

The Promotion Dynamics of American Executives*

Christian Belzil
Centre National de Recherche Scientifique
(Gate, Universite de Lyon 2)
IZA and CIRANO

Michael Bognanno
Department of Economics
Fox School of Business and Management
Temple University, Philadelphia

January 21, 2004

Abstract

We formulate an empirical model of promotion with dynamic self-selection where the current promotion probability depends on the hierarchical level in the firm, individual human capital, unobserved (to the econometrician) individual specific attributes, time varying firm specific variables (firm size and profits) as well as endogenous past promotion histories. We examine the causal effect of previous promotion histories (as measured by realized speed of promotion) on future promotion outcomes. The model is fit on an 8 year panel of promotion histories of 30,000 American executives employed in more than 380 different firms. The stochastic process generating promotions is

*We thank Edward Lazear, Bart Hamilton and Thierry Magnac, for helpful discussions at an early stage of this project. We also thank Jean-Marc Robin, Guy Laroque and Francis Kramarz for useful comments and suggestions. The support of Cirano, IZA and a Temple University Research leave are gratefully acknowledged by Bognanno. The usual disclaimer applies.

weakly correlated with standard human capital endowment variables (age, schooling and tenure). It may be viewed as a series of promotion probabilities which become smaller as an individual moves up in the hierarchy and is primarily explained by individual (or firm) specific factors other than measured human capital. We also find that, conditional on unobservables, the promotion probability is only mildly enhanced, on average, by the speed of promotion achieved in the past (a structural fast track effect). However, we find the existence of a relatively high cross-sectional dispersion in the effect of past promotion histories and we are able to provide an explanation for this relatively high dispersion. In general, the magnitude of the individual specific effect of achieved speed of promotion is inversely related to accumulated human capital (schooling and tenure). We believe that these findings are consistent with the hypothesis that the signaling aspect of past promotions is stronger for those who are less educated and stronger for those who are relatively new in a firm. We also find that a negative correlation between current promotion and past speed of promotion cannot be ruled out for a portion of the population, and we are able to relate this finding to the “Peter Principle”.

JEL Classification: C33, J41, M5, M51

1 Introduction

Personnel economics is one of the fastest growing fields of modern economics (Lazear 1999). It concerns the strategies that corporations use in governing their internal labor markets and the performance incentives created. Because incentives are promotion-driven to an important extent in corporate hierarchies, the subject of promotion is one of key importance in personnel economics. The theoretical connection between promotions and wages appears in tournament (Lazear and Rosen 1981, Rosen 1986), human capital (Carmichael 1983 and Prendergast 1993) and learning (Chiappori and Salanie 1999) models. Other recent models incorporate multiple factors, including learning about workers’ ability over time, acquisition of human capital on-the-job and job assignments to explain empirical results about wage and promotion dynamics (Bernhardt 1995 and Gibbons and Waldman 1999).

In recent theoretical models, the link between early career results and subsequent promotions (often referred to as a “fast track” when early promotions increase the pace of later ones) is a central element. Statistically, the notion

of a fast track is equivalent to the existence of serial correlation in individual promotion histories.¹ The potential existence of fast tracks suggests that the stochastic process that generates promotion outcomes may be inherently dynamic. Fast tracks are suggested in theoretical work by Bernhardt and by Gibbons and Waldman. In Bernhardt, firms exploit able workers that are not identified as able in the outside labor market. Promotion signals ability publicly. The workers who get promoted quickly are publicly identified as able and are awarded subsequent promotions, even before perhaps more able workers whose ability remains hidden to the outside labor market. In this way, fast tracks result, even after controlling for worker ability.

Through a different mechanism, fast tracks result in the modeling of Gibbons and Waldman even when worker ability is common knowledge. Workers with more innate ability are simply promoted more quickly initially and spend less time in subsequent levels because of faster growth in their effective ability. Effective ability is a function of both innate ability and experience. It is ability, as opposed to asymmetric information regarding ability, which is creating the fast track.

While most recent influential contributions to personnel economics remain theoretical, a small but expanding empirical literature on promotions also exists. As of now, the empirical literature is dominated by studies based on personnel records of a single firm. These include papers by Ariga, Ohkusa and Brunello (1999), Baker, Gibbs and Holmstrom (1993, 1994a, 1994b), Chiappori, Salanie and Valentin (1999), Lazear (1992), and Seltzer and Merrett (2000). The promotion-wage growth link is well established in Lazear (1992) and Baker, Gibbs and Holmstrom (1993). Up to now, a major focus of the empirical literature has been the promotion/wage growth link.² The existence of fast tracks has also been investigated empirically. Evidence of promotion fast tracks was found in Ariga, Ohkusa and Brunello's single-firm Japanese study, in Baker, Gibbs and Holmstrom's (1994b) study of a medium sized US service sector firm and in the Seltzer and Merrett's study of the Union Bank of Australia.

¹Because of the link between promotion and wage growth, serial correlation in promotion outcomes will typically imply serial correlation in wage growth.

²Using thirteen years of personnel records from a large durable goods manufacturer, Lazear found that the wages of workers fell after spending seven years in the same position, while job change resulted in higher current wages and later wage growth. Baker, Gibbs and Holmstrom (1994,b) found that increasing position tenure and real pay are inversely related.

Making both theoretical and empirical contributions, Chiappori, Salanie and Valentin (1999) examine promotions in a large French state-owned firm between 1960 and 1982. They develop a testable implication for wage dynamics, dubbed the “late beginner” effect, using a model of learning with downward wage rigidity that does not require observed worker output. They demonstrate that, conditional on an equal first period start-wage and last period end-wage, the worker who reaches the last-period’s end-wage later in his career will have the higher expected ability and should have better future prospects. Chiappori, Salanie and Valentin (1999) find empirical support for this hypothesis. At the same time, without any conditional statement, they find that an “early starter” (i.e., as measured by high early wages) also enjoys better future prospects.

Finally, promotion dynamics have been analyzed in the sociology and management literatures.³ Some stylized findings often cited include: (1) the importance of early career results on later promotion;⁴ (2) the importance of education and advanced degrees on promotion;⁵ (3) the importance of functional area and age in the attainment of a top executive position;⁶ (4) the disadvantage faced by new hires.⁷

Though good descriptive work on firm hierarchies exists, it is probably fair to say that the level of sophistication of the empirical work found in

³For surveys see Forbes and Piercy (1991) and Rosenbaum (1984).

⁴Rosenbaum, J. E. (1979) finds that those promoted first were more likely to receive further promotions and to reach higher levels in the firm. Howard and Bray (1988) find that Bell System managers with more significant job challenges in years one through eight exhibited greater advancement at year twenty.

⁵Howard and Bray (1988) found a college degree to be the best predictor of promotion. Forbes and Piercy (1991, p. 165) find that the time to the CEO position is reduced through higher levels of education. Useem and Karabel (1986) show the importance of earning a degree from an elite institution when the executive is not from elite social origins.

⁶Vroom and MacCrimmon (1968) found that promotion opportunities varied with functional area and were better in finance and marketing. Forbes and Piercy (1991, p.4) find the functional area backgrounds of CEOs to vary by industry. They also find with regards to eventual CEOs that the time to reach various top positions in the organization varied by functional area (p.145) and provided evidence of age varying systematically with career level (p.144). For example, CEOs reached a top management position by age 47 on average and none reached this level after age 58. Out of 230 CEOs, none were promoted to the CEO position later than age 65, the mean age was 50.

⁷Forbes and Piercy (p.5) note that successful top executives spend most of their careers within the same firm. Tuckel and Siegel (1983) find most CEOs to have spent their entire careers within one firm.

the personnel economics literature is not commensurate with the level of sophistication of its theoretical foundations. The empirical literature may be criticized for at least two main reasons. First, as most empirical work is based on a single firm, the empirical results that models are being built around risk being idiosyncratic to the specific firms or occupations studied.⁸ Many of the stylized facts in the literature rely on a hand full of papers that await corroboration. As a result, and without more extensive studies using multiple firms, it is practically impossible to infer whether individual promotion histories are mostly explained by individual specific endowments or by firm heterogeneity.

Secondly, the literature is largely dominated by empirical strategies aimed at finding variables correlated with promotion outcomes. These include relative in-level wages, time since last promotion and measures of past promotion achievements. While finding variables correlated with promotion outcomes may be a useful step in the comprehension of promotion mechanisms in corporate hierarchies, economists should be reluctant to attribute a causal relationship between these indicators and promotion outcomes. In other words, in the presence of dynamic self-selection, the correlation between past promotion histories and subsequent promotion outcomes may be highly spurious.⁹

In the econometric analysis presented here, promotions may be viewed as a sequence of discrete outcomes driven by observed as well as unobserved (to the econometrician) individual and firm characteristics. Statistical analysis is complicated by the fact that unobserved attributes are career persistent and that, in a context where employers have imperfect information about individual workers, past outcomes may be used to “signal” worker ability. This implies that the serial persistence observed in individual promotion histories may be simultaneously explained by persistent unobserved heterogeneity, deemed a non-causal fast track in the sense that rapid promotions themselves are not causing subsequent promotions, as well as state dependence (Heckman 1981). In our analysis, we use the terminology of a “structural fast track” to refer to the statistical notion of state dependence. A rigorous analysis therefore requires that the dynamics of the stochastic process gener-

⁸In one of few multi-firm studies, Bognanno (2001) investigates implications of tournament theory using the same US panel data used in this paper.

⁹Recent papers that model dynamic self-selection in human capital accumulation include Keane and Wolpin (1997), Eckstein and Wolpin (1999), Belzil and Hansen (2002), in a structural framework and Cameron and Heckman (1998 and 2001) in a reduced-form framework.

ating promotions and the endogeneity of the initial conditions be taken into account.

This paper examines the promotion dynamics of American executives across large corporations. We estimate a dynamic reduced-form model of career promotion outcomes using an employer-employee panel of 385 of the largest corporations in the U.S. The paper provides an empirical picture that it is not idiosyncratic to one or a small group of firms. However, using data across firms comes at the cost of introducing multi-dimensional heterogeneity which is often not accounted for in applied work. The econometric model allows for the following features:

- Promotion probabilities are a function of individual observed human capital endowments (age, schooling and tenure)¹⁰;
- Promotion probabilities are potentially affected by firm size and firm profits;
- Promotion probabilities are potentially affected by past promotion histories and, more precisely, by the speed of promotion achieved up to the present period (the notion of a structural fast track);
- Promotion probabilities are potentially affected by the current level, given individual specific endowments;
- Promotion probabilities are affected by time persistent individual unobserved factors such as ability, motivation as well as firm specific unobserved factors (the notion of a non-causal fast track);
- Both the initial speed of promotion and the initial level in the firm (measured at the start of the sampling period) are endogenous. They are affected by individual unobserved heterogeneity correlated with unobserved heterogeneity affecting promotion outcomes;
- The distribution of unobserved heterogeneity is allowed to depend on the number of sampling periods each individual is observed. This allows us to control for panel attrition bias which could arise in the presence of endogenous mobility or other sampling problems.

¹⁰While it is widely recognized that the more educated also experience higher wage growth, the causal effect of schooling on post schooling human capital investments is rarely investigated. However, as we allow education to allow for promotions in the present framework, we implicitly allow schooling to affect wage growth.

After a thorough review of the empirical literature, we can assert that our model is the most general and most comprehensive empirical model of promotion dynamics ever estimated. This econometric structure allows us to answer the following fundamental questions:

1. Are promotion histories significantly affected by standard human capital endowment variables (age, tenure and education)?
2. Do promotion probabilities differ with the executive's level in the hierarchy?
3. Are promotion histories significantly affected by persistent individual unobserved factors?
4. What is the causal effect of having achieved a higher rate of promotion since entering the labor market on subsequent promotion probabilities?
5. Is the causal effect of past promotion histories characterized by a substantial level of heterogeneity?
6. What is the relative importance of individual specific factors (such as human capital) and promotion dynamics in explaining differences in career promotions?
7. Are promotion dynamics correlated with individual specific attributes such as education and tenure?

We believe that answering these questions is a major undertaking and will shed light on one of the most crucial aspects of personnel economics; that is the determinants of promotion dynamics. In particular, our analysis will allow us to distinguish between a structural (or causal) fast track effect, as opposed to fast tracks explained by unobserved persistent factors.

Our results point out the complexity of the stochastic process which governs firm level transitions. We find that among standard individual specific human capital endowment variables (tenure and schooling), only schooling really matters. The stochastic process generating promotions may be viewed as a series of promotion probabilities which become smaller as an individual move up in the hierarchy and which are primarily explained by unobserved (to the econometrician) individual (or firm) specific factors other than measured human capital.

Our results have also allowed us to uncover the dynamic aspects of promotion histories. While the theoretical literature reserves a central place for the notion of a fast track, we find that the stochastic process generating promotions is only mildly positively correlated (structurally), on average, with the speed of promotion achieved in the past. However, we find the existence of a relatively high cross-sectional dispersion in the effect of past promotion histories (speed of promotion) and we are able to provide an explanation for this relatively high dispersion. In general, the magnitude of the individual specific effect of achieving a higher speed of promotion is inversely related to accumulated human capital; that is it is negatively correlated with schooling and tenure. We believe that these findings are consistent with the hypothesis that the signaling aspect of past promotions is stronger for those who are less educated and who are relatively new in a firm. Accumulated schooling and tenure may reduce the risk associated with promotion (from the perspective of the firm), and may therefore lead firms to ignore past histories. We also find that a negative correlation between current promotion and past speed of promotion cannot be ruled out of for a portion of the population, and we are able to relate this finding to the famous “Peter Principle” (Lazear, forthcoming).

The remaining sections of the paper are structured according to the following format. Section 2 describes the panel data that we use. The econometric strategy is discussed in Section 3 and the econometric model is laid out in Section 4. Section 5 is devoted to the presentation of the results. In Section 6, we discuss briefly an alternative model specification in which population heterogeneity is affected by accumulated human capital. Concluding remarks are found in Section 7. Finally, various summary statistics are found in the Appendix (Section 8).

2 The Data

The panel data set analyzed in this paper contains 32,890 executives working at 385 of the largest firms in the United States during the period from 1981 to 1988. These data were obtained by the Center for Advanced Human Resources Studies at Cornell University.¹¹ All of these executives appear in at least two consecutive years while 20,990 are in the data for three years,

¹¹Additional description of the data may be found in Bognanno (2001).

14,427 for four years, 8971 for five years, 4967 for six years, 2964 for seven years and 1619 for eight years.

In order to belong to our sample, the executives had to meet the following conditions: 1) as our focus is on promotion, only executives appearing in at least two consecutive years are kept for analysis; 2) only executives who are first observed no more than six levels beneath the CEO are kept for analysis. Executives seven or more levels from the CEO position are relatively few in number and provide a thin basis on which to extrapolate, as is evident in Table A2. The database reveals information on individual, job and firm characteristics, including age, education, functional area, job title, tenure, base pay, bonus pay, reporting level, industry, firm profits, sales, and employment. Executives in the data range in position from the CEO down to regional sales executives, production superintendents, and plant managers. These are the most common job titles six levels below the CEO.

One advantage of these data over other datasets that have been used to study promotion is that we do not rely on changes in the job code that result in higher consequent pay or movement up an empirically determined job ladder to determine whether a promotion has taken place. Each individual's job level, the number of reporting levels from the CEO position (level 1), is provided by each firm annually, unlike in some other papers where it must be empirically determined. Promotion in this study reflects an upward movement in the firm's hierarchy.

Two tables, found in the Appendix, provide insight into the data. The first five sets of variables in Table A1 pertain to the number of consecutive years an executive of level 7 or higher appears in the data. Attrition results from both firms failing to report data on individuals and from firms leaving the survey. Due to the occurrence of promotions, executives in the data longer tend to be observed at slightly higher levels in the firms. Executives in their first observation have an average age of 46 years and average education of 16.4 years. Only 3.5% of executives were newcomers to their firms in the year they were first observed in the data. Seventeen percent of executives were employed in finance positions, 14.5% were in marketing, 13.8% were profit center heads, and 12.4% held manufacturing positions. Seventy-four percent of the executives were employed in manufacturing firms. Overall, promotions are relatively rare events. The average number of promotions per individual and per year is 0.11 (Table A1).

Table A2 provides more detail on promotion, it shows the fraction of executives promoted between their first and second years in the data by

level, as well as tenure and age by level. It is clear from Table A2 that the rate of promotion diminishes at higher levels in the hierarchy of the firm. For instance, a level 5 executive enjoys a 15% probability of promotion annually, where for a level 3 executive, the rate is only 4.3%. Across all levels, for executives in their first two years in the data, the rate of promotion is 11.7%. To the extent that promotion is a rare event, it is more so the case for senior executives.

3 Econometric Strategy

As a starting point, we consider the estimation of a dynamic promotion probability model, which ideally, would be of the following form

$$\Pr(Y_{ijt} = 1) = F(\varpi_i, \varpi_j, Y_{ijt-1}, Y_{ij,t-2} \dots Y_{ijt-p}, L_{ijt_0})$$

where Y_{ijt} is an indicator recording the event of a promotion for individual i in firm j at time t , ϖ_i and ϖ_j are individual and firm specific relevant attributes, L_{ijt_0} is the starting level (at time t_0) and $(Y_{ijt-1}, Y_{ij,t-2} \dots Y_{ijt-p})$ is a p dimensional vector of relevant past promotion outcomes. In order to implement such a model, we have to address three specific problems. One problem is how to summarize the entire vector of past promotion histories in a reasonable way. The second issue relates to the identification of individual and firm specific unobserved characteristics. Finally, we also face the usual dilemma faced by those estimating dynamic discrete choice models, namely whether to use conditional maximum likelihood techniques (sometimes referred to as fixed effects estimation) or use a random effect specification. These issues deserve some discussion.

First, with respect to the modeling of promotion dynamics, it should be recalled that in the econometric literature devoted to the estimation of dynamic logit models with fixed effects (Chamberlain, 1984 and Magnac, 2000), it is pointed out that non-parametric identification of two lags requires at least seven periods. However, the empirical literature suggests that the role of past promotions goes substantially beyond lags of order two or three. In actual internal labor markets, the promotion cycle is likely to be relatively long, perhaps 7 years as pointed out by Lazear, 1992. Therefore, it would not be realistic to estimate a model that restricted the number of lags to two or

three.¹² For this reason, we disregard the short run dimension of promotion dynamics, and focus on a summary of all past promotion outcomes.

Ideally, we would like a measure of past histories that embodies the signal provided to the labor market regarding the caliber of the executive. The theoretical literature considers the importance of the signal provided by initial promotion, assuming a common starting level. Were the initial placement levels considered as well as promotions, the importance of this signal would be just as relevant. Our measure should therefore be capturing the effect of early promotion history as well as the level of the initial placement in the firm's hierarchy. In order to capture both aspects, we define a speed of promotion variable (referred to as *Speed* below) which is measured as the ratio of the level an executive has risen to by the start of the sample to the executive's years of labor market experience. With level 1 representing CEO's, levels fall with promotions and higher initial assignments. Since it is intuitively easier to think of promotion speed as a positive number, we look at the level an executive has risen to at the start of the sample in reference to level 12, the lowest level reported in the data. The reference level chosen is irrelevant as it changes the number of levels an executive has risen equally across executives. If we were to measure the speed of promotion only by considering the number of promotions, those who entered at a higher level would have fewer promotions due to starting closer to the top of their hierarchies. As such, we would then be confounding these executives with those who started beneath and have a lower promotion probability for other reasons.

The distinction between individual and firm specific attributes is also problematic, given the structure of the sample data. While it is possible to observe a few firm specific variables (to be discussed below), the movement of executives between firms cannot be observed in the data set that we use. Therefore, the data do not allow us to identify the firm specific unobserved term from the individual specific term, unlike what is done in Abowd, Kramarz and Margolies (1999) and Belzil (2000).¹³ Without loss of generality,

¹²To see this argument, consider estimating a model where the current promotion probability depends on the past two or three promotion outcomes. These parameter estimates would turn out to be negative and would imply that simulated promotion histories entail penalizing executives who have been promoted during the sampling period. Indeed, we have verified this assertion by estimating a dynamic promotion model where current promotions depend on up to 3 or 4 past promotion outcomes. All parameters turn out to be negative, although those pertaining to order 3 and order 4 were much weaker (very close to 0).

¹³In these papers, identification is rendered possible by a relatively small number of

we therefore refer to the unobserved factors as individual specific.

Finally, with respect to estimation issues, our choice of an econometric estimation technique is largely dictated by the need to allow for multi-dimensional population heterogeneity in promotion dynamics. Our analysis, as are most microeconomic analyses, is plagued by the presence of unobserved individual and firm factors. For this reason, the effects of past promotions may hardly be seen as a single parameter. Indeed, in many areas of microeconomics, classical models in which marginal effects are summarized by a single parameter are gradually being replaced by more general models with slopes as well as intercepts that are observation specific.¹⁴ In order to assess the relevance of theoretical models predicting positive promotion dynamics, we need to recover the marginal effects associated with the variable capturing the effect of past promotion and, ultimately, compare them to the marginal effects of variables such as education and tenure. A key objective is therefore to recover the population distribution of individual specific promotion dynamics parameters. For these main reasons, we focus on random effect estimation techniques.¹⁵ While random effects techniques require a specification of the initial conditions of the stochastic process analyzed and are often implemented in a fully parametric framework, we propose a random effect estimation strategy based on flexible (semi-parametric) methods. We use a finite mixture model in order to characterize the distribution of individual specific intercepts and slopes.¹⁶ This approach allows us to minimize the impact of distributional assumptions needed in order to implement such a model.

individuals observed, at different points in time, in two or more firms.

¹⁴Recent papers, including Alvarez, Browning and Ejrnaes (2003), have pointed out the particular need for multi-dimensional heterogeneity in the modeling of earnings growth.

¹⁵One of the advantages of conditional likelihood techniques is the fact that statistical inference may be achieved without having to specify a distribution for the individual specific effects, including the initial conditions. However, the conditional approach precludes the estimation of time invariant regressors such as schooling, and does not allow one to recover the marginal effects.

¹⁶This approach is also common in empirical dynamic programming models with unobserved heterogeneity (Keane and Wolpin, 1997, Eckstein and Wolpin, 1999, Belzil and Hansen, 2002). It is largely influenced by the estimation method proposed by Heckman and Singer, 1984.

4 Econometric Model

In this section, we present the econometric model. Modeling promotion dynamics requires i) the definition of a promotion, ii) a parametric form for the promotion probability and, in particular, the allowance for the dependence of current promotion on endogenous past promotion histories, iii) the role of unobserved heterogeneity, iv) a definition of the initial conditions' distribution, namely the initial speed of promotion achieved or the rank level in the firm at the start of the sampling period.

4.1 The Definition of a Promotion

The aim of the model is to make inference about individual promotion histories from a sequence of rank levels (within a firm) occupied by individuals. The sequence contains up to 8 years of data. We define a promotion as a negative change in level (an accession to a higher rank in the hierarchy); that is

$$Y_{ijt} = 1(L_{ijt} - L_{ijt-1} < 0) \quad (1)$$

where L_{ijt} is the rank of individual i , in firm j , at time t and $1(\cdot)$ is the indicator function.¹⁷ In total, this results in 7 potential promotion outcomes per individual. We do not distinguish between demotions and absence of promotions.¹⁸ Similarly, we do not make the distinction between promotions for more than 1 level (a rare event) and standard promotions taking place when an individual reaches the next level. Promotion is coded as a binary variable. A promotion of one or more levels is coded as a one. A demotion or unchanged level is recorded as a zero. In order to minimize the impact of measurement error, if the level in the year subsequent to a promotion reflects a demotion, the original improvement in level is regarded as a coding error and no promotion is recorded. This means that we do not code as promotions improvements in level that last just one year. Similarly, if a worker is demoted in one year and promoted in the next to the original level, this return to the original level is not recorded as a promotion. Last,

¹⁷The reader should remember that a smaller number for the level variable (L_{ijt}) implies a higher rank.

¹⁸Making the distinction would require the use of an ordered model. We initially experimented with this possibility, but estimation was rendered difficult by the low incidence of demotions.

promotions are only registered when they constitute an improvement over the workers initial level in the data. This ensures that promotions register only when they constitute a net upward movement over the span of observations and not just upward movement over the previous year.

4.2 The Promotion Probability

The basic element of our econometric strategy is the following promotion probability

$$\Pr(Y_{ijt} = 1) = \Lambda(\beta_X \cdot X_{it-1} + \beta_W \cdot W_{jt-1} + \beta_{L6} \cdot L_{6it-1} + \dots \beta_{L2} \cdot L_{2it-1} + \beta_i^P \cdot Speed_{it-1} + \alpha_i^P) \quad (2)$$

where

- X_{it} represents a vector of individual specific attributes (education, tenure in the firm, an indicator, newcomer, for which 1 signifies that the executive is in the first year of employment with the firm and age measured at the initial sample period). Of those variables, only tenure is time varying. To achieve as much flexibility as possible, the education variable may be transformed into several class variables. These are i) those who have obtained a high school degree or less (12 years or less), ii) those who have attended some post high school training (13 to 15 years), iii) college graduates (16 years), iv) those who have obtained 17 or 18 years and v) those who had 19 years or more. The high school graduates and drop-outs are the reference group.
- W_{jt} represents a vector of firm specific time varying variables reflecting levels and the changes in variables such as firm size (employment) and profits. As a promotion is defined as the change in level from $t - 1$ to t , profits and employment are the average values of the current and the previous period changes.¹⁹ Profits are measured in millions of 1980 US dollars. Firm size is measured in thousands of employees. Employment changes are calculated as the percentage over the previous year.

¹⁹This type of smoothing is common in the empirical literature on worker reallocation (see Belzil, 2000, for an example).

- L_6, \dots, L_2 are endogenous time varying binary indicators equal to 1 if the individual is at the rank level indicated by the subscript and 0 if not. Level 7 is the reference group. Level 1 (CEO) is not included because CEO's cannot be promoted.
- $Speed_{it}$ measures the speed of promotion achieved up to date t . It is calculated as the ratio of the number of levels reached at any point in time (in reference to level 12) and the difference between age and years of education (minus 5). It is meant to capture the structural fast track hypothesis. As individuals are observed over the sampling period, the speed of promotion is adjusted according to the following law of motion

$$speed_{it} = \frac{\#levels_{t-1} + Y_{it}}{(age_{t-1} - educ - 5) + 1}$$

Note that we assume that the parameter space of the individual specific parameters, β_i^S , is unrestricted. This means that, as such, negative structural fast tracks, are not ruled out. One reason for not restricting the parameter space is that we certainly do not want to rule out the possibility that the mean of β_i^S is 0 in the population. If we restricted the parameters to be positive valued, this would most likely have a key impact on the estimated distribution of β_i^S . Another reason is that a negative fast track, at least for a certain fraction of the population, may be explained economically (see Section 5.1.2).

- α_i is an individual (or perhaps firm) specific unobserved term which represents individual unobserved heterogeneity such as unobserved skills and motivation or unobserved differences in firm structures. It is detailed below.
- The promotion probabilities are assumed to be logistic; that is

$$\Lambda(.) = \frac{\exp(.)}{1 + \exp(.)} \quad (3)$$

Conditional on unobserved individual specific unobserved heterogeneity, the promotion outcomes are assumed to be independent.

- $(\beta_x, \beta_W, \beta_{L6} \dots \beta_{L2})$ are parameters to be estimated and α_i^P and β_i^P are individual specific unobserved terms with distribution functions that have to be estimated (approximated).

4.3 The Initial Condition Problem

In order to resolve the initial condition problem, we formulate a model for the initial speed of promotion. Our solution is therefore in the spirit of Heckman, 1981. The alternative approach is to define the distribution of the unobserved heterogeneity term(s) conditional on the initial conditions.²⁰

We assume that the initial speed of promotion achieved by an individual (by the start of the sampling period) is represented by a linear regression function. That is

$$speed_{it_0} = X_{it_0}^s \cdot \delta + \alpha_i^s + \varepsilon_{it_0} \quad (4)$$

where $speed_{it_0}$ is the initial speed, $X_{it_0}^s$ is a set of regressors containing education and age, and where ε_{it_0} is i.i.d. Normal with mean 0 and variance σ_S^2 . As many individuals have started with a different firm, we cannot really condition on tenure. So only age and education are used as regressors. The endogeneity of the initial speed of promotion is taken into account in the model by assuming that unobserved heterogeneity affecting the current level reached by an individual, α_i^s , is correlated with α_i^P . Note that we are implicitly assuming that conditioning on α_i^s is sufficient to estimate the level specific effects. This assumption appears reasonable as the measured speed of promotion and the initial level are closely correlated.

4.4 Endogenous Sampling and Attrition

A striking feature of our data is the relatively wide range in the number of times individuals are reported. As it is impossible to rule out attrition bias or endogenous reporting by the firm, we recognize that the distribution of unobservable promotability (α_i^P) is not necessarily orthogonal to the frequency at which individuals are reported. To resolve this problem, we interpret the distribution of unobserved heterogeneity as conditional on the number of time an individual is reported.²¹ We assume that

²⁰This entails writing $\alpha_i^P = \alpha_{0i}^P + \alpha_1^P * speed_{it_0}$ and $\beta_i^S = \beta_{0i}^S + \beta_1^S \cdot speed_{it_0}$ where α_{0i}^P and β_{0i}^S are the residual parts of the unobserved heterogeneity terms, assumed to be orthogonal to the initial speed of promotion ($speed_{it_0}$).

²¹Conditioning on number of observations in order to resolve initial condition problems is discussed in Wooldridge (2003).

$$\alpha_i^S = \alpha_{0i}^S + \alpha_2^S \cdot N_i \quad (5)$$

$$\alpha_i^P = \alpha_{0i}^P + \alpha_2^P \cdot N_i \quad (6)$$

and

$$\beta_i^P = \beta_{0i}^P + \beta_2^P \cdot N_i \quad (7)$$

where N_i is a measure of imbalance in the panel. Typically, N_i is the number of times an individual is reported by its firm or, alternatively, it is a set of binary variables which exhaust all the possible number of transitions reported (2 to 7).

4.5 The Distribution of Unobserved Heterogeneity

Our estimation method is based on the premise that α_{0i}^p , β_{0i}^S and α_{0i}^S are jointly distributed according to an unknown cumulative distribution function, $H(\cdot)$, which is approximated using a multivariate discrete distribution of the form:

$$\Pr(\alpha_{0i}^p = \alpha_{0k}^p, \beta_{0i}^S = \beta_{0k}^S, \alpha_{0i}^S = \alpha_{0k}^S) = p_k$$

where $k = 1, \dots, K$. The number of types, K , is assumed to be known, although it is the outcome of various experimentations.

The type probability, p_k , is estimated using a logistic transform, that is

$$p_k = \frac{\exp(\delta_k)}{\exp(\delta_0) + \dots + \exp(\delta_K)} \quad (8)$$

with the obvious restriction that probabilities sum up to 1.

4.6 The Likelihood Function

The likelihood function represents the joint probability of the speed of promotion already achieved at the beginning of the sampling period and the sequence of promotion outcomes observed until the end of the panel. As the panel data cover eight years, we therefore have a maximum of seven potential promotion outcomes per individual. Theoretically, obtaining parameter

estimates requires the maximization of the following log likelihood function, L_i ;

$$L_i = \log \left\{ \int \frac{1}{\sigma_L} \cdot \phi \left(\frac{L_{it_0} - X_{it_0}^S \cdot \delta - \alpha_{i0}^S}{\sigma_L} \right) \cdot \Pr(Y_{ij1}, \dots, Y_{ij7}; \alpha_{i0}^P, \beta_{i0}^P) \cdot dH(\alpha_{i0}^S, \alpha_{i0}^P, \beta_{i0}^P) \right\} \quad (9)$$

where $\Pr(Y_{ij1} \dots Y_{ij7})$ is the joint probability of the observed promotion outcomes. However, in practice, the objective function is a discrete approximation (a weighted sum) of (8), where the weights are given by the population proportion defined in (7).

$$L_i = \log \left\{ \sum_{k=1}^K p_k \cdot \frac{1}{\sigma_L} \cdot \phi \left(\frac{L_{it_0} - X_{it_0} \cdot \delta - \alpha_{0k}^S}{\sigma_L} \right) \cdot \Pr(Y_{ij1}, \dots, Y_{ij7}; \alpha_{0k}^P, \beta_{0k}^P) \right\} \quad (10)$$

4.7 An Overview of the Model Specifications

The results presented below represent only a portion of the empirical work that we have done over the course of this project. We have worked with various model specifications. Basically, the specifications differ in terms of the treatment of the initial conditions as well as in terms of the treatment of the education variable. Finally, as the number of types (K) is not formally estimated, we experimented with different numbers of types.

First, with respect to the treatment of initial conditions, all model specifications reported in the following sections have also been carried out in a framework where the distribution of unobserved heterogeneity is conditional on the initial condition (the initial speed of promotion). The details of the estimation procedure are found in footnote 17. As all fundamental results, especially those related to the promotion dynamics and the marginal effects, were comparable, we report the results of the model where the initial condition is actually modeled.²²

With respect to the choice of variables, we consider two versions of the model. In the first specification, we condition on schooling and assume that schooling is exogenous. In a random effect structure such as the one estimated here, exogeneity implies, among other things, that the distribution of

²²For those interested, the other estimates may be found in a companion working paper (Belzil and Bognanno, 2004).

the unobservable factor is orthogonal to schooling. This assumption is questionable. In an effort to minimize the number of unduly strong assumptions, we re-estimated the model while omitting schooling explicitly (therefore including it implicitly in the individual specific unobserved term). In such a framework, only tenure and age are forced to be orthogonal to the individual effects. These assumptions are certainly less controversial. Both versions are found in the empirical result sections below.

Finally, with respect to the number of types used to estimate the mixture models, all estimates reported herein are for the case with four types. Going up to six types turned out to be unproductive as all the basic results (mean, variance and correlations) may be illustrated with 4 types.

5 Models with Heterogeneous Promotion Dynamics

In this section, we first present the parameter estimates obtained from both model specifications (with and without education). These are found in Section 5.1. In order to illustrate the results, we present the marginal effects corresponding to the main parameters of the model (promotion dynamics, education, age and tenure) and discuss their relative magnitude (in Section 5.2). Finally, using a variance decomposition of the index function, we illustrate the relative importance of unobserved factors.

5.1 Parameter Estimates

The parameter estimates are found in Table 1. The first column is devoted to the model that estimates the effect of education (Model 1). The second contains estimates for the case in which education is part of the unobserved heterogeneity term (Model 2). For both models, the estimates related to the initial condition (the initial speed of promotion) are found in the upper part while those related to the promotion probabilities are in the lower part. At the outset, and as indicated by the asymptotic standard errors, it should be noted that virtually all parameters are estimated very precisely. This is a reflection of the relatively large number of individuals and periods available in the data. Both models are estimated for the case where $K=4$ (four types of executives).

5.1.1 Model with Exogenous Education

The set of estimates for the initial condition equation indicates that the initial speed of promotion decreases with age (-0.0127), and increases with schooling (the estimates range from 0.5751 for those who have attended college to 0.6811 for those who have 19 years or more). With regards to age, this result is expected because younger workers are at a stage in their career where promotions are more likely. With regards to schooling, the results are consistent with both the more educated being promoted more frequently and beginning at a higher initial level. The estimates for the type specific intercept terms (α_{0k}^S) range from 0.8839 (type 1) to approximately 0.32 (for type 2, type 3 and type 4), suggesting that the marginal distribution of the individual specific term of the initial promotion speed equation may be summarized by two types. Obviously, executives of type 1 are those who have achieved a greater speed of promotion by the start of the sampling period. The correlation between this individual specific term and other heterogeneity components of the model are analyzed below.

Turning to the estimates of the promotion probability equation, found in the lower part of Table 1, we find that unobserved heterogeneity (α_{0i}^P) plays an important role as it is possible to identify 4 distinct types; type 1, which corresponds to 23% of the population, has a higher promotion probability (with an intercept of -1.3496) while type 4 individuals (27% of the population) have the lowest promotion probability (with an intercept of -2.9332). It is relatively clear that, given age and education, those who have achieved a higher speed of promotion initially (type 1) are also those who tend to be more likely to be promoted subsequently. A more formal analysis may be performed using various summary statistics of the distribution of unobserved heterogeneity (mean, variance and correlations), found in Tables 3, 4 and 5. Indeed, the implied correlation between unobserved promotability and the initial speed of promotion (between α_{0i}^P and α_{0i}^S) is equal to 0.91 (Table 4).

It is also important to note that, after conditioning on unobserved heterogeneity and age, the effect of tenure tends to be very small. The parameter estimate for tenure is negative (-0.0029). The relatively small magnitude of the parameter (to be discussed below in conjunction with marginal effects), is most likely explained by the relatively high dispersion in the intercept terms and suggests that, given unobservable factors, how long one has served in the firm is practically irrelevant for the purpose of predicting promotion outcomes. However, newcomers to the firm have a lower promotion probability.

This is not surprising. We also find that promotion probabilities increase with age (0.0172). However, as we are conditioning on tenure and level in the firm, the positivity of the parameter estimate is not inconsistent with the negativity of the effect of age on the initial speed of promotion.

Among the standard human capital variables used in the empirical literature, schooling appears to be the only one affecting promotion outcomes. Assuming exogeneity of the education variable with respect to the stochastic process generating promotions, the schooling binary variables, ranging between 0.0147 and 0.3074, indicate clearly that schooling increases promotion (except for schooling levels exceeding 18 years). The positive effect of education on promotion outcomes is certainly consistent with the steeper slope of age earnings profile of the more educated.

In our analysis, firm profits and size, along with their changes, are the only firm variables included. We include them in order to have additional control variables for unmeasured firm factors and but we do not really focus on firm heterogeneity. Indeed, firm unobserved heterogeneity cannot be distinguished from individual unobserved heterogeneity. As is clear from the results, they play practically no role in the promotion process. We find that the effect of firm size and profits on promotion are negative (-0.0015 and -0.0001) and very small, although firm size is relatively precisely estimated. This may reflect the fact that, other things equal, promotions are more likely in smaller firms. The effects of firm employment changes and profit changes are even smaller. They may indicate that the promotion process of American executives is not sensitive to the business cycle.

We also find evidence that promotion outcomes are largely dependent on the current level of the manager. The rank specific dummies (ranging from -0.13 at rank 6 to -2.65 at rank 2) indicate that given all individual and firm specific endowments, promotion probabilities become smaller as one reaches higher ranks. The average promotion probabilities per level, are found in Table 6. They indicate that, although the average promotion probability is around 0.12, the level specific average probabilities range from 0.23 (level 7) to 0.02 (level 2).

As indicated in Section 3 and Section 4, the distribution of unobserved heterogeneity is understood to be conditional on the number of times an individual is observed. As may be seen from the results, there is some evidence that the per-period promotion probability is smaller for those individuals for whom we have more observations. The effect (-0.0040) is however small. At this stage, it is difficult to give a clear interpretation of this result. The re-

porting frequency imbalances may well be explained by a diversity of factors such as outside mobility or inconsistent reporting by firms.²³

Given the focus of this paper, those estimates that generate the most interest pertain to the effects of the speed of promotion on promotion outcomes. As is clear from the speed of promotion equation in Section 4.2, the speed of promotion variable is taking into account early career promotions as well as promotions taking place during the sample period. As it was the case for the promotion probability intercept terms, there is substantial heterogeneity in the slope parameters. There are four distinct types identified in the data. Type 1 individuals, endowed with a high promotion probability, have a slope nearly equal to 0 (namely -0.0046). On the other hand, those of other types, endowed with lower promotion probabilities, are endowed with positive slopes. When averaged over all four types, the effect of past speed of promotion is 0.20 (see Table 3). This means that, on average, there exists a structural positive fast track. However, for 23% of the population (type 1), current promotion probabilities are independent of past promotion histories. In other words, type 1 individuals achieved more frequent promotions but do not derive advantages from their promotion histories.²⁴ These results suggest a negative correlation between the propensity to be promoted and the existence of a structural fast track.

In general, we observe that cross sectional differences in promotability are more important than differences in the effects of past speed of promotion (as indicated by the relative standard deviations found in Table 3). More precisely, the standard deviation of the intercept term (α_i^P) is 0.60 while the standard deviation of the slope (β_i^P) is 0.12. In terms of the index function (the propensity to be promoted), these numbers, coupled with the range of the initial speed variable (an average of 0.37 with a standard deviation equal to 0.21) will imply that differences in the intercept term have a much larger explanatory power than differences in previous promotion histories.

5.1.2 Model without Education

Estimating a promotion model with schooling endogenous is beyond the scope of this paper but we may at least investigate the robustness of the results

²³We have also estimated the model without conditioning on observation frequency and found very similar results.

²⁴This may easily be verified upon simulating the type specific promotion histories, given the parameters of the model. This, and related issues, are discussed below.

by re-estimating a version of the model in which schooling is ignored and implicitly part of the unobservable terms α_{0i}^P and α_{0i}^S . These estimates are found in column 2 of Table 1. As is evident from the results presented in column 2, estimating a model without conditioning on education leads to quite similar results.

Overall, the model specification is capable of replicating the main features of the specification in which schooling is assumed exogenous. For instance, we still find evidence that those who have been promoted at a faster rate in the past (the initial condition) are also more likely to be promoted in the future. This may be verified formally from the correlation found in Table 5 (equal to 0.9056). While there might be a slightly larger variability among the promotion intercepts (as seen by the standard deviation of the distribution in Table 3) when schooling is ignored, we do not note any meaningful changes. As well, the estimates also imply that differences in level constitute the prime determinant of individual differences in promotion histories. While the sign of the tenure variable has flipped side (tenure is now positive), the estimate is even smaller in absolute terms (the estimate is 0.0014).

The major change that deserves some discussion is found at the level of the individual specific promotion dynamics variable. When schooling is ignored, we are now able to identify four distinct slopes for the effect of past promotion. The major difference is that for type 1 (representing only 18% of the population), the effect of the speed of promotion is negative (-0.08). The other type specific parameters are -0.0040 (type 2), 0.2481 (type 3) and 0.1940 (type 4). The population average parameter, 0.0975 (found Table 3), is however still positive but is closer to zero. This decrease in the population average is mirrored in the corresponding marginal effect (to be discussed below).

The negativity of the promotion dynamics parameter is interesting in itself and deserves some attention. Negative fast track effects are, as far as we know, never mentioned in the literature. For this reason, it may be tempting to restrict the parameter space and impose positivity. We chose not to do so. Our reading of the fast track effect is the following; if the speed of promotion raises current and subsequent promotion probabilities, those who have been promoted first will build a comparative advantage in promotions. In practice, negative fast track effects may take the following forms. In a world where individual abilities are eventually known by the firm and where identical individuals achieve the same final level, the realization of an abnormally high rate of early career promotions may simply be compensated

by a lower promotion rate later. Another related explanation could be the incidence of promotions at a level that exceeds one’s level of competence. If a significant fraction of the population reaches their level of incompetence, their subsequent rate of promotion may be negatively correlated with the speed of previous promotions. In personnel economics, this notion is sometimes referred to as the “Peter Principle”.²⁵ This feature of the promotion process may also be explained statistically. In dynamic discrete panel data models, it is by definition impossible to identify those individuals who have reached a maximum level (for whom the promotion probability falls to 0) from those who face a positive probability but have simply not been promoted. Put differently, it is impossible to estimate an individual specific maximum level below which promotion probabilities are positive and above which it is equal to 0. If the promotion process is bounded, the individual specific maximum level would have to be made function of observed human capital as well as unobserved heterogeneity, most likely correlated with the promotion probability intercept. Obviously, the identification of two sources of unobserved heterogeneity from individual series of promotion outcomes would be tenuous.

To summarize, we retain the fact that, on average, structural fast track effects are qualitatively small, positive on average and are not a key determinant of observed promotion histories. Our results has implications for the literature on wages/earnings growth. If promotion process depicts serial correlation, so should the wage growth process. Furthermore, after controlling for persistent individual specific factors, the structural correlation between current and past wage growth should be relatively small. As far as we know, this issue is never investigated in the earnings dynamics literature. Up to now, most of the researchers have investigated the level of serial correlation in wages and ignored wage growth.

5.2 The Marginal Effects

To shed light on the parameter estimates, the corresponding marginal effects are reported in Table 2. The marginal effects are computed for each individual and averaged over the entire sample. As noted before, the level of statistical significance of the parameters is high enough that we are not re-

²⁵The relevance of the Peter principle in economic models of promotions is discussed in Lazear, forthcoming.

ally concerned with the precision of the marginal effects. Instead, we report a standard deviation of the marginal effects which illustrates the cross sectional differences in the marginal effects (for given parameter values). With respect to accumulated schooling, we report all the marginal effects (the effect of completing college, the effect of obtaining 2 years beyond college as well as the marginal effect of going beyond 2 years after college) in reference to those who attended (or graduated from) high school. We do this because high school graduates and high school dropouts constitute only a small fraction of the population.

The estimates indicate that college completion increases annual promotion probabilities by 0.0195 while obtaining 18 years increases it by 0.0287. Reaching the next schooling category (more than 18) increases it by 0.0266. As a comparison, the marginal effect of age is 0.0017 in the model with schooling and 0.0015 in the model where it is treated as unobserved. As conjectured before, the marginal effects of tenure (-0.0003 in the model with schooling and 0.0060 in the model without schooling) are found to be very small. These estimates imply that, after conditioning on unobserved skills explaining promotions, time spent in the firm or in the labor market is irrelevant.

In our model, the marginal effects of past promotion speed depend on types as well as on individual regressors. When computed at the average value of the type specific slopes, which is equal to 0.20, the marginal effect for the model with schooling averages 0.0157. It is therefore comparable, although somewhat smaller, to the marginal effect associated to college completion. In the model where education is ignored, the marginal effect, averaging 0.0089, is smaller and is characterized by a wider dispersion. It appears that statistical inference about the role of past promotion histories on future promotion outcomes is not hindered by the exogeneity assumption of the education variable.

Finally as was made clear already from Table 1, differences in levels correspond to the largest marginal effects. They may easily be inferred from table 6. On average, the promotion probabilities decrease between 0.03 and 0.05 as an executive reaches a higher level. This is obviously true in both model specifications.

At this stage, we can therefore assert that differences in promotion outcomes are explained mostly by differences in level and, to a lesser extent, by individual differences in unobserved individual specific attributes (or perhaps unobserved firm attributes). When compared in terms of their marginal ef-

fects, individual differences in past promotion histories are less important than differences in education.

5.3 Decomposing the Index Function

In order to assess the relative importance of each variable (or group of variables), we simulated promotion histories from the parameters of Model 1 and worked with the single index function explaining the promotion probability propensity. Our objective is to evaluate the explanatory power of each variable (or group of variables). Our measure of explanatory power is reported as a percentage variable. The percentages denote the loss in explanatory power of the explained part of the index function regression for each group of variables. They are computed from the difference in the coefficient of correlation from the regression that includes all factors and a regression that excludes only each particular variable or group of variables. The results are in Table 7.

As suggested by our previous analysis of the marginal effects, differences in promotion probabilities are explained primarily by differences in level. In other words, dropping the level indicators reduce R^2 by 0.46. Omitting the individual specific promotion probability intercept terms reduced it by 0.14 while omitting human capital endowment variables (age, tenure and schooling) reduced it by 0.035. Finally, differences in past promotion histories, as well as differences in the promotion dynamics slopes, are each found to account for less than 1% of the total variation in the index function.

Before concluding that the human capital variables are not important, it should be recognized that they are much more influential in estimating the initial condition. The second column of Table 7 shows that age and schooling explain 35% of the variation in the initial speed of promotion. They therefore must either raise the level of entry or the rate of early promotion, or both. However, since over the sampling period, we observe executives that are both young and of low tenure and find a relatively small causal effect on promotion probabilities, we suggest that age and education primarily influence the level of entry.

Our results illustrate that the finding of “fast tracks” in the empirical literature results almost purely from unobserved heterogeneity (in the promotion probability intercept) among American executives. Serial correlation in promotion histories may hardly be given a structural (causal) interpretation.

6 Interpreting Heterogeneity in Promotion Dynamics: The Role of Human Capital

In this section, we investigate the nature of population heterogeneity, already documented above. To do this, we extend our econometric model to allow for interaction terms between individual attributes and the speed of promotion variable. We pay a particular attention to two variables, education and tenure. To achieve a high level of flexibility, we preserve the same stochastic specification of the promotion dynamics parameters (4 types), which implies that the individual specific slopes are expressed as

$$\beta_{it}^S = \beta_{0i}^S + \beta_2^S \cdot N_i + \beta_3^S \cdot Education_i + \beta_4^S \cdot tenure_{it} \quad (11)$$

Note that estimation of this model specification requires education to be measured in years and that, given the time varying nature of tenure, the slope will automatically vary with time (with tenure). While there might exist more flexible methods to allow for interactions (such as spline functions allowing for the slope to differ at all (or many) possible values of education and tenure), we retain the standard interaction term in order to keep the number of parameters at a manageable level and because our objective is only to infer the sign of the derivative of the slope with respect to tenure and education.

6.1 The Effects of Education on Promotion Dynamics

As it was indicated before, our estimates indicate that the differences in the individual specific intercept terms of the promotion probability index function are negatively correlated with the individual specific parameters measuring the effect of speed of promotion. If more educated workers are more promotable, it is therefore possible that their past promotion histories are less important and that differences in education might account for a portion of the cross-sectional differences in individual specific slopes. Within a behavioral framework, this may be explained by the fact that, from the perspective of the firm, uncertainty about individual abilities is decreasing with education and that, in the event where past promotion histories are

used as a signal, the significance of the signal is therefore decreasing with schooling.

To preserve the compactness of the results, we present a summary of the type specific slope parameters.²⁶ These estimates are found in Table 8. As conjectured before, the estimate for the education interaction term (β_3) is negative and equal to -0.0234. It indicates that, as individuals get more schooling, the effect of past promotion goes toward 0 (or negative numbers). The individual specific slopes average to -0.0582 and the related marginal effect is equal to -0.0062 (found in Table 9). These estimates indicate that, on average, past promotions are not a really important determinant of subsequent promotions.

6.2 The Effects of Tenure on Promotion Dynamics

If past promotion histories are used as a signal by employers, it is also natural to expect the effect of past speed of promotion to decrease with tenure. This decrease could reflect a simple information acquisition from the perspective of the firm. This assertion is verified upon looking at the estimates of Table 8. The parameter estimate for the interaction term between tenure and speed of promotion is also negative (-0.0093). Because the related marginal effect is not as strong as for schooling, the average slope is now positive (0.2014) and the marginal effect is 0.0221 (the highest value found so far).

At this stage, we may conclude that population heterogeneity in the effects of past promotion histories is inversely related to the human capital accumulation process. In other words, the promotion process of more educated workers is less affected by structural fast track effects than those who have less education and, as workers accumulate job specific human capital (as indicated by tenure), the positive fast track effect tends to vanish. Both findings are consistent with the possibility that structural fast track effects are inversely related to the importance of information asymmetries (schooling and tenure reduce the uncertainties about worker's ability and motivation). Accumulated schooling and tenure may reduced the risk associated with promotion (from the perspective of the firm), and may therefore lead firms to ignore past histories.

²⁶Other estimates are found in Belzil and Bognanno (2004).

Table 1- Parameter Estimates

	Model 1	Model 2
	Exogenous schooling	Without Schooling
	Parameter and (asympt. st-errors)	Parameter and (asympt. st-errors)
Initial Condition		
Schooling		
12 years or less	-	-
13-15 years	0.5751 (0.0022)	-
16 years	0.6018 (0.0013)	-
17-18 years	0.6514 (0.0014)	-
19 years or more	0.6811 (0.0018)	-
age	-0.0127 (0.0001)	-0.0228 (0.0003)
α_{0type1}^S	0.8839 (0.0035)	0.9234 (0.0043)
α_{0type2}^S	0.3171 (0.0021)	0.3523 (0.0023)
α_{0type3}^S	0.3203 (0.0022)	0.3220 (0.0020)
α_{0type4}^S	0.3214 (0.0004)	0.3012 (0.0035)
σ_S	0.0762 (0.0021)	0.0923 (0.0006)
Promotion Probability		
Schooling		
12 years or less	-	-
13-15 years	0.0147 (0.0004)	-
16 years	0.2208 (0.0011)	-
17-18 years	0.3074 (0.0005)	-
19 years or more	0.2915 (0.0004)	-
age	0.0172 (0.0018)	0.0105 (0.0019)
newcomer	-0.0477 (0.0004)	-0.1057 (0.0169)
tenure	-0.0029 (0.0013)	0.0014 (0.0016)
firm size	-0.0015 (0.0003)	0.0011 (0.0004)
firm profits	0.0001 (0.0007)	-0.0020 (0.0007)
Δ firm size	0.0001 (0.0012)	0.0001 (0.0019)
Δ firm profits	-0.0001 (0.0005)	-0.0001 (0.0400)

Table 1- continued

	Parameter and (asyp. t-ratio)	Parameter and (asyp. t-ratio)
Promotion Speed		
$\beta_{0,type1}^S$	-0.0046 (0.0012)	-0.0835 (0.0230)
$\beta_{0,type2}^S$	0.2482 (0.0014)	-0.0040 (0.0147)
$\beta_{0,type3}^S$	0.3203 (0.0005)	0.2481 (0.0132)
$\beta_{0,type4}^S$	0.1222 (0.0004)	0.1940 (0.0130)
level 7	-	
level 6	-0.1509 (0.0006)	-0.1305 (0.0204)
level 5	-0.5459 (0.0004)	-0.5808 (0.0072)
level 4	-1.0579 (0.0020)	-1.1686 (0.0135)
level 3	-1.8034 (0.0040)	-1.8559 (0.0109)
level 2	-2.6500 (0.0030)	-2.9491 (0.0252)
$\alpha_{0,type1}^P$	-1.3496 (0.0025)	-2.0652 (0.0515)
$\alpha_{0,type2}^P$	-1.9081 (0.0006)	-2.3039 (0.0270)
$\alpha_{0,type3}^P$	-2.5648 (0.0011)	-2.3784 (0.0376)
$\alpha_{0,type4}^P$	-2.9332 (0.0009)	-2.2587 (0.0427)
Type Probabilities		
Prob <i>type1</i>	0.2276 (0.0070)	0.1812 (0.0145)
Prob <i>type2</i>	0.2604 (0.0033)	0.3514 (0.0967)
prob <i>type3</i>	0.2514 (0.0009)	0.2655 (0.0323)
Attrition		
α_2^S (initial condition)	0.0001 (0.0003)	0.0003 (0.0002)
α_2^P (prom. prob. intercept)	-0.0040 (0.0007)	-0.0082 (0.0080)
β_2^S (prom. prob. slope)	0.0123 (0.0005)	0.0067 (0.0012)
# of individuals	25000	25000
mean # of periods	4.6	4.6
mean log likelihood	-0.302942	-0.398917

Table 2
Some Marginal Effects from Model 1

	Model 1	Model 2
	Estimate (st-deviation)	Estimate (st-deviation)
Individual Specific Variables		
16 years of schooling	0.0195 (0.0114)	-
17-18 years of schooling	0.0287 (0.0165)	-
19 years of schooling or more	0.0266 (0.0154)	-
tenure	-0.0003 (0.0002)	0.0060 (0.0029)
age	0.0017 (0.0009)	0.0015 (0.0012)
Promotion dynamics		
Speed of promotion	0.0157 (0.0133)	0.0089 (0.0144)

Note: The marginal effects (the effect of completing college, the effect of obtaining 2 years beyond college as well as the marginal effect of going beyond 2 years after college) are in reference to those who have attended high school. The marginal effects are averaged over all individuals. The reported standard deviations is a measure of cross-sectional dispersion in the marginal effects, given parameter estimates.

Table 3
Unobserved heterogeneity: Means/St. Deviations

	Model 1	Model 2
	Mean (st. deviation)	Mean (st. deviation)
α_i^S	0.3895 (0.2913)	0.4105 (0.3011)
α_i^P	-2.1884 (0.6051)	-2.2798 (0.7025)
β_i^P	0.2014 (0.1228)	0.0975 (0.0987)

Table 4
Unobserved Heterogeneity Correlations:
Model 1 (with schooling)

	α_i^S	α_i^P	β_i^P
α_i^S	1.0000	0.9154	-0.8990
α_i^P	-	1.0000	-0.8779
β_i^P	-	-	1.0000

Table 5
Unobserved Heterogeneity Correlations:
Model 2 (without schooling)

	α_i^S	α_i^P	β_i^P
α_i^S	1.0000	0.9056	-0.7634
α_i^P	-	1.0000	-0.7723
β_i^P	-	-	1.0000

Table 6
Average Promotions Probabilities across Rank
in Model 1 and Model 2

Rank	average probability (standard deviation)	average probability (standard deviation)
level 7	0.2357 (0.1062)	0.2223 (0.1097)
level 6	0.2128 (0.0988)	0.2011 (0.1002)
level 5	0.1561 (0.0778)	0.1438 (0.0822)
level 4	0.1000 (0.0533)	0.0996 (0.0349)
level 3	0.0517 (0.0290)	0.0612 (0.0300)
level 2	0.0230 (0.0133)	0.0286 (0.0188)
All levels	0.1213 (0.0833)	0.1192 (0.0766)

Note: The promotion probabilities are averaged over all individuals at a particular rank. The reported standard deviations are a measure of cross-sectional dispersion in the promotion probabilities, given the parameter estimates.

Table 7
Variance Decomposition of the Index Function in Model 1:
The Loss in Explanatory Power for each Group of variables

Variables	(1) Promotion probability Explanatory power	(2) Initial Condition Explanatory Power
Human Capital (age, tenure, schooling)	3.5%	35%
Level in the Firm	46.3%	-
Promotion Speed	0.5%	-
Promotion prob. intercepts	14%	-
Promotion prob. slopes	0.5%	-
Initial condition intercept	-	65%

Note: The percentages denote the loss in explanatory power of the explained part of the index function regression for each group of variables. They are computed from the difference in the coefficient of correlation from the regression that includes all factor and a regression that excludes only each particular group. The regression function for the initial condition includes only schooling and age.

Table 8
Models with Interactions:
Summary of the promotion dynamics parameters

variable interacted	(1) education	(2) tenure
	Parameters (standard errors)	Parameters (standard errors)
type probabilities		
Prob <i>type1</i>	0.2034 (0.0037)	0.1956 (0.0047)
Prob <i>type2</i>	0.2532 (0.0055)	0.2743 (0.0054)
Prob <i>type3</i>	0.2412 (0.0067)	0.2498 (0.0037)
Promotion Speed		
$\beta_{0,type1}^S$	0.3853 (0.0056)	0.3196 (0.0070)
$\beta_{0,type2}^S$	0.3415 (0.0004)	0.3330 (0.0040)
$\beta_{0,type3}^S$	0.2558 (0.0040)	0.2551 (0.0004)
$\beta_{0,type4}^S$	0.3354 (0.0033)	0.3363 (0.0006)
<i>education * speed</i>	-0.0234 (0.0002)	-
<i>tenure * speed</i>	-	-0.0093 (0.0003)

Table 9
Model with Interactions
Mean/St-deviations of the promotion dynamics parameters
and the marginal effects

	(1)	(2)
variable interacted	schooling	tenure
Mean (β_0^S)	-0.0582	0.2014
st.deviation (β_0^S)	(0.0645)	(0.1049)
Marginal Effect	-0.0062	0.0221
st-deviation	(0.0078)	(0.0153)

7 Conclusion

In this paper we develop a dynamic model of the promotion process and estimate it using a panel of American executives followed up to eight years. Our model allows promotion probabilities to depend on endogenous past promotion histories, observable human capital endowments, time varying firm specific variables and unobservable individual heterogeneity. The model is the most general and comprehensive empirical model of promotion dynamics estimated to date.

Our results shed light on the complex process that governs hierarchical transitions. The stochastic process that drives promotion might be thought of as a series of probabilities that are smaller for individuals further up in the hierarchy. These probabilities are more dependent upon unobserved individual effects than on observable human capital variables. Among standard human capital variables (age, tenure and schooling), only schooling really matters in predicting current promotion probabilities.

Focusing on the human capital variables, promotion probabilities increase with education and age. Controlling for age and unobserved individual heterogeneity, tenure is virtually irrelevant in predicting promotion outcomes. The effect of human capital variables, in total, on promotion probabilities after controlling for endogenous initial conditions and unobserved individual effects is very weak however. Differences in promotion probabilities are explained, in order of importance, by level, individual specific promotion probability intercept terms, human capital variables and past promotion histories. The explanatory power of these variables differs widely. Age and schooling, while very weak in explaining differences in current promotion probabilities, do explain a good deal of the variation in the historical speed of promotion.

Fast tracks in the theoretical literature have been motivated by differences in ability (Gibbons and Waldman 1999) and by considerations in addition to ability, such as signaling (Bernhardt 1995). We deem fast tracks resulting from difference in ability as “non-structural” or “spurious” and fast tracks resulting from an advantage gained from early promotion as “structural” or “causal,” whether these structural fast tracks result from signaling or from other considerations. Our model is able to distinguish between structural and non-structural fast tracks.

We find that the promotion process is only mildly positively structurally correlated with a measure of the past rate of promotion, which suggests a

weak role for structural fast tracks in the promotion process. We suggest that the fast tracks documented in the empirical literature result largely from unobserved individual heterogeneity and not from rapid early promotions having their own inherent effect on later promotions.

Going beyond this overall result in regards to fast tracks, there is evidence of high cross-sectional dispersion in the effect of past promotion histories on promotion probabilities. Interestingly, we find that the strength of the individual specific effect of achieving a high rate of past promotion on promotion probabilities is inversely related to accumulated human capital. Specifically, structural fast tracks are weaker for individuals with more schooling and tenure. This result may lend some support for a role for signaling. It is consistent with the hypothesis that the signaling aspect of past promotion is stronger for those who are less educated or are relatively new to the firm. Accumulated schooling and tenure may reduce the risk associated with promotion (from the perspective of the firm), and therefore lead firms to rely less on past promotion histories. We also find evidence for a negative correlation between current promotion probability and past promotion rates for a portion of the population and suggest that this result may relate to the well-known “Peter Principle.”

Personnel economics is a growing field and the role of promotion is important within it. Empirical results on promotion in a framework suitably recognizing the endogeneity of past promotion histories, among other things, have been lacking. This paper fills this void and provides many new results on the promotion process that should help to inform the literature. Interesting empirical issues remain. In particular, it would be valuable to investigate the relative importance of human capital and endogenous promotions in explaining lifetime earnings and the nature of serial correlation in wage growth. We believe that our approach is a good starting point in investigating these issues.

References

- [1] Abowd, John, Kramarz, Francis and Margolies, David “High Wage Workers and High-Wage Firms, *Econometrica* 67 (1999),2, 251-333.
- [2] Alvarez, Javier, Browning Martin and Ejrnaes, Mette (2003) “Modeling Income Processes with Lots of Heterogeneity” Working Paper, Copen-

hagen University

- [3] Ariga, Kenn, Ohkusa, Yasushi and Brunello, Giorgio. "Fast Track: Is it in the Genes? The Promotion Policy of a large Japanese Firm," *Journal of Economic Behavior and Organization*, 38 (1999): 385-402.
- [4] Baker, George, Gibbs, Michael and Holmstrom, Bengt. "Hierarchies and Compensation: A Case Study," *European Economic Review*, 37 (1993): 366-378.
- [5] _____. "The Wage Policy of a Firm," *Quarterly Journal of Economics*, CIX (1994 a): 881-919.
- [6] _____. "The Internal Economics of the Firm: Evidence From Personnel Data," *Quarterly Journal of Economics*, CIX (1994 b): 921-955.
- [7] Belzil, Christian "Job Creation and Destruction, Worker Reallocation and Wages, *Journal of Labor Economics*, vol.18 April 2000,183-203 .
- [8] Belzil, Christian "Short Run and Long Run Dynamics in an Empirical model of promotion" IZA Working Paper, 2003.
- [9] Belzil, Christian and Hansen, Jörgen (2002) "Unobserved Ability and the Return to Schooling", *Econometrica*, vol 70, no 6, 575-591.
- [10] Bernhardt, Dan. "Strategic Promotion and Compensation," *The Review of Economic Studies*, Vol. 62, No. 2 (1995): pp 315-339.
- [11] Bognanno, Michael "Corporate Tournaments" *Journal of Labor Economics*, vol 19, No 2 (2001): pp 290-315.
- [12] Cameron, Stephen and Heckman, James (2001) "The Dynamics of Educational Attainments for Black, Hispanic, and White Males" *Journal of Political Economy*, 109 (3), 455-499
- [13] Cameron, Stephen and Heckman, James (1998) "Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males" *Journal of Political Economy*, 106 (2), 262-333.
- [14] Carmichael, Lorne. "Firm-Specific Human Capital and Promotion Ladders," *Bell Journal of Economics*, Vol. 14, No. 1 (1983): 251-258.

- [15] Chamberlain, Gary, "Panel Data " in Handbook of Econometrics, Ed. by Zvi Griliches and M.D. Intrilligator, amsterdam, north holland.
- [16] Chiappori, Pierre-Andre, Salanie, Bernard and Valentin, Julie. "Early Starters Versus Late Beginners," Journal of Political Economy, Vol. 107, No. 4 (1999): pp 731-760.
- [17] Eckstein, Zvi and Kenneth Wolpin (1999) "Why youth Drop Out of High School: The Impact of preferences, Opportunities and Abilities" *Econometrica*, vol. 67, No. 6 (November), 1295-1339.
- [18] Farber, Henry S. and Gibbons, Robert. "Learning and Wage Dynamics," The Quarterly Journal of Economics, Vol. 111, No. 4 (1996): pp 1007-1047.
- [19] Forbes, J. Benjamin and Piercy, James E. (1991). Corporate Mobility and Paths to the Top. New York: Quorum Books.
- [20] Gibbons, Robert and Waldman, Michael. "A Theory of Wage and Promotion Dynamics Inside Firms," The Quarterly Journal of Economics, Vol.114, No. 4 (1999): pp 1321-1358.
- [21] Gibbs, Michael. "Incentive Compensation in a Corporate Hierarchy." Journal of Accounting and Economics, XIX (1995).
- [22] Heckman, James (1981) "Statistical Models for Discrete Panel Data" in Structural Analysis of Discrete Data with Economic Application, edited by Charles Manski and Daniel McFadden, MIT Press.
- [23] Heckman, James and Singer, Burton (1984) "A Method for Minimizing the impact of distributional Assumptions in Econometric models for Duration Data, *Econometrica*, 52 (2): pp 271-320.
- [24] Keane, Michael P. and Wolpin, Kenneth (1997) "The Career Decisions of Young Men" *Journal of Political Economy*, 105 (3), 473-522.
- [25] Howard, A. and Bray, D.W. (1988). Managerial Lives in Transition: Advancing Age and Changing Times. New York: Guilford.
- [26] Lazear, Edward, "Personnel Economics: Past Lessons and Future Directions" Journal of Labor Economics" vol 17, no 2, pp 199-236

- [27] Lazear, Edward P. "The Job as a Concept," in *Performance Measurement, Evaluation, and Incentives*, ed. William J. Bruns, Jr. (Boston: Harvard Business School Press, 1992).
- [28] Lazear, Edward P., and Rosen, Sherwin. "Rank Order Tournaments as an Optimum Labor Contract," *Journal of Political Economy*, 89 (October 1981): 841-864.
- [29] Macleod, Bentley and Macomson, James "Reputation and Hierarchy in Dynamic Models of Employment" *Journal of Political Economy*, 96 (August 1988): 832-854.
- [30] Magnac, Thierry. "Subsidized Training and Youth Employment: Distinguishing Unobserved Heterogeneity from State Dependence in Labour Market Histories" *The Economic Journal* vol 110 (466):pp 805-837.
- [31] Postel-Vinay, Fabien and Robin, Jean-Marc, Equilibrium Wage Dispersion with Heterogeneous Workers and Firms, *Econometrica*, vol 70 (2002), no 6, 2295-2350.
- [32] Prendergast, Canice. "The Role of Promotion in Inducing Specific Human Capital Acquisition." *The Quarterly Journal of Economics*, Vol. 108, No. 2. (May, 1993), pp. 523-534.
- [33] Rosen, Sherwin. "Prizes and Incentives In Elimination Tournaments," *American Economic Review*, 76 (September 1986): 701-716.
- [34] Rosenbaum, James E. "Tournament Mobility: Career Patterns in a Corporation." *Administrative Science Quarterly*, 24 (1979), 220-241.
- [35] Rosenbaum, James E. (1984). *Career Mobility in a Corporate Hierarchy*. Orlando: Academic Press, Inc.
- [36] Seltzer, Andrew and Merrett, David. *Journal of Labor Economics*, Vol. 18, No. 4. (Oct., 2000), pp. 573-613.
- [37] Tuckel, P. and Siegel, K. "The Myth of the Migrant Manager." *Business Horizons*, 26 (1) (1983), 64-70.
- [38] Useem, M. and Karabel, J. "Pathways to Top Corporate Management." *American Sociological Review*, 51 (1986), 184-200.

- [39] Vroom, V. F. and MacCrimmon, K.R. "Toward a Stochastic Model of Managerial Careers." *Administrative Science Quarterly*, 13 (1968), 26-46.
- [40] Wooldridge, Jeffrey (2003) "Simple Solutions to the Initial Conditions Problem in Dynamic, Non-linear Panel Data Models with Unobserved Heterogeneity" forthcoming in *The Journal of Applied Econometrics*

8 Data Appendix

Table A1
Some Summary Statistics

Variables	Mean	Standard Dev.
Level in the firm		
in year 1	4.30	1.27
in year 2	4.29	1.29
in year 3	4.18	1.31
in year 4	4.05	1.29
in year 5	3.97	1.31
in year 6	3.87	1.31
in year 7	3.71	1.27
in year 8	3.52	1.21
Fraction promoted		
in year 2	0.11	-
in year 3	0.09	-
in year 4	0.08	-
in year 5	0.07	-
in year 6	0.09	-
in year 7	0.08	-
in year 8	0.05	-
Firm profits/1,000,000 of 1980\$		
in year 1	130.9	268.1
in year 2	127.4	237.4
in year 3	106.9	231.8
in year 4	106.0	225.5
in year 5	116.5	224.8
in year 6	106.1	209.4
in year 7	120.4	209.7
in year 8	127.9	260.6

Table A1- Continued

Variables	Mean	Std Dev.
Firm size/1000		
in year 1	33.51	40.44
in year 2	33.33	39.58
in year 3	31.73	37.49
in year 4	32.22	38.44
in year 5	32.99	39.96
in year 6	34.07	42.44
in year 7	32.35	40.27
in year 8	32.91	41.35
education (years)	16.37	1.87
age	46.13	8.68
tenure	13.28	10.30
fraction newcomer	0.04	0.19
age above	49.90	4.47
fraction newcomer above	0.21	0.17
promotion per individual (per year)	0.1066	0.2438
Initial Speed of promotion	0.3683	0.2111

Table A2
Promotion Incidence by Level
in the first 2 years of data

Level in the firm	Number of individuals	Fraction promoted	Average tenure	Average age
1	316	-	21.1	55.8
2	1957	0.011	15.0	50.4
3	6480	0.043	12.9	47.4
4	10093	0.081	12.8	45.9
5	8201	0.150	13.1	45.1
6	4361	0.217	13.6	44.6
7	1482	0.266	14.7	44.6
8	369	0.360	14.6	43.6
9	159	0.333	15.1	43.7
10	52	0.385	15.1	41.2
11	12	0.417	18.8	44.1
All	33482	0.117	13.3	46.1