DISCUSSION PAPER SERIES

No. 6623

LEARNING AND SMOOTH STOPPING

Robin Mason and Juuso Välimäki

INDUSTRIAL ORGANIZATION and LABOUR ECONOMICS

Centre for Economic Policy Research

www.cepr.org

Available online at: **www.cepr.org/pubs/dps/DP6623.asp**

ISSN 0265-8003

LEARNING AND SMOOTH STOPPING

Robin Mason, University of Southampton and CEPR
Juuso Välimäki, Helsinki School of Economics, University of Southampton and HECER

Discussion Paper No. 6623
January 2008

Centre for Economic Policy Research
90–98 Goswell Rd, London EC1V 7RR, UK
Tel: (44 20) 7878 2900, Fax: (44 20) 7878 2999
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 6623

January 2008

ABSTRACT

Learning and Smooth Stopping

We propose a simple model of optimal stopping where the economic environment changes as a result of learning. A primary application of our framework is an optimal job search problem when the worker's labour market opportunities are initially uncertain. We distinguish between two interpretations of the model. In the first, a worker learns about common market conditions, such as the number of potential employers, that affect all searchers. In the second, the worker learns about her idiosyncratic productivity distribution across firms. For the first model, we show that learning leads to higher wage demands by the workers. In the second model, we give sufficient conditions so that learning leads to higher wage demands for optimistic workers and lower demands for pessimistic workers due to learning.

JEL Classification: D83
Keywords: learning and stopping

Robin Mason
Department of Economics
University of Southampton
Highfield
Southampton
S017 1BJ
Email: robin.mason@soton.ac.uk

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=145252

Juuso Välimäki
Helsinki School of Economics
Arkadiankatu7
FI-00100 Helsinki
FINLAND
Email: juuso.valimaki@hse.fi

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=148707

Submitted 07 December 2007

1 Introduction

Models of optimal stopping capture the trade-off between known current payoffs and the opportunity cost of future potentially superior possibilities. In most of their economic applications, the environment in which the stopping decisions are taken is assumed to be stationary. In the standard job search model, for example, a rejected offer is followed by another draw from the same distribution of offers. See Rogerson, Shimer, and Wright (2005) for a survey. Similarly, in most models of research and development (R&D), current effort is assumed to have the same probability of yielding a success regardless of the history. See for example Grossman and Shapiro (1986); although also see Malueg and Tsutsui (1997) for an exception to this, in a model of R&D competition.

The predictions of these models are at odds with empirical observations. Reservation wages of unemployed workers decline in the duration of the unemployment spell; see e.g., Lancaster and Chesher (1983). Prices of houses decline as a function of time on the market; see e.g., Merlo and Ortalo-Magne (2004, p. 214). Research projects are sometimes started and later abandoned.

We propose a simple model of optimal stopping where the economic environment changes as a result of learning. From a methodological point of view, our main contribution is the development of a framework that captures essential features of the learning problem and yet is analytically tractable. Rather than using the deterministic model of optimal stopping typically used in optimal job search models, we generalise the other canonical model of stopping in economic theory, where stopping occurs probabilistically (as in e.g., models of R&D). We cast the model in the language of job search, but in the concluding section, we point out alternative interpretations for the model.

In the main model, a worker that is initially uncertain about her job prospects meets potential employers randomly. She posts a reservation wage; potential employers arrive according to a Poisson process and observe the posted offer. They offer employment if and only if the worker's productivity exceeds her wage demand. The worker is initially uncertain about the probability of meeting a suitable employer, and the only information that the worker receives is the hiring decision. The worker updates her beliefs about the distribution of her productivity in a Bayesian fashion, becoming more pessimistic about her future opportunities in each period when she is not hired. The speed of decline of her beliefs about future job opportunities depends on the current wage demand. At very high demands, no employer is willing to hire the worker and hence little information is conveyed by the fact that no job offers were received. At very low wages, all employers are happy to employ the worker and hence the absence of a job offer is merely an indication that no potential employers were present. At intermediate wages, some employers are willing to employ the worker while others are not. If the distribution of a high productivity worker dominates that of a low productivity worker, in an appropriate sense,[1] then the absence of a job offer is an indication that the worker is more likely to be of low productivity. In summary, when setting her current wage offer, the worker controls the immediate expected wages conditional on a firm offering employment, as well as the beliefs about her own productivity if no job is offered.

Even though our model is a standard Bayesian learning model, the belief dynamics may appear unfamiliar at first sight. Conditional on an accepted wage offer, the game is over and continuation beliefs play no further role. When we describe the evolution of beliefs over time, we are implicitly conditioning on the event that no job was offered in the current period. Since this event is bad news about her future employment prospects, the worker become more pessimistic over time. The worker can then control the (downward) drift of her beliefs through her choice of reservation wage. In addition, the process of beliefs is effectively deterministic, again because continuation beliefs play a role in only one of two possible random outcomes. This is in contrast to many learning models where the continuation beliefs are stochastic and form a martingale, and where the agent can control only the variance of her beliefs.

When comparing the model with learning to the model with stationary wage offers, two features change simultaneously. A more pessimistic future implies a lower current value of being unemployed and hence a lower current reservation wage. We call this feature the *controlled stopping effect*. In a model where perceived job opportunities

[1] We show that hazard rate dominance is the relevant relationship for our model.

decline due to reasons unrelated to current wage demands (such as depreciation of skills), the controlled stopping effect would be the only effect.

In our model, the current reservation wage also determines the magnitude of the change in beliefs if the worker is not hired. By setting a more informative reservation wage, the worker can cause her beliefs to fall further in the event of remaining unemployed. A more pessimistic worker has a lower value from job search than a more optimistic worker—there is a capital loss from learning. Holding fixed the probability of a receiving a job offer, then, the worker has an incentive to move her reservation wage in the direction that minimises this capital loss i.e., the change in beliefs if she is not hired. We call this feature the *controlled learning effect*.

Which effect dominates? Does learning cause the worker to raise or lower her reservation wage, relative to the case when no learning occurs? How does this comparison depend on the level of beliefs (i.e., degree of optimism or pessimism) of the worker? Finally, does a more pessimistic worker always set a lower reservation wage; or do the two effects lead to more complicated dynamics for the wage?

We distinguish between two different scenarios. In the first, there is no uncertainty about the distribution of the workers' productivities across the firms but the market conditions (common to all workers) are unknown at the outset. We model this by introducing uncertainty about the Poisson rate of contact between firms and workers; we call this scenario the *common uncertainty model*. A high Poisson rate corresponds naturally to a job market where the workers are in relatively high demand. In the *idiosyncratic uncertainty model*, the rate of contact between workers and firm is known but there is heterogeneity in the workers' productivities. Any given worker may be of high or low productivity. We assume that the productivity distribution (across firms) of a high productivity worker dominates that of the low type worker, in terms of hazard rate dominance.

A key question for the analysis is whether these two effects work in opposite or the same directions. The benchmark is a model where the rate of contact is fixed at its current expected level. Since the value of continuing search decreases as the worker becomes more pessimistic (in the model with learning), the controlled stopping effect on its own leads to lower optimal wage demands. In the common uncertainty model, a higher wage demand always slows down learning, since the difference in the probability of finding a job in the two states is monotonic (and in fact positively multiplicative) in the overall probability of finding a job. We show that the controlled learning effect dominates in this case. Hence the optimal wage demands in the common uncertainty model exceed the wage demands in the equivalent model with no learning.

In the idiosyncratic uncertainty model, the controlled learning effect is not monotonic in wage demands. If all employers are willing to hire at very low wage demands and no employers hire at very high demands, it is clear that intermediate wage demands are the most informative. As a result, the controlled learning effect pulls up wage demands of optimistic workers (those with relatively high reservation wages) and pushes down the demands of pessimistic ones. Since the controlled stopping effect always pushes wage demands downwards, it is clear that the model with learning yields relatively low wage demands for pessimistic workers, in comparison to the model with no learning. We also derive sufficient conditions for the model to yield higher wage demands for optimistic workers in the learning case. As a result, learning accelerates the decline in wage demands as beliefs become more pessimistic.

This work is related to the literature on search with learning. Burdett and Vishwanath (1988) analyse an optimal stopping problem when wage offers come from exogenous sources. Crucially, in Burdett and Vishwanath (1988), learning is independent of the current actions and hence the stopping problem is quite different from the one that we propose to study. Furthermore, the analysis in Burdett and Vishwanath (1988) is complicated by the fact that new wage offers lead to discrete jumps in the posterior belief whereas in our model, the evolution of beliefs is smooth (see e.g., equation (1)). There is a small number of papers on search in models where the state of the system changes exogenously. Van den Berg (1990) and Smith (1999) are examples of such models. Anderson and Smith (2006) analyse a matching model where each match results in a current immediate payoff and reveals new information about the productivity of the partners. In all of these papers, beliefs change i.e., the environment is non-stationary; but the change in beliefs is not affected by the action chosen by the economic agent. Our contribution is to develop a tractable framework that allows agents to affect their learning through their actions.

The paper is organised as follows. Section 2 presents the model of learning and search. Section 3 analyses the model in the common uncertainty case; section 4 deals with the case of idiosyncratic learning. Section 5 discusses alternative interpretations of the model and concludes.

2 Labour Search and Learning

Consider an unemployed worker setting his reservation wage for entering employment. The worker announces at the start of each period his reservation wage for that period. If he receives a job offer with a wage that exceeds the reservation wage, then he accepts that offer and stops searching. The probability that the worker receives such a job offer depends on two factors: whether a vacancy arises during the period; and whether the employer with a vacancy is willing to meet the worker's reservation wage. The worker does not observe whether a vacancy occurs: he observes only when a job offer is made. Assume that employers know perfectly the productivity of the worker, which is specific to each particular employer. Hence an employer will make an offer of exactly the reservation wage if the worker's productivity in that job exceeds the reservation wage; otherwise, no offer is made. A vacancy lasts for one period, and then disappears. (This model also matches a monopoly seller of a single unit facing an infinite horizon of buyers who each live for only one period.) We assume that search is not directly costly. But by continuing search, the worker delays receiving a wage; discounting then makes search costly.

Time is continuous, the worker is infinitely lived and discounts future at rate r. There are two states of the world: 'high' (H) or 'low' (L). The worker is initially uncertain about the state and her s problem is to post wage demands to optimise her expected lifetime wage earnings. The problem can thus be cast as a continuous time discounted dynamic program where the belief about the state of the world serves as the state variable.

The worker receives an offer during a time interval of length Δt in state H when posting a reservation wage $w \in \mathbb{R}_+$ with probability $H(w)\Delta t$; the probability in state L is $L(w)\Delta t$. Both probabilities are decreasing in w (a higher reservation wage makes it less likely that an offer is received). Let the corresponding densities be $l(w)$ and $h(w)$. Assume that the distributions $L(\cdot)$ and $H(\cdot)$ can be ranked by hazard rate dominance:

Assumption 1 (Hazard rate dominance)

$$\frac{l(w)}{1-L(w)} \geq \frac{h(w)}{1-H(w)} \quad \forall\, w \in \mathbb{R}_+.$$

Note that assumption 1 implies first-order stochastic dominance (FOSD) i.e., $\Delta F(w) \equiv L(w) - H(w) \geq 0$.

The state variable in this problem is then the worker's scalar belief $\pi \in [0,1]$ that the state of the world is high. When the worker receives a job offer (at the reservation wage), the problem stops. Hence the only event that is relevant for updating the worker's beliefs is when no job offer has occurred in a period. Denote the posterior probability that the distribution is H after a history that has included no job offers by π. Along the optimal path, offers must occur with strictly positive probability and hence π must be strictly decreasing, by FOSD. The updated posterior after no job offer is, by Bayes' rule,

$$\pi + \Delta\pi = \frac{\pi(1-H(w))}{\pi(1-H(w)) + (1-\pi)(1-L(w))}$$

so that

$$\Delta\pi = -\pi(1-\pi)\Delta G(w) \leq 0$$

where $\Delta G(w) \equiv H(w) - L(w) \geq 0$. Notice that the worker becomes more pessimistic about the state of the world when no job offer is received.

We now consider two different cases of the model. First, we suppose that the workers are learning about the arrival rate of offers. Since it is natural to interpret this type of uncertainty as arising from market conditions (such as the number of active firms) affecting all workers in a similar manner, we call this common uncertainty. In the second model, there is uncertainty about the distribution of a worker's productivity across firms. It is natural to think of this case as learning about idiosyncratic productivity. Even though we could nest the mathematical structure of the first model in the second, we find it useful to analyse them separately. In addition to the differences in interpretation, the analysis of the first model is much more transparent and hence serves as a good guide to the more complicated arguments of the second model.

3 Common uncertainty

In this case, the states of the world are distinguished by the frequency of arrivals of vacancies. In the low state L, the Poisson arrival rate is $\lambda_L > 0$. In the high state H, the Poisson arrival rate is $\lambda_H > \lambda_L$. Employers' wage offers are drawn from a distribution $F(\cdot)$ with a density function $f(w)$, so that, conditional on a vacancy arising, the probability of the worker's reservation wage w being met is $1 - F(w)$. Hence $H(w) = \lambda_H(1 - F(w))$ and $L(w) = \lambda_L(1 - F(w))$. Assume that $1 - F(w) - wf(w)$ is strictly decreasing in w, so that the objective function in the static problem is concave and therefore the problem has a unique solution.

Bayes' rule gives the worker's posterior, after a period when no job offer was received, in the continuous-time limit ($\Delta t \to 0$) as

$$\dot{\pi}(w, \pi) = -\pi(1 - \pi)\Delta\lambda(1 - F(w)) < 0 \tag{1}$$

where $\Delta\lambda \equiv \lambda_H - \lambda_L > 0$. Notice that a *higher* reservation wage is always *less* informative in this case. By setting a high reservation wage, the worker decreases the rate at which his posterior goes down, since $\dot{\pi}_w(w, \pi) = \pi(1 - \pi)\Delta\lambda f(w) > 0$.

The worker's static problem,

$$\max_w \left\{ \lambda(\pi)(1 - F(w))w \right\},$$

has a unique solution w_S given by the first-order condition

$$\Gamma_S(w_S, \pi) \equiv \delta(\pi)(1 - F(w_S) - w_S f(w_S)) = 0$$

where $\delta(\pi) \equiv r/\lambda(\pi)$. (The $\delta(\pi)$ factor appears to aid comparisons later.)

Consider next the benchmark intertemporal model where the worker faces the same arrival rate $\lambda(\pi)$ in all periods. Since the problem remains the same in all periods, the dynamic programming problem of the worker is particularly easy to solve. This solution will serve as a natural point of comparison to the model that incorporates learning about λ i.e., with π changing over time. We refer to this case as the *repeated* problem.

For the stationary problem, the worker's Bellman equation is simply:

$$V_R(\pi) = \max_w \left\{ \lambda(\pi)\Delta t(1 - F(w))w + \frac{1 - \lambda(\pi)\Delta t(1 - F(w))}{1 + r\Delta t} V_R(\pi) \right\}.$$

In the continuous-time limit as $\Delta t \to 0$, this becomes

$$\delta(\pi)V_R(\pi) = \max_w \left\{ (1 - F(w))(w - V_R(\pi)) \right\}.$$

The first-order condition for the repeated problem is

$$\Gamma_R(w_R, \pi) \equiv \Gamma_S(w_R(\pi), \pi) + (1 - F(w_R(\pi)))^2 = 0. \tag{2}$$

Since $(1 - F(w))$ is decreasing in w, there is a unique solution to equation (2). And since $(1 - F(w)) \geq 0$, $w_R(\pi) \geq w_S$, as expected. The value function of the repeated problem is then

$$V_R(\pi) = \frac{(1 - F(w_R(\pi)))w_R(\pi)}{1 - F(w_R(\pi)) + \delta(\pi)}.$$

Consider next the model with learning. The crucial difference compared to the repeated problem is that now π evolves over time according to equation (1). The worker anticipates this and adjusts her optimal wage demand to control the current probability of being hired and the future value of search.

The worker's Bellman equation is now:

$$V_D(\pi) = \max_w \left\{ \lambda(\pi)\Delta t(1 - F(w))w + \frac{1 - \lambda(\pi)\Delta t(1 - F(w))}{1 + r\Delta t} V_D(\pi + \Delta\pi) \right\}.$$

In the continuous-time limit as $\Delta t \to 0$, this becomes

$$rV_D(\pi) = \max_w \left\{ \lambda(\pi)(1 - F(w))(w - V_D(\pi)) - \pi(1 - \pi)\Delta\lambda(1 - F(w))V_D'(\pi) \right\}.$$

It is easy to see that in this problem, $V'(\pi) > 0$ almost everywhere. Since the value in any program is realized only at the stopping moment, the value of any fixed policy must be increasing in the arrival rate and hence π. Hence the optimal policy starting at π yields a higher value when started at state $\pi' > \pi$ and as a consequence $V(\pi') > V(\pi)$. We shall show later (see lemma 2 in the appendix) that $V(\pi)$ is convex and as a result, $V(\pi)$ is differentiable almost everywhere.

After standard manipulations, the Bellman equation can be rewritten as

$$V_D(\pi) = \max_w \left\{ V_R(w, \pi) - \frac{\Delta\lambda(1 - F(w))}{\lambda(\pi)(1 - F(w, \pi)) + r} \pi(1 - \pi)V_D'(\pi) \right\}. \tag{3}$$

This latter equation makes clear the connection between the repeated and dynamic problem. The dynamic value function is equal to a repeated value function, $V_R(w_D, \pi)$ (evaluated at the optimal dynamic action, w_D), minus a term that arises due to learning. Note that $-\Delta\lambda(1 - F(w))\pi(1 - \pi)V_D'(\pi)$ is the capital change (a loss) of the agent's program resulting from the evolution of the agent's posterior. The learning term is the present value of the infinite stream of the capital losses due to learning, using the effective discount rate $\lambda(\pi)(1 - F(w)) + r$.

It is useful to analyse the behaviour of the learning term more carefully. The factor $\pi(1-\pi)V'(\pi)$ in the capital loss term represents the impact of a unit of learning on future payoffs. It is unaffected by the current choice of w and hence we ignore its impact for the moment. We thus concentrate on the behaviour of

$$-\frac{\Delta\lambda(1 - F(w))}{\lambda(\pi)(1 - F(w, \pi)) + r}$$

as a function of w. Its derivative with respect to w has two terms. The first,

$$-\frac{\Delta\lambda(1 - F(w))\lambda(\pi)f(w)}{(\lambda(\pi)(1 - F(w, \pi)) + r)^2},$$

measures the effect of w on the modified stopping rate. We call this effect the *controlled stopping effect.* This effect captures the increasing pessimism of the worker and has always a negative sign. This effect alone leads to lower wage demands for the worker.

The second term

$$\frac{\Delta\lambda f(w)}{\lambda(\pi)(1 - F(w, \pi)) + r},$$

measures the change in learning, normalised by the stopping rate resulting from a change in w. We call this effect the *controlled learning effect.* In the current model, a higher w leads to a smaller change in π, through Bayes' rule; as a result, this effect always directs the worker towards higher wage demands to minimise the bad news conditional on not being employed.

Because learning takes a particularly simple multiplicative form in this model, it is possible to sign the overall effect of learning unambiguously. After cancelling terms in the controlled stopping and learning effects, we can write the first-order condition for the problem as:

$$\Gamma_R(w, \pi) + r\Delta\lambda f(w)\pi(1 - \pi)V_D'(\pi) = 0. \qquad (4)$$

Again, we observe that since $\Gamma_R(w, \pi)$ is decreasing in w and $r\Delta\lambda f(w)\pi(1-\pi)V_D'(\pi) > 0$, all solutions $w_D(\pi)$ to the first-order condition (4) satisfy $w_D(\pi) > w_R(\pi)$ for all π. (Since $f'(w)$ may be positive or negative, it is no longer possible to guarantee the uniqueness of $w_D(\pi)$.)

Hence the reservation wage set by the worker who learns about the state of the world is *above* the reservation wage of the non-learning worker. As a result, the probability that the learning worker receives a job offer is below that of the non-learning; and this despite the fact that the learning worker becomes more pessimistic over time about his employment prospects. Pessimism on its own would cause the worker to set a *lower* reservation wage, since the worker should be more willing to exit search and accept a job offer. But the worker is able to slow down the rate at which he becomes more pessimistic by setting a high reservation wage. This more than off-sets the fall in the worker's posterior, ensuring a higher reservation wage.

We summarise this discussion in the following proposition.

Proposition 1 *In the case of learning about the arrival rate of job offer arrivals, $w_S \leq w_R(\pi) \leq w_D(\pi)$ for all $\pi \in [0,1]$.*

In this case, it is also relatively straightforward to establish monotonicity of the reservation wages in the posterior belief.

Proposition 2 *$w_R(\pi)$ and $w_D(\pi)$ are non-decreasing in π.*

Proof. The first-order condition for $w_R(\pi)$ can be written as

$$1 - F(w_R(\pi)) - f(w_R(\pi))(w_R(\pi) - V_R(\pi)) = 0.$$

The monotonicity of $w_R(\pi)$ follows from the second-order condition, and the fact that the derivative of this first-order condition with respect to π is $f(\cdot)V_R'(\pi)$, which is positive.

The first-order condition for $w_D(\pi)$ is

$$\lambda(\pi)(1 - F(w_D(\pi)) - f(w_D(\pi))(w_D(\pi) - V_D(\pi))) + f(w_D(\pi))\Delta\lambda\pi\langle 1-\pi)V_D'(\pi) = 0. \tag{5}$$

Using the second-order condition, $w_D(\pi)$ is non-decreasing in π iff

$$\Delta\lambda(1 - F(w_D(\pi)) - f(w_D(\pi))(w_D(\pi) - V_D(\pi))) \\ + \lambda f(w_D(\pi))V_D'(\pi) + f(w_D(\pi))\Delta\lambda\frac{d}{d\pi}(\pi\langle 1-\pi)V_D'(\pi))$$

is non-negative. But this can be established, using the Envelope theorem on equation (3), the first-order condition (5), and a few straightforward manipulations of the resulting expressions. □

3.1 Two-point example

We can illustrate these results in a simple case where the worker's productivity in a job can take two values, $\overline{v} > \underline{v} > 0$, with probabilities $\beta \in (0,1)$ and $1-\beta$ respectively. In this case, we are able to provide explicit solutions for the worker's value function and reservation wage.

For the problem to be interesting, assume that $\beta\overline{v} < \underline{v}$, so that in the static problem, the worker would set a wage equal to $\underline{v}$. Also assume that if the worker knows that $\lambda = \lambda_H$, its optimal reservation wage is $\overline{v}$; but if the worker knows that $\lambda = \lambda_L$, its optimal wage is $\underline{v}$. That is, assume

$$\frac{\beta\lambda_H\overline{v}}{r+\beta\lambda_H} \geq \frac{\lambda_H\underline{v}}{r+\lambda_H} > \frac{\lambda_L\underline{v}}{r+\lambda_L} \geq \frac{\beta\lambda_L\overline{v}}{r+\beta\lambda_L}.$$

By standard arguments, the (dynamic) worker's reservation wage strategy then takes a simple form: set the wage at $\overline{v}$ when the posterior is above some cut-off; and $\underline{v}$ for posteriors below this level. Let the cut-off level of the posterior $\pi^* \in (0,1)$. For $\pi \geq \pi^*$, the worker's value function is the solution to the ordinary differential equation (ODE)

$$\pi(1-\pi)\Delta\lambda V'(\pi) + \left(\frac{r}{\beta} + \lambda(\pi)\right)V(\pi) - \overline{v}\lambda(\pi) = 0.$$

The general solution to this ODE is of the form

$$V(\pi) = k_1\left(\frac{1-\pi}{\pi}\right)^{\frac{r+\beta\lambda_H}{\beta\Delta\lambda}}\pi + \frac{\beta\lambda_L\overline{v}}{r+\beta\lambda_L} + \frac{r\beta\Delta\lambda\overline{v}}{(r+\beta\lambda_L)(r+\beta\lambda_H)}\pi$$

for $\pi \geq \pi^* > 0$. The first term represents the option value from learning to the worker. (k_1 is a constant of integration that will be determined below.) It is a convex function of π.

For $\pi < \pi^*$, the worker sets its reservation wage at $\underline{v}$; its value function is then given by the solution to a second ODE:

$$\pi(1-\pi)\Delta\lambda V'(\pi) + (r + \lambda(\pi))V(\pi) - \underline{v}\lambda(\pi) = 0.$$

The general solution to this ODE is of the form

$$V(\pi) = k_2 \left(\frac{1-\pi}{\pi}\right)^{\frac{r+\lambda_H}{\Delta\lambda}} \pi + \frac{\lambda_L \underline{v}}{r+\lambda_L} + \frac{r\Delta\lambda\underline{v}}{(r+\lambda_L)(r+\lambda_H)}\pi.$$

The first term represents the option value from learning to the worker. It is unbounded as π approaches zero; hence the constant of integration k_2 must equal zero. In words, the worker has no option value when $\pi \leq \pi^*$. The reason is that there is no prospect of the worker's posterior increasing above π^*; hence the worker's reservation wage is constant (equal to $\underline{v}$) once the posterior falls below π^*. There is then no option for the worker to value. So, for $\pi < \pi^*$,

$$V(\pi) = \frac{\lambda_L \underline{v}}{r+\lambda_L} + \frac{r\Delta\lambda\underline{v}}{(r+\lambda_L)(r+\lambda_H)}\pi.$$

There are two remaining parameters to determine: k_1 and the optimal boundary π^*. Since π^* is chosen optimally, value matching and smooth pasting apply; these two boundary conditions are sufficient to determine k_1 and π^*.

The worker's value function in this case is illustrated in figure 1. It is a strictly convex function of π in the region $\pi \geq \pi^*$; for $\pi < \pi^*$, it is a linear function. The two portions value match and smooth paste at π^*, given by

$$\pi^* = \frac{\big(-(\underline{v}-\beta\bar{b})r^2 + (\beta(\overline{v}-\underline{v})\lambda_L - (\underline{v}-\beta\overline{v})\lambda_H)r + \beta(\overline{v}-\underline{v})\lambda_L\lambda_H\big)\lambda_L}{\big(-(\underline{v}-\beta\overline{v})r^2 + \beta(\overline{v}-\underline{v})(\lambda_L+\lambda_H)r + \beta(\overline{v}-\underline{v})\lambda_L\lambda_H\big)\Delta\lambda}.$$

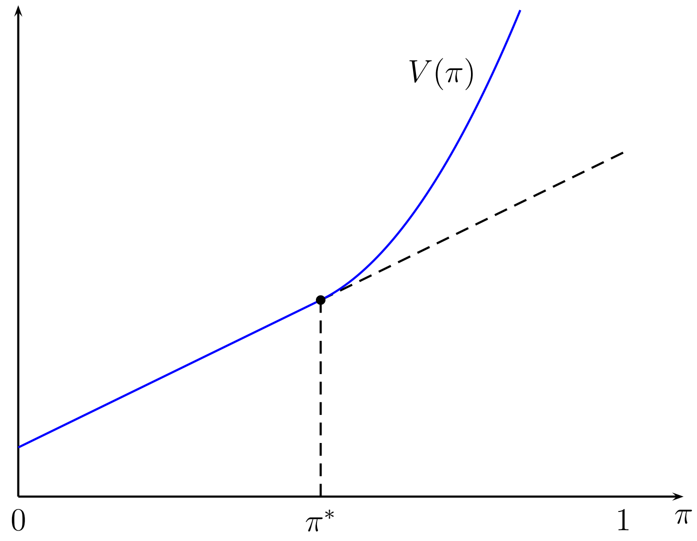

Figure 1: The value function with a 2-point valuation distribution

Let λ_0 be such that

$$\frac{\beta\lambda_0\overline{v}}{r+\beta\lambda_0} = \frac{\lambda_0\underline{v}}{r+\lambda_0}, \qquad \text{i.e.,}\quad \lambda_0 = \frac{r(\underline{v}-\beta\overline{v})}{\beta(\overline{v}-\underline{v})}.$$

For a worker who is certain of the arrival rate, the optimal reservation wage is $\overline{v}$ for all $\lambda \geq \lambda_0$, and $\underline{v}$ for $\lambda < \lambda_0$. Then define π_0 to be such that $\pi_0\lambda_H + (1-\pi_0)\lambda_L = \lambda_0$. If $\pi^* \leq \pi_0$, then uncertainty causes the worker to set a higher reservation wage in this case with a two-point valuation distribution. Lengthy but straightforward calculation shows that this inequality holds if

$$\frac{\underline{v}}{r+\lambda_H} < \frac{\beta\overline{v}}{r+\beta\lambda_H}$$

which was assumed at the outset. In summary: learning about an unknown arrival rate causes the worker to set a higher reservation wage.

4 Idiosyncratic uncertainty

In this second case, the states of the world are distinguished by the distributions from which productivity levels of the workers are drawn (conditional on a vacancy arising). In the high state of the world, productivities are drawn from the distribution $F_H(\cdot)$; in the low state, they are drawn from $F_L(\cdot)$. We assume now that the corresponding densities $f_H(w)$ and $f_L(w)$ are continuously differentiable. Let the union of the supports of the distributions be $[0, 1]$ (without loss of generality). Let $\Delta F(w) \equiv F_L(w) - F_H(w) > 0$ denote the difference between the two distributions; and let $\Delta f(w) \equiv f_L(w) - f_H(w)$ denote the difference between the densities. In both states of the world, vacancies occur randomly according to a Poisson process with parameter λ (the same in both states).

We let π denote the worker's subjective posterior probability that the state is H. Furthermore, we let the current expected distribution of productivities across firms be denoted by

$$F(w, \pi) \equiv \pi F_H(w) + (1 - \pi) F_L(w),$$

and define $f(w, \pi)$ similarly.

The worker's updated posterior after no job offer is calculated similarly to the previous section:

$$\dot{\pi} = -\lambda \pi (1 - \pi) \Delta F(w) < 0.$$

Notice that the worker again becomes more pessimistic about the state of the world when no job offer is received.

We assume again that $1 - F(w, \pi) - wf(w, \pi)$ is strictly decreasing in w for any π. This ensures that the worker's static problem,

$$\max_w \left\{ \lambda (1 - F(w, \pi)) w \right\},$$

has a unique solution w_S given by the first-order condition

$$\Gamma_S(w_S, \pi) \equiv \delta (1 - F(w_S, \pi) - w_S f(w_S, \pi)) = 0.$$

By direct analogy to the previous section, the worker's Bellman equation (in the continuous-time limit) for the repeated problem (i.e. the problem without learning about π) is

$$\delta V_R(\pi) = \max_w \left\{ (1 - F(w, \pi))(w - V_R(\pi)) \right\}$$

where $\delta \equiv r/\lambda$. The first-order condition for the repeated problem is

$$\Gamma_R(w_R, \pi) \equiv \Gamma_S(w_R, \pi) + (1 - F(w_R(\pi), \pi))^2 = 0,$$

and the value function of the repeated problem is then

$$V_R(\pi) = \frac{(1 - F(w_R(\pi), \pi))w_R(\pi)}{1 - F(w_R(\pi), \pi) + \delta}.$$

It will be helpful for subsequent arguments to establish the monotonicity of $w_R(\pi)$ at this stage of the analysis.

Proposition 3 *$w_R(\pi)$ is non-decreasing in π for all $\pi \in [0, 1]$.*

Proof. $w_R(\pi)$ is non-decreasing in π iff

$$\Delta F(w_R(\pi)) + \frac{\Delta f(w_R(\pi))(1 - F(w_R(\pi), \pi))}{f(w_R(\pi), \pi)} + f(w_R(\pi), \pi)V_R'(\pi) \geq 0. \tag{6}$$

We know that $V_R'(\pi) \geq 0$; and $f(w_R(\pi), \pi) \geq 0$. Hence the last term in this expression is non-negative. Hence condition (6) requires that

$$\Delta F(w_R(\pi)) + \frac{\Delta f(w_R(\pi))(1 - F(w_R(\pi), \pi))}{f(w_R(\pi), \pi)} \geq 0.$$

Clearly it is sufficient that

$$\Delta F(w) + \frac{\Delta f(w)(1 - F(w, \pi))}{f(w, \pi)} \geq 0 \ \ \forall \ w \in \mathbb{R}_+ \tag{7}$$

since the term $\Delta F(w_R(\pi), \pi)(1 - F(w_R(\pi), \pi))/(1 - F(w_R(\pi), \pi) + \delta)$ is positive (given first-order stochastic dominance). But condition (7) is ensured by assumption 1. $\square$

The dynamic programming equation in the case with learning is (in the continuous-time limit)

$$V_D(\pi) = \max_w \left\{ V_R(w,\pi) - \frac{\Delta F(w)}{1 - F(w,\pi) + \delta} \pi(1-\pi) V_D'(\pi) \right\}, \tag{8}$$

where

$$V_R(w,\pi) \equiv \frac{(1 - F(w,\pi))w}{1 - F(w,\pi) + \delta}.$$

As in the case of learning about the arrival rate, the dynamic value function is equal to a repeated value function, $V_R(\cdot,\cdot)$, minus the present value of the infinite stream of the capital losses due to learning:

$$\frac{\Delta F(w)}{1 - F(w,\pi) + \delta} \pi(1-\pi) V_D'(\pi).$$

We ignore for the moment the issue of non-concavity of the maximand in equation (8) (but we return to this below). The first-order condition for the dynamic problem is

$$\Gamma_R(w_D(\pi),\pi) - \pi(1-\pi) V_D'(\pi) \hat{\Gamma}_D(w_D(\pi)) = 0 \tag{9}$$

where

$$\hat{\Gamma}_D(w) \equiv f(w,\pi)\Delta F(w) + (1 - F(w,\pi) + \delta)\Delta f(w). \tag{10}$$

Although it is not immediately obvious, $\hat{\Gamma}_D(\cdot)$ does not actually depend on π, even though the posterior appears on the right-hand side of equation (10). In fact, all terms in π cancel out in this expression.

The sign of $\hat{\Gamma}_D$ tells us whether a higher reservation wage increases or decreases the present value of the capital losses from learning. This present value is increasing in w when $\hat{\Gamma}_D$ is positive, and decreasing in w when $\hat{\Gamma}_D$ is negative. As in the previous section, this effect operates through two channels: controlled stopping (the term $f(w,\pi)\Delta F(w)$) and controlled learning (the term $(1-F(w,\pi)+\delta)\Delta f(w)$). It is again easy to see that the controlled stopping effect, on its own, leads to a lower reservation wage for the worker. In fact, this is a feature of all models where continuation beliefs are more pessimistic than current ones and where the the action has a negative effect on the probability of stopping.

In the case of learning about common uncertainty, the controlled learning effect is always positive: a higher reservation wage is always less informative, and hence always decreases the learning capital loss. In the case of learning about idiosyncratic uncertainty, if $\hat{\Gamma}_D(w)$ had the same sign for all w, then the comparison of $w_D(\pi)$ and $w_R(\pi)$ would be equally straightforward. But this is not the case. $\hat{\Gamma}_D(0) = (\delta + 1)\Delta f(0)$ must be non-negative; and $\hat{\Gamma}_D(1) = \delta\Delta f(1)$ must be non-positive (both due to FOSD). Hence $\hat{\Gamma}_D(w)$ must cross zero from above at least once.

This non-monotonicity arises quite naturally in the model of learning about distributions. On an intuitive level, neither very high wage demands nor very low wage demands are very informative about the type of distribution: the former are always rejected by firms and the latter are always accepted.

In the rest of this section, we shall suppose that $\hat{\Gamma}_D(w)$ has a single crossing point i.e., that there is a unique $\hat{w} \in (0, 1)$ such that $\hat{\Gamma}_D(\hat{w}) = 0$ and $\hat{\Gamma}_D(w) > (<) 0$ for all $w < (>) \hat{w}$. In other words, we assume that there is a unique reservation wage that maximises the capital losses from learning. (Things are more complicated when $\hat{\Gamma}_D(w)$ has multiple crossing points. But the same arguments apply, so we concentrate on the case of a single crossing point.) As an illustration, figure 2 plots $\hat{\Gamma}_D(w)$ for the case where $F_L(w) = 2w - w^2$ and $F_H(w) = w^2$, where $w \in [0, 1]$. For this example, $\hat{\Gamma}_D(w) = 2(w^2 + (1 - 2w)(1 + \delta))$ and $\hat{w} = 1 + \delta - \sqrt{\delta(1 + \delta)}$.

Because $\hat{\Gamma}_D(w)$ changes sign, we cannot guarantee that the maximand in equation (8) is quasi-concave. The problem is that learning may be sufficiently important that the worker optimally changes her reservation wage discontinuously as her posterior changes, in order to control the progress of her beliefs. In other words, we cannot ensure in general that the first-order derivative

$$\Gamma_S(w, \pi) + \Gamma_R(w, \pi) - \pi(1 - \pi)V_D'(\pi)\hat{\Gamma}_D(w)$$

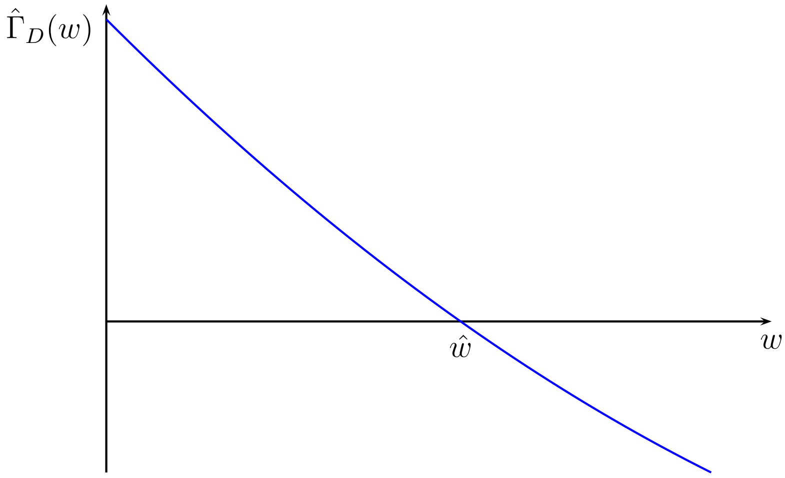

Figure 2: $\hat{\Gamma}_D(w)$ when $F_L(w) = 2w - w^2$ and $F_H(w) = w^2$ on $[0, 1]$

is single downward-crossing in w for any π. The sum of the first two terms $\Gamma_S(w, \pi) + \Gamma_R(w, \pi)$ (the first-order derivative from the repeated problem) is single downward-crossing in w. But since the slope of $\hat{\Gamma}_D$ is ambiguous (and $\pi(1-\pi)V_D'(\pi) \geq 0$), the overall slope of $\Gamma_S(w, \pi) + \Gamma_R(w, \pi) - \pi(1-\pi)V_D'(\pi)\hat{\Gamma}_D(w)$ cannot be determined in general. If $\pi(1-\pi)V_D'(\pi)$ is sufficiently large (i.e., learning is sufficiently important), then it may be that single downward-crossing is violated.

The analytical hurdle is that we are unable to derive $V_D'(\pi)$ in general, and so are unable to determine its size. Our approach is to derive the following bound on the learning effect $\pi(1-\pi)V_D'(\pi)$.

Lemma 1 *For any* $\pi \in [0, 1]$,

$$\pi(1-\pi)V_D'(\pi) \leq B(\pi) \equiv \min\left\{\pi\left(\frac{1 - F(\hat{w}, \pi) + \delta}{1 - F_L(\hat{w}) + \delta}\right)(V(1) - V_R(\pi)), \pi(1-\pi)V(1)\right\}$$

where $V(1) \equiv V_R(1) = V_D(1)$ *is the (common) value function when* $\pi = 1$*:*

$$V(1) \equiv \max_w \left\{\frac{(1 - F_H(w))w}{1 - F_H(w) + \delta}\right\}.$$

The proof of the lemma is in the appendix. In order to interpret the first part of the bound $B(\pi)$, note that equation (8) implies that

$$V_D(\pi) \geq V_R(\pi) - \frac{\Delta F(w_R(\pi))\pi(1-\pi)V_D'(\pi)}{1 - F(w_R(\pi),\pi) + \delta} \geq V_R(\pi) - \frac{\Delta F(\hat{w})\pi(1-\pi)V_D'(\pi)}{1 - F(\hat{w},\pi) + \delta}.$$

The first inequality comes from maximisation. The second inequality follows because

$$\frac{\Delta F(w)\pi(1-\pi)V_D'(\pi)}{1 - F(w,\pi) + \delta}$$

is maximised at $\hat{w}$. The first lower bound for $V_D(\pi)$ is then the value from the repeated problem, $V_R(\pi)$, minus the maximised present value of the stream of the capital losses in the dynamic problem. This lower bound on $V_D(\pi)$, combined with convexity of the value function, then gives the upper bound in the lemma on the effect of learning, measured by $\pi(1-\pi)V_D'(\pi)$.

Next, we make the following assumption. For this assumption, define $\hat{\pi}$ as follows. If $\hat{w} \in (w(0), w(1))$, then $\hat{\pi}$ is such that $w_R(\hat{\pi}) = \hat{w}$. If $\hat{w} \leq w(0)$, then $\hat{\pi} = 0$; if $\hat{w} \geq w(1)$, then $\hat{\pi} = 1$. (Here, $w(1) \equiv w_R(1) = w_D(1)$ and $w(0) \equiv w_R(0) = w_D(0)$.)

Assumption 2 *1. If $\pi \leq \hat{\pi}$, then $\Gamma_R(w,\pi) - B(\pi)\hat{\Gamma}_D(w) \leq 0$ for all $w \geq \hat{w}$.*

2. If $\pi > \hat{\pi}$, then $\Gamma_R(w,\pi) - B(\pi)\hat{\Gamma}_D(w) \geq 0$ for all $w \leq \hat{w}$.

The economic content of this assumption is to ensure that learning cannot be 'too important', by placing a limit on the upper bound of the learning effect. In order to establish any clear results, we need a condition of the form

$$\left|\frac{\partial}{\partial w}V_R(w,\pi)\right| \geq \left|\frac{\partial}{\partial w}\left(\frac{\Delta F(w)}{1 - F(w,\pi) + \delta}\pi(1-\pi)V_D'(\pi)\right)\right|.$$

In words, this requires that changing the reservation wage has a greater effect on the value of the repeated problem than it does on the capital losses arising from learning. But a condition like this has two weaknesses. First, it involves endogenous variables that we cannot solve for: in particular, $V_D(\pi)$. Hence we use the upper bound $B(\pi)$ on $\pi(1-\pi)V_D'(\pi)$ in assumption 2. Secondly, the condition is actually stronger than is needed. In the proof of proposition 4, in fact all we need is for the condition to hold when either π is small and w large (i.e., $\pi \leq \hat{\pi}$ and $w \geq \hat{w}$), or vice versa. Nevertheless, the general purpose of assumption 2 is to ensure that the value from the repeated problem matters more than learning capital losses when setting the dynamic reservation wage.

Ideally, we would like to be able to relate assumption 2 directly to the primitives of the model: δ and the distributions $F_H(\cdot)$ and $F_L(\cdot)$. Unfortunately, we have not found a way of doing this. But at least it is easy to verify whether the assumption is satisfied. For example, consider the set of distributions $F_L(w), F_H(w)$ whose densities are linear and symmetric:

$$f_L(w) = t - 2(t-1)w, \quad f_H(w) = 2 - t + 2(t-1)w$$

for $t \in [1, 2]$ and $a \in [0, 1]$. When $t = 1$, the distributions are both uniforms; when $t = 2$, the distributions are those that we used above in e.g., figure 2. We have verified numerically that assumption 2 is satisfied, for a large number of values of $\delta \geq 0$ and $t \in [1, 2]$.

Armed with the assumption, we are finally able to compare the repeated and dynamic reservation wages.

Proposition 4 *Suppose that assumption 2 holds. If* $\pi \leq (>)\hat{\pi}$*, then* $w_D(\pi) \leq (>) w_R(\pi)$*.*

Proof. We provide the proof of the first part of the proposition (if $\pi \leq \hat{\pi}$, then $w_D(\pi) \leq w_R(\pi)$). The proof of the second part is similar and so is omitted. Since $w_R(\pi)$ is non-decreasing in π (from proposition 3), if $\pi \leq \hat{\pi}$, then $\hat{w} \geq w_R(\pi)$. By assumption 2, $\Gamma_S(w,\pi) + \Gamma_R(w,\pi) - G(\pi)\hat{\Gamma}_D(w) \leq 0$ for all $w \geq \hat{w}$. But, for $w \geq \hat{w}$, $\hat{\Gamma}_D \leq 0$, and so $\Gamma_S(w,\pi) + \Gamma_R(w,\pi) - G(\pi)\hat{\Gamma}_D(w) \geq \Gamma_S(w,\pi) + \Gamma_R(w,\pi) - \pi(1-\pi)V'(\pi)\hat{\Gamma}_D(w)$. Hence $\Gamma_S(w,\pi) + \Gamma_R(w,\pi) - \pi(1-\pi)V'(\pi)\hat{\Gamma}_D(w) \leq 0$ for all $w \geq \hat{w}$. Therefore $w_D(\pi) < \hat{w}$. But for $w < \hat{w}$, $\hat{\Gamma}_D \geq 0$, and so $\Gamma_S(w,\pi) + \Gamma_R(w,\pi) - \pi(1-\pi)V'(\pi)\hat{\Gamma}_D(w) \leq \Gamma_S(w,\pi) + \Gamma_R(w,\pi)$. Hence $w_D(\pi) \leq w_R(\pi)$. □

Proposition 4 shows the difference between this application, in which Γ_D can be of either sign, and the case of common uncertainty, where Γ_D is negative. In the latter case, the dynamic reservation wage is always greater than the repeated wage, whatever the worker's beliefs. This is because the controlled learning effect always dominates the controlled stopping effect in this example; and controlled learning always calls for a higher reservation wage. In this case of learning about idiosyncratic uncertainty, a higher reservation wage can either increase or decrease the capital losses from learning. When a higher wage decreases the capital loss—when Γ_D is negative—the dynamic wage is above the repeated. This occurs when the wage is relatively high: $\hat{\Gamma}_D(w)$ is negative for $w \geq \hat{w}$. By proposition 5, this means for relatively high posteriors. On the other hand, when a higher wage increases the capital loss, the dynamic wage is above the repeated. This occurs when the posterior and the reservation wage relatively low.

This is illustrated in figure 3 (in which $F_H = w^2$ and $F_L = 2w - w^2$, and $\delta = 1$) which plots the difference between $w_D(\pi)$ and $w_R(\pi)$. It shows that $w_D(\pi)$ is greater than $w_R(\pi)$ for π greater than around 0.5. Note also that the absolute difference between $w_D(\pi)$ and $w_R(\pi)$ is greater when π is low. In this region, the controlled stopping and controlled learning effects point in the same direction: both call for a lower reservation wage. At higher values of π , the two effects point in opposite directions (and the controlled learning effect dominates). As a result, the difference between the dynamic and repeated reservation wages is smaller.

Finally, figure 4 plots $w_D(\pi)$, illustrating that it is monotonically increasing in the posterior belief (for the same distributions). Unfortunately, we cannot establish this result fully. Proposition 5 states what we are able to show.

Proposition 5 *$w_D(\pi)$ is non-decreasing in π for $\pi \leq \hat{\pi}$.*

Proof. Differentiation of the first-order condition (9), substitution of the Bellman equation (8), and use of the Envelope theorem shows that the sign of the derivative of $w_D(\pi)$ in π is given by the sign of

$$(\Delta F(w_D(\pi)))^2 + \hat{\Gamma}_D(w_D(\pi))V_D'(\pi).$$

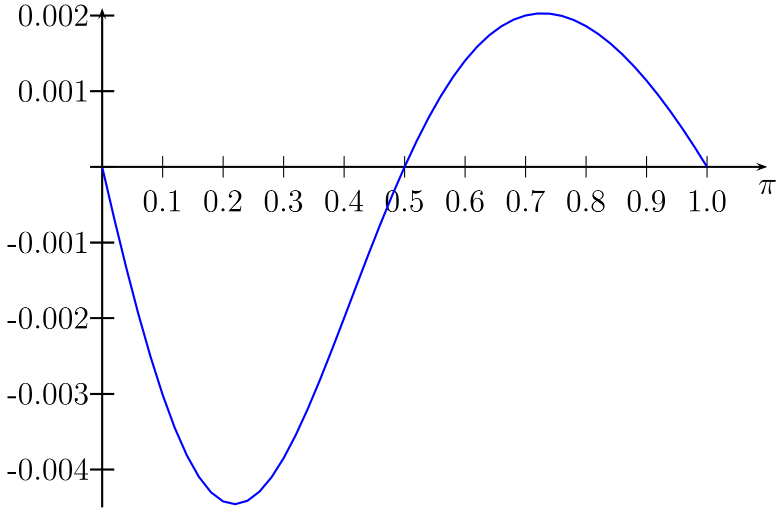

Figure 3: The difference between the repeated and dynamic reservation wages

From proposition 4, if $\pi \leq \hat{\pi}$, then $w_D(\pi) \leq \hat{w}$. If $w_D(\pi) \leq \hat{w}$, then $\hat{\Gamma}_D(w_D(\pi)) \geq 0$. The result follows. □

For π less than $\hat{\pi}$, the controlled stopping and controlled learning effects point in the same direction; and both induce monotonicity in the dynamic reservation wage. But when π is above $\hat{\pi}$, these two effects pull in opposite directions. Proposition 4 shows that we are able to determine the balance in terms of the *levels* of reservation wages. It appears to be much more difficult to resolve how the competing factors affect the *slope* of the dynamic reservation wage. The learning effect can, potentially at least, be non-monotonic in the posterior. We have yet to find a workable bound on the size of the rate of change of the learning effect that will allow us to establish monotonicity for all values of π. Numerical analysis indicates that, for the set of distributions with densities

$$f_L(w) = t - 2(t-1)w, \quad f_H(w) = 2 - t + 2(t-1)w$$

for $t \in [1, 2]$ and $w \in [0, 1]$, $w_D(\pi)$ is monotonic in π.

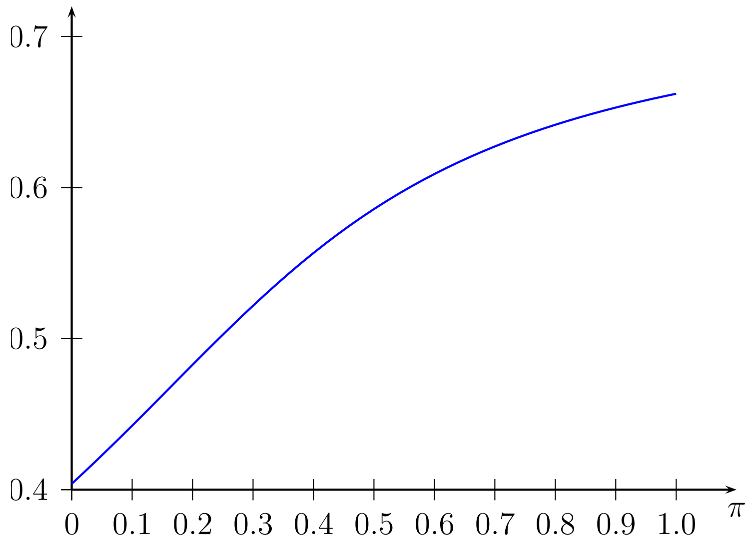

Figure 4: The dynamic reservation wage $w_D(\pi)$

5 Conclusion

We have written the model in terms of a job search problem in order to see the effects of learning within a concrete economic application. The model can easily be generalised to accommodate some other standard economic models. By introducing a flow cost for actions, it is possible to transform the model of learning about arrival rates into a model of R&D where the true success probability is initially unknown. Furthermore, there is no particularly compelling reason to concentrate only on Bayesian learning models. The change in posteriors could equally well reflect accumulation of past accumulated actions. Hence this model can be used as a basis for a non-Poisson model of R&D.[2]

The entire paper is written in terms of ever more pessimistic continuation beliefs. Obviously we could have taken an application where continuation beliefs drift upwards conditional on no stopping. As an example, consider the maintenance problem of a machine whose durability is uncertain. (Or the problem of finding the optimal effort level to demand from an agent or worker). As long as the machine does not break down, beliefs about its longevity become more optimistic. Apart from the obvious change in signs, the

[2] A previous version of this paper includes such an application and is available from the authors upon request.

controlled stopping and learning effects are present in such models as well.

In this paper, we have analysed models of Poisson-type information revelation where conditional on no stopping, the process of beliefs is deterministic. A useful extension of the model could consider beliefs that are stochastic even conditional on no stopping. Models of controlled Brownian motion would be a natural analytical framework for this. Such an extension would be particularly valuable when extending the model to multi-agent situations. While a multi-agent version of the R&D race with learning seems relatively straightforward (since the game ends at the moment of first innovation), models of informational externalities do not seem to fit the Poisson framework that well. In order for our methods to work, beliefs must evolve smoothly over time. A model where the beliefs are driven by controlled Brownian motions, observations on other players could be smoothed by observational noise in such a way that the model is still applicable. These models might lead to even higher current wage demands by the workers. Since the value function is convex in beliefs, information flowing from the decisions of other players tends to increase the continuation value to any player. Hence this would counteract the controlled stopping effect and push the wage demands further up.

We leave the analysis of these extensions to future research.

Proof of lemma 1

To show the first part of the bound, note that from the Bellman equation (8) and maximisation,

$$\delta V_D(\pi) \geq (1 - F(w, \pi))(w - V_D(\pi)) - \Delta F(w)\pi(1 - \pi)V_D'(\pi)$$

for any $w \in [0,1]$. In particular, this inequality holds for $w = w_R(\pi)$. Rearranging, this gives

$$\begin{aligned} V_D(\pi) &\geq \frac{(1-F(w_R(\pi),\pi))w_R(\pi)}{1-F(w_R(\pi),\pi)+\delta} - \frac{\Delta F(w_R(\pi))}{1-F(w_R(\pi),\pi)+\delta}\pi(1-\pi)V_D'(\pi) \\ &= V_R(\pi) - \frac{\Delta F(w_R(\pi))}{1-F(w_R(\pi),\pi)+\delta}\pi(1-\pi)V_D'(\pi). \end{aligned}$$

The following lemma establishes that $V_D(\pi)$ is convex.

Lemma 2 *$V_D(\pi)$ is convex in π.*

Proof. Consider two points, π and π' and let $\pi^\lambda = \lambda\pi + (1-\lambda)\pi'$. We want to show that for all π, π' and π^λ, we have:

$$V_D(\pi^\lambda) \leq \lambda V_D(\pi) + (1-\lambda)V_D(\pi').$$

Denote by $\mathbf{w}^\lambda$ the path of optimal actions starting from π^λ. Denote by $W^H(\mathbf{w})$ the expected payoff from an arbitrary path of prices $\mathbf{w}$ conditional on the true distribution being H and similarly for $W^L(\mathbf{w})$. Then we have:

$$V_D(\pi^\lambda) = \pi^\lambda W^H(\mathbf{w}^\lambda) + (1-\pi^\lambda)W^L(\mathbf{w}^\lambda).$$

We also know that

$$\begin{aligned} V_D(\pi) &\geq \pi W^H(\mathbf{w}^\lambda) + (1-\pi)W^L(\mathbf{w}^\lambda), \\ V_D(\pi') &\geq \pi' W^H(\mathbf{w}^\lambda) + (1-\pi')W^L(\mathbf{w}^\lambda) \end{aligned}$$

since $\mathbf{w}^\lambda$ is a feasible price path.

Hence

$$\begin{aligned}\lambda V_D(\pi) + (1-\lambda)V_D(\pi') \geq \lambda(\pi W^H(\mathbf{w}^\lambda) + (1-\pi)W^L(\mathbf{w}^\lambda))\\ + (1-\lambda)(\pi' W^H(\mathbf{a}^\lambda) + (1-\pi')W^L(\mathbf{w}^\lambda))\\ = (\lambda\pi + (1-\lambda)\pi')W^H(\mathbf{w}^\lambda) + (1-\lambda\pi - (1-\lambda)\pi')W^L(\mathbf{w}^\lambda)\\ = \pi^\lambda W^H(\mathbf{w}^\lambda) + (1-\pi^\lambda)W^L(\mathbf{w}^\lambda) = V_D(\pi^\lambda).\end{aligned}$$

This proves the claim. □

Since $V_D(\pi)$ is convex, $V_D'(\pi) \leq (V(1) - V_D(\pi))/(1-\pi)$. Therefore

$$V_D(\pi) \geq V_R(\pi) - \frac{\Delta F(w_R(\pi))}{1 - F(w_R(\pi), \pi) + \delta}\pi(V(1) - V_D(\pi)).$$

Note that $\Delta F(w)/(1 - F(w,\pi) + \delta)$ is maximised at $\hat{w}$. Hence

$$V_D(\pi) \geq V_R(\pi) - \frac{\Delta F(\hat{w})}{1 - F((\hat{w}, \pi) + \delta}\pi(V(1) - V_D(\pi)).$$

Rearranging this inequality, and using convexity again, means that

$$\pi(1-\pi)V_D'(\pi) \leq \pi(V(1) - V_D(\pi)) \leq \pi\left(\frac{1 - F(\hat{w},\pi) + \delta}{1 - F_L(\hat{w}) + \delta}\right)(V(1) - V_R(\pi)),$$

which completes the first part of the proof. For the second half of the bound, note that maximisation implies that $V_D(\pi) \geq \pi V(1)$. Since this inequality holds with strict equality at $\pi = 1$, it implies that $V_D'(\pi) \leq V(1)$. This completes the proof.

References

Anderson, A., and L. Smith (2006): "Explosive Convexity of the Value Function in Learning Models," May 21.

Burdett, K., and T. Vishwanath (1988): "Declining Reservation Wages and Learning," *Review of Economic Studies*, 55, 655–665.

Grossman, G. M., and C. Shapiro (1986): "Optimal Dynamic R&D Programs," *RAND Journal of Economics*, 17(4), 581–593.

Lancaster, T., and A. Chesher (1983): "An Econometric Analysis of Reservation Wages," *Econometrica*, 51, 1661–1676.

Malueg, D. A., and S. O. Tsutsui (1997): "Dynamic R&D Competition with Learning," *RAND Journal of Economics*, 28(4), 751–772.

Merlo, A., and F. Ortalo-Magne (2004): "Bargaining over Residential Real Estate: Evidence from England," *Journal of Urban Economics*, 56, 192–216.

Rogerson, R., R. Shimer, and R. Wright (2005): "Search-Theoretic Models of the Labor Market: A Survey," *Journal of Economic Literature*, XLIII, 959–988.

Smith, L. (1999): "Optimal job search in a changing world," *Mathematical Social Sciences*, 38, 1–9.

Van den Berg, G. (1990): "Nonstationarity in Job Search Theory," *Review of Economic Studies*, 57, 255–277.