WAGE DYNAMICS AND UNOBSERVED HETEROGENEITY:
TIME PREFERENCE OR LEARNING ABILITY?

Lalith Munasinghe
Nachum Sicherman

Wage Dynamics and Unobserved Heterogeneity: Time Preference or Learning Ability?
Lalith Munasinghe and Nachum Sicherman
NBER Working Paper No. 11031
January 2005
JEL No. J22, J24, J31

ABSTRACT

A large portion of the variation in wages and wage growth rates among individuals is due to "unobserved" heterogeneity, and the source of individual heterogeneity is typically attributed to data limitations and/or the unobservability of certain productivity related factors. In this paper we develop a test that discriminates between two inherently unobservable sources of heterogeneity (both of which can clearly account for the variation in wages and wage growth rates): learning ability and workers' inter-temporal preferences (discounting). We apply this test to the large observed differences in wages and wage growth rates between smokers and non-smokers. The evidence supports the discounting hypothesis.

Lalith Munasinghe
Department of Economics
Barnard College
Columbia University
3009 Broadway
New York, NY 10027
lm25@columbia.edu

Nachum Sicherman
Graduate School of Business
819 Uris Hall
Columbia University
New York, NY 10027
and NBER
ns38@columbia.edu

1 Introduction

Despite several decades of research in labor economics and dramatic improvements in data collection, a substantial fraction of individual distributions of wage levels and wage growth rates still remains unexplained. As a consequence, the study of unobserved sources of individual heterogeneity is important not only for the well-known purpose of correctly estimating wage effects of key observed factors such as schooling, but also for our understanding and assessment of some of these less tractable determinants of wages.

The obvious and most commonly cited source of heterogeneity appeals to differences in individual productivity. Although information on productivity related characteristics is widespread – from intelligence test scores to detailed schooling information – various other dimensions of productivity and learning ability, including drive, motivation, discipline, focus, charisma and communication skills among others, are inherently unobservable to the econometrician. A second source of heterogeneity is related to preferences. For example, differences in the relative valuation of earnings with other on-the-job and off-the-job amenities are likely to affect occupational and job choices, and thus give rise to earnings differentials. While this type of heterogeneity has been extensively discussed in the literature (e.g. Rosen 1972), another important source of preference related heterogeneity – i.e., differences in individual time preference – has received much less attention.

The crux of the issue is that unobserved variations in either learning ability or time preference can account for the unexplained variation in wage levels and wage growth rates. So the question is whether the source of the unexplained variation in wage dynamics is due to unobserved learning ability or time preference. The objective of this paper is to present a simple theory of wage dynamics and derive a discrimating test to empirical assess the relative importance of these two potential sources of unobserved heterogeneity in wage determination over the life cycle.

The discriminating test we propose consists of a two-part strategy. The first part is to

find a variable that is highly correlated with wage levels and especially with wage growth rates (net of other observable effects), but where this observed correlation is not causal. Such a correlation raises the possibility that this candidate variable is a likely proxy for some unobserved factor such as learning ability or time preference that is directly related to wage dynamics. The second part of the strategy is to design a discriminating test to address whether this candidate variable is indeed a proxy for learning ability or time preference.

Our candidate variable is whether people smoke or not. Smoking is highly (negatively) correlated with wages and wage growth. Smokers earn less than non-smokers and have substantially lower wage growth rates. While it is possible that some of the differences between smokers and non-smokers are due to direct effects (e.g., health), it is more likely that these differences are due to other variables that are jointly correlated with smoking and wage growth. Indeed, the estimated difference in wage growth rates across smokers and non-smokers is reduced substantially after including a rich set of variables such as schooling and test scores. However, the net differences in wage growth rates still remain large and highly significant.[1] We contend that time preference and learning ability are two likely sources of unobserved heterogeneity that could account for the net gap in wage growth rates across smokers and non-smokers.

The first explanation is that individuals differ in their rates of time preference, and those who discount the future more heavily will clearly weigh their current wages (first wage) relative to their future wages (wage growth) more heavily than those who discount the future less.[2] Since discount rates are likely to affect all forms of investments, individuals

---

[1] Smokers in the U.S. earn substantially less than non-smokers. For example, Levine et. al. (1997), using the National Longitudinal Surveys of Youth (NLSY), find that smokers earn 11% and 17% less than non-smokers in 1984 and 1991, respectively. After controlling for a host of individual and family characteristics, this wage gap reduces to 4.2% and 6.9%. Using the same NLSY data, we find that the major source of this wage gap is the dramatic differences in wage growth rates between smokers and non-smokers. In another study (Munasinghe and Sicherman, 2003) we report that over a period of about ten years non-smokers experience, on average, wage growth rates that are about 60% greater than those of non-smokers. Controlling for an extraordinary rich set of controls, the differences are reduced substantially, to around 20%, but remain large and significant.

[2] Since individual discount rates are defined in terms of current versus future consumption and not in

2

with high discount rates will be less likely to make other investments, including investing in their health. Hence the discounting hypothesis is that individuals with high discount rates will be more likely to smoke and more likely to have flatter wage profiles.[3] The second explanation is that individuals differ in their learning ability, and that more efficient (able) learners are likely to invest more in schooling as well as in other forms of human capital, including job training. As a consequence, efficient learners will have a higher first wage and a steeper wage profile. If efficient learners are also less likely to smoke because their higher learning ability leads them to better understand the negative effects of smoking then it is possible that this unobserved dimension of learning ability could be the culprit behind the observed negative correlation between smoking and wage growth rates. In sum, both hypotheses lead to the same prediction, namely, a negative correlation between smoking and wage growth rates.

The discriminating test we propose in this paper is based on theoretical implications of time preference versus learning ability on the trade-off between the first wage and subsequent wage increases, two key parameters that describe individual careers. In particular, we show that the correlation between the first period wage and wage growth rate across smokers and non-smokers differ markedly under the two alternative hypotheses. To preview our modeling results: First, holding learning ability constant, the first period wage and wage growth rate are negatively correlated because individual discount rates are unobserved. More importantly, however, if smoking is correlated with individual discount rates

_____

terms of current versus future incomes, we implicitly assume that workers face borrowing constraints against returns on investments in human capital. It should be noted, however, that even if the capital market is perfect, the returns on an investment in schooling, for example, depend on hours of work if schooling raises market productivity by a larger percentage than it raises non-market productivity. Individuals who are more future-oriented desire relatively more leisure at older ages. Therefore, they work more at younger ages and have a higher discounted marginal benefit on a given investment than persons who are more present oriented.

[3]The idea that smoking is a proxy for discount rates is extensively documented in the economics literature (Fuchs, 1982). Empirical studies find correlations between smoking and various other behaviors related to future outcomes, including health status, educational attainment, earnings levels, use of seat belts, physical exercise, and brushing and flossing teeth (Hersch and Viscusi 1990; Hersch 1996; Levine et. al. 1997; Hersch 2000; Viscusi and Hersch 2001).

then this negative trade-off between the first period wage and wage growth rate is larger among smokers than among non-smokers. Second, if non-smokers are simply more efficient learners and the distribution of discount rates is similar across smokers and non-smokers, then the observed negative trade-off will typically be smaller among smokers than among non-smokers. As a consequence, the discounting and learning hypotheses predict a different sign on the interaction term between smoking and the first wage in a wage growth equation.

In our model of career choice individuals face a trade-off between the first wage and wage growth rate as depicted in Figure 1. The technology of on-the-job human capital production is based on the following considerations. Increase in training time and effort (and the corresponding reductions in first period wages) increases future productivity and thus future wages. However, the fixity of the human body and initial level of human capital would clearly imply that this increase occurs at a diminishing rate. Hence, all else constant, the production possibility frontier, or more specifically, the constraint function between the first period wage and wage growth rate will be concave to the origin. Thus the negative slope of this constraint will decrease in absolute value – i.e. a decreasing marginal rate of transformation – as more and more resources are diverted from current wages to training that enhances future wages (and thus the wage growth rate). We use this relatively simple two-period model of career choice to study the ramifications of the discounting and learning hypotheses.

Under the discounting hypothesis, the key assumption of the model that generates the negative interaction effect between smoking and first-period wages is diminishing returns to training or learning – i.e. the fact, as mentioned above, that the constraint function between the first period wage and wage growth rate is concave to the origin. Note that for a given production constraint, the optimal choice point on this constraint will be determined by an individual's time preference. Put differently, heterogeneity of time preference will simply trace this non-linear production constraint. The first model implication is of course that holding initial human capital constant, the first period wage will be negatively correlated

4

with the wage growth rate. Moreover, low discount workers will locate in the flatter region of the constraint where the ratio of second period wage increase to first period wages are relatively high, and high discount workers will locate in the steeper region of the constraint where the ratio of second period wage growth to first period wages is relatively lower. If smokers have relatively higher discount rates than non-smokers then the observed negative correlation between first period wages and second period wage increases will be larger among smokers than among non-smokers (holding initial level of human capital constant). These details are graphically depicted in Figure 2.[4]

On the other hand, if the source of unobserved heterogeneity is differences in learning ability then workers with different learning abilities will locate on different production possibility frontiers. Since we assume that all workers start their careers with the same level of human capital, the production frontier of the more able workers will be higher – the same first period wage intercept but a higher second period wage increase intercept – and thus steeper than the production frontier of their less able counterparts. Therefore, if non-smokers are more able learners than their smoker counterparts then the observed negative correlation between first period wages and second period wage increases will be larger among non-smokers. The key point is that the implications for the observed trade-off between the first wage and wage growth across smokers and non-smokers are reversed under the two alternative hypotheses.

In the empirical section of the paper we test whether the negative relationship between the first wage and wage growth rate is stronger among smokers or non-smokers. Our empirical strategy is to first delineate for each individual the two key variables: the wage at the time of full time entry into the labor market (first wage) and subsequent wage increases over the working career (wage growth rate). The latter variable is based on a coefficient estimate of time in labor market by running a wage regression for each individual in our data sample. We then implement wage growth regression models that include the first wage, smoking, and an interaction of first wage and smoking as the three key explanatory variables. The

---

[4]The specific parameters in Figure 2 pertain to the modeling details in Section 2.

5

sign on this interaction term constitutes a discriminating test. To anticipate our empirical findings: The estimated coefficient of the interaction between smoking and first wage is negative and highly significant. This finding is robust across various model specifications, and lends support to the discounting hypothesis.

Before proceeding to the details of the paper an important caveat needs to be addressed. Our empirical test raises the question of whether we have properly identified smoking as a proxy for individual discount rates. Since each individual in our sample contributes to only one observation in our regression framework, the negative coefficient of the interaction term is based on a cross section of observations. The observed correlation between first wage and wage growth across smokers and non-smokers could in principle be due to other sources of heterogeneity since we do not trace the same individual making choices about first wage and wage growth at different points in time. However, given that each individual represents a single career, this identification problem is not simply a data issue, but rather arises from a conceptual consideration that each individual represents one overall career observation. Although we acknowledge that our discriminating test is not conclusive in the sense that the trade-off between first wages and wage growth is estimated via a cross section of individuals, we hope that it is at least a first step towards getting insight into the intractable issue of what type of unobserved heterogeneity might be at the root of wage differentials.

The remainder of the paper is arranged as follows. The next section presents a model of wage dynamics and derives the discriminating test to address whether the observed correlations between smoking and wage growth is due to time preference or learning ability. In Section 3 we present the empirical analysis. We begin with a description of the NLSY data and an outline of the estimation framework. We then present the findings from the wage growth regressions and show that the estimated coefficient of the interaction between smoking and first wage is negative across various model specifications and wage growth metrics. Section 4 concludes with a brief summary.

6

2 A Theory of Wage Dynamics

Non-smokers have substantially higher wage growth rates compared to their smoker counterparts. Since it can be argued that smoking is a possible correlate of either individual discount rates or low learning ability and because both the discount rate and learning ability are unobserved determinants of wage growth, the observed disparity in wage growth rates across non-smokers and smokers as mentioned earlier in the introduction is consistent with two alternative hypotheses: smoking is a proxy for discount rates or smoking is a proxy for learning ability. To resolve this identification problem, we present a two period intertemporal model of wage dynamics where wages are determined by the interaction of the individual discount rate and learning ability. On the basis of this model, we derive a discriminating test to address whether smoking is a proxy for discounting or learning ability. The model follows the logic of Rosen (1974). In the first (investment) period workers can sacrifice current wages in order to enhance their productivity (and thus wages) in the second (post investment) period. With a few additional assumptions, individual preferences (i.e. the discount rate) will then determine the optimal combination of current and future wages along the production possibility frontier (which itself will be determined by training efficiency). Hence wage dynamics are modeled as a function of individual discount rates and training efficiency.

2.1 Model

Consider the wage possibilities of a worker who lives for two periods. If the worker makes investments in human capital in the first period, she will enhance her future productivity at the expense of her current productivity. As a result, assume that the trade-offs in productivity in the two periods can be summarized by the following production possibilities curve:

$$\Delta W = \frac{\alpha \gamma}{2} - \frac{\gamma}{2\alpha} W^2. \tag{1}$$

7

Assuming that workers are paid their productivity, interpret ΔW as the wage increase in the second period and W as the first period wages. In the absence of training, first period and second period wages are equal and given by α, and wage increase is equal to zero, i.e., $\Delta W = 0$. Hence we interpret α as the wage potential corresponding to a fixed, initial level of human capital, and we assume throughout that it is the same across all workers.[5] The γ parameter is correlated with the rate of transformation of first wage into future wage increases (ΔW) by diverting resources from current production (thus decreasing W) into training that enhances future productivity.[6] Higher values of γ imply more efficient training, and the assumed functional form implies concavity of the production possibility curve to highlight decreasing returns to training.[7] We note once again that our discriminating test (derived below) hinges on this extremely weak and uncontroversial assumption about diminishing returns to training.

We now turn to worker preferences. Assume that workers face borrowing constraints, and thus W and ($W + \Delta W$) mimic consumption in the first period and second period, respectively. Since the resource constraint given in (1) above is concave to the origin, we present analyses with both linear and more standard non-linear Cobb-Douglas utility functions. First consider the following linear utility function:

$$U = W + \beta \Delta W, \tag{2}$$

where β is the time discount factor and hence higher values of β imply more future oriented-ness.[8]

---

[5]In the model we abstract from considering heterogeneity of initial levels of human capital. In the empirical analysis we of course control for many observed dimensions of human capital. Our focus in this theory section is on isolating the effects of unobserved factors such as time preference and learning ability on wage dynamics.

[6]This assumption is similar to Rosen's (1972) idea that different jobs are characterized by the different learning opportunities they offer.

[7]Note that the marginal rate of transformation (MRT) is equal to $\frac{\gamma}{\alpha}W$. Clearly MRT is increasing in γ; it is also increasing in W, which implies diminishing returns to training. Note, as training increases, W decreases and so does the MRT.

[8]Note $\beta = 1/(1 + r)$, where r is the rate of time preference for the present.

The first-order condition for utility maximization implies that the slope of the production possibilities curve (marginal rate of transformation of W for ΔW) equals the marginal rate of substitution (MRS) of W for ΔW in consumption:

$$MRT_{W,\Delta W} \equiv \frac{\gamma}{\alpha}W \equiv \frac{1}{\beta} \equiv MRS_{W,\Delta W}.$$

Hence

$$W \equiv \frac{\alpha}{\gamma\beta}, \tag{3}$$

and by substituting (3) into (1) we obtain

$$\Delta W \equiv \frac{\alpha\gamma}{2} - \frac{\alpha}{2\gamma\beta^2}. \tag{4}$$

Thus (3) and (4) are the solutions to the constrained utility maximization problem in terms of first and second period wages, respectively. Solutions to different discount rates are graphically depicted in Figure 2.[9] As can be seen from Figure 2, heterogeneity of discount rates simply traces the constraint. The same analysis is shown in Figure 3 with the exception that we now include a second constraint to distinguish between high ability and low ability workers.

First note, since our focus is on wage growth, the partial derivatives of ΔW with respect to β and γ are both positive:

$$\frac{\partial \Delta W}{\partial \beta} \equiv \frac{\alpha}{\gamma\beta^3} > 0, \text{ and } \frac{\partial \Delta W}{\partial \gamma} \equiv \frac{\alpha}{2} + \frac{\alpha}{2\gamma^2\beta^2} > 0.$$

These basic comparative static exercises imply that wage growth of non-smokers relative to their smoker counterparts can be higher on account of either non-smokers having higher βs on average or because non-smokers have higher γs on average. Thus the finding that wage growth is higher among a sample of non-smokers than it is among a sample of smokers is

---

[9]Modifying the utility function such that ΔW represents total second period consumption (instead of just increases in second period consumption) will not alter any of the analytical results.

consistent with either hypothesis.[10] So the question is whether this model of wage dynamics also implies a further discriminating test between the discounting hypothesis and learning ability hypothesis.

2.2 Discounting or Learning Ability: A Discriminating Test

First consider the correlation between W and ΔW when β (time preference) is unobserved, and α (initial human capital) and γ (learning ability) are assumed to be constant. Clearly, the observed W and ΔW simply trace the possibility frontier given by equation (1). Thus a regression of W on ΔW will imply a negative correlation between wage growth and first period wages.[11] The question is whether this negative relationship is likely to differ across smokers and non-smokers if smoking is a proxy for individual discount rates. From Figure 2 the answer is obvious: the slope of the production possibility frontier where smokers locate is relatively steeper. Thus the negative trade-off between first period wages and wage growth is larger for smokers than non-smokers. A slightly more formal argument is presented below.

Suppose under the discounting hypothesis that smoking is a "limited" proxy for individual time preference. For simplicity, assume that $S^* = \lambda/\beta$, where S^* denotes smoking propensity, λ some constant, and β the time discount factor. However, we only observe whether an individual smokes or not (hence the sense in which smoking is a "limited" proxy). Hence let $S = 1$ if $S^* \geq k$ and $S = 0$ otherwise. So S is the smoking indicator variable and k some

---

[10]The simple models depicted in Figures 2 and 3 imply that non-smokers will have lower first period wages, a prediction that is not supported by the data. Note, however, we have assumed that the initial stock of human capital is the same across high and low discount workers and across high ability and low ability workers, respectively. That is, we have assumed a constant α across all workers. If however low discount workers (or high ability workers) have made greater pre-market investments then it is likely that these workers may indeed have relatively higher stocks of unobserved skills (since our regression analysis controls for various observed human capital characteristics). Hence if unobservable dimensions of α are higher for low discount and high ability workers then it is indeed possible that they will on average also experience higher levels of first period wages as we see in the data.

[11]The possible objection that W is endogenous and thus should not be included as a RHS variable in a wage growth regression is somewhat artificial in the context of this model. Note that both W and ΔW are a function of β from (3) and (4), respectively. However, since β is unobserved but W is observed, from (3) it is obvious that W is an ideal proxy for β. Thus the inclusion of W in an wage growth regression could be interpreted precisely as a proxy variable for unobserved β.

constant. For illustrative purposes suppose that β is uniformly distributed, i.e., $\beta \in [0,1]$. From the definition of S it implies that $\beta^S \in [0, \beta^*]$ and $\beta^{NS} \in [\beta^*, 1]$, where $\beta^* = \lambda/k$. That is, smokers will have lower βs and non-smokers will have higher βs, respectively. Since β determines both W and ΔW, smokers and non-smokers will locate on the bottom-right region and top-left region, respectively, of the possibility frontier given by (1). Given the concavity of (1) the negative relationship between W and ΔW will be lower in absolute terms for a sample of non-smokers relative to a sample of smokers. Hence, under the hypothesis that smoking is a proxy for discounting, the observed negative trade-off between first period wages and future wage increases will be relatively less for non-smokers. See the illustrative thick solid and dashed lines in Figure 2 that identify the simple correlation between W and ΔW for non-smokers and smokers, respectively.

Now consider the case where smoking is a proxy for learning ability holding the distribution of discount rates among high ability and low ability workers identical. The objective is to isolate any possible effect of learning ability on the observed trade-off between first period wages and wage growth. Given that the overall slope of the production possibility frontier is steeper for the high learning ability workers than it is for the low ability workers, Figure 3 might suggest that the negative trade-off should be relatively larger for high ability workers. However, given our linear utility function the correct answer is that the observed trade-off between first period wage and wage growth is the same across high and low ability workers. To see this more formally, note that the observed trade-off is the ratio of the change in earnings growth to the change in first period earnings, where both changes are due to a change in β. At the limit this is nothing but the partial derivative of ΔW with respect to β divided by the partial derivative of W with respect to β. Therefore at the limit the trade-off between W and ΔW due to a small change in β is given by:

$$\frac{\partial \Delta W / \partial \beta}{\partial W / \partial \beta} \equiv \frac{\alpha/\gamma\beta^3}{-\alpha/\gamma\beta^2} \equiv -\frac{1}{\beta}$$

Hence the trade-off under the linear utility specification is only a function of the discount

rate and not also of the training ability. Recall, that under the assumption that smoking is a proxy for discounting, this negative trade-off will be larger for smokers (because they are presumed to have lower βs) as we claimed earlier; but if smoking is a proxy for learning ability, then there is no difference in the observed negative trade-off between first period wages and second period wage growth across smokers and non-smokers. This "no difference" result is a consequence of the linear utility function, and as the next example illustrates, with convex utility functions the observed negative trade-off is predicted to be larger among non-smokers under the learning hypothesis. The point however is that under the two alternative hypotheses the prediction in the difference in this trade-off between smokers and non-smokers is reversed, implying that we have a discriminating test.

Next we examine this same trade-off using a more standard Cobb-Douglas utility specification. Consider the following utility function:

$$U = W \Delta W^{\beta}. \tag{5}$$

Again β can be interpreted as the discount factor, where a higher value of β implies more future oriented-ness. Of course this type of utility function is more realistic in the sense that unlike linear utility functions, the trade-off an individual is willing to make between future wage growth and current wages is now a function also of the ratio of future wage growth to current wages. The solutions to this constrained utility maximization problem are given by:

$$W = \frac{\alpha}{\sqrt{2\beta + 1}}. \tag{6}$$

and by substituting (6) into (1), we obtain

$$\Delta W = \frac{\alpha \gamma \beta}{2\beta + 1}. \tag{7}$$

The trade-off between first period wages and wage growth due to a change in β is given by:

$$\frac{\partial \Delta W / \partial \beta}{\partial W / \partial \beta} = \frac{\alpha \gamma / (2\beta + 1)^2}{-\alpha / (2\beta + 1)^{3/2}} = -\frac{\gamma}{\sqrt{2\beta + 1}}.$$

12

If we hold learning ability constant and assume that smoking is a proxy for discount rates then the negative trade-off will clearly be stronger for smokers (lower β means a larger negative value in the above expression). On the other hand, if smoking is a proxy for low learning ability and the distribution of preferences is similar across high and low ability workers, then the negative trade-off will be smaller for smokers (lower γ means a smaller negative value in the above expression) than for non-smokers. Figures 4 and 5 present these results in a graphic format under the assumption that smoking is a proxy for discounting and low ability, respectively. In each diagram, the thick solid and dashed lines identify the simple correlation between W and ΔW for non-smokers and smokers, respectively. Note that the slope of the negative thick solid line is flatter than the slope of the thick dashed line in Figure 4, and that this result is reversed in Figure 5.

Thus the sign of the coefficient of an interaction term between smoking and first period wages (in a wage growth regression) constitutes as a discriminating test for whether smoking is a proxy for discounting or whether it is a proxy for learning ability.

3   Empirical Analysis

3.1   NLSY Data

We use the National Longitudinal Surveys of Youth (NLSY) data from 1979 to 1994. The data contain information about smoking behavior of the respondents in their late teens and early twenties. The panel nature of the data and the fact that we observe the entire early work histories of the vast majority of our respondents allow us to directly correlate smoking behavior with individual earnings over the first decade or so of their careers. In addition, the NLSY data contain rich information on a variety of individual, family, geographic and work related characteristics. In the appendix we have a detailed description of the NLSY data.

Our construction of individual wage growth rates exploits the panel nature of the NLSY

by running, for each individual in our sample, a simple OLS wage regression with time since first entering the labor market as the independent variable. We interpret the time coefficient as an estimate of the individual average wage growth rate and implement it as the dependent variable in our wage growth analyses. In the individual OLS regressions we specify the wage rate metric both as the real wage and as the log of real wage.[12]

It should be pointed out that although we use a longitudinal data set, effectively we only observe one career per person. This is a drawback because a complete identification of our model requires multiple observations per person. It is difficult, however, to think of a data set in which individuals are observed making more than one choice concerning their "careers", and thus making more than one choice of a wage profile (i.e., initial wage as well as wage growth rate).

3.2 Estimation Framework

In this section we implement the discriminating test suggested by the model, using the NLSY data set. The test requires the inclusion of an interaction term between first wage and smoking in a wage growth regression (in addition to first wage and smoking as separate variables). Therefore, we estimate the following model:

$$\Delta W_i = X_i \alpha + \alpha_1 S_i + \alpha_2 W_i + \alpha_3 S_i W_i + \varepsilon_i,$$

where ΔW_i is a measure of individual i's wage growth, X_i is a vector of individual characteristics, S is a smoking indicator, W is the individual's first wage, and ε_i is individual i's error term. The estimated coefficient of the parameter α_3 will constitute the discriminating test between the two sources of unobserved heterogeneity. Our prediction of course is that $\partial \Delta W_i / \partial W < 0$, implying that α_2 is negative. But the discriminating test is whether this

---

[12]Standard wage equations of course only use the log specification. The reason for this practice is the universal implementation of Mincer's human capital earnings function. In the context of our paper, where what matters is the perception of wage increases, it is not evident a priori which measure of wage growth is more relevant. Therefore, we present evidence using both measures. In a previous version of the paper we also used nominal wages and its log counterpart and duplicated the same qualitative results.

negative relation between the first wage and wage growth is stronger for smokers or non-smokers – i.e. whether the sign on α_3 is negative or positive. Time preference implies that it should be negative, and learning ability implies that it should be reversed and positive.

3.3 Results

3.3.1 Descriptive Statistics

Table 1 presents smoking rates (reported in 1984) for a select group of occupations. In our entire sample 38% were classified as smokers in 1984.[13] The variation in this gross smoking rate is quite dramatic with "Maids and Housemen" at the high end with a smoking rate of 62%, and "Teachers in Elementary Schools" at the low end with less than 10%. These rates are comparable to estimates found in other studies.[14]

Table 2 presents some descriptive statistics for smokers and non-smokers. These gross mean characteristics are strikingly different. For example, educational attainment is substantially higher for the sample of non-smokers.[15] Non-smokers on average have over one and a quarter more years of completed schooling, are much more likely to have a high school diploma, and score about 10 points higher on the Armed Forces Qualifying Tests (AFQT). The differences in labor market outcomes are even more dramatic. Non-smokers earn higher wages, have more "net" labor market experience, and have lower turnover rates, than smokers. Not surprisingly, a relatively smaller percent of non-smokers report health as a limiting factor to the amount and kind of work they could do.[16]

---

[13] We define a smoker as someone who smokes at least one cigarette per day on average. This definition is more or less comparable to the definition of "current smoker" used by the US Center for Disease Control (See the data appendix for further details). Similar rates have been reported by others (Evans and Montromery (1994), Levine, Gustafson and Velenchik (1997)).

[14] Although the detailed rankings of occupations by smoking has varied over the years (Nelson et.al. 1994), they are broadly similar to ours, where blue collar workers are at the top, and white collar workers are at the bottom, where teachers and sales representative have the lowest smoking rates (Bang and Kim, 2001)

[15] This is a well-documented finding (see Sander 1995).

[16] It is possible that individuals having health problems will select into jobs where their health problems do not limit their work, thus resulting in an underestimation of health problems due to smoking. We hope that the use of the word "could" in the survey instrument reduces such a problem. (See data appendix for the exact wording of the survey instrument.) For additional evidence on the health limitations of smoking

The gross hourly pay is substantially lower for smokers compared to non-smokers. On average, the non-smoker wage premium is over 15%. The differences in first wage and wage growth rates provide further insight into the overall wage disparity between smokers and non-smokers. The mean first wage is lower for smokers, but this difference (7%) is not as substantial as the difference in overall wages. However, the substantial difference in the increases in hourly wages (.34 versus .21) represents a huge wage growth premium of over 60% between smokers and non-smokers, and thus suggests that wage growth differentials are largely responsible for the well-documented fact that smokers earn less than non-smokers.[17]

3.3.2 Regression Results

In Table 3 we report our findings from a variety of wage growth models. We present the coefficient estimates of smoking, first wage, and interaction of smoking and first wage, from the first to the third row of each panel, respectively. The columns from left to right include different control variables: in the first column we do not include any additional variables, and in the second and third columns we include the listed variables cumulatively. In panel I we use the first real wage and in panel II we use the log of first real wage.

The estimated coefficient of the interaction between smoking and first wage (third row of each panel) is negative and highly significant across all specifications. Note in particular that adding a host of control variables, including completed years of schooling and AFQT scores,[18] only seems to strengthen this negative interaction effect. This finding is of course consistent with the time preference hypothesis and inconsistent with the learning hypothesis.

---

on the job see Leigh (1985 and 1986).

[17]If smoking is a proxy for discounting then our measure of wage growth may in fact underestimate the "effective" wage growth rate for non-smokers. Since non-smokers are more likely to have a College degree and thus have considerably higher debt, their wages during the early part of careers will overstate their disposable income. Hence the "effective" wage growth differential between non-smokers and smokers is likely to be even higher than what is suggested by our measured wage growth differences. Similar considerations suggest that the "effective" first wage differences between non-smokers and smokers will be smaller than the gross differences presented in Table 2. These considerations, however, are unlikely to lead to any first order bias of our regression coefficients since we control for completed years of schooling.

[18]If we consider schooling time as part of the trade-off, the gap between smokers and non-smokers will be much larger, given the lower levels of schooling among smokers.

Although this conclusion is unsurprising given that the inclusion of AFQT scores (or the ten ASVAB scores) failed to nullify the negative effect of smoking on wage growth, the robustness of the interaction effect is strong evidence in support of the time preference hypothesis, and a rejection of the hypothesis that an unobserved dimension of learning ability is the culprit behind the observed correlation between smoking and wage dynamics.

We conclude this section with a few noteworthy comments related to some of the other coefficient estimates. First, the correlation between first wage and wage growth is negative as predicted by dynamic equilibrium models of compensation. Second, the positive coefficient estimate of smoking should not be interpreted as a reversal of our earlier finding of a negative correlation between smoking and wage growth. In the context of including an interaction term between first wage and smoking, the positive coefficient on smoking simply says that if the first wage is zero then smokers are predicted to have a higher future wage than their non-smoker counterparts. Of course this hypothetical exercise is out of sample since we exclude all wage observations below a threshold of \$2 (in 1987 dollars) from our analysis. However, the broader implication of this result may be quite consistent with the time preference hypothesis.

A very low first wage is likely to yield greater dis-utility for the present-oriented (smokers) than for the future-oriented (non-smokers), and thus the compensating future wage may in fact be relatively higher for smokers than for non-smokers at very low levels of current wages. None of these considerations, however, negate our key finding that an increase in first wage leads to a higher wage growth penalty for smokers compared to non-smokers.[19]

We attempt to summarize our key findings in the two diagrams given in Figures 6 and 7. Figure 6 simply highlights the fact that smokers have lower and flatter wage profiles based on our analysis of the smoking "effects" on first wage and wage growth. In Figure 7 the graph displays the negative interaction effect of smoking and first wage on wage growth.

---

[19] Of course in our data we do not observe individuals actually making these trade-offs between first wage and wage growth since each individual represents a single data point in all our regressions. Hence our conclusions are clearly based on cross sectional evidence.

17

The Y-axis is wage growth and the X-axis is first wage. The negative relationship shows the predicted trade-off between first wage and wage growth. The negative interaction term is illustrated by the fact that this negative relationship is steeper for smokers than it is for non-smokers. Note that the Y-intercept is higher for smokers than non-smokers (as implied by the positive smoking coefficient), but the two lines intersect at a relatively low first wage of approximately $4.50. We also summarize in Figure 7 our earlier findings depicted in Figure 6 – both, lower average first wage and lower average wage growth among smokers – to illustrate the consistency of our separate analyses. It is clear from Figure 7 that our earlier findings are entirely compatible with either a steeper or flatter negative relationship between first wage and wage growth among smokers.

4   Conclusion

Much of the variation in wages across individuals remains unexplained despite the inclusion of a host of observed variables. This unexplained variation in wages is due either to the lack of data or to the fact that some relevant individual characteristics are inherently unobservable. The question we address in this paper is simple but challenging: Is it possible to say whether unexplained wage growth differentials across individuals is due to one type of unobserved heterogeneity or another? We consider time preference (discounting) and learning ability as two possible sources of unobserved heterogeneity that could in principle explain wage growth differences, and develop a test to address whether wage growth differentials are due to time preference or unobserved learning ability. In particular, we argue on the basis of evidence from our discriminating test that the significant differences in wage growth rates between smokers and non-smokers are due (unobserved) differences in the mean discount rate between smokers and non-smokers.

Our findings raise the possibility that a variety of psychological and personality traits are likely to be more important than simple cognitive ability to process information in explaining

18

differences in labor market outcomes across individuals. As a consequence research endeavors related to the various dimensions and interactions of what we refer to in economic parlance as the "individual discount rate," are likely to provide insights not only to our understanding of career choice, investment in human capital, and wage determination, but more broadly to our understanding of the distribution of wealth, health, and happiness.

5 Data Appendix: National Longitudinal Surveys of Youth, 1979-94

Our data is from the NLSY. This is a panel of 12,686 youth, aged 14-21 in 1978, and sampled continuously since 1979. Our sample includes data up to 1994. We include individuals in our sample when they first report that their "main activity" is "working." Therefore, our "First Wage" variable is recorded accordingly. A key variable of our analysis is whether, in 1984, respondents answered affirmatively to whether they smoked or not. We classify a smoker (in 1984) as a person who at least smoked one cigarette per day on average. Out of 12,584 who responded, 38.8% were classified as smokers. Similar, but not exact, questions were also asked in 1992. A smoker was classified as someone who "smokes daily" (as opposed to "occasionally" or "not at all"). Of the 8,341 who responded in 1992, 28.92% were classified as smokers. We utilized only the 1984 smoking questions for the following reasons: By 1992 the number of individuals that answered the smoking survey dropped dramatically due to sample attrition. Only 7,822 individuals answered the 1992 survey and had valid wage growth measures. More important, the sample attrition was not random. Individuals who smoked in 1984 were more likely to drop out, and even worse, those with lower wage growth were substantially more likely to drop out.

Our principal dependent variable is a wage growth measure for each individual constructed over the first several years in the labor market. We construct a wage growth coefficient for each individual by estimating a wage regression as a function of time only. This estimated coefficient is our measure of individual wage growth rates. For wages we use two measures of hourly payments, real wages (in 1987 dollars) and their natural logarithm. In an earlier version of the paper we also used nominal wages. Since none of our results was affected by the use of nominal wages, we limit the analysis here to real wages. We consider wage reports to be valid only if nominal pay is between $2 and $200. Given the construction of our wage growth measure, a minimum of two valid observations per individual is required

in order to be included in the sample.

Below we discuss the construction of several key variables used in the regressions.

Health. Respondents in the NLSY were asked, in each survey, the following two questions: (1) "hltamt" - whether health limited the amount of work you could do since last survey ("(are you/would you be) limited in the kind of work you (could) do on a job for pay because of your health?"), and (2) "hltknd" - whether health limited the kind of work you could do since last survey ("(are you/would you be) limited in the amount of work you (could) do because of your health?"). Using the answer to these two questions we constructed several other additional health measures: (1) "evera" - if a person ever reported hltamt=1; (2) "everk" - if a person ever reported hltknd=1; (3) "mhlta" - % of times reported hltamt=1; and (4) "mhltk" - % of times reported hltknd=1. We experimented with all these measures but report the regression results using "mhltk" only. None of the results was affected by using any of the other alternative measures of health.

Schooling. We use the respondents report on "highest grade completed" to construct six schooling dummies: 8 years or less, 9-11 years, 12 years, 12-15 years, 16 years, and 17+ years. We report only the results using the dummy variables. Replacing the schooling dummies with the continuous measure didn't affect our results.

AFQT Scores. During the summer and fall of 1980, NLSY79 respondents participated in an effort of the U.S. Departments of Defense and Military Services to update the norms of the Armed Services Vocational Aptitude Battery (ASVAB). The Department of Defense and Congress, after questioning the appropriateness of using the World War II reference population as the primary basis for interpreting the enlistment test scores of contemporary recruits, decided in 1979 to conduct this new study. NLSY79 respondents were selected since they comprised a pre-existing nationally representative sample of young people born during the period 1957 through 1964. This testing, which came to be referred to as the "Profile of American Youth," was conducted by NORC (National Organization for Research at the University of Chicago) representatives according to standard ASVAB procedure guidelines;

21

respondents were paid $50 for their participation. Groups of five to ten persons were tested at more than 400 test sites, including hotels, community centers, and libraries throughout the United States and abroad. A total of 11,914 civilian and military NLSY79 respondents (or 94 percent of the 1979 sample) completed this test: 5,766 or 94.4 percent of the cross-sectional sample, 4,990 or 94.2 percent of the supplemental sample, and 1,158 or 90.5 percent of the military sample.

The ASVAB consists of a battery of 10 tests that measure knowledge and skill in the following areas: (1) general science; (2) arithmetic reasoning; (3) word knowledge; (4) paragraph comprehension; (5) numerical operations; (6) coding speed; (7) auto and shop information; (8) mathematics knowledge; (9) mechanical comprehension; and (10) electronics information. A composite score derived from select sections of the battery can be used to construct an approximate and unofficial Armed Forces Qualifications Test score (AFQT) for each youth. The AFQT is a general measure of trainability and a primary criterion of enlistment eligibility for the Armed Forces. The creation of this percentile score, called AFQT89, involves: (1) computing a verbal composite score by summing word knowledge and paragraph comprehension raw scores; (2) converting subtest raw scores for verbal, math knowledge, and arithmetic reasoning; (3) multiplying the verbal standard score by two; (4) summing the standard scores for verbal, math knowledge, and arithmetic reasoning; and (5) converting the summed standard score to a percentile.

Religion. We use two indicators for religious affiliation and frequency of attending religious services: (1) dummy variables indicating whether the person was raised Protestant, Baptist, Episcopalian, Lutheran, Methodist, Presbyterian, Roman Catholic, or Jewish. (2) The frequency in which respondents attended religious services (never, several times a year, about once a month, three times a month, about once a week, or more than once a week). In the regressions' results reported in the paper we include both measures: religious affiliation and frequency of attendance.

Neighborhood Income. This variable was constructed using the 1980 census of population

22

and later matched to the NLSY sample. It was calculated to describe the expected value of a person's neighborhood per capita income, given their race (white, black, or Hispanic), education level (less than 8 years, 9-11 years, 12-15 years, or 16+ years), and the county they live in, using data from the 1980 Census. Neighborhoods are defined as block groups in the 1980 Census, so each block group's per capita income was calculated by dividing aggregate household income for the block group by the total number of persons living in that block group. (Both of these were variables given to us by the census.) Since the 1980 Census does not include Hispanic as one of their race categories, but rather as an ethnic category separate of race (Hispanic persons were placed into one of five race categories: white, black, Asian, Native American or other.), we needed to approximate the number of Hispanics, non-Hispanic whites and non-Hispanic blacks living in each block group. We did this by generating a beta for each race, which would represent the percent of Hispanic persons that are of that race, and multiplying it by the number of Hispanic persons living in the block group. This would approximate the number of Hispanic persons that were recorded as belonging to each race, so that these categories could be amended. (This process was done for each block group and for each education level. It was repeated several times, replacing the betas, for precision.) The outcomes of this process were numbers for Hispanic, non-Hispanic white, and non-Hispanic black persons in each education level for each block group. "Neighborhood Income" was then found by creating the expected value of a person of race X and education level Y for each county.

References

[1] Barsky, Robert, F. Thomas Juster, Miles S. Kimball, and Matthew D. Shapiro, "Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Study," The Quarterly Journal of Economics, May 1997, pp: 537-580.

[2] Bang, K. M., and Kim, J. H. Prevalence of Cigarette Smoking by Occupation and Industry in the United States. American Journal of Industrial Medicine, 40, September 2001, 233-239.

[3] Evans, William N., and Edward Montgomery, "Education and Health: Where There's Smoke There's an Instrument," NBER working paper No. 4949, December 1994.

[4] Fuchs, Victor R (1982). "Time Preference and Health: An Exploratory Study," in Economic Aspects of Health. edited by V.R. Fuchs, Chicago; University of Chicago Press, 1982.

[5] Hersch, Joni. "Smoking, Seat Belts, and Other Risky Consumer Decisions: Differences by Gender and Race," Managerial and Decision Economics, Vol. 17, 471-481 (1996)

[6] Hersch, Joni. "Gender, Income Levels, and the Demand for Cigarettes," Journal of Risk and Uncertainty. vol. 21, no. 2. 2000: 263-282.

[7] Hersch, Joni. and W. Kip Viscusi. "Cigarette Smoking, Seatbelt Use, and Differences in Wage-Risk Tradeoffs'. The Journal of Human Resources. vol. 25, no. 2, Spring 1990.

[8] Jamner, Larry D., Laura Cousino Klein, Carol K. Whalen, and Judi Profant, "Nicotine, Gender, and Time Preference. Working paper, Department of Psychology and Social Behavior, School of Social Ecology, University of California, Irvine, 1998.

[9] Leigh, J. P. "An empirical analysis of self-reported, work-limiting disability." Med Care, 23(4), 1985, 310-319.

[10] Leigh, J. P. Correlates of Absenteeism. Human Relations, vol. 39, 1986, 81-100.

[11] Levine, Phillip B., Tara A. Gustafson, and Ann D. Velenchik, "More Bad News for Smokers? The Effects of Cigarette Smoking on Labor Market Outcomes" Industrial and Labor Relations Review, vol. 50, no. 3, April 1997.

[12] Loewenstein, George, and Nachum Sicherman. "Do Workers Prefer Increasing Wage Profiles?." Journal of Labor Economics Vol. 9, No. 1, February 1991.

[13] Munasinghe, Lalith, and Nachum Sicherman. "Why Do Dancers Smoke? Smoking, Time Preference, and Wage Dynamics." Working paper, 2003.

24

[14] Munasinghe. Lalith, "Wage Growth and the Theory of Turnover". Journal of Labor Economics, vol.18 (April 2000): 204-220.

[15] Nelson, D. E., Emont, S. L., Brackbill, R. M., Cameron, L. L., Peddicord, J., and Fiore, M. C. Cigarette smoking prevalence by occupation in the United States. A comparison between 1978 to 1980 and 1987 to 1990. J. Occup. Med., May 1994, 516-525.

[16] Rosen. Sherwin. "Learning and Experience in the Labor Market". The Journal of Human Resources, Volume 7. Issue 3 (Summer 1972), pp: 326-342.

[17] Rosen, Sherwin. "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition" The Journal of Political Economy, Vol. 82. No. 1. (Jan. - Feb., 1974), pp. 34-5 5.

[18] Sander, William, "Schooling and Smoking" Economics of Education Review, vol. 14, no. 1, pp: 23-33, 1995.

[19] Viscusi, W. Kip, and Joni Hersch, "Cigarette Smoking as Job Risk Takers", The Review of Economics and Statistics, May 2001, 83(2): 269-280.

Table 1: Smoking Rates by Select Occupations
NLSY 1984

Occupation	Smoking Rate	Observations
All	.376	9,501
Maids and Housemen	.621	66
Roofers	.600	40
Kitchen Workers	.566	53
Waiters	.554	233
Heavy Truck Drivers	.509	106
Laborers and Construction	.476	504
Carpenters	.451	82
Janitors and Cleaners	.413	305
Housekeepers and Butlers	.386	171
Truck Drivers (light)	.365	126
Teachers (n.e.c.)	.349	43
Cashiers	.348	391
Sales Workers, other Commodities	.325	351
Secretaries	.283	297
Athletes	.226	31
Computer Programmers	.208	101
Teachers, Elementary Schools	.094	53

Note: Occupations are randomly selected across the spectrum of smoking rates with the exception that we omit occupations with very low sample sizes.

Table 2: Descriptive Statistics across Smokers and Non-smokers
NLSY 1979-1994

	Non-Smokers	Smokers
Variables		
Age (in 1984)	23.49	23.66
Non White	.43	.37
Sex (% males)	.49	.54
Married/with Spouse	.438	.376
Schooling	13.3	11.9
High School Diploma	.86	.66
AFQT (1989)	45.3	34.7
Father Schooling	9.7	9.1
Mother Schooling	10.4	9.9
Health limit type of work	.035	.045
Net years of experience (in 1994)*	11.5	10.4
Potential years of experience (in 1994)*	14.1	15.8
Hourly (nominal) pay (79-94)	13.29	11.54
Changed employer since last interview (79-94)	.354	.431
Quit Job (79-94)	.245	.288
Was Laid-off (79-94)	.068	.086
Was fired (79-94)	.016	.031
First wage (real)	5.91	5.55
First wage (nominal)	4.77	4.36
Increase in hourly wage (nominal)	.67	.49
Increase in hourly wage (1987 dollars)	.34	.21

* "Net" refers to actual years of experience, while "potential" experience is calculated as (age-schooling-6). Note: Number of observation varies across variables. For fixed individual characteristics the number is about 6700, and for means taken over the whole 79-94 period, the number is about 50,000 valid observations. All mean differences are significantly different from zero, using 99% confidence level (the t-test was performed assuming equal variance).

Table 3: Wage Growth Regressions

OLS Specification including Interaction between First Wage and Smoking

(Standard Errors in Parentheses)

	No Controls	Limited Controls[1]	Full Controls[2]
Panel I			
Smoking (S)	0.117	0.245	0.181
	(.036)	(.035)	(.070)
First real wage (FRW)	-.0.062	-0.074	-0.075
	(.002)	(.002)	(.002)
SxFRW	-0.049	-0.053	-0.054
	(.006)	(.005)	(.005)
Panel II			
Smoking (S)	0.186	0.188	0.144
	(.086)	(.083)	(.099)
ln of first real wage (FLRW)	-0.372	-0.590	-0.608
	(.030)	(.030)	(.030)
SxFLRW	-0.204	-0.141	-0.151
	(.051)	(.049)	(.050)

1. Limited Controls include age, race, sex, schooling, and AFQT scores.

2. Full Controls include in addition to the above also health status, average per capital neighborhood income, religion affiliation, and frequency of attendance of religious services. See the Data Appendix for a more detailed description of all variables.

Figure 1

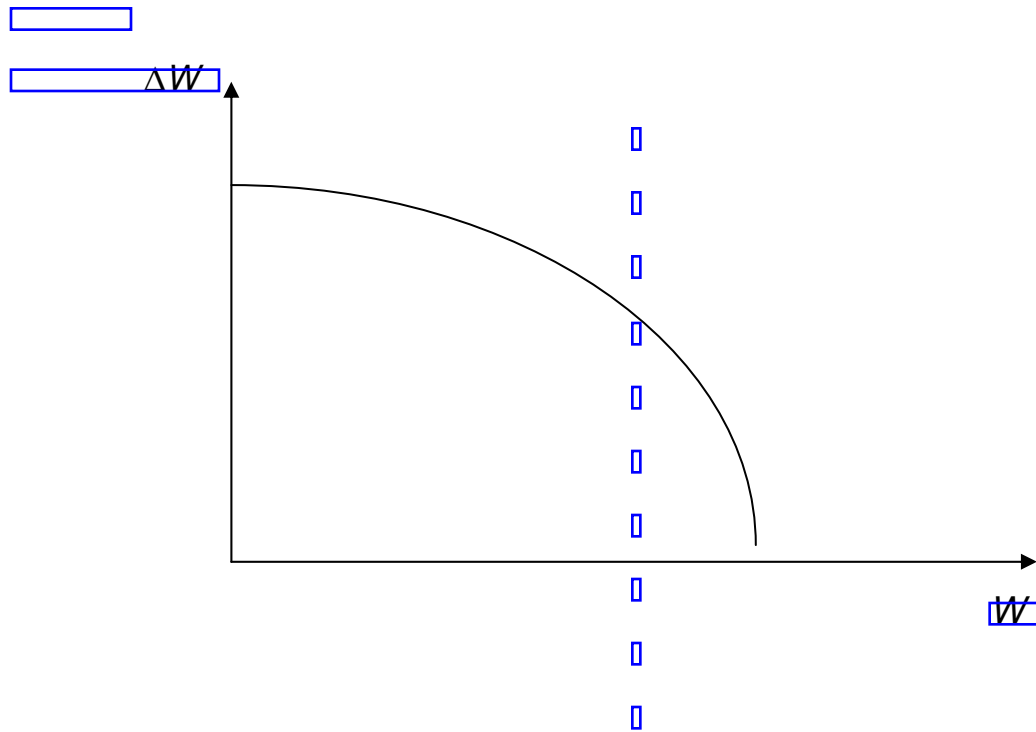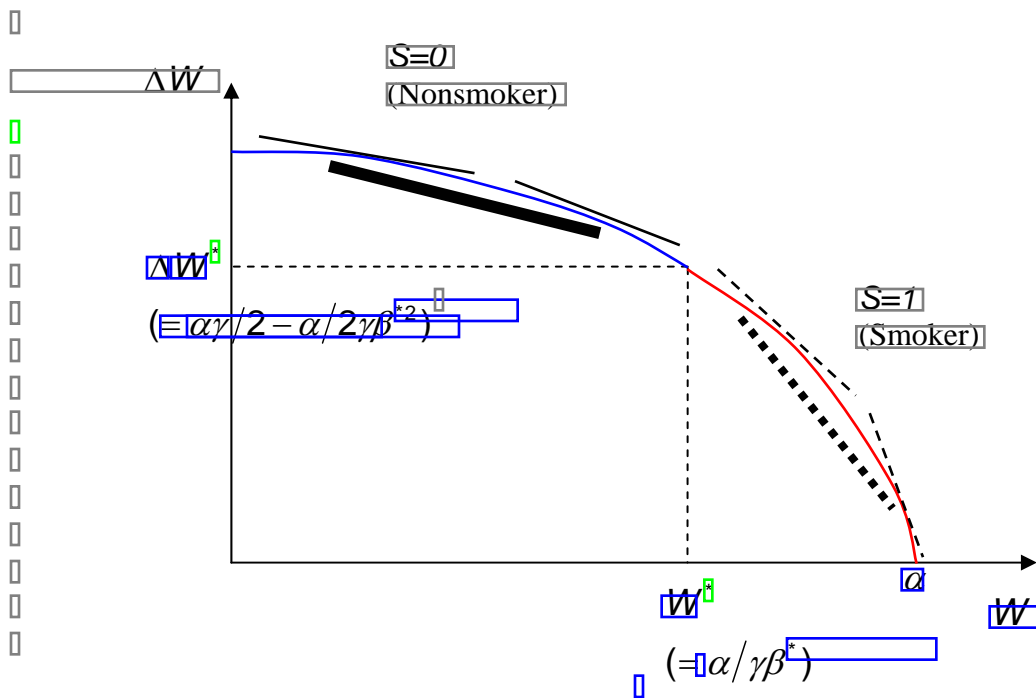**Career Choice: Trade-off between First Wage and Wage Growth**

ΔW

W

Figure 2

ΔW

$S=0$
(Nonsmoker)

$S=1$
(Smoker)

ΔW^*

$(= \alpha\gamma/2 - \alpha/2\gamma\beta^{*2})$

W^*

$(= \alpha/\gamma\beta^*)$

α

W

Figure 3

ΔW

$S=0$
(Nonsmoker)

$\frac{\alpha}{2\gamma}$

$\frac{\alpha}{2\gamma}$

$S=1$
(Smoker)

α W
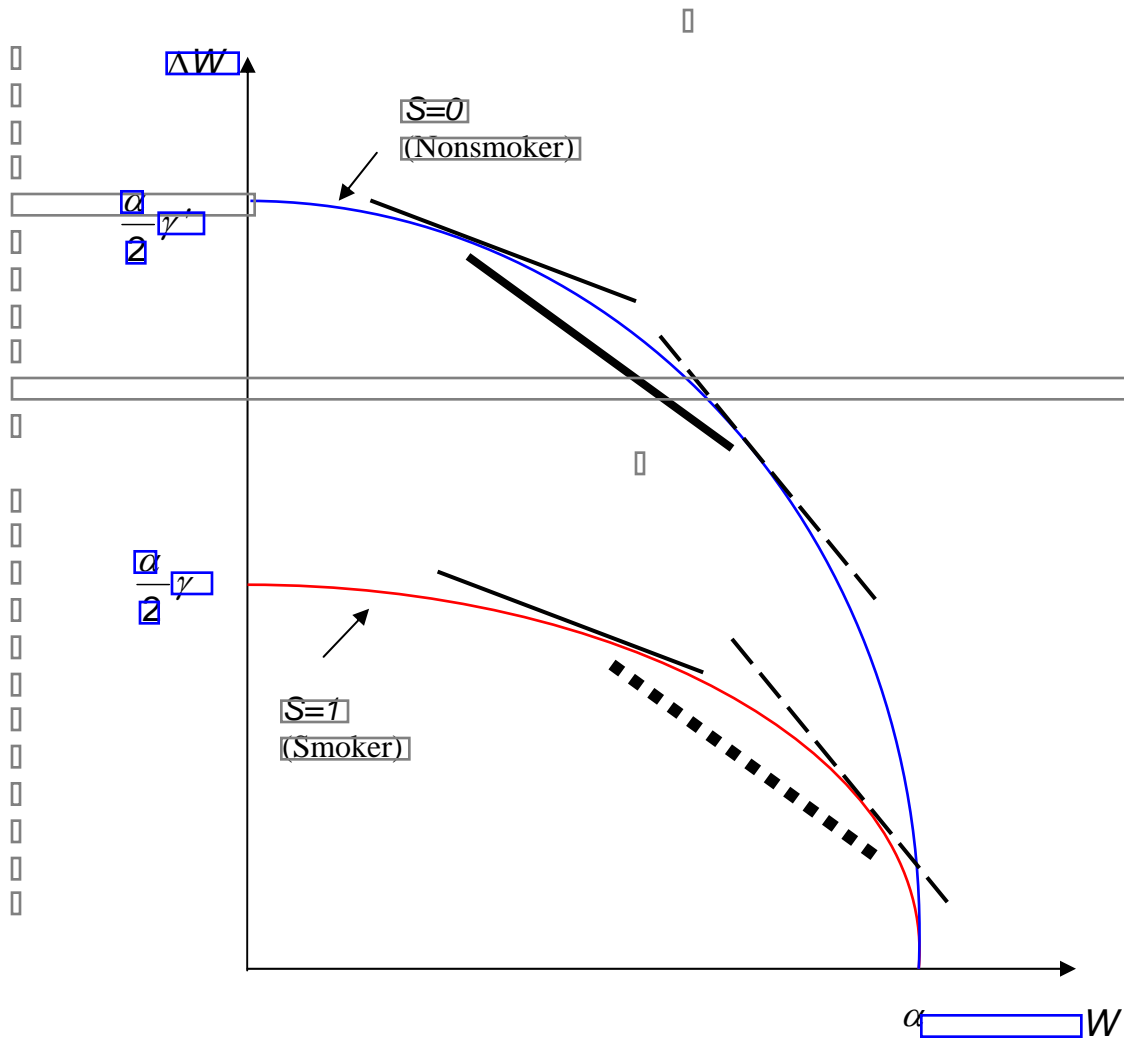
Figure 4

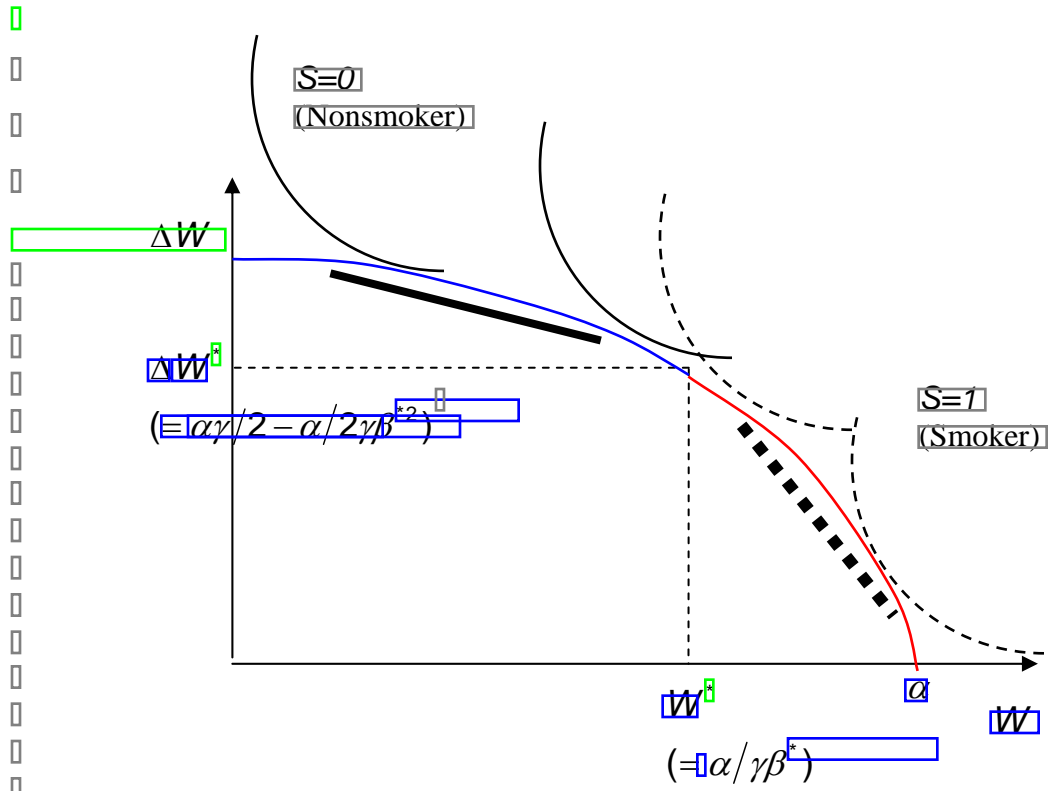## Discounting Hypothesis

S=0
(Nonsmoker)

ΔW

S=1
(Smoker)

ΔW^*

$(= \alpha\gamma/2 - \alpha/2\gamma\beta^{*2})$

W^*

α

W

$(= \alpha/\gamma\beta^*)$

Figure 5

Learning Hypothesis

Legend/labels on figure:

ΔW

$\frac{\alpha}{2} \gamma'$

$\frac{\alpha}{2} \gamma$

$S=0$ (Nonsmoker)

$S=1$ (Smoker)

α W

Figure 6

Wage Profiles across Smokers and Non-smokers

Wage

NS

S

β_{NS}

β_S

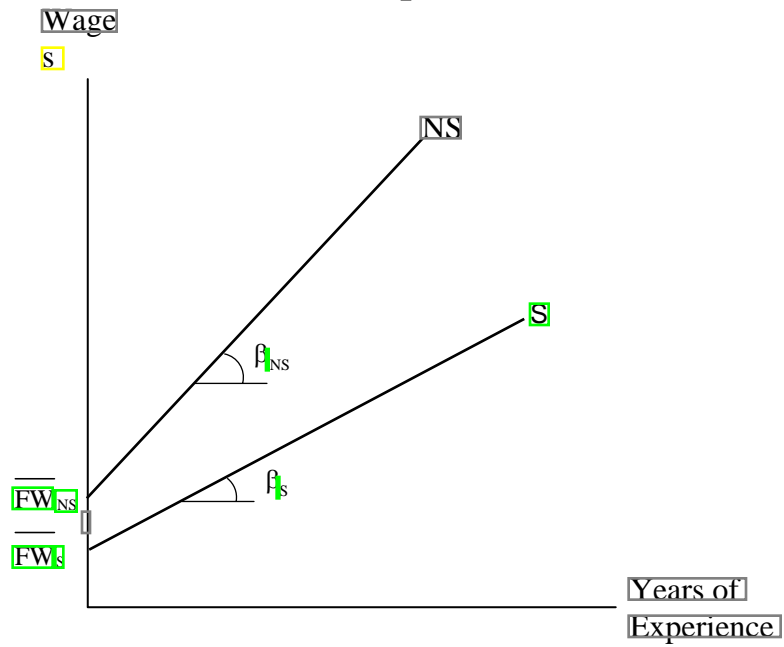$\overline{FW}_{NS}$

$\overline{FW}_S$

Years of Experience

Figure 7

Trade-off Between First Wage (FW) and Wage Growth (ΔW) for Smokers & Non-smokers

ΔW

ΔW$_{NS}$

ΔW$_S$

NS

S

~ $4.50

$\overline{FW}_S$ $\overline{FW}_{NS}$

FW

33