. Also note that we may underestimate occupational mobility.

- no distinction between quits and layoffs. We need to assume that all separations are voluntary and thus the outcome of worker maximization. Alternatively, we can introduce an exogenous probability of separation in addition to the worker's quit decision.
- censoring of wages (ceiling in social security contributions)
- describe definition of occupations either from Tarifvertraege (Tarifarchiv BSI) or from chamber of commerce (*Ausbildungsberufsklassifizierung*)
- account for changes in the occupational structure and thus occupational ranking over time by readjusting weights.

B Extensions of the Basic Model

We can easily modify the model in order to allow workers to learn about their true skill at each task. The more time they spend at each task, the more they learn about their true productivity at that task. Does this provide an additional motive for diversification in the first period? Suppose workers' true productivity at the two tasks is equal to A and B. Both A and B are normally distributed with mean μ and σ_μ^2. At the beginning of the first period, workers receive two signals about their productivity at the two tasks:

$$z_A^1 = A + \varepsilon_A^0, \text{ and } z_B^1 = B + \varepsilon_B^0.$$

ε_A and ε_B are normally distributed with mean 0 and $\sigma_{\varepsilon_0}^2$. We assume that A and B are independent. Hence, signal B and signal A provide no information about the other task. The analysis would be more complicated if A and B were correlated. The model still should be manageable as long as A and B are linearly dependent.

Workers use these signals to update their productivity at the two tasks:

$$\begin{aligned}
E\left[A|z_A^1\right] &= \frac{\mu\sigma_{\varepsilon_0}^2 + z_A^1\sigma_\mu^2}{\sigma_{\varepsilon_0}^2 + \sigma_\mu^2} := A_0, \\
V\left[A|z_A^1\right] &= \frac{\sigma_{\varepsilon_0}^2\sigma_\mu^2}{\sigma_{\varepsilon_0}^2 + \sigma_\mu^2} := \sigma_0^2, \\
E\left[B|z_B^1\right] &= \frac{\mu\sigma_{\varepsilon_0}^2 + z_B^1\sigma_\mu^2}{\sigma_{\varepsilon_0}^2 + \sigma_\mu^2} := B_0, \text{ and} \\
V\left[B|z_B^1\right] &= \frac{\sigma_{\varepsilon_0}^2\sigma_\mu^2}{\sigma_{\varepsilon_0}^2 + \sigma_\mu^2} := \sigma_0^2.
\end{aligned}$$

How much workers learn about their productivity in each task during the first period depends on how much time they spend at each task during the first period. Suppose at the end of the first period, workers and firms observe $y_A^2 = pA + \varepsilon_A^1$, and $y_B^2 = (1-p)B + \varepsilon_B^1$. ε_A^1 and ε_B^1 are normally distributed with mean 0 and variance $\sigma_{\varepsilon_1}^2$. From this they can compute

$$\frac{y_A^2}{p} = A + \frac{\varepsilon_A^1}{p} := z_A^2, \text{ and } \frac{y_B^2}{1-p} = B + \frac{\varepsilon_B^1}{1-p} := z_B^2.$$

Workers use z_A^2 and z_B^2 as new signals for their true productivity A and B. Note that if workers spend little time an task A, signal z_A^2 will be very noisy and workers learn little about their true productivity at task A.

The same holds if workers spend little time on task B. Workers compute

$$\begin{aligned}
E\left[A|z_A^1, z_A^2\right] &= \frac{A_0\frac{\sigma_{\varepsilon_1}^2}{p^2} + z_A^1\sigma_0^2}{\frac{\sigma_{\varepsilon_1}^2}{p^2} + \sigma_0^2} := A_1, \\
V\left[A|z_A^1, z_A^2\right] &= \frac{\frac{\sigma_{\varepsilon_1}^2}{p^2}\sigma_0^2}{\frac{\sigma_{\varepsilon_1}^2}{p^2} + \sigma_0^2} := \sigma_1^2, \\
E\left[B|z_B^1, z_B^2\right] &= \frac{\mu\frac{\sigma_{\varepsilon_1}^2}{(1-p)^2} + z_B^1\sigma_0^2}{\frac{\sigma_{\varepsilon_1}^2}{(1-p)^2} + \sigma_0^2} := B_1, \text{ and} \\
V\left[B|z_B^1, z_B^2\right] &= \frac{\frac{\sigma_{\varepsilon_1}^2}{(1-p)^2}\sigma_0^2}{\frac{\sigma_{\varepsilon_1}^2}{(1-p)^2} + \sigma_0^2} := \sigma_1^2.
\end{aligned}$$

In the second period, workers maximize

$$\max_{p_2} \quad \sqrt{p_2 A_1} + \sqrt{(1-p_2)B_1},$$

and choose

$$p_2 = \frac{A_1}{A_1 + B_1}.$$

In the first period, workers take into account the optimal occupational choice in the second period, and recognize that they learn more about their true productivity in each task the more they spend on each task. They maximize

$$\max_{p_1} \sqrt{p_1 A_0} + \sqrt{(1-p_1)B_0} + E[\sqrt{A_1 + B_1}|A_0, B_0].$$

We need to derive the distribution of $A_1 + B_1$, conditional and A_0 and B_0. Conditional an A_0, A_1 is normally distributed with mean A_0 and variance $\frac{\sigma_0^4}{\frac{\sigma_{\varepsilon_1}^2}{p^2}+\sigma_0^2}$. Note that the higher p, the higher the variance - i.e. the more there is to learn about the true productivity. Conditional on $B_0,$ B_1 is normally distributed with mean B_0 and variance $\frac{\sigma_0^4}{\frac{\sigma_{\varepsilon_1}^2}{(1-p)^2}+\sigma_0^2}$. Hence, $A_1 + B_1$ is normally distributed with mean $A_0 + B_0$, and variance $\frac{\sigma_0^4}{\frac{\sigma_{\varepsilon_1}^2}{(1-p)^2}+\sigma_0^2} + \frac{\sigma_0^4}{\frac{\sigma_{\varepsilon_1}^2}{p^2}+\sigma_0^2} := \widetilde{\sigma}^2$. Workers' objective function becomes

$$\sqrt{p_1 A_0} + \sqrt{(1-p_1)B_0} + \int_{-\infty}^{\infty} \sqrt{|A_0 + \varepsilon_A + B_0 + \varepsilon_B|}\frac{1}{\sqrt{4\pi\widetilde{\sigma}^2}}\exp(-0.5\frac{(\varepsilon_A + \varepsilon_B)^2}{2\widetilde{\sigma}^2})d(\varepsilon_A + \varepsilon_B).$$

Again, it is assumed that A_1 and B_1 are almost always positive so that the above expression is a good approximation. Whether learning leads to more or less diversification in the first than in the second period, depends on $\frac{\partial\tilde{\sigma}^2}{\partial p}$.

$$\frac{\partial\tilde{\sigma}^2}{\partial p} = \sigma_0^4\sigma_{\varepsilon_1}^2(\frac{1}{p^3(\frac{\sigma_{\varepsilon_1}^2}{p^2}+\sigma_0^2)^2} - \frac{1}{(1-p)^3(\frac{\sigma_{\varepsilon_1}^2}{(1-p)^2}+\sigma_0^2)^2}).$$

This expression is 0 if $p = 0$. Hence, workers with $A_0 = B_0$ continue to choose $p_1 = 0.5$. For other values of p, the sign of this expression depends on σ_0^2 and $\sigma_{\varepsilon_1}^2$. If $\sigma_0^2 = \sigma_{\varepsilon_1}^2$, then $\tilde{\sigma}^2$ is minimized if $p = 0.5$. Hence, there should be more specialization in the first period than in the second period.

This depends on the particular functional form. Suppose instead that workers and firms observe $\sqrt{p}A_0 + \varepsilon_1$, and $\sqrt{1-p}B_0 + \varepsilon_1$. Then $\tilde{\sigma}^2$ equals

$$\tilde{\sigma}^2 = \frac{\sigma_0^4}{\frac{\sigma_{\varepsilon_1}^2}{(1-p)}+\sigma_0^2} + \frac{\sigma_0^4}{\frac{\sigma_{\varepsilon_1}^2}{p}+\sigma_0^2}.$$

This expression is maximized when $p = 0.5$ (although $\tilde{\sigma}^2$ does not seem to be to dependent on p). Overall, it seems that learning does not provide a strong motive for diversification in the first period and this is independent of the underlying technology chosen.