Career Consequences of Hyperbolic Time Preferences

Francesco Drago

May 2006

Career Consequences of Hyperbolic Time Preferences

Francesco Drago

University of Naples, Parthenope,
University of Siena and IZA Bonn

ABSTRACT

Career Consequences of Hyperbolic Time Preferences [*]

In this paper I address theoretically and assess empirically the effect of impatience on
workers' on-the-job behavior. Theoretically, short-run impatience explains 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 low absence rate and high effort
on the other, strongly affect mobility and individual wage growth. On-the-job search results in
higher wages with the new employer while collaboration leads 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. Impatient workers are more likely to be movers and to experience wage
increases by switching jobs. The model rests on the empirical findings that the long term
wage increases of stayers are in general larger than those of the movers, and the benefits
resulting from collaboration are not as immediate as the rewards from search conditional on
the arrival of a better job offer. I use a large longitudinal data set (NLSY 79) to test the
predictions of the model. Various measures of impatience are positively correlated to the job
arrival rate and negatively correlated to collaboration. Finally, using some theoretical
predictions I am able to show empirically that the results are driven by variation in short-run
impatience within the hyperbolic model rather than by variation in long-run impatience within
the exponential model.

JEL Classification: C23, C70, J63

Keywords: job mobility, hyperbolic discounting, wage growth

Corresponding author:

Francesco Drago
Dipartimento di Studi Economici
Universitá di Napoli, Parthenope
Via Medina 40
80133 Napoli
Italy
Email: drago2@unisi.it

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? Despite the increasing interest in the behavioral determinants of labor market outcomes, economists have not yet considered psychological factors to explain why workers change jobs. Indeed, the three main models on job mobility, i.e. the mover-stayer model (Blumen et al., 1955), the "search good" (Burdett, 1978; Jovanovic, 1979b) and the "experience good" models of job matching (Johnson, 1978; Jovanovic, 1979a), are based on heterogeneity in workers productivity, search activity for better matches and learning about the quality of the current match, respectively.

Recent empirical findings suggest that these arguments do not exhaustively explain why workers change jobs voluntarily. While job mobility is found to be an important channel through which workers experience wage increases (Topel and Ward, 1992), frequent mobility in the early stage of career is associated with poor wage growth performances in the future (e.g. Light and McGarry, 1998). In fact, during the first 8 years of career the wage growth rate of a worker who stays on the same job exceeds by 30 percent that of a worker who holds 4-6 jobs in the same period. This difference is not driven by less productive workers self-selecting into mobility as "immobile workers fare better than their mobile counterparts regardless of how we control for observed and unobserved heterogeneity" (Light and McGarry, 1998, p. 284). In sum, the long-run monetary reward of the stayers, who experience wage increases with the same employer through promotion and position change, is generally greater than the long-run expected gain from moving (e.g. McCue, 1996). Why do workers move then? Is it possible to complete the standard theories to take account of the above patterns? In this paper I argue that an important difference between staying and moving is that wage increases on the same job are deferred and paid with delay.[1] Relying on the fact that the benefits from promotion and position change are not as "immediate" as the ones from job-to-job movements, I propose a new explanation for job mobility that is based on workers' self-control problems. Following DellaVigna and Paserman (2005) [2] I focus on the general class of time preferences that generate self-control problems - hyperbolic discounting -, and I ask whether for identically productive individuals heterogeneity in time preferences accounts for different career paths.

The theoretical part models the optimal choice of a worker with hyperbolic time preferences in a finite period model. I restrict my attention to quasi-hyperbolic approximation of hyperbolic

---

[1]This is a common feature of compensation schemes. As Topel and Ward (1992) put it: "Inspection of the quarterly data revealed a strong tendency for within-job earnings changes to occur at annual intervals" (p. 456). Another strand of literature rationalizes deferred compensation schemes with incentive concerns. (Lazear, 1979). Also, the serial correlation in promotion rates (Baker et al., 1994; Gibbons and Waldmann, 1999) implies that some of the benefits of a promotion are materializing in the long-run.

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

time preferences (Phelps and Pollack, 1968; Laibson, 1997). Such preferences imply a lower discount factor when the delay horizon is short (short-run impatience) than when the delay horizon is long (long-run impatience) and generate self-control problems and time-inconsistency in a simple fashion. Motivated by strong experimental and field evidence, 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, 1999a).

The worker has to decide whether to climb the wage ladder by staying or by moving. What I call collaborative behaviors, e.g. high effort, good relationships with managers and supervisors and low absence rate, increase the chance of wage increases on the same job (e.g., see Flabbi and Ichino, 2001 and McCue, 1996). On the other hand, on-the-job search increases the arrival rate of a new job offer. In the basic setting collaboration and search are assumed to be substitutes. In particular a worker has to choose how to allocate one unit of time between search and collaboration (e.g., interviews with prospective employers may lead to absenteeism). The basic assumptions of the decision problem of the worker capture the stylized facts mentioned above: the long-run benefit from collaboration is greater than the one from search, and benefits that result from collaboration are not as immediate as the rewards from search conditional on the arrival of a better job offer. While movers are employed immediately at the new wage once they change jobs, a worker who collaborates and is promoted receives the wage increases with a certain delay.

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. Search 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). 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 workers with more severe self-control problems to search more and to switch job. This theoretical prediction is robust to general timing assumptions on on-the-job search effort choices and wage payoffs.

The first step of the empirical strategy is to look for cross-sectional variation in some measures of impatience and self-control problems and to show that it predicts cross-sectional variation in behavior. I use the NLSY 79 dataset that allows to control for a large number of individual variables and job specific attributes. 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, heavy drinking, not having a life insurance are some of the variables that proxy for impatience. The results of the empirical investigation support the theoretical model. The arrival rate of better jobs is strongly

and positively affected by various measures of impatience even having controlled for variables that are presumably correlated with ability, such as health status and human capital measures. It is remarkable that the results run against the predictions for some of the unobservables that are negatively correlated with impatience (e.g. the uncontrolled part of ability). For example, low ability workers have a lower hazard rate into better jobs[3] so the effect of impatience is not only not driven by omitted ability but it is probably downward biased. Moreover, these results differ from those obtained by DellaVigna and Paserman (2005) where for unemployed workers the search effort and hazard rate into employment are negatively correlated to the same measures of self-control problems. These opposite findings are not contradictory once the contextual differences are taken into account. Here employed workers have two alternatives to raise their wages and the fact that search brings benefits that materialize sooner explains the different results. Also, I find a negative relation between one of the main collaborative behaviors, namely low absence rate, and impatience.

In the theoretical model, heterogeneity of impatience within the exponential model could in principal explain variation in the behaviors mentioned above. Then, the empirical results could be driven by variation in long-run impatience rather than by variation in short-run impatience. However, there are good reasons to focus on the hyperbolic model. Most importantly, by empirically testing some theoretical predictions, I am able to distinguish between the hyperbolic and the exponential models and I show that the exponential one fails to explain some regularities in the data.

This problem of hyperbolic versus exponential discounting is addressed in the second step of the empirical strategy. Following O'Donoghue and Rabin (1999a) I distinguish between sophisticated and naive agents, where the firsts are aware of their self-control problems and the seconds are not. For the 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. In the absence of a perfect commitment technology, sophisticated are even further away from the long-run optimum than their naive counterparts (they search more than the naive do as search in the present and in the future are complements). The prediction that in the absence of commitment devices sophisticated search more is the key in empirically distinguishing the hyperbolic from the exponential model. The lack of commitment devices in the labor market is generally assumed in the literature, which, on the one hand, justifies the approach. On the other hand it also implies that it is very difficult to proxy for sophistication.

Fortunately, the NLSY 79 dataset includes data on retirements assets like IRA that are found to be perfect commitment devices for agents that have self-control problems regarding consumption (Benhabib and Bisin, 2005). In an attempt to proxy for sophistication, I construct a dummy equal to one if respondents in the NLSY 79 have money assets in retirement accounts

---

[3]See the positive coefficients of education and AFQT score in the hazard rate regressions in the empirical section.

such as IRA and zero otherwise, and I adjust it for cumulative earnings and income. Under the identifying assumption of no commitment power in the labor market, it is possible to test the theoretical prediction that sophisticated search more: the proxy for sophistication should have a positive impact on the job arrival rate. Note, however, that this is also a test to distinguish hyperbolic discounting from exponential discounting. Indeed, in the exponential setting where agents are time-consistent, investment in retirements accounts reflect long-run patience which should have a negative impact on the job arrival rate. The results confirm that the impact of IRA is positive, strong and significant even having included a large set of controls in the regressions. I show that several alternative explanations, like differences in human capital, risk preferences and search costs, fail to explain the overall results.

The main contribution of this paper is twofold. I propose a new explanation for job mobility in which for the first time a behavioral trait such as impatience is shown to be a determinant factor of workers' career. As other recent papers show, preferences, attitudes and personality are at least as important in explaining labor market outcomes as standard factors such as education (Bowles et al. 2001, Heckman et al. 2001). The second contribution concerns the hyperbolic discounting literature. Field evidence on time-inconsistency sometime concerns decision making of agents (e.g. unemployed workers in DellaVigna and Paserman (2005), recipients of food stamps in Shapiro (2005)) that are likely to have low levels of human capital. Self-control problems are more frequent among these agents for whom the detrimental effects of time-inconsistency are more severe on a variety of outcomes. This paper provides evidence that also agents with high cognitive ability may suffer from self-control problems. In this respect, the model has the advantage to deliver some testable predictions that empirically are not confounded by elements such as low human capital or low ability.

In the next section I summarize the empirical evidence on which the model is built. In section 3 I develop a simple model that captures the effect of impatience on workers' behaviors. The empirical part of this study is presented in Section 4 and section 5 concludes.

2 Stylized Facts

2.1 Wage Growth for Stayers and Movers

In this paper collaboration is certain kinds of worker's activity that leads to promotion and more generally to wage increase on the same job, e.g. high effort for the task assigned, absence of misconduct behaviors, low absence rate (Flabbi and Ichino, 2001; Audas, et al., 2004).[4]

Wage growth within and between jobs has been extensively analyzed empirically especially concerning young workers in the US. In an influential paper on this issue Topel and Ward (1992) find that job mobility is very frequent for young individuals. A job change with a different

---

[4]Low absence rate is generally associated with a high performance-evaluation which is a strong predictor for a promotion.

employer is on average associated to a twelve percent increase in the wage rate but this increase is lower for workers with more experience. On the other hand, in the absence of job mobility, the average rate of annual wage growth for new entrants into the labor force is about fourteen percent (Topel and Ward, 1992, p. 456).[5] However, in most of the studies on job mobility, the growth rates of stayers do not focus on promotion and have to be considered as conservative estimates of wage increase deriving from collaboration. McCue (1996) analyzes 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. The positive difference between the payoffs from staying and moving is even larger if we look at the long-run consequences of mobility and promotion. There is strong evidence that in the long-run stayers have higher wage growth rates than workers who voluntary initiate mobility (Mincer and Jovanovic, 1981; Munasinghe and Sigman, 2004; Light and McGarry, 1998). Moreover, a promotion in the present predicts future promotions. The serial correlation in promotion rates is a robust finding in the personnel economics literature (see Gibbons and Waldman 1999). Hence, I shall assume that average wage increase from staying and being promoted is higher than the average wage increase from moving to a new job.

The assumption that wage increases from promotion are higher than the ones from moving, however, raises the problem of unobserved counterfactual, i.e. we do not have information on what wage increase the worker who moved would have receive if she had stayed on the same job. However, the results of a recent paper of Light and McGarry (1998) support the view that the unobserved counterfactual does not invalidate this assumption. Utilizing the panel structure of their data set they are able to control for unobservables and compare the wage gains of "identical" movers and stayers. They conclude that stayers are better off than movers in terms of wage gains payoffs. They provide strong evidence against the mover-stayer model according to which movers are less capable than stayers. To this extent, differences in wage growth performances between movers and stayers are not driven by less productive workers self-selecting into mobility. Immobile workers experience a higher wage growth "regardless of how we control for observed and unobserved heterogeneity" (p. 284). Also the empirical results in this paper support the view that there is no reason for movers not to be eligible for at least the same magnitude of wage increase as their peers who are staying. The empirical section, in line with the results of some papers such as Royarlty (1998) and Light and Omori (2004), reveals that measures of human capital such as education and AFQT score have a positive impact on mobility.

Concerning the hazards rates of promotions and job-switches, Topel and Ward (1992) find that in the US the annualized arrival rate of a new job for young workers is high at the beginning of their career and it stabilizes, on the value of 0.2 after a couples of years of tenure. Jolivet,

---

[5]In the National Longitudinal Survey, Light (2005) finds very small or insignificant difference in annual wage growth rates among job stayers and job movers.

Postel-Vinay and Robin (2004) estimate a job arrival rate between 0.09 and 0.22.[6] Similar estimates are found for the arrival rate of a promotion that is also higher at the first years of the career (e.g. McCue, 1996, Lazear, 1995).

There exists an important qualitative difference between wage increases within jobs and wage increases due to job changes. Unlike movers that conditional on the arrival of a better job offer switch job and receive the new wage, stayers get bonuses and wage increases from promotion and collaboration with some delay. Within-jobs earnings changes generally occur at annual intervals (Topel and Ward, 1992). Also, the literature on deferred compensation offer strong arguments of why employers find it convenient to postpone wage increases. Moreover, the presence of serial correlation in the arrival rates of wage increases and of promotions implies that the benefits of a present promotion are materializing in the long-run.

2.2 An Overview of the Data

As a preliminary step to address why a model of quasi-hyperbolic discounting contributes to explain job mobility, in this sub-section I present some descriptive statistics about movers and stayers in the National Longitudinal Survey of Youth (NLSY) for the period 1985-1996. The definition of the employment spell and the detailed description of the sample analyzed is offered in the empirical section. Here it is important to emphasize that: 1) the job changes considered do not include the job terminations for reasons such as firing, layoff, end of temporary job[7], and 2) the period (1985-1996) was chosen for the reason that all measures of impatience are taken prior to 1985. Table 1 presents descriptive statistics about job changes. The sample consists of 2821 stayers (immobile workers) who held only one job from the beginning of their career until 1996 and of 7724 movers (or 10307 movers if I consider *also* the spells associated to wage decreases). Table 1 reveals a large heterogeneity in the amount of mobility. Moreover, the majority of the transitions are "instantaneous" job changes: out of all transitions, those that occur in less than 2 weeks (i.e. the new job starts in the same or one week later than the old job terminates) represent about 70 percent of the total transitions.

Along the lines of DellaVigna and Paserman (2005), in the NLSY I use several proxies for short-run impatience such as not having a bank account, smoking and heavy drinking (see the empirical section). Table 2 presents means and standard deviation of these measures. The table shows that movers have higher and statistically different measures of impatience. One could argue that these movers jump from job-to-job because are less capable than stayers. It is well known that more impatient individuals invest less in human capital and the positive correlation between measures of impatience and mobility could be explained through this channel.

---

[6]However, the majority of the studies on on-the-job arrival rates do not take into account individuals' search effort that is usually not observable.

[7]Some previous studies do not attempt to distinguish between voluntary transitions and involuntary separations and consider as movers all workers who change job (Munasinghe and Sigman, 2004). The definition of the employment spell (in section 4) in this paper aims at detecting voluntary transitions.

However, Table 2 shows that movers have similar and in many cases statistically not different years of education from those of stayers. Moreover, movers have generally higher AFQT scores. Table 2 also reports for movers higher measures of educational parental background which is to a large extent exogenous to the individuals.

A model of job mobility and hyperbolic discounting can accommodate well this preliminary evidence together with the finding that "mobility does little to 'help' but can do a significant amount to 'hurt' wages" (Light and McGarry, 1998). In the presence of self-control problems, (short-run) impatient workers initiate frequent job changes at the beginning of their career even if it is sub-optimal for their long-run selves. In such a model, a key theoretical prediction is that sophistication is associated to higher mobility rates. In fact, in the empirical section, AFQT score and savings in IRAs account that represent to a certain extent a measure of sophistication, have a positive impact on the arrival rate of new jobs as well as several measures of short-run impatience.

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 are back to the standard exponential discounting case. When $\beta \in (0,1)$, time preferences are quasi-hyperbolic. The discount factor between the current period t and the next one $t+1$ is $\beta\delta$, while the discount factor between any two periods in the future is δ. These preferences capture and stylize a simple fact documented by dozens of experiments: individuals do not behave according to the exponential discounting model in which the rate of time preferences is constant and preferences are time-consistent. In particular, individuals are more impatient in the short-run than in the long-run (e.g. see Thaler 1981; Ainslie, 1992). With the structure of quasi-hyperbolic discounting, suppose $\beta = 0.9$ and $\delta = 0.99$. In current period t, a reward of 10 that starts materializing in period $t+3$ is preferred to a reward of 9 that starts at period $t+2$ ($\beta\delta^3 10 > \beta\delta^2 9$). When period $t+2$ arrives, this preference is reversed by the same individual as $9 > \beta\delta 10$. In other words, quasi-hyperbolic time preferences involve self-control problems.[8]

Following the tradition of the hyperbolic discounting literature I study the different selves

---

[8]Many empirical regularities on consumption (Angeletos et al. 2001; Laibson, 1997), drug addiction (Gruber and Koszegi, 2001), unemployed worker behavior (Paserman, 2004) among others, are explained better in terms of quasi-hyperbolic preferences rather than with time-consistent preferences.

of the individual that in each period take some decisions (Laibson, 1997; O'Donoghue and Rabin, 1999a). 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, that is, to face in the future periods $\beta = 1$. (O'Donoghue and Rabin, 1999a). The seconds are aware of their self-control problems, i.e. each self knows how future selves will discount future payoffs.

3.2 Three period model

Consider a worker employed at the current wage w. Such a worker can experience a wage increase either by being promoted or by changing job (employer). In the first case a promotion implies a wage increase $w^p - w$ according to a deterministic function $\Phi(w) = w^p$.[9] The generic wage w with the superscript p denotes the wage after promotion for a worker whose wage before the promotion is w and $\Phi(w)$, $\Phi'(w) > 0$, maps continuously the wage w into $[w, \overline{w}']$. The function $\Phi(w)$ is the same across all jobs, i.e. all jobs offer the same opportunities in terms of wage growth. The only restriction on $\Phi(w)$ is stated in Assumption 1 (see below). In the second case (changing job), the worker switches job if from the known c.d.f $F(w')$ she receives a new wage offer, w', that she prefers to the current one, w. The c.d.f $F(w')$ is assumed to be continuous on the support $[\underline{w}', \overline{w}']$, with the first moment $F'(w') > 0$. I denote with $E(w'|w' > w)$ the expectation of w' (i.e. the expectation of the wage of a new job offer) conditional on $w' > w$, where w is the current wage.

Moreover, $s \in [0, 1]$ is the on-the-job search effort chosen by the worker that affects the probability to receive a job offer. I assume that the worker has to allocate one unit of time between search and collaboration. Hence, collaboration equals $1 - s$ and affects the probability to receive a promotion. Denote λs and $\lambda(1 - s)$, where $\lambda < 1$ is a constant parameter, as the probability to receive a new wage offer and to be promoted, respectively. Then, the expected gain from collaboration is

$$\lambda(1 - s)(w^p - w).$$

Since with probability $1 - F(w)$ the wage offer is greater than the current one w, the expected gain from search is

$$\lambda s[1 - F(w)](E(w'|w' > w) - w)$$

In this setting, I develop a three period model to show the basic behavioral effects of different forms of impatience. The periods are labeled with $t = -1, 0, 1$ and in each period the worker chooses the intensity of her behavior concerning collaboration and search. Based on the previous discussion the main assumptions are the following:

---

[9]The wage has to be interpreted as a measure of monetary compensation and amenities (e.g. fringe benefits, work place).

Assumption 1 *For any w, the c.d.f. $F(w')$ and Φ are such that $[1 - F(w)]E(w' | w' > w) < w^p$.*

Assumption 2 *A worker has to allocate one unit of time between search and collaboration. Hence, collaboration equals $1 - s$.*

Assumption 3 *The relative cost of search, $c(s)$, is a U-shaped, convex cost function of s.*

Assumption 4 *The worker employed at wage w may experience a promotion in period t and is paid w^p in period $t + 1$. The worker that searches at period t, conditional on the arrival of a job offer, may switch job at period t and is paid the new wage in the same period.*

Assumption 1 states that the expected reward from being promoted is higher than that from moving. Assumption 2 implies that search and collaboration are substitutes. Search requires time that could be invested in the task assigned and more generally in collaboration (e.g. interviews with prospective employers may lead to absenteeism). Assumption 3 specifies the functional form of the search cost, but the results also hold for an increasing, convex cost function.[10] Finally, Assumption 4 captures the fact that wage increases from promotion are not as immediate as the wage increases from search because those deriving from collaboration are deferred. Moreover, it implies that the costs and benefits associated to searching for and finding a new job belong to the same period. In sub-section 3.5 I relax this assumption and show that the results derived here carry over to more general timing and preference structures.

I assume that for each period the worker can receive only one wage offer (so that she can switch job only once per period) and that the joint probability of a promotion and of a job offer is negligible in the interval of time considered. Then, in period -1 the self -1 of the worker employed at the wage w chooses the search effort to solve:

$$\max_{s \in [0,1]} w - c(s) + \lambda s \int_w^{\overline{w}'} (w' - w) dF(w') +$$

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

where $V_0(w)$ is the lifetime utility of employment at wage w at period 0 from the perspective of self -1 (see Appendix A for a detailed construction of expression (1)). Expression (1) uses the fact that a worker moves to another job if and only if it pays a higher wage (see Lemma 1 in Appendix A). The first line of eq. (1) 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

---

[10]Derivations can be found in the earlier version - Drago (2005).

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 $V_0(w)$.

At period 0, a worker employed at the generic wage w', may again switch job, be promoted (and receive the wage increase at period 1) or staying without changing irrespective of what she experienced at period -1. Therefore, from the perspective of self -1, the continuation payoffs in expression (1) from staying without a promotion, changing job, and staying experiencing a promotion are, respectively:

$$V_0(w) = w - c(s) + \lambda s \int_w^{\overline{w}'} (w' - w)dF(w') +$$

$$\delta \left\{ \lambda s \int_w^{\overline{w}'} w'dF(w') + \lambda(1-s)w^p + [1 - \lambda(1-sF(w))]w \right\}, \tag{2}$$

$$V_0(w'') = w'' - c(s) + \lambda s \int_{w''}^{\overline{w}'} (w' - w'')dF(w') +$$

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

$$V_0(w^p) = w^p - c(s) + \lambda s \int_{w^p}^{\overline{w}'} (w' - w^p)dF(w') +$$

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

where in (3) $V_0(w'')$ is continuation payoff from moving and w'' is the realization of the new better wage offer when the worker is employed at the wage rate w. Note that in these three expressions, the continuation payoffs at period 1, i.e. those multiplied by δ, consist of the wages only. This is equivalent to say that in the last period self 1 cannot experience a job change.[11] The short-term discount rate β is not present in (2), (3) and (4) because these are the continuation payoffs from the perspective of self -1 once the drop in the discount rate has already occurred.

I start by signing the effect of impatience on search effort for self -1. In doing so I abstract for the moment from sophistication and naivete (the analysis of the solution of the game implied

---

[11]In reality, in the last period the worker does want to jump to another job because this is the only way through which she can experience a wage increase. However, since the current wage is a sufficient statistics for the value of a job, the results are the same if we assume as in (2), (3) and (4) that the continuation payoffs consist of wages only.

by hyperbolic time preferences is offered in the next subsection). From (1) the f.o.c. is:

$$c'(s) = \lambda \int_w^{\overline{w}'} (w' - w) dF(w') + \beta \delta \lambda \left\{ \int_w^{\overline{w}'} [V_0(w') - V_0(w)] dF(w') + [V_0(w) - V(w^p)] \right\} \quad (5)$$

Technical assumptions that guarantee an interior solution are provided in Appendix A. From Assumption 1 and the fact that expression (1) is increasing in the current wage, it is immediate to see that in expression (5) the part multiplied by $\beta\delta$ is negative. The latter can be termed as the marginal long-run cost of moving. A simple exercise of comparative statics, with two workers identical but in the degree of short-run impatience, delivers the following basic result.[12]

Proposition 1 *Let worker 1 and worker 2 have the same δ and different β, with $\beta_1 < \beta_2 = 1$. Then, for any $w \in [\underline{w}', \overline{w}']$, the optimal search effort of worker 1 is higher than the optimal search effort of worker 2.*

Therefore, for identical discount factors δ, a hyperbolic discounter searches more than an exponential one. Proposition 1 can be generalized by saying that more impatient workers in general search more and are more likely to be movers. Although the observational equivalence between δ and β, i.e. in (5) the effect of a variation in δ on s is the same of that of a variation in β, will be addressed later, it is useful to note that most estimates of δ lie in a rather narrow range ($[0.95, 9/10]$), while estimates of β have a much broader range of variation ($[2/5, 1]$) (Fang and Silverman, 2002 and Paserman, 2004). Therefore, the considerable large variation in mobility rates, that has been observed by many studies (e.g. Light and McGarry, 1998), other things equal, should be due to variation in β.

3.3 Equilibrium for Sophisticated

In the rest of the theoretical analysis I distinguish between sophisticated and naive agents. I start from the sophisticated that are those who are aware of their self-control problems, i.e. they perfectly predict future selves behavior. Sophistication involves a conflict between self -1 and self 0. 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.

Hence, a sophisticated may solve the adverse consequences of her self-control problems by precommiting future selves behavior. This is possible when there exists a commitment device that forces future selves to follow a behavior that is optimal from the point of view of the current self. 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 is optimal respect to the choices of future selves. More precisely, in this case there exists an

---

[12]The proofs of all results are derived in Appendix A.

intrapersonal game with perfect information. Here I first study the case in which there is no commitment power, and then the possibility of commitment.

The solution to the game implied by the absence of commitment device is found by backward induction. Note that assuming, as in the previous sub-section, that self 1 cannot move in the last period is innocuous and simplifies the analysis (there is no conflict between self 1 and selves -1 and 0 because at time 1 there is not an opportunity cost of moving). Analyzing the case in which self 0 at period 0 earns the wage w, the first order condition for self 0 at period 0 is the following:

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

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

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

From these expressions it easy to see that for any current wage w, self -1 disagrees on the optimal level of search effort that self 0 will supply.

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 expression (5) the search effort of self 0 implied by the self 0's f.o.c. (6). The equilibrium concept is a Nash equilibrium in Markov strategies that for each self and for each wage specifies the optimal search effort. 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.*

In the three period model this proposition applies to self -1 and self 0. However, this proposition can be generalized to a N finite period model (see sub-section 3.5). The existence and the uniqueness of an equilibrium for the three period model is based on standard arguments.

Now consider the possibility of commitment devices. From expressions (5) and (6), it is easy to see that for any wage rate self -1 believes that self 0 will search too much. As a consequence, the sophisticated worker is willing to pay a positive price for a commitment device that forces future selves to stick to a behavior that is optimal from her point of view.

Proposition 3 *For self -1 there exists an $\alpha > 0$ such that 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 herself to pay a penalty for a job change in the next period 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. In the rest of the paper I assume that there are not commitment devices for sophisticated.

3.4 Exponentials, Naives and Sophisticated

Naivete implies that the agent believes that her continuation payoffs is the same of an exponential agent, i.e. she believes that future selves will behave according to $\beta = 1$. At present a naive believes that for the same wage rate in the future she will search less than today. However, when the future arrives she faces again the self-control problems of past selves. For naives there is no intrapersonal game among the selves (expression (5) and (6) are identical for a naive). In this simple setting the optimal solution for the search effort is trivial as we need to consider how self -1 chooses the search effort when she believes that self 0 will choose as an exponential agent. By strict convexity of the cost function, standard arguments lead to the existence and uniqueness of the equilibrium. The same reasoning for the existence and the uniqueness applies to the case of the exponential agent. In this case, 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 to compare them with those of sophistication. Some algebra leads to the following Proposition.

Proposition 4 *Consider a naive and a sophisticated with identical β and δ. For the self -1 of a sophisticated worker s^*_{-1} is higher than the one supplied by the self -1 of a naive worker.*

To build intuition for this result notice that search in the present and in the future are complements. By searching more tomorrow one increases the probability to receive an immediate wage increases and, given that the wage distribution has an upper bound, in turn one decreases scope for future collaboration. In other words, searching more tomorrow implies a lower long-run cost of search today. The self -1 of a sophisticated, by anticipating that self 0 will exert a search effort that is higher than the one predicted by the same self of a naive, perceives a lower long-run cost of search and therefore search more.[13]

Define $\hat{\beta}$ as the worker's belief about the degree of self-control she will have in the future, β. While for naives $\hat{\beta} = 1$ and for sophisticated $\hat{\beta} = \beta$, partially naives are those for which $\hat{\beta} \in (\beta, 1)$ (O'Donoghue and Rabin, 2001).

Proposition 5 *For a hyperbolic discounter, the optimal search effort s^*_{-1} of self -1 is increasing in the degree of sophistication, i.e. it is decreasing in $\hat{\beta} - \beta$.*

This proposition generalizes the result of Proposition 4. An important implication of Proposition 5 is that for given β any departure from the naivete assumption monotonically increases the job arrival rate of the worker.

---

[13]The result of Proposition 4 is in line with the "sophistication effect" found by O'Donoghue and Rabin (1999a) in the case of short-term rewards and delayed costs.

3.5 Robustness

Assumption 4 and Timing

The model built in the previous section rests on the assumption that workers consider the search process and its payoffs as belonging to the "present" and the benefit from collaboration to the "future". In reality, however, search takes time and it precedes the arrival of job offers. Therefore, it is important to investigate whether the theoretical predictions derived earlier carry over to more general assumptions about the decision problem. Suppose that in the framework discussed above the search cost was born at $t = -1$, wage increase from job-change was received at $t = 0$ while that from promotion at $t = 1$. I show in Appendix B that the functional form of hyperbolic discounting, of which quasi hyperbolic discounting is a limit case, leads to the same theoretical predictions established before.[14] The only exception is quasi-hyperbolic discounting, for which the results do not hold under this alternative timing structure. I conjecture that a model based on pure hyperbolic discounting delivers the results described above under various, general formulations of the workers' decision problem. Whether the limit case had any relevance in this context is addressed and rejected through the empirical analysis.

General equilibrium

Another important point is summarized in the following questions. Would the main implication be supported in a model where the decision problem of the employer is considered? Why an employer would be willing to hire a worker with self-control problems? It is reasonable to believe that the answer to the first question is affirmative. Indeed, it is difficult to discriminate between workers with self-control problems and workers with exponential preferences since past mobility is a noisy signal of time-inconsistent preferences. Second and more importantly, the marginal benefit from search is decreasing in the wage: impatient workers do not change jobs continuously as they stop searching when they receive a high critical wage (see the proof of Lemma 3 in Appendix A). Moreover, the model implicitly assumes that firms do not make counter-offers to retain workers that find new jobs. Counter-offering might affect some results. Matching alternative offers is important from the theoretical point of view to generate wage increases on the same job (Postel-Vinay and Robin, 2002) but it is very limited in reality. Mortensen (2003) offers a number of reasons to argue that it is not reasonable to assume counter offering unless one considers particular labor markets as the academic one (on this see also Moscarini, 2004).

N period model

Is it possible to generalize the model to a finite $N = t, t+1, ..., N-1$ period model? The answer depends on whether Proposition 2 applies also to the N finite period model. Concerning the

---

[14]In Appendix B, I use the functional form proposed by Loewenstein and Prelec (1992) which is, in a limit case, the discrete counterpart of the continuous setting of Harris and Laibson (2004).

existence, the extension to the N period model does 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 (Peleg and Yarii, 1973). Moreover, such an equilibrium in a finite period model is generally unique. Since the nature of the conflict between self $-i$ and $-i+1$ in a N finite period model is the same of the one between self -1 and self 0 in the three period model, all other propositions hold.

4 Empirical Results

The empirical strategy consists of two steps. I first look for cross-sectional variation in some measures of self-control and impatience I show that it predicts cross-sectional variation in behavior as predicted by Proposition 1. In the second step, by using some theoretical predictions I provide evidence in support of the hyperbolic model versus the exponential one. Finally, I analyze several alternative explanations.

4.1 Employment Spells in the NLSY

I use the National Longitudinal Survey of Youth (NLSY) for the empirical analysis. 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. 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. I retain all jobs held by individuals that were reported in 1979 or later until 1996 by respondents that were not enrolled in school and were not part of the military subsample. I define an employment spell to detect a voluntary job change. This employment spell is defined as an employer change such that:

1. the difference between the start date of the new job and the stop date (in weeks) of the previous one is between zero and four[15],

2. the new job is associated to a higher wage,

3. the termination of the previous is not due to reasons such as firing, laid off, plant closed, quit for pregnancy/family reasons, end of temporary or seasonal jobs, program ended.

---

[15]In a previous version of the paper - Drago (2005) - I performed the analysis considering the spells such that the difference between the start date of the new job and the stop date of the previous one is between zero and two. Very similar results are obtained.

By considering only transitions to better paid jobs, it is very likely that the analysis takes into account only voluntary transitions. At the same time it does not induce any selection bias that affects the results. Indeed, when condition 2 is relaxed the results are very robust (see sub-section 4.5). I follow standard procedures in excluding jobs for which abnormal wages or extremely low wages are reported. The entire procedure rules out jobs where individuals work less than 20 hour per week and jobs hold by individuals that are less than 16. I retain only the employment spells that were reported in 1985 or later. The measures of impatience that I consider in the next sub-section (e.g. smoking) could depend on the labor market history of individuals. The first restriction - 1985 - is necessary because the majority of the measures of impatience are reported prior to 1985. Information on the jobs held by individuals include the wage rate, size of the firm (just for some jobs), types of occupation and industry.

Table 1 gives summary statistics that was in part discussed in section 2. There are 2821 stayers defined as those workers that held only one job until 1996. Among movers, there are 2808 workers with one completed spell and 1155 and 423 workers with two and three completed spells, respectively. There is also a considerable amount of workers that changed job 4(140), 5(43) times (number of workers in parenthesis) in the period considered. Despite the definition of an employment spell is strict respect to the literature, mobility seems to be a strong phenomenon as many other have found.[16] The mean duration of a completed spell is 84 weeks, while the median is equal to 42 weeks and the 25th percentile is equal to 18 weeks. In the second column of Table 1 are reported the job-to-job transitions according to the definition above but including also the transitions associated to wage decreases.

4.2   Measures of Impatience

In the NLSY several variables can be used to proxy for impatience. On this point I closely follow the procedure of DellaVigna and Paserman (2005) (DV-P (2005) thereafter) who construct on the same dataset several proxies for impatience to test for its effect on unemployed workers labor outcomes in the period 1985-1996. Hence, I briefly report the discussion in DV-P (2005) on the construction of the proxies. The variables used considered are the following: 1) whether or not the respondent is a smoker, 2) alcohol consumption (number of hangovers in the last 30 days), 3) contraceptive use, 4) no life insurance on the job, 5) not having a bank account, 6) NLSY assessment of impatience. The first five variables involve a trade-off between immediate rewards and delayed costs. In DV-P(2005) a detailed discussion of why these variables capture impatience is provided. The last measures (NLSY assessment) derives from a question to the interviewer that reports the behavior of the respondent during the interview. Among the

---

[16]Note that the sample period here does not consider job changes prior to 1985 so that the youngest worker in the sample is 20 years old. The decline of job mobility with age and experience is a very well-known pattern that explains some differences in the amount of mobility with previous studies. The empirical results, however, do not change considering job changes prior to 1985.

possible answers there is also "impatient". By taking into account contraceptive use as in DV-P (2005), in my case missing values reduce the sample to less than 2000 observations (e.g. also women would be excluded). Hence I replace contraceptive use with another measure, i.e. number of times respondent have used cocaine, that also captures a preference of the type of today versus tomorrow and self-control problem. In this way I retain about 10000 observations.

The raw measures are adjusted to partial out confounding elements, e.g., marital status, children and age for the life-insurance, individual's age and cumulative past earnings for the bank account. Each proxy is standardized to have mean zero and variance equal to one in the sample population. Descriptive statistics about these proxies is presented in Table 2 and discussed in section 2. Movers with higher employment spells are associated to higher measures of impatience.

Following the procedure of DV-P (2005), I find out an aggregate measure of impatience with factor analysis by using the six standardize proxies described above (smoking, drinking, no life insurance, no bank account, NLS assessment, and drug use). I estimate a factor model via maximum likelihood. There are just two factors to be retained. One of this accounts for more than 60 percent of the variance. For this factor all factor loadings have positive sign. The conjecture is that this factor is impatience. In my procedure, the Cronbach reliability measure is 0.269 that reflects the average correlation equal 0.059. Both measures (Cronbach reliability and average correlation) are very similar to those of DV-P (2005). Low correlation in this case is an expected result as the variables taken into account are noisy measures of impatience (see DV-P(2005)). Table 3 lists the factor loadings and the score coefficient to create the aggregate measure of impatience. The measures that receive most weight are drug use, smoking, and not having a bank account.

Finally, the identifying assumptions maintained in the empirical analysis 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 higher measures of impatience may be associated with both higher short-run $(1 - \beta)$ and higher long-run $(1 - \delta)$ impatience. In the context of this application, however, it is important to note that one of the robust results of DV-P (2005) is that variation in these variables capture variation in short-run impatience and not variation in long-run impatience.

4.3 Impatience and Exit Rate

In order to assess the effect of impatience on the hazard rate of new jobs I estimate a Cox proportional hazard model (Cox, 1972). The Cox model estimates the rate at which an employed worker moves to a better job and captures how the measures of impatience, as well as other covariates, shifts the hazard rate in a proportional manner. The hazard rate that I estimate is written as

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

where t_i is the observed duration of an employment spell (defined in sub-section 4.1) in weeks, λ_0 is the baseline hazard rate and ϕ is the vector of coefficients on the explanatory variable x_i for individuals $i = 1, 2, ...N$. In this specification a spell is censored if it remains in progress when the observation period ends.[17] An important feature of the Cox model is that it is not necessary to assume any parametric structure for the baseline hazard rate. As in DV-P (2005), where completed spells are unemployment spells, I treat each of multiple spells by the same individual as separate observations. This form of clustering is handled by estimating the model with robust standard errors (see Lin and Wei, 1989).

In table 4 I report the results of separate estimations of the model above. In particular, in the first three columns, I report the results of 7 different models that differ from each other only for the measure of impatience included among the regressors, x_i. In the first column I report the results of the models that include among the covariates only the measures of impatience. All coefficients in the separates models are positive, have a strong impact and are statistically significant at conventional levels. A two standard deviation increase in the aggregate measure of impatience, for example, leads to a 30 percent increase in the job arrival rate. The variable on the bank account, drug use and smoking have similar large effects, while not having a life insurance and alcohol consumption have somewhat smaller effects. Hence, all these measures predict mobility into better paid jobs.

In the second column of Table 4, I report the results of the models including a small set of covariates. In particular, I include a set of dummies (sex and race), the wage of the current job, AFQT score, and the age of the worker (the results of for these variables are reported in Table 5 for a Cox model that includes the aggregate measure of impatience). The effects of the various measures of impatience are still strong and highly significant. In this case a two standard deviation increase in the aggregate measure of impatience leads to a 25 percent increase in the job arrival rate. By including a measure of cognitive ability like AFQT score we should have expected a higher coefficient of impatience, as AFQT and impatience are negatively correlated and AFQT has a positive effect on the hazard rate (see Table 5). On the other hand, however, the wage is also negatively correlated to impatience but has a negative impact on the hazard rate. For this reason in the second column we do not observe in general higher coefficients on the measures of impatience.

In the third column of Table 4, I report the results of the models by including a large set of covariates such as years of education, parental educational background, geographical dummies, occupation, industry (for the entire set see Table 5). The effect of impatience is slightly lower than in the previous regressions but still strong and significant.[18] In this case a two standard

---

[17]Hence, I do not consider those workers who switch to unemployment.

[18]The fact that conditioning on occupation the coefficient on impatience remains large and statistically significant provides some indication that sorting across occupations based on the degree of impatience does not explain the empirical results. To provide further support on this point, I also calculated the average measure of the proxy for the aggregate measure of impatience for each occupation. I then used among the regressors these

deviation increase in the aggregate measure of impatience leads to a 22 percent increase in the job arrival rate. Also, note that in Table 5, in line with the literature, the coefficients on human capital measures are positive. In the last column of Table 4, I report the results of a single model including the different measures of impatience simultaneously with the large set of controls.

To sum up, since the model predicts that more impatient workers search more and many studies show that higher on-the-job search effort leads to higher job arrival rate, it is plausible to conclude that the key theoretical prediction is supported by the empirical results.[19]

4.4   Long-run and Short-run Impatience

Although there is evidence that variation in the proxies captures variation in short-run impatience rather than variation in long-run impatience (DV-P, 2005), by testing some theoretical predictions it is possible to provide further support for the hyperbolic model. Proposition 4 and 5 establish that when in the labor market commitment devices are not available, sophistication induces worker to exert a higher search effort. It is possible to test this prediction using the NLSY dataset that provides information on some other respondents' activities where a market for commitment devices clearly exists. As it will be clear this test also helps to distinguish hyperbolic discounting from exponential discounting, i.e. to assess whether the results are driven by variation in β or by variation in δ, which is the object of the second step of the empirical strategy.

There is a question in the NLSY 79 from which it is possible to construct a variable that reflects a degree of sophistication. In 1994 individuals are asked whether have any money assets in IRAs or Keough retirement accounts. These retirement programs are widely diffused in the US and unlike 401k plans for which participation is set sometimes by default, IRA and Keough require an active decision to save each year.[20] I construct a dummy equal to one if the individuals have some money assets in these retirement accounts and zero otherwise; I adjust this dummy for cumulative earnings and income and I term this variable as SOPH. These type of assets are perfect commitment devices for agents with self-control problems in consumption that believe they are "saving too little of their income". Agents investing in IRA accounts are penalized for early withdrawals (before the age of $59\frac{1}{2}$), and for borrowing against the content of the accounts (Benhabib and Bisin, 2005). IRAs or Keough retirement accounts are, indeed, illiquid assets and they have been introduced to encourage savings. I interpret

average measures in place of the dummy for occupation. While it is likely that there exists some kind of sorting, this did not affect the basic results that the measure of impatience has a strong and statistically significant coefficient.

[19]The NLSY 79 does not provide on-the-job search information (that however are difficult to measure and to observe) of employed workers in the sample period considered.

[20]Benjamin and Shapiro (2005) consider IRA and Keough participation as sophisticated decision making (in this case they estimate the effects of cognitive ability on the event of participation).

the SOPH as a proxy for some degree of sophistication. This interpretation relies on a simple intuition: when there are agents with self-control problems concerning consumption and when there are commitment devices to consume less, just those with some degree of sophistication demand such commitment devices.[21] Under the identifying assumption that workers do not have perfect commitment devices to stay in the same job, using the SOPH variable I can test the propositions of the model that state that sophisticated have higher job arrival rates.

Considering a model where preferences are exponential, the SOPH variable should be negatively correlated to long-run impatience (negatively correlated to $(1 - \delta)$). Indeed, it is reasonable to believe that agents with greater orientation toward the future invest money in retirement assets. Therefore, by estimating the job arrival rate the coefficient on the SOPH should be negative if variation in δ is driving the results obtained in the previous sub-section. If instead, according to the theoretical prediction of Proposition 4 and 5, it is the variation in β to drive the results, the coefficient on the SOPH should be positive. In Table 6 I report the results of several Cox models that include among the covariates the measures of impatience, the SOPH and a variable that reflects the interaction between SOPH and impatience. The coefficient on the SOPH is always positive and in most of the cases statistically significant at conventional levels. The effect of sophistication seems to be very large. A two standard deviation in the measure of SOPH leads to a between 14-26 percent increase in the job arrival rate. I performed the analysis by also including a variable that reflects the interaction between SOPH and impatience. The results on this variable are inconclusive (not reported). It is also worth noting that the SOPH is also significant when no measure of impatience is included in the regressions.

Finally, to the extent that cognitive ability improves decision making (e.g. see Benjamin and Shapiro, 2005), and that AFQT scores is positively associated to sophistication, additional evidence for the fact sophisticated are more mobile derives from the previous analysis. For example, in table 5 the AFQT score has a strong and highly significant impact on the job arrival rate. The coefficient on AFQT is always positive and highly significant whatever measure of impatience is included in the regressions and whatever definition of employment spell (only associated to wage increases or associated also to wage decreases) is considered.

4.5   Robustness Checks and Alternative Explanations

Definition of employment spell

The empirical results obtained concerned transitions into better paid jobs. Table 7 reports the coefficients on the measures of impatience from the Cox models that estimate the hazard rate of transitions associated also to wage decreases (i.e. in this case an employment spell is defined such that: 1) the difference between the start date of the new job and the stop date (in weeks) of the previous one is between zero and four, 2) the termination of the previous job is not due

---

[21]As, for example, O'Donoghue and Rabin (1999b) show, naive agents may procrastinate any financial operation that they believe to be beneficial for their welfare after retirement.

to reasons such as firing, laid off, plant closed, etc). As it is clear from Table 7, the results are very similar to the previous ones. It is important to note that also the coefficients on human capital variables (education and AFQT) remain positive and significant.

Size of the firm

The effects of the measures of impatience are large and strong even when I control for the size of the firm. In principle it possible that more impatient 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. The size of the firm is an indicator of the promotion opportunity for a job. Such a robustness check can be done for a small part of completed spells. The NLSY provides information on the size of the firm just for some years and for some jobs ("the CPS job" or the most recent job). Therefore, by performing such a robustness check a large number of observations is lost.[22] As an indicator of the firm size in the NLSY I take the variable "number of employee at location of current job". The results from this robustness check show still positive, large and significant positive effects of impatience on the job arrival rate.

Risk preferences

Since the measures for impatience may be correlated with low risk aversion, let me analyze if the results can be explained by heterogeneity in risk preferences.[23] It is problematic to address this point because it is not clear if staying is more risky than moving or vice versa. Hence, I consider both cases. First, suppose that staying is less risky than moving. The results in the first step of the empirical strategy might be explained by risk preferences (i.e. risk lovers have higher hazards rates). However, since savings in IRAs should also reflect a high degree of risk aversion, we should also expect savings in IRAs to be negatively correlated to hazard rates. Second, suppose that moving is less risky than staying. In this case the explanation based on risk preferences is consistent with the results on IRAs but it is not consistent with the results obtained in the first step (i.e. proxies are positively correlated to hazard rates).

Search costs

Variation in the proxies might reflect variation in search costs or taste for change. This explanation, however, is not consistent with the results of DV-P (2005). If this explanation was correct we should also expect these proxies to be positively correlated to higher exit rates from unemployment and to unemployed search effort (assuming that employed and unemployed workers

---

[22]In particular, all jobs the duration of which is less than one year, and as a consequence, all employment spells with duration less than one 52 weeks, do not enter the regressions. Therefore, the results are likely to be biased (in particular, for the particular type of loss of observations, we have a downward estimation of the effect of impatience on the job arrival rate).

[23]It should be noted, however, that scholars who attempted to proxy for low risk aversion in the NLSY used other variables. Fairlie (2002) uses questions on drug dealing in the NLSY and he argues that this variable is a proxy for attitudes toward risk, entrepreneurial ability, and preferences for autonomy. Running regressions including the measure used by Fairlie (2002) do not affect the coefficients of the proxies for impatience.

are not different in some relevant unobserved variables). Instead, in the NLSY DV-P (2005) obtain results in the opposite direction. Unlike search costs, the hypothesis that these proxies reflect short-run impatience is consistent with the results of DV-P (2005): coherently with the theoretical models an impatient worker searches more when he is employed and searches less when unemployed.

Mean of the wage offer distribution

Measures of impatience might capture differences in the mean of the wage offer distribution (with higher mean associated to higher measures). In this case this would explain the positive correlation between exit rates to better jobs and these measures. However, if this was the case, we should expect workers with higher levels of these measures (the movers) to have higher individual wage growth rates. As documented in section 2, strong empirical evidence documents that stayers have wage growth rates higher than those of the movers.

Human capital

The alternative explanation that unobserved low human capital accounts for the empirical results is rejected by the findings of the empirical works discussed in section 2. It is also rejected by the results of the empirical analysis that, in line with previous studies, show that higher level of education and AFQT score have a strong impact on the job arrival rate.

4.6 Impatience and Collaboration

The other main prediction of the theoretical model is that more patient workers collaborate and stay on the job by experiencing wage increases. In this subsection, to test this prediction I analyze the effect of impatience on the absence rate that recently has received attention in literature as an important indicator of collaboration (Flabbi and Ichino, 2001 and Audas et al., 2004). Absenteeism is 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 et al., 2004). A low absence rate is positively related to the probability of a promotion or to the probability of a wage increase. The NLSY allows to estimate the relationship between impatience and absenteeism.

In the NLSY there is a question in the period 1985-1993 on how many hours the respondent was absent at work during the survey week. The answer is reported for the most recent job held by the individual. In the context of this application, the coefficient of interest is on impatience that is constant over time. Therefore, I exploit the panel structure of the data to estimate the following random effects model:

$$y_{it} = X_{it}b + Imp_i\gamma + \nu_i + \epsilon_{it},$$

where $i = 1, 2, ...n.$ refers to individuals, $t = 1, 2, ...T.$ to the time period and y is the number of hours the workers was absent during the survey week. The set of controls (time variant and time invariant) and the explanatory variable of impatience are denoted by X and Imp, respectively, while ν is the time-invariant unobserved component and ϵ is a standard time variant

error term. Note that when we include among the regressors the dummy for occupation and industry, we control for unobserved characteristics that are valued differently across industry and occupation. The results of the model with a small set of controls and the large set of controls are reported in table 8. Bad health status is an important factor affecting absenteeism. The coefficient on impatience is positive and significant including in the regressions the large set of controls (second column of Table 8).[24] This means that more impatient workers have higher absence rates as predicted by the model. It is remarkable that by introducing more controls the coefficient on impatience increases and becomes significant. These results, however, must be interpreted cautiously. It could be possible, for example, that heavy drinking or drug use leads to absenteeism for reasons that are not addressed by the model.

5   Conclusion

Workers with more severe self-control problems initiate frequent job changes at the beginning of their career even if it is sub-optimal for their long-run selves. As emphasized in the introduction, the fact that the empirical results cannot be explained by low human capital gives further support to the model proposed. Many important results, indeed, run against the predictions for some of the unobservables. This distinguishes this paper from previous contributions, where it is difficult to disentangle the effect of hyperbolic discounting from unobservables such as low human capital and ability. A key theoretical prediction in the model is that sophisticated workers are even further away from the long-run optimum than their naive counterparts. As it is shown in the empirical section, this prediction allows also to provide evidence that hyperbolic time preferences rather than exponential discounting explain the overall pattern of results. Note that the technique used to distinguish the two types of time preferences is peculiar to the situation studied in this paper, i.e. short-term rewards and long-run costs. This suggests that more empirical research should focus on this kind of trade-offs to shed light on the empirical discrimination between exponential and hyperbolic discounting.

This paper has some implications for firms' hiring policies. Employed workers with high education and ability may be attractive but these workers may suffer from self-control problems on-the-job. This paper joins the nascent field of behavioral labor economics and it is in line with the works that show the importance of behavioral traits not related to cognitive skills for individuals' labor market outcomes (e.g. see Bowles, Gintis and Osborne, 2001).

---

[24]Similar results are obtained when the random effect model is estimated with a Poisson regression and with pooled OLS.

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.

Benhabib, J. and A. Bisin (2005). Modeling Internal Commitment Mechanisms and Self-Control: A Neuroeconomics Approach to Consumption-Saving Decisions, *Games and Economic Behavior*, 52, 460-492

Benjamin, D. and J. Shapiro (2005). Who is "Behavioral"? Cognitive Ability and Anomalous Preferences, mimeo, University of Chicago.

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

Blumen, I., Kogan, M., and McCarthy, P., (1955). *The Industrial Mobility of Labor as a Probability Process*. Cornell Univ. Press, Ithaca, NY.

Burdett, K., (1978). A Theory of Employee Job Search and Quit Rates, *American Economic Review*, 68, 212-220.

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.

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

Fairlie, R., (2002). Drug Dealing and Legitimate Self-Employment, *Journal of Labor Economics*, 20, 538-567.

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

Harris, C. and Laibson, D. (2004). Instantaneous Gratification, mimeo.

Heckman, J. and Rubinstein, Y. (2001). The Importance of Noncognitive Skills: Lessons from the GED Testing Program" *AEA Papers and Proceedings*.

Johnson, W. R., (1978). A Theory of Job Shopping, *Quarterly Journal of Economics*, 92, 261-277.

Jovanovic, B., (1979a). Job matching and the theory of turnover. *Journal of Political Economy*, 87 (5), 972 990.

Jovanovic, B., (1979b). Firm-specific capital and turnover. *Journal of Political Economy*, 87 (6), 12461260.

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

Light, A., (2005). Job mobility and wage growth: evidence from the NLSY79, *Monthly Labor Review*, 128, 33-39

Light, A. and Omori, Y., (2004). Unemployment Insurance and Job Quits, *Journal of Labor Economics*, 22, 1159-89.

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.

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

Lazear, E. P., (1979). *Personnel Economics*, The MIT Press, Cambridge, Massachusetts.

Loewenstein, G. and D. Prelec, (1992). Anomalies in Intertemporal Choice; Evidence and an Interpretation. *Quarterly Journal of Economics*, 107 (2), 573-597.

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), 1074-1078.

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, (1999a). Doing Now or Later, *American Economic Review*, 89(1), 103-133.

O'Donoghue, T. D. and M. Rabin, (1999b). Procrastination in Preparing for Retirement, in *Behavioral Dimensions of Retirement Economics*, Henry Aaron (eds) The Brookings Institution.

O'Donoghue, T. D. and M. Rabin, (2001). Choice and Procrastination, *Quarterly Journal of Economics*, 116, 121-159.

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.

Phelps, E. S., and Pollak, R., (1968). On Second-Best National Saving and Game-Equilibrium Growth, *Review of Economic Studies*, 35, 185 - 199.

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.

Royarlty, A. B., (1998). Job-to-Job and Job-to-Nonemployment Turnover by Gender and Education Level, *Journal of Labor Economics*, 16(2), 392-443.

Shapiro, J. (2005). Is there a Daily Discount Rate? Evidence from the Food Stamp Nutrition Cycle, *Journal of Public Economics*, 89, 303-325.

Thaler, R. H., (1981). Some Empirical Evidence on Dynamic Inconsistency, *Economics Letters*, 8, 201207.

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

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

A. 1 $c'(0) < H$; $c'(1) > K$

Derivation of the value function 1. At any period the worker employed at the wage w chooses s to solve:

$$\max_{s \in [0,1]} -c(s) + \lambda s \beta \delta \max \left\langle \int [V_0(w') + (1/\beta\delta)w'] dF(w') - V_0(w) - (1/\beta\delta)w, 0 \right\rangle +$$
$$+ \lambda(1-s)\beta\delta V_0(w^p) + (1-\lambda)\beta\delta[V_0(w) + (1/\beta\delta)w], \qquad (8)$$

where $V_0(w)$ is the lifetime utility of employment at wage w at period 0 from the perspective of self -1. A new job offer arrives with probability λs. If the worker accepts she receives the new wage in the same period and from the next period she experiences the continuation payoff $\int V_0(w') dF(w')$. A promotion arrives with probability $\lambda(1-s)$ the benefit of which start materializing in the next period with the continuation payoff $V_0(w^p)$. With probability $1 - \lambda$ the worker do not receive either a new job offer or a promotion.

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

Therefore, an employed worker moves to another job if and only if it pays a higher wage. This conclusion rests on the fact that all jobs offer the same opportunities in terms of wage growth. The objective function of the worker may be written as follows:

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

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 nor promoted neither 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 nor promotion neither 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 (2), (3) and (4) of self -1 are obtained using the same procedure. However, the difference with respect to the value function (1) is that for self -1 at time 0 the drop in discounting has already occurred so that β does not appear in expression (2), (3) and (4). An interior solution to the problem of the worker is guaranteed by assumption A1. From the f.o.c. (5) the following Lemma are derived.

Lemma 2 *There exists a $\epsilon > 0$ such that in the interval $[\underline{w}', \underline{w}' + \epsilon]$ the marginal benefit from search is decreasing in $w \in [\underline{w}', \underline{w}' + \epsilon]$.*

Assumption 5 *Assume that the individual is sufficiently impatient that at $w = \underline{w}'$ the marginal gain from search, i.e. the RHS of equation (5), is greater than zero. Assume also that a critical wage w^R such that the RHS of (5) is zero exists.*

Lemma 3 *The optimal level of search effort is strictly decreasing in $w \in [\underline{w}', w^*]$, where $w^* > w^R$. The optimal level of search effort is increasing in w for any $w \in [w^*, \overline{w}']$.*

Lemma 2, Lemma 3

The proofs of Lemma 2 and Lemma 3 are based on the study of the sign of the derivative of the marginal benefit from search with respect to the wage. Let me 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 that is less than one. It is immediate to see that for $w = \underline{w}'$ expression (10) is equal to -1 and that there exists an $\epsilon > 0$ such that, for $w = \underline{w}' + \epsilon$, (10) is negative. The first part of Lemma 3 derives from standard arguments: as the cost function is convex, an increase in the RHS of eq. (5) leads to an increase in the search effort. Given Assumption 5 and the fact that for $w' = \overline{w}'$ the RHS of eq. (5) equals zero, there must be an interval $[w' > w^R, \overline{w}']$ where the marginal benefit is negative and increasing. It follows that it is not possible to have a $w' > w^R$ such that the marginal gain from search is positive. Accordingly, most of job-to-job changes for jobs that pay a wage that lies in the left tail of the wage distribution.

Proposition 3 This proposition derives from the fact that current self does not agree on the search effort supplied by future self. Consider the value function (1) in period -1 and consider a small deviation from the optimal search effort for sophisticated agents. Formally, denote the value function (1) of self -1 with Γ and for the moment consider only the interval of the wage such that the marginal benefit from search is positive. 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. 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, 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 from Lemma 3.

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 $d\Gamma/ds < 0$. In other words,

at the equilibrium level for sophisticated it would be optimal to decrease the search effort supplied by future self of a negative amount exactly equal to $d\Gamma/ds$.

Now consider the range of the wage where the marginal benefit is negative and decreasing. The proposition is proved by recognizing that from the perspective of self -1, in this range the marginal benefit at time 0 for each wage is always lower than the marginal benefit from the perspective of self 0. The same argument applies to the range of wages for which the marginal benefit from search is increasing and negative.

Proposition 4 and 5 The proof of Proposition 4 uses expression (6). Let me again focus on the part multiplied by $\beta\delta$. The current self of a sophisticated cannot control the search effort supplied by future self. However, the current knows the effort level the future self will exert for each wage. This is reflected in the continuation values $V_0(w)$, $V_0(w^p)$ and $V'_0(w)$. 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 according to her current preferences. The naive self believes that s^{i+1} will be exactly the one an exponential agent (and herself) supplies in the future. Therefore, by Proposition 1, the s^{i+1} the naive takes into account for the maximization is lower than that the sophisticated considers. Hence, in order to prove Proposition 4 is sufficient to show that a higher s^{i+1} implies a higher effort. The previous proof has showed that a higher search effort supplied in the next period implies that the part multiplied by $\beta\delta$ decreases. However, since 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} considered in the maximization.

Appendix B

Consider the case of hyperbolic hyperbolic discounting and assume that time in the model is $t = 0, 1, 2$. A general functional form for the discount function was proposed by Loewenstein and Prelec (1992):

$$\phi(t) = (1 + \alpha t)^{-\beta/\alpha} \qquad \alpha, \beta > 0$$

In this specification impatience is decreasing in α and increasing in β. When $t = 0$ the function equals one, from $t > 0$ it decreases continuously as a hyperbola. It is possible to relax Assumption 4 in the following way. A worker that searches at time 0 can switch job at period 1 and in this case she earns the new wage at that period. A worker that collaborates at time 0 can be promoted at period 1 and in this case she earns the new wage at period 2. With this timing structure we still have that benefit from moving arrive before than benefit from promotion but we do not have that benefits from moving are "immediate". Using the discount function above, the f.o.c. in expression (5) for search is:

$$c'(s) = (1 + \alpha)^{-\beta/\alpha} \lambda \int_{w}^{\overline{w}'} (w' - w) dF(w') +$$

$$(1 + 2\alpha)^{-\beta/\alpha} \lambda \left\{ \int_{w}^{\overline{w}'} [V_0(w') - V_0(w)] dF(w') + [V_0(w) - V(w^p)] \right\} \tag{11}$$

It is easy to check that also in this case, the key theoretical predictions holds, i.e. higher impatience is associated to higher on-the-job search effort. In particular, a lower α (higher impatience) implies a higher

RHS of (11). Suppose that $\beta = 1$, $\alpha = 5$. We have that the short-run gain from search is discounted by $(1 + 5)^{-1/5} = 0.6988$, while the long-run cost from search is discounted by $(1 + 2 \cdot 5)^{-1/5} = 0.6190$. Consider now $\beta = 1$ and $\alpha = 2$. In this case the agent is more impatient. The short-run gain from search is discounted by 0.5773 while the long-run cost is discounted by 0.4472. From $\alpha = 5$ to $\alpha = 2$, a 21.04 perent drop in the evaluation of the short-term gain occurs ($0.6988/0.5773$-1). However, the drop in the evaluation of the long-run cost is 38.41 percent. In other words, for $\alpha = 2$ the RHS of (11) is higher than that for $\alpha = 5$. This result generalizes to all possible value of α, i.e. more impatient workers search more. An exception is when α tends to infinite. In this case the discount function drops immediately at time 0 and is equal from any periods onwards.

TABLE 1

Summary Statistics, Job Changes

Number of spells	13881	16914
Number of completed spells	7274	10307
Mean duration[*]	96.14	99.92
Duration 25th percentile	20	21
Median	48	50
Duration 75th percentile	112	118
Job changes in zero weeks (the difference between the start date of the new job and the stop date of old one is equal to zero)	2652	3888
job changes in 1 week	3025	4227
job changes in 2 weeks	782	1027
job changes in 3 weeks	443	628
job changes in 4 weeks	372	413
Spells by individuals:		
Stayers	2821	2821
movers with 1 completed spell	2808	2799
movers with 2 completed spells	1155	1470
movers with 3 completed spells	423	630
movers with 4 completed spells	140	317
movers with 5 completed spells	43	140
movers with 6 completed spells	14	65
movers with 7 or more completed spells	4	41

Note. − The employment spells reported refer to the period 1985-1996, for individuals who were not in school, and in the military sub-sample. Completed employment spells are defined as job changes such that 1) difference between the start date of the new job and the stop date (in weeks) of the previous one is between zero and four, 2) the new job is associated to a higher wage, 3) the termination of the previous is not due to reasons such as firing, laid off, plant closed, quit for pregnancy/family reasons, end of temporary or seasonal jobs, program ended. In column (2) job-to-job-transitions include also those that are not associated to wage increases (relaxing condition 2).

[*] The mean refers only to completed spells.

TABLE 2

TABLE 2

Summary statistics, Stayers and Movers

	stayers	movers (1 spell)	movers (2 spells)	movers (3 spells)	movers (4 spells)
Heavy drinking: in the last 30 days (1984) number of hangovers	-.0162 (1.158)	.0331* (1.171)	.0365** (.9387)	.0416* .7428	.1484** (1.266)
	-.0162 (1.158)	.0208 (1.093)	.0279* (1.010)	.0639** .8782	.0910** (1.069)
Smoking before the employment spells in 1984	-.0923 (.9735)	-.0265** (.9934)	.0417** (1.009)	.0766** (1.015)	.1541** (1.025)
	-.0923 (.9735)	-.0549* (.9854)	.0291** (1.010)	.0953** (1.018)	.0853** (1.016)
Bank Account not having a bank account before the employment spells	-.0475 (.0977)	-.0403 (1.001)	-.0436 (1.025)	.0729** (1.025)	.0746** (1.035)
	-.0475 (.0977)	-.0390 (.9969)	-.0425 (1.013)	-.0163 (1.023)	.0487** (1.045)
Drug Use number of time respondent used cocaine before 1988	-.0812 (.9065)	.0281** (1.011)	.1134** (1.090)	.2719** (1.183)	.3730** (1.321)
	-.0812 (.9065)	.0202** (1.024)	.0793** (1.051)	.1863** (1.122)	.2911** (1.222)
Life Insurance no life insurance at the job before the employment spells	-.0507 (.9973)	.0807** (1.051)	.1062** (1.065)	.0871** (1.069)	.2261** (1.118)
	-.0507 (.9973)	.0808** (1.052)	.0859** (1.054)	.1088** (1.079)	.1270** (1.074)
NLS Assessment measure of impatience during the interview (1981-1985)	-.0450 (.9973)	.0390** (1.810)	-.0586 (.8751)	-.1251* (.6164)	.0134** (.9307)
	-.0450 (.9973)	.0040** (.9977)	.0093** (1.018)	-.0342 (.9247)	-.1475** (.6101)
Years of education	13.03 (2.51)	12.96 (2.42)	12.84** (2.29)	12.72** (2.39)	12.48** (1.98)
	13.03 (2.51)	12.98 (2.43)	12.89** (2.42)	12.70** (2.31)	12.59** (2.13)

TABLE 2

(Continued)

	stayers	movers (1 spell)	movers (2 spells)	movers (3 spells)	movers (4 spells)
Mother's years of education	10.69	11.00**	10.97**	10.86*	11.34**
	(3.29)	(3.15)	(3.06)	(3.26)	(2.78)
	10.69	10.97**	11.06**	10.94**	10.78**
	(3.29)	(3.20)	(3.11)	(2.97)	(3.30)
Father's years of education	10.80	11.11**	11.03**	11.16**	11.36**
	(4.02)	(3.99)	(3.76)	(3.89)	(3.44)
	10.80	11.05**	11.08**	11.15**	10.88
	(4.02)	(3.98)	(3.94)	(3.87)	(3.71)
AFQT score	39.77	40.76	40.82**	38.69	36.89**
	(28.70)	(28.56)	(27.69)	(27.40)	(26.82)
	39.77	41.09**	40.87*	39.46	39.18
	(28.70)	(28.47)	(27.95)	(27.68)	(27.26)

Note. − Entries are means (standard deviations in parenthesis). Each measure of impatience has mean zero and standard deviation one in the entire population. For each variable, the first two rows refer to employment spells associated only to wage increases, the second two rows refer to spells associated *also* to wage decreases.
*: Significantly different from the measure of stayers at the 0.1 level.
**: Significantly different from from the measure of stayers at the 0.05 level.

TABLE 3

Factor Analysis, Aggregate Measure of Impatience

	Factor Loadings	Uniqueness	Score coefficient
Alcohol	.2127	.9791	.1149
Drug Use	.6051	.5805	.5355
NLS Assessment	.0972	.9791	.0510
Bank Account	.2695	.6681	.2072
Life Insurance	.0454	.9979	.0233
Smoking	.3571	.8636	.2124

Note. − Entries in the table represent maximum likelihood estimates for a factor analysis model with one factor retained.

TABLE 4

Cox Models

	No controls	Small set of controls	Large set of controls	Impat. measures included simultaneously
Aggregate	.1498**	.1170**	.1079**	
Measure	(.0120)	(.0137)	(.0179)	
of Impatience	[10351]	[10351]	[6952]	
Alcohol	.0188*	.0274**	.0120	.0063
Number of hangovers	(.0101)	(.0101)	(.0113)	(.0127)
in the past 30 days (1984)	[12293]	[12293]	[8745]	[6979]
Drug Use	.0950**	.0943**	.0831**	.0670**
number of times respondent	(.0109)	(.0131)	(.0165)	(.0181)
used cocaine before 1988	[11420]	[11420]	[7656]	[6979]
NLS Assessment	.0226**	.0133	.0254)	.0332*
Measure of impatience	(.0114)	(.0136)	(.0171)	(.0192)
during the interview	[12685]	[12685]	[8542]	[6979]
Bank Account	.0824**	.0550	.0655*	.0564**
no bank account	(.0114)	(.0134)	(.0167)	(.0192)
before employment spells	[12898]	[12898]	[8685]	[6979]
Life Insurance	0340**	-.0209*	-.0161	−.0147
no life insurance	(.0106)	(.0116)	(.0146)	(.0163)
	[12684]	[12684]	[8525]	[6979]
Smoking	.1147**	.0715**	.0710**	.0452**
smoked before	(.0117)	(.0128)	(.0162)	(.0192)
the employment spells (1984)	[12293]	[12293]	[8745]	[6979]

Note. − Robust standard errors in parenthesis, number of observations in brackets. Entries in the first, the second, and the third columns are coefficients on the standardize measures of impatience from separate Cox proportional hazards models. The first column reports results of the coefficients on impatience in models estimated without other control variables. In the second column, in addition to the singles measures of impatience, the control variables include in the models are sex, race, wage, age, AFQT scores. In the third the results are reported from a model estimated with a large set of controls (see Table 5). Observations with missing values for any of the control variables were discarded.
*: Significantly different from 0 at the 0.1 level.
**: Significantly different from 0 at the 0.05 level.

35

TABLE 5

Cox Model with the Large Set of Covariates

Controls	$Coeff.$	$Std.Err.$	$Coeff.$	$Std.Err.$
Aggregate measure of impatience	.1170**	.0137	.1079**	.0179
Age	-.1954**	.0036	-.1853**	.0049
Wage	-.0013**	.0000	-.0010**	.0001
Sex Dummy	.1706**	.0291	.1395**	.0410
Race Dummy	$-$.0439	.0325	-.0368	.0425
AFQT score	.0059**	.0006	.0044**	.0008
Education	-		.0427**	.1080
Fathereduc	-	-	.0026	.0039
Mothereduc	-	-	-.0056	.0053
Marital Status	-	-	.1016**	.0365
Health status	-	-	.0256	.0355
Dummy Occupation-Professional	-	-	omitted	
Dummy Occupation-Manager	-	-	.1433**	.0716
Dummy Occupation-Saleworker	-	-	.2239**	.0889
Dummy Occupation-Clerical	-	-	.0132	.0686
Dummy Occupation-Craftsmen	-	-	-.1469	.0841
Dummy Occupation-Operatives	-	-	.1311*	.0794
Dummy Occupation-Laborers	-	-	.0771	.0953
Dummy Occupation-Farmers	-	-	-.3638	.4538
Dummy Occupation-Foremen	-	-	.0204	.2468
Dummy Occupation-Service	-	-	.2478**	.0727
Dummy Industry-Agriculture	-	-	omitted	
Dummy Industry-Private	-	-	.1964	.2097
Dummy Industry-Mining	-	-	-.1072	.2137
Dummy Industry-Construction	-	-	.3808**	.1520
Dummy Industry-Manufacturers	-	-	-.0387	.1436
Dummy Industry-Transports	-	-	-.0217	.1533
Dummy Industry-Wholesale Trade	-	-	.1589	.1403
Dummy Industry-Finance	-	-	-.0062	.1495
Dummy Industry-Business Services	-	-	.3346**	.1473
Dummy Industry-Personal Services	-	-	.0680	.1512
Dummy Industry-Entertainment	-	-	.1926	.1799
Dummy Industry-Professional Services	-	-	.0455	.1564
Dummy Industry-Public Adm.	-	-	-.3722**	.1609
Number of Observation	10351		6952	

Note. $-$ All the entries in first and the third columns represent coefficients on the variables from Cox proportional hazards models. The second and the fourth column report robust standard errors. In the third column the equation includes also geographical dummies, a dummy for urban status and dummies for central city or SMSA residence. Observations with missing values for any of the control variables were discarded..
*: Significantly different from 0 at the 0.1 level.
**: Significantly different from 0 at the 0.05 level.

TABLE 6

Cox Models with SOPH

	Coeff.	Std. Err.	Coeff.	Std. Err.
	small set of controls		large set of controls	
Impatience	.1145**	.0171	.1036**	.0224
SOPH	.0994*	.0581	.1261*	.0674
Observations	[7707]		[5101]	
Drug Use	.1005**	.0165	.0829**	.0211
SOPH	.0887*	.0513	.1190**	.0602
Observations	[7927]		[5563]	
Bank account	.0481**	.0175	.0590**	.0218
SOPH	.0872	.0559	.1148*	.0689
Observations	[9504]		[6289]	
NLSY Assessment	.0219	.0187	.0274	.0252
SOPH	.1078**	.0483	.1244**	.0557
Observations	[9353]		[6191]	
Alcohol	.0349**	.0129	.0195	.0144
SOPH	.1011**	.0472	.1224**	.0552
Observations	[9557]		[6613]	
Smoking	.0582**	.0166	.0606**	.0202
SOPH	.0933*	.0484	.1126**	.0577
Observations	[9557]		[6313]	
Life insurance	−.0376**	.0148	−.0114	.0178
SOPH	.0909*	.0478	.1135**	.0551
Observations	[9339]		[6167]	

Note. − Entries in the first and the third columns are coefficients from separates Cox hazard models. The second and the fourth column report robust standard errors. All equations include a variable that reflects the interaction between impatience and SOPH. The small set of controls include wage, sex, race, AFQT score and age. The large set include the variables reported in Table 5 in addition to geographical dummies, a dummy for urban status and dummies for central city or SMSA residence. The SOPH is a dummy equal to one if the respondent have an IRA account and zero otherwise and it is adjusted for cumulative earnings and income. Observations with missing values for any of the control variables were discarded.
*: Significantly different from 0 at the 0.1 level.
**: Significantly different from 0 at the 0.05 level.

TABLE 7

Cox Models, Spells Associated also to Wage Decreases

	No controls	Small set of controls	Large set of controls
Aggregate	.1533**	.1319**	.1144**
Measure	(.0107)	(.0117)	(.0148)
	[12600]	[12600]	[8573]
Alcohol	.0185*	.0163	.0034
	(.0098)	(.0109)	(.0123)
	[15882]	[15822]	[10826]
Drug Use	.0984**	.0982**	.0796**
	(.0097)	(.0110)	(.0132)
	[13912]	[13912]	[9052]
NLS Assessment	.0236**	.0147	.0146)
	(.0102)	(.0131)	(.0140)
	[15741]	[15471]	[10593]
Bank Account	.0871**	.0697**	.0646*
	(.0103)	(.0115)	(.0137)
	[15699]	[15699]	[10741]
Life Insurance	.0332**	-.0118*	−0.0253*
	(.0094)	(.0104)	(.0150)
	[15439]	[15439]	[10544]
Smoking	.1162**	.0872**	.0880**
	(.0100)	(.0110)	(.0163)
	[15882]	[15882]	[10826]

Robust standard errors in parenthesis. Number of observations in brackets. The entries in first, the second, and the third columns represent coefficients on the measures of impatience from separate Cox proportional hazards models. Employment spells here comprise also job-to-job transitions associated to wage decreases. The first column reports results of the coefficients on impatience in models estimated without control variables. In the second column the control variables included in the models are sex, race, wage, age, and educational attainment. In the third the results are reported from a model estimated with a large set of controls (see Table 5). Observations with missing values for any of the control variables were discarded.
*: Significantly different from 0 at the 0.1 level.
**: Significantly different from 0 at the 0.05 level.

TABLE 8

Impatience and Absenteeism

Controls	Coeff.	Std.Err.	Coeff.	Std.Err.
Aggregate Impatience	.0107	.0089	.0241**	.0097
Age	-.0014	.0027	-.0024	.0030
Wage	.0001	.0002	.0001	.0002
Sex Dummy	.0049	.0184	.0263	.0226
Race Dummy	−.0221	.0205	-.0021	.0223
AFQT	.0011**	.0003	.0008*	.0004
Health Status	.0799**	.0203	.0803**	.0205
Observations	50134		44892	
Groups	8119		8044	

Note. − Entries in first and the third columns are coefficients from random-effects models. The entries in the first and the third columns refer to the model estimated with a small set (AFQT, sex, race, wage, age, health) and the large set of coefficients, respectively. Observations with missing values for any of the control variables were discarded.
*: Significantly different from 0 at the 0.1 level.
**: Significantly different from 0 at the 0.05 level.