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 biases OLS estimates towards finding permanent negative effects of early displacements. (JEL J63, J65)

---

* Contact: vw2112@columbia.edu. Till von Wachter would like to thank David Card for constant guidance and support throughout this project. We would also like to thank Ken Chay, Olivier Deschenes, David Lee, David Levine, Justin McCrary, Daniel McFadden, Tiago Ribeiro, Jesse Rothstein, Emmanuel Saez, Diane Whitmore, three anonymous referees, and seminar participants at Berkeley, Chicago, Rochester, Columbia, Harvard, University of California San Diego, London School of Economics, Oxford, University College London, and the 2004 IAB Summer Symposium, the 2004 SOLE meetings, and the 2003 EALE and EEA meetings for helpful comments and suggestions. Till von Wachter thanks the Institut für Arbeitsmarkts- und Berufsforschung (IAB) for hosting him and the Institute for Labor and Employment for financial support. All remaining errors are our own.

1. Introduction

The first years of young workers' careers are characterized by high rates of earnings growth and high job mobility.[1] Many economists interpret this as evidence of beneficial job search (Topel and Ward 1992). Yet, some argue high mobility also exposes young workers to negative events in a crucial phase of their career with possibly long-term consequences (Ryan 2001). Young workers indeed have high displacement rates (Farber 1993) and suffer the largest wage declines in recessions (Blanchflower and Oswald 1994). Moreover, important recent studies replicating classic analyses of mature job losers for the case of young displaced workers, have found large and persistent wage losses (Kletzer and Fairlie 2003, Gustafson 1998). However, multiple confounding factors implied by the main mechanisms behind young workers' wage and job dynamics suggest common identification strategies valid for mature workers may fail in the analysis of young job losers.

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, comparisons of leavers and stayers suggest that there are persistent 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 pool of displaced worker is nonrandom even controlling for selection within firms. 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

[1] Seventy to eighty percent of life-time earnings growth occurs within ten years in the labor market (Murphy and Welch 1990). During the same time, workers on average hold 7 jobs; 5 of which during the 5 five years (Topel and Ward 1992).

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 the universe of graduates from the German apprenticeship system in the period from 1992 to 1994 that are observed working at least once in the first five years after training.[2] 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. These estimates can be interpreted as the local average treatment effect from a job loss for young workers induced to move by temporary demand shocks relative to similar workers at the same firm at risk of moving in other

[2] In Germany, more than two thirds of recent cohorts of school graduates participate in apprenticeship training programs that last on average two years and include both formal and practical training.

periods. 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 and to gain insights on the importance of basic models of early job and wage mobility.

The estimates draw a complex picture of the labor market for young workers where job search leads to voluntary mobility and true temporary wage losses from displacements, but sorting among firms and negative selection are important sources of job and wage dynamics as well. This implies important insights on early careers may be lost if either of these components is ignored. The results also speak to potential biases affecting studies of displaced workers lacking information on the demand side. First, since in the presence of adverse selection or firm training pre-displacement wages of young workers do not reflect productivity, worker fixed effects cannot be used to control for selection. Second, if workers sort themselves into firms by their turnover rates displacement is not a random event even controlling for selection within firms, and simple IV strategies retain a bias.

To make the comparison between estimators and theoretical implications explicit, the next section presents a model of wage determination that captures the basic theories of early job mobility in a unified framework. We then use this model to interpret the bias of OLS and to present our estimation strategy. The third section describes the matched worker-firm data set and the German apprenticeship system. 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. However, most of the separate mechanisms emphasized by alternative theories of job and wage mobility 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 in this complex environment.

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.[3] The goal of the 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

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

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.

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 first firm (leaving out other control variables for simplicity), the probability limit of the estimated effect on wages of an early job move after t years is

---

[3] 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.

$$p\lim \hat{\delta}_{It}^{OLS} = \frac{\operatorname{cov}(a_i - \bar{a}_{j(i)}, D_{i0})}{\operatorname{var}(D_{i0})} + \frac{\operatorname{cov}(\bar{a}_{j(i)}, D_{i0})}{\operatorname{var}(D_{i0})} + \delta_{It} + (\delta_{Vt} - \delta_{It}) \frac{\operatorname{cov}(V_{i0}, D_{i0})}{\operatorname{var}(D_{i0})}. \tag{2}$$

Several theories of wage and job mobility have implications for different components of Equation (2), and this will be helpful in the interpretation of the empirical results.[4] In summary, if movers are negatively selected, then $\operatorname{cov}(a_i - \bar{a}_{j(i)}, D_{i0}) < 0$; similarly, if less able workers are sorted at initial hiring into firms with high turnover, then $\operatorname{cov}(\bar{a}_{j(i)}, D_{i0}) < 0$. In both cases OLS tends to be biased toward finding a negative effect even if $\delta_{It} = 0$. On the other hand, the presence of voluntary mobility suggests that $\delta_{Vt} - \delta_{It} > 0$. In this case, since $\operatorname{cov}(V_{i0}, D_{i0}) > 0$, this implies that OLS would tend to underestimate the true effect of an involuntary move. Together with these confounding elements, the OLS estimate may also pick up true negative effects of a displacement.

The most common explanation for negative selection of displaced workers has been adverse selection in the labor market (Gibbons and Katz 1991). Recent evidence suggests that adverse selection may be less of a problem for older workers (Krashinsky 2002), which is not surprising if markets continuously learn about workers' ability (Farber and Gibbons 1996). However, wages and career histories of younger workers do not yet reflect their skills, and a displacement may signal additional information to the market. Acemoglu and Pischke (1998) 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 application.[5] An alternative source of negative selection is wage rigidity that may lead firms to fire less able workers instead of lowering their wages.

If not only workers but also firms differ systematically, observed mobility and wage changes may be driven by a sorting process of heterogeneous workers into heterogeneous firms. Sorting is a particular problem for the study of displaced workers if less able workers are hired by firms with

---

[4] 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.

[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.

higher turnover rates. That workers sort 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. A growing recent literature also 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). As suggested by Margolis (1995), it is likely that some sorting of workers occurs along firms' average turnover rates, and 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).

If, as the case in Germany, 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 particularly important. In terms of the model in Equation (1), initial assignment of less-able workers into high turnover firms implies that $\operatorname{cov}(\bar{a}_{j(i)}, D_{i0}) < 0$. As in the case of adverse selection workers are always paid their marginal product, and thus $\delta_{It} = \delta_{Vt} = 0$ and workers' mobility status has no causal impact on wages.[6]

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 earnings losses from an early job change may *underestimate* the effect of a job loss on wages.[7] [8] This is particularly relevant for

---

[6] If firms and workers themselves only gradually learn about their abilities and preferences, the process of sorting becomes sequential as in Gibbons and Katz (1992) and Gibbons, Katz, Lemieux, and Parent (2002). 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.

[7] A recent paper by Bender, Dustmann, Margolis, and Meghir (1999) 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.

young workers who have high voluntary mobility rates (Topel and Ward 1992). Job search also offers an explanation for true temporary wage losses from job displacement. A displacement destroys accumulated 'search capital' because workers have to start looking for good jobs from scratch (Manning 2003). Gradually, workers again find better job matches, and the initial wage losses from displacements should be temporary. In terms of Equation (1), job search implies that $\delta_{Vt} > 0$, leading to a positive bias in simple OLS estimates.[9] Transitory wage losses imply that $\delta_{It} < 0$ and $\partial \delta_{It} / \partial t > 0$.

True temporary wage losses from an early job displacement could also arise within the standard neo-classical human capital model. Displaced workers may lose skills specific to their previous employer or occupation (e.g., Kletzer 1998, Neal 1995) or lose opportunities to acquire general skills due to non-employment (Giuliano and von Wachter 2004). However, in the application of this paper, the German apprenticeship system is meant to provide mostly general skills (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, and this holds for German apprentices as well.[10]

Standard models of career development do not imply permanent effects of job displacements on earnings. The most common explanations of long-term 'scarring' effects of early job losses, losses in experience accumulation, and negative signaling to employers, require unduly strong assumption on wage and job dynamics.[11] However, models of career-ladders with specific entry-level

---

[8] This problem may also arise in more conventional data sets. Using the NLSY, Gardecki and Neumark (1998) 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. However, using local unemployment rates as an instrument for early job mobility, Neumark (2002) finds that early job mobility has a significant negative impact, suggesting that mobile workers in his sample are positively selected.

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

[10] The median duration of unemployment of a worker leaving the training firm at the end of an apprenticeship in our sample is six days. Since workers spend no more than 40%, but more likely 20% of their time actually working at their training firm, effective tenure at the end of training is 0.5-1 year at the modal training duration. In future work the available detailed data on career histories enables us to assess aspects of these alternative explanations directly.

[11] For example, under asymmetric information some form of continuous learning by employers would predict that eventually workers get paid their true marginal product. That is not to say that imperfect employer learning (for example due to ranking of applicants Blanchard and Diamond (1994) or due to statistical discrimination Machin and Manning

jobs can imply persistent wage losses of early displacements.[12] Similarly, permanent losses arise if workers lose premiums paid above market wages. If entry-level and high-paying jobs are scarce, or if entry into internal labor markets is restricted by age, then young job losers are likely to permanently lose career prospects or rents. Workers exiting large firms are particularly at risk of permanently losing career chances and wage premiums, and will be analyzed separately in the empirical analysis.

2.2. Estimates of Wage Losses and Confounding Factors

The paper proposes an estimation strategy that allows both estimation of the true causal effect of early job losses 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 we employ within-firm changes in labor demand for young workers as an instrument for involuntary mobility. Specifically, we use shocks to the retention rate of young workers at the end of apprentice training as instrument for mobility. 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. Since a key point of the paper is that the retention rate (z_{ijc}) may systematically differ across firms and that this may attract different types of workers, the final instrument used for the probability of moving will be the deviation of z_{ijc} from its firm specific average.[13] This isolates as closely as possible the group of

---

(1999)) is not possible. 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). 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).

[12] This could occur if initial jobs are 'stepping stones' in a sequential accumulation of knowledge (Jovanovic and Nyarko 1997), if experience accumulation is limited to certain types of jobs or firms (Rosen 1972, Okun 1973, Gibbons and Waldman 2002), or if jobs are ports of entry into well-defined career paths (Doeringer and Piore 1971, Baker, Gibbs, and Holmstrom 1994).

[13] 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. The final instrument used for the probability of moving will be

workers who would not have moved under normal business conditions and thereby best approximates an exogenous displacement. Moreover, we use training firm fixed effects to control for the effects of sorting, since non-displaced workers at the same training firm who are at risk of moving in other periods act as comparison group for the wage-outcomes of displaced workers.

Shocks to the retention rate of cohorts of young graduating apprentices are powerful and valid instruments. The variation in retention rates of graduating apprentices at the cohort and establishment level in Germany is high, and only 62% of it can be explained by establishment and cohort effects.[14] Moreover, 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.[15] Two observations suggest that changes in retention rate represent genuine economic shocks and not variation in apprentice quality. First, firm employment changes at all age levels are significantly positively correlated, suggesting a common economic cause. Second, the final estimate implies only a temporary wage loss from job displacement, instead of the permanent differences that would be implied by cohort-heterogeneity. 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.[16]

---

the deviation of z_{ijc} from its firm specific average $\tilde{z}_{ijc} = z_{ijc} - \bar{z}_j$, where $\bar{z}_j = \frac{1}{C}\sum_{c=1}^{C}\frac{1}{n_{jc}}\sum_{i=1}^{n_{jc}} z_{ijc}$. 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.

[14] Davis and Haltiwanger (1990) were the first to extensively document that changes in plant-level employment demand are frequent, large, and heterogeneous. 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.

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

[16] Systematic correlation of average turnover rates with sample characteristics disappears when controlling for firm fixed effects, i.e., the sample becomes balanced on observables (see Appendix E available from authors' website). Ideally, we would have access to data on sales at the plant and year level to assess the quality of the instrument. 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.

A remaining concern is that variation in the external labor market may induce changes in the fraction of workers leaving voluntarily, inducing a negative correlation between retention shocks and mobility among firms with few apprentices. To address this problem we restrict our sample to establishments with a minimal number of graduating apprentices. Moreover, first stage results suggest a significant positive (not negative) correlation of retention shocks and mobility. To also directly isolate demand side variation in employment, we 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 (but less than 10 months to exclude military leavers). Since we are certain to exclude most voluntary movers, involuntary movers should drive the main variation in the second instrument, yielding additional implications exploited in the sensitivity analysis.

The IV estimate with training firm fixed effects 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. 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 relative to similar workers at the same training firm at risk of moving involuntarily in other periods.[17] This is the relevant causal effect for the group of workers who are at risk of moving due to temporary demand fluctuations but who would stay at the firm under normal business conditions. Since we exclude small training firms from our sample, this is the effect on wages for workers displaced due to transitory shocks from stable medium-sized training firms. If these firms pay higher wages, losses of movers from small establishments may be smaller.

---

[17] 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). Defiers could arise if contrary to what is typically assumed in the literature on adverse selection, the market does not know which firm has temporarily low retention rates. Since leaving such a firm may yield a negative signal to the market, some who would move may decide to stay and wait for better times. The effect on the estimates depends on the ability of defiers relative to those who always stay at the firm. If the ability of voluntary movers (potential defiers) is higher than that of 'always stayers,' this would imply a permanent negative bias of the IVFE, contrary to our findings of catch-up. If voluntary movers are a random draw among 'always stayers' (as in a search interpretation of mobility) IVFE is not affected by the presence of defiers.

An advantage of our estimation procedure is that access to a matched employer-employee panel provides us with at least five estimators of the wage loss from displacement: OLS with firm fixed effects (OLSFE), simple instrumental variables (IV1), and IV with firm fixed effects (IVFE1). In addition, we present two additional IV estimates based on our second instrument (IV2 and IVFE2). Besides delivering the true long-term effect of an early job loss, the comparison of different estimators provides important information on the biases underlying the simple OLS estimator. Moreover, since economic theory has separate implications regarding initial assignment and various forms of selection, stepwise estimation gives a way to assess the relative importance of various mechanisms underlying wage and job changes.[18]

The simplest of the alternative estimates, OLS with firm fixed effects (OLSFE) is identified by deviations from firm averages. By only comparing workers who graduated at the same training firm, OLSFE accounts for the bias from initial assignment.[19] 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 $\mathrm{cov}(\bar{a}_{j(i)}, z_{ij}) < 0$, and the resulting IV estimator is biased. Since the denominator of the bias is now smaller than in the case of OLS ($\mathrm{cov}(D_{i0}, z_{ij}) < \mathrm{var}(D_{i0})$), initial sorting could imply that $\hat{\delta}_{It}^{IV} < \hat{\delta}_{It}^{OLS}$, i.e., the IV estimator can be more negative than OLS.[20]

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

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

[19] Under the model in Equation (1), the probability limit of the resulting estimate is

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

[20] The probability limit is $p\lim \hat{\delta}_{It}^{IV} = \delta_{It} + \mathrm{cov}(\bar{a}_{j(i)}, z_{ij}) / \mathrm{cov}(D_{i0}, z_{ij})$. Alternatively, IV1 might be more negative than OLS if the effect of negative selection on OLS is more than offset by the positive bias from voluntary mobility.

yields a consistent and unbiased estimate of the true effect of involuntary displacement, i.e., the probability limit is $p\lim \hat{\delta}_{It}^{IVFE} = \delta_{It}$. 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.

This is the first study to introduce firm fixed effects to control for permanent differences in average retention rates, and to use continuous occupation- and age specific labor demand shocks at the firm level as an instrument for job loss. Absent panel data on workers and employment flows at the firm level, researchers have to be aware of the potential remaining bias from sorting and selection. Most importantly, including worker fixed effects cannot control for these biases, since asymmetric information, gradual employer learning, and training wages imply that for young workers wage histories are unlikely to reflect productivity. A common alternative approach is to use plant closings as instrument; this may also retain a bias, since controlling for initial sorting becomes difficult.[21] Moreover, job loss from plant closings may estimate a different treatment effect than job losses induced by less drastic and more common adjustments in employment.

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.[22] 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. The institutional structure of the German apprenticeship system is ideal to analyze the persistence of early labor

---

[21] Krashinsky (2002) has analyzed problems with the use of plant closings as instrument for displacements.
[22] 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).

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. Although this eliminates several difficulties in the study of displaced workers,[23] all three of the 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 at the end of training and remains 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 2002, 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. Moreover, although institutional environments clearly differ, it can be shown that enough basic similarities in the labor market for young workers between the two countries exist to make the results relevant for the understanding of U.S. labor markets.[24] Another advantage is the accessibility of the relevant establishment-level data matched to detailed career information for a large sample of workers.

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.[25] Since the notification procedure for social security requires employers to record any permanent or temporary change of employment relationships and takes stock of existing employees

---

[23] For young displaced workers, difficulties arise specially from the definition of labor market entry (Kletzer and Fairlie 2002), heterogeneity in previous labor market experience, or the definition of the type of job transition studied.

[24] In Appendix C available from our website, 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. 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.

[25] An overview of the data is given in Bender, Haas, and Klose (2000), a detailed description can be found 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.

at each establishment twice a year, the employment register contains detailed histories for each worker's time in covered employment. In addition, 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 to a limit, but top coding is very rare for younger workers. In addition, the data contain basic demographic information as well as information on occupation, industry, job-status, and education.[26] 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 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.[27]

The sample used for this paper consists of information on the universe of trainees graduating from an apprenticeship in 1992 to 1994 in West Germany. The timeframe is chosen such that several cohorts can be observed for at least 5 years after entering the labor market.[28] For each graduating apprentice the sample contains information on the establishment where training takes place, on training itself, and basic demographic characteristics. 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 (see Table 1). 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., shorter vocational training spells are not included), it requires a minimum

[26] 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). This is less of a problem for workers in our sample who enter apprenticeships upon graduation.
[27] Unfortunately, it is currently not possible to link establishments that belong to a common parent firm.
[28] Before 1992 exact date of apprentice graduation is not identifiable in the data.

length of continuous training of 450 days, and it excludes workers who have prior labor market experience, 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.[29] 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 1 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 two and a half 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. 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.

[29]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.

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 low relative to the first wage workers earn (about 50%), although higher than apprenticeship standards due to the focus on larger firms.

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

Table 1 also shows that mobility at the end of training is high and non-random. Those who move at the end of training tend to be less educated, more likely to be male, more likely to be trained at smaller firms, and receive lower training wages.[30] 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 2. Movers have 9-10% lower wages than stayers, and this difference is basically unchanged after 5 years in the labor market. 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 firms all standard errors are clustered at the level of the training firm.[31] In case of OLS with training firm fixed effects,

[30] 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.
[31] 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), this equivalence fails and estimating the stacked model is more efficient (Ruud 2000, p.703).

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.

The remaining columns of Table 2 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. The evidence in Table 2 confirms that the difference in wages among movers and stayers remains significantly negative and stable over time even if we control for all additional observable characteristics. Thereby, only including the log training wage has a large effect on the wage difference. This is not surprising, since if we cannot include individual workers' fixed effects, including training wages partially controls for productive ability. However, the coefficient on training wage is only about 0.2, confirming the suspicion that it is an imperfect control for ability.[32] The preferred OLS specification used throughout the analysis is that of Column 5, which excludes industry controls. Industry of training is another dimension along which sorting and selection could occur, and is further discussed in the sensitivity analysis.

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. The wage differences among movers and stayers after controlling for firm fixed effects alone are shown in the third column of Table 3. 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 basic OLS 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 3 does not further

[32] The results on other regressors are summarized in Appendix D available from our website. Note that two thirds of the difference in training wage among stayers and movers is explained by differences among their establishments. Thus, differences in the explanatory power from training wages partly comes from permanent differences among training establishments.

affect the differences in wages (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. If adverse selection is the main source of wage losses, then workers leaving firms with higher turnover rates should be less negatively selected. Table 4 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-sample instrumental variables.[33] Column 1 of Table 4 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 3 (coefficients on other variables are shown in Appendix D available

---

[33] 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).

from our website). Using simple IV, 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, negative selection alone cannot seem to explain the observed wage differences. On the other hand, the increase in the gap among movers and stayers is 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 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 4 (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.[34] The F-test statistics for the hypothesis of joint insignificance of the instruments are again well beyond the critical level for weak instruments suggested by Stock and Watson (1997).

[34] 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).

The final IV estimates (IVFE) are shown in the last column of Table 3. Including establishment fixed effects the estimated wage difference among stayers and movers in the first year after entry into the labor market is -10.8%. However, using firm fixed effects the difference is not persistent and decreases to –3.5% and 0.9% 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 3, 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. Thus, although all other estimates suggest permanent negative effects, an IV strategy with firm fixed effects implies large initial effects that converge to zero within five years.

These results, summarized in Figure 1, suggest that initial sorting and negative selection are an important characteristic of job and wage dynamics of young apprentices. Only if we control for both permanent differences across firms and firm-specific shocks wage differences between movers and stayers are no longer permanent but show strong signs of catch-up. In other words, those workers induced to move involuntarily by temporarily low retention rates relative to similar workers at the same firm experience only temporary wage losses. Since we estimate the expected loss for workers at risk of moving at any moment in time, this is exactly the treatment effect of interest. Given our focus on workers leaving from larger training firms, it should be kept in mind that the effects may overstate losses for workers exiting small firms. Consistent with 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 an important fraction of mobility is voluntary. As further discussed below, these results have also important implications for interpreting OLS and IV estimates in situations in which researcher cannot control for both selection and sorting.

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 2 shows the simple average of real wages by intervals of the fraction 'other' movers (the instrument in levels) for three experience years. The three panels show that the linearity assumption underlying the results in Table 3 is justified.[35] Moreover, the relationship is negative and does not change over time. Panel B of Figure 2 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.[36] 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 Composition

The main results of Table 3 and Figure 1 are robust to several important specification checks. The IV estimates with fixed effects for the alternative specifications are shown in Table 5, whereas Figure 3 displays the full range of estimates for selected specifications. A 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. To

---

[35] 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 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.

[36] 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 (.0007), -.0024 (.0009), and -.0015 (.0007) for experience years 1, 3, and 5, respectively. Intercepts are .022 (.004), .012 (.006), and .007 (.004). 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.

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 2). As the results from the first stage regression in Column1 of Table 4 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 Figure 2). 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 4, 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. Using this definition, the relationship 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 this 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 4 displays the first stage coefficients. Figure 3 (Panel A) and Column 1 of Table 5 show the main estimates. The basic patterns shown in Table 3 and seen in Figure 1 are clearly confirmed. Those movers coming from firms from which a high fraction of workers who move 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 of movers across firms these estimates are reduced and the wage losses of movers decline even more strongly than before. The bottom panel of Table 5 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 a lower bound of the initial effect of involuntarily leaving the training firm.

Another concern expressed above is that changes in the sample composition might induce part of the observed reversion of losses. This might happen for example if among movers the best workers leave for military service 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. Second, since the panel-sample excludes many men leaving for military service, as an alternative we restricted total time spent out of covered labor force to be at most 6 months per year. Panel B and D of Figure 3 show the main results for these samples, while Columns 2 and 3 of Table 5 contain the final IV estimates. For the panel sample, 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.[37] The estimates for the less restrictive sample tell a very similar story.

Another 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.[38] Column 6 of Table 5 shows the basic IV estimate with fixed effects if all workers returning to their

---

[37] If we repeat this analysis with the narrowly defined instrument (Figure 3 Panel C, Table 5 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 likely to be due to sampling variation, since estimates tend to be clustered around zero.

[38] One can indeed show that retention rates are positively correlated with the probability of returning to the training firm, so the main specifications all include a dummy for whether a mover works at the training establishment. 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, in general displaced workers who return to their original employer are not.

original training firms are excluded. Compared to Column 5 of Table 3, 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. Results with the alternative instrument are similar, albeit initially even more negative and with a higher standard error (Column 7). Reversion remains a key feature of the data. The 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 a strong firm-specific component in 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 entry into a temporarily depressed regional labor market.

While the regional dimension plays only a small role, part of the permanent differences between firms can be explained by industry differences. Fixed effects for 3-digit industry explain 17.4% of variation in firms' retention rates in our sample; the respective number is 21.7% for training wages and 30.1% for wages one year after the end of training.[39] In contrast, fixed effects for training establishments explain 56.1%, 44.2%, and 42.9% of variance for retention rates, training wages, and wages, respectively. Explanatory power of industry effects is relatively low for retention rates and training wages, but high for wages in employment, not surprising in an environment of collective bargaining. Consistent with these numbers, 3-digit industry effects can explain part of the

[39] This holds for a 3-digit industry classification. Adding industry-cohort-interactions does little to raise explanatory power, consistent with prevalence of firm-level shocks. Detailed region effects play a minor role in levels or interacted.

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 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 characteristic 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 6 summarizes the results. The first stages are strong both with and without fixed effects (not shown).[40] 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, consistent with the fact that 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 may be stronger than in the full sample. Third, while the estimated wage differences from IV with firm fixed

[40] The first stage is shown at brackets of the demeaned instrument in Figure 4, Panel D.

effects show a clear sign of decline over time it remains significantly negative even after five years in the labor market (Column 5).[41]

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, wage losses of movers with respect to stayers should be driven partially 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 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.[42] These results further underscore the importance of controlling for firm characteristics when analyzing job displacements.

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.

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. Appendix E available from the our website indeed shows

[41] These results are robust to the full range of sensitivity checks. Note that although the basic IVFE estimator also implies a lesser degree of voluntary mobility out of large training firms, the alternative instrument shown in Column 6 is again more negative than OLS.

[42] This conclusion is consistent with results found by Krashinsky (2002) who pointed out the potential of firm size premiums in explaining losses of displaced workers and with Dustmann and Schoenberg (2002) who discuss differences in OLS estimates of wage losses by firm size in the German apprenticeship system.

that these firms generally pay lower wages, have higher turnover rates, and grow slower.[43] This result adds to a growing literature based on matched employer-employee data 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 2002) and on young German workers in general. 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

[43] 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 also explain these phenomena. Note that union representation among the firms in our sample is very high (about 90%) and thus cannot explain the differences in retention rates.

‘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.[44]

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 random process generating job mobility as in Acemoglu and Pischke (1998) would have to be invoked to explain the coexistence of 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 rates towards more desirable jobs. Since IV tends to be more negative than OLS rather than less, this is a prediction not borne out by the data.

Last, persistent effects from involuntarily leaving large training firms are 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 or rigid wage contracts. 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

This paper shows that sorting and selection mechanisms in the labor market for young workers biases OLS estimates of the cost of an early job loss, and presents new unbiased estimates

---

[44] 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.

of the short- and long-term effects of early displacements. A permanent negative bias arises if less able workers tend to be hired by firms with higher turnover rates, or if employers selectively displace their least able workers. A positive initial bias may arise if 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 exploits variation in firms' hiring rates of young apprentices 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. We interpret our estimates as local average treatment effects for those workers induced to move by temporarily low retention rates relative to the outcomes of similar workers at the same training firm at risk of moving involuntarily in other periods. These results are robust to changes in sample composition 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 premiums or internal labor 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 less able workers 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 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 estimates presented here suggest that crucial information is gained about the mechanisms of job and wage dynamics when sorting and selection are considered jointly. The paper thus 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 provide guidance for the case in which matched panel data on workers and firms is not available. Generally, since for young workers individual fixed effects will not remove the biases implied by selection or sorting, only a strategy of both instrumental variables and firm fixed effects will yield unbiased results. If this is not possible, OLS with and without firm fixed effects will understate the initial losses and overstate the long-run impact. An IV estimator without fixed effects, on the other hand, should capture some of the reversion but similarly overstate short and long-term losses. Last, additional measures of involuntary mobility or firm characteristics may help to gauge some of the potential biases of simple OLS estimates. The latter is particularly relevant, since sorting of workers along firms' turnover rates implies displacement is not a random event even controlling for selection within firms.

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).

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.

Bauer, Thomas K. und Stefan Bender (2004). 'Technological Change, Organizational Change, and Job Turnover.' *Labour Economics:* 2 (11) 265-291.

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 (1999). '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.

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 (2002). 'Mobility after apprenticeship – Evidence from register data.' Applied Economics Quarterly (Konjunkturpolitik) 48: 256-278.

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, Ohio State.

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.

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 (2003). 'The Long-Term Costs of Job Displacement Among Young Workers.' Industrial Labor Relations Review 56(4): 682-698.

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 (2003). *Monopsony in Motion: Imperfect Competition in Labor Markets*. Princeton: Princeton University Press.

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.

Murphy, Kevin M. and Finis Welch (1990). 'Empirical Age-Earnings Profiles.' *Journal of Labor Economics* 8(2): 202-229.

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, 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 (2002). 'Youth Labor Market in the U.S.: Shopping Around vs. Staying Put.' Review of Economics and Statistics, August, 462-482..

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.

Schwerdt, Wolfgang and Stefan Bender (2003). 'Was tun Lehrlinge nach ihrer Ausbildung? - Eine Analyse mit dem Linked Emloyer-Employee-Datensatz des IAB.' (What do Apprentices do after Training? An analysis with the IAB Linked Employer-Employee Data). *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung*: 36 (1) 46-59.

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.

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: 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	0.36	0.38	0.34
Fraction German	0.87	0.88	0.86
Average Training Duration	2.57	2.56	2.57
Fraction 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 difference in means is significantly different at a 1% level for all characteristics). The only characteristic changing over time is the wage, all other variables pertain to the training period.

Table 2: 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. Standard errors clustered at the establishment level are in parentheses.

Table 3: 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	-0.065	-0.065	-0.121	-0.108
		(0.0038)	(0.0029)	(0.0025)	(0.0077)	(0.0268)
	2	-0.093	-0.069	-0.065	-0.113	-0.064
		(0.0014)	(0.0013)	(0.0027)	(0.0072)	(0.0215)
	3	-0.093	-0.071	-0.064	-0.108	-0.035
		(0.0011)	(0.0011)	(0.0028)	(0.0071)	(0.0203)
	4	-0.091	-0.070	-0.062	-0.101	-0.005
		(0.0010)	(0.0011)	(0.0028)	(0.0071)	(0.0209)
	5	-0.092	-0.072	-0.063	-0.098	0.009
		(0.0011)	(0.0011)	(0.0030)	(0.0075)	(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 4: 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	0.135	0.604	0.151
		(0.0064)	(0.0169)	(0.0135)	(0.0128)
	2	0.739	0.152	0.654	0.173
		(0.0023)	(0.0171)	(0.0046)	(0.0126)
	3	0.748	0.161	0.675	0.189
		(0.0016)	(0.0172)	(0.0030)	(0.0127)
	4	0.750	0.164	0.683	0.195
		(0.0014)	(0.0172)	(0.0026)	(0.0127)
	5	0.754	0.167	0.696	0.205
		(0.0013)	(0.0174)	(0.0024)	(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 5: 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	-0.166	-0.155	-0.207	-0.228	-0.160	-0.333
		(0.0331)	(0.0279)	(0.0264)	(0.0317)	(0.0318)	(0.0413)	(0.0617)
	2	-0.135	-0.103	-0.098	-0.154	-0.166	-0.088	-0.189
		(0.0249)	(0.0237)	(0.0213)	(0.0269)	(0.0235)	(0.0297)	(0.0355)
	3	-0.064	-0.041	-0.046	-0.040	-0.063	-0.050	-0.088
		(0.0225)	(0.0233)	(0.0199)	(0.0246)	(0.0214)	(0.0275)	(0.0314)
	4	-0.013	-0.024	-0.005	-0.003	0.006	-0.011	-0.008
		(0.0242)	(0.0242)	(0.0207)	(0.0269)	(0.0244)	(0.0281)	(0.0349)
	5	0.015	-0.009	0.010	0.052	0.051	-0.002	0.001
		(0.0244)	(0.0261)	(0.0223)	(0.0302)	(0.0269)	(0.0295)	(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 6: 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	-0.093	-0.089	-0.152	-0.095	-0.188	-0.048
		(0.0058)	(0.0051)	(0.0045)	(0.0128)	(0.0137)	(0.0277)	(0.0191)
	2	-0.108	-0.097	-0.087	-0.145	-0.076	-0.163	-0.029
		(0.0021)	(0.0053)	(0.0049)	(0.0120)	(0.0127)	(0.0222)	(0.0144)
	3	-0.109	-0.098	-0.087	-0.139	-0.065	-0.107	-0.023
		(0.0017)	(0.0054)	(0.0052)	(0.0115)	(0.0125)	(0.0198)	(0.0133)
	4	-0.105	-0.093	-0.082	-0.131	-0.050	-0.084	-0.010
		(0.0016)	(0.0054)	(0.0050)	(0.0116)	(0.0127)	(0.0215)	(0.0130)
	5	-0.106	-0.094	-0.084	-0.126	-0.041	-0.054	-0.012
		(0.0016)	(0.0057)	(0.0054)	(0.0121)	(0.0137)	(0.0222)	(0.0131)
T-Statistics	**H0: 1=3**	0.42	1.68	0.85	2.42	-3.1	-3.5	-1.8
	H0: 3=5	1.28	1.59	1.33	3.14	-3.0	-3.6	-2.6
	H0: 1=5	0.50	0.36	1.56	3.92	-4.5	-4.9	-2.3

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: OLS and IV Estimates of Wage Losses from Leaving Training Firm At Graduation, With and Without Firm Fixed Effects - Main Sample

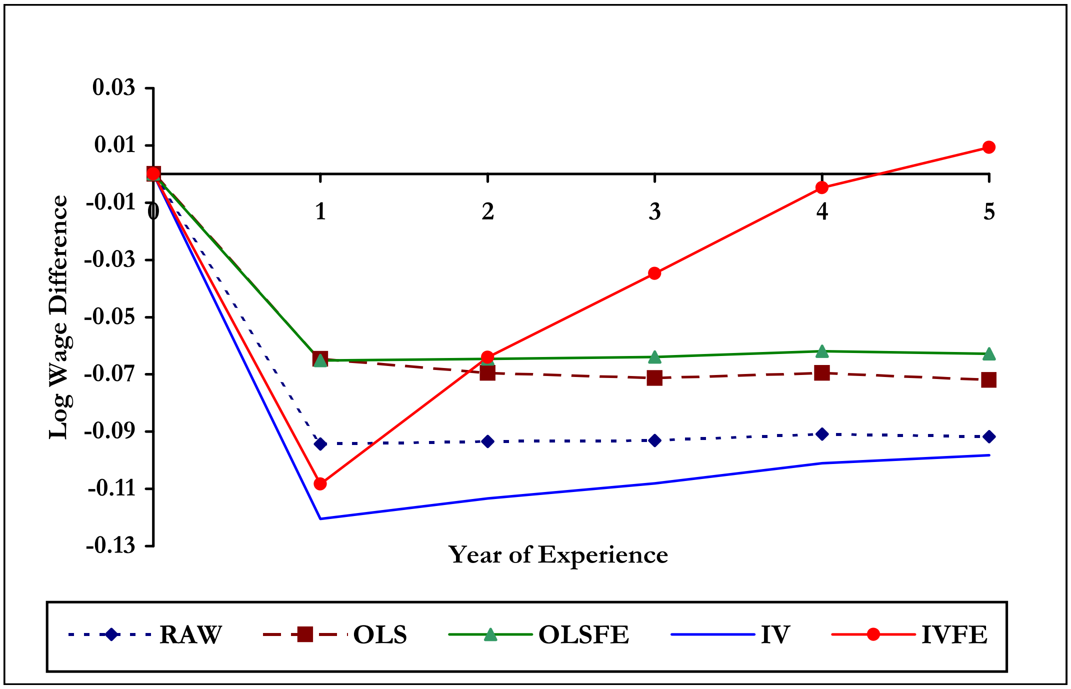


Figure 2: Average Real Wages By Intervals of Fraction 'Other' Apprentices Moving from Training Firm at End of Training

Panel A: Without Firm Fixed Effects | **Panel B: With Firm Fixed Effects**

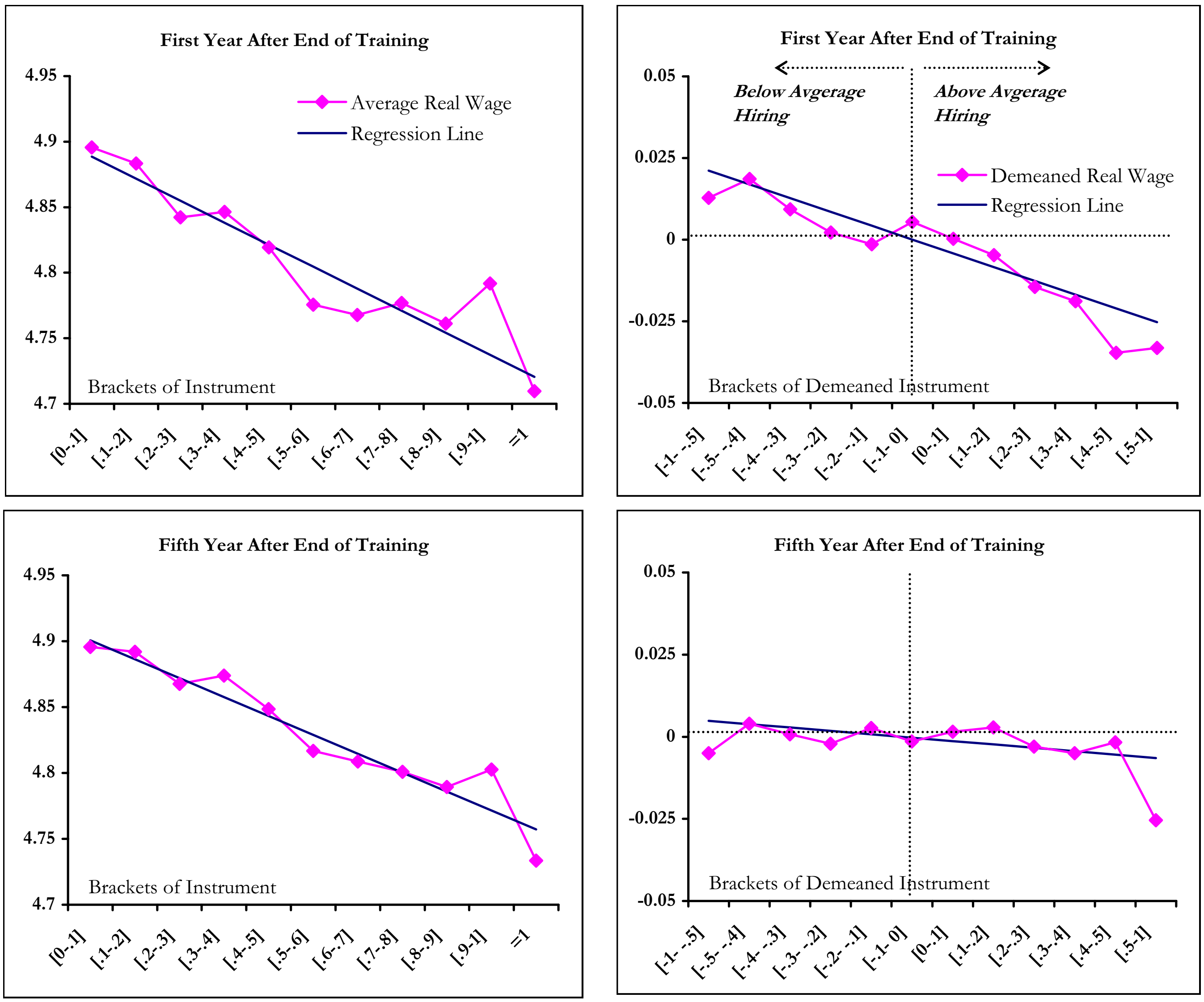
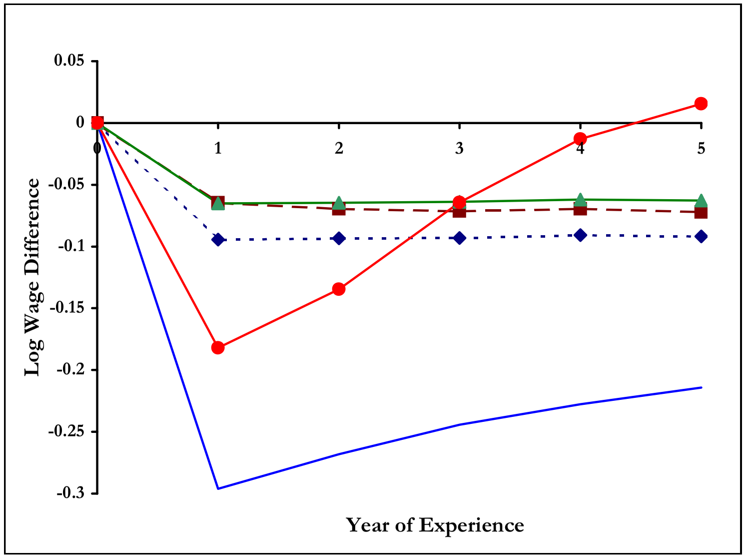
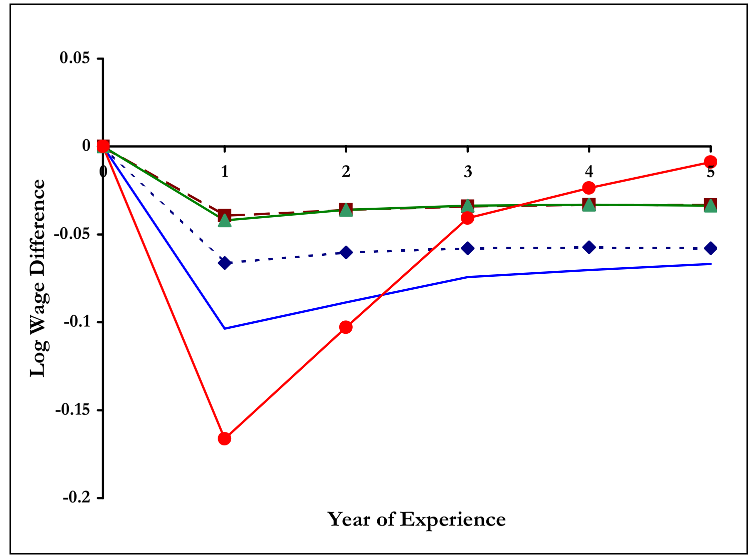


Figure 3: Sensitivity - Effects on Wages of Leaving Training Firm at Graduation, Various Specifications

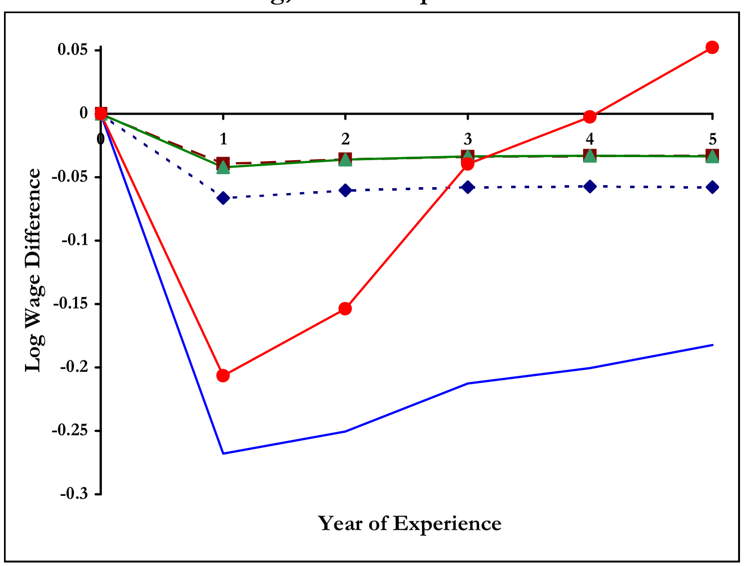
Panel A: Effect of Moving, Fraction Unemployed (IV2)



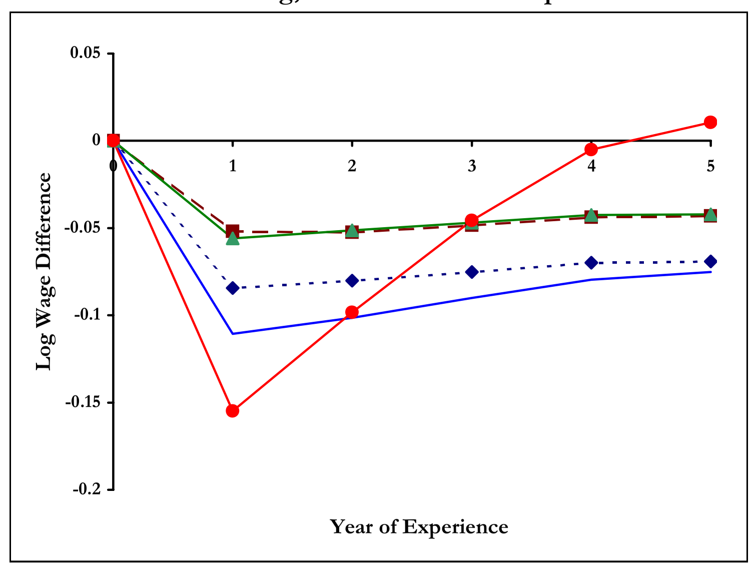
Panel B: Effect of Moving, Panel Sample



Panel C: Effect of Moving, Panel Sample with IV2



Panel D: Effect of Moving, Restricted OLF Sample



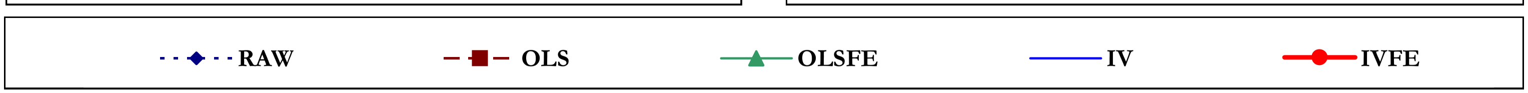
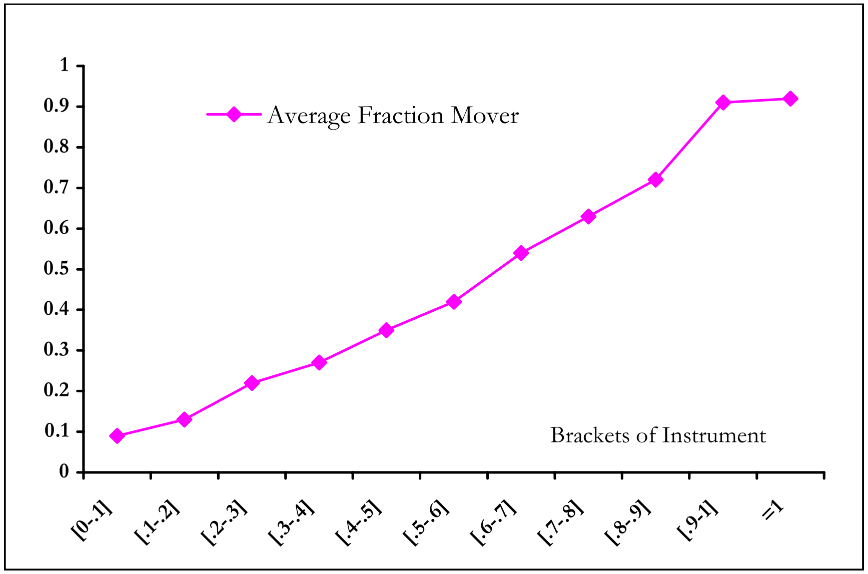
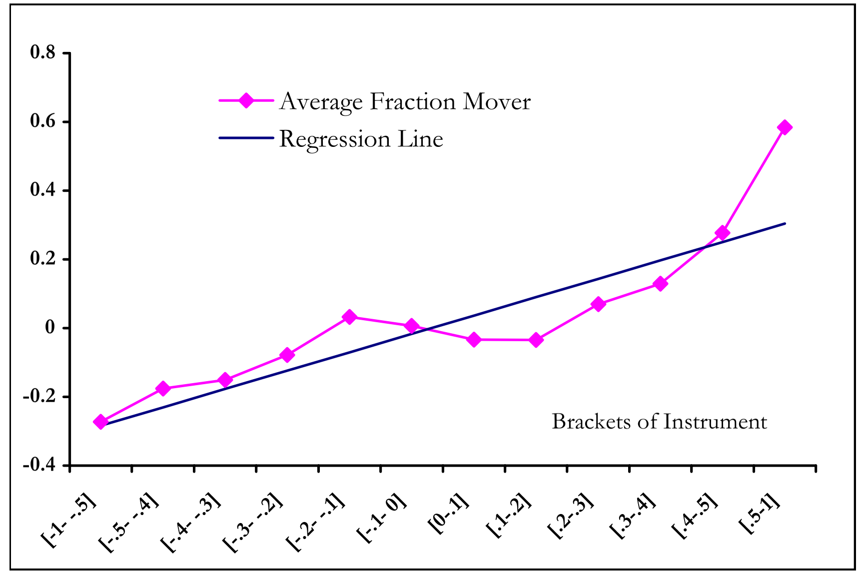


Figure 4: Fraction Movers by Intervals of Instrument, With or Without Firm Fixed Effects

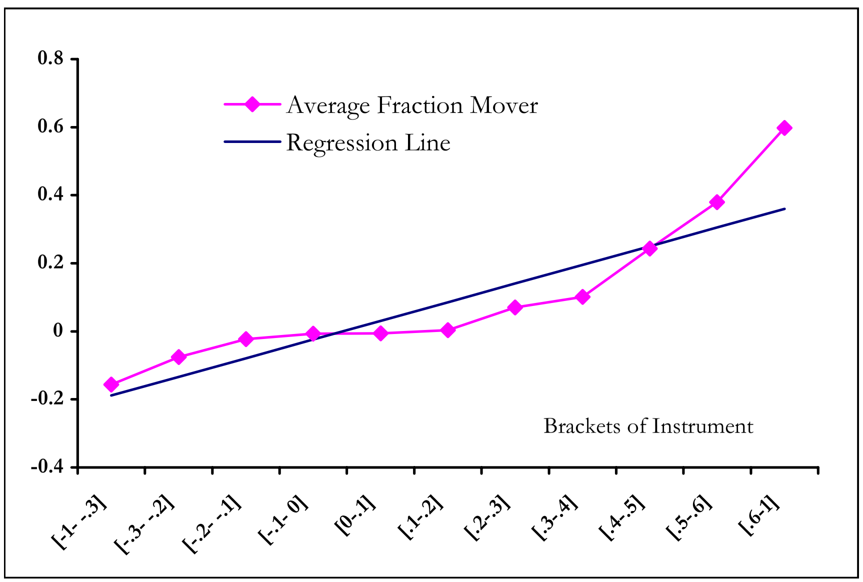
Panel A: Without Fixed Effects, Main Sample



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