Executive Compensation in a Matching Model

Douglas DeJong
University of Iowa

Zhejia Ling
University of Iowa

Elena Pastorino
Federal Reserve Bank of Minneapolis
University of Minnesota

B. Ravikumar
Federal Reserve Bank of St. Louis
Washington University

May 15, 2012

PRELIMINARY – DO NOT QUOTE

Abstract

In this paper we empirically revisit the question of the relative importance of returns to firm-specific tenure and to general labor market experience in the market for executives. We do so by exploiting a new rich matched employer-employee dataset on executive careers. We shed light on the importance of explicitly accounting for an executive's firm-to-firm and job-to-job mobility, within and across firms, over the course of the executive's career in order to accurately measure the magnitude of each type of returns. Treating the allocation of firm value among executives and other stakeholders as a standard joint consumption problem, we prove that a measure of the implied value sharing rule, as embedded in the observed total compensation of an executive, can be recovered.

1 Introduction and Motivation

Starting from the seminal work of Altonji and Shakotko (1987) and Topel (1991), a long-debated issue in the labor economics literature pertains to the relative importance of experience in the labor market and seniority in a firm in explaining wage growth over the life cycle and the resulting earnings differences across individuals. Traditionally, earnings differences across individuals have been attributed to differences in their education and, more recently to differences in their family background and skill endowment (see Cunha and Heckman, 2007). However, in addition to the role of schooling and other individual and family characteristics, growing empirical evidence has pointed out the importance of labor market participation for wage growth, in particular its timing over the life cycle, and the centrality of job mobility to explaining returns to labor market experience. For instance, Rubinstein and Weiss (2007) have documented that wage growth, which happens mainly early in the life cycle, is associated with increasing labor force participation and high job mobility. They estimate that wage growth during the first decade in the labor market is approximately 50% for high school graduates and approximately 80% for individuals with a college degree or more.

Despite its quantitative significance, no clear consensus exists as to the extent to which life-cycle wage growth can be attributed to general experience in the labor market or to specific experience in a given industry or occupation or within a firm. Further, in light of emerging evidence on differences in such returns among individuals of different skill and in different occupations (see Dustmann and Meghir (2005) and Kambourov and Manovski (2009)), a natural question is the extent to which differences in the returns to general experience and firm seniority arise due to differences in their magnitude within a given occupation.

A common difficulty to addressing this question is data availability, on one hand, and endogeneity of sample information, on the other. Indeed, a challenge to empirically documenting the importance of firm seniority for wage growth in any given occupation is the lack of detailed information regarding firm and job characteristics in commonly used panel datasets, like the Panel Study of Income Dynamics (PSID) (for an illustration of the issues in the context of the PSID, see Altonji and Williams (2005) and Buchinsky, Fougère, Kramarz, and Tchernis (2010)). Moreover, as decisions about labor market participation and employment in a given occupation, industry, or firm are made by individuals in their best interests, the non-randomness of information on employment histories poses well-known challenges to the measurement of returns to experience and seniority.

In this study we plan to pursue three goals. We consider the market for executives and, using an original dataset that combines detailed characteristics on the employment histories of executives and the firms employing them, we document the earnings patterns of a large sample of U.S. executives. We then turn to empirically investigate their sources. Specifically, we first examine the empirical determinants of job mobility and of the dynamics of compensation of executives in order to assess the magnitude of their returns to firm tenure and the importance of turnover for compensation growth. Second, we contrast results obtained by employing standard tools of applied analysis with those derived from applying more advanced techniques that take into account the endogeneity of mobility decisions on the part of executives, the cumulative effect of mobility on compensation, and the selection of executives to top positions within a firm based on unobserved (to the econometrician) and possibly time-varying characteristics. Third, treating the allocation of firm value among executives and other stakeholders as a standard joint consumption problem (see Bourguignon, Browning, and Chiappori (2009) and Browning, Chiappori, and Lewbel (2010)), we prove that a measure of the implied value sharing rule, as embedded in the observed total compensation of an executive, can be recovered. We then explore the extent to which variation in this sharing rule over time for the same firm and across firms is responsible for the observed time profile of executive compensation and for differences in executive pay across firms of different market capitalizations.

Our preliminary results show that common estimates of the importance of individual and firm characteristics for executive pay that do not take explicitly into account mobility or individual and firm heterogeneity are severely biased. We suggest a number of estimators that attenuate this bias and evaluate their performance.

2 Model of Returns to Tenure and Experience

2.1 Topel (1991)

We start by reviewing Topel (1991)'s seminal work on measuring returns to firm tenure and labor market experience, as the paper provides an influential implementation of a standard model of wage determination. We treat his work as the benchmark against which we compare more recent contributions to the literature on returns to tenure and experience as well as our work.

2.1.1 Setup

Consider the following prototype model of wage determination

$$y_{ijt} = X_{ijt}\beta_1 + T_{ijt}\beta_2 + \epsilon_{ijt} \tag{1}$$

where y_{ijt} denotes the (log) wage for individual i on job j at time t, X_{ijt} is total labor market experience, and T_{ijt} is current job tenure (seniority). The parameters $\beta_1$and β_2 represent average returns to an additional year of either experience or tenure, respectively. The most popular interpretation of (1) is that β_1 represents the return on general human capital (training and the like) that accumulates with experience, while β_2 represents the return on accumulated job-specific capital that would be lost if a job were to end. Biases in estimating these returns are generated by the covariance between the regressors and the unobservables, ϵ. Topel's main concern is with covariance that is the outcome of optimizing behavior, as workers seek to locate and maintain a productive (high-wage) employment relationship. Thus, one can decompose the unobservables as

$$\epsilon_{ijt} = \phi_{ijt} + \mu_i + \upsilon_{ijt} \tag{2}$$

where ϕ_{ijt} represents the stochastic component of wages that may be specific to a worker-firm pair, and μ_i is a person-specific effect that accounts for unobserved differences in earning capacity across individuals (e.g., 'ability'). The terms υ_{ijt} account for marketwide random shocks as well as measurement error that are known to plague survey data. Topel assumes that the components of (2): (a) are mutually orthogonal; (b) μ_i and υ_{ijt} are orthogonal to the regressors in (1) (for now).

Notice that fixed 'job effects' ($\phi_{ijt} = \phi_{ij}$) are a special case of (2) in which the specific value of a job does not evolve over time. This component captures the notion of a 'good match' in the sense of wages that are higher than what a worker could obtain elsewhere. It will generate bias in estimating (1) if ϕ is correlated with experience or job tenure. Correspondingly, let the auxiliary regression of ϕ on the observables be

$$\phi_{ijt} = X_{ijt}b_1 + T_{ijt}b_2 + \mu_{ijt} \tag{3}$$

In light of (3), since

$$y_{ijt} = X_{ijt}\beta_1 + T_{ijt}\beta_2 + \epsilon_{ijt} = X_{ijt}\beta_1 + T_{ijt}\beta_2 + \phi_{ijt} + \mu_i + \upsilon_{ijt}$$

we can express y_{ijt} as

$$y_{ijt} = X_{ijt}\beta_1 + T_{ijt}\beta_2 + X_{ijt}b_1 + T_{ijt}b_2 + u_{ijt} + \mu_i + \upsilon_{ijt}$$

$$= X_{ijt}(\beta_1 + b_1) + T_{ijt}(\beta_2 + b_2) + \mu_{ijt} + \mu_i + \upsilon_{ijt} \tag{4}$$

Least squares applied to (1) will yield biased estimates of $\beta_1$and β_2 since $E\widehat{\beta_1} = \beta_1 + b_1$and $E\widehat{\beta_2} = \beta_2 + b_2$ (however, $\beta_1 + b_1$ and $\beta_2 + b_2$ are consistently estimated if μ_i and ν_{ijt} are orthogonal to X_{ijt} and T_{ijt}). Topel (1991) suggests the following two-step procedure to correct for the bias in $\widehat{\beta_2}$, the main parameter of interest.

First Step. Within-job wage growth can be analyzed from the first differences of (1) for persons who do not change jobs, which eliminates fixed job and individual effects. Specifically,

$$y_{ijt} - y_{ijt-1} = X_{ijt}(\beta_1 + b_1) + T_{ijt}(\beta_2 + b_2) + \mu_{ijt} + \mu_i + \upsilon_{ijt} -$$

$$X_{ijt-1}(\beta_1 + b_1) - T_{ijt-1}(\beta_2 + b_2) - \mu_{ijt-1} - \mu_i - \upsilon_{ijt-1} \tag{5}$$

$$= \beta_1 + \beta_2 + b_1 + b_2 + \mu_{ijt} + \upsilon_{ijt} - \mu_{ijt-1} - \upsilon_{ijt-1} \tag{6}$$

since $\Delta X = X_{ijt} - X_{ijt-1} = 1$ and $\Delta T = T_{ijt} - T_{ijt-1} = 1$. From the model in (1) in first differences

$$y_{ijt} - y_{ijt-1} = \beta_1 + \beta_2 + \epsilon_{ijt} - \epsilon_{ijt-1} \tag{7}$$

it follows that if job effects are fixed, that is, $\phi_{ijt} = \phi_{ijt-1}$, then (6) specializes to

$$y_{ijt} - y_{ijt-1} = \beta_1 + \beta_2 + \mu_{ijt} + \upsilon_{ijt} - \mu_{ijt-1} - \upsilon_{ijt-1} \,.$$

Remark 1. *The first step of Topel's two-stage estimation would not be warranted in general if job effects were time-varying*.

If $\epsilon_{ijt} - \epsilon_{ijt-1}$ has mean zero ($\phi_{ijt} = \phi_{ij}$ but Topel's favored interpretation is that ϕ_{ijt} follows a random walk with mean-zero innovations), then least squares applied to (7) will yield a consistent estimate of average within-job wage growth (note, however, that mobility decisions may also generate selection in (7) because only acceptable values of $\epsilon_{ijt} - \epsilon_{ijt-1}$ are observed).

Remark 2. *The endogeneity of mobility decisions may affect the estimation of $\beta_1 + \beta_2$, that is, estimates may not be consistent*.

Second Step. Given (7), an estimate of β_1 can be obtained from initial wages on new jobs,

$$y_{ijt} = X_{0jt}\beta_1 + \phi_{ij} + \mu_i + \upsilon_{ijt} \qquad (8)$$

where X_0 is initial experience on the job. Observe that the error term in (8) is nonrandom because only acceptable new job offers are observed. For example, ϕ and X_0 are positively correlated if expected match quality rises with time in the market. One approach to this problem is to explicitly model the mobility decisions that underlie this selection bias, in which case standard sample selection corrections (e.g., Heckman 1976) might be applied. With this strategy, identification relies crucially on distributional assumptions (wage offers must be normally distributed), as well as on other (strong) restrictions (Topel 1991).

According to Topel, a more robust alternative is simply to acknowledge the selection bias implicit in (8) and to treat $\widehat{\beta_1 + \beta_2} - \widehat{\beta_1}$ as an estimate of the return to seniority. In particular, since $X \equiv X_0 + T$, letting $B = \beta_1 + \beta_2$ implies that (1) can be rewritten as

$$y_{ijt} = X_{ijt}\beta_1 + T_{ijt}\beta_2 + \epsilon_{ijt} = (X_{0jt} + T_{ijt})\beta_1 + T_{ijt}\beta_2 + \epsilon_{ijt}$$

$$= X_{0jt}\beta_1 + T_{ijt}(\beta_1 + \beta_2) + \epsilon_{ijt} = X_{0jt}\beta_1 + T_{ijt}B + \epsilon_{ijt}$$

or

$$y = X_0\beta_1 + TB + \epsilon \tag{9}$$

By using the first-step estimate $\widehat{\beta_1 + \beta_2}$ of $\beta_1 + \beta_2$ from (4), we then obtain

$$y - T\hat{B} = X_0\beta_1 + TB + \epsilon - T\hat{B} = X_0\beta_1 + e \tag{10}$$

where $e = T\left(B - \hat{B}\right) + \epsilon$. Topel notices that (10) is preferable to (8) because it makes use of data from all periods of all jobs.

2.1.2 Implementation

The first step is implemented as follows. Note first that if the evolution of wages within jobs follows a random walk, then the residuals of the wage growth model are serially independent and least squares applied to (7) is an efficient estimator. As in Topel and Ward (1992), Topel's examination of the time-series properties of within-job wage changes yields two important conclusions:

(1) Topel finds no evidence of positive serial correlation in within-job wage innovations, $\epsilon_{ijt} - \epsilon_{ijt-1}$. This is a strong finding since one might expect that some types of jobs offer steeper

wage profiles than others. This lack of serial correlation implies that heterogeneity in permanent rates of wage growth among jobs is empirically unimportant;

(2) Topel finds that the within-job evolution of the wage has a strong permanent component that closely approximates a random walk, so the residuals satisfy

$$\phi_{ijt} = \phi_{ijt-1} - \eta_{ijt} \tag{11}$$

where η_{ijt} is serially independent with mean zero (see the details in Topel (1991). Then, values of η_{ijt} reflect 'permanent' changes in a worker's expected lifetime wealth. For example, these may reflect uncertain returns on investments in human capital or simply new information about a worker's productivity (*note that this latter interpretation is supportive of a learning model to explain the increase of wages with seniority*). If these changes are firm-specific rents, they will affect future job-changing decisions. In contrast, if they mainly represent changes in general human capital, then future job mobility will be unaffected by them (*note that this interpretation of Topel may not be warranted in a model of Betrand competition with firms of heterogeneous productivity*).These possibilities have different implications for interpreting the estimated returns to seniority .

Remark 3. *Under (11), the first step in Topel's two-stage estimation is warranted if job effects are time-varying but follow a random walk.*

In implementing the second-step, consistent estimates of $\beta_1 + \beta_2$ and the parameters of higher-order terms in experience and tenure (plausibly included for reasons of fit) are obtained from the within-job growth model, that is, the first-step model (1). Denote these terms by $\chi\hat{\Gamma}$. Recall (9), $y = X_0\beta_1 + TB + \epsilon$, now re-interpreted as $y = X_0\beta_1 + \chi\Gamma + \epsilon$. Let F denote the vector of other factors (education, etc.) that affect wages, so $y = X_0\beta_1 + \chi\Gamma + \mathrm{F}\gamma + \epsilon$. Subtracting $\chi\hat{\Gamma}$ from both sides of the wage equation yields the second-step model,

$$y - \chi\hat{\Gamma} = X_0\beta_1 + \chi\Gamma + \mathrm{F}\gamma + \epsilon - \chi\hat{\Gamma} = X_0\beta_1 + \mathrm{F}\gamma + e \tag{12}$$

where now $e = \epsilon + \chi(\Gamma - \hat{\Gamma})$. Topel's estimated value of β_1 from implementing (12) is about 7 percent (7.13%). This estimate is substantially smaller than the value of $\beta_1 + \beta_2$ estimated from within-job growth, which is 12.58%. The remainder is the main effect of job tenure on wages,

$$\widehat{\beta_2} = \widehat{\beta_1 + \beta_2} - \widehat{\beta_1} = 12.58\% - 7.13\% = 5.45\%$$

That is, Topel estimates that in the first year of the typical new job, the real wage rises by over 5 percent ($\widehat{\beta_2} = 0.0545$) because of the accumulation of job-specific experience alone. Cumulative returns to various lengths of job tenure are based on the main effect of $\widehat{\beta_2} = 0.0545$, together with the concavity of the wage profile implied by the effects of higher-order terms (χ). The returns to seniority are large: *Topel estimates that 10 years of job seniority increase the wage of the typical worker by 28 percent* ($e^{2459} - 1$) *relative to alternatives.* Compared to the estimates of the wage profile generated by ordinary least squares (OLS) applied to (1), these effects are larger, though not dramatically so. Since Topel argues that the two-step procedure generates a lower bound on the true returns, his conclusion is that the OLS estimates may actually be close to the truth.

A final point relates to the estimation of the bias in β_1 and β_2. Though the two-step procedure cannot identify the bias terms b_1 and b_2 separately, their sum is clearly identified since $\beta_1 + \beta_2$ is consistently estimated. In fact, $b_1 + b_2$ is the component of wage growth that is caused by systematic job changing (compare $y = X_0\beta_1 + TB + \epsilon$ when $T = 0$ and when $T > 0$).

Since $E\phi = X_0 b_1 + T(b_1 + b_2)$, the notion that 'good jobs survive' is equivalent to $b_1 + b_2 > 0$. Observe that the sum $b_1 + b_2$ can be estimated directly by reinserting the term $T(b_1 + b_2)$ on the right side of equation (12) and applying least squares. Conceptually, from

$$y = X\beta_1 + T\beta_2 + \epsilon = X_0\beta_1 + TB + \epsilon$$

recalling that $B = \beta_1 + \beta_2$ and, by construction, $X = X_0 + T$ we obtain

$$\begin{aligned} y &= X_0\widehat{\beta_1} + T(\widehat{\beta_1 + \beta_2}) + \epsilon = X_0\widehat{\beta_1} + T(\widehat{\beta_1 + \beta_2}) + \phi + \mu + \upsilon \\ &= X_0\widehat{\beta_1} + T(\widehat{\beta_1 + \beta_2}) + Xb_1 + Tb_2 + u + \mu + \upsilon \\ &= X_0\widehat{\beta_1} + T(\widehat{\beta_1 + \beta_2}) + X_0 b_1 + T(b_1 + b_2) + u + \mu + \upsilon \\ &= X_0(\widehat{\beta_1} + b_1) + T(\widehat{\beta_1 + \beta_2} + b_1 + b_2) + u + \mu + \upsilon \end{aligned}$$

As noted, if the evolution of wages within jobs follows a random walk, then the residuals of the wage growth model are serially independent and least squares applied to (7) delivers an efficient estimator of $\beta_1 + \beta_2$. Therefore,

$$\widehat{b_1 + b_2} = \widehat{\beta_1 + \beta_2 + b_1} + b_2 - \widehat{\beta_1 + \beta_2}$$

(and $\widehat{b_1} = \widehat{\beta_1 + b_1} - \widehat{\beta_1}$). In practice, a similar argument applies and, *under the assumption of a linear relationship between tenure and job specific unobservables,*

$$\begin{aligned} y &= X_0\widehat{\beta_1} + F\hat{\gamma} + \chi\hat{\Gamma} + \epsilon = X_0\widehat{\beta_1} + F\hat{\gamma} + \chi\hat{\Gamma} + Xb_1 + Tb_2 + u + \mu + \nu \qquad (13) \\ &= X_0\widehat{\beta_1} + F\hat{\gamma} + \chi\hat{\Gamma} + X_0 b_1 + T(b_1 + b_2) + u + \mu + \upsilon \\ &= X_0\left(\widehat{\beta_1} + b_1\right) + T(b_1 + b_2) + F\hat{\gamma} + \chi\hat{\Gamma} + u + \mu + \upsilon \end{aligned}$$

The resulting estimate is a wage growth bias of about 0.2% per year. Finally, it can be shown (see Topel (1991)) that the bias in the two-step estimators of β_1 and β_2 is $b_1 + \gamma_{X_0 T}(b_1 + b_2)$, where $\gamma_{X_0T} = (X_0'X_0)^{-1}X_0'T$ is the least-squares coefficient from a regression of tenure on initial experience, X_0. Topel reports that a regression of current tenure on initial experience yields $\gamma_{X_0T} = -0.25$, so $\gamma_{X_0T}(b_1 + b_2) = -0.25 \times 0.002 = -0.005$, which is one-twentieth of one percentage point per year. Note that this implies that the bias in the two-step estimator of β_2, the return to job tenure, is virtually independent of any covariance of job tenure with the unobservables, that is, of the unsigned value of $b_2$since $b_1 \geq 0$, the downward bias in the estimated return to seniority is solely due to improvement in match quality with total labor market experience.

2.1.3 Discussion

Topel estimates job-specific wage premiums that would be earned by a typical worker as he accumulates seniority. According to Topel’s opinion, the most popular interpretation of these returns is that workers anticipate rising compensation over the life of a job, as in contract models such as Becker (1964), Salop and Salop (1976), or Lazear (1981). A second interpretation is also possible, however, since jobs that yield high wage growth may be more likely to survive. In this case returns to seniority are realized period by period, though they may not be anticipated at the start of a job. This would generate *selection bias in wage growth.*

Yet an alternative rationale for the positive relationship between job tenure and wages is that workers' unobserved productivities are negatively related to mobility, which would generate *ability bias in the returns to job tenure*. For example, more able (high-wage) persons may change jobs less often, so tenure and wages will be positively correlated in survey data even if $\beta_2 = 0$. Evidence suggestive of this possibility is that education, an observed element of human capital, is negatively related to job changing. Alternatively, if turnover is costly to employers, then the net productivity of stable workers will be greater, and employers will pay more to obtain them. In either case, unobserved characteristics that raise wages (μ_i) are positively correlated with observed tenure, which raises the estimated returns to job seniority. Topel corrects for potential correlation between 'ability', μ_i, and initial experience, X_0, through an instrumental variable (IV) scheme that relies on the existence of a variable the is uncorrelated with the fixed effect but correlated with X_0 in order to 'net out' the correlation between μ_i and X_0. Topel argues that a plausible candidate instrument is total experience. In particular, Topel assumes that the distribution of μ_i is unrelated to experience (successive cohorts of workers are equally able and equally mobile) so that $E(X'\mu) = 0$. Under this condition, X may be used as an instrumental variable for X_0 in estimating the second-step model.

Remark 4. *Topel's IV correction for the potential correlation between the individual fixed effect and initial experience requires earning capacity to be uncorrelated with experience.*

If recorded experience varies with unobserved individual earning capacity, Topel's scheme is no longer valid. A natural question is to what extent this restriction is plausible. Two facts seem likely to undermine Topel's IV scheme: (1) correlation between unobserved ability and total labor market experience, and (2) the presence of cohort effects.

2.1.4 Comparison with Related Papers

Here we briefly compare the approach of Topel (1991) with the one by Altonji and Shakotko (1987), henceforth AS, who provide a much lower estimate for the returns to firm tenure, and other papers that followed. Specifically, Topel reports an estimate (see Tables 1 and 2 of Topel (1991)) of the cumulative returns to experience at 10 years of experience of 0.354.

The estimates of AS at this level of experience range between 0.372 and 0.442. Our discussion is largely based on Buchinsky, Fougère, Kramarz, and Tchernis (2010).

We start by remarking that, under the assumption that experience at the entry level is exogenous and, hence, uncorrelated with the error terms, Topel (1991) obtains an unbiased estimate for $\beta_1 + \beta_2$ and an upward biased estimate for β_1 (due to the selection bias induced by not modelling mobility decisions on the part of workers). Hence, Topel argues that his estimate of β_2, $\widehat{\beta_2} = 0.0545$, provides a lower bound for the returns to seniority. Topel (1991) also examined two additional sources of potential biases in the estimates of $\beta_1 + \beta_2$ but finds that accounting for these potential biases had a very small effect on the estimate for β_2. Of course, if experience is not exogenous and is positively (negatively) correlated with ϕ_{ijt} because most mobile workers voluntarily (involuntarily) change jobs for better (worse) matches, then the estimate of β_1, say $\widehat{\beta_1}$, will be upward (downward) biased.

In contrast, AS use an instrumental variables approach in which it is assumed (in Topel's notation) that $\phi_{ijt} = \phi_{ij}$, that is, the individual job-specific term is time-invariant. Under this assumption, deviation of seniority from its average in a specific job is a valid instrument for seniority. Since this method is a variant of Topel's two-step approach, it is not surprising that AS obtain an estimate for $\beta_1 + \beta_2$ that is similar to that obtained by Topel. Yet, AS's procedure appears to induce an upward bias in the IV estimate for β_1, and hence a downward bias in the estimate for β_2. The problem is potentially magnified by two other factors: (a) measurement error problem in the tenure data used by AS; and (b) differences in the treatment of time trends in the regression. Namely, Topel uses a specific index for the aggregate changes in real wages by using data from the current population survey, while AS used a simple time trend. As a result, the growth in the quality of jobs, due to better matches over time, would cause an additional downward bias in the estimate of β_2.[1]

Altonji and Williams (2005) specify a model that is closer in spirit to Topel's model but their approach differs in some meaningful way. AW crucially rely on the assumption that the match effect ϕ and time are independent, that is, $Cov(t, \phi) = 0$, conditional on experience (or experience and tenure). This assumption may be questionable, especially in cases where workers have had more time to find jobs with higher match value ϕ. Additionally, t may also be correlated with (the person-specific effect in Topel (1991)) because of changes in the sample composition. The estimates of AW of the cumulative returns to experience at 10 years of experience range between 0.310 and 0.374.

Overall, one important conclusion from both Topel (1991) and Altonji and Williams (2005) is that individual heterogeneity is an important factor of the wage growth process. It appears that some of the reduction in the upward bias in the estimate for β_1 in Topel (1991) is due to a reduction in the bias that stems from individual heterogeneity. Topel reports an estimate (see Tables 1 and 2 of Topel, 1991) of the cumulative returns to experience at 10 years of experience of 0.354. The estimates of AS at this level of experience range between 0.372 and 0.442, while those of AW range between 0.310 and 0.374.

In another recent paper, Dustmann and Meghir (2005) (DM, hereafter) allow for three different sources of returns due to the accumulation of human capital, namely experience, sector-

[1] Abraham and Farber (1987) use a somewhat different set of assumptions. In particular, they use completed tenure to proxy for the unobserved dimensions of the individual's, or job's, quality. A problem with their approach is that many of the workers in their data extract have censored spells of employment. Also, they use a quadratic polynomial in experience when estimating the log wage equation, whereas AS and Topel use a quartic specification.

specific seniority, and firm-specific seniority. In order to estimate the returns to experience, they use data on displaced workers in their new jobs, assuming that such workers could not predict closure of an establishment more than a year in advance. Furthermore, under the assumption that displaced workers have preferences for work similar to those that induced their sectorial choices, controlling for the endogeneity of experience also controls for the endogeneity of sector tenure. In a subsequent step, DM estimate two reduced-form equations, one for experience and another one for participation. The residuals from these two regressions are used as regressors in the wage regression of displaced workers. This allows DM to account for possible sample selection biases induced by restricting attention to only the individuals staying with their current employer. Using data from Germany and the United States, DM find that the returns to tenure for both skilled and unskilled workers are large. The estimated returns to sector-specific tenure are much smaller but (statistically) significant.[2]

Finally, Farber (1999) notes the importance of modeling some specific features of the mobility process. First, he shows that in the first few months of a job there is an increase in the probability of job separation, which decreases steadily thereafter. Farber provides strong evidence that contradicts the simple model of pure unobserved heterogeneity, suggesting that one must distinguish heterogeneity from duration dependence. He also finds strong evidence that: (a) firms tend to lay off less senior workers who have lower specific firm capital; and (b) job losses result in substantial permanent earnings losses. On this latter point, see also Gibbons and Katz (1991).

2.1 Buchinsky, Fougère, Kramarz, and Tchernis (2010)

[2] See also in Farber (1999) a discussion of the empirical findings in the literature on displaced workers. Further references are Addison and Portugal (1989) and Jacobson, LaLonde, and Sullivan (1993).

We now focus on a paper that in spirit and specification is closest to our approach. In the literature, much of the focus on the returns to seniority has concentrated on the possible endogeneity of job changes and its effect on the estimated returns to tenure. Buchinsky, Fougère, Kramarz, and Tchernis (2010) (BFKT, henceforth) contribute to the debate by also considering the possible endogeneity of labor market experience, and its potential effects on the estimated returns to tenure and experience. To address this issue, they develop a model in which individuals make two key decisions, namely employment (or participation) and inter-firm mobility. In turn, these decisions influence the observed outcome of interest, namely wages.

Within this model, they revisit the issue regarding the magnitude of the returns to seniority in the United States. They use data from the PSID (a slightly different sample extract that the one used by Altonji and Shakotko (1987) and Topel (1991)) and estimate their model for three separate education groups: high school drop-outs, high school graduates, and college graduates. They adopt a Bayesian approach and employ Markov Chain Monte Carlo methods for estimating the joint posterior distribution for the model's parameters. (Note that one difference with respect to DM is that BFKT do not model sectoral choices, hence they abstract from sector-specific returns. See Neal (1995) and Parent (1999, 2000) for the importance of sector- and firm-specific human capital.)

The authors find that the returns to seniority are higher than those previously estimated in the literature, including those reported by Topel (1991). Specifically, their results indicate that, while the estimated returns to experience are somewhat higher than those previously found in the literature, they are of similar magnitude. In contrast, the estimates of the returns to seniority are much higher than those previously obtained, including those obtained by Topel (1991). Consequently, their estimates of total within-job wage growth are significantly higher than

Topel's estimates, and those reported by Abraham and Farber (1987). This result holds true for all three education groups analyzed.

Their study also sheds light on several important factors which lead to the differences between their estimates and those obtained in previous studies. First, their study highlights that it is important to explicitly model the employment and mobility decisions, which, in turn, define experience and seniority. Second, they establish the need to account for unobserved heterogeneity in the participation and mobility decisions, and in the wage function. Third, they demonstrate the need to explicitly control for job-specific components in the wage function, through the introduction of a function that serves as a summary statistic for what they term an individual's specific career path.

This function captures the overall effect of the worker's specific career path on the worker's market wage. In particular, they find that the magnitude of the estimated returns changes markedly when they account for this factor, but qualitative results remain similar. This strongly indicates that the timing of a job change during the course of an individual's career is important for an individual's wage trajectory.

2.2.1 Setup

BFKT's model builds on a specification of the wage function common in the literature. They specify the observed log wage equation for individual i in job j at time t as

$$w_{ijt} = w^*_{ijt} I(y_{it} = 1) = (x'_{wijt}\delta_0 + \varepsilon_{ijt}) I(y_{it} = 1) \qquad (14)$$

where, by definition, $w^*_{ijt} = x'_{wijt}\delta_0 + \varepsilon_{ijt}$. In (14), x'_{wijt} is a vector of observed characteristics, including education, labor market experience and firm tenure, of an individual in the current job,

$I(\cdot)$ is an indicator function that equals one if, and only if, $y_{it} = 1$, that is, if, and only if, the i-th individual participates in the labor market at time $t,$ and the Dirac delta function δ_0, which equals one in this situation. So, the wage offer, w^*_{ijt} , is observed only if the individual chooses to work. BFKT decompose the error term ε_{ijt} into three components,

$$\varepsilon_{ijt} = J^W_{ijt} + \alpha_{ijt} + \xi_{ijt}$$

where α_{ijt} is a person-specific correlated random effect, analogous to μ_i in Topel (1991), and ξ_{ijt} is a contemporaneous idiosyncratic error term. The term J^W_{ijt} is analogous to the term ϕ_{ijt} in Topel (1991) with the important difference that in BFKT it explicitly provides a summary statistic for the individual's work history and career. Namely, J^W_{ijt} captures the timing and magnitude of all discontinuous jumps in the individual's wages that resulted from all job changes experienced by the individual until date t.

In principle, this function can be viewed as a full set of dummy variables capturing all observed jumps in the data. However, BFKT argue that in their empirical application this would require estimation of a prohibitively large number of parameters. Thus, they approximate J^W_{ijt} by a piece-wise linear function of experience and seniority at the time of a job change, which is given by

$$J^W_{ijt} = (\phi^s_0 + \phi^e_0 e_{i0})d_{i1} + \sum_{l=1}^{M_{it}} \left[\sum_{k=1}^{4} \left(\phi_{k0} + \phi^s_k s_{it_l-1} + \phi^e_k e_{it_l-1} \right) d_{kit_l} \right]$$

where $d_{1it_l} = 1$ if the l-th job of the i-th individual lasted less than a year and equals 0 otherwise, $d_{2it_l} = 1$ if the l-th job of the i-th individual lasted between 2 and 5 years and equals 0 otherwise, $d_{4it_l} = 1$ if the l-th job lasted between 6 and 10 years and equals 0 otherwise,

$d_{4it_l} = 1$ if the l-th job lasted more than 10 years and equals 0 otherwise. Finally, M_{it} denotes the number of job changes experienced by the i-th individual at time t (not including the individual's first sample year). If an individual changed jobs in the first sample year, then $d_{i1} = 1$, otherwise $d_{i1} = 0$. The quantities s_{it_l-1} and e_{it_l-1} denote the individual's seniority and experience in year t_l, respectively, when individual i leaves job l. Note that while the ϕ's are fixed parameters, the size of the jumps (within each of the four brackets of seniority) may differ depending on the level of seniority and labor market experience at the time of a job change. Overall the function J_{ijt}^{W} contains thirteen identifiable parameters, corresponding to the four brackets of seniority and the first sample year.

Observe that the J_{ijt}^{W} function generalizes the term ϕ in Topel (1991) and captures the initial conditions specific to the individual at the start of a new job. Equivalently, this function provides a measure of the opportunity wage of the worker if the worker were to move to a new job at that point in the worker's career. Note also that inclusion of actual rather than potential labor market experience as a determinant of initial earnings at a new job allows BFKT to distinguish between displaced workers, who experienced a period of non-employment after displacement, and workers who moved directly from one job to another. This difference is all the more critical since BFKT do not distinguish between participation and employment. Instead, the inclusion of the seniority level at past jobs allows BFKT to control for the quality of past job matches. Whether the frequency of changing jobs and the individual's labor market attachment matters is an empirical question that BFKT can explicitly address based on J_{ijt}^{W}.

Lastly, note that J_{ijt}^{W} is individual-job specific. In general, there are several ways to define a job. BFKT define a job as a particular employment spell in an individual's career. Hence, it is possible that different individuals will have the same values for J_{ijt}^{W} even though they may not be employed at the same firm. This definition of a job is consistent with their modeling approach. However, our data allows us to rely on a finer, much more precise definition of a job based on administrative records as corresponding to an executive's title within the hierarchy of titles of a firm.

3 Contribution

Our research will explore the following questions: (1) Is there any evidence that the pattern of labor market experience differs among executives with higher education and/or more successful careers? (2) Do we detect any important cohort effect? One reason for this could be due to the recent changes in the legislation surrounding executive pay.

Observe that if the answer to either question is positive, then Topel's IV scheme would be inapplicable. The purpose of our research is to follow this line of argument to first replicate and then improve on Topel's measures of the relative magnitude of returns to tenure and to labor market experience. Observe also that Topel assumes that the level of experience at a new job is exogenous, which may be a problematic assumption. Indeed, one of the reason between the difference in estimates between BFKT and Topel (1991) is due to the fact that BFKT explicitly consider the participation decision as endogenous.

We plan to do so in four steps: (1) by making use of new detailed matched firm-executive data containing information on firm characteristics (absent from Topel's analysis) and job characteristics (poor in Topel's analysis, especially with respect to an executive's title and position within the hierarchy of jobs of a firm); (2) by allowing for a more flexible functional form specification; (3) by explicitly accounting for the endogeneity of the mobility and participation decisions in our empirical specification; and (4) by relying on a semi-parametric estimation approach.

We build on BFKT by accounting for the endogeneity of the participation and mobility decisions, by allowing for unobserved heterogeneity among executives, and by explicitly controlling for the effect of past mobility on current wages. We augment their work in two ways. First, from a modeling point of view, we account for the endogeneity of job mobility decisions both *within firms and between firms*, we allow for a flexible *nonparametric* specification of unobserved heterogeneity, and we control for the effect of past mobility on current wages via *alternative specifications* that are consistent with different structural interpretations.

Second, from an empirical point of view, by using a *matched employer-employee* database with rich information on firm productivity and financial characteristics, we are able to assess the separate contribution of individual and firm characteristics to returns to tenure and labor market experience. The availability of firm data also allows us to incorporate firm characteristics in controlling for the impact of the quality of past job matches on executives' mobility decisions, rather than merely relying on the inclusion of the seniority level at past jobs as BFKT do.[3]

Further, based on information on an executive's mobility between jobs within a firm, we can isolate the effect of mobility across jobs *within a firm* on the returns to firm tenure and contrast its importance to the importance of mobility across jobs *between firms* for total wage growth.

[3] In this sense, we view our work as building on the analysis of Abowd, Kramarz, and Roux (2006), who, based on French matched employer-employee data, document the importance of mobility for wages in that they find that entry wages depend upon seniority in the previous job, as well as the number of previous jobs held by the individual.

Our key source of identification of returns to tenure and experience is the time-series dimension of our data. Specifically, our data contain observations on many individuals who changed jobs and firms over the sample period, and did so at different points in their life-cycle (given our focus on executives, necessarily our data contain information on older workers than in more representative samples like the PSID. Nonetheless, we show below that descriptive statistics from our data are comparable to those from the PSID if one restricts attention to individuals with college degree; see, for instance, Table 1 in BFKT).[4] Naturally, the time series dimension also allows us to control and pin down individual-specific effects.

4 Data and Sample Construction

4.1 Database

The data in this study are collected from different sources. Executive biographic information is from BoardEx of Management Diagnostics Limited. Compensation data are from Standard and Poor's ExecuComp database.

Management Diagnostic Limited is a private research company specializing in collecting and disseminating social network data on company officials (including top executives and board of directors) of US and European public and private companies. The BoardEx database collects annual information beginning in 2000 and is organized as a time series of individual curriculum vitae. At a specific point in time (i.e., the "report date" in BoardEx), an individual curriculum vitae is constructed based on the most recent disclosure information obtained by analysts at the Management Diagnostic Limited. The curriculum vitae contains college, graduate and

[4] As in BFKT, since experience and seniority are fully endogenized, we need not impose any further restrictions on the data extract (e.g. restricting attention to only exogenously displaced workers) as is done in Dustmann and Meghir (2005) or Topel (1991).

professional education and degree information, past employment history (including beginning and ending dates of various roles), current employment status (including primary employment and outside roles), and social activities (club memberships, positions held in various foundations and charitable groups, among others).

ExecuComp provides executive compensation data collected directly from each company's annual proxy (DEF14A SEC form). It collects up to nine executives for a given year per company, though most companies only report five (SEC only requires disclosure of compensation details of the top five earners among all executives). The number of executives in excess of five for which details are provided is therefore the company's own decision. Detailed information on salary, bonus, options and stock awards, non-equity incentive plans, pensions and other compensation items are disclosed in the proxy statement and provided by ExecuComp. The universe of firms covers the S&P 1500 plus companies that were once part of the 1500 (but removed from the index) that are still trading, and some client requests. Data collection on the S&P 1500 began in 1994. However, there are data back to 1992 but it is not the entire S&P 1500 – it is mostly for the S&P 500. We focus on total compensation earned by executives, which is the sum of salary, bonus, other annual, total value of restricted stock granted, total value of stock options granted (using Black-Scholes), long-term incentive payouts, and all other total (data item *TDC1* in ExecuComp).

4.2 Sample Construction

For the purpose of this study, we merge BoardEx data and ExecuComp data together so each executive in the sample has both biographical and compensation information. In the BoardEx database, the unique company identification code is "Company ID" and the unique executive identification code is “Director ID”. In the ExecuComp database, the unique company identification code is GVKEY and the unique executive identification code is EXEID. However, there is no existing link between “Company ID” in BoardEx and GVKEY in ExecuComp, and there is no existing link between “Director ID” in BoardEx and EXEID in ExecuComp. Therefore, for each executive in ExecuComp, we obtain their unique identifier from BoardEx to extract their biographical information.

We start with all executives from ExecuComp database. To find each executive from the most recent BoardEx data available, we require that: (1) the person in BoardEx has the same last name and first name as the one in ExecuComp; (2) the firm-year observed for the executive in ExecuComp matches the firm-year observed for the executive in BoardEx. To match firms in these two databases, we follow the procedure below. First, for active companies, BoardEx provides the International Security Identification Number (ISIN). We derive each firm’s CUSIP from ISIN and match with firms in ExecuComp by using CUSIP.[5] Second, for the inactive companies, BoardEx does not always keep the ISIN. If the ISIN is not provided, we match the company name recorded by BoardEx with the name of the company in ExecuComp using the built-in algorithm in SAS. For all executives from the ExecuComp database, we find a match from BoardEx for 18,209 executives.[6]

For the 18,209 executives in the sample, 1,179 people were born before year 1940, 4,773 people were born during the 1940s, 7,470 people were born during the 1950s, 4,144 people were born during the 1960s, only 320 people were born after 1970, and 323 people have unknown

[5] CUSIP is another common identifier often used in ExecuComp database.
[6] Note that since BoardEx started to collect data for executives and directors who can be observed in public firms after year 2000, executives who can be observed in ExecuComp only before 2000 should not be found in BoardEx. In other words, managers cannot be observed in public firms after 2000 will not be in our sample.

dates of birth. Using the employment history information from BoardEx data, the 18,209 executives were hired by 6.38 firms (including both private and public firms) on average. We observe 115,936 firm-executive pairs and 47,786 firms in the BoardEx data for these individuals. In terms of compensation information from ExecuComp, when these individuals are hired by S&P1500 firms, they earn an annual average (median) total compensation of $2,483,210 ($1,162,380).

5 Preliminary Results

The seminar presentation.

References

[1] Addison, J.T., and P. Portugal (1989): "Job Displacement, Relative Wage Changes, and Duration of Unemployment", *Journal of Labor Economics* 7, 281-302.

[2] Abowd, J.M., F. Kramarz, and S. Roux (2006): "Wages, Mobility, and Firm Performance: Advantages and Insights from Using Matched Worker-Firm Data", *Economic Journal* 116, 245-285.

[3] Abraham, K.G., and H.S., Farber (1987): "Job Duration, Seniority and Earnings", *American Economic Review* 77, 278-297.

[4] Altonji, J., and R. Shakotko (1987): "Do Wages Rise with Job Seniority?", *Review of Economic Studies* 54, 437-459.

[5] Altonji, J., and N. Williams (2005): "Do Wages Rise with Job Seniority? A Reassessment", *Industrial and Labor Relations Review* 58, 370-397.

[6] Becker, G. (1964): "Human Capital: a Theoretical Analysis with Special Reference to Education". New York: Columbia University Press for NBER.

[7] Bourguignon, F., M. Browning, and P.-A. Chiappori (2009): "Efficient Intra-Household Allocations and Distribution Factors: Implications and Identification", *Review of Economic Studies* 76, 503-528.

[8] Browning, M., P.-A. Chiappori, and A. Lewbel (2010): "Estimating Consumption Economies of Scale, Adult Equivalence Scales, and Household Bargaining Power", mimeo.

[9] Buchinsky, M., D. Fougère, F. Kramarz, and R. Tchernis (2010): "Interfirm Mobility, Wages and the Returns to Seniority and Experience in the United States", *Review of Economic Studies* 77, 972-1001.

[10] Cunha, F., and J. Heckman (2007): "The Technology of Skill Formation", *American Economic Review* 97, 31-47.

[11] Dustmann, C., and C. Meghir (2005): "Wages, Experience and Seniority", *Review of Economic Studies* 72, 77-108.

[12] Farber, H.S. (1999): "*Mobility and Stability*", in Ashenfelter, O. and Card, D. (eds.) Handbook of Labor Economics, Volume 3B, Chapter 37 (New York: North-Holland), 2441-2483.

[13] Gayle, G.-L., L. Golan, and R.A. Miller (2011): "Promotion, Turnover and Compensation in the Executive Labor Market", mimeo.

[14] Gibbons, R., and L. Katz (1991): "Layoffs and Lemons", *Journal of Labor Economics* 9, 351-380.

[15] Heckman, J. (1979): “Sample Selection Bias as a Specification Error”, *Econometrica* 47, 153-161.

[16] Jacobson, L., L. LaLonde, and D. Sullivan (1993): “Earnings Losses of Displaced Workers”, *American Economic Review* 83, 685-709.

[17] Kambourov, G., and I. Manovski (2009): “Occupational Specificity of Human Capital”, *International Economic Review* 50, 63-115.

[18] Lazear, E. (1981): “Why is there Mandatory Retirement?”, *Journal of Political Economy* 87, 1261-1284.

[19] Neal, D. (1995): “Industry-Specific Human Capital: Evidence from Displaced Workers”, *Journal of Labor Economics* 13, 653-677.

[20] Parent, D. (2000): “Industry-Specific Capital and the Wage Profile: Evidence from the National Longitudinal Survey of Youth and the Panel Study of Income Dynamics”, *Journal of Labor Economics* 18, 306-323.

[21] Parent, D. (1999): “Wages and Mobility: The Impact of Employer-Provided Training”, *Journal of Labor Economics* 17, 298-317.

[22] Rubinstein, Y., and Y. Weiss (2007): “Post Schooling Wage Growth: Investment, Search and Learning”, in Handbook of the Economics of Education, Volume I, E.A. Hanushek and F. Welch Eds., North-Holland, The Netherlands.

[23] Salop, J. and S. Salop (1976): "Self-Selection and Turnover in the Labor Market," *Quarterly Journal of Economics* 90, 619-627.

[24] Topel, R.H. (1991): “Specific Capital, Mobility and Wages: Wages Rise with Job Seniority”, *Journal of Political Economy* 99, 145-176.

[25] Topel, R.H., and M.P. Ward (1992): “Job Mobility and the Careers of Young Men”, *Quarterly Journal of Economics* 107, 439-479.