CIRANO

Centre interuniversitaire de recherche
en analyse des organisations

Série Scientifique
Scientific Series

Nº 95s-27

WAGES AND MOBILITY: THE IMPACT OF EMPLOYER-PROVIDED TRAINING

Daniel Parent

Montréal
Avril 1995

CIRANO

Le CIRANO est une corporation privée à but non lucratif constituée en vertu de la Loi des compagnies du Québec. Le financement de son infrastructure et de ses activités de recherche provient des cotisations de ses organisations-membres, d'une subvention d'infrastructure du ministère de l'Industrie, du Commerce, de la Science et de la Technologie, de même que des subventions et mandats obtenus par ses équipes de recherche. La *Série Scientifique* est la réalisation d'une des missions que s'est données le CIRANO, soit de développer l'analyse scientifique des organisations et des comportements stratégiques.

CIRANO is a private non-profit organization incorporated under the Québec Companies Act. Its infrastructure and research activities are funded through fees paid by member organizations, an infrastructure grant from the Ministère de l'Industrie, du Commerce, de la Science et de la Technologie, and grants and research mandates obtained by its research teams. The Scientific Series *fulfils one of the missions of CIRANO: to develop the scientific analysis of organizations and strategic behaviour.*

Les organisations-partenaires / The Partner Organizations

- Ministère de l'Industrie, du Commerce, de la Science et de la Technologie.
- École des Hautes Études Commerciales.
- École Polytechnique.
- Université de Montréal.
- Université Laval.
- McGill University.
- Université du Québec à Montréal.
- Bell Québec.
- La Caisse de dépôt et de placement du Québec.
- Hydro-Québec.
- Banque Laurentienne du Canada.
- Fédération des caisses populaires de Montréal et de l'Ouest-du-Québec.
- Téléglobe Canada.
- Société d'électrolyse et de chimie Alcan Ltée.
- Avenor
- CIDEM

ISSN 1198-8177

Wages and Mobility: The Impact of Employer-Provided Training*

Daniel Parent†

Abstract / Résumé

Using data from the National Longitudinal Survey of Youth (NLSY) for the period spanning the years 1979-1991, this essay examines the impact of employer-provided formal training on the wage profile and on the mobility of young Americans making their transition to the labor market. By exploiting the longitudinal aspect of the data set, we are able to provide some control for unobserved individual and job-match heterogeneity by making use of the methodology proposed by Altonji and Shakotko (ReStud '87). The results show that (i) training with the current employer has a statistically and economically significant positive effect on the wage; (ii) employers seem to reward skills acquired through training with previous employers as much as skills they provide themselves; (iii) workers undergoing training have 18 percent lower starting salaries than other workers; this result is obtained by setting up a starting wage equation and by making use of a variable called "on-the-job training still in progress at the time of the interview" ; (iv) with a hazard model which makes use of multiple employment spells by the same worker (thereby allowing the implementation of fixed-effects methods akin to the conditional logit method), skills acquired through formal training programs provided by the current employer seem to be fairly specific. The upshot from these results is that formal on-the-job-training in the current job contains both a general component which the employer rewards up to its market value and a specific component which reduces mobility while not being rewarded.

En utilisant des données américaines du National Longitudinal Survey of Youth (NLSY), cette étude s'attarde à examiner l'impact de la formation dispensée par l'employeur sur le profil salarial ainsi que sur la mobilité des jeunes travailleurs faisant leur entrée sur le marché du travail. En exploitant l'aspect longitudinal de l'échantillon de façon à tenir compte de l'hétérogénéité non observée, les résultats montrent (i) un impact économiquement et statistiquement significatif de la formation sur le salaire dans l'emploi courant, (ii) un impact substantiel sur le salaire de la formation acquise avec les employeurs précédents, (iii) une réduction d'environ 18 % du salaire de départ pour les travailleurs en formation, et (iv) par un modèle de durée qui tient compte des épisodes multiples d'emploi (permettant alors l'utilisation de méthodes de type « effets fixes »), un degré substantiel de spécificité du capital humain acquis par le biais de programmes de formation dispensés par l'employeur. La conclusion à tirer de ces résultats est que le capital humain acquis contient à la fois une composante générale rémunérée également par tous les employeurs ainsi qu'une composante spécifique qui réduit la mobilité tout en n'étant pas rémunérée.

---

* This paper is a revised version of chapter 2 of my Ph. D. dissertation submitted to the Université de Montréal in December 1994. Special thanks to my supervisors, Thomas Lemieux and Bentley MacLeod, for their comments, suggestions and general guidance. I am also grateful to Joe Altonji and Bart Hamilton for their comments and suggestions. Financial support from the SHRCC is also gratefully acknowledged.

† Université de Montréal, CIRANO, C.R.D.E.

I. Introduction

Much of the debate surrounding the issue of whether wages rise with years of seniority has focused on the magnitude of the tenure effect and on the possible explanations for this effect, including the theory of human capital. Only recently have there been attempts at measuring directly the effect of accumulating human capital through training. Mincer (1988), Brown (1989), Barron, Black and Loewenstein (1989), Barron, Berger and Black (1993), and Levine (1993) have attempted to do so with training data that were of a qualitative and subjective nature while Lynch (1992) and Blanchflower and Lynch (1994) used a data set, the National Longitudinal Survey of Youth (NLSY), that has the actual number of weeks spent training.[1]

The objective of this paper is to provide yet another look at this question using the NLSY for the period spanning the years 1979 to 1991. The longitudinal aspect of the larger data set is then exploited to analyze the impact of explicitly taking into account unobserved heterogeneity on (i) the effect of various forms of training on the wage profile, ii) the extent to which workers changing jobs benefit from having been trained by previous employers, and (iii) the propensity to change jobs. Also, the question of whether workers pay for their training by accepting lower starting wages will be examined by using a variable , on-the-job training still in progress, that seems best suited for picking up that effect, if indeed workers do pay for the training expenses (which are shown to be incurred largely by employers).

The estimating strategy is the following: after estimating with ordinary and generalized least squares a standard wage equation augmented with the training variables, these variables as well as the tenure and experience variables will be instrumented in order to alleviate correlation problems between them and the individual and job-match fixed effects, following the procedure proposed by Altonji and Shakotko (1987). The results show that there are substantial returns to training, more particularly on-the-job training, whether it be with the current employer or with subsequent employers. In fact, the results indicate that the human capital acquired through training appears to be just as valuable (to the worker) with subsequent employers than it is with the current employer. This apparent portability of the skills acquired through on-the-job training suggests that we should be observing lower starting wages for workers who are undergoing training. Indeed, the results show that workers implicitly pay for these returns by having lower starting wages. This is shown by estimating a starting wage equation with ordinary least squares where some control is provided for unobserved heterogeneity. Next, a proportional hazard model is

[1] See also Altonji and Spletzer (1991) for another example of a study which makes use of data on the duration of training spells. Their paper, however, focuses on the determinants of employer-provided training.

estimated and the results show a substantial degree of firm-specificity for human capital acquired through training received with the current employer. This result holds even if control is provided for unobserved person-specific heterogeneity through the use of fixed-effect methods. Note, however, that when I control for such heterogeneity, training received with previous employers is shown to have no significant positive impact on mobility, contrary to the case in which there is no control for heterogeneity.

These results point toward the following conclusion: for this sample of young workers, employers react to market pressures by setting wages at the market level with no premium paid for the firm-specific productivity of the trained workers. To be more precise, the results from the wage equation and those from the hazard model indicate that formal on-the-job training in the current job contains at once a general component which the employer rewards up to its market value and a firm-specific component which reduces mobility but which does not appear to be rewarded. Thus, there is little evidence that these workers are paid much in excess of their outside option.

The rest of the paper is organized as follows. Section II briefly reviews the theoretical models underpinning the estimation. Section III provides a thorough description of the data set and of the various definitions of the training variables used in the estimation procedures. Particular attention is paid to the timing of training over the course of an employment relationship and on who pays the direct costs of training, observation that will help in the interpretation of the results. Section IV presents the estimating framework as well as a discussion of the results. Section V concludes the paper.

II. Theoretical Framework

According to the theory of general human capital as formulated in Becker (1975), workers who invest in general on-the-job training should pay the full costs and reap the full return from their investment. If firms pay the direct costs of general training, then workers should be expected to reimburse the firm by accepting a lower starting wage. In other words, during the training period, workers are paid below their value on the market. When training is completed, their value to the firm as well as their market value increase and competitive forces insure that they are paid at their market value. An immediate implication is that wages should rise with experience in the labor market since productivity increases with time in the labor market for those who get general training. Moreover, the theory predicts that investment in on-the-job training should decline as the workers age for basically two reasons: 1) the period over which the rewards will accrue to the workers gets shorter and 2) the opportunity cost of training increases as the wage increases over time. On the other hand, if workers

invest in entirely firm-specific on-the-job training, the theory suggests that the firm and the workers should share in the returns and the costs. The motivation for sharing the investment stems from the uncertainty surrounding the parties' post-investment behavior. For instance, were the workers paying the full costs of the investment, they would face the possibility of an employer-initiated separation in the period following the investment. They would therefore be incurring a capital loss since, by definition, the value of marginal product outside the firm is the same whether training has occurred or not. As for the firm, if it were to bear the full costs of training, the possibility that the workers would leave while the firm was in the process of harvesting the return on its investment would imply a capital loss for the latter. Consequently, both parties share the costs and the returns on the investment. The empirical implications are that 1) wages should be rising with seniority because productivity increases with seniority and 2) turnover should decrease when there is investment in firm-specific human capital.

Hashimoto (1981) subsequently provided a formalization of Becker's sharing hypothesis. He showed that sharing arises in response to transaction costs, not because of uncertainty concerning the behavior of parties in the post-investment years. More precisely, since it is costly for both parties to evaluate and agree on the worker's productivity in the firm and outside the firm in the years following the investment, there will be situations where the worker may quit or the firm may dismiss the worker even if separation entails a net loss. Therefore, parties will determine the optimum sharing ratio prior to undertaking the investment, thus minimizing the loss from non-optimal separations. Hashimoto shows that if there were no transaction costs, the sharing ratio would be irrelevant to the separation decisions. Although the main reason for sharing is different than in Becker's original argument, the empirical implication for the wage profile is the same, namely that wages should rise with seniority.

An even more efficient contract was derived by Carmichael (1983). Like Hashimoto, he assumed that after an initial period of training, information is revealed asymmetrically: only the firm knows the value of marginal product while only the worker knows his degree of job satisfaction. Since information cannot be exchanged ex post due to transaction costs, both parties make their separation decision independently. The worker leaves if his job satisfaction is below a certain threshold while the firm fires the worker if his revealed productivity is below a critical level. Carmichael argues that welfare can be improved by making both parties internalize the entire expected losses from a separation, regardless of who initiates it, instead of merely internalizing their share of the return, as in Hashimoto's contract. To achieve such a constrained optimum, Carmichael modified Hashimoto's structure by introducing seniority-based promotions. Suppose that there are two jobs in the second period, type 1 jobs and type 2 jobs. The productivity of the workers is identical in the

two types of jobs. After separation decisions are made at the start of the second period, all workers still with the firm enter type 1 jobs where they are paid w_2. Later in period, some of these workers are promoted to type 2 jobs on the basis of their seniority where they are paid $w_2 + B$ where B is a bonus which is determined ex ante when the parties agree on the contract. By assumption, there is a fixed number of type 2 jobs, therefore a layoff can only save the firm w_2 since a junior worker will immediately fill the vacant position. Thus, the incentive for the firm to fire any high wage worker is reduced. In effect, junior workers are acting as a third party by benefitting from any separation since the layoff of a senior worker increases their own seniority, thus improving their chance of eventually getting a type 2 job. Having determined the second period wage schedule, the first period wage is adjusted so that the expected value of the contract equals that available to the worker in the market. Carmichael shows that this system of promotions based on seniority makes the separation decisions more efficient. Two implications of Carmichael's model are worth emphasizing. The first one is that the more senior workers will typically be earning more than their value of marginal product to the firm. Thus the theory of specific human capital can respond to the challenges posed by Lazear (1979) concerning mandatory retirement and by Medoff and Abraham (1980) who provided evidence that wages might rise with seniority even if productivity did not. Secondly, his optimal contract implies that workers pay the full cost of the investment in the form of lower starting wages. It should be noted Carmichael implicitly assumes that the firm can commit to the wage schedule derived ex ante. This is an important assumption because it assures that the firm will not change the number of type 2 jobs available, even though it would be profitable to do so once a junior worker is due for a promotion.

By explicitly taking into account the possibility of contract renegotiation, Macleod and Malcomson (1993a,b) show that firms making specific investments need not offer above market clearing wages to their workers. Since it is impossible or too costly to write a complete contract ex ante that would prevent renegotiation, there will be situations in which the parties renegotiate an incomplete contract. This happens if and only if one party prefers to break the contract though it is still efficient for the parties to continue their match. In this case, one party is prepared to give up some of its gains in order to induce the other party not to breach. To be more precise, a firm and a worker have formed a match in period one and the firm has made a specific investment that affects second-period payoffs, and the worker invests only in general skills. Random shocks to the productivity of the worker and to his outside market wage become known to both parties at the end of the first period. If the draw the worker receives makes the outside option preferable to the negotiated wage at the beginning of the relation, assuming that it is efficient to continue the match, the firm renegotiates the worker's wage to the level of the outside option. Similarly, if the firm prefers to fire the worker rather than pay her the contract wage, the worker is prepared to take

a pay cut in order to avoid a separation. Naturally, there are situations where no renegotiation occurs after the realization of the random shocks. In that case, both parties prefer to trade at the contract wage that involves some sharing of the return on the firm's investment. When only the firm invests in the worker, a contract which avoids the "holdup" problem (see Williamson et al. (1975)), is one in which the worker receives her market alternative each period, and hence wages do not rise with tenure. When both parties make relation-specific investments, Macleod and Malcomson show that if the worker accepts the terms of the efficient contract, then the measured effect of tenure on wages is positive.

III. The Data

The National Longitudinal Survey of Youth data set surveyed 12,686 young males and females who were between the age of 14 and 21 in 1979[2]. It contains detailed employment histories of the respondents thereby permitting the construction of relatively error-free variables for tenure as well as for the total experience accumulated since the beginning of one's full time transition to the labor market. Of particular interest is the existence of a detailed account of the training histories of the respondents, with the starting and ending dates of up to four spells[3] of training lasting at least four weeks since the last interview. While the data give a good account of the formal aspect of skill acquisition it is missing out on the informal aspect and, possibly, on formal programs of short duration. The training spells are broken down into up to 12 categories (the actual number varies from year to year, although it stays at twelve from 1988 onward) which were then classified in three groups in the same fashion as Lynch (1992). Categories 8 and 9 (company-provided training and seminars or training programs at work but not run by employer) were classified as on-the-job training, category 3 represents apprenticeship programs and categories 1 (business colleges), 2 (nurses programs), 4 (vocational-technical institutes), 5 (barber and beauty), 6 (flight school), 7 (correspondence courses), 10 (seminars or training programs outside of work), 11 (vocational rehabilitation center) and 12 (other) were classified as off-the-job training. It is also known whether the programs were completed, not completed or still in progress at the time of the interview. Since these workers can be followed as they move from employer to employer, it is then straightforward to calculate the accumulated weeks of training with all previous

[2] The response rate was at 71% in 1991.

[3] Actually, three spells from 1979 to 1986 and four spells from 1988 to 1991. The fact that no questions concerning training were asked in the 1987 interview does not pose any particular problems (except possibly accuracy of recall) because the starting and ending dates given in 1988 allow for the appropriate assignments with respect to the 1987 and 1988 interview dates.

employers[4] by each type of training. I also know the identity of the party who incurred the direct costs of the training programs for the years 1988-1991. Even though this information is not available for the whole period considered, the four years for which it is available provide strong evidence that employers are paying for OJT.

Lynch (1992) could only rely on a 1983 cross-section of the data set to get most of her results while Blanchflower and Lynch (1994) followed 1979,1980 and 1981 subsamples of 18-year-olds until they were 25 years old. Here the major restriction that I impose is in terms of entering the labor market on a full-time basis.[5] The people who were considered as meeting that criterion were (i) those whose primary activity was either working full-time, on a temporary lay-off or looking actively for a job, (ii) those who did not return to school on a full-time basis within six years[6] and (iii) those who had worked at least half the year since the last interview and who were working at least 20 hours per week. Individuals excluded from the sample are those younger than 18, those that had been in the military at any time, the self-employed, the ones whose jobs were part of a government program and the ones working without pay, those who were in the farming business and also all public sector employees. Another restriction was to exclude those that occupied two jobs (or more) on a full-time basis because it is impossible to determine with which of these employers the training is taking place, thereby possibly falsifying the results. For the earnings equation, I am then left with 29,020 observations while the starting wage model has 13,394 observations. For the hazard model, only those that had at least two completed spells of employment were retained to allow the implementation of a fixed-effect estimation methodology, which leaves me with 8,097 spell-workers.[7]

Some summary statistics of the sample are provided in table 1 while table 2 gives us a breakdown of the number of employees who were trained by types of training and by industrial aggregates.

---

[4] Note that all training undergone by the individuals while they were not employed has been excluded from the calculations.

[5] We also limit the sample to workers who are at least 18 years of age.

[6] The choice of six years as a cutoff point is arbitrary, and hence debatable. The idea is to exclude those that make "quasi-permanent" transitions and who might be considering returning to school a few years down the road. It seems reasonable to assume that few people would enter the labor market while planning to leave it in six years or more to go back to school. The same could not be said if we were considering a one to three year (say) horizon. In any event, the results were left unchanged if all school returners were excluded.

[7] Note that people were allowed to work less than 20 hours per week only if they were undergoing training. Also, to avoid left-censoring, individuals who had started working prior to January 1, 1978 were eliminated.

Next, table 3 shows some summary statistics comparing those that have received training (and who have completed their programs) with those that have not. To be more precise, the sample is restricted to workers who had positive values for the training variables at any moment in the employment relationship. For example, workers those that had no training in the first two years, but who received training thereafter are included in the sample. This way of calculating the means avoids any upward bias that would have resulted had these people been excluded. Beside the wage differentials, it is very interesting to note that the trained workers have enjoyed longer careers with their respective employers compared to those that have never been trained. Also, it is worth noting that the additional experience that the trained workers have is usually accounted for by the additional tenure. Therefore, taking these numbers at face value, it would seem that all types of training are of a fairly specific nature.

If all training costs are paid by the employers and the skill enhancement programs are, at least to some degree, portable, then you would expect the workers to bear some portion of the costs by receiving lower starting wages. As is shown in table 4, it appears that employers are bearing much of the costs of all types of training programs, especially the programs corresponding to on-the-job training where between 1988 and 1991, 96.5% of the direct costs were paid by the employers. Although that information is not available for the whole length of the panel, it seems very unlikely that a dramatic shift would be observed in the event that it was. Interestingly, even the categories of training programs which are labelled as off-the-job were paid for in large part by the employers, suggesting that many of these programs were considered as relevant to the job, thereby straining a little bit the usefulness of distinguishing between on-the-job and off-the-job programs.

The issue of job relevance of the training programs can also be examined by looking at the timing of the programs i.e. how soon after starting working with their current employer do workers begin their training. One expects that the sooner the worker receives training after being hired, the more relevant to the current job is the program. A firm making specific investments maximizes its profits by pocketing the returns from these investments as soon as possible. Figures 1 to 4 allow us to see the extent to which training starts early in the employment relationship. In figure 1, the number of months elapsed since the beginning of the job has been plotted against the number of workers undergoing their first (and only the first) on-the-job program with their new employer[8]. It is clear from this graph that much of the training occurs very early in the relationship. The same can be said for off-the-job and apprenticeship

[8] Note that not all workers undergoing their first program are represented here; I have set the cutoff point at 28 months.

programs although to a lesser extent regarding OFT (see figures 2 and 3). This pattern of early training is important for estimation purposes since most if not all studies that have used training data of a qualitative and subjective nature have been unable to find convincing evidence of workers paying for their training through lower starting wages. As Levine (1993) recognizes, it may be that such data are too coarse to pick up that effect if it is there. In the next section, I find evidence for this hypothesis by focusing on training that is still going on at the moment of the interview.

The tenure-training (OJT) profile is illustrated in figure 4. To construct the training profile, the sample is restricted to workers who had (OJT) training with their current employer. The average number of weeks of completed training for these workers by tenure intervals of one year are then computed[9]. As the theory of human capital predicts (e.g. Ben-Porath (1967)), all investment is heavily frontloaded. The profile fluctuations following the first year reflect the turnover of workers in the sample.

IV. Results

IV.1 Earnings Equation Estimates

Consider the following log wage equation augmented with training variables:

$$\ln w_{ijt} = \beta_1 T_{ijt} + \beta_2 Exp_{it} + \beta_3 TCJ_{ijt} + \beta_4 TPJ_{it} + \alpha_i + \theta_{ij} + \epsilon_{ijt} \qquad (1)$$

where w_{ijt} represents the real hourly wage of person i in job j at time t, T is tenure, Exp is total labor market experience, TCJ is training in current job, and TPJ is training in previous jobs. All other covariates, including higher order terms for experience and tenure, are ignored for ease of presentation. The unobserved heterogeneity components can be decomposed into an individual effect (α_i) and a job-match effect (θ_{ij}).[10] The person-specific effect can be seen as representing unmeasured aspects of each individual's earning ability while the job-match component represents the unknown (to the econometrician) quality of the employment relationship stemming from search activity, for example. Both of these effects are assumed to be time-invariant. It is important to reiterate that the training variables available in the NLSY reflect the formal aspect of the process by which workers accumulate human capital, in that only

[9] For the first year of tenure, it is also the case that by the six-month mark, workers have on average over 12 weeks of completed OJT training.

[10] See Chamberlain ('82) for a discussion of nuisance parameters in panel data.

training programs which last at least one month are accounted for. No doubt that we miss out on the informal aspect of the training as well as on the formal aspect for short (i.e. less than one month) programs.

The training covariates for each category of programs (on-the-job, off-the-job and apprenticeship) are partitioned into two parts: training received with the current employer and all training received with previous employers, and all these variables have been converted from weeks to years (and fractions thereof). Column (1) of table 5 shows the results obtained by using OLS.[11] As can be seen, all three types of training are rewarded by the current employer with completed OJT (OJTCC) being rewarded the most. It also appears that human capital accumulated through formal training programs with previous employers is quite portable from firm to firm. However, the same sort of individual and job-match heterogeneity biases which affect the coefficients on tenure and experience are likely to be present regarding the training variables. More able individuals (those with high α's) may have enjoyed careers that were interrupted less frequently by unemployment spells and they may have received more training, both with the current employer and with previous employers, while better matches (high θ's) are likely to be formed if you have more experience and more training (due to human capital effects and search effects). Also, tenure and training with the current employer are possibly correlated with the job-match component. Dropping for the moment the assumption of a time-invariant job-match component, let's suppose that it can be written as

$$\theta_{ijt} = \psi_1 T_{ijt} + \psi_2 Exp_{it} + \psi_3 OJTCC_{ijt} + \psi_4 OJTCP_{it} + \omega_{ijt} \qquad (2)$$

where ω_{ijt} is assumed to be orthogonal to the regressors. The other training variables are not shown for ease of presentation but the same analysis applies to them. The discussion above suggests that ψ_2 and ψ_4 are positive. In the context of maximizing behavior on the part of a worker who faces a wage distribution, Topel(1991) argues that ψ_1 is negative once we control for experience, assuming there is a tenure effect. If there is no tenure effect, then ψ_1 equals zero. Regarding ψ_3, given that we control for both experience and tenure, it seems reasonable to argue that the job-match component of wages is non-negatively correlated with on-the-job skill acquisition. Substituting equation (2) into equation (1) we get

$$\begin{aligned} \ln w_{ijt} = {} & (\beta_1+\psi_1)T_{ijt} + (\beta_2+\psi_2)Exp_{it} + (\beta_3+\psi_3)OJTCC_{ijt} \\ & + (\beta_4+\psi_4)OJTCP_{it} + \alpha_i + \omega_{ijt} + \epsilon_{ijt} \end{aligned} \qquad (3)$$

[11] All results are calculated using the weighted sample.

We see from equation (3) that although we are interested in the β's, using ordinary least-squares will produce estimates of composite effects and the regressors would still be correlated with α. To provide some correction for these problems, the instrumental variable (IV) methodology of Altonji and Shakotko (1987) is used and is extended to the training variables[12]. All variables pertaining to training with the current employer, as well as tenure, are instrumented with their deviations from job-match means, whereas experience and training with previous employers are instrumented with their deviations from individual means. The instruments for training with the current employer and tenure are, by construction, uncorrelated with the job-match component while the instruments for training with previous employers and experience are, also by construction, uncorrelated with individual component. This methodology is preferred to first-differencing on the ground that differencing has greater potential to further enhance any measurement errors present in the data as compared with using deviations from means.

Another estimation problem pertains to the fact that residuals are likely to be serially correlated given that the same individuals are followed over time. Consequently, to provide correction for this problem, equation (1) is reestimated with generalized least-squares (under the assumption of a fixed individual effect along with an i.i.d. term), both with and without the IV methodology proposed. The results are reported in columns (2) and (3) of table 5.

Focusing on on-the-job training first, we can see that most of the results are qualitatively the same as with OLS, except that the coefficients are smaller both with GLS and IV-GLS. Therefore, the returns estimated with OLS cannot be attributed solely to job-match or individual heterogeneity. Curiously, for off-the-job training programs, GLS and IV-GLS estimates are much larger than with OLS. Thus, it seems that correcting for heterogeneity has the effect of removing the downward bias that poorly matched or low-ability workers cause to the estimated coeeficients. This contrasts with the results on apprenticeship programs where the results suggest that well matched and high-ability individuals are selected into the programs. Given the few workers who undertake apprenticeship (63), a word of caution seems warranted here regarding the results pertaining to this type of training. The basic conclusion that is to be drawn from these results is the portability of formal on-the-job and off-the-job training programs.[13] It appears that all completed programs of OJT and OFT are fairly general. Also, since the OJTCC (OFTCC) and OJTCP (OFTCP) coefficients are not significantly different, we are led to conclude from these results that current employers

[12] See also Finnie (1993) for an extension of the method of Altonji and Shakotko to the experience variable.

[13] In fact, a simple F-test on the equality of the coefficients associated with OJTCC and OJTCP could not reject the hypothesis that the two are indeed equal.

set the wage of their trained workers at pretty much the market level following completion of the programs. Already the summary statistics in table 3 provided a clue that employers reacted to market pressures. There we saw that workers who had received OJT stayed with their employers a full year and a half longer on average. This fact and the fact that OJT seems to be, at least to some degree, general (thereby increasing outside job opportunities) is not really contradictory; rather, it just means that employers manage to keep their trained workers by paying them their worth on the outside market.

IV.2 Specification of the Hazard Model

To gain a better perspective on the degree of firm-specificity of the human capital acquired with the current employer through formal training, a proportional hazard model is estimated by using Cox's partial likelihood approach.[14] The advantage of this approach is that it avoids the need to specify a baseline hazard function (actually, all common factors, including the baseline hazard, cancel from the partial likelihood). Another advantage of Cox's methodology is that it is possible to eliminate all individual-specific factors, including individual-specific baseline hazards, provided that multiple spells of employment are available for each worker. To be more specific, let's suppose that for worker i we have n_i spells (ordered by their increasing length) and that the duration for each spell is denoted t_{ij}, where j stands for the spell number. Assuming all spells for the same person are independently distributed given her heterogeneity parameter, I can write the hazard functions as:

$$\lambda_{ij}(t) = \exp(\beta' X_{ij}(t) + \alpha_i)\lambda_{i0}(t), \quad j = 1,\ldots,n_i \; ; \; i = 1,\ldots,N. \tag{4}$$

Then, it can be shown that the partial log-likelihood function is equal to:

$$L_P = \sum_{i=1}^{N} \sum_{j=1}^{n_i - 1} \ln\left(\frac{\exp(\beta' X_{i(j)}(t_{i(j)}))}{\sum_{k=j}^{n_i} \exp(\beta' X_{i(k)}(t_{i(j)}))}\right) \tag{5}$$

where the denominator corresponds to the risk set of worker i. Note that both α_i and λ_{i0} do not appear in equation (5). I should point out that a rigorously specified hazard model dealing with time-varying covariates such as the training variables ought to condition on the complete time-paths (past and future) of the covariates, especially if

[14] See Kiefer (1988) for a survey of hazard models or Lancaster (1990) for a detailed textbook exposition.

these covariates are considered endogenous to the leaving process. This issue is a difficult one to treat and I approximate the paths by using the current values (as of t) for these variables.

The results are presented in table 6. The effect of correcting for unobserved individual heterogeneity, while changing the magnitude of the coefficients (they are actually larger in absolute value), does not alter the basic conclusion that on-the-job training with the current employer seems to have a sizeable impact on the conditional probability of leaving.[15] Consequently, it can inferred from these results that the skills accumulated are fairly firm-specific. On the other hand, it seems that skills acquired with previous employers are not a significant factor of increased mobility once control is provided for unobserved individual heterogeneity. However, we should not be so surprised at such a result. It is a well known empirical fact that mobility tends to decrease with experience, presumably because workers, as time goes by, are able to find increasingly better matches.[16] Yet, labor market experience serves as a proxy for general human capital in a standard earnings equation. Given that skills accumulated with previous employers, as measured by the number of weeks spent training, are positively correlated with experience, then it is likely that this variable is correlated with the quality of the match in much the same way as experience is. Therefore, even if the coefficient associated with previous training were to be negative and significant, it cannot be concluded that skills acquired through previous training are specific to the current firm. By focusing on the probability of exiting from the first job, thereby circumventing to a degree the problem created by the positive correlation between match quality and experience, Lynch (1991) finds evidence of a positive impact of prior off-the-job training on mobility. Thus, she concludes the skills acquired acquired through this type of training is fairly general.

The results in table 6 coupled with those in column (2) of table 5 indicate that formal on-the-job training in the current job contains at once a general component which the employer rewards up to its market value and a specific component which reduces worker mobility. It is then plausible to suggest that firms tend to keep their trained workers longer as compared with other workers because they are more productive.

---

[15] Note that the effect of failing to control for heterogeneity is to bias the coefficients downward in a partial likelihood framework (see Lancaster (1990)). Therefore, it is not possible to conclude that the increase in the coefficient associated with OJTCC is due solely to the fact that I control for the presence of low-ability individuals.

[16] See Topel and Ward (1992) on that subject.

IV.3 The Impact of Training on the Starting Wage

The NLSY does not contain any information concerning the costs of these training programs and hence we cannot determine if the workers pay the totality of the costs in the form of lower starting wages. However, given that the trained workers seem to be paid at a level approaching their market-wide marginal product and given that from table 4, it is known that between 1988 and 1991 employers paid over 95% of the direct costs of OJT programs, we should be finding that the workers indirectly pay for part of these programs in the form of lower starting wages.

The way to capture that effect is by setting up a starting wage equation and by using a variable indicating that training is still going on. OJT which has not been completed (workers quit on the programs or the program was not completed for other reasons) will not do for the following reason. Assuming that workers should be paying for their training by having lower wages at the start, the results above show that market pressure forces employers to pay their trained workers at their marginal product following completion of OJT.[17] Then, there is no reason to believe that things are much different for workers who quit on a program. More precisely, their wage should be reset at the value of marginal product just like it would have been had they not undertaken any training program[18]. If I use instead the number of weeks of OJT training undertaken by the worker and still going on at the time of the interview, there lies the best hope of capturing any effect. There remains the problem of unobserved heterogeneity. Authors (*e.g.* Barron, Black, Loewenstein (1989)) have generally acknowledged that the effect of job-match or individual heterogeneity biases will be to underestimate the impact on the starting wage, even possibly masking it totally, the reason is that more able persons may receive more training and may be paid more even if they are undergoing training as compared with workers who are not being

---

[17] To quote Mincer (1962) p.53: "A direct computation of foregone earnings of workers engaged in on-the-job training would be possible if data were available on their earnings during and after the period of training, and on earnings of a comparison group of workers who have the same amount of formal schooling and are otherwise similar to the trainees, but do not receive any on-the-job training. Presumably, the latter would have a flatter age-earnings profile than the former."

[18] Uncompleted OJT never came out as significant in any of the regressions although it did show up as negative whether it was instrumented or not. Perhaps this reflects some sluggishness on the part of the wage as it may not be reset immediately after quitting training. Another possible explanation is that employers revise (downward) their expectation regarding the productivity of the workers. In any event, these problems do not occur if we use ongoing training at the time of the interview.

trained.[19] Therefore, any result found should probably be seen as a lower bound on the true result. However, some control is provided for unobserved heterogeneity with a variable indicating the maximum amount of training that person will have received by the end of her relationship with the employer.[20] If it is true that higher ability individuals are likely to be paid more and to receive more training, then that variable should be picking up some of that effect. Just to make clear the methodology employed here, I used only the first wage observations on each jobs and then used OLS. Results are reported in table 7. Even though only 134 observations had a positive value for OJT still going on (OJTOC)[21], the effect is relatively large and significant: workers do seem to pay for their OJT through lower starting wages. The same cannot be said for OFT and APP but this is not really that surprising given that over 25% of OFT programs undertaken between 1988 and 1991 were not payed by the employer (35% for apprenticeship) thereby introducing more noise in the model. Also, when the control for unobserved heterogeneity is added to the equation, the effect is larger by about one percent.

To summarize the results I have obtained, it seems that OJT and OFT with the current employer is neither entirely general (otherwise there would be no reason for it to reduce mobility) nor is it entirely specific (since then, having been trained by previous employers should have no impact on the current wage), but is rather a blend of these two extremes. Also, there is no evidence that employers pay their trained workers above their market-wide marginal product. Finally, workers appear to implicitly pay for the degree of portability of their newly acquired skills by having lower starting wages.

V. Conclusion

The idea that general skills accumulated through training should be paid for by the employee is a central prediction of the theory of human capital. However, all efforts directed at verifying that prediction have been practically fruitless, so much so

---

[19] Actually, this hypothesis goes back to Mincer (1962). On page 51 (footnote 4) he writes "...Greater learning from experience is characteristic of workers with greater motivation and ability, and their earnings at the early stages of the career may in some cases be as high or higher than those of other workers. But such finding that people with greater ability have higher productivity than others at any given stage of experience does not negate the existence of investment in on-the-job training, though it may bias the estimation of its magnitude."

[20] See Abraham and Farber (1987) for another application of a similar control.

[21] Given that much of the training takes place very early in the employment relationship (see fig. 1), we should not be surprised to find so few respondents still in training at the time of the interview.

that some researchers (e.g. Barron, Berger and Black (1993)) have started interpreting the lack of evidence of workers paying for their training through lower starting wages as possible evidence of dual labor markets. While the latter might indeed exist, this paper has provided direct evidence that workers do in fact have lower starting wages while they are going through training. Also, it has been shown that firms respond to market pressures by setting the wage for trained workers at the market level even if, as the hazard model results show, a sizeable portion of the investment in on-the-job training is specific. Therefore, the evidence concerning training is not consistent with the firm and the worker sharing the rent on the specific part of the investment. I should mention again that these results hold for formal training programs which naturally represent only a portion of the process by which workers accumulate skills on their current job. It could be that the informal process of skill acquisition makes the workers more indispensable to the firm which could then ill-afford to lose them and would consequently, perhaps, be willing to share the rent over the return on the specific capital embodied in these workers.

REFERENCES

Abraham, K. & H. Farber, "Job Duration, Seniority, and Earnings." *American Economic Review* 77 (June 1987): 278-97.

Altonji, J., & R. Shakotko, "Do Wages Rise with Seniority?" *Review of Economic Studies* 54 (July 1987): 437-59.

Altonji, J.G., J. Spletzer, "Worker Characteristics, Job Characteristics, and the Receipt of On-the-Job Training", *Industrial and Labor Relations Review*, vol. 45, No. 1 (October 1991): 58-79.

Barron, J., M. Berger, D. Black, "Do Workers Pay for On-the-Job Training?" Center for Business and Economic Research, College of Business and Economics, University of Kentucky. Working paper no. E-169-93. Lexington, Kentucky, 1993.

Barron, J., D. Black, Loewenstein, M., "Job Matching and On-the-Job Training", *Journal of Labor Economics*, 1989, vol. 7, no.1.

Becker, G.S., *Human Capital*, The University of Chicago Press, 1975.

Blanchflower, D.G., L.M. Lynch, "Training at Work: A Comparison of U.S. and British Youths", in Lisa M. Lynch ed., *Training and the Private Sector*, NBER Comparative Labor Markets Series, The University of Chicago Press, 1994: 233-260.

Brown, J., "Why Do Wages Increase with Tenure? On-the-Job Training and Life-Cycle Wage Growth Observed within Firms." *American Economic Review* 79 (December 1989): 971-91.

Carmichael, L., "Firm-Specific Capital and Promotion Ladders", *Bell Journal of Economics* 14, 1983: 251-258.

Chamberlain, G., "Multivariate Regression Models for Panel Data", *Journal of Econometrics*, 18, 1982: 5-46.

Finnie, R., "Tenure, Experience, and Men's and Women's Wages: Panel Estimates from the National Longitudinal Survey of Youth". Groupe de Recherche en Politique Economique, cahier de recherche 9305, Université Laval, 1993.

Hashimoto. M., "Firm-Specific Human Capital as a Shared Investment."*American Economic Review* vol.71, no. 3, June 1981: 475-482.

Kiefer, N.M., "Economic Duration Data and Hazard Functions", *Journal of Economic Literature*", vol. XXVI (June 1988): 646-679.

Lancaster, T., *The Econometric Analysis of Transition Data*, Econometric Society Monographs No. 17, Cambridge University Press, 1990.

Lazear, E.P., "Why Is There Mandatory Retirement?", *Journal of Political Economy* 87, December 1979: 1261-84.

Levine, D.I., "Worth Waiting for? Delayed Compensation, Training, and Turnover in the United States and Japan. *Journal of Labor Economics*, vol. 11, no.4, October 1993

Lynch, L., "The Role of Off-the-Job vs. On-the-Job Training for the Mobility of Women Workers." *A.E.R. Papers and Proceedings* May 1991: 151-56.

Lynch, L., "Private-Sector Training and the Earnings of Young Workers." *American Economic Review* 82 (March 1992): 299-312.

MacLeod, W. Bentley & J. M. Malcomson, "Specific Investment and Wage Profiles in labour Markets." *European Economic Review* 37, 1993a: 343-354.

MacLeod, W. Bentley & J. M. Malcomson, "Investment, Holdup, and the Form of Market Contracts", *American Economic Review* 83(4), 1993b: 811-837.

Medoff, J., K. Abraham, "Experience, Performance, and Earnings", *Quarterly Journal of Economics* 95, December 1980: 703-36.

Medoff, J., K. Abraham, "Are Those Paid More Really More Productive?", *The Journal of Human Resources*, 41, 1981: 186-216.

Mincer , J. "On-the-Job Training: Costs, Returns, and Some Implications", *Journal of Political Economy*, Supplement, October 1962: 50-79.

Mincer, J., "Job Training, Wage Growth, and Labor Turnover." Working Paper no. 2690. Cambridge, Mass.: National Bureau of Economic Research, August 1988.

Topel, R.H., M.P. Ward, "Job Mobility and the Careers of Young Men", *Quarterly Journal of Economics*, 1992: 439-479.

TABLE 1

MEAN SAMPLE CHARACTERISTICS (Weighted)-NLSY

Real Hourly Wage ($1979)	5.75
Hours Worked	41.7
Tenure	2.44
Experience	5.82
Years in School	12.44
Number of Workers with On-the-Job Training	890
Number of Weeks Trained (OJT-Completed)*	13.68
Number of Workers with Off-the-Job Training	721
Number of Weeks Trained (OFT-Completed)*	15.85
Number of Workers with Apprenticeship Training	63
Number of Weeks Trained (APP-Completed)*	31.89
Percentage Nonwhite	12.6
Percentage Married	44.7
Percentage Female	45.4
Age	25.1
Number of Observations	29,020
Number of Individuals	5,649
Number of Jobs	13,590

Note-*Conditional on having positive value for these variables.

TABLE 2

NUMBER OF JOB-WORKER OBSERVATIONS
WITH TRAINING BY INDUSTRIAL AGGREGATES

INDUSTRIES	ON-THE-JOB	OFF-THE-JOB	APPRENTICESHIP
Fisheries (282)	7	21	0
Mining (398)	24	16	0
Construction (2,129)	43	77	115
Manufacturing (7,838)	590	494	94
Transportation, Communication and Public Utilities (1,890)	354	163	19
Wholesale and Retail Trade (7,233)	474	482	42
Finance, Insurance and Real Estate (2,282)	382	226	8
Business and Repair Services (1,969)	160	143	24
Personal Services (956)	56	67	5
Entertainment and Recreation Indus. (317)	20	19	1
Professional and Related Services (3,726)	425	492	11

Note-Number of observations indicated in parentheses.

TABLE 3

SUMMARY STATISTICS-TRAINING (COMPLETED) VS NO TRAINING

	OJT	No OJT	OFT	No OFT	APP	No APP
Real Wage ($1979)	7.52	5.47	6.92	5.61	8.63	5.71
Hours Worked	42.23	41.57	41.68	41.65	42.65	41.64
Tenure	3.62	2.25	3.58	2.29	4.58	2.41
Experience	7.07	5.63	6.71	5.71	7.18	5.80
Training Weeks	13.68	0	15.85	0	31.89	0

Note-All statistics are weighted.

TABLE 4

TRAINING PROGRAMS UNDERTAKEN BETWEEN 1988 AND 1991

	ON-THE-JOB	OFF-THE-JOB	APPRENTICESHIP
Programs Undertaken	1159	734	54
Number of Programs Payed by Employers	1119	532	35
Percentage Paid By Employers	96.5	72.5	65

TABLE 5

EARNINGS FUNCTIONS ESTIMATES: THE IMPACT OF TRAINING

(Dependent Variable: Log of Real Hourly Wages ($1979))

Independent Variable	(1) (OLS)	(2) (GLS)	(3) (IV-GSL)
Tenure	0.0508	0.043	0.0210
	(0.0032)	(0.0030)	(0.0042)
Tenure Squared	-0.003	-0.003	-0.0019
	(0.0003)	(0.0003)	(0.0003)
Total Experience	0.0532	0.0768	0.0945
	(0.0044)	(0.0044)	(0.0049)
Total Experience Squared	-0.0011	-0.0021	-0.0025
	(0.0003)	(0.0003)	(0.0003)
OJTCC	0.1692	0.1666	0.1216
	(0.0255)	(0.0274)	(0.0372)
OJTCP	0.1984	0.2049	0.1319
	(0.0278)	(0.0355)	(0.0439)
OFTCC	0.0750	0.0931	0.1420
	(0.0212)	(0.0216)	(0.0300)
OFTCP	0.0089	0.0820	0.1260
	(0.0173)	(0.0247)	(0.0335)
APPCC	0.0637	0.0414	0.0187
	(0.0242)	(0.0253)	(0.0316)
APPCP	0.2146	0.1464	0.0733
	(0.0224)	(0.0308)	(0.0380)
Indust. Dummies	YES	YES	YES
Occup. Dummies	YES	YES	YES
R-Squared	0.3431	0.8014	0.7949
F-TEST:			
OJTCC Coef.=OJTCP Coef.	0.59	0.84	0.04
p-value	0.44	0.36	0.84

Notes-Standard errors are in parentheses. Sample size is 29,020. Other covariates not shown in table include race, gender, marital status, union membership, SMSA, urban/rural, health, an intercept, unemployment rate in corresponding region, number of jobs ever held by respondent, four region and three education dummies. Aggregate effects are controlled for by having a dummy variable for each year. OJTCC: on-the-job training completed with current employer. OJTCP: on-the-job training completed with previous employers. The same definitions apply to off-the-job (OFT) and apprenticeship (APP) programs.

TABLE 6

HAZARD MODEL RESULTS
(Cox Partial Likelihood Approach)

Independent Variable	(1) No Control for Heterogeneity	(2) With Control for Heterogeneity
Experience (0-1)	0.3082	0.3576
	(0.0638)	(0.1030)
Experience (1-2)	-0.1592	-0.1958
	(0.1809)	(0.2898)
Experience (2-3)	-0.1858	-0.1048
	(0.1799)	(0.2885)
Experience (>3)	-0.7406	-0.2882
	(0.1819)	(0.2955)
OJTCC	-1.3516	-2.9917
	(0.2350)	(0.5236)
OJTCP	0.2599	-0.1513
	(0.1257)	(0.3741)
OFTCC	-1.5107	-2.4151
	(0.2056)	(0.4172)
OFTCP	0.4439	0.3932
	(0.1068)	(0.2971)
APPCC	-0.2926	-0.0781
	(0.2274)	(0.3863)
APPCP	0.0258	0.2231
	(0.1660)	(0.3205)
Number of Previous jobs	0.1186	0.2058
	(0.0033)	(0.0128)
Indust. Dummies	YES	YES
Occup. Dummies	YES	YES
Log Likelihood	-54,738.01	-4,717.03

Notes-Sample size: 8,097 spells of which 1,019 are censored. Standard errors in parentheses. Unshown covariates are the same as those in table 5.

TABLE 7

OLS REGRESSION ESTIMATES OF THE EFFECT OF
ONGOING TRAINING ON STARTING WAGES
(Dependent Variable: log of Real Hourly Wages ($1979)

Independent Variable	(1) Coefficients (Standard Errors)	(2) Coefficients (Standard Errors)
Tenure	0.0438	0.0440
	(0.0047)	(0.0047)
Total Experience	0.0255	0.0257
	(0.0151)	-0.0015
(OC)	-0.1699	-0.1811
	(0.0594)	(0.0595)
Maximum OJT	-	0.1413
During Relationship		(0.0467)
OJTCC	0.3013	0.1454
	(0.0774)	(0.0928)
OJTCP	0.2261	0.2274
	(0.0351)	(0.0351)
OFTOC	-0.0641	-0.0622
	(0.0495)	(0.0495)
OFTCC	0.0201	0.0232
	(0.0725)	(0.0724)
OFTCP	0.0845	0.0833
	(0.0268)	(0.0268)
APPOC	0.0797	0.0769
	(.0609)	(0.0608)
APPCC	0.0293	0.0308
	(0.1489)	(0.1489)
APPCP	0.0051	0.0051
	(-0.0008)	(-0.0008)
Indust. Dummies	YES	YES
Occup. Dummies	YES	YES
R-Squared	0.264	0.2645

Notes-Sample size is 13,394. Unshown covariates are the same as those in table 5. OJTOC: Number of weeks of ongoing OJT with the current employer. The same definition applies to OFT and APP.

Figure 1. Timing of First Training Program (OJT)

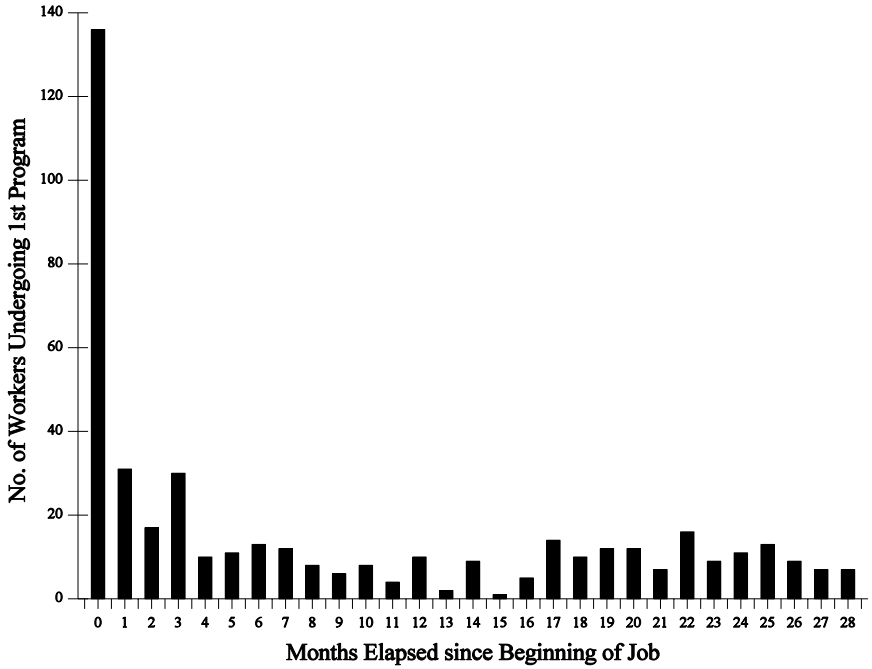

Figure 2. Timing of First Training Program (OFT)

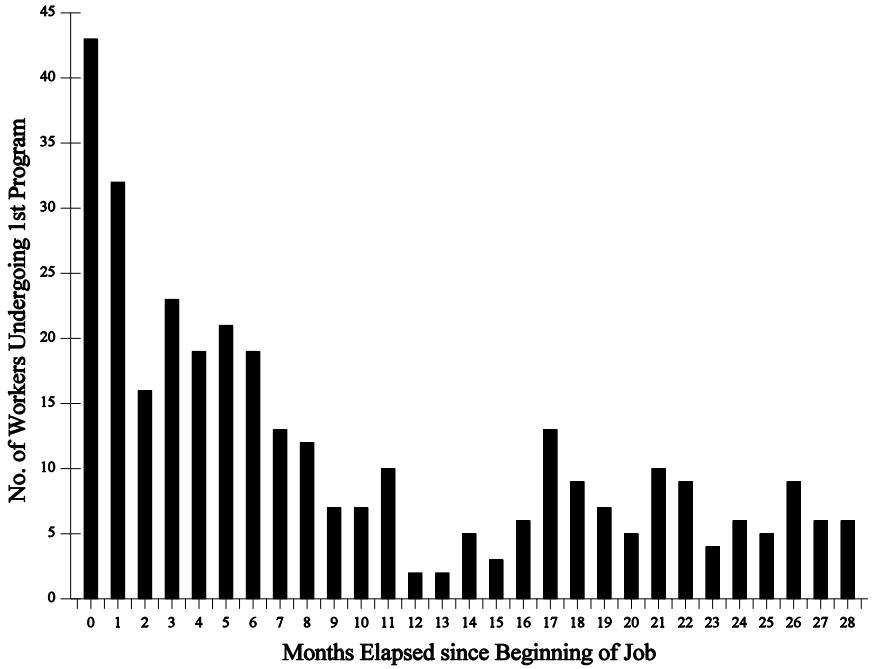

Figure 3. Timing of First Training Program (APP)

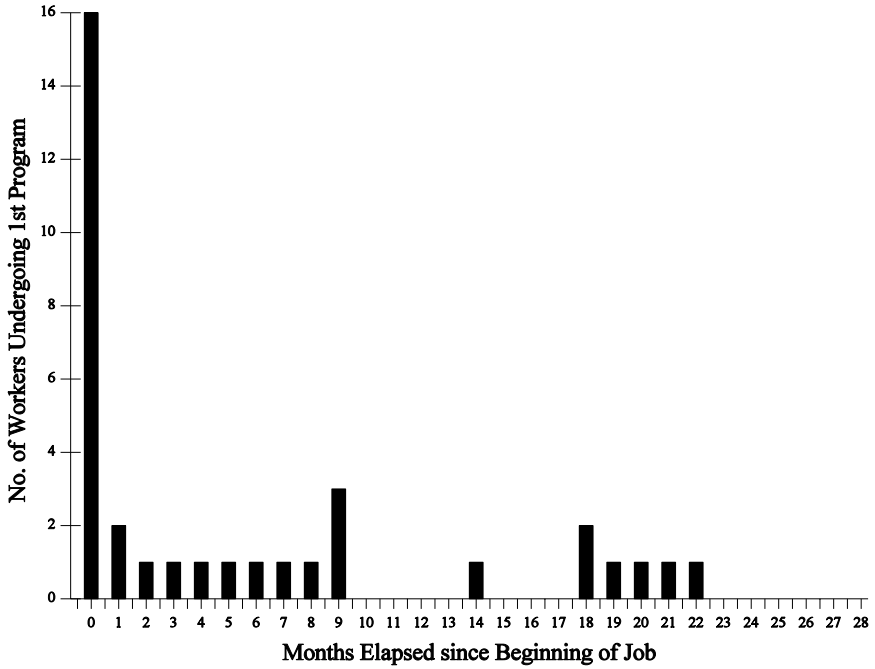

Figure 4. Tenure-Training Relationship (OJT)