

THE DUAL AVENUES OF LABOR MARKET SIGNALING

By

Michael Waldman
Johnson Graduate School of Management
Cornell University
323 Sage Hall
Ithaca, NY 14853
(607) 255-8631
mw46@cornell.edu

November 2014

* I would like to thank seminar and conference participants at Cornell University, the University of Rochester, the 2014 Personnel Economics Conference at the University of Cologne, and the 2014 Trans Pacific Labor Economics Conference at the University of New South Wales for helpful comments.

ABSTRACT

This paper explores interactions between education signaling as first explored in Spence (1973) and promotion signaling as first explored in Waldman (1984). In order to establish a foundation for later results, the paper starts with two analyses each of which captures one of these avenues through which worker ability can be signaled to the labor market. The paper then builds on these early analyses by constructing and analyzing a model characterized by both education and promotion signaling, where education signals a worker's "academic" ability while promotion signals a worker's "productive" ability. This main analysis yields three returns to education signaling: i) a higher starting wage as in standard analyses of education signaling; ii) higher wages for non-promoted workers late in workers' careers; and iii) a higher probability that workers are on high level jobs late in workers' careers. The paper concludes with a discussion of what these theoretical results imply both for the social welfare aspects of education signaling and for recent studies focused on estimating the returns to education signaling.

I. INTRODUCTION

Starting with Spence's (1973) seminal contribution, it is well understood that one way that signaling affects labor market outcomes is through the schooling decision. And starting with Waldman (1984), it is also understood that a second way that signaling affects labor market outcomes is through the promotion decision. But there has been little research analyzing how these dual avenues through which signaling affects labor market outcomes are related. This paper investigates how education signaling and promotion signaling are related and, in particular, argues that this relationship is important for understanding the nature of returns to education signaling.

The basic education signaling argument is well known.¹ In models that capture the basic argument workers have private information about their own abilities and firms infer worker ability levels from publicly observable schooling decisions. In turn, in order to signal high ability and earn higher wages, workers overinvest in education, i.e., many workers invest in education beyond the level at which the marginal social return to investing equals the marginal social cost. And starting with Altonji and Pierret (1997) a number of papers extend this argument by incorporating the idea that firms start to learn about worker abilities after workers enter the labor market, so the signaling role of education should become less important as workers gain labor market experience. An important result in these papers is that the returns to education signaling should be concentrated early in workers' careers. Further, empirical papers such as Lange (2007) that take this into account find a limited role for signaling in real world education decisions.

In the basic promotion signaling argument it is firms that acquire private information about their own employees and prospective employers infer information about a worker's ability by observing the current employer's decision concerning whether or not to promote the worker.²

¹ In addition to Spence (1973), papers that investigate theoretical aspects of the education signaling argument include Riley (1975,1979a) and Cho and Kreps (1987). See Riley (2001) and Spence (2002) for surveys that discuss the theoretical and empirical literatures on this topic.

² In addition to Waldman (1984), papers that investigate theoretical aspects of the promotion signaling argument include MacLeod and Malcomson (1988), Ricart i Costa (1988), Waldman (1990), and Bernhardt (1995). See Waldman (2012) for a survey that discusses the theoretical and empirical literatures on this topic.

The result is that prospective employers are willing to bid more for a worker when they observe a promotion, so in equilibrium firms pair large wage increases with promotions in order to stop promoted workers from being bid away and also promote fewer workers than is efficient in order to reduce compensation costs. Further, a number of papers such as Bernhardt (1995) and DeVaro and Waldman (2012) build on this basic argument to show that in this type of setting workers with higher levels of education will be favored in the promotion process. The logic is that the signaling role of promotion is less important for more highly educated workers so there is less incentive to distort the promotion decision for such workers.

In this paper I construct and analyze a model that combines these two ideas. One innovation in the model is that, in order for workers and current employers to both have private information about worker abilities, I make a distinction between “academic” ability and “productive” ability. Academic ability captures an individual’s ability to do well in school, while productive ability captures an individual’s ability to be productive in employment. In previous theoretical models no distinction is made between these two concepts, but clearly from a real world standpoint a worker’s ability to do well in school is related to but not exactly the same as the worker’s ability to do well in a work setting. So I assume that academic ability and productive ability are positively correlated but that they are not identical. Further, I assume that an individual knows his or her own academic ability, while a worker’s current employer privately observes output which allows the firm to infer productive ability.

I construct and analyze a model in which individuals choose education levels at the beginning of their careers and then work in the labor market for two periods. I assume that schooling makes workers more productive, that the labor market is competitive, and that there are two job levels, where workers with low levels of productive ability produce more on the low level job while those with high levels are more productive on the high level job. One main result is that the model exhibits both education signaling as in Spence (1973) and promotion signaling as in Waldman (1984). Education signaling means that the private returns to higher levels of education exceed the direct extra productivity associated with higher education, so workers

invest more in education than in the first best. Promotion signaling means there is a wage increase associated with promotion because of the higher wage bids prospective employers make to promoted workers, so firms promote fewer workers than in the first best.

In addition, consistent with the earlier literature on promotion signaling, I also find that workers with higher levels of education are favored in the promotion process. That is, there can be pairs of workers where the worker in the pair with higher education has lower productive ability, but yet the worker with more education is the one promoted. The result is that there are three returns to education signaling in this model. First, there is a higher starting wage as in standard models of education signaling. Second, there is a higher wage for non-promoted old workers. Third, workers with more education are favored in the promotion process beyond the amount that is justified by the higher productive abilities of these workers. Note that in contrast to the first or standard return associated with education signaling, the second and third returns translate into changes in compensation levels paid to workers late in their careers.

These results have two important implications concerning how one should think about education signaling. The first concerns whether education signaling is good or bad from the standpoint of social welfare. In the standard education signaling model the increase in education levels due to signaling unambiguously reduces social welfare. The reason is that from a social welfare standpoint the cost of the increased education levels is higher than the aggregate increase in worker productivity generated by the higher education levels. In contrast, in this model characterized by both education and promotion signaling, the increase in education levels due to education signaling has both positive and negative effects on social welfare. On the one hand, the increase in education levels due to the first two signaling returns which are higher starting wages and higher wages for non-promoted old workers reduces social welfare as in the standard argument. On the other hand, the increase in education levels due to the third signaling return which is a higher probability of promotion serves to increase social welfare because it reduces the distortion in the promotion decision due to promotion signaling.

The other important implication concerns studies focused on measuring the returns to education signaling. Recent papers that measure this return such as Lange (2007) typically assume a single job and that after workers enter the labor market firms learn about worker abilities in a symmetric fashion, i.e., any information generated about a worker's ability is public so at any point in time all firms have the same beliefs about the worker's ability. In such a world the signaling role of education becomes less important for compensation as a worker gains labor market experience. The reason is that as experience increases less and less weight is placed on the education signal in the formation of beliefs about the worker's ability level. This is the fundamental insight driving the limited returns to education signaling in Lange's empirical analysis. But if there are important returns to education signaling that occur late in careers, as is true in the model investigated here, then the type of approach employed by Lange and others which implicitly assumes that returns are concentrated early in careers may significantly understate the returns to education signaling. I provide a detailed discussion of these issues in Section VI.

The paper in the literature closest to this one is Ishida (2004). He also considers a model that combines education signaling with promotion signaling but in his model, in contrast to the one considered here, education does not have a direct effect on worker productivity and so the only role of education is signaling. Also, in his model there is no distinction between academic ability and productive ability which is an important component of the modeling approach here. As a result, Ishida does find that education signaling can be used to reduce the promotion signaling distortion, but he does not find the same three returns to education signaling found here and the social welfare implications of education signaling in his model are different. In addition, he does not discuss what the approach of combining education with promotion signaling implies for studies focused on measuring the returns to education signaling.³

³ Perri (1994) also provides related results although that paper focuses on how ability testing by individuals interact with promotion signaling rather than the interaction between education signaling and promotion signaling.

The outline for the paper is as follows. Sections II and III provide basic analyses of the education and promotion signaling arguments, where Section II focuses on education signaling and Section III on promotion signaling. Section IV presents the main model which combines education and promotion signaling and provides preliminary analyses. Section V provides the main analysis. Section VI discusses the social welfare implications of the model and also what the analysis implies for recent papers that estimate the returns to education signaling. Section VII presents concluding remarks.

II. EDUCATION SIGNALING

In this section I start with a simple one-period model that captures the basic education signaling argument. I then extend the model to multiple sub-periods in order to capture the argument in the more recent literature that symmetric learning after workers enter the labor market serves to limit the role of education signaling.

Note that the goal of Sections II and III is to capture the main arguments in the education and promotion signaling literatures in easy to follow analyses. This will allow readers to more easily follow the main analysis in Sections IV and V which combines various arguments from these two literatures in a model characterized by both education and promotion signaling.

A) A One-Period Model

There is a pool of N workers where worker i has ability θ_i and each worker's value for θ_i is drawn from a probability density function $h(\cdot)$ and distribution function $H(\cdot)$ which takes the form of a uniform distribution over the interval $[0, \theta^H]$. Note that the assumption that the distribution is uniform is not required for the qualitative nature of the results. But rather it is imposed so that there are simple closed form solutions for various equilibrium values which makes the logic of the analysis easier to follow.

There is free entry of risk neutral firms into production so wages are determined by a zero expected profit condition and it is also assumed that wages are the only costs of production.

Worker i produces θ_i if $s_i=0$ and produces $\theta_i+\Delta$ if $s_i=1$, where s_i is worker i 's schooling level. Following Spence (1973), it is assumed that it is less costly for high ability workers to choose schooling. Specifically, let $c(\theta_i)$ be the cost to worker i of choosing $s_i=1$, where $c'<0$, $c(\theta^H)<\Delta$, and $c(0)>\Delta+(\theta^H/2)$.

The information assumptions are the standard ones in this literature. Each worker i knows his or her own value for θ_i but firms only know $H(\cdot)$ while they directly observe each worker's choice of s_i . The timing of the game is also standard. First, each worker's value for θ_i is drawn from the distribution function $H(\cdot)$. Second, workers simultaneously choose schooling levels. Third, firms simultaneously offer wages to workers. Fourth, workers choose firms, produce, and get paid. Also, the focus is on Perfect Bayesian equilibria, where we also assume that beliefs concerning off-the-equilibrium path actions are consistent with each such action being taken by the type with the smallest cost of choosing that action. This assumption is similar to the notion of a Proper equilibrium first discussed in Myerson (1978). This assumption concerning off-the-equilibrium path beliefs in combination with the assumptions $c(\theta^H)<\Delta$ and $c(0)>\Delta+(\theta^H/2)$ ensures that some workers choose schooling and some no schooling in equilibrium.

I start by analyzing the benchmark case of what happens when there is full information concerning the θ_i s, i.e., θ_i s are publicly observed at the beginning of the game. Competition and zero profits yield that each worker i is paid $\theta_i+s_i\Delta$, so the net return to choosing $s_i=1$ is $\Delta-c(\theta_i)$. Thus, each worker maximizing utility, i.e., the wage minus the cost of schooling, yields that there is a critical value θ^* that satisfies $c(\theta^*)=\Delta$, where worker i chooses $s_i=1$ if $\theta_i\geq\theta^*$ and chooses $s_i=0$ if $\theta_i<\theta^*$.⁴

I now consider the main analysis which assumes asymmetric information about the θ_i s, i.e., workers know their own values for θ_i while firms know $H(\cdot)$ and only directly observe each

⁴ Throughout the paper I assume a worker chooses schooling whenever the worker is indifferent between schooling and no schooling. Similarly, throughout the paper I assume a firm chooses to promote a worker whenever the firm is indifferent between promotion and no promotion.

worker's value for s_i . This case is described in Proposition 1. Note, all proofs are in the Appendix.

Proposition 1: If workers have private information about their own abilities and there is a single period, then i) through iii) describe equilibrium behavior.

- i) There exists a critical value for θ_i , denoted θ' , such that worker i chooses $s_i=1$ if $\theta_i \geq \theta'$ and $s_i=0$ if $\theta_i < \theta'$.
- ii) Worker i is paid $\Delta + ((\theta^H + \theta')/2)$ if $\theta_i \geq \theta'$ and is paid $\theta'/2$ if $\theta_i < \theta'$.
- iii) $\theta' < \theta^*$.

Proposition 1 captures a number of important aspects of equilibrium behavior. The first result captured in ii) is that, since firms can only directly observe schooling levels and there are only two levels, in equilibrium there are only two wages where one is for workers with schooling and one is for workers without schooling. In turn, this first result immediately leads to the second result which is that, since the cost of schooling is falling in θ_i and the benefit to schooling is the same for all workers, there is a critical value for ability, θ' , such that a worker chooses schooling when θ_i is greater than this critical value and chooses no schooling when it is less.

The wages stated in the proposition follow from the idea that competition means wages must satisfy expected productivity. In turn, since θ' is the value for θ_i such that the net return to schooling equals zero, we have $[\Delta + ((\theta^H + \theta')/2)] - (\theta'/2) = c(\theta')$ or $c(\theta') = \Delta + (\theta^H/2)$. This, in turn, yields $\theta' < \theta^*$ given $c' < 0$ and $c(\theta^*) = \Delta$. This last result tells us that this model is consistent with the standard education signaling result that from a social welfare standpoint there is overinvestment in education. Further, the logic for the result is the standard logic. That is, because the increase in the wage associated with schooling includes being grouped with higher θ_i types, the private return to schooling exceeds the social return and “too many” individuals choose schooling.

B) T Sub-periods and Symmetric Learning

The more recent literature on education signaling argues that the basic Spence signaling result captured above exaggerates the importance of signaling. The argument is that Spence's original argument ignores that firms learn about worker abilities directly after labor market entry, so the importance of the signal should decrease as workers gain labor market experience. Further, basically all the papers in the literature that make this point assume that this learning is symmetric.

In this subsection I capture this argument by extending the model in the previous subsection to allow for this learning effect. And to make the argument transparent I make a simple assumption concerning how this learning occurs. The new assumptions are as follows. First, the single period now consists of T sub-periods each of which is $1/T$ periods long. Second, in each sub-period worker i produces $(1/T)(\theta_i + s_i \Delta)$. Third, at the end of each sub-period each worker's output for that sub-period is publicly observed. Fourth, contracting and wage determination occur at the beginning of each sub-period. Fifth, there is no discounting. Note that there is no stochastic element concerning output determination which means learning is complete by the beginning of sub-period 2. It would be more realistic to assume that learning is gradual which could be captured by adding a stochastic element, but the argument is more transparent without it.

The timing of the game is now as follows. First, each worker's value for θ_i is drawn from the distribution function $H(\cdot)$. Second, workers simultaneously choose schooling levels. Third, firms simultaneously offer wages for sub-period 1. Fourth, workers choose firms for sub-period 1, produce, and get paid. Fifth, the third and fourth steps repeat for all subsequent sub-periods through sub-period T .

Notice that a benchmark analysis which assumes the θ_i s are publicly observed at the beginning of the game yields an equilibrium very similar to the one derived for the full information benchmark in the previous subsection. Competition and zero expected profits yields that in each sub-period worker i is paid $(1/T)(\theta_i + s_i \Delta)$. Further, the aggregate wage paid to worker

i over the T sub-periods is $\theta_i + s_i \Delta$ which is the same as in the full information benchmark analysis of the previous subsection. In turn, the net return to schooling equals $\Delta - c(\theta_i)$, so just like in that earlier benchmark analysis worker i chooses $s_i = 1$ when $\theta_i \geq \theta^*$, and chooses $s_i = 0$ when $\theta_i < \theta^*$.

Now consider asymmetric information. This case is described in Proposition 2.

Proposition 2: If workers have private information about their own abilities and there are T , $T \geq 1$, sub-periods, then i) through iv) describe equilibrium behavior.

- i) There exists a critical value for θ_i , denoted θ_T' , such that worker i chooses $s_i = 1$ if $\theta_i \geq \theta_T'$ and $s_i = 0$ if $\theta_i < \theta_T'$.
- ii) In sub-period 1 worker i is paid $(1/T)[\Delta + ((\theta^H + \theta_T')/2)]$ if $\theta_i \geq \theta_T'$ and is paid $(1/T)(\theta_T'/2)$ if $\theta_i < \theta_T'$.
- iii) In sub-period t , $t = 2, \dots, T$, worker i is paid $(1/T)(\theta_i + s_i \Delta)$.
- iv) $\theta_T' < \theta^*$ for all T , $T \geq 1$, θ_T' increases with T , and θ_T' approaches θ^* as T gets large.

Proposition 2 tells us that equilibrium behavior when there are T sub-periods has some similarities with equilibrium behavior when there is only a single period (or more precisely, a single sub-period). For example, in sub-period 1, which is the only sub-period in the Proposition 1 analysis, there are two wages – a schooling wage and a no-schooling wage – and the wages are determined by a zero expected profit condition. Also, there is overinvestment in schooling again since $\theta_T' < \theta^*$ for all T , $T \geq 1$.

But there are also a number of new results. One new result follows from the idea that at the end of sub-period 1 firms observe sub-period 1 outputs and are able to infer ability levels from these observations. Specifically, combining this idea with competition and zero profits yields iii) of the proposition which states that the wage for each worker i in every sub-period t , $t \geq 2$, equals $(1/T)(\theta_i + s_i \Delta)$.

Given these results concerning wages, we have that θ_T' must satisfy equation (1).

$$(1) \quad (1/T)[\Delta + ((\theta^H + \theta_T')/2)] + ((T-1)/T)(\Delta + \theta_T') - (1/T)(\theta_T'/2) - ((T-1)/T)\theta_T' = c(\theta_T')$$

Equation (1) yields equation (2).

$$(2) \quad \Delta + (1/T)(\theta^H/2) = c(\theta_T')$$

The first thing to note about equation (2) is that, as it should, it reduces to the equation that defines θ' when $T=1$. Given $c' < 0$ and $c(\theta^*) = \Delta$, equation (2) also tells us that $\theta_T' < \theta^*$ for all T , $T \geq 1$, i.e., there is overinvestment in schooling for every T . But note further that, since the left hand side of (2) falls with T , θ_T' increases with T . And since $(1/T)(\theta^H/2)$ approaches zero as T gets large, θ_T' approaches θ^* as T gets large.

If we interpret the single period as the length of a worker's career, then increasing T is analogous to increasing the speed of learning in a symmetric learning model where workers' outputs are publicly observable after labor market entry. The model in this subsection thus captures the idea that signaling becomes less important the faster that firms learn given that learning after labor market entry is symmetric.

C) Evidence

I end this section with a brief discussion of the empirical literature concerning education signaling. There is an extensive literature on the topic. Many of the early papers on the topic such as Layard and Pacharopoulos (1974) and Hungerford and Solon (1987) show that attaining a degree is correlated with higher compensation which was put forth as evidence for the signaling argument. But this result is consistent with a world where education's sole role is acquiring human capital since completing a degree can be a proxy for higher human capital. So these early papers do not provide strong evidence for the signaling idea. But various other papers such as Wolpin (1977), Riley (1979b), Lang and Kropp (1986), and Bedard (2001) present evidence consistent with education signaling not subject to this criticism.

More recently attention has turned from whether or not education signaling exists to how important it is in real world labor markets. Altonji and Pierret (1997) argue that signaling should be more important in terms of compensation early rather than late in careers, where the logic they employ is that of symmetric learning after labor market entry. That is, if firms learn about

worker abilities in a symmetric fashion after labor market entry, then as workers age less and less weight should be placed on the education signal in determining compensation so the role of signaling should be limited because it only has a significant effect on wages early in workers' careers. This is the logic captured in the analysis above.

In an important paper Lange (2007) builds on analyses in Farber and Gibbons (1997) and Altonji and Pierret (2001) and constructs a structural estimation model focused on this idea. He finds in an analysis of the National Longitudinal Survey of Youth that the speed of employer learning is fast which suggests that signaling is not very important. In particular, in his preferred specification approximately ten percent of the total gain to an additional year of schooling is due to education signaling.

III. PROMOTION SIGNALING

In this section I capture the main arguments in the promotion signaling literature. I start with a simple model with workers who look *ex ante* identical and show how promotions can serve as a signal of worker ability. I then extend the model to multiple schooling groups to show how higher levels of education affect the promotion decision.

A) Ex Ante Identical Workers

As in the previous section, there is a pool of N workers where worker i has ability θ_i and each worker's value for θ_i is drawn from a probability density function $h(\cdot)$ and distribution function $H(\cdot)$ which takes the form of a uniform distribution over the interval $[0, \theta^H]$. It is further assumed that $\theta^H > 1$ which given additional assumptions below means that it is efficient for a θ^H worker to be promoted in the second period. Also, workers and firms are risk neutral and there is no discounting.

As in the previous model the full length of a worker's career is one period, where it is assumed this single period consists of two sub-periods. At the end of the analysis I briefly discuss what happens when there are more sub-periods. In this two sub-period model there is

free entry and wages are the only costs. So wages are determined by zero expected profits. All workers can be assigned to either of two jobs – an unskilled job denoted job 1 and a skilled job denoted job 2 – although workers in sub-period 1 produce zero if assigned to job 2. If worker i is assigned to job 1 in a sub-period, the worker produces $(1/2)(B+\theta_i)$, $B>1$, if it is sub-period 1 or sub-period 2 and the worker is at a different employer than the one which employed the worker in sub-period 1. If it is sub-period 2 and the worker is employed at the same firm that employed the worker in sub-period 1, then the worker produces $z(1/2)(B+\theta_i)$. z thus captures firm specific human capital.

On job 2, as indicated, worker i produces zero if it is sub-period 1, while if it is sub-period 2 the worker produces $(1/2)(B-1+2\theta_i)$ at a new employer and produces $z(1/2)(B-1+2\theta_i)$ when the worker is at the same employer as in sub-period 1. In Waldman (1984) there are no promotions in the absence of firm specific human capital. In the current model there are no promotions unless z is sufficiently large, so that is what is assumed.⁵

The information assumptions are the standard ones found in promotion signaling models. First, at the beginning of sub-period 1 each worker's ability is known by neither the worker nor the firms, but everyone knows $H(\cdot)$. At the end of each sub-period the realization of each worker's output is privately observed by the worker's current employer. Since there is no stochastic element determining the amount of output produced, this means each worker's ability is learned with certainty by the worker's sub-period 1 employer at the end of sub-period 1.⁶ Further, job assignments are publicly observable.

The timing of the game is as follows. First, at the beginning of sub-period 1 each worker's value for θ_i is realized. Second, each firm then offers a wage/job assignment pair for sub-period 1 employment. Third, workers choose firms, produce, and get paid. Fourth, at the end of sub-period 1 each firm privately observes the outputs of its sub-period 1 employees. Fifth,

⁵ The specific condition is $z > (2\theta^H - 1) / (\theta^H - 1)$.

⁶ If a worker is assigned to job 2 in sub-period 1, even though output is zero it is still assumed that the sub-period 1 employer learns the worker's ability. One way to justify this assumption is to assume that output in this case equals $q\theta_i$ and our focus is equilibrium behavior in the limit as q approaches zero from above.

at the beginning of sub-period 2 each firm offers a wage/job assignment pair to each of its sub-period 1 employees. Sixth, firms then make wage/job assignment offers to workers other than their sub-period 1 employees. Seventh, each firm then makes a wage counter-offer to each of its sub-period 1 employees. Eighth, workers choose firms, produce, and get paid.

Note that the counter-offer assumption employed here can be found in early related papers such as Greenwald (1986) and Milgrom and Oster (1987).⁷ Also, the overall timing of this model is similar to the timing found in Zabojnik and Bernhardt (2001). As in the previous section, we focus on Perfect Bayesian equilibria in which beliefs concerning off-the-equilibrium path actions are consistent with each such action being taken by the type with the smallest cost of choosing that action. Imposing this assumption here means that our model is characterized by a strong winner's curse as found in Milgrom and Oster's analysis.⁸

I start the analysis of this model by considering a benchmark case. Specifically, suppose there is symmetric rather than asymmetric learning which means that outputs are publicly observed at the end of sub-period 1 rather than each worker's output being privately observed by the worker's sub-period 1 employer. In this case in sub-period 1 all workers are assigned to job 1 and are paid $(1/2)[B+(\theta^H/2)]+\lambda$, $\lambda>0$, where $(1/2)[B+(\theta^H/2)]$ is the expected sub-period 1 output and λ is the expected sub-period 2 profit associated with hiring a worker in sub-period 1. As discussed next, the reason there is such a profit is because of the presence of firm specific human capital.

Now consider sub-period 2. Because of firm specific human capital, all workers remain with their sub-period 1 employers. Further, because of symmetric learning and the counter-offer assumption, worker i is assigned to job 1 and is paid $(1/2)(B+\theta_i)$ if $\theta_i<\theta^*$, $\theta^*=1$, and is assigned

⁷ See Barron, Berger, and Black (2006) for an empirical analysis that shows that counter-offers are common in real world labor markets. Also, note that Golan (2005) argues that the promotion signaling distortion does not arise in models characterized by the counter-offer assumption, but Waldman and Zax (2014) show that Golan's results rely on her assumption that output on the low level job is a constant and that there is a distortion if, as is the case here, output on the low level job increases with worker ability.

⁸ Milgrom and Oster (1987) did not impose this additional assumption. The main reason this extra assumption is required here for the strong winner's curse result but was not in that earlier paper is that our model is characterized by firm specific human capital which was not the case in Milgrom and Oster's paper.

to job 2 and is paid $(1/2)(B-1+2\theta_i)$ if $\theta_i \geq \theta^*$. In other words, workers are efficiently assigned to jobs in sub-period 2 and each worker's pay equals the worker's productivity at an alternative employer. Given worker i 's productivity in sub-period 2 at the worker's sub-period 1 employer equals $z(1/2)\max\{B+\theta_i, B-1+2\theta_i\}$, we now have $\lambda > 0$.

Now consider what happens given asymmetric learning, i.e., outputs are privately observed by current employers. It is again the case that in sub-period 1 all workers are assigned to job 1 and are paid $(1/2)[B+(\theta^H/2)]+\lambda$, although the specific value for λ is different than under symmetric learning. Under symmetric learning there is positive expected sub-period 2 profits associated with hiring a worker in sub-period 1 because of firm specific human capital. Under asymmetric learning there is positive expected sub-period 2 profits both because of firm specific human capital and because of the winner's curse result described below which serves to lower sub-period 2 wages.

There are also similarities between what happens under asymmetric learning in sub-period 2 and the benchmark analysis. First, every worker remains with the worker's sub-period 1 employer. Second, there is a critical value for θ_i , now call it θ' , such that worker i is assigned to job 2 if $\theta_i \geq \theta'$ and is assigned to job 1 if $\theta_i < \theta'$. The difference now is that sub-period 2 job assignments are no longer efficient. That is, $\theta' > \theta^* = 1$ which means fewer workers are promoted than is first best efficient. The logic, which is explained in more detail below, is that firms want to avoid paying the higher wage associated with promotion with the result that the probability of promotion is lower than the efficient level.

I now present a more complete description of equilibrium behavior given asymmetric learning.

Proposition 3: Given firms privately observe the output levels of their workers, there exist values θ' and λ such that equilibrium behavior is characterized by i) through iv).

- i) In sub-period 1 all workers are assigned to job 1 and are paid $(1/2)[B+(\theta^H/2)]+\lambda$.
- ii) In sub-period 2 worker i is assigned to job 1 and is paid $(1/2)B$ if $\theta_i < \theta'$.

- iii) In sub-period 2 worker i is assigned to job 2 and is paid $(1/2)(B-1+2\theta')$ if $\theta_i \geq \theta'$.
- iv) $1 = \theta^* < \theta' < \theta^H$, θ' is a decreasing function of z , and $\lambda > 0$.

Proposition 3 provides more detail concerning equilibrium behavior given asymmetric learning. First, sub-period 2 wages are determined by the winner's curse. That is, because of the counter-offer assumption, a prospective employer in sub-period 2 is only willing to offer a worker a wage equal to the lowest productivity associated with the worker's signal/job assignment and then the current employer matches this wage offer and the worker stays. This is why workers assigned to job 1 in sub-period 2 are paid $(1/2)B$ while those assigned to job 2 are paid $(1/2)(B-1+2\theta')$. Second, because the wage paid in sub-period 2 to a promoted worker is higher than the wage paid to a non-promoted worker, a firm promotes fewer workers than is efficient, i.e., $\theta' > \theta^*$. Third, this distortion falls with an increase in firm specific human capital. The logic here is that the firm receives a higher benefit from an efficient sub-period 2 job assignment when firm specific human capital is more important, so the severity of the distortion falls with the degree of firm specific human capital.

B) M Schooling Groups

The more recent literature on promotion signaling argues that the promotion distortion captured above is smaller for workers with higher levels of schooling. The logic is that the signal associated with promotion is smaller for more highly educated workers so there is less incentive to distort the promotion decision for such workers. In this subsection I capture this argument by introducing multiple schooling groups into the model analyzed in the previous subsection.

There are now two sub-periods and M schooling groups, denoted $1, \dots, M$, where there are N_m workers in group m . The θ_i s for workers in group m are drawn from a probability density function $h_m(\cdot)$ and distribution function $H_m(\cdot)$, where the distribution is uniform over the interval $[x(m), x(m) + \theta^H]$. It is further assumed that $x(1) = 0$, $x' > 0$, and $B + x(M) > B - 1 + 2x(M)$. That is, for the lowest schooling group the ability distribution is the same as in the previous subsection, the

range of the ability distribution increases with the schooling group, and even for the highest schooling group it is efficient to assign some workers to job 1 in sub-period 2. It is also assumed that a worker's schooling group is publicly observed at the beginning of the game.

The first thing to note about this model is that the benchmark case of symmetric learning after workers enter the labor market is very similar to the benchmark analysis of symmetric learning in the previous subsection in which all workers were ex ante identical. First, all decisions are efficient which means there is no turnover and workers are efficiently assigned to jobs in each sub-period. Second, in sub-period 1 wages equal expected productivity plus expected sub-period 2 profits associated with hiring in sub-period 1. Third, each worker's wage in sub-period 2 equals the worker's productivity at an alternative employer. Additionally, the only aspect of this benchmark which is new is that the probability of promotion rises with education which follows given job assignments are efficient and $x'(m) > 0$ for all m .

I now consider the case of asymmetric learning. Equilibrium in this case is described in Proposition 4.

Proposition 4: Given firms privately observe the output levels of their workers and there are M schooling groups, there exist functions $\lambda(m)$ and $\theta'(m)$ such that equilibrium behavior is described by i) through iv).

- i) In sub-period 1 all workers are assigned to job 1 and worker i in group m is paid $(1/2)[B+x(m)+(\theta^H/2)]+\lambda(m)$.
- ii) In sub-period 2 worker i in group m is assigned to job 1 and is paid $(1/2)[B+x(m)]$ if $\theta_i < \theta'(m)$.
- iii) In sub-period 2 worker i in schooling group m is assigned to job 2 and is paid $(1/2)[B-1+2\theta'(m)]$ if $\theta_i \geq \theta'(m)$.
- iv) $\theta'(m) > \theta^* = 1$ for all m , $\theta'(m+1) < \theta'(m)$ for all $m=1, \dots, M-1$, and $\lambda(m) > 0$ for all m .

Proposition 4 tells us that the asymmetric learning analysis with multiple schooling groups works similarly to the single group case. For example, sub-period 1 wages are again determined by sub-period 1 expected productivity plus the expected sub-period 2 profit associated with hiring a worker in sub-period 1. Second, the winner's curse means that sub-period 2 wages equal the productivity at an alternative employer of the worst worker with the same labor market signal. Third, because wages are higher for promoted workers, firms promote fewer workers than is efficient, i.e., $\theta'(m) > \theta^* = 1$ for all m .

The new result here concerns the size of the promotion distortion. Specifically, the size of this distortion falls with the schooling level, i.e., $\theta'(m)$ gets closer to θ^* as m increases. The logic for this result is that the non-promotion wage increases with the schooling level which, if the critical value for θ_i did not change with an increase in m , would mean a smaller promotion premium as m increases. But this means there would be less incentive to distort the promotion decision as m increases, so what happens in equilibrium is that this critical value falls as m increases. Another way to describe this result is that, as m increases the wage increase upon promotion due to the signal becomes smaller, so the incentive for the employer to distort the promotion decision is also smaller.

One last point about the theory of promotion signaling concerns what happens when the number of sub-periods is allowed to grow. In the previous section which dealt with education signaling, the signal became unimportant as the number of sub-periods grew large. That is, the education signaling distortion became vanishingly small and, more generally, equilibrium wages and investment levels grew closer and closer to the equilibrium values under full information. Although I will not show it formally for length reasons, in this section's model allowing the number of sub-periods to grow large does not cause the signaling distortion to become vanishingly small and the equilibrium does not approach the full information equilibrium. The reason is that, because learning is private rather than public, firms have an incentive to hide their information and so the promotion signaling distortion continues to be significant even with a

large number of sub-periods and the equilibrium continues to look quite different than the full information equilibrium.⁹

C) Evidence

Gibbons and Katz (1991) was the first paper to empirically test for the presence of asymmetric learning in labor markets. Their focus was Greenwald's (1986) adverse selection argument concerning labor market turnover and its implications for differences between laid off workers and those fired in a plant closing. Using the Current Population Survey, they found support for the adverse selection argument. However, there have been a number of follow-up studies such as Acemoglu and Pischke (1998), Krashinsky (2002), and Song (2007) and the evidence is mixed concerning whether the original Gibbons and Katz findings are in fact due to adverse selection. But there are a number of more recent papers that take alternative approaches to test for asymmetric learning such as Schoenberg (2007), Pinkston (2009), and Kahn (2013). And in general these more recent papers find evidence consistent with asymmetric learning being important.¹⁰

There are also a number of recent papers that directly consider the promotion-as-signal hypothesis. Most of these papers focus on tests derived from the basic idea first put forth in Bernhardt (1995) and also investigated above that the signal associated with promotion should be smaller for workers in higher education groups, so workers in these groups should be favored in the promotion process. In general, the papers that take this approach such as Belzil and Bognanno (2010), DeVaro and Waldman (2012), Cassidy, DeVaro, and Kauhanen (2014), and Bognanno and Melero (Forthcoming) find evidence that supports the argument.¹¹

⁹ See Bernhardt (1995) for an analysis of a promotion signaling model with a large number of periods.

¹⁰ Of these more recent papers, Schoenberg's is the only one that finds weak evidence for asymmetric learning. But as I argue in Waldman (2012), it is unclear that the test for which she finds weak evidence for asymmetric learning is in fact a valid test of the asymmetric learning argument. See also Kim and Usui (2014) for a recent related discussion and analysis.

¹¹ Gibbs (2003) employs alternative tests of the promotion signaling hypothesis that do not depend on how promotion signals vary with education and he also finds evidence that supports the hypothesis.

IV. EDUCATION AND PROMOTION SIGNALING – PART 1

In this section I present a model which exhibits both education and promotion signaling and then provide preliminary analyses. These analyses describe how the model works when there is no signaling and when there is only education signaling. In the next section I analyze equilibrium behavior given both education and promotion signaling.

One innovation in the model presented below is that it introduces two distinct types of worker ability – academic ability which affects the cost of schooling and productive ability which determines a worker’s productivity on the job. As discussed in more detail below, having these two distinct types of ability is important for the two types of signaling to co-exist.

A) The Model

There is a pool of N workers where worker i has academic ability θ_i and each worker’s value for θ_i is drawn from a probability density function $h(\cdot)$ and distribution function $H(\cdot)$ which takes the form of a uniform distribution over the interval $[0, \theta^H]$, $\theta^H > 1$. There is also an attribute α_i drawn from a probability density function $g(\cdot)$ and distribution function $G(\cdot)$ which takes the form of a uniform distribution over the interval $[-\alpha, \alpha]$, where $\alpha > \theta^H$ and the realizations of θ_i and α_i are independent. The assumption that the θ_i distribution is uniform is not necessary but serves to simplify the exposition.

Each worker also has a value for productive ability, denoted θ_i^P , where $\theta_i^P = \theta_i + \alpha_i$. That is, productive ability and academic ability are correlated but not perfectly. As described in detail below, a worker’s productivity is also affected by whether or not the worker chooses schooling, where s_i , $s_i = 0, 1$, is the schooling level of worker i . Also, the cost of schooling for worker i is given by $c(\theta_i)$, where $c' < 0$ and $c'' < 0$.¹² That is, the cost of schooling is determined by a worker’s academic ability.

¹² The assumptions $\alpha > \theta^H$, that the α distribution is uniform, and $c'' < 0$ ensure that there is a critical value for academic ability such that all workers with higher academic ability levels choose schooling and all those with lower academic ability levels choose no schooling. See the proof of Proposition 5 in the Appendix for details.

There is free entry of risk neutral firms into production so wages are determined by a zero expected profit condition and it is also assumed that wages are the only cost in the production process. There are two sub-periods, where in each sub-period each worker can be assigned to either of two jobs – an unskilled job again denoted job 1 and a skilled job again denoted job 2. If worker i is assigned to job 1 and it is either sub-period 1 or sub-period 2 and the worker is at a different employer than the worker's employer in sub-period 1, then the worker produces $(1/2)(B+\theta_i^P+s_i\Delta)$. However, if it is sub-period 2 and the worker is employed at the same employer as in sub-period 1, then the worker produces $z(1/2)(B+\theta_i^P+s_i\Delta)$. We also assume $c(\theta^H) < (z+1)(\Delta/2)$, $c(0) > (z+1)(\Delta/2) + (\theta^H/2)$, and that z is sufficiently large that there are promotions in equilibrium both for workers with schooling and those without.¹³

Now suppose that worker i is assigned to job 2. If it is sub-period 1, then the worker produces zero. If it is sub-period 2 and the worker is at a new employer, then the worker's output equals $(1/2)(B-1+2\theta_i^P+s_i\Delta)$. And if it is sub-period 2 and the worker is at the same employer as in sub-period 1, then the worker produces $z(1/2)(B-1+2\theta_i^P+s_i\Delta)$. It is also assumed that there is no discounting by either the firms or workers.

The information assumptions are as follows. First, at the beginning of the game each worker privately observes his or her own value for θ_i while each worker's value for α_i is not known by anyone, although the functions $H(\cdot)$ and $G(\cdot)$ are common knowledge. Second, at the end of each sub-period a worker's output is privately observed by the worker's current employer, so a worker's sub-period 1 employer learns the worker's productive ability at the end of the first sub-period.¹⁴ Third, schooling levels and job assignments are public information.

¹³ As in Section III, the specific condition is $z > (2\theta^H - 1)/(\theta^H - 1)$. It is also assumed that $B > \alpha$ to avoid negative wages in equilibrium.

¹⁴ As was true in the model analyzed in Section III, if a worker is assigned to job 2 in sub-period 1, even though output is zero it is still assumed that the sub-period 1 employer learns the worker's productive ability. Similar to what was true earlier, one way to justify this assumption is to assume that output in this case equals $q\theta_i^P$ and our focus is equilibrium behavior in the limit as q approaches zero from above.

The timing of moves in the game is as follows. First, at the beginning of sub-period 1 values for θ_i and α_i are realized. Second, workers then simultaneously choose schooling levels. Third, this is followed by each firm offering a wage/job assignment pair for sub-period 1 employment. Fourth, workers then choose firms, produce, and get paid. Fifth, at the end of sub-period 1 each firm privately observes the outputs of its sub-period 1 employees. Sixth, sub-period 2 starts with each firm offering a wage/job assignment pair to each of its sub-period 1 employees. Seventh, each firm is then allowed to make wage/job assignment offers to workers it did not employ in sub-period 1. Eighth, this is followed by firms being allowed to make a wage counter-offer to each of its sub-period 1 employees. Ninth, workers choose firms, produce, and get paid.

B) Preliminary Analyses

In this subsection I consider two preliminary analyses. I start with the benchmark case where θ_i s are publicly observed at the beginning of the game and outputs are publicly observed at the end of each sub-period. Given these information assumptions, there will be neither education signaling which requires private information concerning the θ_i s nor promotion signaling which requires outputs to be privately observed by employers. As discussed below, the result is that schooling decisions and promotion decisions are both efficient.

Consider first wages and job assignments. In sub-period 1 all workers are assigned to job 1 given assignment to job 2 results in zero output. Also, each worker i is paid $(1/2)(B+\theta_i+s_i\Delta) + \lambda(\theta_i, s_i)$, where $(1/2)(B+\theta_i+s_i\Delta)$ is worker i 's expected sub-period 1 output given the α_i s are not observed at the beginning of the game and $\lambda(\theta_i, s_i)$ is the expected sub-period 2 profit associated with hiring a worker in sub-period 1 with academic ability θ_i and schooling level s_i . This wage is determined by competition across firms at the beginning of sub-period 1 and the resulting zero expected profit condition.

Now consider sub-period 2. If $\theta_i^P \geq (<)1$, then worker i stays with the worker's sub-period 1 employer, is assigned to job 2(1), and is paid $(1/2)\max\{B+\theta_i^P+s_i\Delta, B-1+2\theta_i^P+s_i\Delta\}$. The logic

here is that, because outputs are publicly observed at the end of sub-period 1, firms are able to infer the α_i s and thus the θ_i^P s. So as a result in sub-period 2 workers are assigned to jobs efficiently and pay is determined by the worker's productivity at a prospective employer which is what these alternative employers offer.

The final aspect of equilibrium behavior in this benchmark case is the schooling decisions made at the beginning of the game. Similar to what was true in the benchmark analysis in Section II, there is a critical value for academic ability, denoted θ^* , such that worker i chooses $s_i=1$ if $\theta_i \geq \theta^*$ and chooses $s_i=0$ if $\theta_i < \theta^*$. The logic is that θ^* is the value for θ_i such that the expected return to schooling just equals the cost of schooling, so since the cost of schooling falls with θ_i individuals choose schooling when θ_i exceeds this value and no schooling when θ_i is less than this value. Further, since the added productivity associated with schooling is $\Delta/2$ in sub-period 1 and $z\Delta/2$ in sub-period 2 given there is no turnover in equilibrium, θ^* now satisfies $c(\theta^*)=(z+1)(\Delta/2)$.

Note that in this benchmark analysis schooling decisions, promotion decisions, and turnover decisions are all efficient. There is no distortion concerning schooling decisions because θ_i s are publicly observed at the beginning of the game, so schooling does not have a signaling role. Further, there is no distortion concerning promotion decisions because outputs are publicly observed at the end of sub-period 1 which means that all firms correctly infer each worker i 's value for θ_i^P at the end of sub-period 1. And there is no turnover which is efficient given the presence of firm specific human capital.

I now consider the second preliminary analysis which is that each worker privately observes his or her own value for θ_i at the beginning of the game but it is still the case that outputs are publicly observed at the end of each sub-period. This case is similar to the symmetric learning extension of the basic education signaling model in Section II. That is, in this case education serves as a signal in sub-period 1 in terms of beliefs firms hold concerning a worker's academic ability and thus expected productive ability since the two are equal given no information about a worker's realization for α_i . Thus, the education choice affects equilibrium

pay in sub-period 1 in a fashion consistent with education being a signal. But at the end of sub-period 1 firms observe outputs and correctly infer realizations for θ_i^P , so there is no signaling role for education in sub-period 2. Additionally, since the θ_i^P s are learned by all firms at the end of sub-period 1, there is also no signaling role for promotions.

I now describe equilibrium in this case in more detail. I start by focusing on the beginning of the game. Because only schooling is publicly observable at the beginning of sub-period 1, in sub-period 1 there are two wages – a schooling wage and a no-schooling wage – and a critical value for academic ability, call it θ' , such that at the beginning of sub-period 1 worker i chooses schooling if $\theta_i \geq \theta'$ and chooses no schooling if $\theta_i < \theta'$. Also, in sub-period 1 all workers are assigned to job 1 as in the earlier benchmark case. Given these results, we can derive expressions for wages in sub-period 1. The schooling wage is given by $(1/2)[B + ((\theta^H + \theta')/2) + \Delta] + \lambda^S$ while the no-schooling wage equals $(1/2)[B + (\theta'/2)] + \lambda^{NS}$, where λ^S is the expected profit in sub-period 2 associated with hiring a worker with schooling in sub-period 1 and λ^{NS} is the expected profit in sub-period 2 associated with hiring a worker without schooling in sub-period 1. In other words, the wages in sub-period 1 equal expected sub-period 1 output plus expected sub-period 2 profits.

Now consider sub-period 2. Because outputs are publicly observed at the end of sub-period 1 and so everyone knows each worker's value for θ_i^P at the beginning of sub-period 2, in sub-period 2 equilibrium behavior is exactly the same as in the first preliminary analysis. First, every worker stays with his or her sub-period 1 employer. Second, if $\theta_i^P \geq (<) 1$, then the worker is assigned to job 2(1) and is paid $(1/2)\max\{B + \theta_i^P + s_i\Delta, B - 1 + 2\theta_i^P + s_i\Delta\}$. Or another way to put this is that, because productive abilities are learned by all firms at the end of sub-period 1, job assignments in sub-period 2 are efficient and a worker's wage in sub-period 2 equals the worker's output at an alternative employer given the worker is assigned efficiently.

The last aspect of equilibrium behavior in this case is the value for θ' . Because schooling serves a signaling role for sub-period 1, there is overinvestment in education just like in Section II's analysis, i.e., $\theta' < \theta^*$. The logic is the same as in that earlier analysis. Choosing schooling

means that, in terms of the sub-period 1 wage, a worker is pooled with workers with high values for θ_i . So the increase in the sub-period 1 wage due to schooling exceeds the direct increase in productivity due to schooling and as a result workers overinvest in education relative to the first best. In other words, similar to what was true in the earlier analysis in Subsection II.B, there is overinvestment in education but that overinvestment is due solely to the higher schooling wage in the first sub-period because with symmetric learning after labor market entry there is no role for signaling after the first sub-period.

V. EDUCATION AND PROMOTION SIGNALING – PART 2

In this section I analyze the model described in Subsection IV.A – at the beginning of the game each worker privately observes his or her own realization for θ_i and at the end of each sub-period each firm privately observes the outputs of its workers. I start this analysis with some preliminary results. Consider first sub-period 1. Behavior in this sub-period has a number of similarities with behavior in sub-period 1 in the second preliminary analysis of the previous section. First, there is again a critical value for academic ability, in this case call it θ'' , such that worker i chooses schooling when $\theta_i \geq \theta''$ and no schooling when $\theta_i < \theta''$. Second, because in sub-period 1 only the schooling decision is publicly observable, in sub-period 1 there are again two wages – a schooling wage and a no-schooling wage. Third, these wages take the same form as in that second preliminary analysis. That is, the schooling wage is given by $(1/2)[B + ((\theta^H + \theta'')/2) + \Delta] + \lambda^S$ while the no-schooling wage equals $(1/2)[B + (\theta''/2)] + \lambda^{NS}$, although the specific values for λ^S and λ^{NS} are different.

Now consider sub-period 2. Because of firm specific human capital there is no turnover. But other than this feature of the equilibrium, sub-period 2 equilibrium behavior is quite different than equilibrium behavior in either of the preliminary analyses of the previous subsection. The reason is that, because of asymmetric learning after labor market entry, only a worker's current employer observes the worker's output in sub-period 1 and so only the current employer learns the worker's true value for θ_i^P at the end of sub-period 1. The result is that, as was the case in

Section III, a firm's promotion decisions are used as a signal of worker ability by prospective employers where in this case it is productive ability that is being signaled.

From the analyses in the previous section we know what constitutes efficient behavior in the second sub-period. That is, worker i is assigned to job 2 if $\theta_i^P \geq 1$, while the assignment is to job 1 if $\theta_i^P < 1$. I now describe actual equilibrium behavior in sub-period 2 given asymmetric learning. There are two critical values for θ_i^P , denoted $\theta_S^{P'}$ and $\theta_{NS}^{P'}$, that determine which workers get promoted, where $\theta_S^{P'}$ determines promotion decisions for workers with schooling and $\theta_{NS}^{P'}$ determines promotion decisions for workers without schooling. Specifically, if $s_i=1$, then worker i is promoted at the beginning of sub-period 2 when $\theta_i^P \geq \theta_S^{P'}$ and remains in job 1 in sub-period 2 if $\theta_i^P < \theta_S^{P'}$. Similarly, if $s_i=0$, then worker i is promoted at the beginning of sub-period 2 when $\theta_i^P \geq \theta_{NS}^{P'}$ and remains in job 1 in sub-period 2 if $\theta_i^P < \theta_{NS}^{P'}$. Further, analysis yields that these critical values satisfy $1 < \theta_S^{P'} < \theta_{NS}^{P'}$. That is, fewer workers are promoted than is efficient and this inefficiency is larger for workers without schooling.

These results build on Section III's analysis. In Section III it was shown that when promotions serve as a signal fewer workers are promoted than is efficient and the size of the distortion falls with the worker's education level. Here we find the same qualitative result – fewer workers are promoted than is efficient and the severity of the inefficiency is smaller for workers with schooling than for workers without schooling. Also, the logic for the finding here is the same as for the similar result in Section III. First, the signaling role of promotion results in promoted workers receiving a higher wage (which I discuss further below), so in order to reduce the aggregate wage bill firms promote fewer workers than is efficient. Second, the signaling role of promotion is smaller for workers with more education because the non-promotion wage for these workers increases due to the education signal, so the incentive to distort the promotion decision is smaller for these workers.

In the following proposition I provide a fuller characterization of equilibrium behavior in this setting.

Proposition 5 : Given each worker privately observes his or her own realization for θ_i at the beginning of the game and at the end of each sub-period firms privately observe their own workers' outputs, there exist values θ'' , $\theta_S^{P'}$, $\theta_{NS}^{P'}$, λ^S , and λ^{NS} such that i) through v) describe equilibrium behavior.

- i) Worker i chooses $s_i=1$ if $\theta_i \geq \theta''$ and chooses $s_i=0$ if $\theta_i < \theta''$.
- ii) In sub-period 1 all workers are assigned to job 1, where worker i is paid $(1/2)[B + ((\theta^H + \theta'')/2) + \Delta] + \lambda^S$ when $s_i=1$ and is paid $(1/2)[B + (\theta''/2)] + \lambda^{NS}$ when $s_i=0$.
- iii) In sub-period 2 worker i is assigned to job 1 and is paid $(1/2)(B - \alpha)$ if $s_i=0$ and $\theta_i^P < \theta_{NS}^{P'}$, while the worker is assigned to job 2 and is paid $(1/2)(B - 1 + 2\theta_{NS}^{P'})$ if $s_i=0$ and $\theta_i^P \geq \theta_{NS}^{P'}$.
- iv) In sub-period 2 worker i is assigned to job 1 and is paid $(1/2)(B + \theta'' - \alpha + \Delta)$ if $s_i=1$ and $\theta_i^P < \theta_S^{P'}$, while the worker is assigned to job 2 and is paid $(1/2)(B - 1 + 2\theta_S^{P'} + \Delta)$ if $s_i=1$ and $\theta_i^P \geq \theta_S^{P'}$.
- v) $0 < \theta'' < \theta^*$, $1 < \theta_S^{P'} < \theta_{NS}^{P'}$, $\lambda^S > 0$, and $\lambda^{NS} > 0$.

Proposition 5 provides additional results. One additional result is that in sub-period 2 wages are determined by the winner's curse as was the case in the main analyses in Section III. For example, consider sub-period 2 and worker i for whom $s_i=0$ and $\theta_i^P < \theta_{NS}^{P'}$. As indicated in iii) of the proposition, this worker is paid $(1/2)(B - \alpha)$ which is the amount an alternative employer is willing to bid for the worst worker with this labor market history, i.e., a worker with effective productive ability equal to $-\alpha$. The logic is that in sub-period 2 for every such worker the worker's sub-period 1 employer matches the market wage offer equal to $(1/2)(B - \alpha)$ and the worker stays. The sub-period 2 wages for other workers also follow from this basic logic.

Note that these wage results are consistent with the earlier discussion concerning why there is a promotion signaling distortion and why this distortion is reduced for workers who choose schooling, i.e., $1 < \theta_S^{P'} < \theta_{NS}^{P'}$. As discussed, the distortion arises because promotions are associated with higher wages and so firms do not promote workers who are only slightly more productive in job 2 in order to reduce the aggregate wage bill. As an example that promotion

wages are higher consider sub-period 2 and workers without schooling. The non-promotion wage is $(1/2)(B-\alpha)$ while the promotion wage is $(1/2)(B-1+2\theta_{NS}^{P'})$. Given $\theta_{NS}^{P'} > 1$, we have that the promotion wage exceeds the non-promotion wage.

As indicated, we can also use the results concerning wages to see why workers with schooling are favored in the promotion process. As just mentioned, the non-promotion wage for workers without schooling is $(1/2)(B-\alpha)$, while the non-promotion wage for workers with schooling is $(1/2)(B+\theta''-\alpha+\Delta)$. Suppose, for example, that workers with schooling are not favored in the promotion process and instead both sets of workers are subject to the same minimum value for the productive ability required for promotion, i.e., $\theta_S^{P'} = \theta_{NS}^{P'}$. The logic of the winner's curse then tells us that the promotion wage in sub-period 2 would be $\Delta/2$ higher for workers with schooling than for those without. But then, given the non-promotion wage for workers with schooling exceeds the non-promotion wage for workers without schooling by more than $\Delta/2$, i.e., $(1/2)(B+\theta''-\alpha+\Delta) - (1/2)(B-\alpha) > \Delta/2$, the return to promoting a worker with schooling whose productive ability equals $\theta_S^{P'}$ would exceed the return to promoting a worker without schooling who has productive ability equal to $\theta_{NS}^{P'}$ which contradicts the supposition that $\theta_S^{P'} = \theta_{NS}^{P'}$. Further, since there is a similar contradiction given $\theta_S^{P'} > \theta_{NS}^{P'}$, it must be the case that $1 < \theta_S^{P'} < \theta_{NS}^{P'}$.

The last result in Proposition 5 I will discuss is the first part of condition v) which states that $\theta'' < \theta^*$, i.e., there is again overinvestment in education. This result is similar to the overinvestment result in the second preliminary analysis of the previous subsection in which there was symmetric learning after labor market entry. That is, because there are signaling returns to choosing schooling, the total return to schooling is higher than the direct productivity increase, $(z+1)(\Delta/2)$, associated with schooling. As a result, there is overinvestment in schooling relative to the first best.

The last step of the analysis is to consider the education signaling returns in more detail.

Corollary 1 to Proposition 5: When a worker chooses schooling, there are three education signaling returns which are captured in i) through iii).

- i) The sub-period 1 wage increases by more than $(\Delta/2)+\lambda^S-\lambda^{NS}$.
- ii) The sub-period 2 wage for non-promoted workers increases by more than $\Delta/2$.
- iii) The probability of promotion in sub-period 2 rises.

Corollary 1 to Proposition 5 states that there are three returns to education signaling in this model. The first is that the starting wage increases more with schooling than the sum of the direct productivity increase associated with schooling plus the change in sub-period 2 expected profits associated with hiring a worker with schooling in sub-period 1 rather than a worker without schooling, i.e., $(1/2)[B+((\theta^H+\theta'')/2)+\Delta]+\lambda^S-(1/2)[B+(\theta''/2)]-\lambda^{NS}>(1/2)\Delta+\lambda^S-\lambda^{NS}$. The logic is that by choosing schooling a worker pools with workers with higher values for θ_i and this increases the sub-period 1 wage. This return is analogous to the standard return to education signaling found in numerous prior theoretical analyses and found in the analysis with symmetric learning after labor market entry in the previous subsection.

The second return to education signaling in Proposition 5 is that the wage for non-promoted old workers increases more with schooling than the direct productivity increase associated with schooling, i.e., $(1/2)[B+(\theta''-\alpha)+\Delta]-(1/2)(B-\alpha)=(1/2)(\theta''+\Delta)>\Delta/2$. As was the case for the first return, the logic is that by choosing schooling a worker pools with higher θ_i types and the result is a higher non-promotion wage for old workers.¹⁵

The third return to education signaling is that by choosing schooling a worker lowers the minimum value for productive ability required to earn a promotion, i.e., $\theta_S^{P'}<\theta_{NS}^{P'}$. Another way to put this is that, because of the signaling role of education, choosing schooling increases the probability of promotion. As discussed above, the logic for this result is that education signaling

¹⁵ For promoted old workers the wage is determined by the promotion signal and the idea that higher θ_i types choose schooling does not directly affect the wage.

causes the wage for non-promoted workers with schooling in sub-period 2 to be higher and this reduces the incentive for the current employer to distort the promotion decision.^{16,17}

As already mentioned, an important aspect of the additional returns to education signaling is that they translate into increased compensation late rather than early in workers' careers. That is, in standard education signaling models the return to education signaling is higher wages early in workers' careers and little or no return late in careers. This is because by late in workers' careers firms will have learned workers' abilities directly by observing outputs and thus in terms of compensation little weight is placed on the education signal. But in the model analyzed in this section there are additional educational signaling returns – higher wages for non-promoted old workers and higher promotion probabilities for old workers – that translate into higher compensation late in careers.¹⁸ As I discuss in detail in the next section, this result has important implications for studies that attempt to measure the returns to education signaling.

VI. DISCUSSION

In this section I discuss two issues. In the first subsection I discuss the social welfare implications of the analysis in Section V which combines education signaling with promotion signaling. In the second subsection I discuss what this analysis implies for studies focused on measuring the returns to education signaling.

¹⁶ Remember that, as was found in the benchmark analysis of the previous subsection, given the way schooling enters the production function, in the efficient outcome choosing schooling does not affect a worker's promotion probability.

¹⁷ One might think that, because there are additional education signaling returns in this model relative to the second preliminary analysis of the previous section, there should be more overinvestment in education here, i.e., $\theta'' < \theta' < \theta^*$. But that logic is incomplete. Education signaling in the analysis here affects the difference $\lambda^S - \lambda^{NS}$ and can make this difference smaller than in the second preliminary analysis of the previous section. The result is an ambiguous relationship between θ' and θ'' .

¹⁸ Part of the third return to education signaling in this section's model can be a higher wage in the first sub-period. The logic is that the promotion signal results in more efficient job assignments in sub-period 2. If these more efficient sub-period 2 job assignments translate into higher firm profits in sub-period 2, then this third return to education signaling will also result in higher sub-period 1 wages for workers with schooling due to the competition in sub-period 1 for the hiring of these workers.

A) Social Welfare Implications

In standard models of education signaling such as the models investigated in Section II the increase in educational investments due to education serving as a signal of worker ability unambiguously reduces social welfare. This is obviously the case when comparing the equilibrium outcome to a first best outcome. For example, compare the benchmark case of full information for either of the models analyzed in Section II with the equilibrium outcome when there is asymmetric information concerning workers' ability levels and schooling serves as a signal. In the benchmark case workers choose schooling if and only if the increase in worker productivity due to schooling is greater than or equal to the worker's cost of schooling. In contrast, when there is asymmetric information some workers choose schooling even though the increase in productivity is less than the cost. So social welfare which I am defining to be the sum of outputs produced minus aggregate costs spent on schooling is lower when there is signaling.

Now suppose that instead of comparing social welfare given signaling to the full information outcome, we compared it to what happens when there is asymmetric information but workers are constrained to choose the same education levels as in the full information outcome. It would still be the case that in the signaling outcome social welfare would be lower because the costs of the higher education levels are smaller than the productivity increases due to those higher education levels. Another way to put this is that, if we introduced a social welfare maximizing social planner into the asymmetric information case and that social planner was only able to choose education levels, the planner would make the same education choices as in the full information analysis, so the higher education levels seen in the absence of the planner must reduce welfare.

Now consider the model investigated in Sections IV and V. Let's start with the two preliminary analyses considered in Section IV. In the first preliminary analysis there was no private information, while in the second the only private information was about workers' academic abilities. So in the first preliminary analysis there was neither education signaling nor promotion signaling, while the second preliminary analysis was characterized by education

signaling but not promotion signaling. Further, in that first preliminary analysis both education and promotion decisions were first best efficient.

A comparison of those two preliminary analyses yields the same conclusion as the comparison between the full information and signaling analyses of Section II. That is, education levels are higher when education serves as a signal and these higher education levels reduce social welfare because the direct productivity increase due to the higher education levels is less than the costs of the increased education levels. So, given only education signaling, social welfare is lower than in a first best outcome. Further, this conclusion is basically unchanged if we compare the signaling equilibrium to the case of asymmetric information concerning academic ability levels in which a social planner chooses the education levels. Specifically, the social planner would choose the first best education levels and the higher education levels in the signaling equilibrium would again serve to lower social welfare.

Now consider the analysis in Section V in which education serves as a signal of academic ability while promotions serve as a signal of productive ability. A comparison with the first preliminary analysis still yields that signaling reduces social welfare, where this is now the case for both types of signaling. Education signaling reduces social welfare because the increased productivity due to signaling is less than the increase in education costs, while promotion signaling reduces social welfare because some workers are not promoted due to promotion signaling and for these workers a promotion would increase productivity.

But now consider a second best comparison which focuses on the social welfare effects of education signaling given that the setting is characterized by promotion signaling. That is, suppose there is a social welfare maximizing social planner who chooses education levels but promotions are chosen by employers. In contrast to the cases discussed above, this social planner would not choose the first best education levels. Rather, this social planner would have more workers choose education than in the first best because the promotion distortion is smaller for workers with education. In other words, from a second best standpoint the optimal investment in education is not determined solely by the direct productivity increase associated with education

but also includes the more efficient job assignments associated with more education. So the higher education levels found in Proposition 5 relative to the first best do not unambiguously decrease social welfare.

Another way to describe this result is that, in contrast to a standard education signaling model, in Section V's analysis the increase in education levels due to education signaling relative to the first best education levels has both positive and negative effects on social welfare. For workers who do not earn a promotion in the Proposition 5 equilibrium the increases in education due to education signaling decreases social welfare because in aggregate the increased costs of education exceed the increase in productivity. Similarly, for workers who are promoted in the Proposition 5 equilibrium who would also be promoted given the lower first best education levels, the higher education levels given signaling again reduce social welfare because in aggregate the added education costs exceed the higher productivity. In other words, for these two groups of workers the social welfare effects of the increase in education levels due to signaling is basically the same as in standard education signaling models. But now there is another effect which is that the higher education levels due to signaling causes some workers to be promoted who would not be promoted at the lower first best education levels. And these promotions enhance efficiency where the reason is that promotion signaling causes too few workers to be promoted and so when the number of promotions increases due to education signaling the efficiency of the promotion process improves.

In summary, from a social welfare standpoint the returns to education signaling given there is both education and promotion signaling are different than in a standard model of education signaling. In a standard model the increased education levels due to education signaling have an unambiguous negative effect on social welfare. In contrast, given a setting characterized by both education and promotion signaling, a second best perspective yields that

the increased education levels due to education signaling have both positive and negative effects on social welfare.¹⁹

B) Measuring the Returns to Education Signaling

As discussed briefly in the Introduction and at the end of Section II, the empirical literature on education signaling has moved from testing whether real world labor markets are characterized by education signaling to estimating the returns to education signaling in real world labor markets. In this subsection I discuss the implications of Section V's analysis for this type of estimation.

The key paper in this literature is Lange (2007). Building on earlier analyses in Farber and Gibbons (1997) and Altonji and Pierret (2001), Lange builds a model of education signaling similar to the model investigated in Section II.B in which there is symmetric learning concerning worker ability levels after labor market entry. The main difference is that, in contrast to the analysis in II.B in which learning is complete by the end of the first sub-period, in Lange's model the learning is gradual because of a stochastic element in the production function. Lange's goal is to investigate the implications of the argument in Altonji and Pierret (1997) that, because of learning after labor market entry, education signals should be relatively unimportant in terms of compensation late in workers' careers because firms will have learned each worker's ability directly by observing the worker's performance on the job.

Lange (2007) employs a structural estimation approach using data from the National Longitudinal Survey of Youth and finds two related results. First, he finds that learning is quick.

¹⁹ One question of interest is, if we employ this second best perspective, is the net effect on social welfare due to the higher signaling education levels relative to the first best levels ambiguous or is the net effect always positive or always negative? In a slightly enriched version of the model analyzed in Section V it is possible to show that this net effect can be either positive or negative. For example, if production functions are allowed to vary, then it can be shown that the net effect is negative when the two jobs are sufficiently similar because then inefficiencies due to the promotion signaling distortion are limited. Also, if one relaxes the assumption that the academic ability distribution is uniform, then it is possible to construct examples where the net effect is positive where one situation in which this occurs is when there is a large population weight on academic ability levels just below the first best critical value for academic ability, i.e., values just below θ^* . The logic is that for these workers the net return is positive because the direct increase in productivity due to education is close to the cost, so if these workers are sufficiently important in the overall population then the aggregate net effect will also be positive.

For example, he finds that on average it takes about three years for initial expectation errors by employers concerning workers' ability levels to decline by half and after five years it is roughly a third of the initial value. Second, he finds across a wide range of parameterizations that less than twenty five percent of the overall return to schooling is due to education signaling and in his preferred specification approximately ten percent of this overall return is due to signaling.²⁰

My point is that these results concerning the overall returns to education signaling are dependent on the assumption of symmetric learning after labor market entry and that results might be quite different given asymmetric learning and promotion signaling. With the symmetric learning assumption the education signal becomes less and less important as workers gain labor market experience because learning is cumulative and the return to the education signal falls with the amount of other information firms have about workers' ability levels. Lange finds that learning is quick and that firms have little remaining uncertainty about workers' ability levels by the end of workers' careers. So education signaling returns in this approach are found to be heavily concentrated early rather than late in careers and, because learning is quick, even early in careers these returns drop off quickly.

But the analysis in Section V tells us that returns to education signaling are quite different given asymmetric learning after labor market entry and promotion signaling (and as discussed at the end of Section III most of the recent empirical literature supports the asymmetric learning and promotion signaling approach). The reason is that with asymmetric learning and promotion signaling there are important returns to education signaling that manifest themselves as higher compensation late rather than early in workers' careers. These returns are not captured in a study that assumes symmetric learning because, in a sense, the assumptions of the symmetric learning approach force the estimation procedure to find a small return to education signaling late in

²⁰ Kaymak (2014) uses a similar approach but allows the speed of employer learning to vary across occupations and finds a somewhat larger role for signaling.

careers. So if those returns are significant, then the type of approach employed by Lange is likely to produce an underestimate of the returns to education signaling.

Let me end this discussion by posing a thought experiment. Consider resume padding which is the practice of adding false or exaggerated information to a resume such as claiming to have earned an MBA when this is not the case. If the Altonji and Pierret (1997) argument that Lange (2007) builds on is correct, that would mean there is little return to education signaling after the first few years of workers' careers. And this, in turn, would mean little return to resume padding concerning education credentials late in careers because resume padding should only be beneficial to the extent that current and prospective employers use the exaggerated information as signals. In contrast, Section V's analysis suggests that resume padding concerning education credentials can be beneficial even late in careers if the probability of getting caught is sufficiently small. What does the evidence suggest? It says that resume padding is common and there are a number of well known examples of resume padding by high level executives with significant labor market experience such as Scott Thompson who was Yahoo's CEO, Kenneth Lonchar who was CFO at Veritas Software, Ronald Zarella who was CEO at Bausch and Lomb, and Jeffrey Papows who was president of IBM's software maker Lotus Development.²¹

VII. CONCLUSION

Previous literature suggests that the two main avenues through which worker ability is signaled in the labor market are education decisions and promotion decisions. But there has been little attention paid to how these dual avenues of labor market signaling interact. In this paper I have constructed and analyzed a model characterized by both education and promotion signaling,

²¹ See Winstein (2008) for a discussion. In the first two of these cases the executive was forced to resign when the padding was discovered, while in the other cases the executive was allowed to stay on at least until the next scandal (Jeffrey Papows resigned a year after the resume padding was discovered after a sexual discrimination complaint was made). One possible way to reconcile these examples with Lange's (2007) findings is that sometimes an individual starts padding his or her resume early in the individual's career and decides not to remove the false information later because removal might draw attention to the earlier misinformation. With this in mind, it might be interesting for some of these more prominent examples of resume padding by high level executives to identify at what point in time the incorrect information was added to the individual's resume.

where much of the focus has been on how the presence of promotion signaling affects the returns to education signaling.

One innovation in the model investigated here is that there are two types of worker ability – academic ability and productive ability – where higher academic ability reduces the cost of schooling while higher productive ability means the worker is more productive in the workplace. In the model the two types of ability are correlated but not perfectly, so even when academic ability is known with certainty there can still be uncertainty concerning a worker's productive ability. The result is that in equilibrium both types of signaling are operative. Workers are privately informed about their own academic abilities and education choices serve as a signal of this private information, where as in standard education signaling models there is overinvestment in education relative to the first best. Firms, on the other hand, privately learn their own employees' productive abilities by observing output realizations and promotion decisions serve as a signal of this private information, where as in standard promotion signaling models too few workers are promoted relative to the first best.

In addition to showing that the two types of signaling can co-exist, the analysis shows that there are three returns to education signaling in this model. First, there is the standard return that more education signals higher ability (in this case academic ability) and this translates into higher wages early in workers' careers. Second, because learning after labor market entry is asymmetric rather than symmetric, the idea that more education signals higher ability also translates into higher wages for non-promoted workers late in workers' careers. Third, because of the interaction between education signaling and promotion signaling, a third education signaling return is higher promotion probabilities that also translate into higher wages late in workers' careers.

I also argued that these results have two important implications concerning how one should think about education signaling. The first is that, in contrast to what is true in standard models of education signaling, the higher educational investments due to education signaling are not unambiguously negative from a social welfare standpoint because from a second best

perspective the increase in promotion probabilities improve efficiency by reducing the social welfare costs of promotion signaling. The second is that estimates of the returns to education signaling such as in Lange (2007) are likely underestimates because they implicitly assume that the returns to education signaling are close to zero late in workers's careers while the analysis here shows this is not the case given asymmetric learning and promotion signaling.

There are a number of directions in which the analysis in this paper can be extended. In terms of theory one possibility is to add a stochastic element to the production functions while a second is to introduce more more sub-periods and more job levels. The former would make the model more realistic by making the learning process more gradual, but my belief is that such an extension would not change the qualitative nature of the results.

On the other hand, I think that adding more sub-periods and more job levels could turn out to be quite important because this addition would likely reinforce the finding that there are important education signaling returns late in workers' careers. The idea is that, if more highly educated workers are favored in the promotion process as was true in Section V's analysis, then it should be the case that more highly educated workers are promoted to the top rungs of job ladders much more often than is justified by their higher average level of productivity. In turn, if very high levels of compensation are associated with these high level positions which seems frequently to be the case, then the education signaling returns due to higher promotion probabilities might be even more important than is suggested by the analysis in Section V.

In terms of empirical work it would be interesting to conduct a structural estimation exercise similar to the one in Lange (2007) but allow for asymmetric learning and promotion signaling. In this paper I have identified a potential bias in the type of empirical exercise conducted by Lange in which a structural estimation approach is used to estimate the returns to education signaling where Lange assumes symmetric learning after labor market entry. Taking a similar structural estimation approach but assuming asymmetric learning and promotion signaling after labor market entry should be a way to learn about the size of this bias.

APPENDIX

Proof of Proposition 1: Since only schooling levels are publicly observable, there are only two wages in equilibrium where one is for workers with schooling and the other for workers without schooling. Further, combining this with the fact that the cost of schooling falls with worker ability means there is a critical value for θ_i , denoted θ' , such that worker i chooses schooling when $\theta_i \geq \theta'$ and chooses no schooling when $\theta_i < \theta'$ (see footnote 4). This proves i).

The next step is to prove that $0 < \theta' < \theta^H$. Suppose $\theta' \geq \theta^H$. Starting from this equilibrium, if a worker did choose schooling, given our off-the-equilibrium path assumption concerning beliefs, firms would anticipate expected productivity equal to $\Delta + \theta^H$ so this is the schooling wage, while the zero profit condition yields that the wage for a worker with no schooling is $\theta^H/2$. A $\theta^H - \varepsilon$ worker would thus have a net return to choosing schooling that approaches $\Delta + (\theta^H/2) - c(\theta^H)$ as ε approaches zero. This expression is strictly positive since $c(\theta^H) < \Delta$. So $\theta' < \theta^H$.

Suppose $\theta' \leq 0$. Then the schooling wage is $\Delta + (\theta^H/2)$. Further, our off-the-equilibrium path assumption concerning beliefs yields that the no-schooling wage is zero. So the net return to schooling for worker i equals $\Delta + (\theta^H/2) - c(\theta_i)$. But for θ_i sufficiently close to zero this expression is strictly negative since $c(0) > \Delta + (\theta^H/2)$. This contradicts $\theta' \leq 0$. So $\theta' > 0$.

Given free entry, wages are determined by zero expected profits. Further, given the uniform distribution over the interval $[0, \theta^H]$ for θ_i now yields that the schooling wage equals $\Delta + ((\theta^H + \theta')/2)$ and the no-schooling wage equals $\theta'/2$. This proves ii).

The value for θ' is the value for θ_i such that the worker is indifferent between schooling and no schooling. Given the wages just derived, this means θ' satisfies $\Delta + ((\theta^H + \theta')/2) - (\theta'/2) = c(\theta')$. This reduces to $\Delta + (\theta^H/2) = c(\theta')$. Since $\Delta + (\theta^H/2) > \Delta$ and $c' < 0$, we now have $\theta' < \theta^*$. This proves iii).

Proof of Proposition 2: At the end of sub-period 1 all firms observe each worker's sub-period 1 output. Given output in sub-period 1 is deterministic and equals $(1/T)(\theta_i + s_i \Delta)$ for worker i and s_i

is publicly observable, at the beginning of sub-period 2 each worker i 's value for θ_i is known by all employers. Given free entry, wages are determined by zero expected profits. Given this and that both s_i and θ_i are known by all firms at the beginning of each sub-period t , $t > 1$, the wage for worker i in each sub-period t , $t > 1$, equals $(1/T)(\theta_i + s_i \Delta)$. This proves iii). Note that this means that the benefit associated with schooling for each sub-period t , $t > 1$, is independent of θ_i .

Now consider sub-period 1. Because at the beginning of sub-period 1 only the schooling decision is publicly observable, in sub-period 1 there are two wages where one is for workers with schooling and one is for workers without schooling. Combining this with the wage results for later sub-periods yields that the benefit of schooling is independent of θ_i . Given that the cost of schooling falls with θ_i , we again have that there is a critical value for θ_i , now called θ_T' , such that worker i chooses $s_i = 1$ if $\theta_i \geq \theta_T'$ and chooses $s_i = 0$ if $\theta_i < \theta_T'$ (see footnote 4). This proves i). Further, using arguments like those in the proof of Proposition 1, it must be the case that $0 < \theta_T' < \theta^H$ for all T , $T \geq 1$.

Free entry again means wages are determined by zero expected profits. Given the schooling decisions just derived, that only the schooling decisions are observable at the beginning of sub-period 1, and $0 < \theta_T' < \theta^H$ for all T , $T \geq 1$, the schooling wage in sub-period 1 is given by $(1/T)[\Delta + ((\theta^H + \theta_T')/2)]$ while the no-schooling wage is given by $(1/T)(\theta_T'/2)$. This proves ii).

The value for θ_T' is the value for θ_i such that worker i is indifferent between schooling and no schooling. Given the wages just derived, this means θ_T' satisfies equation (2). Given $(1/T)(\theta^H/2) > 0$ and $c' < 0$, we have $\theta_T' < \theta^*$ for all T . Also, since $(1/T)(\theta^H/2)$ decreases with T and $c' < 0$, we have that θ_T' increases with T . And since $(1/T)(\theta^H/2)$ approaches zero as T gets large, θ_T' approaches θ^* as T gets large. This proves iv).

Proof of Proposition 3: Consider a firm that hires a worker in sub-period 1 and assigns the worker to job 1. The worker produces $(1/2)(B + \theta_i)$ and when the firm observes the worker's output at the end of sub-period 1 the firm is able to infer θ_i . Given this, consider sub-period 2.

Given the counteroffer assumption and the presence of firm specific human capital, the firm matches the wage offers of alternative employers and the worker stays. Further, since alternative employers observe sub-period 2 job assignments but not worker ability levels, there are two market wage offers where one is for workers assigned by the sub-period 1 employer to job 1 and one for workers assigned to job 2. Further, since this means the cost to the initial employer of assigning a worker to job 2 rather than job 1 is independent of θ_i while the change in productivity associated with this behavior is increasing in θ_i , there must be a critical value, denoted θ' , such that the sub-period 1 employer assigns the worker to job 2 at the beginning of sub-period 2 if $\theta_i \geq \theta'$ and assigns the worker to job 1 if $\theta_i < \theta'$ (see footnote 4).

The next step is to show that $1 < \theta' < \theta^H$. Suppose a worker is assigned to job 1 in sub-period 2 with probability one. In this case alternative employers bid $(1/2)B$ due to the winner's curse and the initial employer earns $z(1/2)(B+\theta_i)-(1/2)B$ from employing worker i in sub-period 2. Suppose $\theta_i = \theta^H - \varepsilon$ and let ε approach zero. If the initial employer instead assigned the worker to job 2, then the market wage offer would be $(1/2)(B-1+2\theta^H)$ given the assumption made concerning off-the-equilibrium path moves. The amount the initial employer therefore earns from assigning the worker to job 2 approaches $(z-1)(1/2)(B-1+2\theta^H)$ as ε approaches zero. Given $\theta^H > 1$ and z sufficiently large, this expression is larger than the profit from assigning the worker to job 1 (see footnote 5). This shows that, starting from a situation in which a worker is assigned to job 1 in sub-period 2 with probability one, an initial employer would have an incentive to deviate and assign a worker with ability $\theta^H - \varepsilon$, ε small, to job 2 at the beginning of sub-period 2. So $\theta' < \theta^H$. A similar argument yields $\theta' > 0$.

We now have that $0 < \theta' < \theta^H$. Wages are again determined by the winner's curse and because of firm specific human capital there is no turnover in equilibrium. We thus have that if a worker with ability θ' is not promoted the worker is paid $(1/2)B$ and if the worker is promoted the worker is paid $(1/2)\max\{B+\theta', B-1+2\theta'\}$. Suppose $\theta' \leq 1$. Then a firm earns $z(1/2)(B+\theta')-(1/2)B$ from assigning the worker to job 1 and earns $z(1/2)(B-1+2\theta')-(1/2)(B+\theta')$ from assigning the worker to job 2. If $\theta' \leq 1$, the first expression is strictly larger which is a contradiction. So $\theta' > 1$

and $\max\{B+\theta', B-1+2\theta'\}=B-1+2\theta'$. Assuming i) is correct which we finish the proof of below, we have now shown ii) and iii).

Consider again sub-period 1. If a firm hires a worker and assigns the worker to job 2, then sub-period 1 output equals zero and behavior in sub-period 2 is exactly as above (see footnote 5). So after hiring a worker in sub-period 1 the firm maximizes its profits by assigning the worker to job 1 in sub-period 1. Further, free entry means that the sub-period 1 wage is determined by zero expected profits. Given from above we know that there are strictly positive expected sub-period 2 profits from hiring a worker in sub-period 1, the result that workers are assigned to job 1 in sub-period 1 yields that the sub-period 1 wage can be expressed as $(1/2)[B+(\theta^H/2)]+\lambda$, $\lambda>0$, where $(1/2)[B+(\theta^H/2)]$ is expected sub-period 1 output and λ is the expected sub-period 2 profit. This completes the proof of i), ii), and iii).

The last step is to finish the proof of iv). I have already shown that $1=\theta^*<\theta'<\theta^H$ and $\lambda>0$. θ' is the value for θ_i such that the initial employer is indifferent between assigning worker i to jobs 1 and 2 in sub-period 2. Given earlier results, θ' satisfies (A1).

$$(A1) \quad z(1/2)(B-1+2\theta')-(1/2)(B-1+2\theta')=z(1/2)(B+\theta')-(1/2)B$$

(A1) yields $\theta'=(z-1)/(z-2)$ which decreases with an increase in z . This completes the proof of iv).

Proof of Proposition 4: With slight modification all the parts of Proposition 4 follow from arguments in the proof of Proposition 3 except the result in iv) which states that $\theta'(m+1)<\theta'(m)$ for all $m=1,\dots,M-1$. Given this, consider group m , where $1\leq m\leq M-1$. In the case of multiple schooling groups equation (A1) translates into (A2).

$$(A2) \quad z(1/2)[B-1+2\theta'(m)]-(1/2)[B-1+2\theta'(m)] \\ =z(1/2)[B+\theta'(m)]-(1/2)[B+x(m)] \text{ for all } m, m=1,\dots,M$$

(A2) yields $\theta'(m)=(z-1-x(m))/(z-2)$. Given $x'>0$, we now have that $\theta'(m+1)<\theta'(m)$ for all $m=1,\dots,M-1$. This completes the proof.

Proof of Proposition 5: Using arguments similar to arguments in the proof of Proposition 3 it can be shown that a firm that hires a worker in sub-period 1 will assign the worker to job 1. Let θ'' denote the minimum value for academic ability such that the worker chooses schooling.

Arguments similar to those given in the proof of Proposition 1 yield $0 < \theta'' < \theta^H$.

Suppose $s_i = 0$ and consider what happens in sub-period 2. At the end of sub-period 1 the sub-period 1 employer observes the worker's sub-period 1 output and from this observation learns the worker's value for θ_i^P . Further, because prospective employers only observe the worker's schooling level and job assignment prior to making wage offers to this worker, there are two such wage offers where one is for workers assigned to job 1 and one is for workers assigned to job 2. This means the cost in terms of the sub-period 2 wage associated with the initial employer assigning the worker to job 2 rather than job 1 is independent of θ_i^P . Given the production functions for the two jobs, this means there must be a critical value for θ_i^P , call it $\theta_{NS}^{P'}$, such that the worker is assigned to job 2 if $\theta_i^P \geq \theta_{NS}^{P'}$ and is assigned to job 1 if $\theta_i^P < \theta_{NS}^{P'}$ (see footnote 4). Using a similar logic there must also be a similar critical value for workers with schooling. Call this critical value $\theta_S^{P'}$.

Based on the logic of the winner's curve, we can now derive sub-period 2 wages. Alternative employers always bid the lowest possible productivity given the worker's labor market history and then the sub-period 1 employer matches and the worker stays. Consider a worker without schooling. This logic yields that if the worker is assigned to job 1 then the wage equals $(1/2)(B - \alpha)$, while if the worker is assigned to job 2 then the wage equals $(1/2)\max\{B + \theta_{NS}^{P'}, B - 1 + 2\theta_{NS}^{P'}\}$. Similarly, consider a worker with schooling. If the worker is assigned to job 1 then the wage equals $(1/2)\max\{B + \theta'' - \alpha + \Delta, B - 1 + 2(\theta'' - \alpha) + \Delta\}$, while if the worker is assigned to job 2 then the wage equals $(1/2)\max\{B + \theta_S^{P'} + \Delta, B - 1 + 2\theta_S^{P'} + \Delta\}$. This proves iii) and iv) if $\theta_{NS}^{P'} > 1$ and $\theta_S^{P'} > 1$. I prove these extra conditions below.

I now turn to condition v). Based on the sub-period 2 wages derived above, a firm that hires a worker with schooling in sub-period 1 and a firm that hires a worker without schooling in

sub-period 1 both earn strictly positive expected sub-period 2 profits from this hiring. Thus, $\lambda^S > 0$ and $\lambda^{NS} > 0$.

Now consider the condition $1 < \theta_S^{P'} < \theta_{NS}^{P'}$. $\theta_{NS}^{P'}$ is the value for θ_i^P such that a firm is indifferent in sub-period 2 between assigning a worker without schooling that it employed in sub-period 1 to jobs 1 and 2. Based on sub-period 2 wages derived above, we now have equation (A3).

$$(A3) \quad z(1/2)(B-1+2\theta_{NS}^{P'})-(1/2)\max\{B+\theta_{NS}^{P'}, B-1+2\theta_{NS}^{P'}\}=z(1/2)(B+\theta_{NS}^{P'})-(1/2)(B-\alpha)$$

If $\theta_{NS}^{P'} \leq 1$, then $z(1/2)(B-1+2\theta_{NS}^{P'}) \leq z(1/2)(B+\theta_{NS}^{P'})$. We also know that $(1/2)\max\{B+\theta_{NS}^{P'}, B-1+2\theta_{NS}^{P'}\} \geq (1/2)(B-\alpha)$. And we know both conditions cannot simultaneously hold as equalities.

Together these relationships tell us that (A3) cannot be satisfied if $\theta_{NS}^{P'} \leq 1$, so $\theta_{NS}^{P'} > 1$.

The equation that defines $\theta_S^{P'}$ is given in (A4)

$$(A4) \quad z(1/2)(B-1+2\theta_S^{P'}+\Delta)-(1/2)\max\{B+\theta_S^{P'}+\Delta, B-1+2\theta_S^{P'}+\Delta\} \\ =z(1/2)(B+\theta_S^{P'}+\Delta)-(1/2)\max\{B+(\theta''-\alpha)+\Delta, B-1+2(\theta''-\alpha)+\Delta\}$$

Given $z > (2\theta^H-1)/(\theta^H-1)$ (see footnote 12), a comparison of (A3) and (A4) tells us that $\theta_S^{P'} < \theta_{NS}^{P'}$.

Suppose $\theta_S^{P'} \leq 1$. Then $z(1/2)(B-1+2\theta_S^{P'}) \leq z(1/2)(B+\theta_S^{P'})$. We also know that $(1/2)\max\{B+\theta_S^{P'}, B-1+2\theta_S^{P'}\} \geq (1/2)\max\{B+(\theta''-\alpha), B-1+2(\theta''-\alpha)\}$. And we know both conditions cannot

simultaneously hold as equalities. Together these relationships tell us that (A4) cannot be satisfied if $\theta_S^{P'} \leq 1$. We thus have $1 < \theta_S^{P'} < \theta_{NS}^{P'}$. This completes the proof of iii) and iv).

The next step is to show that $s_i=1$ for every worker i such that $\theta_i \geq \theta''$. The cost of schooling for a θ_i worker is $c(\theta_i)$ while the expected increase in lifetime compensation associated with choosing schooling can be expressed as the sum of five components. One component is $(\Delta/2)+z(\Delta/2)$ which is the direct increase in lifetime productivity due to schooling. A second component is the increase in the sub-period 1 wage above $\Delta/2$ due to the education signal. This equals $(1/2)(\theta^H/2)+\lambda^S-\lambda^{NS}$. A third component is a decrease in the sub-period 2 promotion wage not taking into account the direct increase in productivity due to schooling multiplied by the probability of promotion when the worker chooses no schooling. Call this probability $r(\theta_i)$ for worker i . This third component equals $(1/2)r(\theta_i)(2\theta_S^{P'}-2\theta_{NS}^{P'})$. The fourth component is an

increase in the sub-period 2 non-promotion wage not taking into account the direct increase in productivity due to schooling multiplied by the probability of non-promotion when the worker chooses no schooling. This fourth component equals $(1/2)[1-r(\theta_i)]\theta''$. The fifth component is the increase in the sub-period 2 expected wage that results due to the increase in the probability of promotion, i.e., $1 < \theta_S^{P'} < \theta_{NS}^{P'}$. Let $V(\theta_i)$ denote the increase in the sub-period 2 expected wage due to the increase in the probability of promotion by a worker with academic ability θ_i when the worker chooses schooling.

Let $\rho(\theta_i)$ denote the expected net return to schooling for a worker with academic ability θ_i . A worker with academic ability θ'' must be indifferent between schooling and no schooling. Thus, equation (A5) must be satisfied.

$$(A5) \quad \rho(\theta'') = (\Delta/2) + z(\Delta/2) + (1/2)(\theta^H/2) + \lambda^S - \lambda^{NS} + (1/2)r(\theta'')(2\theta_S^{P'} - 2\theta_{NS}^{P'}) \\ + (1/2)[1-r(\theta'')]\theta'' + V(\theta'') - c(\theta'') = 0$$

By the definition of θ'' the derivative of $\rho(\theta_i)$ with respect to θ_i evaluated at $\theta_i = \theta^H$ must be strictly positive (in taking this derivative the third-to-last θ'' term is taken as a constant because the two non-promotion wages are taken as given from the standpoint of a worker making a schooling choice).

Now consider a worker with academic ability θ_i such that $\theta_i > \theta''$. Equation (A6) shows the expected net return to schooling for this worker.

$$(A6) \quad \rho(\theta_i) = (\Delta/2) + z(\Delta/2) + (1/2)(\theta^H/2) + \lambda^S - \lambda^{NS} + (1/2)r(\theta_i)(2\theta_S^{P'} - 2\theta_{NS}^{P'}) \\ + (1/2)[1-r(\theta_i)]\theta'' + V(\theta_i) - c(\theta_i)$$

Given $\alpha > \theta^H$, for every θ_i there are realizations for θ_i^P below zero and realizations above θ^H and in combination with the α distribution being uniform this means that $V(\theta_i)$ is a constant and that the derivative of $r(\theta_i)$ is also a constant. Given these results, $c'' < 0$, and that the derivative of $\rho(\theta_i)$ with respect to θ_i evaluated at $\theta_i = \theta^H$ is strictly positive, a comparison of (A5) and (A6) yields that the derivative of right hand side of (A6) is strictly positive for all $\theta_i > \theta''$. Given $\rho(\theta'') = 0$, we now have that $\rho(\theta_i) > 0$ for all $\theta_i > \theta''$ which means $s_i = 1$ for all $\theta_i > \theta''$. This proves i). Further, ii) now

follows given that free entry means that sub-period 1 wages are determined by expected productivity.

The last step of the proof is to show that $\theta'' < \theta^*$. Remember that θ^* which is the first best critical value for θ_i for the schooling decision was defined in the first preliminary analysis. It is the value for θ_i such that the benefit from schooling just equals the cost of schooling when the θ_i s are observable. It is thus defined in equation (A7).

$$(A7) \quad (\Delta/2) + z(\Delta/2) = c(\theta^*)$$

Given other results already derived and that θ'' is the value for academic ability such that a worker is indifferent between schooling and no schooling, θ'' must satisfy equation (A8).

$$(A8) \quad (1/2)[B + ((\theta^H + \theta'')/2) + \Delta] + \lambda^S - (1/2)[B + (\theta''/2)] - \lambda^{NS} + (1/2)[r^S(B-1+2\theta_S^{P'} + \Delta) + (1-r^S)(B + (\theta'' - \alpha) + \Delta)] - (1/2)[r^{NS}(B-1+2\theta_{NS}^{P'}) + (1-r^{NS})(B-\alpha)] = c(\theta'')$$

In equation (A8) r^S is the probability a worker with academic ability θ'' with schooling is assigned to job 2 in sub-period 2 while r^{NS} is the probability a worker with academic ability θ'' without schooling is assigned to job 2 in sub-period 2. We know $(1/2)[B + ((\theta^H + \theta'')/2) + \Delta] - (1/2)[B + (\theta''/2)] > \Delta/2$, so a comparison of (A8) with (A7) yields $\theta'' < \theta^*$ if $\lambda^S - \lambda^{NS} + (1/2)[r^S(B-1+2\theta_S^{P'} + \Delta) + (1-r^S)(B + (\theta'' - \alpha) + \Delta)] - (1/2)[r^{NS}(B-1+2\theta_{NS}^{P'}) + (1-r^{NS})(B-\alpha)] > z(\Delta/2)$.

Let $\lambda^S(\theta^k)$ ($\lambda^{NS}(\theta^k)$) be the expected sub-period 2 profit earned by a firm from employing a worker in sub-period 2 with academic ability θ^k and schooling (no schooling) that the firm initially hired in sub-period 1. Also, let $\gamma^S(\theta^k)$ ($\gamma^{NS}(\theta^k)$) be the corresponding sub-period 2 expected productivity for a θ^k worker with (without) schooling. By definition $\lambda^S(\theta'') + (1/2)[r^S(B-1+2\theta_S^{P'} + \Delta) + (1-r^S)(B + (\theta'' - \alpha) + \Delta)] = \gamma^S(\theta'')$, while $\lambda^{NS}(\theta'') + (1/2)[r^{NS}(B-1+2\theta_{NS}^{P'}) + (1-r^{NS})(B-\alpha)] = \gamma^{NS}(\theta'')$. Further, since $1 < \theta_S^{P'} < \theta_{NS}^{P'}$ means sub-period 2 job assignments are more efficient when a worker chooses schooling, we have $\gamma^S(\theta'') - \gamma^{NS}(\theta'') > z(\Delta/2)$. This, in turn yields (A9).

$$(A9) \quad \lambda^S(\theta'') - \lambda^{NS}(\theta'') + (1/2)[r^S(B-1+2\theta_S^{P'} + \Delta) + (1-r^S)(B + (\theta'' - \alpha) + \Delta)] - (1/2)[r^{NS}(B-1+2\theta_{NS}^{P'}) + (1-r^{NS})(B-\alpha)] > z(\Delta/2)$$

Since a firm's sub-period 2 profit from employing a continuing worker with schooling is increasing in θ_i^P and given the manner in which the realizations for θ_i^P increase with θ_i , it must be the case that $\lambda^S(\theta^k) > \lambda^S(\theta'')$ for all $\theta^k > \theta''$. Since λ^S is the expected value for $\lambda^S(\theta^k)$ averaged over all θ^k , $\theta^k \geq \theta''$, we now have that $\lambda^S > \lambda^S(\theta'')$. A similar argument yields $\lambda^{NS} < \lambda^{NS}(\theta'')$. Combining these results with (A9) yields (A10).

$$(A10) \quad \lambda^S - \lambda^{NS} + (1/2)[r^S(B-1+2\theta_S^{P'}+\Delta) + (1-r^S)(B+(\theta''-\alpha)+\Delta)] \\ - (1/2)[r^{NS}(B-1+2\theta_{NS}^{P'}) + (1-r^{NS})(B-\alpha)] > z(\Delta/2)$$

This completes the proof of v).

Proof of Corollary 1 to Proposition 5: Education signaling returns are increases in expected compensation due to schooling changing beliefs about the worker's value for θ_i or θ_i^P . When a worker chooses $s_i=1$ then firms believe the minimum possible value for $\theta_i=\theta''$ rather than zero and the result, as captured in ii) of the proposition, is that the sub-period 1 wage increases by $(\theta^H/4) + (\Delta/2) + \lambda^S - \lambda^{NS}$ and only part of this increase is due to the accumulation of human capital. This proves i). As captured in iii) and iv) of the proposition, the sub-period 2 wage for non-promoted workers increases with schooling by an amount $(1/2)\theta'' + (\Delta/2)$ where again only part of this increase is due to the accumulation of human capital. This proves ii). The change in beliefs concerning the minimum θ_i^P associated with schooling means that the non-promotion wage in sub-period 2 is higher by more than $\Delta/2$ for workers with schooling as just indicated. In turn, as shown in the proof of Proposition 5, this results in $\theta_S^{P'} < \theta_{NS}^{P'}$ which means that holding θ_i fixed the probability of promotion increases when worker i chooses schooling and this is an education signaling return since the first best probability of promotion is independent of the schooling decision. This proves iii).

REFERENCES

Acemoglu, D. and J.S. Pischke (1998), "Why do Firms Train? Theory and Evidence," *Quarterly Journal of Economics*, 113, pp. 79-119.

Altonji, J. and C. Pierret (1997), "Employer Learning and the Signaling Value of Education," in I. Ohashi and T. Tachibanaki (eds.), *Industrial Relations, Incentives and Employment* (Macmillan Press Ltd., London), pp. 159-195.

Altonji, J. and C. Pierret (2001), "Employer Learning and Statistical Discrimination," *Quarterly Journal of Economics*, 116, pp. 313-350.

Barron, J., M. Berger, and D. Black (2006), "Selective Counteroffers," *Journal of Labor Economics*, 24, pp. 385-409.

Bedard, K. (2001), "Human Capital versus Signaling Models: University Access and High School Dropouts," *Journal of Political Economy*, 109, pp. 749-775.

Belzil, C. and M.L. Bognanno (2010), "The Promotion Dynamics of American Executives," *Research in Labor Economics*, 30, pp. 189-231.

Bernhardt, D. (1995), "Strategic Promotion and Compensation," *Review of Economic Studies*, 62, pp. 315-339.

Bognanno, M. and E. Melero (Forthcoming), "Promotion Signals, Experience and Education," *Journal of Economics and Management Strategy*.

Cassidy, H., J. DeVaro, and A. Kauhanen (2014), "The Signaling Role of Promotions: New Evidence from Europe," Mimeo, Department of Economics, California State University, East Bay.

Cho, I. and D. Kreps (1987), "Signaling Games and Stable Equilibria," *Quarterly Journal of Economics*, 102, pp. 179-221.

DeVaro, J. and M. Waldman (2012), "The Signaling Role of Promotions: Further Theory and Empirical Evidence," *Journal of Labor Economics*, 30, pp. 91-147.

Farber, H. and R. Gibbons (1996), "Learning and Wage Dynamics," *Quarterly Journal of Economics*, 111, pp. 1007-1047.

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

Gibbs, M. (2003), "Evidence on Wage and Promotion Dynamics in an Internal Labor Market," Mimeo, University of Chicago.

Golan, L. (2005), "Counteroffers and Efficiency in Labor Markets with Asymmetric Information," *Journal of Labor Economics*, 23, pp. 373-393.

- Greenwald, B. (1986), "Adverse Selection in the Labour Market," *Review of Economic Studies*, 53, pp. 325-347.
- Hungerford, T. and G. Solon (1987), "Sheepskin Effects in the Returns to Education," *Review of Economics and Statistics*, 69, pp. 175-177.
- Ishida, J. (2004), "Signaling and Strategically Delayed Promotion," *Labour Economics*, 11, pp. 687-700.
- Kahn, L. (2013), "Asymmetric Information Between Employers," *American Economic Journal: Applied Economics*, 5, pp. 165-205.
- Kaymak, B. (2014), "Quantifying the Signaling Role of Education," Mimeo, Department of Economics, University of Montreal.
- Krashinsky, H. (2002), "Evidence on Adverse Selection and Establishment Size in the Labor Market," *Industrial & Labor Relations Review*, 56, pp. 84-96.
- Lang, K. and D. Kropp (1986), "Human Capital versus Sorting: The Effects of Compulsory Attendance Laws," *Quarterly Journal of Economics*, 101, pp. 609-624.
- Lange, F. (2007), "The Speed of Employer Learning," *Journal of Labor Economics*, 25, pp. 1-35.
- Layard, R. and G. Psacharopoulos (1974), "The Screening Hypothesis and the Returns to Education," *Journal of Political Economy*, 82, pp. 985-999.
- MacLeod, W.B. and J.M. Malcomson (1988), "Reputation and Hierarchy in Dynamic Models of Employment," *Journal of Political Economy*, 96, pp. 832-854.
- Milgrom, P. and S. Oster (1987), "Job Discrimination, Market Forces, and the Invisibility Hypothesis," *Quarterly Journal of Economics*, 102, pp. 453-474.
- Myerson, R.B. (1978), "Refinements of the Nash Equilibrium Concept," *International Journal of Game Theory*, 7, pp. 73-80.
- Perri, T. (1994), "Testing for Ability When Job Assignment is a Signal," *Labour Economics*, 1, pp. 365-381.
- Pinkston, J. (2009), "A Model of Asymmetric Learning with Testable Implications," *Review of Economic Studies*, 76, pp. 367-394.
- Ricart i Costa, J. (1988), "Managerial Task Assignment and Promotion," *Econometrica*, 56, pp. 449-466.
- Riley, J. (1975), "Competitive Signalling," *Journal of Economic Theory*, 10, pp. 174-186.

- Riley, J. (1979a), "Informational Equilibrium," *Econometrica*, 47, pp. 331-359.
- Riley, J. (1979b), "Testing the Educational Screening Hypothesis," *Journal of Political Economy*, 87, pp. S227-S252.
- Riley, J. (2001), "Silver Signals, Twenty-Five Years of Screening and Signaling," *Journal of Economic Literature*, 39, pp. 432-478.
- Schoenberg, U. (2007), "Testing for Asymmetric Employer Learning," *Journal of Labor Economics*, 25, pp. 651-692.
- Song, Y. (2007), "Recall Bias in the Displaced Workers Survey: Are Layoffs Really Lemons?," *Labour Economics*, 14, pp. 335-345.
- Spence, M. (1973), "Job Market Signaling," *Quarterly Journal of Economics*, 87, pp. 355-374.
- Spence, M. (2002), "Signaling in Retrospect and the Informational Structure of Markets," *American Economic Review*, 92, pp. 434-459.
- Waldman, M. (1984), "Job Assignments, Signalling, and Efficiency," *Rand Journal of Economics*, 15, pp. 255-267.
- Waldman, M. (1990), "Up-or-Out Contracts: A Signaling Perspective," *Journal of Labor Economics*, 8, pp. 230-250.
- Waldman, M. (2012), "Theory and Evidence in Internal Labor Markets," in R. Gibbons and J. Roberts (eds.), *The Handbook of Organizational Economics* (Princeton University Press, Princeton, NJ), pp. 520-574.
- Waldman, M. and O. Zax (2014), "An Exploration of the Promotion Signaling Distortion," Mimeo, Johnson Graduate School of Management, Cornell University.
- Winstein, K.J. (2008), "Inflated Credentials Surface in Executive Suite," *Wall Street Journal*, November 13, 2008.
- Wolpin, K. (1977), "Education and Screening," *American Economic Review*, 67, pp. 949-958.
- Zabojnik, J. and D. Bernhardt (2001), "Corporate Tournaments, Human Capital Acquisition, and the Firm Size-Wage Relation," *Review of Economic Studies*, 68, pp. 693-716.