In the Right Place at the Wrong Time –

The Role of Firms and Luck in Young Workers' Careers

Till von Wachter[*]
Columbia University

Stefan Bender
Institut für Arbeitsmarkt- und Berufsforschung

We exploit administrative data on young German workers and their employers to study the long-term effects of an early job loss. To account for non-random sorting of workers into firms with different turnover rates and for selective job mobility, we use changes over time in firm- and age-specific labor demand as an instrument for displacement. We find that wage losses of young job losers are initially 15% but fade to zero within five years. Only workers leaving very large establishments suffer persistent losses. A comparison of estimators implies that initial sorting, negative selection, and voluntary job mobility may have biased previous U.S. studies finding permanent effects of early displacements. (JEL J63, J65)

1. Introduction

There are two competing views of the role of early job mobility in young workers' careers. One view, shared by many economists, is that early job mobility plays an important role in career development and wage growth. An alternative view that has informed policy proposals in the past highlights the potential costs of early job mobility. For young Americans the rate of job change is indeed very high, a fact that has been interpreted as evidence of beneficial job search (Topel and Ward 1992). But it has also been argued that the unstructured transition from school to work in the US labor market leads to excess mobility and slows the rate of human capital accumulation (Ryan 2001). In fact, we know that young workers have high displacement rates (Farber 1993) and suffer the largest wage declines in recessions (Blanchflower and Oswald 1994). Consistent with the more negative view, studies of early job displacement typically find persistent wage losses for young displaced workers (Kletzer and Fairlie 2001, Gustafson 1998).

The paper presents estimates of the long-term wage losses suffered by young German workers who leave their training firm at the end of an apprenticeship. Similar to what has been found for the United States, simple comparisons of leavers and stayers suggest that there are large permanent costs of displacement – on the order of 10 percent after 5 years. These comparisons, however, ignore two critical issues suggesting that simple estimates overstate wage losses. First, it is widely recognized that leavers may be adversely selected (Gibbons and Katz 1991). A second issue that has received less attention in the literature is that the sample of leavers is disproportionately drawn from firms with high turnover rates. To the extent that high-turnover firms attract lower-quality apprentices, or offer lower-quality training, the nonrandom nature of the displaced worker pool is a problem. A third issue – particularly important for young workers – is that leavers include both involuntary movers and those who moved voluntarily. Since voluntary movers tend to benefit from mobility, this would lead simple estimates to understate the effects of displacements.

Ideally, what is needed to identify the causal effect of displacement in this environment is exogenous variation in firm-specific demand for apprentices. As a proxy for this, we use the fraction

of apprentices in the same cohort at the same firm who leave the firm at the end of training. By pooling data for several cohorts and adding firm fixed effects, the instrument represents year-to-year variation in the fraction of apprentices retained by each firm. This instrument is clearly orthogonal to permanent characteristics of the firm, and to any individual-specific demand side shocks, such as adverse selection or learning effects. It may still reflect some variation in supply side opportunities for the apprentices of a given firm in a given cohort. Thus, we consider a second instrumental variable, based on the fraction of the trainees' cohort that experiences a spell of unemployment at the end of their apprenticeship. The inclusion of firm fixed effects also controls for any bias from initial sorting of workers into firms based on unobserved ability.

The sample consists of all graduates from the German apprenticeship system in the period from 1992 to 1994 who are observed working at least once in the first five years after training. In Germany, more than two thirds of recent cohorts participate in apprenticeship training programs that last on average two years and include both formal and practical training. About 35% of apprentices leave their training firm at graduation, suggesting adverse selection of workers is potentially an important problem. Initial sorting of workers into different types of firms is relevant as well, since firms provide different amounts of training and offer different career prospects as evident from variation in turnover rates. Moreover, high mobility of apprentices in the years following training suggests that some of those leaving their training firms move voluntarily. Thus, post-training mobility occurs in a rich environment with voluntary and involuntary mobility, adverse selection, and nonrandom sorting of workers into their training firm.

Using an instrumental variables (IV) estimator based on random firm-level fluctuations in retention rates, we find that involuntarily displaced trainees have initially lower wages than those who stayed, but that these losses disappear within five years after the end of training. Only the wage losses of workers leaving very large training firms have a persistent component, consistent with the presence of firm-size wage differentials or internal labor markets. Estimates accounting for nonrandom selection into and out of firms thus do not imply permanent negative effects of job losses. Understanding the discrepancy between these and the simple OLS estimates requires closer

examination of the different confounding factors. Alternative estimates of wage losses given by OLS with fixed effects, IV, or IV with fixed effects address different sources of selection within firms or initial sorting between firms. Comparison between these estimates therefore helps to disentangle the separate impacts of sorting and selection. Moreover, each of these confounding factors is closely related to a different theory of job and wage mobility. Thus, the comparison of different estimators also provides a way to assess the importance of the basic models of early job mobility among young workers.

To make the comparison between estimators and theoretical implications explicit, we present a straightforward model of wage determination that captures the basic theories of early job mobility in a unified framework. Using this model to interpret the empirical results, we conclude that standard job search theory provides a good explanation for the incidence of voluntary mobility and for the patterns of wage losses observed for involuntarily displaced workers. In addition, the fact that training firm fixed effects matter for both OLS and IV estimates suggest that higher ability workers are sorted into lower-turnover firms at the start of apprenticeships. Lower ability workers are also more likely to be released by their training firms at the end of training. This suggests a potentially important role of adverse selection in the labor market for young apprentices as suggested by Acemoglu and Pischke (1998) and Gibbons and Katz (1998).

The estimates draw a rich picture of the labor market for young workers where sorting, selection, and voluntary mobility occur simultaneously. This implies important insights on early careers may be lost if either of the components is ignored. By proposing a unified approach based on a rich set of data including information on firms, the paper complements and extends previous studies that focused only on single aspects of job and wage dynamics. The results also speak to potential biases affecting previous studies of early job mobility that found long-term effects of early job losses but lacked information on the demand side. First, individual fixed effects cannot be used to control for negative selection in the presence of adverse selection, training wages, or employer learning, because wage histories do not reflect workers' productivity. Second, if young workers sort themselves into firms by their turnover rates displacement is not a random event even controlling

for selection within firms. Once firm level information is used to take into account selection, sorting, and voluntary mobility in the market for German apprentices early job losses have initially strong but temporary effects.

The next section outlines the model of wage determination, relates it to theories of job mobility, and uses it to interpret the bias of OLS and to outline the estimation strategy. The third section describes the matched worker-firm data set, and compares the German apprenticeship system to the US labor market. The fourth section presents the basic empirical results and a detailed sensitivity analysis. The fifth section discusses the empirical findings in light of models of job and wage mobility and the last section concludes.

2. Estimation of Wage Losses and Theories of Job Mobility

Even though it is a well-documented feature of job mobility, standard models of the labor market do not predict that job losers experience wage declines. Several explanations have been proposed in the literature, each focusing on a separate aspect of job mobility. However, most of the mechanisms emphasized by different theories are likely to occur simultaneously in the labor market. The following statistical model of wage determination helps distinguish causal effects of displacements from potential confounding factors possibly affecting studies addressing only single mechanisms of wage and job mobility.

2.1. Wage Determination and Theories of Job Mobility

Consider a class of models in which young workers' real log wages are a function of their innate skills, a_i, and of their mobility status after their last job, $D_{i0} = V_{i0} + I_{i0}$. Mobility can be either voluntary ($V_{i0} = 1$) or involuntary ($I_{i0} = 1$); denote the gain or loss from voluntary and involuntary mobility t periods after a job change as δ_{Vt} and δ_{It}, respectively.[1] The goal of the

---

[1] Note that both voluntary and involuntary mobility could lead to gains or losses for different workers. In this case, one can reinterpret δ_{Vt} and δ_{It} as the *average* gain or loss from voluntary and involuntary mobility, respectively. As further discussed below, the main estimated coefficient then is the local average treatment effect.

4

analysis is to obtain an estimate of δ_{It}, the wage loss from a job displacement over time. However, as in many applications, suppose it is not known whether a job change was voluntary or not, that is, only D_{i0} is known, and neither V_{i0} nor I_{i0} is observed separately. The process determining wages t periods after a job change then is $w_{it} = \delta_{It}D_{i0} + (\delta_{Vt} - \delta_{It})V_{i0} + a_i + \varepsilon_{it}$. To capture that workers may be sorted among their *initial* employers, the firms that in the present application provide training, this can be rewritten as

$$w_{it} = \delta_{It}D_{i0} + (\delta_{Vt} - \delta_{It})V_{i0} + (a_i - \bar{a}_{j(i)}) + \bar{a}_{j(i)} + \varepsilon_{it}, \qquad (1)$$

where $\bar{a}_{j(i)}$ is average ability of workers at firm j that trained individual i, and ε_{it} is a random disturbance term. In this formulation, wages are determined by mobility status, an individual component of ability relative to the training firm's average, $a_i - \bar{a}_{j(i)}$, and a firm specific component of ability, $\bar{a}_{j(i)}$, neither of which is usually observed by the econometrician. This basic model is able to incorporate several theories of job and wage mobility. Each theory has implications for different components of Equation (1), and this will be helpful in the interpretation of the empirical results.[2]

A widely cited theoretical explanation for wage losses of displaced workers has been the presence of adverse selection in the labor market. The basic idea, put forward by Gibbons and Katz (1991) and already present in Greenwald (1986), is that in a world in which only current employers are informed about a worker's true ability a displacement may be perceived as a negative signal about a worker by other employers. In equilibrium, the firm displaces less able workers and these workers get paid according to their lower expected ability. Thus, displaced workers suffer wage losses even if job changes themselves have no direct effect on wages (i.e., $\delta_{It} = 0$). In terms of equation (1), adverse selection implies that movers are likely to be the least able workers within a firm, i.e., $\text{cov}(a_i - \bar{a}_{j(i)}, D_{i0}) < 0$. By raising the average ability of the pool of changing workers, equilibrium

---

[2] While most of these theories could be integrated into richer models explaining a broader set of facts, the following discussion concentrates on the main contribution of each theory.

is sustained by the presence of individuals moving for exogenous reasons. Exogenous mobility has been featured prominently in tests of the adverse selection hypothesis as well. For example, Gibbons and Katz find that wage losses are higher for workers displaced by lay-offs compared to those displaced from plant closings, as the latter should be less selected. Their basic insight can be generalized, and this will be taken up in the next section.

In a recent study, Krashinsky (2002) finds that among mature workers the differences in wage losses between workers displaced by plant closings and lay-offs found by Gibbons and Katz (1991) is partly driven by differences in firm size of pre-displacement employers.[3] Krashinsky's findings suggest that adverse selection may matter less for older workers. This is not surprising if markets continuously learn about workers' ability, as suggested by Farber and Gibbons (1996). In this case wages and career histories of older workers reflect their skills, and the information contained in any additional single signal such as a displacement is small.[4] However, it remains an important problem for young workers, since for them the market should still be learning about skills and motivation. This is exploited by Acemoglu and Pischke (1998), who develop a model of the German apprenticeship system in which employers' monopsony rents generated by private information about young workers encourage them to pay for general training. Thus, the hypothesis of adverse selection is particularly relevant for the present paper.[5]

Krashinsky's results also highlight the need to control for firm characteristics in the study of displaced workers. This point is related to a deeper issue raised by Gibbons and Katz (1991) themselves and further developed in a sequence of papers (Gibbons and Katz 1992, Gibbons, Katz, Lemieux, and Parent 2002) – namely that firms may differ systematically and that observed mobility and wage changes may be driven by a sorting process of heterogeneous workers into heterogeneous

---

[3] Krashinsky (2002) argues that small firms are more likely to close in face of economic shocks, while larger firms are more likely to reduce their work force. Laid-off workers should thus tend to loose any wage premia or rents they earned from working at larger firms.

[4] While Farber and Gibbons (1996) concentrate on symmetric learning, a similar result holds for asymmetric learning (von Wachter 2001a).

[5] Adverse selection has featured centrally in the debate on why firms pay for general training in Germany. The main alternative explanation has been the role of labor market institutions such as unions (Dustmann and Schoenberg 2002) or firing costs.

firms. Sorting is a particular problem for the study of displaced workers if less able workers are hired by firms with higher turnover rates. If workers are initially assigned to firms in this way, then not only are movers not comparable to stayers in general, but movers and stayers are not comparable across different types of firms.

That workers select themselves into firms based on ability is suggested by Abowd, Kramarz, and Margolis (1999), who find that differences in workers' ability levels explain a large fraction of wage-differences among firms.[6] Similarly, some firms may value job stability more than others and try to structure wage incentives accordingly. Since firms with low turnover rates offer better career opportunities and have greater incentives to invest into their work force in the form of high-quality training, they are likely to attract the most able workers. This is a key hypothesis in several theoretical models of turnover and wage-profiles (e.g., Salop and Salop 1976, Weiss and Wang 1998, Neal 1998). Yet, until recently little was known about differences in average mobility rates, sorting, and tenure profiles between firms. While a growing recent literature indeed documents a considerable degree of heterogeneity in turnover rates and growth rates among establishments (e.g., Davis and Haltiwanger 1990, Anderson and Meyer 1994, Abowd, Corbel, and Kramarz 1999), few studies address the interaction between turnover and sorting of workers.[7]

If firms thoroughly screen young workers during the hiring process, better firms may be able to attract the most skilled and motivated young workers and initial assignment is perfect. This is a likely scenario for Germany, where training firms thoroughly screen potential new apprentices by school grades, internships, and entry exams. To capture the effects of efficient initial assignment, equation (1) allows for differences in average ability across training firms, $\bar{a}_{j(i)}$.[8] In terms of this model, initial assignment of less-able workers into high turnover firms implies that

---

[6] That firms differ in key characteristics and that heterogeneous workers may select into firms based on these permanent differences is not a new idea. Groshen (1991) argues that differences in average firm compensation practices can explain a large fraction of cross-sectional wage variation. That part of these differences could be explained by differences in workers' ability has been raised in the literature on inter-industry wage differentials (e.g., Gibbons and Katz 1992).

[7] An exception is Margolis (1995), who taking into account the potential bias from selection of workers into firms shows that the return to seniority indeed varies considerably between firms.

[8] Note that this could also be interpreted as capturing differences in the quality of training across firms, or differences in firms' wages more generally.

$\text{cov}(\bar{a}_{j(i)}, D_{i0}) < 0$; in simple estimates displaced workers may then spuriously appear to obtain lower wages. As in the case of adverse selection, in this model workers' mobility status has no direct impact on wages because workers are always paid their marginal product, and thus $\delta_{It} = \delta_{Vt} = 0$.

Note that if one observes a fully informative pre-displacement wage, the bias from perfect initial sorting or negative selection can be eliminated by analyzing wage changes. While this is often possible for mature workers, as for example in the case of Jacobson, Lalonde, and Sullivan (1993), it might not be a good strategy for young workers. First, in many occupations young workers receive training wages that are below their ability levels. Second, if there is asymmetric information between employers about workers' ability pre-displacement wages do not reflect productivity (Gibbons and Katz 1991). Thus, given adverse selection is a particular problem for young workers, including individual fixed effects is not a valid strategy to control for selection. This is a key methodological difference distinguishing the present analysis from the current U.S. literature on early job loss (e.g., Gustafson 1998, Kletzer and Fairlie 2002).

Another reason for why wages of young workers may not fully reflect ability is if firms and workers themselves only gradually learn about their abilities and preferences. The process of sorting then becomes sequential as in Gibbons and Katz (1992). Because worse workers get down-ranked over time as employers learn about their true ability, this implies that displacements are associated with wage losses even controlling for initial assignment, i.e., there is both negative selection and initial sorting. However, now more able workers should leave less attractive firms once their ability becomes known – the opposite implication than from perfect initial assignment. Note that since every worker is paid according to their marginal product at all times, as before job changes by themselves have no direct effect on wages, i.e. $\delta_{It} = \delta_{Vt} = 0$.

The research on young displaced workers aims to focus exclusively on involuntary lay-offs. However, in an environment of high job-to-job fluctuations the distinction between involuntary and voluntary job change may be hard to draw. If measures of job displacement pool voluntary and involuntary movers, as many administrative data sets do, simple estimates of the earnings losses

from an early job change may *underestimate* the effect of a job loss on wages.[9] A recent study by Neumark (1998) using the NLSY shows that this problem may also arise in more conventional data sets when alternative measures of early job mobility are used.[10] Thus, while both adverse selection and sorting among firms imply that simple estimates overstate earnings losses of displaced workers, a high degree of voluntary mobility implies the opposite.

This is particularly relevant for young workers, since at least since Topel and Ward (1992) it is widely accepted that voluntary job mobility is an important feature of early careers in the U.S. Comparable estimates suggest that early mobility plays an important role in other countries as well (e.g., Euwals and Winkelman 2001, von Wachter and Giuliano 2004).[11] While not predicted by asymmetric information or initial assignment, a high degree of beneficial job mobility is consistent with models of job search, in which workers are homogeneous, but repeatedly draw job offers from a distribution of wages. Over time, workers searching on the job should obtain more favorable job matches, such that their wages grow with experience even in the absence of general human capital accumulation (Burdett 1978). However, a displacement destroys this 'search capital' because workers have to start looking for good jobs from scratch (Manning 2003). Thus, job search is a promising explanation for true temporary wage losses from job displacement. Gradually, workers again find better job matches, and the initial wage losses from displacements should be temporary. These predictions can be easily incorporated in the basic model of Equation (1). Job search implies that

---

[9] A recent paper by Bender, Dustmann, Margolis, and Meghir (1995) estimating wage losses of mature displaced using administrative data from France and Germany tries to circumvent this problem by defining a displacement to have occurred when workers spend at least 30 days out of the labor force after terminating a job. By focusing on displaced workers who became unemployed, this risks imposing part of the final outcome ex ante. To counter the same problem, Jacobson et al. (1993) construct a special 'mass-layoff' sample of workers who leave firms who experience large reduction in workforce.

[10] Gardecki and Neumark (1997) find that high job mobility early in a career (as measured by the number of jobs held or the highest job tenure attained in the first five years since market entry) has no impact on wages later in life, and they argues that such an estimate must be a lower bound due to the presence of negative selection. However, using local unemployment rates as an instrument for early job mobility, Neumark (1998) finds that early job mobility has a significant negative impact, strongly suggesting that mobile workers in his sample are positively selected.

[11] Topel and Ward (1992) find that for young men, during the first ten years in the labor market job changes lead to 10% wage increases on average, and that 30-40% of wage growth occurs at job changes. More generally, wage changes of voluntary movers have been found to be higher than those of involuntary lay-offs (Mincer 1986, Bartel and Borjas 1981).

$\delta_{Vt} > 0$, leading to a positive bias in simple OLS estimates.[12] Transitory wage losses imply that

$$\delta_{It} < 0 \text{ and } \frac{\partial \delta_{It}}{\partial t} > 0 .$$

Other models than job search imply true wage losses from an early job displacement, too. Within the standard neo-classical human capital model, displaced workers may loose skills specific to their previous employer or occupation (e.g., Kletzer 1998, Neal 1995). Alternatively, they could loose opportunities to acquire general skills if displacements increase time spent out of employment (Giuliano and von Wachter 2004). However, these explanations carry less weight in the present context. First, the German apprenticeship system is meant to provide mostly general skills. Thus, curricula at apprentice schools are set at the national level and on-the-job training is monitored by public agencies. Second and more importantly, both tenure spells and unemployment spells are generally short for young workers. In future work the available detailed data on career histories enables us to assess aspects of these alternative explanations directly.

None of the models discussed so far, and standard models of career development more generally, do imply permanent effects of job displacements on earnings. Apart from losses in labor market experience, empirical papers concerned with long-term 'scarring' effects of early job loss typically cite some form of permanent negative signaling to employers as explanation.[13] However, the idea that early displacements or unemployment experiences irrespective of duration 'scar' young workers relies on strong implicit assumptions on the learning process.[14] An alternative group of models of career development predicting permanent effects of a job loss rely on the importance of

---

[12] Over time, these gains are stable, or increasing if the new job has a steeper career-profile ($\partial \delta_{Vt} / \partial t \geq 0$).

[13] Most empirical papers discussing 'scarring' effects of early unemployment spells or displacements do not specify the precise economic mechanism behind permanent or highly persistent effects. A notable exception is Machin and Manning (1999), who discuss how true duration dependence could arise among the long-term unemployed. For workers of all ages see, e.g., Heckman and Borjas (1980), Arulampalam, Gregg and Gregory (2001). For young workers see, e.g., Ellwood (1982), Margolis, Simonnet, and Vilhuber (2000).

[14] Even under asymmetric information, continuous learning by employers would predict that eventually workers get paid their true marginal product. The idea of such strong signaling imposes long-term effects almost by assumption. That is not to say that imperfect employer learning (e.g., employers stop learning at some point) is not a realistic possibility. Machin and Manning (1999) suggest that statistical discrimination by employers could lead to such an outcome. A particular example was suggested by Blanchard and Diamond (1994) who construct a model in which employers rank workers by latest arrival into the pool of unemployed.

the initial job for future career outcomes. For example, in a model of 'stepping stone' human capital accumulation the first job provides access to human capital accumulation crucial for advancement to the next higher level (Jovanovic and Nyarko 1997). If there are entry restrictions into 'career' jobs leading to different paths of human capital accumulation, early shocks have permanent effects (e.g., Baker, Gibbs, and Holmstrom 1994, Gibbons and Waldman 2002).[15] This is particularly likely if firms provide different rates of experience accumulation on the job as suggested by Rosen (1972). Alternatively, labor markets within larger firms are said to provide well-defined career paths. Internal labor markets are said to protect workers from external market conditions and to restrict entry to specific jobs (also referred to as "ports of entry").[16] If such entry-level jobs are scarce, or if entry into internal labor markets is restricted by age, then workers leaving large firms are likely to permanently lose career prospects or rents associated with firm size. Workers exiting large firms are thus particularly at risk of persistent wage losses and will be analyzed separately in the empirical analysis.

2.2. Estimates of Wage Losses and Confounding Factors

If one ignores the model of Equation (1) and estimates a simple OLS regression of log real wages on a dummy D_{i0} for moving out of the training firm (leaving out other control variables for simplicity), the probability limit of the estimated effect on wages of a move out of the training firm at the end of training after t years is

$$p \lim \hat{\delta}_{It}^{OLS} = \frac{\text{cov}(a_i - \bar{a}_{j(i)}, D_{i0})}{\text{var}(D_{i0})} + \frac{\text{cov}(\bar{a}_{j(i)}, D_{i0})}{\text{var}(D_{i0})} + \delta_{It} + (\delta_{Vt} - \delta_{It}) \frac{\text{cov}(V_{i0}, D_{i0})}{\text{var}(D_{i0})}.$$

Negative selection implies that $\text{cov}(a_i - \bar{a}_{j(i)}, D_{i0}) < 0$, whereas initial assignment implies that

$\text{cov}(\bar{a}_{j(i)}, D_{i0}) < 0$. In both cases OLS tends to be biased toward finding a negative effect even if

---

[15] With a similar argument Okun (1973) suggested cyclical effects could have permanent adverse or beneficial consequences at the cohort level by pushing young workers into better or worse jobs, thereby affecting their human capital accumulation.

[16] Internal labor markets have been defined by Doeringer and Piore (1971), loosely speaking, as a set of institutions within firms that determine wages (not necessarily by productivity), shelter workers from outside market conditions, and have a clear connection to that labor markets through well-defined jobs ('ports of entry'). Baker, Gibbs and Holmstrom (1994) test several predictions of the internal labor market paradigm.

$\delta_{It} = 0$. On the other hand, search models predict that $\delta_{Vt} - \delta_{It} > 0$. Since $\mathrm{cov}(V_{i0}, D_{i0}) > 0$, this implies that OLS would tend to underestimate the true effect of an involuntary move from the training firm. Together with these confounding elements, the OLS estimate may also pick up true negative effects of a displacement implied by job search, sequential human capital accumulation, or institutional models. Clearly, without further information it is hopeless to disentangle the various pieces of information contained in the OLS-estimate and obtain the true effect of mobility.

The paper proposes an estimation strategy that allows both estimation of the true causal effect of an early job loss as well as an assessment of the biases and mechanisms underlying early job mobility. To solve the problems introduced by the presence of sorting into and selection out of firms the paper uses firm level data on the training firms of young German apprentices to implement the following two-tiered strategy. First, it uses firm fixed effects to control for systematic differences of workers between firms. Thereby, non-displaced workers at the same training firm function as comparison group for the wage-outcomes of displaced workers. Second, the paper uses firms' retention rates of other young graduates finishing apprenticeship in the same year as a displaced worker as an instrument for the probability of a displacement. To account for the fact that workers may sort into firms based on average retention rates (i.e., the career prospects firms offer), the preferred instrument will be the deviation of the retention rate from the firm's average. This isolates as closely as possible the group of workers who would not have moved under normal business conditions and thereby best approximate an exogenous displacement.

The retention rate of a firm is measured by the fraction of workers *other* than the young trainee in question that finished apprenticeship training in the same year who left the training firm. Thereby, we use the mobility behavior of other graduates in the same firm as a proxy for the individual trainee's probability of moving. Let D_{ijc} be a dummy variable denoting the event that worker i in graduating in cohort c leaves firm j. Then for each worker the fraction movers among other trainees graduating from the same firm during the same year is $z_{ijc} = m_{jc(-i)} / (n_{jc} - 1)$, where n_{jc} is the number of graduates at firm j in cohort c and $m_{jc(-i)} = \sum_{l \neq i}^{n_{jc}} D_{ljc}$ is the number of movers

among a young graduate's peers. Since a key point of the paper is that z_{ijc} may systematically differ

across firms, the final instrument used for the probability of moving will be the deviation of z_{ijc}

from its firm specific average $\widetilde{z}_{ijc} = z_{ijc} - \overline{z}_j$, where $\overline{z}_j = \frac{1}{C}\sum_{c=1}^{C}\frac{1}{n_{jc}}\sum_{i=1}^{n_{jc}} z_{ijc}$.[17]

The empirical strategy is to use within-firm changes in labor demand for young apprentices as

measured by $\widetilde{z}_{ijc}$ as an instrument for involuntary mobility. That changes in plant-level employment

demand are frequent, large, and heterogeneous has been suggested by Davis and Haltiwanger

(1990).[18] Similarly, the variation in retention rates of graduating apprentices at the cohort and plant

level in Germany is high: plant and cohort effects by themselves only explain 62% of overall

variation in retention rates. Moreover, as further discussed below, retention shocks are highly

correlated with an individual workers' propensity to move. The instrumental variable strategy is valid

if shocks to the retention rate are due to unexpected changes in labor demand and not correlated

with average ability of apprentice cohorts.[19] If changes in retention rate were driven by cohort

quality, they should not be correlated with employment changes at other age levels. Moreover, those

displaced should be of lower ability and thus have permanently lower wages. Neither is the case in

the present application: firm employment changes at all age levels have a significant positive

correlation, and the final estimate implies only a temporary wage loss from job displacement.

Confidence in the instrument is further strengthened by the fact that there is no systematic

correlation of observable characteristics of apprentice cohorts with retention rates controlling for

firm fixed effects.[20]

---

[17] The average is taken across cohorts and workers. In a full sample, this will be exactly equal to the average retention rate of the firm across cohorts. In the final sample, this won't hold exactly due to sample restrictions.

[18] These findings have been confirmed for other countries, e.g., see Bauer and Bender (2004) for Germany and Abowd, Corbel, and Kramarz (1999) for France.

[19] Dropping the cohort subscript, the assumptions necessary for the instrumental variable approach are $\text{cov}(a_i - \overline{a}_{j(i)}, z_{ij} - \overline{z}_{j(i)}) = 0$ and $\text{cov}(V_{i0}, z_{ij} - \overline{z}_{j(i)}) = 0$; since the main variation is at the establishment-cohort level, correlations at that dimension determine validity of the instrument.

[20] The systematic correlation of average turnover rates with sample characteristics shown in Table 7 disappears when controlling for firm fixed effects, i.e., the sample becomes balanced on observables. Ideally, we would have access to

A remaining concern is that variation in external labor market may induce changes in the fraction of workers leaving voluntarily, inducing a negative correlation between retention shocks and mobility. This is particularly relevant for firms with very few apprentices where the mobility of an apprentice might be directly influenced by the mobility decisions of individual colleagues. In part, the question will be resolved by considering the first stage. In addition, we restrict our sample to firms with a minimal number of graduating apprentices. To isolate demand side variation in employment, we also consider a second instrument (henceforth IV2), which treats as 'movers' only those workers who have a spell of unemployment of at least 30 days at the end of training. Since we are certain to exclude most voluntary movers, involuntary movers should drive most of the variation in the second instrument. Thereby, it should yield a valid second set of estimates and a useful sensitivity check on the approach.

Other studies on displaced workers have used plant-level changes in employment to identify unexpected shocks to labor demand. In particular, Gibbons and Katz (1991) use plant closing as an instrument for displacement. Their insight is crucial, since in the presence of asymmetric information pre-displacement wages are not informative about workers' ability. Thus, a strategy of worker fixed effects is not appropriate to control for biases from selection in an environment of asymmetric information. As suggested above, this is particular important for the study of young workers for whom wage histories not reflect productivity. Another seminal study using the fraction laid-off at a given firm as determinant of who is counted as displaced worker is Jacobson et al. (1993). Since they use a sample of mature workers with a long career history, this is done to exclude voluntary job movers rather than to control for selection. No study on displaced workers controls for the 'normal' level of turnover at a firm, and therefore these papers do not directly control for sorting that occurred prior to any 'shocks' to firms' employment. Here, we not only introduce firm fixed effects to control for permanent differences in average retention rates, but are also able to use occupation-specific labor demand shocks at the firm level to control for negative selection. This is

data on sales at the plant and year level. While survey information exists for part of the plants in the sample from the IAB-Establishment panel, sales, investment, and profit data have a high degree of missing observations.

the first study to use the identifying variation from continuous employment shocks at a very detailed demographic and occupational as explicit instrument for job displacement.

Besides delivering the true long-term effect of an early job loss, the comparison of various estimators provides important information on the biases underlying simple OLS estimators. Access to a matched employer-employee panel provides us with at least three additional estimators of the wage loss from displacement at our disposition to estimate the true effect of a displacement from the training firm: OLS with firm fixed effects (OLSFE), instrumental variables (IV1), and IV with firm fixed effects (IVFE1). Moreover, we present two additional IV estimates based on our second instrument (IV2 and IVFE2). To see the advantage of interpreting multiple estimators, suppose again that all confounding factors are present in the data, such that the process determining wages can be captured by Equation (1). The simplest of the alternative estimates, OLS with firm fixed effects (OLSFE) is identified by deviations from firm averages. The probability limit of the resulting estimate is

$$p\lim \hat{\delta}_{It}^{OLSFE} = \delta_{It} + (\delta_{Vt} - \delta_{It}) \frac{\text{cov}(V_{i0}, D_{i0})}{\text{var}(D_{i0})} + \frac{\text{cov}(a_i - \overline{a}_{j(i)}, D_{i0})}{\text{var}(D_{i0})}.$$

By only comparing workers who graduated at the same training firm, OLSFE accounts for the bias from initial assignment. Yet, it is still affected by negative selection and by voluntary mobility.

The next more sophisticated estimator is IV using the fraction of 'other' movers as an instrument (IV1). If there is no initial sorting, IV in levels identifies the true effect of involuntary mobility δ_{It}. However, if the least able workers are sorted into the firms with the lowest retention rate (the highest fraction 'other' movers), then we have that $\text{cov}(\overline{a}_{j(i)}, z_{ij}) < 0$. The resulting IV estimator is biased, i.e., the probability limit is

$$p\lim \hat{\delta}_{It}^{IV} = \delta_{It} + \frac{\text{cov}(\overline{a}_{j(i)}, z_{ij})}{\text{cov}(D_{i0}, z_{ij})}.$$

Note that since the denominator is now smaller than in the case of OLS $(\text{cov}(D_{i0}, z_{ij}) < \text{var}(D_{i0}))$, initial sorting could imply that $\hat{\delta}_{It}^{IV} < \hat{\delta}_{It}^{OLS}$, i.e., the IV estimator can be *more* negative than OLS.

15

Alternatively, this could occur if the effect of negative selection on OLS is more than offset by the positive bias from voluntary mobility.

To account for the remaining bias, the last step is to introduce firm fixed effects into the basic IV regression (IVFE). Since firm fixed effects now control for initial assignment the IV estimate should yield a consistent and unbiased estimate of the true effect of involuntary displacement, i.e., the probability limit is $p \lim \hat{\delta}_{It}^{IVFE} = \delta_{It}$. By using firm fixed effects, IVFE is identified by wage losses of workers moving because the retention rate at their firm was lower than average relative to similar workers within the same training firm who are at risk of moving in other periods. Those workers that never move or those that always move do not help to identify the estimate. In other words, if treatment effects are heterogeneous the resulting estimator is an estimate of the *local average treatment effect* for those workers induced to move by a temporarily low retention rate.[21] This is the relevant causal effect for those workers who are at risk of moving due to temporary demand conditions.

If there are no confounding factors, then OLS, OLSFE, IV, and IVFE should all yield similar estimates of the effect of moving out of the training firm. However, in the presence of selection and sorting at the firm level, only the IV estimator with firm fixed effects will yield an unbiased estimate of wage losses from an early displacement. An additional advantage of the chosen approach is that the difference between the various estimates be used to gauge the importance of the biases affecting the simple OLS estimator. Moreover, economic theory has separate implications regarding initial assignment and various forms of selection. Since these predictions translate differently into the various estimators, the stepwise estimation procedure gives a way to assess the relative importance of various mechanisms underlying wage and job changes.[22]

---

[21] This assumes that there are no defiers – i.e., workers that would have left the firm under normal business conditions but stay because the retention rate is low (Angrist, Imbens, and Rubin 1996).

[22] This is further elaborated while discussing the results. Mathematical derivations are presented in Appendix A available on the authors' website.

3. Data and Institutional Background

3.1. The German Apprenticeship system

The application in the present paper is concerned with the wage losses of young German apprentice leaving their training firm. Figure 1 depicts a stylized representation of the German apprentices system.[23] Two-thirds of young Germans follow an apprenticeship in the German "Dual System", during which they receive both formal state-sponsored schooling as well as training on the job. Most apprentices start training right after junior-high school, and the majority fully participates in the labor force at the end of the apprenticeship. Apprenticeships last on average two and a half years, after which about 40% of workers leave the training firm immediately. Training is mainly general and employment rates of graduating apprentices are very high, and these are the two features most often cited by proponents of large-scale apprentice systems in other countries. The institutional structure of the German apprenticeship system is ideal to analyze the persistence of early labor market shocks, since we can study the effects of a well-defined event (transition from training into the labor market at the same or at a different firm) for a large group of workers with relatively homogeneous labor market experience and background.

However, all three of potential confounding factors are potentially present in the German case. Employers are likely to learn about workers and try to retain only the best of them (Acemoglu and Pischke 1998). Mobility is high even for workers who stay at the training firm, suggesting that young graduates from the Dual System have other options and move voluntarily (Euwals and Winkelmann 1996 and Schwerdt and Bender 2003). Moreover, firms differ in their turnover rates and possibly in the quality of training they provide (Winkelmann 1996), and actively try to screen among applicants to their apprenticeship programs. Thus, the application of the theoretical and conceptual framework outlined above is appropriate for the German case. However, it is the

---

[23] A detailed account of the German education and apprenticeship system can be found in Franz et al. (2000), Winkelmann (1996), and the annual employment/qualification reports of the German government (Berufsbildungsbericht 2001).

availability of the relevant establishment-level data matched to detailed information on workers' careers that makes this exercise most exciting

Before describing the data, the paper will benefit from briefly establishing that, while Germany does clearly differ in its institutions, enough basic similarities in the labor market for young workers exist to make the results relevant for the understanding of U.S. labor markets.[24] we have used the administrative data to replicate several of the main results of Topel and Ward's (1992) seminal study on career patterns of young American workers. As in Topel and Ward's study, the sample consists of all men between 18 and 34 years who are in stable employment.[25] The results of the exercise suggest that while institutions and the degree of job mobility clearly differ, it is reasonable to suppose that, in the sense of Ryan (2001), the 'fundamental economic mechanisms' operating in the labor markets for young workers in the two countries bear some basic similarities.

The main points, shown in Table 1 can be summarized as follows. First, labor force attachment is slightly higher among Germans over the first two years of potential experience (the period during which a large fraction of German workers participates in apprenticeship training), but evolves similarly in Germany and the U.S. afterwards (Panel A of Table 1). Second, while Americans do transit through more jobs in the first years of the labor market, job attachment is similar for jobs lasting at least six quarters.[26] Third, Panels B and C of Table 1 show that both wage growth within and between employers is important and of similar magnitude in Germany and the US.[27] Fourth, it is the first years in the labor market that matters most in both countries; in the US, two thirds of earnings growth occurs in the first ten years of the career (Murphy and Welch 1992). In Germany,

---

[24] Critics of policy proposals based on the German apprenticeship scheme argue that the system is fundamentally linked to German labor market institutions absent in the US. For an overview of the arguments for and against establishing large-scale apprentice programs in the US, see Stern et al. (1994) or Heckman et al. (1996).

[25] The study is based on a 1% sample of the data set used in the main analysis provided by the Institut fuer Arbeitmarkt-und Berufsforschung (IAB). The data, sample used, and approach are described in the notes to the tables and figures in Appendix C available from the authors' website.

[26] The difference in the overall hazard of job leaving is concentrated in the first six quarters of job attachment. This pattern is driven from differences in job-to-job transition rates; job-to-non-employment transition rates are similar. Similar differences in mobility rates among young workers are reported by Ryan (2001).

[27] Average within job wage growth for a restricted sample with at least six quarters of tenure in the US is 7.1%, average completed job duration is 7 years, and average rate of job change is 28% (Topel and Ward 1992, Table VI). The corresponding values in Germany are 9%, 7.3 years, and 25%. Average wage growth occurring between jobs (controlling for experience) is 19.9% in the US (TW, Table VII) and 20.6% in Germany.

counting the apprentice period, 41% of earnings growth occurs in the first 5 years, and 80% of earnings growth occurs in the first ten years.[28] Last, as will be seen in the next section, raw wage losses from leaving the training firm are a similar order of magnitude of wage losses of young displaced workers in the US.[29] In addition, Harhoff and Kane (1997) and Blau and Kahn (1997) find that apprentices occupy a similar position in the wage distribution, have similar wages, and similar wage-experience profiles as high-school graduates in the US.[30] Thus, while specific numerical results may not be directly transferable across labor markets, this comparison suggests that the mechanisms of early job and wage dynamics and the potential biases of estimation methods are similar enough for the current approach to provide useful insights into the study of young job losers' careers in the U.S..

3.2. German Social Security Data

The data used in this paper are drawn from the German employment register containing information on all employees covered by social security, representing around 80% of the German workforce.[31] The notification procedure for social security requires employers to record any permanent or temporary change of employment relationships, and in addition takes stock of existing employees at each establishment twice a year. Therefore, the employment register contains detailed histories for each worker's time in covered employment. Besides period of coverage, the key information contained in the register for administrative purposes (and therefore the most reliable) are gross daily wages subject to social security contributions. Contributions have to be paid only up

---

[28] If wage growth during apprenticeship is excluded, 70% of wage growth occurs in the first 10 years.
[29] No estimates for *young* displaced workers exist so far for Germany. As shown by Couch (1996), wage losses for mature displaced workers are similar to wage losses of workers in the US.
[30] See Harhoff and Kane (1997) Table 5 and Figures 3 and 4, or Blau and Kahn (1997) Figures 1b and 2b. The main difference between the US and Germany is the labor force participation rate and unemployment rate of very young workers (Ryan 2001, Blau and Kahn 1997). These results are consistent with the idea that the main impact of the German apprenticeship system is to reduce unemployment rate among young workers.
[31] The employment register was established in 1973 to integrate the notification procedures for health insurance, pensions and employment insurance. The data is described briefly in Bender, Haas, and Klose (2000) and in more detail in Bender et al. (1997). Coverage includes full- and part-time employees of private enterprises, apprentices, and other trainees, as well as temporarily suspended employment relationships. The self-employed, civil servants, and students are excluded.

to a limit, but top coding is very rare for younger workers. In addition, the data contain basic demographic information (gender, age, nationality) as well as information on occupation, industry, job-status, and education.[32] Most important for the present purpose, the data also contain unique establishment identifiers. These were used to create a separate data set of establishment characteristics that were aggregated up from the employment register and merged back onto the individual level data. Characteristics include among others establishment size, employment growth, number of graduating apprentices, and average wages. The relevant entity throughout the empirical analysis is thus the establishment. Despite the inaccuracy it entails in some cases, we will keep using the terms establishment and firm interchangeably for the rest of the analysis.[33]

The working sample consists of information on the universe of trainees graduating from an apprenticeship in 1992 to 1994 in West Germany drawn from the employment register. The timeframe is chosen such that several cohorts can be observed for at least 5 years after entering the labor market. (Before 1992 exact date of apprentice graduation is not identifiable in the data.) For each graduating apprentice the sample contains information on the establishment where training takes place (size, employment growth rate, number of apprentices graduating, number of graduating apprentices staying at the firm, total number of apprentices, average training wage, average overall wage), on training itself (duration in days, training wage, occupation, industry), and basic demographic characteristics (gender, age, education prior to training). Moreover, for each apprentice the sample contains daily gross wages for the first five completed years of potential labor market experience after the end of training. We also know whether apprentices have spells of non-employment and will use it to make further sample restrictions in the sensitivity analysis.

To ensure the sample consists of 'core' apprentices, it is restricted to occupations participating in the "Dual System" (i.e., other vocational training is not included), it requires a minimum length of continuous training of 450 days, and it excludes workers who have prior labor market experience,

---

[32] The entity reporting is the establishment for which an employee works and can thus change over time. This can lead to mistakes in the coding of some demographic variables (e.g., nationality or marital status) and in particular education (which tends to reflect required rather than actual qualification).

[33] Unfortunately, it is currently not possible to link establishments that belong to a common parent firm.

who have more than one apprenticeship spell, and who are older than 30 at the end of training. To make the study of retention rates useful, an additional crucial restriction is that establishments with less than 50 covered employees and less than 5 graduating apprentices in a given year are excluded from the sample. A large fraction of apprenticeships occur at very small establishments, such that this restriction reduces our sample by roughly 50%. This limits the representativeness of the sample with respect to the German apprenticeship system as a whole. On the other hand, very small training firms seem to follow different incentives than larger firms (Winkelmann 1996), and the economic mechanisms studied in the paper are more likely to apply to larger firms.[34] Finally, workers are required to have a minimal amount of attachment to covered employment (i.e., they must have at least one appearance in covered employment after their third year of potential experience) and daily real wages are required to be above 30DM in 1996 prices (about $15).

Table 2 shows the basic characteristics of the sample for all graduating apprentices in the final sample (Column 1) as well as separately for workers moving and staying at the training establishment the day after the end of training. The main sample consists of 295653 observations on graduating apprentices. Since it is restricted to larger training firms, the sample is slightly older, slightly better educated, and has a higher fraction men than the full sample of apprentices. A high fraction of the sample is concentrated among very large firms (with more than 500 employees) as these have larger training programs. Most training lasts at least two years, but longer training spells are not uncommon. The fraction moving from the training firm is 40%, which is slightly lower than in the raw sample and slightly higher than tabulations from the German Socio-Economic Panel or Qualifications- and Careers Survey suggest.[35] This could be due to the fact that the current sample counts even brief separations from the training firm as moving and that the sample is more recent.

---

[34]For the same reason, Acemoglu and Pischke (1998) also focus on firms with at least 50 employees. The comparison of training programs and the fate of apprentices graduating from larger vs. very small training firms is interesting in its own right and is a question for future research. Very large training firms will be discussed separately below.
[35] The raw tabulations are similar to samples from other data sets (e.g., see Winkelmann (1996) for the German Socio-Economic Panel, Acemoglu and Pischke (1998) for the Qualification- and Career Survey of the Bundesinstitut fuer Berufsbildung (BIBB), and Euwals and Winkelmann (1996) for the IAB Employment Subsample), with the main exceptions noted in the text.

The standard deviation of fraction 'other' movers and of training firms' annual growth in overall employment is high – a first indication that there is a high degree of variation in firm characteristics within the sample. The training wage, set by collective bargaining, is very low. Although only half the size of the first wage it is higher than apprenticeship standards due to the focus on larger firms.

4. The Long-Run Effects of Leaving the Training Firm

Mobility at the end of apprenticeship training is high. Table 2 also shows that mobility is not random – movers are systematically different from stayers. Those who move at the end of training tend to be less educated, more likely to be male, and more likely to be trained at smaller firms. Movers also receive lower training wages.[36] Moreover, they work at firms that pay slightly lower training wages, that have lower employment growth, and that have a much higher average fraction of their apprentices moving. This is suggestive of the strong correlation between the probability of moving and the firm's retention rate of apprentices exploited below. Movers tend to be concentrated in the service sector, transport and communications, and are more likely to be blue-collar workers. Movers are more likely than stayers to graduate in 1994 – a recession year.

Raw wage differences, controlling for cohort and experience effects and their interaction, are shown in the first column of Table 3. Movers have 9-10% lower wages than stayers, and this difference is basically unchanged after 5 years in the labor market. The remaining columns of Table 3 try to explain this persistent difference by controlling for observable characteristics. The data consists of an unbalanced panel of apprentices observed annually during the first five years of potential labor market experience since the end of their apprenticeship. Thus, for the purpose of estimation the observations are stacked into a panel, and all estimates are obtained from the stacked model. Since error terms will be correlated across individuals and potentially also within training

---

[36] Note that on average workers with lower training wages have higher wages in the labor market, i.e., there appears to be mean reversion. Thus, without controlling for mean reversion movers will appear to have slightly higher wage growth than stayers relative to training wages.

firms all standard errors are clustered at the level of the training firm.[37] In case of OLS with training

firm fixed effects, the fixed effects are restricted to be the same across periods. As further discussed

below, changes in the sample composition occur due to military service or exit into other forms of

employment. To ascertain that these changes over time do not affect our results, we run all our

specifications on both balanced and unbalanced panels in the sensitivity analysis.

Table 3 shows the differences in real wages between movers and stayers after controlling for

the characteristics of the worker, the firm, and of the apprenticeship. All regressions also include

interactions between cohort and experience dummies, which effectively controls for year effects.

Including individual worker characteristics such as gender, prior education, or nationality (Column

2) does little to change the effects. Adding employment size and employment growth rates of the

training establishment reduces the differences only by about 1% (Column 3). The gap is significantly

reduced to about 7% when the individual log real training wage, dummies for training duration, and

a dummy for whether movers work at their training firms are included (Column 4). Since we cannot

include individual workers' fixed effects, including training wages partially controls for productive

ability (in a limited fashion as training wages are not necessarily reflective of workers' skills). The

coefficient on training wage, shown with other selected coefficients in the Appendix Table, is only

about 0.2, confirming the suspicion that it is an imperfect control of ability.[38] Conditional on

characteristics of apprenticeships and training establishments, training occupation and industry result

in only a small decline in the differences (Columns 5 and 6). The preferred OLS specification used

throughout the analysis is that of Column 5, which excludes industry controls. Training industry is

another dimension along which sorting and selection could occur, and is thus further discussed in

the sensitivity analysis. The evidence in Table 3 strongly suggests that while the overall difference in

---

[37] The cluster is not interacted with period. It therefore includes all observations on an individual and takes care of cross-individual correlation as well. Given that the regressors are the same across the five periods, in case of OLS it would be equivalent to estimate the model separately for each period. However, if there are cross-equation restrictions as in the case of OLSFE (or IVFE), estimating the stacked model is necessary to obtain correct estimates of the standard errors (Ruud 2000, p.703).

[38] Note that two thirds of the difference in training wage among stayers and movers is explained by differences among their establishments. Thus, differences the explanatory power from training wages partly comes from permanent differences among training establishments.

wages among movers and stayers declines significantly when additional characteristics are controlled for, it remains significantly negative and stable over time at all specifications.

As discussed in Section 2, these estimates may not reflect true 'causal' effects. If they are due to initial sorting of less able workers into establishments with lower retention rates (and lower wages), then controlling for firm fixed effects should solve the problem, since they force the comparison to be done relative to workers trained at the same firm. The wage differences among movers and stayers after controlling for firm fixed effects alone are shown in the third column of Table 4. The results suggest that firm fixed effects alone can go a considerable way in explaining the raw difference in wages. Thus, part of the individual and training characteristics in the previous regressions may simply pick up sorting among firms. However, adding firm fixed effects to the full OLS specification including training occupation in Column 2 of Table 4 does not further affect the differences in wages (regression not shown). Thus, sorting alone does not appear to be able to explain the remaining difference between movers and stayers.

Another explanation for the remaining wage differences is that movers are negatively selected with respect to stayers. To control for the possibility of negative selection, we use the fraction of movers among other apprentices graduating at the same firm in the same year as an instrument for the probability of moving. As discussed above, if adverse selection is the main source of wage losses, then workers leaving firms with higher turnover rates should be less negatively selected. Table 5 shows the first stage regressions of a dummy for leaving the training firm at the end of apprenticeship on the instrument and a rich set of observable characteristics on individuals, firms, and apprenticeship training. To implement the IV estimators, we chose to estimate five separate first stages for the five periods in a model of seemingly unrelated regressions, and use the resulting coefficients on the instrument to obtain the IV estimator. This ensures that the samples from which we estimate the first and second stage coefficients are exactly equal and avoids the problems of two-

24

sample instrumental variables.[39] Column 1 of Table 5 shows that the fraction 'other' movers is a strong determinant of the probability of moving. The coefficient on the instrument changes somewhat across experience years in response to changes in the sample composition, but overall is quite similar (however, given the sample sizes, most of the differences are statistically significant). As shown in the panel-sample of the sensitivity analysis, these differences do not measurably affect the results. The F-statistic for a test of joint exclusion of the instruments does not indicate a problem of weak instruments.

The IV estimates, obtained by dividing the coefficients from the reduced form with those of the first stage, and the correct IV standard errors, clustered at the establishment level, are shown in the fourth column of Table 4 (selected coefficients on other variables are shown in the Appendix Table). The estimated wage differences among stayers and movers become *more* negative than the basic OLS estimates and are on the order of magnitude of the 'raw' wage differences. In addition, using the level of firms' fraction 'other' movers as instrument one still obtains persistent differences over time – from the first to the fifth year of potential labor market experience, the difference declines only by 2.3%. Thus, adverse selection alone cannot seem to explain the observed wage differences. On the other hand, the increase in the gap among movers and stayers is again consistent with the hypothesis that less able workers sort themselves into firms with lower retention rates, or that these firms provide training of lower quality.

Both the results from OLSFE and IV suggest that workers sort themselves into firms according to turnover rates. However, the fact that OLSFE is still negative suggests that firms also selectively displace their worse workers. To control for permanent differences *between* firms as well for differences among movers and stayers *within* firms we add firm fixed effects to the basic IV model. This means the instrument for the probability of moving is now the deviation of the fraction

---

[39] If the panel were balanced, to obtain IV estimates we would estimate a single first stage for the endogenous variable (mover status at the end of training) and five reduced form equations. The latter again form a SUR (or restricted SUR in the case of IVFE). The IV estimate would then be obtained by dividing the reduced form by the first stage coefficients. However, since the sample changes between periods, this approach would be akin to a two-sample IV estimator in which the samples used to obtain the first stage and reduced form coefficients are not independent (Angrist and Krueger 1992). Therefore, particular care had to be taken in estimating the standard errors (Murphy and Topel 1985).

of movers in a cohort from the firm specific average across cohorts. As discussed above, by focusing on movers who under normal business conditions would have stayed at the firm, this should mimic the event of a random displacement and should be free of a bias due to adverse selection. Moreover, firm fixed effects control for sorting by comparing workers induced to move by a firm-specific shock to similar workers graduating from the same firm. The first stage is again shown in Table 5 (Column 2). The coefficients estimates on the fraction 'other' movers are now smaller than in Column 1, but still highly significant and of more reasonable size.[40] The F-test statistics for the hypothesis of joint insignificance of the instruments are well beyond the critical level of 10 suggested by Stock and Watson (1997). Weak instruments do not appear to be a problem in the present application even in the presence of establishment fixed effects.

The final IV estimates are shown in the last column of Table 4. Including establishment fixed effects, the estimated wage difference among stayers and movers in the first year after entry into the labor market is -.108%. However, using firm fixed effects the difference is not persistent and decreases to –.035% and .009% in the third and fifth year of the labor market, respectively. These are clear signs of a 'catch-up' of movers towards wage levels of stayers. While the estimate after three years is still significantly different from zero (p-value of 8%), the estimates after four and five years are not. Unfortunately, the standard errors on these estimates do not allow excluding a remaining negative effect. However, as shown in the bottom panel of Table 4, the estimates after three and five years are significantly different from the initial gap and significantly different from each other at a 1% significance level. Unbiased estimates of the effects of moving involuntarily from the training firm show initially large losses that fade to zero within the first four to five years of the labor market.

---

[40] Consider a firm with 20 apprentices. Suppose the average fraction of workers who move after apprenticeship training from the firm is 40%. If instead of 12 workers the firm only retains 4, that fraction rises to 80% (=16/20). Since the coefficient in the first stage is roughly 0.15, a temporary increase in the fraction 'other' movers of 40% implies an increase in the probability of moving of 6% (an increase of 15% relative to the baseline). If the coefficient is 0.75 as in the case of IV without fixed effects, the implied increase is 30% (an increase of 75% relative to the baseline).

Figure 2 plots the basic estimates of the effect of moving in a single graph. The main patterns shown in the figure, which will repeat themselves throughout the empirical analysis, can be summarized as follows:

- OLS differences are smaller than raw differences, but persistent and significant

- OLS with only firm fixed effects reduces difference as much as all other covariates

- IV without fixed effects tends to imply larger differences than OLS

- IV with firm fixed effects leads to larger initial differences than OLS and a significant catch-up.

These results suggest that controlling for both permanent differences across firms and firm-specific shocks is crucial in the analysis of the effects of worker mobility. There seems to be both sorting of workers between firms and sorting of workers within firms, such that neither simple IV estimates nor firm-fixed effects alone can sufficiently account for the underlying selection mechanism. Once this is done, the wage differences between movers and stayers are no longer permanent but show strong signs of catch-up. Moreover, consistent with separate results on the role of mobility in the German labor market, the fact that the difference estimated by IVFE is initially bigger than the OLS estimate suggests that a large portion of mobility is voluntary. Simple estimates miss this distinction as they group several types of workers together. By averaging over positively and negatively selected workers they tend to underestimate the initial effect of moving and overstate the degree of persistence.

To see that these results are robust and not driven by some peculiarities of the data or by choice of particular specifications, it is useful to consider some simple graphs of the reduced form. Controlling only for experience-cohort-effects, Panel A of Figure 3 shows the simple average of real wages by intervals of the fraction 'other' movers (the instrument in levels) for the three experience years. The three panels show that the linearity assumption underlying the results in Table 4 is justified.[41] Moreover, the relationship is negative and does not change over time. Panel B of Figure 3

---

[41] The brackets are defined as {[0-.1),[.1-.2),[.2-.3), …, [.9,1), '=1'}, i.e., the 11th 'bracket' is for fraction 'other' mover equal to one. The graphs also show regression lines of regressions of the averages weighted by their standard errors on a

shows the same graph but now controlling for firm fixed effects. This is the average *deviation* of wages from firm means by twelve brackets of the demeaned instrument. Again, linearity cannot be rejected, and for the first experience year there is a clear significantly negative relationship – larger shocks to the firms' average retention rate induce higher wage gains or losses. However, the estimated slopes rotate around zero from being significantly negative to almost a flat line in the fifth period.[42] Thus, those apprentices that graduated from firms shedding more apprentices than usual have only temporarily lower wages than they could have expected based on average firm outcomes. These patterns reinforce the finding that for graduating German apprentices negative early career shocks fade over time.

4.1. Sensitivity – Measurement and Sample Decomposition

The results of Table 4 and Figure 2 are robust to several important specification checks. The IV estimates with fixed effects for the alternative specifications are shown in Table 6. Figure 4 displays the full range of estimates for selected specifications. An important concern mentioned at the outset is that the instrument as defined might still include some variation due to voluntary movers. Although it does not affect the IV estimates per se since predictive power and sign of the first stage are as expected, it could introduce some measurement error in the instrument for smaller firms (where job offers and mobility of single workers may influence each other). To gauge this possibility, the first panel of Figure 4 shows the fraction of movers by the same intervals of the fraction 'other' movers as before (this is the analogue to Panel A of Figure 3). As the results from the first stage regression in Table 5 (Column1) suggest, there is a strong positive relationship. Panel B of Figure 4 shows the same figure for deviations from firm means (the analogue to Panel B of

---

constant and a linear trend. The slopes are all significantly different from zero but not significantly different from each other. Moreover, the linearity assumption cannot be rejected by a simple Chi-Squared goodness of fit tests.

[42] The brackets are defined as {[-1,-.5),[-.5,-.4),[-.4,-.3), [-.3,-.2), [-.2,-.1), [-.1,0), [0,.1), [.1,.2), [.2,.3), [.3,.4), [.4,.5), [.5,1)}. The distribution is concentrated among small deviations around the mean, but there are still a sizable number of large positive and negative deviations. The estimated slope coefficients from a regression of the cell-averages on a trend and a constant weighted by the inverse of their standard errors are -.0044 (.0008), -.0024 (.0007), and -.0015 (.0008) for experience years 1, 3, and 5, respectively. Intercepts are .022 (.005), .012 (.005), and .007 (.005). The changes from year 1 to year 3 and 5 in both slopes and levels are statistically significant at least at the 10% and 5% level, respectively. Chi-squared goodness of fit test cannot reject the hypothesis of linearity in either year.

Figure 3). On average, a higher 'shock' to the fraction 'other' movers leads to higher than average probability of moving, and this is what the first stage in Table 5, Column 2 picks up. However, the assumption of linearity works less well in this setting – for small deviations of the retention rate around the mean (between –10% and 10%) the relationship seems to be negative. To exclude this variation from the analysis, we use the second version of the instrument based only on those movers who spent at least 30 days out of covered employment (but not less than 10 months to exclude military leavers). Albeit this is too crude an approximation to capture all involuntary movers, it is likely to exclude most voluntary movers from the sample. Using this definition, the relationship between the fraction 'other' movers and the likelihood of moving is again close to linear even controlling for firm fixed effects (Figure 4, Panel C). This suggests that the non-linearities seen in Panel B of the same Figure are indeed due to voluntary mobility. The relationship for large firms shown in Panel D confirms that is a phenomenon affecting small firms.

Using the more narrowly defined instrument, one obtains similar results as with the main instrument. Columns 3 and 4 of Table 5 displays the first stage coefficients and Figure 4 (Panel A) and Column 1 Table 6 show the main estimates. The basic patterns shown in Table 4 and seen in Figure 2 are clearly confirmed. Those movers coming from firms from which a high fraction of workers move and spend some time in non-employment now have much more negative wages then stayers (i.e., IV is more negative). This is expected since those firms always releasing a high fraction of workers into unemployment are less desirable employers and likely to attract less able workers. However, once we control for permanent differences in wages and fraction movers across firms these estimates are reduced and the wage losses of movers decline even more strongly than before. The bottom panel of Table 6 again shows that the decline in the effect over time is again highly significant. Consistent with the suspicion that the original instrument was affected by variation of voluntary mobility, the initial effect is now significantly more negative (from -.108 to -.182). The basic IVFE estimator should thus be best understood as lower bound of the initial effect of involuntarily leaving the training firm.

Another concern expressed above is that changes in the sample decomposition might induce part of the observed reversion of losses. This might happen for example if among movers the best workers leave for the military or if the worst workers sequentially drop out over time. Note that since the OLS estimates are stable, this had to occur only for those induced to move by the deviation of firms' retention rates of trainees from average. To address this problem, we first restrict the sample to include only workers who have valid wage observations for each period in the sample. Panel B of Figure 4 shows the main results for this sample, while Column 2 of Tables 6 contains the final IV estimates. The figure shows that for this sample of workers with high labor force attachment the OLS estimates of the wage difference among movers and stayers is smaller than for the full sample. Moreover, the initial decline estimated by IV with fixed effects is larger, suggesting that among these more stable workers there is an even bigger fraction of voluntary movers. Again only the IV estimator with establishment fixed effects suggests initially large but steadily declining differences. Catch-up is again strong and complete within four to five years, confirming the initial results. If we repeat this analysis with the narrowly defined instrument (Figure 4 Panel C, Table 6 Column 4), as expected we again find larger initial losses and a stronger effect from sorting. Interestingly, we now also find some overshooting in the fifth period. This might be due to sampling variation. It would also be consistent with Acemoglu and Pischke's (1998) hypothesis that training firms in Germany have monopsony power due to asymmetric information.

One might argue that the panel-sample is overly restrictive in that it is likely to exclude many men leaving for military service. As an alternative restriction on labor force attachment that does not automatically exclude those in military or social service we therefore restricted total time spent out of covered labor force to be at most 6 months per year. The corresponding estimates are shown in Panel D of Figure 4 and Column 3 of Table 6. The results are again striking – while all other coefficients predict only slightly changing wage differences among movers and stayers, the IV estimate shows a larger initial difference with a strong following decline towards zero. The estimate is again close zero, although as before the standard errors do not exclude some positive or negative effects. The alternative IV in Column 5 again shows large initial effects, catch up, and overshooting.

Another interesting feature of the German data is that it allows us to identify those movers who later return to the training firm. These are included in the main sample since the probability of returning to the training firm after a temporary shock is a valid determinant of the expected cost of job loss. However, the presence of recalls could lead to underestimation of the true effects of an actual job loss if firms systematically respond to a short-term negative demand shock by temporary layoffs.[43] One can indeed show that retention rates are positively correlated with the probability of returning to the training firm even when controlling for fixed effects. Thus, the main specifications all include a dummy for whether a mover works at the training establishment. But recalled worker not only have higher average earnings, their wage losses (i.e., the slope) differ systematically as well. Column 6 of Table 6 shows the basic IV estimate with fixed effects if all workers returning to their original training firms are excluded. Compared to Column 5 of Table 4, the estimate is initially more negative, confirming the presence of some form of temporary layoffs that respond systematically to temporary demand shocks. However, the loss again reverts to zero within 5 years of labor market experience. Using the alternative instrument (Column 7) the initial effect is much more negative but with a higher standard error; as before, even these strong effects quickly revert back to zero. While reversion remains a key feature of the data, these results again suggest that the basic IVFE estimate may be best interpreted as a lower bound for initial wage losses.

The main results have been obtained by using training firm fixed effects. However, there might be shocks to retention rates of young trainees along other dimensions, most notably sectors and geographical regions. Instead of taking firm fixed effects, we also conditioned alternatively on industry effects, region effects, and their interaction. The basic result is unchanged: wage differences between movers and stayers in all the estimated models are again permanent (not shown). It appears that only deviations of retention rates from firm means are able to isolate unselected involuntary movers. This means that employment shocks have an important firm-specific component in

---

[43] In U.S. survey data such as the NLSY or the Current Population Survey's Displaced Worker Survey (DWS) workers are explicitly asked whether they lost their jobs due to temporary layoffs. While these are commonly excluded from the pool of displaced workers, those who did not consider themselves temporary layoffs and return to their original employer are not (in the DWS no further information usually indicates whether that is the case).

addition to common impulses at the region or industry levels. As an important corollary, it also implies that the initial decline in earnings and reversion following a move out of the training firm is not simply due to the entry into a temporarily depressed regional labor market.

The results suggest that workers sort themselves into firms according to average retention rates of apprentices at the end of training. Table 7 confirms that average turnover rates are an important property of training establishments that is systematically related to other basic firm characteristics.[44] The table displays additional average characteristics of training firms at high, medium, and low average retention rates. Training establishments with lower average retention rates also differ in basic observable characteristics such as size or average training wages. Firms with higher average retention rates also seem to be more attractive employers after apprenticeship training, providing higher wages and longer job attachment for workers staying after the end of training.[45] Moreover, it can be shown that there is a linear monotonic relationship between most of the characteristics shown in Table 7 and the fraction of workers leaving a training establishment at graduation. Thus, if workers choose where to do an apprenticeship taking into an account their career prospects and firms screen for workers, turnover rates are a principal characteristic by which sorting will occur.

While the regional dimension plays only a small role, part of the permanent differences between firms can be explained by industry differences. For example, effects for 3-digit industry explain 17.4% of variation in firms' retention rates between successive cohorts; the respective number is 21.7% for training wages and 30.1% for wages one year after the end of training. Fixed effects for training establishments explain 56.1%, 44.2%, and 42.9% of variance, respectively. Explanatory power of industry effects is relatively low for retention rates and training wages, but high for wages in employment. This is not surprising in an environment of collective bargaining.

---

[44] That turnover rates and wages are related has been suggested by the literature on efficiency wages, and Krueger and Summers (1988) provide some supporting evidence. Neal (1998) shows how sorting could explain the same phenomena.
[45] Note that among firms with 50 or more employees unionization rates and coverage by regional or national wage-agreements are high (90% and above); differences in industrial relations cannot be used to explain differences in turnover rates.

32

Consistent with these numbers, 3-digit industry effects can explain part of the OLS differences in wages of movers and stayers. Thus, while industry shocks cannot help to control for selection, they may help to account for part of initial sorting in the German labor market. A full analysis of the role of sorting, industry, worker, and firm fixed effects for mature workers in the style of Abowd and Kramarz (1999) for Germany still stands out.

4.2. Sensitivity – Large Firms

Large firms may provide exceptional career chances to young workers and pay higher wages to all their employees (e.g., Brown and Medoff 1989, Oi and Idson 1999). The question whether mobility has different effects for those leaving large firms is thus of particular interest. We therefore restrict the sample of apprentices to those 54% who graduate from establishments that employ at least 500 workers. While this is not representative of the German apprenticeship system as a whole, it is representative of large training programs that other countries have sought to emulate. With respect to the full sample, among graduates from larger firms one finds fewer women, slightly longer training durations, higher training wages, and a smaller fraction of movers.

The full analysis is repeated as before, and Table 8 summarizes the results. The first stages are strong both with and without fixed effects (not shown).[46] However, the main results imply some important differences with respect to the full sample. First, the effect of moving is now more negative in the main OLS specifications (Columns 2 and 3). The explanatory power of observable characteristics is weak, and weaker than that of establishment fixed effects – there appears to be more homogeneity in other observable characteristics than in turnover rates among large firms. These large negative effects are a first indicator of possible losses of firm-size wage-premiums since movers from large firms often find jobs at smaller establishments. Second, the simple IV estimate (Column 4) is now more negative than the raw differences, suggesting that initial sorting among larger firms with different average retention rates is stronger than in the full sample. Third, while the

---

[46] The first stage is shown at brackets of the demeaned instrument in Figure 4, Panel D. As suspected, for large firms the graphs show no significant non-linearity for small deviations of the instrument in the mean.

estimated wage differences from IV with firm fixed effects show a clear sign of decline over time it remains significantly negative even after five years in the labor market (Column 5).[47]

These basic patterns recur if we restrict the instrument to be calculated with movers initially exiting the main labor force (the 'narrow' definition of the instrument in Column 6), if we restrict workers to be in the sample all periods, if we impose stronger restrictions on labor force attachment, or if exclude those recalled to the training firm (not shown). This suggests that for workers graduating from large establishments initial luck seems to matter even after five years in the labor market. If large firms provide a special career-environment, additional wage loss of movers with respect to stayers should be driven by a decline in the size of the employing establishment relative to the training firm. To gauge this possibility, Column 7 includes the size of the current establishment interacted with experience as an additional regressor in the IV specification with firm fixed effects. Since this is only correct if size of the current employer is not endogenous, the results should be only taken as indicative. With this reservation in mind, it appears that size of current firm has the potential to fully explain the permanent effects of job loss found for workers exiting from large firms. Movers leaving larger establishments are likely to switch to smaller firms (not shown) and thereby lose their firm size wage premium. Controlling for both size of training and employing firm the initial losses are smaller than before but still significantly negative, and revert to zero within four years of labor market experience. This conclusion is consistent with results found by Krashinsky (2002) who first pointed out the potential of firm size premia in explaining losses of displaced workers.[48] It further corroborates the importance of controlling for the characteristics of the firms displacing job losers suggested by the main results of the paper. The results also suggests that there are permanent losers from early job losses. Large training firms pay higher wages, and those unlucky to lose a job at such a firm suffer permanent wage losses.

---

[47] The basic IVFE estimator also implies a lesser degree of voluntary mobility, suggesting that there may be fewer voluntary movers out of large training firms. However, the alternative instrument shown in Column 6 is again more negative than OLS.

[48] These results are also consistent with Dustmann and Schoenberg (2002) who discuss differences in OLS estimates of wage losses by firm size in the German apprenticeship system.

5. Interpretation – Sorting and Job Search

Different theories of early job and wage dynamics have separate implications regarding confounding factors arising from initial assignment and selection. As shown in Section 2, these biases affect the various estimators to different degree; thus, the stepwise estimation procedure allows us to assess the relative importance of various mechanisms underlying wage and job changes. Since we do not have an instrument for voluntary mobility, we cannot separately identify all bias components of the OLS estimator. However, a qualitative comparison of theories and estimates suggests several recurring patterns.

First, the role of training firm fixed effects suggests an important role of initial assignment of young workers into training firms. Firms with permanently low retention rates appear to be less attractive and attract less able workers. This result adds to a growing literature based on matched employer-employee data sets suggesting that workers are sorted into firms based on their ability (Abowd et al. 1999, Margolis 1995). It also extends the recent literature on heterogeneity in firms' growth and turnover rates (Davis and Haltiwanger 1990, Abowd et al. 1998, Anderson and Meyer 1994). First, it exploits temporary employment shocks in its identification strategy and demonstrates how they can be an important determinant of young workers' careers. Second, it shows that permanent heterogeneity in turnover rates leads to sorting of workers among firms and thereby confounds simple estimates of job losses. An important open question for future research is how much of the average difference in workers' productivity among firms with different turnover rates arises because of differences in training quality.

Second, the pattern of the preferred estimate, IVFE, can most readily be explained by a simple search model. A model of job search predicts an immediate loss of the rents accrued from the initial search for a good training firm; it also implies a gradual reversion of initial losses as workers again search for better jobs over time. Moreover, the fact that the effect of a random displacement is initially more negative than what is estimated by OLS and that the second instrument yields larger initial effects both indicate that there is voluntary job mobility among young apprentices, again

consistent with job search. This explanation is highly consistent with the role of job search found for young American workers by Topel and Ward (1992), and matches related evidence on young apprentices in Germany (Euwals and Winkelmann 2001) and on young German workers in general (see Table 1). However, search models also predict a decline in wage losses estimated by OLS, OLSFE, or IV, and this does not appear to be the case in the full sample. This is because workers displaced from a 'random' displacement are a minority among all workers leaving their training firm, and therefore receive a low weight among estimates pooling all groups of workers.[49]

Third, the high degree of negative selection implied by the permanent negative estimates of the various OLS models (with and without fixed effects) are consistent with an important role of adverse selection. Thus, while recent estimates suggest asymmetric information may matter less for older workers, Gibbons and Katz's (1991) hypothesis appears relevant for younger workers. This is also consistent with Acemoglu and Pischke's (1998) analysis of the German apprenticeship system. A process matching workers to firms as in Acemoglu and Pischke (1998) would have to be invoked to explain the coexistence of a voluntary mobility and firms' monopsony power induced by adverse selection. An alternative model predicting negative selection in the absence of asymmetric information is one where initial assignment is imperfect and occurs gradually over time. However, sequential sorting suggests that better workers move out of firms with high turnover towards more desirable jobs. Since IV tends to be more negative than OLS rather than less, a prediction not borne out by the data.

Last, highly persistent effects from involuntarily leaving the training firm arise for workers trained in establishments with at least 500 employees. This is consistent with models of defined career paths, either within firms (internal labor markets) or between firms ("stepping stone" human capital accumulation). Internal labor markets in particular would also explain presence of negative selection if wages of less able workers cannot be lowered due to the presence of unions (Dustmann

---

[49] One can see that some convergence arises in the panel sample or among those graduating from larger firms. Since these are both cases in which workers are likely to be generally of higher ability, this is consistent with the hypothesis that in the main sample random movers are outnumbered by selected movers.

and Schoenberg 2001). However, since these losses can almost fully be explained by a decline in employment size, it would also be consistent with losses of more narrowly defined firm-size wage premiums simply arising from rent-sharing.

6. Conclusion

Economists and policy makers have long been concerned that displacements and wage losses early in a career might lead to permanent disadvantages. A recent literature indeed suggests that young workers suffer persistent wage losses from an early job displacement. However, this paper has argued that simple estimates might overstate earnings losses if less able workers tend to be hired by firms with higher turnover rates. Such a process of initial assignment of workers to firms adds to any bias arising from negative selection because employers selectively displace their least able workers; or from positive selection because young workers leave voluntarily to take better jobs. Using longitudinal data on German apprentices and their training firms the paper addresses these complex selection and sorting mechanisms directly at the firm level. To measure the true long-run effect of leaving the training firm at the end of an apprenticeship, it uses changes in firms' retention rates of young apprentices as an instrument for displacement. This exploits variation in firms' hiring rates over time to best approximate a 'random' displacement. Moreover, by introducing firm fixed effects it uses workers at the same training firm as comparison group for displaced workers and thereby controls for initial assignment of workers to firms.

Simple comparisons of wages of those leaving the training firm and those staying suggest long-term effects of a displacement in a similar order of magnitude as found among American workers. However, controlling for selection within firms and sorting between firms the estimates show no permanent effects of an initial displacement. Instead, wage losses of leaving the training firm are initially large and then gradually revert to zero within the first four to five years of labor market experience. These results are robust to changes in sample decomposition or definition of the instrument. Permanent effects arise only for workers who leave large training firms to work in smaller establishments, consistent with the presence of firm-size wage premia or internal labor

37

markets. Since each estimate is affected by different confounding factors, the results can also be used to learn about the importance of various mechanisms underlying job and wage mobility. In particular, the importance of firm fixed effects suggests an important role for initial sorting of workers into firms. Less able workers appear to sort into firms with low retention rates and possibly lower quality of training. Second, a simple model of job search could explain the presence of temporary wage losses and an initial upward bias of OLS estimates. Voluntary and beneficial job mobility appears very common among young apprentices, consistent with what has been found for young workers in the United States. Last, there is negative selection as well, consistent with presence of adverse selection as employers learn about their apprentices and release information to the market by a displacement (Acemoglu and Pischke 1998). This suggests asymmetric information may be a problem for younger workers despite a high degree of voluntary job mobility (Gibbons and Katz 1991).

The labor market of young workers is simultaneously characterized by negative selection, sorting among heterogeneous employers, and a high degree of voluntary job mobility. The estimates presented here show that crucial information is gained about the mechanisms of job and wage dynamics when these aspects are considered jointly. The paper thus considerably enriches the characterization of young workers' career paths by Topel and Ward (1992) and others who focus on single aspects of job and wage mobility. The estimates also imply that particular care has to be taken to address the role of heterogeneity among firms and sorting of workers into firms when analyzing job changes of young workers. The fact that at least part of sorting occurs along firms' turnover rates is particularly relevant for the study of young displaced workers, since it implies displacement is not a random event even controlling for selection within firms. Similarly, the role of negative selection cannot be ignored when estimating the cost of job loss. Since young workers' wages might not reflect productivity, either because of asymmetric information, employer learning, or training wages, fixed effects at the worker level cannot be used to address this problem. Both sorting and selection in the labor market for young workers suggest information on firms' labor demand is crucial to obtain reliable estimates of the costs of early job loss.

References

Abowd, John, Patrick Corbel, and Francis Kramarz (1999). The Entry and Exit of Workers and the Growth of Employment: An Analysis of French Establishments'. *Review of Economics and Statistics* 81(2): 170-87.

Abowd, John, Francis Kramarz, and David Margolis (1999). 'High Wage Workers and High Wage Firms.' *Econometrica* 67(2): 411-45.

Acemoglu, Daron and Jörn-Steffen Pischke (1996). 'Why do Firms Train? Theory and Evidence.' *Quarterly Journal of Economics* 113(1): 79-119.

Anderson, Patricia M. and Bruce D. Meyer (1994). 'The Extent and Consequences of Job Turnover.' *Brookings Papers on Economic Activity. Microeconomics*: 177-236.

Angrist, Joshua D. and Alan B. Krueger (1992). 'The Effect of Age at School Entry on Educational Attainment: An Application of Instrumental Variables with Moments from Two Samples.' *Journal of the American Statistical Association* 87 (418): 328-336.

Angrist, Joshua, Guido Imbens, and Donald Rubin (1996). "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association* 91 (434).

Antel, John J. (1985). 'Costly Labor Contract Renegotiation and the Labor Mobility of Young Men.' *American Economic Review.* 75: 976-991.

Arulampalam, Wiji, Paul Gregg, and Mary Gregory (2001). 'Unemployment Scarring.' *Economic Journal* 111: F577-F584.

Baker, George, Michael Gibbs, and Bengt Holmstrom (1994). 'The Internal Economics of the Firm: Evidence from Personnel Data.' *Quarterly Journal of Economics* 109: 881-919.

Bartel, Ann and George J. Borjas (1981). 'Wage Growth and Job Turnover: An Empirical Analysis.' In: S. Rosen (ed.) *Studies in Labor Markets.* Chicago, IL: University of Chicago Press.

Bauer, Thomas K. und Stefan Bender (2002). 'Technological Change, Organizational Change, and Job Turnover: A Descriptive Analysis for Germany.' IZA. Mimeo.

Beaudry, Paul and John DiNardo (1992). 'The Effect of Implicit Contracts on the Movements of Wages over the Business Cycle' *Journal of Political Economy:* 99(4).

Bender, Stefan, Anette Haas, and Christoph Klose (2000). 'IAB Employment Subsample 1975-1995: Opportunities for Analysis Provided by the Anonymised Subsample.' IZA Discussion Paper No. 117.

Bender, Stefan, Christian Dustmann, David Margolis, and Costas Meghir (1995). 'Worker Displacement in France and Germany.' The Institute for Fiscal Studies Working Paper Series No. W99/14.

Bender, Stefan, Jurgen Hilzendegen, Gotz Rohwer, and Helmut Rudolph (1996). 'Die IAB-Beschaftigtenstichprobe 1975-1990.' *Beitraege zur Arbeitsmarkt- und Berufsforschung* 197.

Berufsbildungsbericht (2001). Bonn, Germany: Bundesministerium fur Bildung und Forschung.

Blanchard, Olivier and Peter Diamond (1994). 'Ranking, Unemployment Duration, and Wages.' *Review of Economic Studies* 61: 417-34.

Blanchflower, David G. and Andrew J. Oswald (1994). *The Wage Curve.* Cambridge, MA and London: MIT Press.

Blau, Francine D. and Lawrence M. Kahn (1997). 'Gender and Youth Employment Outcomes: The U.S. and West Germany, 1984-1991.' NBER Working Paper No. 6078.

Buechtermann, Cristoph (1989). 'Befristete Arbeitsvertraege nach dem Beschaeftigtenfoerderungs-gesetz.' *Forschungsvorhaben im Auftrag des Bundesministerium fuer Arbeit und Soziales.*

Brown, Charles and James Medoff (1989). 'The Employer Size-Wage Effect.' *Journal of Political Economy* 97(5): 1027-59.

Burdett, Kenneth (1978). 'A Theory of Employee Job Search and Quit Rates.' *American Economic Review* 68(1): 212-20.

Couch, Kenneth A. (2001). 'Earnings Losses and Unemployment of Displaced Workers in Germany.' *Industrial and Labor Relations Review* 54(3): 559-72.

Davis, Steven J. and John Haltiwanger (1990). 'Gross Job Creation, Gross Job Destruction, and Employment Reallocation.' *Quarterly Journal of Economics* 107(3): 819-863.

Doeringer, Peter and Michael Piore (1971). *Internal Labor Markets and Manpower Analysis.* Lexington, MA: D. C. Heath and Company.

Dustmann, Christian and Uta Schoenberg (2002). 'Why do German Firms Offer Apprenticeship Training?' University College London. Mimeo.

Ellwood, David (1982). 'Teenage Unemployment: Permanent Scars or Temporary Blemishes?' In: Freeman, Richard and David Wise (eds.) *The Youth Labor Market Problem: its Nature, Causes and Consequences.* Chicago: U. Chicago Press.

Euwals, Rob and Rainer Winkelmann (2001). 'Why do Firms Train? Empirical Evidence of the First Labour market Outcomes of Graduated Apprentices.' IZA Discussion Paper No. 319.

Fallick, Bruce C. (1996). 'A Review of the Recent Empirical Literature on Displaced Workers.' *Industrial and Labor Relations Review* 50(1): 5-16.

Farber, Henry (1993). 'The Incidence and Cost of Job Loss: 1982-1991.' *Brookings Papers of Economic Activity: Microeconomics.* 1993: 73-132.

Farber, Henry (1994). 'The Analysis of Interfirm Worker Mobility.' *Journal of Labor Economics* 12: 554-93.

Farber, Henry and Robert Gibbons (1996). 'Learning and Wage Dynamics.' *Quarterly Journal of Economics* 111(4).

Franz, Wolfgang, Joachim Inkmann, Winfried Pohlmeier, and Volker Zimmermann (2000). 'Young and Out in Germany: On the Youth's Chances of Labor Market Entry in Germany.' In: D. Blanchflower and R. Freeman (eds.) *Youth Employment and Joblessness in Advanced Countries.* Chicago: Chicago University Press.

Gardecki, Rosella and David Neumark (1998). 'Order from Chaos? Effects of Early Labor Market Experiences on Adult Labor Market Outcomes.' *Industrial and Labor Relations Review* 51(2): 299-322.

Gibbons, Robert and Lawrence Katz (1991). 'Layoffs and Lemons.' *Journal of Labor Economics* 9(4): 351-380.

Gibbons, Robert and Lawrence Katz (1992). 'Does Unmeasured Ability Explain Inter-Industry Wage Differences?' *Review of Economic Studies* 59: 515-35.

Gibbons, Robert and Michael Waldman (1999). 'A Theory of Wage and Promotion Dynamics Inside Firms.' *Quarterly Journal of Economics* 114: 1321-58.

Gibbons, Robert, Lawrence Katz, Thomas Lemieux and Daniel Parent (2002). 'Comparative Advantage, Learning, and Sectoral Wage Determination.' NBER Working Paper No. 8889.

Greenwald, Bruce (1986). 'Adverse Selection in the Labor Market.' *Review of Economic Studies* 53: 325-47.

Groshen, Erica L. (1991). 'Sources of Inter-Industry Wage Dispersion: How Much Do Employers Matter?' *Quarterly Journal of Economics* 106(3): 869-84.

Grund- and Strukturdaten 2000/2001 (2001). Bonn, Germany: Bundesministerium fur Bildung und Forschung.

Gustafson, Cynthia K. (1998). 'Job Displacement and Mobility of Younger Workers.' Center for Labor Economics Working Paper No. 8.

Harhoff, Dietmar and Thomas J. Kane (1997). 'Is the Apprenticeship System a Panacea for the U.S. Labor Market?' *Journal of Population Economics* 10: 171-96.

Hashimoto, Masanori and Ross A. Miller (1997). 'How do Training and Early Labor Market Experience Affect the Economic Well Being of Youths?' Mimeo.

Heckman, James J. and George J. Borjas (1982). 'Does Unemployment Cause Future Unemployment? Definitions, Questions and Answers From a Continuous Model of Heterogeneity and State Dependence.' *Economica*, New Series 42(187): 247-83.

Heckman, James J., Rebecca Roselius and Jeffrey Smith (1994). 'US Education and Training Policy: A Re-Evaluation of the Underlying Assumptions of the 'New Consensus.'' In: Solomon, Lewis and Alec Levenson (eds). *Labour Markets, Employment Policy and Job Creation*. Boulder: Westview Press.

Jacobson, Louis, Robert LaLonde and Daniel Sullivan (1992). 'Earnings Losses of Displaced Workers.' *American Economic Review* 83(4): 685-709.

Jovanovic, Boyan and B. Nyarko (1997). 'Stepping Stone Mobility.' *Carnegie-Rochester Conference Series in Public Policy* 46: 289-337.

Kletzer, Lori G. (1998). 'Job Displacement.' *Journal of Economic Perspectives* 12(1): 115-36.

Kletzer, Lori G. and Robert W. Fairlie (2001). 'The Long-Term Costs of Job Displacement for Young Adult Workers.' University of California Santa Cruz. Mimeo.

Krashinsky, Harry (2002). 'Evidence on Adverse Selection and Establishment Size In the Labor Market.' *Industrial and Labor Relations Review 56(1): 84-96*.

Krueger, Alan B. and Lawrence H. Summers (1987). 'Efficiency Wages and the Inter-Industry Wage Structure.' *Econometrica* 56: 259-94.

Machin, Stephen and Alan Manning (1999). 'The Causes and Consequences in Long-Term Unemployment in Europe.' In: O. Ashenfelter and D. Card (eds.) *Handbook of labor economics*. Volume 3A. Amsterdam, New York, and Oxford: Elsevier Science.

Manning, Alan (2001). *Monopsony in Motion.* (forthcoming).

Margolis, David (1995). 'Firm Heterogeneity and Worker Self-Selection Bias Estimated Returns to Seniority.' CIRANO Working Paper No. 95s-4.

Margolis, David, Veronique Simonnet and Lars Vilhuber (2000). 'Using Early Career Experiences and Later Career Outcomes to Distinguish Between Models of Labor Market Behavior Under Institutional Constraints.' CNRS Working Paper. Universite Paris 1.

McLaughlin, Kenneth J. (1991). 'A Theory of Quits and Layoffs with Efficient Turnover.' *Journal of Political Economy* 99: 1-29.

Mincer, Jacob (1986). 'Wage Changes in Job Changes.' In: R.G. Ehrenberg (ed.) *Research in Labor Economics.* Vol. 8.Greenwich, CT: JAI Press.

Mincer, Jacob and Boyan Jovanovic (1981). 'Labor Mobility and Wages.' In: Rosen, Sherwin (ed), *Studies in Labor Markets.* Chicago: University of Chicago Press.

Murphy, Kevin M. and Finis Welch (1992). 'The Structure of Wages.' *Quarterly Journal of Economics* 107(1): 285-326.

Murphy, Kevin M. and Robert H. Topel (1985). 'Estimation and Inference in Two-Step Econometric Models.' J*ournal of Business and Economics Statistics* 3(4): 370-79.

Neal, Derek (1995). "Industry-Specific Human Capital: Evidence From Displaced Workers," *Journal of Labor Economics*, October, vol. 13: 653-77.

Neal, Derek (1998). 'The Link Between Ability and Specialization: An Explanation for Observed Correlations Between Wages and Mobility Rates.' *Journal of Human Resources* 33(1): 173-200.

Neumark, David (1998). 'Youth Labor Market in the U.S.: Shopping Around vs. Staying Put.' NBER Working Paper No. 6581.

Oi, Walter Y. and Todd L. Idson (1999). 'Firm Size and Wages.' In: O. Ashenfelter and D. Card (eds.) *Handbook of labor economics.* Volume 3A. Amsterdam, New York, and Oxford: Elsevier Science.

Okun, Arthur M. (1973). 'Upward Mobility in a High-Pressure Economy.' *Brookings Papers of Economic Activity* 1: 207-52.

Rosen, Sherwin (1972). 'Learning and Experience in the Labor Market.' *Journal of Human Resources* 7(3).

Ruud, Paul A. (2000). *An Introduction to Classical Econometric Theory.* New York, Oxford: Oxford University Press.

Ryan, Paul (2001). 'The School-To-Work Transition: A Cross-National Perspective.' *Journal of Economic Literature* 39 (March): 34-92.

Salop, Joanne and Steven Salop (1976). 'Self-Selection and Turnover in the Labor Market.' *Quarterly Journal of Economics* 90.

Stern, David, N. Finkelstein, J. R. Stone III, J. Latting, and C. Dornsife (1995) *School to Work: Research on Programs in the United States.* London and Washington: Falmer Press.)

Topel, Robert and Michael Ward (1992). 'Job Mobility and the Careers of Young Men.' *Quarterly Journal of Economics* 107(2): 439-479.

von Wachter, Till (2001a). 'A Model of Learning and Adverse Selection.' University of California Berkeley. Mimeo.

von Wachter, Till (2001b). 'The Persistence of Bad Luck – Long Term Effects on Earnings of Beginning to Work in a Recession.' University of California Berkeley. Mimeo.

Waldman, Michael (1984). 'Job Assignment, Signaling, and Efficiency.' *Rand Journal of Economics* 15: 255-67.

Weiss, Andrew and Ruqu Wang (1998). 'A Sorting Model of Labor Contracts: Implications for Layoffs and Wage-Tenure Profiles.' *Labour Economics* 5(3): 359-83.

Winkelmann, Rainer (1996). 'Employment Prospects and Skill Acquisition of Apprenticeship-Trained Workers in Germany.' *Industrial and Labor Relations Review* 49(4): 658-672.

Table 1: Early Labor Market Experience of Young Men - US and Germany

		Germany	United States
Panel A: Men, Age 18-34			
Actual Years of Experience at 5 (10) Years of Potential Experience	5 Years	4.30	3.73
	10 Years	8.91	8.19
Average Number of Jobs at 5 (10) Years of Potential Experience	5 Years	2.27	4.56
	10 Years	3.65	6.96
Fraction Leaving Job at 2 (6) Quarters of Job Tenure	2 Quarters	0.15	0.22
	6 Quarters	0.10	0.11
Panel B: Men, Age 18-34, Jobs Lasting 6+ Quarters			
Average Completed Job Duration		7.29	7.00
(Standard Deviation)		(4.13)	(3.67)
Fraction Changing Job		0.25	0.28
(Standard Deviation)		(0.43)	(0.45)
Average Wage Growth on Job		0.09	0.07
(Standard Deviation)		(0.22)	-
Panel C: Controlling for Experience and Tenure			
Average Wage Growth Within Jobs		0.20	0.14
(Standard Error)		(0.00)	(0.01)
Average Wage Growth Between Jobs		0.21	0.20
(Standard Error)		(0.03)	(0.02)
Fraction Growth After 10 Years		80%	60%

Notes: Left Column- own calculations from IAB employment subsample including apprentices. Right Column - Topel & Ward (1992). To be consistent with Topel and Ward, the sample consists of West-German males who were at least 18 years old at entry into the sample and who entered the labor force between 1975 and 1980. Workers who were older than 34 at the end of the sample period were dropped. Moreover, it was required that workers did not spend more than two years out of the labor force or in unemployment.

Table 2: Average Characteristics of Main Sample of Apprentices

	All Graduates	Workers Staying at Training Firm	Workers Leaving Training Firm
Age at End of Training	20.9	20.9	21.0
Fraction High School	0.17	0.18	0.15
Fraction Male	0.87	0.88	0.86
Fraction German	0.36	0.38	0.34
Training Duration >2 Years	0.83	0.84	0.81
Training Firm Size 500+	0.54	0.56	0.50
Training Firm Annual Employment Growth	-0.03	-0.02	-0.05
	(0.30)	(0.30)	(0.29)
Fraction Moving at Graduation	0.40	0	1
Average Fraction Movers Among Other Apprentices	0.40	0.28	0.59
	(0.30)	(0.23)	(0.30)
Log Training Wage (1996 Deutsche Mark)	3.91	3.96	3.85
	(0.30)	(0.30)	(0.29)
Log Real Daily Wage (1996 Deutsche Mark)	4.88	4.92	4.80
	(0.26)	(0.23)	(0.29)
White Collar Occupation	0.46	0.48	0.43
Manufacturing	0.47	0.51	0.42
Services and Trade	0.19	0.16	0.23
Banking, Insurance	0.14	0.16	0.10
Transport, Communications	0.09	0.05	0.13
Cohort 1992	0.37	0.39	0.34
Cohort 1993	0.32	0.32	0.33
Cohort 1994	0.30	0.28	0.34
Number of Observations	295653	177855	117798

Notes: Sample of apprentices who graduated in 1992 to 1994 from establishments with at least 50 employees and at least 5 graduating apprentices. See text for additional sample restrictions. The first column shows sample statistics for the entire sample of graduating apprentices. The last column shows the same characteristics for apprentices who stayed and moved from their training firm the day after the end of training. The only characteristic changing over time is the wage, all other variables pertain to the training period.

Table 3: OLS-Estimates of the Effect of Leaving Training Firm at Graduation on Log Real Wages, Various Specifications

	Exp.	(1)	(2)	(3)	(4)	(5)	(6)
Effect of Leaving Training Firm on Wages By Year of Potential Experience	1	-0.094 (0.0038)	-0.097 (0.0036)	-0.081 (0.0030)	-0.063 (0.0030)	-0.065 (0.0029)	-0.058 (0.0027)
	2	-0.093 (0.0014)	-0.096 (0.0014)	-0.083 (0.0013)	-0.068 (0.0013)	-0.069 (0.0013)	-0.062 (0.0013)
	3	-0.093 (0.0011)	-0.096 (0.0011)	-0.084 (0.0011)	-0.070 (0.0011)	-0.071 (0.0011)	-0.062 (0.0011)
	4	-0.091 (0.0010)	-0.094 (0.0011)	-0.082 (0.0010)	-0.069 (0.0011)	-0.070 (0.0011)	-0.060 (0.0011)
	5	-0.092 (0.0011)	-0.096 (0.0011)	-0.084 (0.0011)	-0.072 (0.0011)	-0.072 (0.0011)	-0.063 (0.0011)
Demographics		-	Yes	Yes	Yes	Yes	Yes
Firm Controls		-	-	Yes	Yes	Yes	Yes
Training Controls		-	-	-	Yes	Yes	Yes
Occupation		-	-	-	-	Yes	Yes
Industry		-	-	-	-	-	Yes
R2		0.06	0.16	0.20	0.24	0.25	0.29
MSE		0.252	0.239	0.233	0.227	0.225	0.220

Notes: Dependent variable is the log real daily wage. All specifications include interactions of experience dummies with cohort dummies. Demographic characteristics consist in age at the end of training and dummies for German, male, and high school graduate. Firm variables include employment growth and three dummies for employment size of training establishment. Training variables include log real training wage, three dummies for training duration, and a dummy for whether a mover works at the training establishment. All observable characteristics are interacted with five experience dummies. Each regression has 991004 observations and 13009 establishments. Standard errors clustered at the establishment level are in parentheses.

Table 4: Different Estimates of Wage Losses of Apprentices Who Leave Training Firm at Graduation - Main Sample

	Year of Exp.	Raw Differences	OLS with Controls	OLS only Firm Fixed Effects	IV without Firm Fixed Effects	IV with Firm Fixed Effects
		(1)	(2)	(3)	(4)	(5)
Effect of Leaving Training Firm on Wages By Year of Potential Experience	1	-0.094 (0.0038)	-0.065 (0.0029)	-0.065 (0.0025)	-0.121 (0.0077)	-0.108 (0.0268)
	2	-0.093 (0.0014)	-0.069 (0.0013)	-0.065 (0.0027)	-0.113 (0.0072)	-0.064 (0.0215)
	3	-0.093 (0.0011)	-0.071 (0.0011)	-0.064 (0.0028)	-0.108 (0.0071)	-0.035 (0.0203)
	4	-0.091 (0.0010)	-0.070 (0.0011)	-0.062 (0.0028)	-0.101 (0.0071)	-0.005 (0.0209)
	5	-0.092 (0.0011)	-0.072 (0.0011)	-0.063 (0.0030)	-0.098 (0.0075)	0.009 (0.0221)
T-Statistics	H0: 1=3	-0.7	4.1	-0.8	-3.4	-3.6
	H0: 3=5	-1.0	0.5	-0.8	-3.4	-3.5
	H0: 1=5	-1.3	3.7	-1.2	-5.0	-4.7

Notes: The dependent variable is the log real daily wage. The first rows report estimates of a dummy for moving out of training firm after end of training interacted with experience-dummies. The last rows report t-test statistics for equality of these coefficients. All specifications include interactions of experience dummies with cohort dummies. The regression models of columns 2, 4, and 5 also include age at the end of training and dummies for German, male, and high school graduate; employment growth rate and three dummies for employment size of the training establishment; log real training wage, three dummies for training duration, and a dummy for whether a mover works at the training establishment; dummies for training occupation. All observable characteristics are interacted with five experience dummies. Each regression has 991004 observations and 13009 establishments. Standard errors clustered at the establishment level are in parentheses.

Table 5: First Stage Regressions - Linear Models of Probability of Leaving Training Firm at Graduation

	Year of Exp.	Instrument 1: Fraction 'Other' Graduates Leaving Firm at the End of Training		Instrument 2: Fraction 'Other' Graduates Leaving Firm with Non-Employment Spells	
		No Firm Fixed Effects	Firm Fixed Effects	No Firm Fixed Effects	Firm Fixed Effects
Instrument Interacted With Year of Potential Experience	1	0.718 (0.0064)	0.135 (0.0169)	0.604 (0.0135)	0.151 (0.0128)
	2	0.739 (0.0023)	0.152 (0.0171)	0.654 (0.0046)	0.173 (0.0126)
	3	0.748 (0.0016)	0.161 (0.0172)	0.675 (0.0030)	0.189 (0.0127)
	4	0.750 (0.0014)	0.164 (0.0172)	0.683 (0.0026)	0.195 (0.0127)
	5	0.754 (0.0013)	0.167 (0.0174)	0.696 (0.0024)	0.205 (0.0128)
R2		0.27	0.30	0.32	0.36
MSE		0.405	0.397	0.391	0.380
F- Statistic		3410.41	41.21	500.06	70.49

Notes: The dependent variable is a dummy for moving out of training firm at end of training. All specifications include interactions of experience dummies with cohort dummies; as well age at the end of training and dummies for German, male, and high school graduate; employment growth rate and three dummies for employment size of the training establishment; log real training wage, three dummies for training duration, and a dummy for whether a mover works at the training establishment; dummies for training occupation. All observable characteristics are interacted with five experience dummies. Each regression has 991004 observations and 13009 establishements. The last row shows the test-statistic for an F-test for the hypothesis that the coefficients on the instruments are jointly equal to zero. Standard errors clustered at the establishment level are in parentheses.

Table 6: Different Estimates of Wage Losses of Apprentices Who Leave Training Firm at Graduation

	Year of Exp.	Alternative Instrument	Panel Sample	Restricted Labor Force Attachment	Alternative Instrument, Panel	Alternative Instrument, Restricted	Excluding Recalls	No Recalls, Alternative Instrument
		(1)	(2)	(3)	(4)	(5)	(6)	(7)
Effect of Leaving Training Firm on Wages By Year of Potential Experience	1	-0.182 (0.0331)	-0.166 (0.0279)	-0.155 (0.0264)	-0.207 (0.0317)	-0.228 (0.0318)	-0.160 (0.0413)	-0.333 (0.0617)
	2	-0.135 (0.0249)	-0.103 (0.0237)	-0.098 (0.0213)	-0.154 (0.0269)	-0.166 (0.0235)	-0.088 (0.0297)	-0.189 (0.0355)
	3	-0.064 (0.0225)	-0.041 (0.0233)	-0.046 (0.0199)	-0.040 (0.0246)	-0.063 (0.0214)	-0.050 (0.0275)	-0.088 (0.0314)
	4	-0.013 (0.0242)	-0.024 (0.0242)	-0.005 (0.0207)	-0.003 (0.0269)	0.006 (0.0244)	-0.011 (0.0281)	-0.008 (0.0349)
	5	0.015 (0.0244)	-0.009 (0.0261)	0.010 (0.0223)	0.052 (0.0302)	0.051 (0.0269)	-0.002 (0.0295)	0.001 (0.0337)
No. of Observations		991004	562670	908603	562670	908603	945192	945192
No. of Firms		13009	12402	129754	12402	129754	12528	12528
T-Statistics	H0: 1=3	-6.1	-4.8	-4.2	-6.3	-6.0	-3.5	-4.4
	H0: 3=5	-2.0	-4.0	-5.7	-4.3	-5.4	-3.3	-4.3
	H0: 1=5	-5.9	-5.5	-4.9	-7.0	-7.3	-4.1	-5.1

Note: The dependent variable is the log real daily wage. The first rows report estimates of a dummy for moving out of training firm after end of training interacted with experience-dummies. The last rows report T-test statistics for equality of these coefficients. All specifications include interactions between experience and cohort dummies. The regression models of columns 2, 4, and 5 also include age at the end of training and dummies for German, male, and high school graduate; employment growth rate and three dummies for employment size of training establishment; log real training wage, three dummies for training duration, and a dummy for whether a mover works at the training establishment; dummies for training occupation. All observable characteristics are interacted with five experience dummies. Standard errors clustered at the establishment level are in parentheses.

Table 7: Characteristics of Training Firms with High, Medium and Low Average Retention Rates of Trainees at Graduation

	Low Fraction Movers	Medium Fraction Movers	High Fraction Movers
Range of Fraction Moving at Graduation	**0-20%**	**20-60%**	**60-100%**
Average Fraction Moving at Graduation	0.06	0.37	0.79
Median Training Firm Size	264	240	160
Median Number of Graduates	7	7	7
Median Number of Trainees	19	20	20
Average Fraction High School	0.16	0.14	0.12
Average Fraction German	0.90	0.86	0.85
Average Fraction Male	0.58	0.57	0.58
Average Log Real Training Wage	3.96	3.89	3.80
Average Training Duration in Years	944	938	946
Average Training Firm Growth	0.01	-0.01	-0.05
Fraction Manufacturing	0.46	0.44	0.37
Fraction Services and Trade	0.21	0.33	0.45
Fraction FIRE	0.17	0.10	0.05
Average Fraction White Collar Workers	0.54	0.56	0.51
Average Job Duration of Hired Trainees (Stayers)			
One Year of Labor Market Experience	335	325	309
Three Years of Labor Market Experience	756	717	655
Five Years of Labor Market Experience	1157	1084	952
Average Log Real Wage of Hired Trainees (Stayers)			
One Year of Labor Market Experience	4.85	4.81	4.73
Three Years of Labor Market Experience	4.91	4.89	4.83
Five Years of Labor Market Experience	4.97	4.95	4.90

Notes: Characteristics of firms training apprentices by lowest 25%, inter-quartile range, and highest 25% of average fraction of apprentices who leave the firm at graduation. The number of firms is 13009. The relationships of training wages, firm size, firm employment growth, job duration, regular wages and fraction services is linear in the average fraction moving.

Table 8: Estimates of Wage Losses of Apprentices Who Leave Large Training Firms at Graduation

	Year of Exp.	Raw Differences	OLS with Controls	OLS only Firm Fixed Effects	IV without Firm Fixed Effects	IV with Firm Fixed Effects	Alternative IV with Firm Effects	IVFE with Firm Size
		(1)	(2)	(3)	(4)	(5)	(6)	(7)
Effect of Leaving Training Firm on Wages By Year of Potential Experience	1	-0.108 (0.0058)	-0.093 (0.0051)	-0.089 (0.0045)	-0.152 (0.0128)	-0.095 (0.0137)	-0.188 (0.0277)	-0.048 (0.0191)
	2	-0.108 (0.0021)	-0.097 (0.0053)	-0.087 (0.0049)	-0.145 (0.0120)	-0.076 (0.0127)	-0.163 (0.0222)	-0.029 (0.0144)
	3	-0.109 (0.0017)	-0.098 (0.0054)	-0.087 (0.0052)	-0.139 (0.0115)	-0.065 (0.0125)	-0.107 (0.0198)	-0.023 (0.0133)
	4	-0.105 (0.0016)	-0.093 (0.0054)	-0.082 (0.0050)	-0.131 (0.0116)	-0.050 (0.0127)	-0.084 (0.0215)	-0.010 (0.0130)
	5	-0.106 (0.0016)	-0.094 (0.0057)	-0.084 (0.0054)	-0.126 (0.0121)	-0.041 (0.0137)	-0.054 (0.0222)	-0.012 (0.0131)
T-Statistics	**H0: 1=3**	-6.1	-4.8	-4.2	-6.3	-6.0	-3.5	-4.4
	H0: 3=5	-2.0	-4.0	-5.7	-4.3	-5.4	-3.3	-4.3
	H0: 1=5	-5.9	-5.5	-4.9	-7.0	-7.3	-4.1	-5.1

Notes: The dependent variable is the log real daily wage. The first rows report estimates of a dummy for moving out of training firm after end of training interacted with experience-dummies. The last rows report t-test statistics for equality of these coefficients. All specifications include interactions of experience dummies with cohort dummies. The regression models of columns 2, 4, and 5 also include age at the end of training and dummies for German, male, and high school graduate; employment growth rate and three dummies for employment size of the training establishment; log real training wage, three dummies for training duration, and a dummy for whether a mover works at the training establishment; dummies for training occupation. All observable characteristics are interacted with five experience dummies. Each regression has 525510 observations and 3236 establishments. Standard errors clustered at the establishment levelare in parentheses.

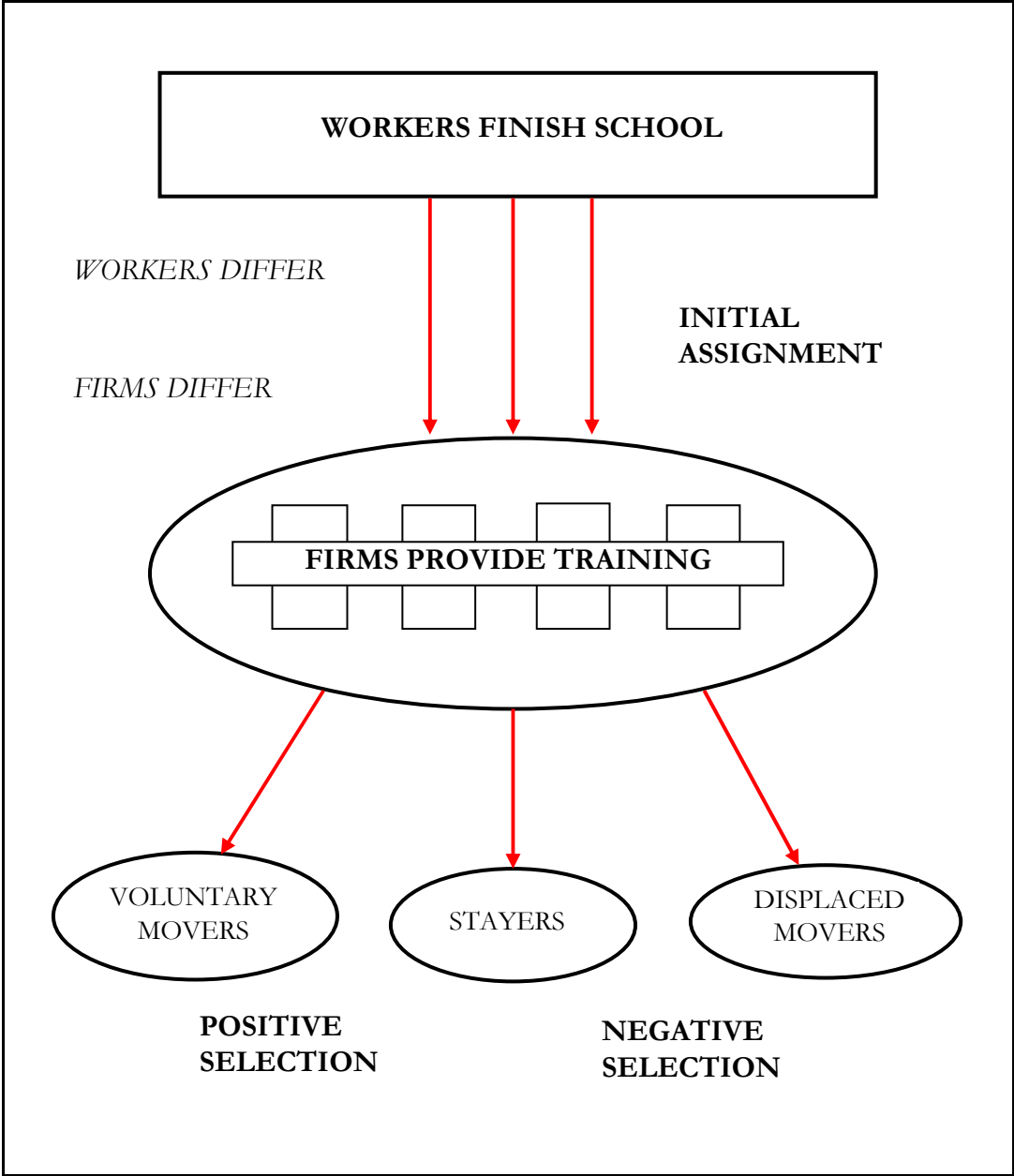**Figure 1: The Structure of Apprentice Training in Germany**

WORKERS FINISH SCHOOL

WORKERS DIFFER

INITIAL ASSIGNMENT

FIRMS DIFFER

FIRMS PROVIDE TRAINING

VOLUNTARY MOVERS

STAYERS

DISPLACED MOVERS

POSITIVE SELECTION

NEGATIVE SELECTION

**Figure 2: OLS and IV Estimates of Wage Losses from Leaving Training Firm At Graduation, With and Without Firm Fixed Effects - Main Sample**
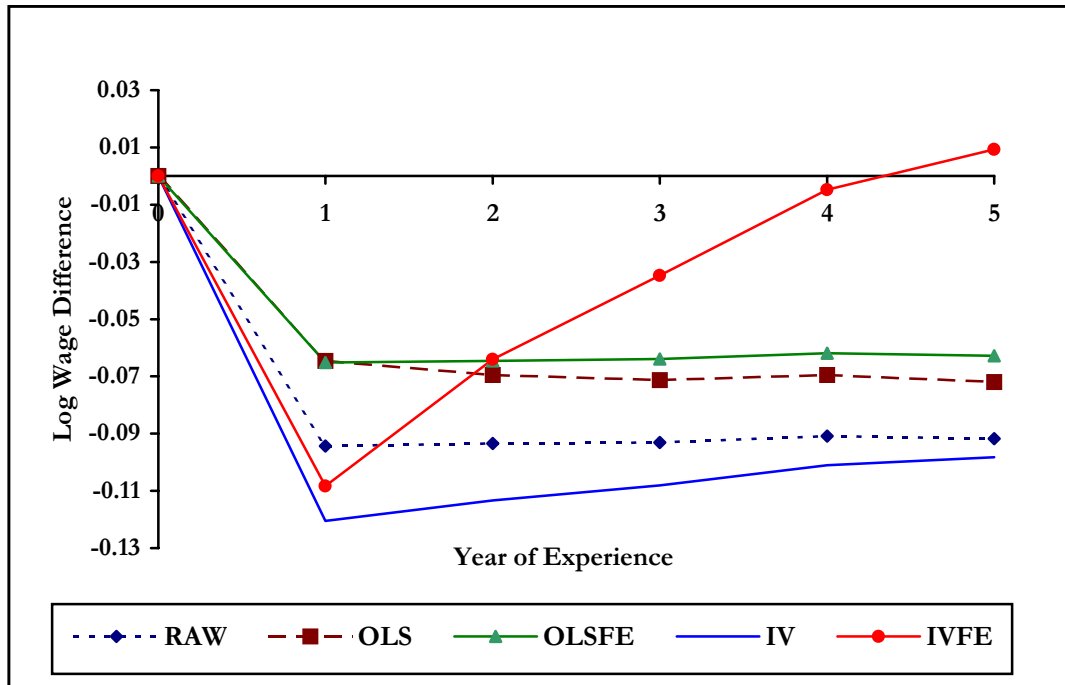
Figure 3: Average Real Wages By Intervals of Fraction 'Other' Apprentices Moving from Training Firm at End of Training

Panel A: Without Firm Fixed Effects
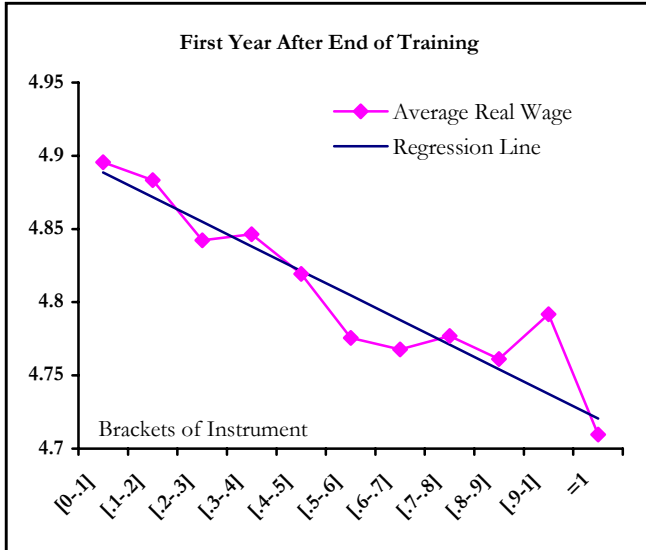
Panel B: With Firm Fixed Effects
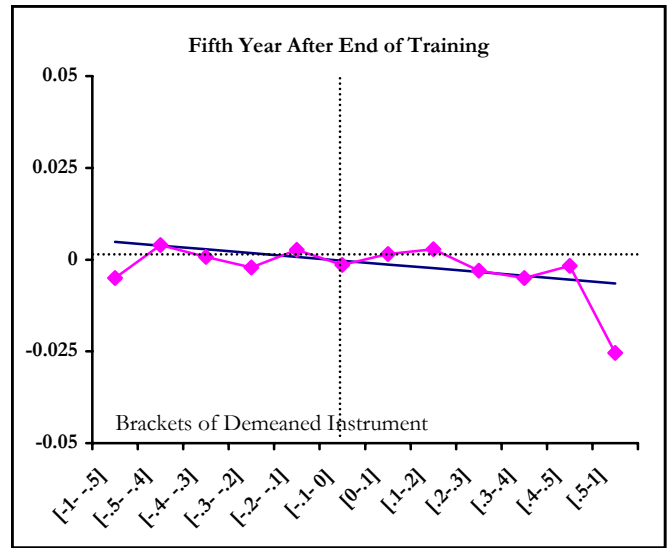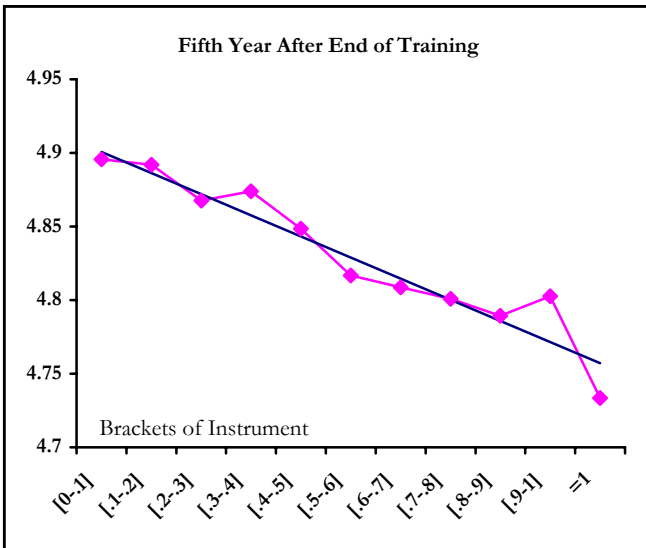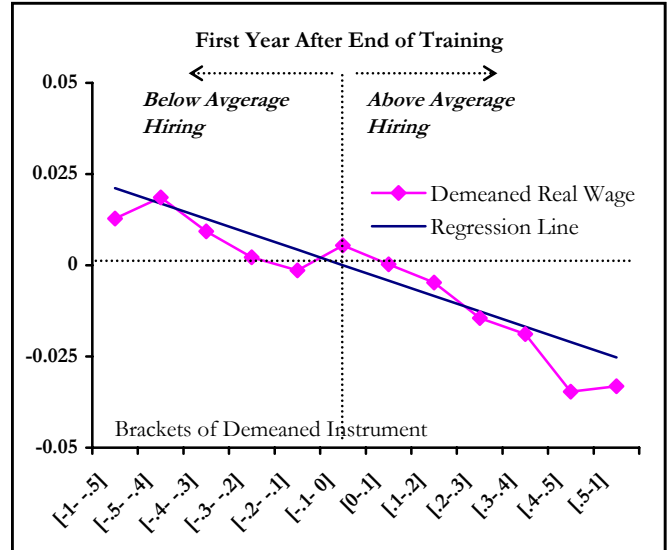
Figure 4: Sensitivity - Effects on Wages of Leaving Training Firm at Graduation, Various Specifications

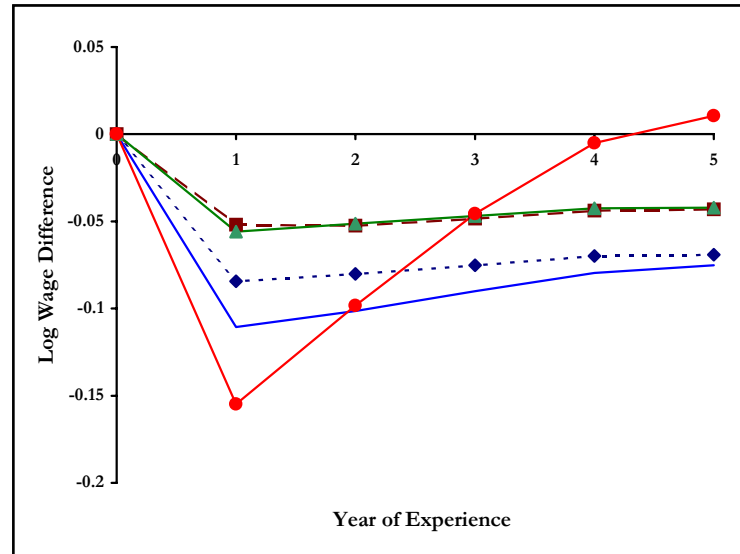**Panel A: Effect of Moving, Fraction Unemployed (IV2)**
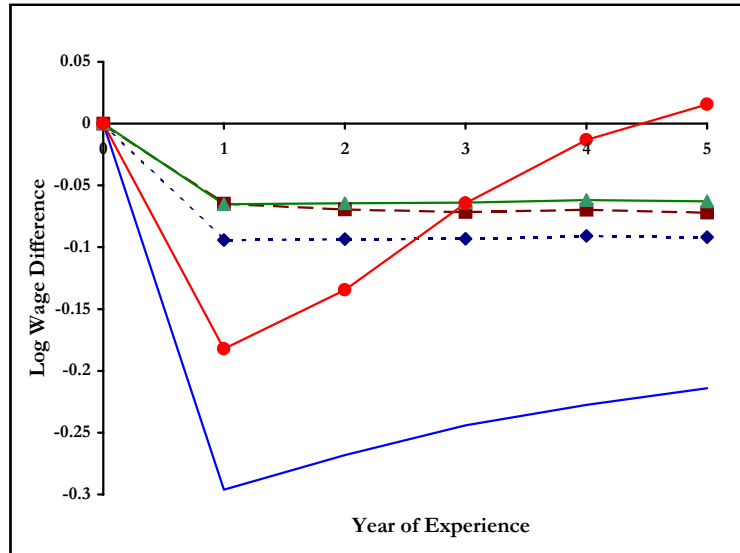
Panel B: Effect of Moving, Panel Sample
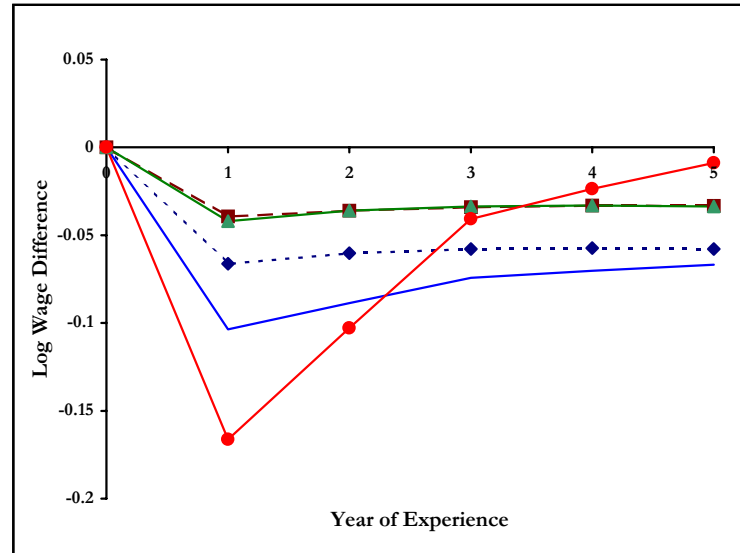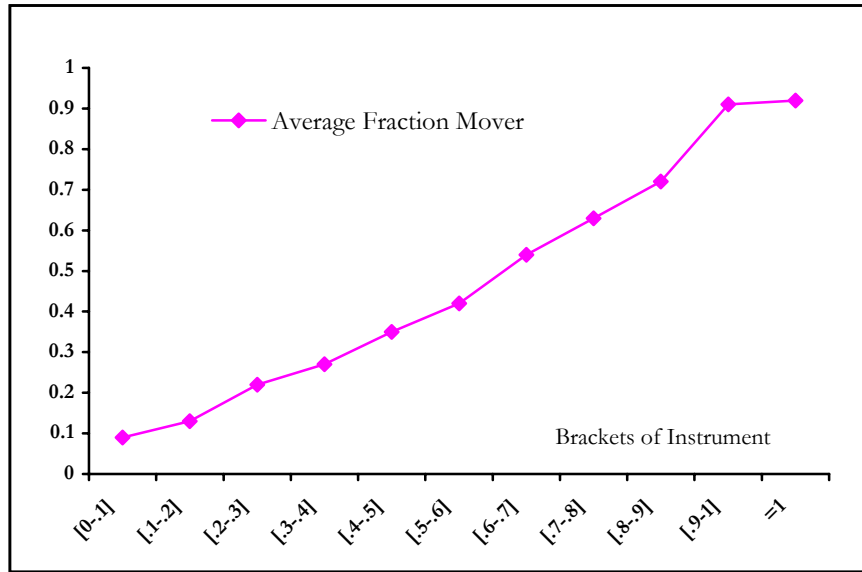

Panel C: Effect of Moving, Panel Sample with IV2
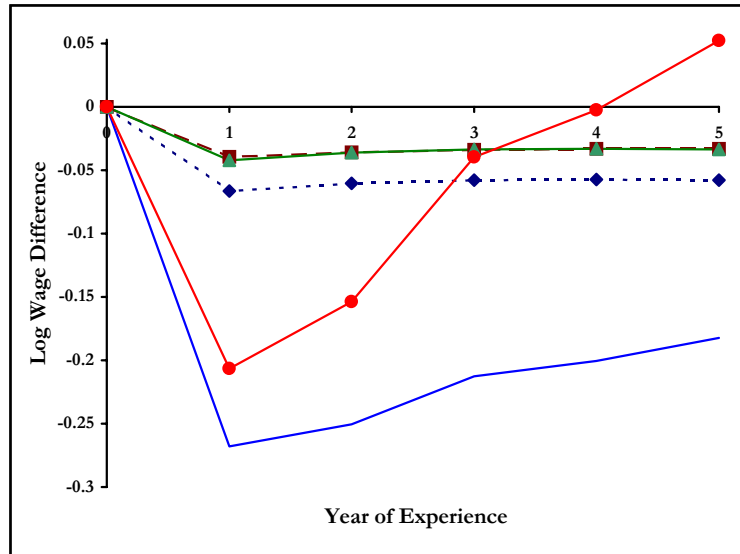
Panel D: Effect of Moving, Restricted OLF Sample
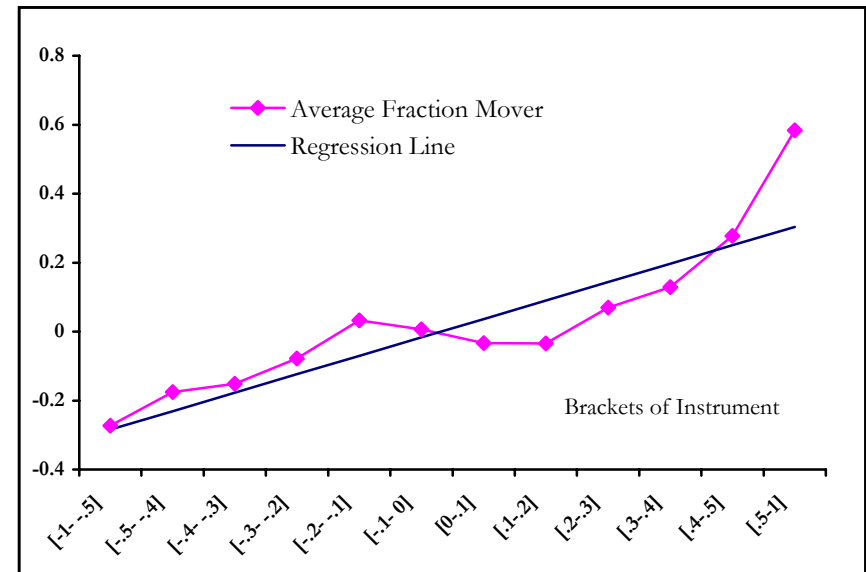

Legend: RAW, OLS, OLSFE, IV, IVFE

Figure 5: Fraction Movers by Intervals of Instrument, With or Without Firm Fixed Effects
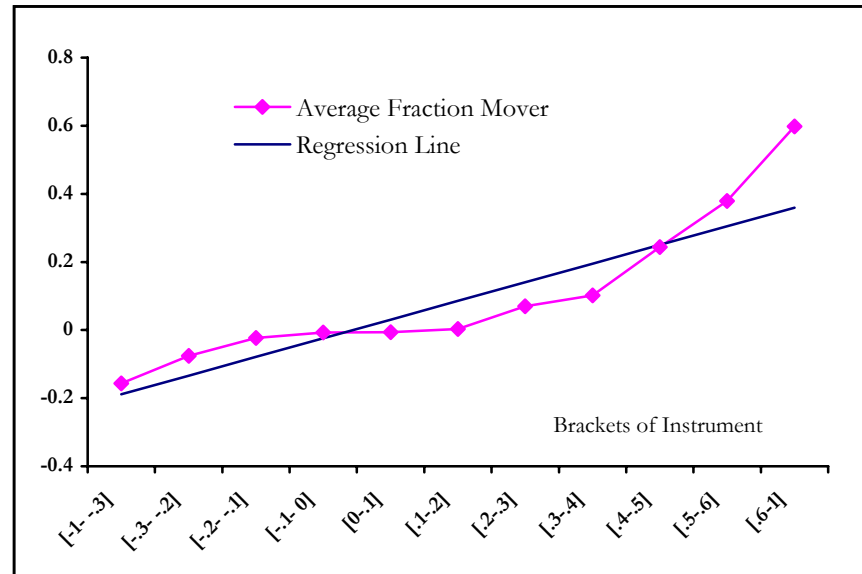
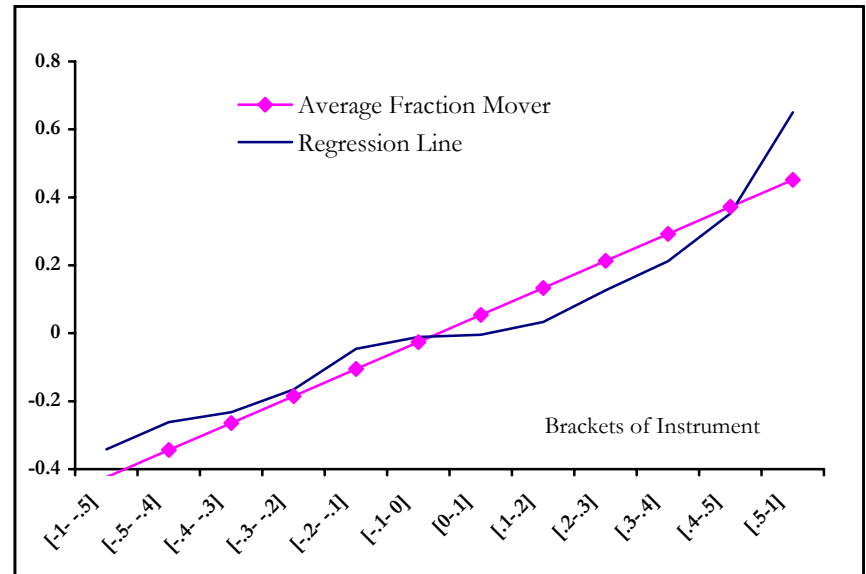**Panel A: Without Fixed Effects, Main Sample**

Panel B: With Firm Fixed Effects, Main Sample

Panel C: With Firm Fixed Effects, Unemployed Movers (IV2)

Panel D: With Firm Fixed Effects, Large Firms

Appendix Table: Selected Coefficient Estimates for Tables 4 and 5

	Year of Exp.	OLS	First Stage		Reduced Form	
			No Fixed Effects	With Fixed Effects	No Fixed Effects	With Fixed Effects
		(1)	(2)	(3)	(4)	(5)
Constant		3.614	0.484	0.970	3.623	4.053
		(0.021)	(0.031)	(0.046)	(0.021)	(0.021)
Fraction High-School	1	0.104	-0.009	-0.017	0.103	0.048
		(0.003)	(0.004)	(0.005)	(0.003)	(0.003)
	3	0.122	-0.013	-0.021	0.122	0.068
		(0.003)	(0.004)	(0.005)	(0.003)	(0.003)
	5	0.142	-0.011	-0.018	0.142	0.088
		(0.003)	(0.004)	(0.005)	(0.003)	(0.003)
Fraction Male	1	0.075	0.004	0.008	0.074	0.048
		(0.002)	(0.003)	(0.003)	(0.002)	(0.002)
	3	0.068	0.007	0.010	0.067	0.045
		(0.002)	(0.003)	(0.003)	(0.002)	(0.002)
	5	0.089	0.010	0.013	0.088	0.068
		(0.002)	(0.003)	(0.003)	(0.002)	(0.002)
Log Real Training Wage	1	0.216	-0.127	-0.168	0.218	0.123
		(0.004)	(0.005)	(0.008)	(0.004)	(0.003)
	3	0.170	-0.114	-0.157	0.173	0.084
		(0.003)	(0.005)	(0.008)	(0.003)	(0.003)
	5	0.152	-0.117	-0.160	0.157	0.070
		(0.004)	(0.005)	(0.008)	(0.004)	(0.003)
Employment Size of Training Establishment 100-500	1	0.056	-0.017	-0.031	0.053	0.001
		(0.004)	(0.004)	(0.018)	(0.004)	(0.006)
	3	0.122	-0.013	-0.021	0.122	0.068
		(0.003)	(0.004)	(0.005)	(0.003)	(0.003)
	5	0.142	-0.011	-0.018	0.142	0.088
		(0.003)	(0.004)	(0.005)	(0.003)	(0.003)
Employment Size of Training Establishment 500+	1	0.129	-0.033	-0.056	0.126	0.001
		(0.005)	(0.004)	(0.024)	(0.005)	(0.008)
	3	0.116	-0.020	-0.047	0.114	-0.009
		(0.004)	(0.004)	(0.023)	(0.004)	(0.008)
	5	0.115	-0.019	-0.045	0.114	-0.008
		(0.005)	(0.004)	(0.023)	(0.005)	(0.008)
Employment Growth of Training Establishment	1	0.005	-0.032	-0.216	0.004	0.008
		(0.003)	(0.015)	(0.024)	(0.002)	(0.006)
	3	-0.005	-0.035	-0.214	-0.006	0.002
		(0.003)	(0.017)	(0.024)	(0.003)	(0.006)
	5	-0.007	-0.027	-0.213	-0.007	0.000
		(0.004)	(0.013)	(0.024)	(0.004)	(0.006)

Notes: See Tables 4 and 5.