Career Consequences of Hyperbolic Time Preferences

Francesco Drago *
University of Siena

September 2, 2005

Abstract

In this paper I address theoretically and assess empirically the effect of impatience on workers' on-the-job behavior. Theoretically, workers' hyperbolic time preferences and the implied self-control problems explain several empirical regularities concerning job mobility and account for different on-the-job behaviors. On-the-job search on one hand and "collaborative behaviors" such as number of working hours, low absence rate, and effort on the other, are crucial aspects of mobility and individual wage growth. On-the-job search results in higher wages with the new employer while "collaborative behaviors" lead to permanent wage increases with the same employer, mainly through promotion or position change. I provide a model that shows that, for identically productive individuals, heterogeneity in hyperbolic time preferences accounts for different mobility and career patterns. Patient workers undertake behaviors that lead to promotions, they are more likely to be stayers and to follow fast-track-career paths with the same employer. Impatient workers are more likely to be movers and to experience wage increases by switching jobs. Hence, differences between stayers and movers are explained in terms of time preferences. The results of the model mainly rest on the fact that the benefits that result from collaboration are not perceived as immediate as the rewards from job-to-job movements conditional on the arrival of a better job offer. I use a large longitudinal data set (NLSY 79) to test the main conclusion of the model implementing logit, panel and duration estimations. Various measures of impatience are positively correlated to the job arrival rate and negatively correlated to collaboration. Finally, hyperbolic discounting rather than the exponential one is shown to be crucial to explain different career patterns.

KEYWORDS: job search, hyperbolic discounting, self-control problems, personnel economics, job mobility, duration models, longitudinal data.

* E-mail address: drago2@unisi.it. This is a preliminary draft, comments are welcome. I would like to thank Marianna Belloc, Armin Falk, Roberto Galbiati, Botond Koszegi, Nicola Meccheri, Matthew Rabin, and Neri Salvadori for conversations and useful comments and especially Sam Bowles, Dora Kadar, and Pietro Vertova for comments and discussions during all the steps of this project. I would like also to thank the participants at the Zeuthen Workshop on Behavioral Economics in Copenhagen, the third Budapest Workshop on Behavioral Economics and seminars participants at Santa Fe Institute, University of Padua (COFIN 2002), University of Siena, and University of Pisa for comments. Financial support from the Italian Ministry of University and Technological and Scientific Research is gratefully acknowledged (COFIN 2002 Project "Growth and Income Distribution"). The usual disclaimer applies.

1 Introduction

Why do some workers jump from job to job while others strive for promotion? Are there any psychological patterns that distinguish movers from stayers? Do self-control problems affect career paths? Empirical evidence points out that differences between movers and stayers, as well as between winners and losers on the job, especially at the beginning of their career, are only in part explained by "standard" characteristics. By extending the work of DellaVigna and Paserman (2005) to employed workers[1] this paper explores the effects of self-control problems that stem from time inconsistency in order to explain some empirical regularities that concern job mobility and career paths.

Suppose that an employed worker may experience significant permanent wage increases either by changing employer or with the same employer, for example through promotion or position change.[2] The behavior of such a worker affects the arrival rates of these two events (a new job with a higher wage or on-the-job promotion). On the one hand, what I call *collaborative behaviors*, e.g. high effort, good relationships with managers and supervisors, low absence rate, currying favor with evaluator, increase the chance of permanent wage increases (e.g., see Flabbi and Ichino, 2001 and McCue, 1996).[3] On the other hand, on-the-job search increases the arrival rate of a new job offer (e.g., see Bloemen, 2005). In this paper I focus on the general class of time preferences that generate self-control problems - hyperbolic discounting - and I ask whether for identically productive individuals and other things equal, heterogeneity in time preferences accounts for different behaviors, i.e. collaboration and search, and thus for different career paths. To address this point theoretically and empirically the analysis brings together the hyperbolic discounting and the personnel economics literature using some standard findings on promotion, wage and recruitment policies (e.g., see Baker, Gibbs and Holmstrom, 1994; Gibbons and Waldmann, 1999; Lazear and Oyer, 2004; Lazear, 1979; Flabbi and Ichino, 2001).

Evidence on significant job-to-job mobility (Topel and Ward, 1992) suggests that the dilemma of whether to move or to stay is encountered in particular by young workers. The question of whether a behavioral trait such as impatience could contribute to explain different behaviors and labor market histories is important for two reasons. First, empirical evidence has detected that some workers are more likely to move between jobs than others (e.g. Munasinghe and

[1] DellaVigna and Paserman (2005) analyze hyperbolic time preferences of unemployed workers to explain heterogeneity in search effort and reservation wages.

[2] Promotions are shown to be the main channel through which employee experience wage increases (McCue, 1996).

[3] In this paper job is synonymous with employer and firm. Moreover, all the jobs are organized hierarchically in several levels and the wage, besides to be the monetary compensation, is a measure of all the "good characteristics" of a job, e.g. work place, health insurance.

Sigman 2004) but differences between individuals' labor market histories regarding mobility and wage growth have been in large part attributed to unobserved characteristics.[4] Second, in terms of wages and future wage growth the long-run expected gain from permanent wage increase on the same job (resulting from promotion, position change, bonuses, etc.) is generally much greater than the one from changing employer. Hence, it is not clear why a worker should exert a positive level of on-the-job search effort sacrificing effort for collaboration.[5] A crucial aspect of this point is the strong evidence that the arrival rates of wage increases and the promotion rates are positively serially correlated (e.g., see Baker, 1997; Baker, Gibbs and Holmstrom, 1994; Gibbons and Waldmann, 1999).[6]

The theoretical part concerns a worker that has to decide how to climb the wage ladder, i.e. staying or moving. In the basic setting collaboration and search are assumed to be substitutes.[7] If we consider that on-the-job search requires time that could be invested in the task assigned and more generally in collaboration this is not a strong assumption (e.g., interviews with new employers may lead to absenteeism). Moreover, if search activity is observable, the argument that search and collaboration are substitutes is also supported by loyalty concerns since a worker who searches would not be promoted because her loyalty would be in question.

An important aspect of the worker's choice is that benefits that result from collaboration are not perceived as immediate as the rewards from job-to-job movements conditional on the arrival of a better job offer. In order to enjoy benefits that stem from collaborative behaviors such behaviors need to be observed for a considerable amount of time. Moreover, while movers are employed immediately at the new wage once they change jobs, a worker who is promoted receives the wage increases with a certain delay as for example Topel and Ward (1992) noted.[8] Also, the serial correlation in promotion rates implies that some of the benefits of a present

[4]Analysis of company data set shows also that differences between winners and losers respect to promotion rates are hardly explained in terms of observable characteristics (Baker, Gibbs and Holmstrom, 1994).

[5]Evidence is given in the next section. Obviously, this point is critical when firms do make counter-offers to retain workers that find new jobs. The practice of counter-offering to match alternative offers is important from the theoretical point of view (Postel-Vinay and Robin, 2002) but it is very limited in reality. Mortensen (2003) widely discusses this point arguing that it is not reasonable to assume counter offering unless one studies the US academic labor market (see also Moscarini, 2004).

[6]Some studies based on company data sets show also serial correlation for wage increases (Dohmen, 2004; Baker, Gibbs and Holmstrom, 1994). See also the discussion in Munasinghe (2000).

[7]In particular a worker has to choose how to allocate one unit of time between search and collaboration. An extension of the model where this assumption is relaxed is provided.

[8]As they point out: "Inspection of the quarterly data revealed a strong tendency for within-job earnings changes to occur at annual intervals" (p. 456). This timing generally depends on the process of supervisors evaluation. In the company data set of Baker, Gibbs and Holmstrom (1994), the timing of the rating is "probably year-end given that bonuses are awarded on February 1 of the next year". Another strand of literature rationalizes the deferred gain from promotion with incentives concerns. (Leazer, 1979; Dohmen, 2004.)

promotion are materializing in the long-run.

In this setting I restrict my attention to quasi-hyperbolic approximation of hyperbolic time preferences (Phelps and Pollack, 1968; Laibson, 1997; O'Donoghue and Rabin, 1999). Such preferences generate self-control problems and time-inconsistency in a simple fashion. Hyperbolic discounting captures the common tendency "to grab immediate rewards and to avoid immediate costs in a way that our 'long-run selves' do not appreciate" (O'Donoghue and Rabin, 1999). The example provided by O'Donoghue and Rabin (1999) is the one of an individual who between doing seven hours of an unpleasant activity on April 1 versus eight hours of the same activity on April 15, on February 1 prefers the first option. But when April 1 gets closer the same individual can reverse his preferences preferring to do eight hours of the unpleasant activity on April 15 instead of doing seven hours on April 1. This preference reversal is accounted by hyperbolic discounting but it is not compatible with the standard exponential discounting setting.

Hyperbolic discounting is important in this framework because it makes on-the-job search attractive and explains why workers search despite the fact that they would gain more by staying. The basic intuition builds on the two facts mentioned above. First, rewards (in terms of wage increase) that result from collaboration are not perceived as immediate as the rewards from job-to-job movements conditional on having received a better job offer. Second, since search implies to forgo future wage increases resulting from collaboration, it is an activity with a short-term reward (net of the search cost) and a long-run opportunity cost (renouncing large wage increases on the same job). Hence, if individuals cannot borrow from future wage increases, higher degrees of short-run impatience make immediate wage increases more valuable relative to the career opportunities within the job, the benefits of which are delayed. In other words "the short-run selves" of the workers tend to put more weight on the short-term reward from search and to put less weight on the implied long-run opportunity cost. This induces more impatient workers to search more and to switch job, thus differences between stayers and movers are explained in terms of time preferences.

Although heterogeneity in the degrees of impatience within the exponential model in principle could explain variation in the behaviors mentioned above, there are good reasons to focus on the hyperbolic model. Following O'Donoghue and Rabin (1999) I distinguish between sophisticated and naive agents, where the firsts are aware of their self-control problems and the seconds are not. For naives the "irresistible desire" of a job change is partly driven by the wrong belief that once they have found a job they will behave as exponential and thus will not search or change again. The self-control problems for sophisticated are even worse as they lead them to search more than the naives do (this result is in line with the "sophistication effect" found by O'Donoghue and Rabin (1999) in the case of immediate rewards and delayed costs). In other words, any departure from the naivete assumption implies a multiplicative effect on the propensity to change job. This is a testable feature of the hyperbolic model and it shows the potentially large behavioral effect of time-inconsistency in time preferences: variation in the discount factor within the exponential model can not replicate variation in behavior implied by heterogeneity in the short-run discount factor within the hyperbolic model.[9]

In the empirical part of the paper, I provide empirical evidence in support of the hyperbolic model vs the exponential one and I test the main predictions of the model using the large data set of the National Survey Longitudinal Data of Youth (NLSY 79). This dataset allows to control for a large number of individual variables as well as for some characteristics of the occupation and the firm. To proxy for impatience I closely follow the most recent applied works on this issue - in particular DellaVigna and Paserman (2005) who use the same dataset - and the psychological evidence on impatience. Smoking, drinking alcohol, not having a bank account and a life insurance are some of the variables that proxy for impatience. Note that one of the conclusion of DellaVigna and Paserman (2005) is that heterogeneity in these proxies captures variation in short-run impatience. When it is possible I exploit the panel structure of the data to control for permanent unobserved heterogeneity. I identify a relation between one of the main collaborative behaviors, namely low absence rate, and impatience. Impatience affects negatively the decision to collaborate where the latter is found to be an important channel through which workers experience promotion (Flabbi and Ichino, 2001, see also Audas, Barmby and Treble, 2004). I show that the job arrival rate is positively affected by various measures of impatience. In order to identify voluntarily job changes I include in the analysis only those job-to-job transitions associated to wage increases that are preceded by job terminations for reasons other than layoff, firing or end of temporary jobs. Since impatience could capture some types of unobserved productive skills, I run regressions including many variables presumably correlated with ability such as father and mother education, family background, AFQT scores and years of schooling. Moreover, I control for the type of occupation, the type of industry, the wage, the size of the firm. The impact of impatience is significant and strong even having controlled for these variables, suggesting that that empirics seems to support the theory developed here.

The main contribution of this paper is twofold. For the labor turnover and personnel economics literature I sign theoretically and I test empirically the effect of impatience on workers' on-the-job behaviors finding that it is very important for the analysis of career paths. In this respect I start from the work of DellaVigna and Paserman (2005) who analyze unemployed workers'

[9] Also, it is worth noting that estimated long-run discount factors vary much less than the corresponding short-run factors across individuals (see Paserman, 2004, Fang and Silverman, 2004).

behavior concerning search effort.[10] I extend their analysis to a situation with multiple choices facing employed workers that is more similar to the short-term reward and long-run cost case. The second contribution is related to this point. For the hyperbolic discounting literature I find an application with immediate costs and short-term and delayed rewards where the theoretical results are confirmed by the empirical analysis.

In the next section I summarize the empirical evidence on which the model is built. Next, I develop a simple model that captures the effect of impatience on workers' behaviors. Then, I present the empirical part of this study. Finally I draw conclusions.

2 Empirical evidence

2.1 Background

The decision to promote or not a worker is mainly driven by performance evaluation and eventually by worker's behaviors, those termed as collaborative. Here the the focus is on the casual link between collaborative behaviors and the arrival of large and permanent wage increases, that in general occur through promotion.[11] For the empirical evidence I consider promotions that are extensively studied in the literature. However, looking only at the wage gains after promotion can be misleading for assessing the extent of the rewards from collaboration. As mentioned before, promotion at present predicts future promotions and the collaborative behaviors that lead to promotion result also in substantial and systematic annual bonuses (e.g. see Gibbs, 1995). Hence, although the discussion that follows is mainly on promotions, it is important to note that the "true" benefit from collaboration is greater than the one from promotion only.

The amount of search effort and the decision to move should be driven not only by the gain in terms of wage but also on the eventual gain in terms of future wage growth on the new job. For instance, a worker can move to a new job that offers a lower wage but greater opportunities of future wage growth. In the theoretical part I posit that the wage is a sufficient statistics for the value of the job by assuming that there is no heterogeneity in wage growth rates offered by employers. This assumption is supported by Topel and Ward (1992) and more recently

[10]Related works are Munasinghe and Sicherman (2005a) and (2005b) who show that heterogeneity in time preferences accounts in part for differences in wage levels and wage growth rates. Their approach however assumes that time preferences may capture some form of unobserved investment in human capital and it is based on the exponential discounting setting.

[11]I do not consider automatic wage increases deriving from seniority that arrives independently from performance evaluations. Such annual return, when it exists, it is clearly much smaller than the return from a promotion or position change. Considering this aspect of seniority would complicate the analysis without changing the results.

by Postel-Vinay, Jolivet and Robin (2004) among others. I later relax this assumption and I show how the main proposition of the model is robust to different specifications about the heterogeneity of wage growth rates.

2.2 Average Gains

Wage growth within and between jobs has been extensively analyzed empirically especially among young workers. The most prominent work on this issue is Topel and Ward (1992). In this paper the authors find that job mobility is very frequent for young individuals. A job change with a different employer is usually associated to a twelve percent of increase in the wage rate but this increases is lower for workers with more experience. However, for the wage growth within the jobs they do not control for promotion and collaborative behaviors. Topel and Ward (1992) find that the average rate of annual wage growth for new entrants into the labor force was about the fourteen percent (p. 456). Considering this estimation as the wage increase deriving from promotion[12], a new entrant in the labor force should expect a wage gain within the job greater than a wage gain from a job-to-job movement. McCue (1996) takes explicitly into account promotions and finds that the magnitude of wage growth due to job-to-job mobility is about the half of the one due to internal promotions. Similar estimates are also found by Pergamit and Veum (1999).[13] Other studies from which it is possible to infer similar evidence are Gerarth and Milkovich (1989) and Lazear (1992). Hence, I shall assume that average wage gain from moving to a new job is lower than the average wage gain deriving from a promotion.

2.3 Arrival Rates of Promotion and Job Change

As many have documented, job arrival rates for employed workers are mainly determined by the search intensity supplied (e.g. Bloemen, 2005). On the other hand, the arrival of a promotion is driven by what I have termed as collaborative behaviors. The emerging literature on personnel economics and on internal labor markets sheds light on this issue and provides some evidence. The arrival of promotions are positively correlated to the absence of misconduct behaviors and absenteeism (Flabbi and Ichino, 2001; Audas, Barmby and Treble, 2004). The absence rate is found to proxy (inversely) the effort exerted by the worker to perform the task assigned. Low absence rate is generally associated to a high performances evaluation that is a strong predictor for a promotion (Audas, Barmby and Treble, 2004). Besides these, there are number of collaborative behaviors. Working hard in terms of hours, e.g. more than 45 hours per week

[12]Since they do not control for the event of promotion and for some collaborative behavior, it is without saying that such gains must be considered very conservative estimations for wage increases deriving from promotions.

[13]Similar findings are documented also for Europe.

in McCue (1996), increases substantially the likelihood to be promoted. Also, when effort is not observable and performance evaluation are subjective, the worker may spend time to curry favor with his supervisor to get a good rating (Prendergast, 1999).

To sketch the trade-off the worker faces, let $s \in [0,1]$ be the search effort and suppose that on-the-job search and collaborative behaviors have either the same cost functions or that are both costless. One of the assumption of the basic model is that on-the-job search and collaboration are substitutes. In this framework $1-s$ measures the intensity of collaborative behaviors. The majority of the studies on on-the-job arrival rates do not control for individuals' search effort that is usually not observable. Hence, estimations of these arrival rates reflect optimal choices of (especially young) workers.[14] Empirical evidence suggests that for a given wage and for large firms that offer the possibility to be promoted, the probability of a promotion for a worker with the highest performance evaluation and with zero absence rate, that in our case corresponds to $s = 0$, is higher than the probability to find a new job when search effort is high, say $s = 1$. Also, the marginal effect of an increase in any of the collaborative behaviors on the probability to be promoted is higher than the marginal effect of search effort on the job arrival rate.[15]

Suppose that $\gamma(1-s)$, is the probability to be promoted and $\lambda\overline{F}(w)s$ is the probability to find a job that pays a higher wage than the current one, w (that is a sufficient statistic for the value of the job), where λ is a constant, $\overline{F}(w') = 1 - F(w')$ and $F(w')$ is the c.d.f. the worker faces in the search process (details are given in subsection 3.3). Moreover let assume that all the jobs offer the same wage growth opportunities and that all the jobs are identical. A rough calibration would require to set $\gamma \cong \lambda$. Let denote with A and B the gain at job change and the gain from promotion, respectively, in terms of lifetime utilities. If for the wage associated to A and B we have $w_A < w_B$, for the gain in lifetime utilities we have also $A < B$ The expected gains are $\lambda\overline{F}(w)sA$ and $\gamma(1-s)B$ and it is immediate to see that it is not clear why a worker should exert a positive level of search effort.

Note that this point may arise even if $\gamma > \lambda\overline{F}(w)$ were not satisfied. When the arrival rate of a promotion depends on the intensity of collaboration, the existence of the serial correlation in

[14] Topel and Ward (1992) find that annualized arrival rate to find a new job for young workers is very high at the beginning of their career and it stabilizes, after a couples of years of tenure, on the value of 0.2. Jolivet, Postel-Vinay and Robin (2004) estimate a job arrival rate between 0.09 and 0.22. The same holds for the arrival rate of a promotion, being higher at the firsts years of the career. According to McCue (1996), the annualized baseline arrival rate of a position change that brings about a wage increase is about 0.15.

[15] For the former marginal effect I take into account McCue (1996) that finds that working more than 45 hours brings about a 20 percent increase in the arrival rate of a promotion (see also Flabbi and Ichino, 2001 for absenteeism and misconduct episodes). For the latter marginal effect I consider the one measured by DellaVigna and Paserman (2005) that find that an increase in the search effort for unemployed workers, measured by the number of search methods, leads to a 15 percent increase in the job arrival rate for unemployed workers (usually unemployed workers receive more job offers than employed workers).

the promotion rates (Baker, 1997; Baker et al. 1994; Gibbons and Waldman 1999; Rosenbaum, 1994; Bruderl, Diekmann, and Preisendorfer, 1991; Podonly and Baron, 1997) increases the value of staying and the long-run average gain from collaboration.

Therefore, in this setting there exists an opportunity cost associated to on-the-job search. The idea that switching job for a higher wage is not so profitable as staying is very old in economics. "The wastefulness of this 'try and try again' process of advancing to a better position is self-evident" (Slichter, 1919, p.216) and it was recognized by many economists.[16] Recent progress in personnel economics such as the findings on the persistence of wage increases strengthens this idea.

3 The Model

3.1 Time preferences

I adopt the structure of time preferences of Laibson (1997) as a simplification of hyperbolic time preferences. Consider the stream of per period utility deriving from consumption as $u(c_t)$. The present value of a flow of future utilities at time t is $u(c_t) + \beta \sum_{s=t+1}^{T} \delta^s u(c_s)$, where δ, is the discount factor that captures long-run patience and $\beta \in (0, 1]$ is a parameter that reflects the short-run patience. When $\beta = 1$ we turn back to the exponential discounting standard case. When $\beta \in (0, 1)$, time preferences are quasi-hyperbolic as the discount function approximates an hyperbole. As a consequence of quasi-hyperbolic discounting, the discount factor between the current period and the next one is $\beta\delta$, while the discount factor between any two periods other than the current and the next ones is δ. These preferences capture and stylize a simple fact documented by dozen of experiments (Ainslie, 1992; Kirby and Herrnstein, 1995; see also McClure, Laibson, Loewenstein and Cohen, 2004). Individuals do not behave according to the exponential discounting model where the rate of time preferences and the implied discount factor is constant. Individuals may suffer self-control problems that result from time inconsistency. Let suppose $\beta = 0.9$ and $\delta = 0.99$. In period t a reward of 10 that starts materializing in period $t + 3$ is preferred to a reward of 9 that starts at period $t + 2$ ($\beta\delta^3 10 > \beta\delta^2 9$). In period $t + 2$, this preference is reversed by the same individual as $9 > \beta\delta 10$. It is immediate to see that these type of preferences involve self-control problems.[17]

[16] For example Mincer and Jovanovic (1981) recognize that frequent job changing may result in flat wage profiles. The sentence quoted is taken from Topel and Ward (1992) and proceeds: "The worker does not know in detail the nature of the job which he is obtaining nor does he know his own capacities. Nevertheless it is the principal method by which workers at the present time improve their working conditions".

[17] Many empirical regularities on consumption (Angeletos, Laibson, Repetto, Tobacman, and 2001; Laibson, 1997), drug addiction (Gruber and Koszegi, 2001), saving plans (Madrian and Shea, 2001), unemployed worker behavior (Paserman, 2004) among others, are explained better in terms of quasi-hyperbolic preferences rather

Following the tradition of the hyperbolic discounting literature we shall consider the different selves of the individual that in each period must take some decisions (Laibson, 1997; O'Donoghue and Rabin, 1999; Diamond and Koszegi, 2003). In particular in each period a self of the individual takes a decision over the control variable, e.g. self $t+1$, self $t+2$ at periods $t+1$ and $t+2$, respectively, and keeps the control over her period. In subsection 3.3 I distinguish between naive and sophisticated agents. The firsts are those who are not aware of their self-control problems, i.e. each self believes that future selves will have exponential time preferences. In other words the naives are overconfident in the ability to solve their self-control problems in that they think they will face in the next periods the parameter $\beta = 1$. The seconds are aware of their self-control problems, i.e. each self knows how future selves will discount events and behaves accordingly.

3.2 Timing of the events

For the core of the results, it is necessary to recall that wage gains from promotion are not perceived as immediate as wage gains at job changes because wage increases from promotion are deferred, for example paid at the end of the year. This is a very general feature of the compensation schemes adopted by firms in case of wage increases, mainly through promotion (see also the discussion in the introduction and footnote 8). In order to capture this fact, in the rest of the paper, I will assume that at time t, conditional on the arrival of a job offer, effort s may lead to an immediate job change, while the collaborative behaviors implied by effort s may lead to a promotion the benefits of which start materializing at time $t+1$.

3.3 Three period model

The analysis concerns identical individuals possessing the same productive skills, i.e. those skills that enter the production function and increase the firm's profit, because I want to address the question of whether heterogeneity in time preferences affects workers' behavior and labor market histories for reasons other than those related to investments in specific or general human capital.[18]

I develop the simplest model that shows the core of the results: a three period time model where the first and the second period the worker chooses the intensity of her behavior concerning collaboration and search. The periods are labeled with t= -1,0,1. An employed worker searches for a job by drawing a sequential random wage sample from a stationary c.d.f $F(w')$, continuous on the support $[\underline{w'}, \overline{w'}]$, with the first moment, $F'(w')$, positive. I denote $1-F(w')$ with $\overline{F}(w')$,

than time-consistent preferences.

[18]Indeed, it is well known that patience may capture some form of unobserved ability since more patient individuals are more willing to invest in their productivity (e.g, recently see Munasinghe and Sicherman, 2005).

where, clearly, $\overline{F}(w')$ is the probability that a wage offer is at least a greater as w', and $\overline{F}(\underline{w'}) = 1$ and $\overline{F}(\overline{w'}) = 0$. Without loss of generality I take the promotion as the only way through which workers experience wage increases. For each job, the worker who is employed at the wage w and is promoted experience a wage increase $w^p - w$ according to a deterministic function $\Phi(w) = w^p$, where the generic wage w with the superscript p denotes the wage after promotion for a worker that before promotion earned w and $\Phi(w)$, with $\Phi'(w) > 0$, maps continuously the wage w into $[w, \overline{w}]$. The wage w^p is the same across all the jobs, that is all the jobs offer the same opportunities in terms of wage growth. The only restriction on $\Phi(w)$ is stated in Assumption 1 (see below).

The worker must choose the search effort $s \in [0, 1]$ that determines how often the distribution of wage offer is sampled. Let $E(w' \,|w' > w)$ the expectation of w' conditional on $w' > w$. Based on the discussion in the previous sections the main assumptions follow:

Assumption 1 *For any wage rate w, $F(w')$ and Φ are such that $E(w' \,|w' > w) < w^p$.*

Assumption 2 *An increase in worker's search effort reduces the intensity of collaborative behaviors.*

Assumption 3 *The worker employed at wage w that experiences a promotion in period t is paid according to the wage w^p in period $t + 1$. The worker that moves to a new job at period t is paid the new wage in the same period.*

Assumption 4 *Search is more costly than collaborative behaviors. The relative cost of search respect to collaborative behaviors, $c(s)$, is increasing and convex in s.*

The first three assumptions accounts for the discussion of the previous section.[19] Assumption 2 implies that the effort for collaborative behaviors is $1 - s$, i.e. the worker allocates one unit of time between search and collaboration. In sub-section 3.5. the results of the model are shown to hold in the case the worker is allowed neither to search nor to collaborate and to enjoy a benefit from this type of activity. Assumption 4 is nor necessary neither sufficient for the results of the model but it is seems reasonable. Unlike search there is no evidence that all the collaborative behaviors are costly. For example it is difficult to believe that the absence of absenteeism episodes has a positive cost other than the opportunity cost of doing something else during the absence period.[20] Assumption 4 states that for example an additional paid hour of work is less costly than an hour of on-the-job search. The same results hold in a more general setting where collaborative behaviors are assumed to be costly as search, more

[19] For the Assumption 1 the majority of the studies find that w^p is about two times $E(w' \,|w' > w)$.

[20] Indeed, a reasonable interpretation is that collaboration is a by-product activity of non-search and to this extent is costless.

costly than search (although this case does not really seem realistic), and costless (additional Appendix upon request).

Denote with λs the probability to find a job, where λ and $s \in [0,1]$ are the search effort and the constant search efficiency parameter, respectively. Note that s determines just how often the c.d.f $F(w')$ is sampled and that the arrival rate of an job offer with a higher wage depends on the current wage, w, and is equal to $\lambda\overline{F}(w)s$. Based on the discussion of the previous section, for the sake of simplicity I posit $\lambda(1-s)$ to be the probability that the worker experiences a promotion without considering serial correlation in promotion rates.

Assuming that the joint probability of a promotion and of a job offer is negligible in the interval of time considered, the worker in period -1 (self -1), employed at the wage w, chooses the search effort to solve:

$$\max_{s\in[0,1]} w - c(s) + \lambda s\beta\delta[max\left\langle \int [V_0(w') + (1/\beta\delta)w' - V_0(w) - (1/\beta\delta)w]dF(w'), 0\right\rangle] + \\ +\lambda(1-s)\beta\delta V_0(w^p) + (1-\lambda)\beta\delta V_0(w), \qquad (1)$$

where $V_0(w)$ is the lifetime utility of employment at wage w at period 0 from the perspective of self -1.

Lemma 1 *The objective function (1) is an increasing function of the current wage, w.*

This lemma derives from the assumption that all the jobs offer the same opportunities in terms of wage increases. Therefore an employed worker move to another job if and only if it pays a higher wage. The objective function of the worker may be written as follows:

$$\max_{s\in[0,1]} w - c(s) + \lambda s \int_w^{\overline{w}} (w' - w)dF(w') + \\ +\beta\delta\left\{\lambda s \int_w^{\overline{w}} [V_0(w')]dF(w') + \lambda(1-s)[V_0(w^p)] + [1 - \lambda(1 - sF(w))]V_0(w)\right\}, \qquad (2)$$

(all the proof are in the Appendix). The first line of eq. (2) is the immediate payoffs in the current period. Namely the wage rate minus the cost associated to the search effort s, $c(s)$, plus the immediate gain from finding a new job. The second line is the continuation payoff in the next period that is discounted by the factor $\beta\delta$ and it is composed of three terms. With probability λs, if the offer is acceptable, the worker experiences the continuation payoff from moving to a new job, $\int_w^{\overline{w}}[V_0(w')]dF(w')$. With probability $\lambda(1-s)$ the worker experiences a promotion and the utility from period 0 onwards is $V_0(w^p)$. Finally, with probability $1 - \lambda(1 - sF(w))$ the worker remains employed at the same job, without experiencing a promotion, with continuation payoff $V_0(w)$. Let w^{pp} be the wage of a worker who experiences the second promotion and w'' the realization of a new wage offer when the worker is employed at the wage rate w. From the perspective of self -1, the continuation payoffs from remaining at the same job without a promotion, from changing job, and from remaining experiencing a promotion are, respectively:

$$V_0(w) = w - c(s) + \lambda s \int_{w}^{\overline{w}} (w' - w) dF(w') +$$
$$\delta \left\{ \lambda s \int_{w}^{\overline{w}} w' dF(w') + \lambda(1-s) w^p + [1 - \lambda(1 - sF(w))] w \right\}, \tag{3}$$

$$V_0(w'') = w'' - c(s) + \lambda s \int_{w''}^{\overline{w}} (w' - w'') dF(w') +$$
$$\delta \left\{ \lambda s \int_{w''}^{\overline{w}} w' dF(w') + \lambda(1-s) w''^p + [1 - \lambda(1 - sF(w''))] w'' \right\}, \tag{4}$$

$$V_0(w^p) = w^p - c(s) + \lambda s \int_{w^p}^{\overline{w}} (w' - w^p) dF(w') +$$
$$\delta \left\{ \lambda s \int_{w^p}^{\overline{w}} w^p dF(w') + \lambda(1-s) w^{pp} + [1 - \lambda(1 - sF(w^p))] w^p \right\}, \tag{5}$$

In these three expressions, the continuation payoffs at period 1, i.e. those multiplied by δ, consist of the wages only because period 1 is the terminal period. Moreover, in (3), (4) and (5) the short-term discount rate β is not present because these are the continuation payoffs from the perspective of self -1 once the drop in the discount rate has already occurred. Note that this structure of time preferences implies that self 0 at period 0 will solve her maximization problem applying a discount factor equal to $\beta\delta$.

I start by signing the effect of impatience on search effort for self -1 by deriving the first order condition for self -1 from equation (2) abstracting for the moment from sophistication and naivete (the analysis of the solution of the game implied by hyperbolic discounting is offered in the next subsection):

$$c'(s) = \lambda \int_{w}^{\overline{w}} (w' - w) dF(w') + \beta\delta\lambda \left\{ \int_{w}^{\overline{w}} [V_0(w') - V_0(w)] dF(w') + [V_0(w) - V(w^p)] \right\} \tag{6}$$

Technical assumptions that guarantee an interior solution are provided in the Appendix A. It is immediate to see that in expression (6) the part multiplied by $\beta\delta$ is negative. Despite this, the first order condition may deliver a positive level of search effort. Moreover, from (6) the following Lemma is derived.

Lemma 2 *The marginal benefit from search is decreasing in w.*

Assumption 5 *Assume that there exists critical wage w^R such that the gain from search is zero, i.e. the RHS of equation (6) is zero.*

This assumption implies that the employed job seekers are only a share of the total labor force, and that this portion is mainly concentrated in the left tail of the wage distribution. This is something on which there is strong evidence. Moreover, note that the literature on hyperbolic discounting finds that the values of β across individuals lies approximately in the interval $[2/5, 1]$ while for δ the range is $[0.95, 9/10]$ (Fang and Silverman, 2002 and Paserman, 2004)[21]. An important point is that δ cannot vary as β across individuals.

Lemma 3 *The optimal level of search effort is positive and strictly decreasing with respect to w for any $w < w^R$. For $w \geq w^R$ the optimal level of search effort is equal to zero. The critical wage w^R is decreasing with respect to β and δ.*

Now it possible to construct a very simple exercise of comparative statics by considering two workers.

Proposition 1 *Let worker 1 and worker 2 have the same δ and different β, with $\beta_1 < \beta_2 = 1$ and let denote by $w^R(1)$ and by $w^R(2)$ the critical wages of worker 1 and 2, respectively. Then, for any $w < w^R(2)$ the search effort of worker 1 is higher than the search effort of worker 2.*

Therefore, for identical discount factors δ, an hyperbolic discounter searches more than an exponential one. Proposition 1 can be easily generalized by saying that more impatience workers in general search more and are more likely to be movers (see sub-section 3.5).

For this and the next results equation (6) is fundamental. In particular for Proposition 1 to hold it is necessary that the term multiplied by the $\beta\delta$ is negative. I argued that indeed in general this is the case as long as it represents the opportunity cost of moving. It is intuitive that considering serial correlation in promotion rates and in wage increases reinforces the results conveyed by Proposition 1.

3.4 Equilibrium for sophisticated

In this subsection I distinguish between sophisticated and naive agents. As it was shown for several aspects of decision making, differences between naives and sophisticated have different behavioral implications. I start from the sophisticated that are those who are aware of their self-control problems, i.e. they predict how future selves behavior. Sophistication involves a conflict between self -1 and self 0 as the current self has rational expectations on the behavior

[21] Paserman (2004) finds similar values from the structural estimation of the model of Della Vigna and Paserman (2005). The dataset used is the same adopted here (NLSY 79).

of future selves. To see the conflict note that from the perspective of self -1, the evaluation at period 0 of the payoff accruing at period 1 is different than the one that is made by self 0. Self -1 and self 0 may disagree on the decision to be taken because self 0 cares about self 1 relatively less than self -1 does. As originally Strotz (1956) pointed out, an individual affected by time inconsistent preferences and aware of her self-control problems can attenuate and eventually solves her problem in two ways. The first possibility solves the self-control problem and consists of precommiting future selves behavior. For example, self -1 could commit self 0 to behave according to her time preferences. This is possible when there exists a commitment device that forces future selves to follow some behavior. As a second possibility, when there is no commitment device, the individual attenuates her self-control problem by finding a "consistent planning" where in every period the choice of the current self anticipates, and it is optimal respect to, the choices of future selves. In this case the different selves of the individual are involved in a strategic environment and the solution is therefore game theoretic in that the current self plays a game against future selves. More precisely, this is called an intrapersonal game that is basically a stochastic sequential game with perfect information.

In order to sketch how to find the solution to such a game, let solve the model by backward induction in the case self 0 earns wage w. For example, this case can be the one in which self 0 has the same job of self -1, i.e. the worker at period -1 has not experienced either a promotion or a job change. I start by analyzing the choice of self 0 as for self 1 the there is no choice. The first order condition for self 0 at period 0 is the following:

$$c'(s) = \lambda \int_{w}^{\overline{w}} (w' - w) dF(w') + \lambda\beta\delta[\int_{w}^{\overline{w}} (w' - w) dF(w') + (w - w^p)] \tag{7}$$

From the perspective of self -1, however, the f.o.c at period 0 is:

$$c'(s) = \lambda \int_{w}^{\overline{w}} (w' - w) dF(w') + \lambda\delta[\int_{w}^{\overline{w}} (w' - w) dF(w') + (w - w^p)] \tag{8}$$

From these expressions it easy to see that if w is such that the RHS of (8) is positive, self -1 disagrees on the optimal level of search effort that self 0 will supply. Note that this argument applies also to a N finite period model where the conflict exists between the selves i and $i+1$ of the individual. Let define by $w^R_{t,i}$ the wage rate from the perspective of self i according to which self t at period t should not search. Hence, for the three period model, $w^R_{0,0}$ and $w^R_{0,-1}$ are the wages according to which the RHS of eq. (7) and (8) are equal to zero, respectively. We have that $w^R_{0,0} > w^R_{0,-1}$ and, as a consequence of Lemma 2, for any wage rate $w \in [w^R_{0,-1}, w^R_{0,0}]$ self -1 wants self 0 not to search. For any wage rate $w < w^R_{0,-1}$ self -1 thinks that self 0 will search too much. For any wage rate $w > w^R_{0,0}$ self -1 and self 0 both agree on not searching.

Analytically, the optimal behavior of self -1 is found by substituting in the continuation payoffs

$V_0(w), V_0(w')$ and $V_0(w^p)$ of eq. (6) the search effort of self 0 implied by the self 0's f.o.c. (7). The equilibrium concept is a Nash equilibrium in Markov strategies that for each self (player) and for each wage (state) specifies implicitly the optimal search effort. Let define the optimal search effort policy function for self i as $s_i^* : [\underline{w}, \overline{w}] \to [0,1]$.

Proposition 2 *A unique Markov-perfect subgame-perfect equilibrium of the above game exists for each self.*

Clearly in the three period model this proposition applies to self -1. However this proposition can be generalized to a N finite period model (see sub-section 3.5). The existence of an equilibrium for the three period model is immediate to see and it is based on standard arguments. Now let consider the possibility of commitment. If self -1 in period -1 were employed at a wage that lies in the interval $[w^R_{0,-1}, w^R_{0,0}]$ she would commit self 0 not to supply a positive level of search effort, whatever the worker experiences in period 0. The following proposition states that in general the sophisticated worker is willing to pay a positive price for a commitment device that attenuates her self-control problem.

Proposition 3 *For self -1 there exists an $\alpha > 0$ such that for $w < w^R_{0,0}$ a decrease in the search effort of self 0 from s_0^* to $s_0^* - \alpha$ increases the net present utility of the welfare of the selves of the individual evaluated in period -1.*

Intuitively, such a commitment can be a contract signed by self -1 that commits the worker to pay a penalty for a job change of an amount such that the net gain from search induces the next self to exert the search effort self -1 would like self 0 exerts.

3.5 Exponentials, naives, and sophisticated

The definition of naivete implies that agents believe that their continuation payoffs are the same of an exponential agent, that is to say they believe that future selves will behave according to a $\beta = 1$. In other words, they are overconfident that future selves will not have the self-control problem that they are currently facing. Formally, taking into account the continuation payoffs (3), (4) and (5) of the three period model, the naive at period -1 believes that the search effort in those payoff is the one supplied by an exponential agent. Such agents are not consistent in their plans as they revise continuously their optimal choices. Hence, there is no intrapersonal game among the selves. For our simple setting the solution is trivial as we need to take into account how self -1 chooses the search effort when he believe that self 0 will choose as an exponential agent. When the selves are more than 3, the solution to the problem of the worker in each period is not found by solving a standard dynamic programming problem. The solution is found by considering the optimal choice of the single self i at time i, for all i, facing a per period discount factor equal to $\beta\delta$. Therefore, by strict convexity of the cost function, continuity of the c.d.f. of wage offers, standard arguments lead to the existence and uniqueness of the equilibrium. The same argument for the existence and the uniqueness applies to the case of the exponential agent. In such a case, however, the derivation of the optimal search effort policy function is made from the first period onwards.

It is interesting to derive the behavioral implications of naivete and comparing them with those of sophistication and exponential. Some algebra leads to the following Proposition.

Proposition 4 *The self -1 of a sophisticated worker supplies a s^* that is higher than the one supplied by the self -1 of a naive worker. For the self -1 of a sophisticated the critical wage w^R is lower that the one of a naive.*

As the naive, the sophisticated has a present bias in time preferences that induce her to exert a high level of effort. However, exactly because the sophisticated predicts correctly future selves behaviors exerts a higher effort than the one she would exert if she did not predict correctly what is the discount rates of the next self. The intuition is that the sophisticated is embedded with a sort of "pessimism" for future selves behavior that induces her to behave in this way. This is similar to the "sophistication effect" found by O'Donoghue and Rabin (1999) in the case where the sophisticate individual faces immediate rewards.

Let define $\hat{\beta}$ the worker's belief about the degree of self-control he will have in the future, β. While for naives $\hat{\beta} = 1$ and for sophisticates $\hat{\beta} = \beta$, *partially naives* are those for which $\hat{\beta} \in (\beta, 1)$ (O'Donoghue and Rabin, 2001). The following Proposition generalizes the result of Proposition 4.

Proposition 5 *For hyperbolic discounters, the optimal search effort s^* of self -1 exerts increases with the degree of sophistication, i.e. it is decreasing in the interval $[\hat{\beta}, 1]$. The critical wage w^R is increasing with the degree of sophistication.*

A major implication of Proposition 5 is that for given β any departure from the naivete assumption monotonically increases the job arrival rate of the worker, because of a higher search effort and a lower critical wage, w^R. From Proposition 4 and 5 it follows that a variation in β across individuals implies a variation in s^* that is greater than the one implied by the same variation in δ. This is one of the most important feature that distinguishes the hyperbolic model from the exponential one.

3.6 Robustness and Extensions

In this section I show the robustness of the the model to different specifications. Is it possible to generalize the model to a finite $N = i, i+1, ..., N-1$ period model? The answer depends on whether Proposition 2 applies also to the N finite period model. Concerning the existence, the extension to the N period model do not cause difficulties. Such difficulties would arise if for some self there were two optimal choices. In this case there would be a jump from a maximum to another one and the earlier self's maximization problem might not have solution.[22]. Such an equilibrium in a finite period model is generally unique as at the terminal history there is no self that is indifferent between two different states. Given that Proposition 2 holds for the N finite period model, also do Proposition 3, 4 and 5. Indeed, the conflict between self $-i$ and $-i+1$ is the same between self -1 and self 0 in the three period model.

A major concern is that in the previous section the result may appear to depend crucially on the fact that the worker is not allowed to choose to do neither search nor collaborative behaviors. This case is analyzed assuming also that "doing nothing" has a positive and immediate value for the worker. Note this is equivalent to relax Assumption 2 as the worker is allowed to contemporaneously decrease the search effort and the intensity of collaborative behaviors. In this case an additional condition that guarantee the same results of the model presented above is required (additional Appendix upon request). It also reasonable to assume that low collaboration implies an endogenous probability of separation for reasons other than from those that lead to voluntarily job-to-job movements, e.g. firing. It is possible to show how the results extends to this framework. In this case more impatient workers are the movers that switch job and get separated.

Finally, although there is evidence for absence of important variation in wage growth rates across jobs, it is useful to check the robustness of the model considering the heterogeneity in the wage profile offered by employers. The formalization of this kind of heterogeneity would complicate the model and the reservation wage (i.e. the wage according to which the worker is indifferent between moving or staying) would not be longer the current wage. Suppose that the wage within the job evolves as $w^p = w + \epsilon(\overline{w} - w)$ where $0 < \epsilon < 1$. Heterogeneity of wage growth rates could be introduced by assuming a continuous distribution for ϵ. By looking at condition (6) it is possible to understand whether the results conveyed by Proposition 1 change in such enriched framework. Consider four possibilities. First, a worker that is employed at a low wage and high wage growth (high ϵ). In this case the long-run benefit from promotion is still greater than the long run benefit from changing job and in eq. (6), i.e. the term multiplied by $\beta\delta$ would be still negative and the results of Proposition 1 hold. The same holds in the second case - a worker that is employed in a job at a high wage and high wage growth - as long as the wage is not the one that makes the marginal benefit from searching equal zero. In the third one where the worker is employed at a high wage but with a low wage growth the result could be ambiguous. However, even if the long run benefit from promotion is low, the

[22]Peleg and Yarii (1973) were the first to point out that an equilibrium strategy may fail to exist when the discount rate is not constant across periods and when there are more than three periods

long-run benefit from changing job can be even smaller because the job arrival rate is small due to the fact that the current wage is high. Therefore, also in this case the more impatient would search more. Finally, consider a worker employed at a low wage and low wage growth. In this case Proposition 1 might not hold. However, the relevance of this case is quite limited. Notice that observing patient workers sorted in low wage growth is very unlikely as Munasinghe and Sicherman (2005) argue and show empirically.

3.7 Exponential or Hyperbolic Model?

Note that from Proposition 1 the behavioral effects of a variation in β and δ are observationally equivalent. This is a shortcoming as long as it is not possible to distinguish exponential agents from hyperbolic agents. As a consequence the validity and the novelty of the hyperbolic model would be in question. However, as I mentioned in the introduction it is possible to provide evidence showing that time-inconsistent preferences are necessary to explain the regularities addressed.

Besides those already mentioned, there are three points that are useful to identify the hyperbolic model vs the exponential one. First, by Proposition 4 and 5 we know that any departure from the naivete assumption amplifies the search effort chosen by individuals and it lowers the critical wage that induces workers to search. Accordingly, if the theoretical model explains better the empirical facts than the exponential one, we shall expect more sophisticated individuals to have higher job arrival rate. Second, job changes for modest wage increases are better explained by time-inconsistent preferences: if one is very impatient in the short-term, he will be willing to change job even for a small wage gain. This allows patient long-run behavior and reasonable long-run discount factor (Rabin, 2002). Therefore, analysis of the average gains from job changes in terms of wages can be informative for the validity of the hyperbolic model. Third, if job changes for modest wage gains were driven by time-consistent preferences we should expect workers to change job with a certain persistence (continuously). Indeed, an exponential agent does not regret from past actions as he consistently chooses the best action for the current self that is not in conflict with future selves. Also in this case the analysis of the frequency of job changes can be informative for the validity of the hyperbolic model.

The empirical part tests the main implication of the model and provides (preliminary) evidence for the three points above. A more detailed framework where the hyperbolic model is distinguished from the exponential one is provided in Drago and Kadar (2005).

4 Empirical Results

I use the National Longitudinal Survey of Youth (NLSY) to test the main conclusions of the model. In the NLSY there are many variables that can proxy impatience. On this point I closely follow the procedure of DellaVigna and Paserman (2005) (DV-P (2005) thereafter) that construct on the same dataset an aggregate measure of impatience. To sum up, the goal of this section is to identify the effect of impatience on workers' on-the-job behaviors regarding search and collaboration through which workers experience different career paths. In the following subsections I will also provide evidence to distinguish the hyperbolic model from the exponential one.

4.1 Employment spells in the NLSY

The NLSY 79 is a representative sample of 12,686 young men and women residents in the US who were 14-22 years old when they were first surveyed in 1979. These individuals were interviewed annually through 1994 and are currently interviewed on a biennial basis. Individuals were interviewed on a large number of questions on their labor market history, family and social background, education, attitudes etc. From the work file history information on employment are available on a weekly basis. I construct the labor market history of each single individual from the 1979 to the 1996. In particular it is possible to link the employers of each individual from two contiguous years and to detect a job-to-job movement with the corresponding wage gain. Information on the jobs held by individuals include the corresponding wage rate, size of the firm (just for some jobs), types of occupation and industry. I classify as job-to-job transition each employer change that occurred with a wage increase and such that the difference between the start date of the new job and the stop date (in weeks) of the previous one is between zero and two. This procedure rules out multiple jobs holdings, jobs where individuals work less than 20 hour per week and jobs hold by individuals that are less than 16. In this case the assumption is that voluntary job-to-job mobility are driven by higher wages and to this extent wages approximate the value of the jobs. There are precedents for this approach (more recently Postel-Vinay, Jolivet and Robin, 2004) which validity is based on the fact that the wage rate besides to be an important variable for worker's decision is a good measure for the amenities that a job offers and for wage growth opportunities[23]. Moreover, I retain only those spells that were reported in 1985 or later until 1996 by respondents that were not enrolled in school and were not part of the military subsample. The measures of impatience (e.g., smoking) could depend on the labor market history of individuals. The first restriction - 1985 - is necessary because all the relevant measures of impatience are reported prior to 1985. In

[23]For example, Hwang, Mortensen and Reed (1998) find that the jobs that pay higher wages are those that offer higher amenities.

this way exogeneity of the measures that proxy for impatience is preserved. Accordingly, there is only right censoring where the jobs that were ongoing on 1996 are censored. Among the the job-to-job transitions it is possible to distinguish those that derive from reasons such as firing, laid off, plant closed, pregnancy, end of temporary or seasonal jobs etc. (I shall exclude these transitions, see below). Table 1 gives summary statistics. It is worth noting that the general results are very similar to those of previous studies on the same sample (e.g. see Postel-Vinay, Jolivet and Robin, 2004).

4.2 Measures of Impatience

As anticipated I use the same measures of impatience of DV-P (2005), the same procedure to construct the aggregate measure of impatience and the same assumptions. Therefore, I briefly report the discussion in DV-P (2005) on these empirical matters. The identifying assumptions are the following. "The individual's discount rate is the same across different activities. [...] The ranking of individuals with respect to impatience does not vary over time." (pp. 16, DV-P (2005)). Unlike DV-P (2005), here it is not necessary to assume that "higher measures of impatience may be associated with either higher short-run $(1-\beta)$ or higher long-run $(1-\delta)$ impatience". Among the variables that are used to proxy for impatience, DV-P (2005) individuate the following: 1) smoking behavior, 2) alcohol behavior, 3) contraceptive use, 4) no life insurance on the job and 5) having a bank account, 6) NLS assessment of impatience. It is important to note that one of the robust result of DV-P (2005) is that such measures capture short-run impatience rather than long-run impatience. For the first two variables, the intuition of why such behaviors can capture impatience is very simple. These behaviors point out a preference for today versus tomorrow. For the smoking behavior, Fuchs (1982) finds evidence that high rates of time preference are associated to smoking.[24] In the NLS individuals are asked about smoking and alcohol behavior in 1984 and in 1983. I create two dummy variables that are equal to one if the individuals smoke at least more than one cigarette per day and if they use to have heavy drinking. For the contraceptive use it is reasonable to believe that more impatience individuals have sexual intercourse without contraceptives. The logic is that contraceptives have a cost today and that the non-use in sexual intercourse may be preferable for people that are impatient or have self-control problems. Following DV-P (2005) I assign a missing value to those that are married and that did not use any birth-control method. For the fourth measure, those that have no life insurance on the job are considered as impatience (cf. DV-P (2005)). Also in this case, it is necessary to adjust the raw measure by partialling out the effects of marital status, children and age. For the fifth measure, those that have a bank account are considered patient (see DV-P (2005)) for a discussion on this point). The

[24] Also Munasinghe and Sicherman (2005) show that smoking is a proxy for impatience.

intuition here is built on an example from O'Donoghue and Rabin (2001) that show that an naive individual with a a low short-run discount rate may procrastinate for ever a financial operation with small cost and delayed benefits. The last measures (NLS assessment) derives from a question to the interviewer that reports the behavior of the respondent. Among the possible answers there is also "impatient". For this measure it is necessary to adjust it for the length of the interview.

Unlike DV-P (2005) I retain also workers of all the races and women. By taking into account contraceptive use as they do, in my case attrition shrinks the sample from about 9000 individuals to 2000 individuals. Moreover, unlike DV-P (2005) I add another variable, drug use, by weighting the use of hashish, cocaine and marijuana (as for smoking, drug use captures a preference for today versus tomorrow). Following their procedure, I find out an aggregate measure of impatience with factor analysis. I estimate a factor model via maximum likelihood. There is just one factor to be retained that accounts for more than the 60 percent of variance and for which all the factor loadings have positive sign. The conjecture is that this factor is impatience. In my procedure, the Cronabach reliability measure is 0.3327 that reflects the average correlation equal 0.0652. Both measures (Cronabach reliability and average correlation) are higher than those of DV-P (2005). It is important to notice that I obtain very similar effects (in magnitude and significance) of impatience on the relevant variables by using the aggregate measure of impatience by closely following DV-P (2005) also on contraceptive uses and not including drug use. Low correlation in this case is an expected result as the variables taken into account are noisy measures of impatience. Table 2 lists the average measures, the factor loadings and the score coefficient to create the aggregate measure of impatience. Each measure is standardized to one in order to obtain an aggregate measure with mean zero and standard deviation equal to 1. In table 3 I report the mean of each adjusted measures of impatience for individuals that have at least one, two and three spells. With the increase of the spells the mean for each measures is increasing. This is an approximate form of evidence for one of the implication of the model: completed spells are associated with levels of variables that proxy for impatience above the mean of the sample.

4.3 Impatience and Exit Rate

I take into account only the spells that do end for reasons other than those that have to do with layoff, plant closed, end of temporary/seasonal job, discharged of fired, program ended, quit for pregnancy/family reasons. The focus is on voluntarily job-to-job transitions and it is possible that more impatience workers change jobs more frequently because for example are more frequently fired or sorted in temporary jobs. I estimate a Cox proportional hazard model (Cox, 1972). The Cox model is particularly convenient for our purpose and it is useful to focus on how a particular trait or independent variable affects the probability that a certain event occurs. Here the Cox model estimates the rate at which an employed worker move to a better job and captures how the aggregate measure of impatience, as well as other covariate, shifts the hazard rate in a proportional manner. The hazard rate I estimate is written as

$$\lambda(t_i|x_i, \phi) = \lambda_0 exp(x_i' \phi)$$

where t_i the observed duration of an employment spell, λ_0 is the baseline hazard rate and ϕ is the vector of coefficients on the explanatory variable x_i for individuals $i = 1, 2, ...N$. An important feature of the Cox model is that it is not necessary to assume any parametric structure for the baseline hazard rate. This makes in a certain sense the model non-parametric. As in DV-P (2005) I treat each of multiple spells by the same individual as separate observations but at the same time I handle this form of clustering by estimating the model with robust standard errors (see Lin and Wei, 1989). The set of covariates x_i contain the aggregate measure of impatience and other explanatory variables.

In table 4 the effect of the aggregate measure of impatience on the hazard rate by running the regression with the wage, race, sex, education is reported. The coefficient on impatience is significant, positive and large. A two standard deviation increase in the aggregate measure of impatience leads to a 42 percent increase of the job arrival rate. Education attainment does not have always a significant coefficient while the coefficient on wage and on age is significant and negative as expected. The coefficients on the dummies on race and sex are positive and significant. These results are in line with previous studies on job-to-job mobility.

However, these estimates could be biased if the aggregate measure of impatience are positively or negatively correlated with variable that affect the probability of a job change and if these variables are omitted from the regression (see also DV-P, 2005). Therefore in the second part of the Table 4 I report the results by estimating the model including a large set of variables such as AFQT score, parental education background, region, occupation, industry (for the entire set see table). The effect of impatience is still strong and significant. The fact that the effect of impatience is strong and significant even when controlling for education, AFQT score and parental education rules out the possibility that coefficient on impatience is biased because it is capturing some form of human capital. In particular it does not seem reasonable to explain the positive coefficient on impatience with the fact that it captures low level of investment in human capital. If this was the case the effect of impatience should disappear when we control for measures of human capital. Moreover, the definition of the completed spells implies that workers move voluntarily to better jobs, i.e. paying higher wages, for reasons other than those related to firing, layoff etc.

In table 5 I report the coefficient of the Cox model by running the regression on all the measures that proxy for impatience. Again, except for the NLS assessment, the coefficients are strongly significant and have positive sign. I follow the same previous procedure: I first run the regression with a small set of controls and then with all the controls. In the first case a two standard deviation increase in the bank account and smoking measures leads to 15 percent and 11 percent increase in the job arrival rate, respectively.

All these effects are large and strong even when I control for the size of the firm. Such a robustness check can be done for a small part of completed spells as the NLSY provides information on this point just for some years and for some jobs ("the CPS job"). This robustness check is important because in principle it can be that more impatience workers are sorted in bad jobs, i.e. in jobs that do not offer good opportunities in terms of career, and therefore they search and change jobs more frequently than more patient workers. Presumably the size of the firm is an indicator of the promotion opportunity for a job. The results from this check show still significant and large positive impact of impatience on the job arrival rate.

To sum up, since the model predicts that more impatient workers search more and casual empiricism and empirical evidence show that higher on-the-job search effort leads to higher job arrival rates, it is plausible to conclude that more impatient workers search more. The ideal and final test would be to regress on-the-job search effort on impatience. The NLSY 79 however does not provide information on the activities (that however are difficult to measure and to observe) of employed job seekers in the sample period considered and the information available for few other years are scarce and not sufficient to perform such a regression.

It is worth noting that here opposite results respect to DV-P (2005) are obtained regarding exit rates and search effort. DV-P (2005) study unemployed workers' behavior where individuals face an immediate cost of effort and the delayed reward of employment. In their case, one of the prediction of the model is that more impatient search less and have lower exit rates. In DV-P (2005), these results are confirmed by the empirical tests from the NLSY 79. In this paper the focus is on employed workers that face a short-term reward (changing job net of the search cost) and a long-term opportunity cost (renouncing large wage increases on the same job) so that the theoretical results regarding search effort are different from those of DV-P (2005). As the model shows, in this case impatience affects positively the exit rate and this result is confirmed by estimation of Cox models from the same data source of DV-P (2005).

4.4 Impatience, Collaboration and Wage Increases

The other main conclusion of the theoretical model is that more patient workers collaborate and stay on the job by experiencing wage increases. In this subsection in order to test this conclusion I first consider one of the more important indicator of collaboration that recently has received attention in literature: the absence rate (Flabbi and Ichino, 2001 and Audas, Barmby and Treble, 2004). Absenteeism is assumed to be inversely related to the worker's output and to her performance and to this extent it can be used as a proxy for low effort (cf. Audas, Barmby and Treble, 2004). A low absence rate is shown to be positively related to the probability of a promotion or wage increase. The empirical works on absence rates, that are primarily methodological, are based on company data sets. In this paper where a large data set that includes a number of type of jobs, industries and occupation is used, it is possible to identify the relationship between impatience and absenteeism. However, given the nature of the data set used here, it is more problematic to identify the impact of absenteeism and of other collaborative behaviors on the arrival of promotion and of a wage increase. The main goal is therefore to estimate the impact of impatience on the event of absence across individuals. The last step of the empirical analysis, i.e. to check whether more patient workers are stayers and experience wage increases, is deferred at the end of this sub-section.

In the NLSY there is a question on how many hours the respondent was absent at work during the survey week and the answer is reported for the most recent job held by the individual. The model to estimate is

$$y^*_{it} = X_{it}b + Imp_i\gamma + \epsilon_{it},$$

$$y_{it} = 1 if y^*_{it} > 0.$$

Where $i = 1, 2, ...n.$ refers to individuals and $t = 1, 2, ...T.$ to the time. The independent variable y_{it} is the dummy equal to one if the worker was absent at least one hour and equal to zero otherwise. The set of controls and the explanatory variable of impatience are denoted by X_{it} and Imp where b and γ are the corresponding coefficients, respectively. In this case, in the large set of coefficient it is also included the tenure as the duration of the job (in weeks) until the date of the interview. I exploit the panel structure of the data to estimate a random effects logit model that accounts for unobserved heterogeneity. The assumption implied by the random effects model is that the unobserved component that is constant over time is not correlated with the regressors. Note that when we include among the regressors the dummy for occupation and industry, we control for unobserved characteristic that are valued differently across industry and occupation. The results of the logit model with a small set of controls and the large set of controls are reported in table 6. The coefficient on impatience is positive and significant including in the regressions the large set of controls. Most of the other coefficients such as sex, race and education are not significant, whereas many occupation and industries dummy (not reported and available upon request) are strongly significant.

One of the claim of this work is that more patient workers are also stayers and experience wage increases. To check whether this claim pass the empirical test I take into account all the workers that at time t were on job j and I create a dummy variable that is equal to one if at time $t + 1$ the worker experienced a wage increase of ten percent on job j and zero otherwise.

Notice that information on wage rates in the NLSY are available on yearly basis. Among the observations with zero I include the workers that experienced a job-to-job movement as the one defined above from job j to another job in the interval $[t, t+1]$.

The model to estimate is

$$\omega_{it}^* = X_{it}b + Imp_i\gamma + \epsilon_{it},$$

$$\omega_{it} = 1if\omega_{it}^* > 0.$$

Where $i = 1, 2, ...n.$ refers to individuals and $t = 1, 2, ...T.$ to the time and ω_{it} is the dummy equal described above. The set of controls and the explanatory variable of impatience are denoted by X_{it} Imp where b and γ are the corresponding coefficients, respectively. Finally the residual, $\epsilon_{it} = \mu\chi_i + u_{it}$, is equal to the sum of a permanent specific component that can reflect ability or other permanent characteristics and an error term. In particular μ is the return of the specific component χ_i that is invariant over time. From the model we expect to find coefficient $\gamma < 0$. The problem here is that estimates of γ can be biased if we do not take into account the permanent specific component that reflects ability because the covariance between the unobserved component and impatience may differ from zero (presumably, $cov(Imp_i, \chi_i) < 0$). Indeed workers that experience more frequently large wage increase can be those that are permanently more productive and this unobservable effect can be picked up by the measure of impatience. As before, in order to solve this problem I exploit the panel structure of the data. By observing individuals over time it is possible to absorb $\mu\chi_i$ in the equation to estimate. Results from a random-effects logit model are reported in table 7. The coefficient on the aggregate measure of impatience are negative as expected and significant including in the model the small and the large set of controls. The coefficient on age is positive as expected and significant. Males experience wage increases more frequently than women as well as individuals with high levels of education relatively to those with lower levels. The coefficients on impatient are larger and significant when the wage increase considered is greater than zero and greater than 5 percent.

4.5 Observational equivalence

As Laibson (1997) points out is not always possible to distinguish a hyperbolic model from an exponential one. Only some comparative statics and some descriptive statistics might help. Many empirical papers on job-to-job mobility find that workers move to better job even for small wage gains. If the wage approximates the value of the job and if these workers (as it seems reasonable) are given the possibility to experience wage increases on the same job, then short-run rather than than long-run impatience explains better such a behavior. If one is impatient in the short-run is more willing to change job. But at the same time moving and jumping from job to job is compatible with other long-run patient behavior and reasonable long-run discount factors. If we inferred from the trade-off implied by moving the discount factor with the exponential model we would get too low discount rates (Drago and Kadar, 2005) that are not in line the empirical evidence that shows long-run patient behavior with respect to education investment, housing or investment in financial assets. Moreover, if variation of discount factors in the exponential model were the culprit for the regularities addressed, we would observe workers to change job quite persistently (or continuously), at least in the left tail of the wage distribution. As also from Table 1 it is possible to observe, this is not the case.

Some findings in the theoretical part of the paper can guide to provide more robust evidence in support of the hyperbolic model. The implication of Proposition 4 and 5 may guide to distinguish the two models. Let take into account a measure that can proxy for sophistication. For example, the "propensity to plan" or investment in illiquid assets might proxy for some degree of sophistication in decision making. For the exponential model, controlling for the large set of covariates, the proxy for sophistication should be orthogonal to the job arrival rate whereas for the hyperbolic model we should expect a positive coefficient for this measure. Other robustness checks can be done by using the recent finding of Benjamin and Shapiro (2005) that investigate whether cognitive ability and human capital reduces psychological bias, it is possible to provide evidence in favor of the hyperbolic model.[25] Preliminary results in this research support the hyperbolic model as the proxies for sophistication have a positive impact on the job arrival rate.

5 Conclusion

The goal of this paper is to address theoretically and assess empirically the effect of impatience on workers' on-the-job behavior. Theoretically, workers' hyperbolic time preferences and the implied self-control problems explain several empirical regularities concerning job mobility and account for different on-the-job behaviors. Empirically, the effect of various measures of impatience on the relevant variables confirm the prediction of the model. More patient workers undertake behaviors that lead to promotions, they are more likely to be stayers and to follow fast-track-career paths with the same employer. Impatient workers are more likely to be movers and to experience wage increases by switching jobs. Hence, differences between stayers and movers, which in the literature have so far been attributed to unobservable characteristics, are explained in terms of time preferences. Various measures of (short-run) impatience are

[25]Benjamin and Shapiro (2005) show that different measures of cognitive ability predict more sophistication in the decision making (but they do not directly adopt "sophistication" for hyperbolic discounting in the sense of O'Donoughue and Rabin, 2001).

positively correlated to the job arrival rate and negatively correlated to collaborative behaviors that lead to promotion and to large wage increase on the same job. This is the case even when I control for variables that reflect productivity. The econometric results are robust to different estimation procedures. The hyperbolic model rather than the exponential one seems to important to explain different career and mobility patterns. However, more effort on the empirical investigation of this point is still necessary. This paper offers some interesting results for the behavioral labor economics approach and the personnel economics literature. Hence, this paper contributes to the strand of literature that shows the importance of the behavioral determinants of earnings and individuals' labor market histories (Bowles, Gintis and Osborne, 2001a and 2001b; Cawley, Heckman and Vytlacil, 2001). Finally, the paper offers a new contribution to the applied behavioral economics literature as it provides an application with a short-term reward and long-term cost.

References

Audas, R., Barmby, T. and Treble, J. (2004). Luck, Effort, and Reward in an Organizational Hierarchy, *Journal of labor Economics*, 22, pp. 379-395.

Angeletos, G. M., D. Laibson, J. Tobacman, A. Repetto and S. Weinberg, (2001). The Hyperbolic Consumption Model: Calibration, Simulation, and Empirical Evaluation, *Journal of Economic Perspectives*, 15(3)

Ainslie, G., (1992). *Picoeconomics: The strategic interaction of successive motivational states within the person*. Cambridge: Cambridge University Press.

Baker, G., M. Gibbs and B. Holmstrom, (1994). The Wage Policy of a Firm, *Quarterly Journal of Economics*, 109, 921-955.

Bloemen, H.G., (2005). Job search, search intensity, and labour market transitions: an empirical analysis, *Journal of Human Resources*, 40, 231-269.

Bowles, S., H. Gintis, and M. Osborne. (2001). "The Determinants of Earnings: a Behavioral Approach," *Journal of Economic Literature* 39, 1137-1176

Brunderl, J., A. Diekmann and P. Preisendorfer, (1997). Patterns of of Intraorganizational Mobility: Tournament Models, Path Dependency, and Early Promotion Effects, *Social Science Research*, 20, 197-216

Cawley, J., J. Heckman and E. Vytlacil. (2001). "Three observations on Wages and Measured Cognitive Ability," *Labour Economics*, 8, 419-442.

Cox, D.R., (1972). Regression Models and Life-tables (with Discussion), *Journal of the Royal*

Statistical Society, 34(2), 187-220.

DellaVigna, S. and D. Paserman, (2005). Job Search and Impatience, *Journal of Labor Economics*, forthcoming.

Diamond, P. and B. Koszegi, (2003). Quasi-Hyperbolic Discounting and Retirement, *Journal of Public Economics*, 87, 1839-1872.

Fang, H. and D. Silverman, (2004). Time-Inconsistency and Welfare Program Participation: Evidence from the NLSY, mimeo, Yale University.

Flabbi, L. and A. Ichino, (2001). Productivity, Seniority and Wages, *Labour Economics*, 6,

Gibbons, R. and M. Waldman, (1999). A Theory of Wage Growth and Promotion Dynamics inside Firms, *Quarterly Journal of Economics*,

Gruber, J. and B. Koszegi, (2001). Is Addition "Rational"? Theory and Evidence, *Quarterly Journal of Economics*, 116(4), 1261-1305

Fuchs, V.R. , (1982). *Time Preference and Health: An Exploratory Study*, in V.R. Fuchs, ed., Economic Aspects of Health, Chicago: University of Chicago Press, 1986, pp. 21442.

Hwang, H, D. T. Mortensen and W. R. Reed, (1998). Hedonic Wages and Labor Market Search, *Journal of Labor Economics*, 16, 815-847.

Laibson, D. I., (1997). Golden Eggs and Hyperbolic Discounting, *Quarterly Journal of Economics*, 112, 443-477.

Light, A. and K. McGarry, (1998). Job Changes Patterns and the Wages of Young Men. *The Review of Economics and Statistics*, 80, 2, 276-286.

Jolivet, G., F. Postel-Vinay and J. Robin, (2004). Wage Distribution and Wage Dynamics in Europe and in the U.S.: Lessons from a Simple Job Search Model, mimeo.

Kirby, Kris N. and Herrnstein, R.J. (1995). Preference Reversals Due to Myopic Discounting of Delayed Reward. *Psychological Science*, 6(2), 8389.

Koszegi, B. and P. Diamond, (2003). Quasi-Hyperbolic Discounting and Retirement, *Journal of Public Economics*, 87, 1839-1872.

Lazear, E., (1979). Why is there Mandatory Retirement, *Journal of Political Economy*, 87, 1261-84

Lazear, E. and P. Oyer, (2004a). Internal and External Labor Markets, *Labour Economics*

Lazear, E. and P. Oyer, (2004b). Internal and External Labor Markets, *American Economic Review*

Light, A. and K. McGarry, (1998). Job Changes Patterns and the Wages of Young Men. *The Review of Economics and Statistics*, 80, 2, 276-286.

Lin, D. Y and Wei, L.J. (1989). The Robust Inference for the Cox Proportional Hazards Model. *Journal of the American Statistical Association*, 84(408), 10741078.

Madrian, B. and D.F. Shea, (2001). The Power of Suggestion: Inertia in 401(k) Participation and Saving Behavior, *Quarterly Journal of Economics*, 116, 1149-1525.

Medoff, J, and K. Abraham, (1980). Experience Performance, and Earnings, *Quarterly Journal of Economics*, 95, pp. 703-736.

McClure, M, D. Laibson, G. Loewenstein and J. D. Cohen, (2004). Separate Neural Systems Value Immediate and Delayed Monetary Rewards, *Science*, 306, 503-507.

McCue, K., (1996). Promotions and Wage Growth, *Journal of Labor Economics*, 14, 175-209.

Mortensen, D., (2003). *Wage Dispersion.* Cambridge: The MIT Press.

Moscarnini, G., (2004). Job-to-Job Quit and Corporate Culture, mimeo, Yale University.

Munasinghe, L. and N. Sicherman, (2005). Wage Dynamics and Unobserved Heterogeneity: Time Preferences or Learning Ability?, *NBER working paper* n. 11031.

Munasinghe, L. and K. Sigman, (2004). A Hobo Syndrome? Mobility, Wages, and Turmover, *Labour Economics*, 11, 201-338.

O'Donoghue, T. D. and M. Rabin, (1999). Doing Now or Later, *American Economic Review*, 89(1), 103-133.

Paserman, D., (2004). Job Search and Hyperbolic Discounting: Structural Estimation and Policy Evaluation, IZA Discussion paper No. 997.

Peleg, B. and M. E. Yarii, (1973). On the Existence of a Consistent Plan of Action when Taste are Changing, *The Review of Economic Studies*, 40, 391-401.

Pergamit, M. R. and J. R. Veum, (1999). What is a Promotion? *Industrial and Labor Relations Review*, 52, 4, 581-601.

Podonly J. and J. Baron, (1997). Resources and Relationships: Social Networks and Mobility in the Workplace, *American Sociology Review*, 673-693.

Postel-Vinay, F. and J.-M. Robin, (2002). The Distribution of Earnings in an Equilibrium Search Model with State-Dependent Offers and Counter-Offers, *International Economic Review*, 43(4), 989-1016.

Prendergast, C., (1999). The Provision of Incentives in Firms, *Journal of Economic Literature*, 37, 7-63.

Rosenbaum, J., (1984). *Career Mobility in a Corporate Hierarchy*, New York, NY: Academic Press.

Slicther, S., (1919). *The Turnover of Factory Labor*, New York: D. Appleton

Strotz, R. H., (1956). Myopia and Inconsistency in Dynamic Utility Maximization, *The Review of Economic Studies*, 23,

Topel, R. H. and M. P. Ward, (1992). Job Mobility and the Careers of Young Men. *Quarterly Journal of Economics*, 107, 2, 439-479.

Appendix A

Let define $K = \lambda \int_{w}^{\overline{w}} (w' - w) dF(w') + \beta\delta\lambda \left\{ \int_{w}^{\overline{w}} [V_0(w') - V_0(w)] dF(w') + [V_0(w) - V(w^p)] \right\}$. I introduce the following technical assumptions.

A. 1 $c'(0) < K < c'(1)$.

A. 2 *There exists a* $\overline{\delta}$ *such that* $\delta < \overline{\delta} \leq 1$.

Derivation of the value function 2. At any period the worker employed at the wage w chooses s to solve:

$$\max_{s \in [0,1]} -c(s) + \lambda s [\int_{w}^{\overline{w}} w' dF(w') + F(w) w] + \lambda(1-s) w + [(1 - \lambda s - \lambda(1-s)] w +$$
$$+ \beta\delta \left\{ \lambda s [\int_{w}^{\overline{w}} [V_0(w')] dF(w') + F(w) V_0(w)] + \lambda(1-s)[V_0(w^p)] + [1 - \lambda(1-s) - \lambda s] V_0(w) \right\}. \tag{9}$$

In the first line there are the immediate costs and the immediate rewards. Namely, the search cost plus the expected return to find a new job. The third term is the probability to be promoted times w, because the compensation in the current period does not change if the worker is promoted. The next term is the probability that the worker is neither promoted nor finds a new job times the current wage w. The second line refers to the next period that is discounted by the factor $\beta\delta$. The first term in the second line in the brackets is the probability to find a new job times the associated lifetime utility. The second term is the probability to be promoted times the lifetime utility from promotion. The last term is the probability of neither promotion nor job change times the lifetime utility to stay at the current job. Rearranging this expression it is easy to get the value function 1. The continuation values (3), (4) and (5) of self -1 are obtained using the same procedure. However, the difference with respect to the value function (2) is that for self -1 at time 0 the drop in discounting has already occurred so that β does not appear in expression (3), (4) and (5). An interior solution to the problem of the worker is guaranteed by A1.

Lemma 2, Lemma 3
To study the sign of the derivative of the marginal benefit from search with respect to the wage, we have to analyze the following expression:

$$-1 + F(w) + \beta\delta[F(w) V_0(w)_w], \tag{10}$$

where $V_0(w)_w$ is the derivative of $V_0(w)$ respect to w. The first term $(-1 + F(w))$ is negative. Let consider the simple case where the derivative is with respect to the wage only, i.e. the derivative of the f.o.c, (7) in the text, with respect to the wage for self 0. The sign of this derivative leads to analyze $-1 + F(w) + \beta\delta F(w)[1 - F(w)]$. That under bounded discount factors (Assumption A2) has always negative sign for any $w \in [\underline{w'}, \overline{w'}]$. Since $V_0(w)_w < 1$, expression (10) has negative sign as well. Lemma 3 derives from standard arguments: as the cost function is convex, an in increase in the RHS of eq. (6) leads to an increase in the search effort.

Proposition 1 The proof derives directly from Lemma 3: from the theorem of implicit function and for $w < w^R$, a decrease in β results in an increase in the RHS of equation (6) because for $w < w^R$ the term multiplied by $\beta\delta$ is negative. By convexity of the cost function this leads to an increase in the optimal search effort.

Proposition 3 This proposition derives from the fact that current self does not agree on the search effort supplied by future self. Let consider the value function (2) in period -1 that is embedded with the optimal search effort and consider a small deviation from the optimal search effort. Formally, let denote the value function (2) of self -1 with Γ. We have $d\Gamma/ds = -c'(s) + \lambda \int_w^{\overline{w}} (w' - w)dF(w') + \beta\delta\lambda \left\{ \int_w^{\overline{w}} [V_0(w') - V_0(w)]dF(w') + [V_0(w) - V_0(w^p)] \right\} + \beta\delta\lambda \frac{\partial \int_w^{\overline{w}} [V_0(w') - V_0(w)]dF(w')}{\partial s} + \beta\delta\lambda \frac{\partial [V_0(w) - V_0(w^p)]}{\partial s}$. The sum of the first three terms is equal to zero by the f.o.c. (6). However the sum of the last two term is not zero as well. From the perspective of self -1 $\frac{\partial V_0(w)}{\partial s}$ is equal to $-c'(s^{i+1}) + \lambda \int_w^{\overline{w}} [(w') - (w)]dF(w') + \delta[\int_w^{\overline{w}} (w' - w)dF(w') + (w - w^p)]$, where s^{i+1} is the known (by sophistication) search effort supplied by the next self.

Step 1 From the perspective of self -1, let take into account $\frac{\partial V_0(w)}{\partial s}$.
Claim 1 $\frac{\partial V_0(w)}{\partial s}$ is negative and equal to $\delta(1 - \beta)[\int_w^{\overline{w}} (w' - w)dF(w') + (w - w^p)]$.
Add $-\delta\beta[\int_w^{\overline{w}} (w' - w)dF(w') + (w - w^p)]$ and $\delta\beta[\int_w^{\overline{w}} (w' - w)dF(w') + (w - w^p)]$ to $-c'(s^{i+1}) + \lambda \int_w^{\overline{w}} (w' - w)dF(w') + \delta[\int_w^{\overline{w}} (w' - w)dF(w') + (w - w^p)]$. It follows that $-c'(s^{i+1}) + \lambda \int_w^{\overline{w}} (w' - w)dF(w') + \delta\beta[\int_w^{\overline{w}} (w' - w)dF(w') + (w - w^p)] = 0$ because s^{i+1} is the search effort chosen such that this latter f.o.c. equals zero. Therefore $\frac{\partial V_0(w)}{\partial s} = \delta(1 - \beta)[\int_w^{\overline{w}} (w' - w)dF(w') + (w - w^p)]$ and from Assumption 1 in the text it is negative. From the first claim it follows also that $\frac{\partial \int_w^{\overline{w}} V_0(w')dF(w')}{\partial s} = \delta(1 - \beta)[\int_{w''}^{\overline{w}} (w' - w'')dF(w') + (w'' - w^p)]$ and that $\frac{\partial V_0(w^p)dF(w')}{\partial s} = \delta(1 - \beta)[\int_{w^p}^{\overline{w}} (w' - w^p)dF(w') + (w^p - w^{pp})]$ have both negative sign.

Step 2
Claim 2 $\frac{\partial V_0(w)}{\partial s} < \frac{\partial \int_w^{\overline{w}} V_0(w')dF(w')}{\partial s} < \frac{\partial V_0(w^p)}{\partial s}$.
This is immediate and intuitive by Ass. 1.
It follows that $\beta\delta\lambda \frac{\partial \int_w^{\overline{w}} [V_0(w') - V_0(w)]dF(w')}{\partial s} + \beta\delta\lambda \frac{\partial [V_0(w) - V_0(w^p)]}{\partial s} < 0$ and that $d\Gamma/ds < 0$.

Proposition 4 and 5 The proof of this Proposition 4 derives from eq. (6) and it is based on the previous proof. Let again focus on the part multiplied by $\beta\delta$. In this part the sophisticated current self cannot control the search effort supplied by future self. However, by sophistication assumption, she knows the level the future self will supply. This is reflected in the continuation values $V_0(w)$, $V_0(w^p)$ and $V_0'(w)$. Let write these continuation value as $V_0(w)(s^{i+1})$, $V_0(w^p)(s^{i+1})$ and $V_0(w'')(s^{i+1})$, that is to say that the continuation values are a function of the search effort of the future self. Recall that for sophisticated s^{i+1} is greater than the one the current self would supply in the future. The naive self believes that s^{i+1} will be exactly the one an exponential agent supplies in the future. Therefore, by Proposition 1, the s^{i+1} the naive takes into account for the maximization is lower than the one the sophisticated takes into account. Hence, in order to prove the Proposition is sufficient to show that a higher s^{i+1} implies a higher effort. The previous proof has showed that a higher search effort implies that the part multiplied by $\beta\delta$ decreases. However, since for the relevant range of the wages this part is always negative this implies that the marginal gain from search is increasing in s^{i+1}. As a consequence a sophisticated puts more effort than a naive. Proposition 5 is easily proved by recognizing that the better one predicts future self behavior the higher is s^{i+1} and the higher the search effort supplied.

Table 1: Summary Statistics on Job-to-Job Movements

Number of spells	12471
Number of completed spells	5944
Number of individuals	7749
Mean duration*	98
Duration 25th percentile	21
Median	49
Duration 75th percentile	116
Spells by individuals:	
with one completed spell	2677, 34.5
with two completed spells	938, 12.1
with three completed spells	323, 4.1
with four completed spells	69, 0.8
with five completed spells	23, 0.3
with six completed spells	4, 0.05
with seven completed spells	1, 0.01

* The mean refers only to completed spells. The completed spells reported in this table do not contain job separation due to layoff, plant closed, end of temporary/seasonal job, discharged of fired, program ended, quit for pregnancy/family reasons.

Tab. 2: Factor Analysis for the the Aggregate measure of impatience

	Factor Loadings	Uniqueness	Score coefficient
Alcohol	0.3520	0.8286	0.2301
Drug Use	0.4264	0.7221	0.3200
NLS Assement	0.0388	0.9760	0.0215
Bank Account	0.3983	0.6827	0.3161
Life Insurance	0.2545	0.7993	0.1795
Smoking	0.3329	0.8571	0.2104

Tab. 3: First evidence

	Individual with at least one spell Mean of standard adjusted measure	Individual with at least two spells Mean of standard adjusted measure	Individual with at least three spells Mean of standard adjusted measure
Alcohol	0.0469(1.0704)	0.1184(1.1669)	0.0745(1.1114)
Drug Use	0.1775(1.1404)	0.2564(1.1894)	0.3136(1.2236)
NLS Assessment	0.0099(0.7580	-0.058(0.2385)	0.006(0.3934)
Bank Account	0.0434(1.0053)	0.0367(0.8763)	0.0604(1.008)
Life Insurance	-.0342(0.7890)	0.010(0.8669)	0.4556(1.1865)
Smoking	0.0472(1.006	0.0974(1.011)	0.1750(1.0145)

Tab. 4: Exit Rate Regressions

Controls	$Coeff.$	$Std.Err.$	$Coeff.$	$Std.Err.$
Impatience	0.2128^{**}	0.03329	0.2198^{**}	0.02813
Age	-0.1645^{**}	0.099	-0.1877^{**}	0.0099
Wage	-0.0134^{**}	0.0000	-0.0015^{**}	0.0001
Sex Dummy	0.2137^{**}	0.0034	0.2041	0.0374
Race Dummy	0.1171^{**}	0.0333	0.0094	0.0414
Edu0 Dummy 1-4th grade	omitted	-	omitted	-
Edu1 Dummy 5-8th grade	-0.0192	0.2231	0.1630	0.2620
Edu2 Dummy 9-13th grade	-0.0869	0.0593	-0.0741	0.0708
Edu3 Dummy 14-17th grade	0.1360^{**}	0.0671	0.0994	0.0734
Edu4 Dummy 18-20th grade	0.4070^{**}	0.1091	0.3230^{**}	0.1067
AFQT score	-	-	0.0057^{**}	0.0008
Fathereduc	-	-	0.0019	0.0059
Mothereduc	-	-	0.0098	0.0769
Marital Status	-	-	0.0278	0.0385
Region: South	-	-	omitted	-
Region: NorthEast	-	-	0.0702^{*}	0.0464
Region: NorthCen	-	-	-0.1035^{**}	0.0059
Region: West	-	-	0.0985^{**}	0.0483
Dummy Occupation-Professional	-	-	omitted	-
Dummy Occupation-Manager	-	-	-0.2164^{**}	0.0688
Dummy Occupation-Saleworker	-	-	0.2165^{**}	0.0821
Dummy Occupation-Clerical	-	-	-0.0111	0.0623
Dummy Occupation-Craftsmen	-	-	-0.1559^{**}	0.0759
Dummy Occupation-Operatives	-	-	0.0251	0.0738
Dummy Occupation-Laborers	-	-	0.0320	0.0841
Dummy Occupation-Farmers	-	-	0.7693^{*}	0.4216
Dummy Occupation-Service	-	-	-0.1204^{**}	0.06882
Dummy Industry-Agriculture	-	-	omitted	-
Dummy Industry-Private	-	-	0.0324	0.1890
Dummy Industry-Mining	-	-	0.3361	0.2171
Dummy Industry-Construction	-	-	0.5158^{**}	0.1307
Dummy Industry-Manufacturers	-	-	0.1411	0.1205
Dummy Industry-Transports	-	-	0.2898^{*}	0.1290
Dummy Industry-Wholesale Trade	-	-	0.3471^{**}	0.1185
Dummy Industry-Finance	-	-	0.1964^{*}	0.1099
Dummy Industry-Business Services	-	-	0.5490^{**}	0.1257
Dummy Industry-Personal Services	-	-	0.1711	0.1356
Dummy Industry-Entertainment	-	-	0.2395	0.1652
Dummy Industry-Professional Services	-	-	0.1140	0.1212
Dummy Industry-Public Adm.	-	-	-0.3039^{**}	0.1417
Number of Observation	12451	-	11976	-

All the entries in first and the third columns represent coefficients on the variables from Cox proportional hazards models. The models includes five dummies indicating that the following variable have missing value: AFQT score, occupation, industry, father education , mother education.

*: Significantly different from 0 at the 0.1 level.**: Significantly different from 0 at the 0.05 level.

Tab. 5: Exit rates: regression with all the measures of impatience

	Coeff. (small set of controls)	Std.Err. Std.Err.	Coeff. (large set of controls)	Std.Err. Std.Err.
Alcohol	0.005	0.171	0.0045	0.0134
Drug Use	0.0548**	0.0165	0.0833**	0.0144
NLS Assement	-0.0348	0.0252	-0.033	0.0220
Bank Account	0.0792**	0.0203	0.0591**	0.0189
Life Insurance	0.0728**	0.0275	0.0583*	0.0235
Smoking	0.0584**	0.0179	0.0637**	0.0157

All the entries in first and the third columns represent coefficients on the variables from single Cox proportional hazards models. The model includes five dummies indicating that the following variable have missing value: AFQT score, occupation, industry, father education , mother education *: Significantly different from 0 at the 0.1 level.**: Significantly different from 0 at the 0.05 level.

Tab. 6. Impatience and Absenteeism

Controls	*Coeff.*	*Std.Err.*	*Coeff.*	*Std.Err.*
Impatience	0.0417*	0.0228	0.0514*	0.0264
Age	−0.0033	0.0263	−0.0061	0.040
Wage	−0.0029	0.0120	−0.0037	0.0279
Sex Dummy	−0.0037	0.0104	0.0041	0.0258
Race Dummy	0.0064	0.0235	0.0129	0.0818
Edu0 Dummy 1-4th grade	omitted	-	omitted	-
Edu1 Dummy 5-8th grade	-0.0042	0.0535	0.0633	0.0921
Edu2 Dummy 9-13th grade	0.0374	0.0893	0.0211	0.0808
Edu3 Dummy 14-17th grade	0.0203	0.1270	0.0544	0.0280
Edu4 Dummy 18-20th grade	0.5521	0.1091	0.0271	0.0167

All the entries in first and the third columns represent coefficients on the variables from random-effects logit model. The entries in the first and the third columns refer to the model estimated with a small and a a large set of coefficients, respectively. The dependent variable is equal to one if the worker was absent at work at least one time during the survey week and zero otherwise. The model includes five dummies indicating that the following variable have missing value: AFQT score, occupation, industry, father education , mother education. *: Significantly different from 0 at the 0.1 level.**: Significantly different from 0 at the 0.05 level.

Tab. 7. Wage increases: Results from Random-Effect Logit Model

Controls	*Coeff.*	*Std.Err.*	*Coeff.*	*Std.Err.*
Impatience	−0.1073**	0.0044	−0.1332**	0.0954
Age	0.0076**	0.0011	0.0087**	0.0038
Wage	−0.0005	0.0023	−0.0036	0.0213
Sex Dummy	0.2122**	0.0221	0.2041**	0.0320
Race Dummy	0.0030	0.0261	0.0037	0.0472
Edu0 Dummy 1-4th grade	omitted	-	omitted	-
Edu1 Dummy 5-8th grade	-0.1426	0.1488	-0.1286	0.1624
Edu2 Dummy 9-13th grade	-0.0196	0.0435	-0.0041	0.0206
Edu3 Dummy 14-17th grade	0.1395**	0.0456	0.1139**	0.0534
Edu4 Dummy 18-20th grade	0.1864**	0.0624	0.2120**	0.0434

All the entries in first and the third columns represent coefficients on the variables from random-effects logit model. The entries in the first and the third columns refer to the model estimated with a small and a a large set of coefficients, respectively. The dependent variable is equal to one if the worker has experienced a ten percent wage increase and zero otherwise. The model includes five dummies indicating that the following variable have missing value: AFQT score, occupation, industry, father education, mother education. *: Significantly different from 0 at the 0.1 level.**: Significantly different from 0 at the 0.05 level.