Specialized Careers

Heski Bar-Isaac* and Johannes Hörner†

August 2011

Abstract

An agent has different abilities in two types of tasks, which are revealed through his performance over time. He initially decides whether to engage in only one task (specialize) or to take on any task that arises (be a generalist). This decision trades off the cost of being idle against staying available for relatively lucrative tasks. We compare specializing with acting as a generalist in an infinite-horizon model and provide complete characterizations of efforts. We show how specializing acts as a means of committing to exert more effort. In a two-period version of the model, this implies that positive fees for switching strategies are desirable.

JEL: C73(Stochastic and Dynamic Games; Evolutionary Games; Repeated Games); D83 (Search; Learning; Information and Knowledge; Communication; Belief); J24 (Human Capital; Skills; Occupational Choice; Labor Productivity); M50 (Personnel Economics: General)

Keywords: careers, specialization, incentives, career concerns

1 Introduction

The idea that there can be private and social benefits from specialization is well-known and much-celebrated. Indeed, in March 2007, the Bank of England began to issue a banknote featuring Adam Smith and a pin factory. In the famous example of Smith (1776) and in much subsequent literature, gains from specialization arise for technological reasons. The literature has tended to abstract from effort incentives and to focus instead on the

*Corresponding Author: Bar-Isaac: Department of Economics, Stern School of Business, NYU, 44 West 4th street 7-73, NYC, NY 10012 USA. heski@nyu.edu; 212 998 0533 (tel); 212 995 4218 (fax).

†Hörner: Department of Economics, Yale University, 30 Hillhouse Ave., New Haven CT 06517, USA. johannes.horner@yale.edu; 203 432-5352 (tel); 203 432-6167 (fax).

market demand and supply of specialized skills (often simply contrasting between skilled and unskilled labor) to determine the extent of specialization in equilibrium.[1]

In this paper, we take a more partial approach and highlight a complementary mechanism. We consider a single individual's decision to specialize, and we highlight that specialization (which is relatively easy to monitor and enforce) can provide incentives for effort that might otherwise be absent. Specializing entails spending relatively more time engaged in a given task, and so the incentives to improve reputation in that task are greater than they would be for a generalist. The cost of specialization is that the agent's lower flexibility might lead to taking fewer assignments.[2]

Of course, it may, in fact, turn out that the agent has little ability in that task; moreover, these incentives diminish over time. Therefore, specialization may require commitment. This is in contrast to purely technological reasons and other rationalizations for specialization, such as comparative advantage or exploiting returns for individual specific human-capital investments (as in Rosen, 1978, 1983). Enforcing specialization, which is easier (but less effective) than enforcing effort directly, can be done formally or informally. Professional organizations often directly constrain the type of work in which a professional can be engaged. For example, the Law Society of England and Wales (2004) imposes considerable restrictions on the non-legal activities that a solicitor can undertake, especially from the same physical premises or under the same business name.[3] Informal and cultural constraints can also play a role in limiting the scope and variety of an agent's work. Here, we simply take as given that such formal and informal institutions provide a commitment device that can generate greater overall efficiency but may introduce some inefficiency ex-post.[4] Alternatively, our results can be viewed as illustrating consequences of such rules,

[1]For a recent overview, see Acemoglu and Autor (forthcoming). In addition to considering the extent or scale of the market, the literature has also addressed the costs of coordination as limiting the extent of specialization (Becker and Murphy, 1992) and extensively explored complementarities with information and communication technologies.

[2]In the model, an agent who is not engaged in a specialized task is idle; in application this "idle" time may reflect less lucrative and less specialized work that is always freely available: for example in a consulting firm this may be data-gathering, fact-checking etc in addition to time without engagement.

[3]It is hard to imagine that clients would not know of such activities, particularly at the same premises. Thus, in ascribing such rules to conflicts of interest, one would have to assume some paternalism on the part of the Law Society. This is a plausible, though not entirely satisfactory explanation—in particular, with respect to commercial work, where one might suppose that clients are fairly sophisticated and where there may be benefits to relaxing these restrictions and enjoying economies of scope.

[4]In the model, we largely take the somewhat extreme view that agents either fully specialize or fully generalize and cannot switch strategies. In practice, things may not be so extreme. In a two-period analog of the model, considered in Section 5, we are able to characterize optimal penalties or switching costs from

even if they arise for other reasons.

Specialization in professional services has typically been considered as a strategic choice at the firm rather than the individual level (Chatain and Zemsky, 2007; Garicano and Hubbard, 2009), and incentives for effort have not been addressed. However, for the medical profession specifically, there is a long-standing literature that seeks to address the extent to which physicians' choices of specialization respond to lifetime earnings in these specialities (Sloan, 1970; Nicholson, 2002; Gagné and Léger, 2005). In part, this interest reflects a policy concern over the aggregate provision of care and shortages in certain practice areas. Courty and Marschke (2008) present an interesting analysis based on modeling different specialities as differing in their exposure to moral hazard; the analysis relates the failure of the market to efficiently allocate health care services with the recent growth in medical specialization and sub-specialization. Our analysis, rather than being a full market analysis, is more modest in scope, focusing on an individual's decision to specialize or not. In this sense, it is perhaps closer to the literature on career concerns, which typically does not address a full market equilibrium—indeed, it is typically predicated on the assumption that the agent has market power and earns all the surplus that he generates.

Formally, we build on the career-concern model introduced in Fama (1980) and Holmström (1982/99). In this seminal model, an agent's current performance influences potential employers' assessment of his ability and, so, his future wages. As a consequence, an agent exerts effort to boost his future wages, even in the absence of outcome-contingent contracts. We adapt the model in a number of respects. First, different kinds of opportunities arise, and the agent's abilities in these different kinds of work may differ. Second, opportunities are not always available, either because the agent is occupied, or because no appropriate task is on offer (thus, the agent may spend some time relatively idle). Finally, we suppose that the agent initially chooses between acting as a generalist who undertakes any available task and as a specialist who undertakes only one kind of task.

Having set up the model, our contribution lies in characterizing and contrasting the expected lifetime values and efforts associated with acting as a specialist or as a generalist. Most importantly, we show how committing to be specialized can be optimal, as it enhances the reputational incentives, despite a lower workload. In our analysis, we provide a closed-form solution for an agent's effort in the basic Holmström (1982/99) framework and replicate some intuitions that arise in the standard career-concerns model: Holding all

moving from specialization to acting as a generalist, and we highlight that similar intuitions apply.

else constant, efforts decrease over time and in the precision of potential employers' beliefs about the agent's ability, and efforts increase in the discount factor. The same is true for an agent's lifetime value when efforts are inefficiently low, and an agent's lifetime value is, also, increasing in an agent's reputation.

Decreasing precisions, increasing discount factors and increasing reputation raises the value both of acting as a generalist and acting as a specialist. Thus, to compare the two, and to see when each is more attractive, it is not sufficient to consider the direction of the comparative statics. We must also quantify the effects. In general, this is a difficult problem—in part because there are several forces at play. First, independent of any effort considerations, specialization may be preferable when the agent has a better reputation in the specialized task. In considering specialization, an agent has to trade off taking more opportunities to work in the lucrative task against spending more time idle.[5] Turning to effects on effort, in general, effort may be above the efficient level (since current effort affects the perceptions about ability in the future), and so factors that boost effort can be harmful. More subtly, note that the extent to which factors boost effort depends on the current level of effort and the slope of the cost-of-effort function. In the general case, the analysis is complicated since current effort has a perpetual influence that is difficult to quantify explicitly, and comparative statics affect not only current efforts, but also future efforts.

While our baseline model allows us to highlight the trade-off between effort incentives and idle time in a rich environment, it is too general to provide clear-cut comparative statics. To do so, we add further structure in specifying a functional form for the cost of effort and simplifying the effects on the future. We do this by first considering a stationary model and then by analyzing a two-period analog of the model. In the two-period analog, where effects on the continuation value are relatively simple, we extend our analysis to suppose that, at some cost, an agent can switch from a specialist to a generalist strategy.

The forces at play suggest that comparative statics on fundamentals can push towards specializing in some circumstances and towards acting as a generalist in others. However, we show that in the symmetric stationary case and in the two-period analog, specialization is preferred if effort is a relatively important part of the fee (the lower is the initial reputation). In the stationary case, increasing the rate of time discount or the time it takes to complete a task, decreasing the frequency with which projects are available, or reducing

[5] In particular, therefore, market or organizational institutions that ensure little idle time will benefit specializing over acting as a generalist.

the noise in the monitoring all favor acting as a generalist over specializing (Proposition 1). In the two-period model, the optimal cost of switching from a specialist strategy in the first period to a generalist strategy in the second increases in the cost of effort and the initial reputation in the specialized task, but decreases in the discount factor and the precision of beliefs about the agent's ability in the specialized task. Loosely, the intuition for these results is that the commitment to exert more effort implied by specialization is more valuable the more important effort is as compared to ability in production, and where boosting incentives to effort is most valuable. In addition, however, there are effects independent of effort; for example, it is very costly to give up opportunities to work if these are few and far between.

Dewatripont, Jewitt and Tirole (1999a and b) have highlighted a different mechanism that suggests that career-concerns incentives might be strengthened through "focus."[6] Most of their analysis assumes that ability is scalar and identical for all tasks, and they highlight that when the agent focuses on one of the many tasks available in the first period, the signal-to-noise ratio is reduced and the agent's effort increased. In Section 6 of Dewatripont, Jewitt and Tirole (1999b), the authors briefly consider task-specific ability; however, they do not consider different types of production but always assume that the agent works in both tasks. In our context, this would necessitate the agent acting as a generalist; instead, in their context, "specialization" refers to the degree of correlation in the task-specific abilities.

Holmström and Milgrom (1991) provide a further benefit of specialization in a moral-hazard environment—removing the opportunity to exert effort in one task, leads to higher (and more-valued) effort in other tasks. Such a benefit does not arise in the symmetric environment that we consider, where the cost of effort in each task is independent of the effort in the other task.

Related particularly to the two-period version of the model, Martinez (2009), Casas-Arce (2010) and Kovrinykh (2007) consider career concerns with one dimension of ability, but allow for agents to switch between multiple employment sectors. The focus in these papers is on how the non-linear returns to reputation that arise imply that career concerns vary with the current level of reputation. Finally, the literature on optimal switching

[6] Dewatripont, Jewitt and Tirole, working with a two-period model, address rich production technologies (allowing effort and ability to be both substitutes and complements) and information structure. As in Holmström (1982/99), we simplify considerably in some respects by considering a linear-normal model, but this allows us to explore and highlight different dynamic considerations and differing abilities in different tasks.

costs in labor markets is also relevant. Most closely related is the discussion of such switching costs as a means of protecting investments in firm-specific human capital, as discussed in Gibbons and Waldman (1999). Current productive effort in our model, can be seen as an investment in task-specific reputation, and the restriction to specialization in our general model can be viewed as an infinite cost of switching tasks. Other rationales for endogenous switching costs in labor markets include severance as a means of risk-sharing (Hart, 1983; Malcomson, 1999), of rent extraction (Burguet, Caminal and Matutes, 2002), and of ensuring more-efficient allocation in the presence of information asymmetries (Burguet, Caminal and Matutes, 2002).

2 Model

An agent has ability $\eta = (\eta_A, \eta_B)$ where η_i denotes his ability in task of type $i = A, B$.[7] Initially, the market and the agent share common beliefs about η. These beliefs are given by a bivariate normal distribution with mean $m_0 = (m_{0A}, m_{0B})$ and precision matrix[8]

$$H_0 = \begin{pmatrix} h_{0A} & 0 \\ 0 & h_{0B} \end{pmatrix}. \tag{1}$$

The agent engages in tasks over time, and as he undertakes tasks, learning about his ability occurs through the observation of his output. In particular, if an agent undertakes task i for the t^{th} time, then the output produced is given by

$$y_{it} := \eta_i + a_{it} + \varepsilon_{it}, \tag{2}$$

where a_{it} is the effort that the agent undertakes in the task at cost $c(a_{it})$, and ε_{it} is a noise term that is normally distributed and independently drawn for each new task; without loss of generality, we suppose that the distribution of ε_t^i has mean 0 and precision 1. We suppose that $c(a)$ is strictly convex, increasing and twice continuously differentiable, with $c(0) = 0$, $c'(0) = 0$ and $\lim_{a\to\infty} c'(a) = \infty$.

Time is continuous and denoted by $s \geq 0$. Opportunities are not always available,

[7] For ease of comparison, we adopt the notation of Holmström (1982/99) where possible.

[8] Note that since the off-diagonal elements are set to 0, the abilities in the two tasks are independent. The model can be extended to allow for correlation, though expressions are somewhat cumbersome. Moreover, there is continuity, and so allowing for some minimal correlation does not affect qualitative results. See Appendix B.

and completing tasks is time-consuming. For ease of exposition (though this can be easily relaxed at the cost of some additional notation), we suppose that A and B tasks are identical with respect to their arrival rates and durations. Specifically, both A and B tasks arrive at the constant Poisson rate $\pi > 0$, and a task of either type takes time $l > 0$ to complete.[9] It is convenient to write $\delta := e^{-lr}$ where $r > 0$ represents the agent's rate of time discounting and to take as a primitive this discounted delay cost until the agent becomes available. Tasks are mutually exclusive. That is, if the agent chooses to undertake some task, he cannot undertake other tasks until completion. This is a crucial aspect of the model since without this realistic feature, there is no trade-off in the model; specialization would simply entail foregoing lucrative B tasks. We further assume that tasks are time-sensitive, so the tasks cannot be stored: If a task arrives while an agent is busy with another task, he cannot store the opportunity. Again, this seems a realistic assumption in the context of service industries. Available tasks, completed tasks, and performance in these tasks are publicly observable.

Turning to compensation, we suppose that each time he undertakes a task, the agent receives an outsider's expectation of output for that task (which, in particular, will depend not only on the prior on the agent's ability in the task, but also on the expected effort).[10] The agent is risk-neutral and discounts future payoffs at the rate r, aiming to maximize the present discounted sum of payments, net of costs of effort.

3 Analysis

3.1 Specialist

First, we consider the case of specialization, which, without loss of generality, we take to be in task A. We first describe how beliefs are updated and then use this to characterize the agent's effort and the value of specialization.

We can restrict attention to the beliefs and updating on the A task (since the agent under specialization undertakes only the A task). Suppose that an agent begins a period in which the common belief is that the agent's ability in A is distributed with mean m_A

[9] Thus, in our model, we consider random arrival of projects. This is intended as a shortcut to model that opportunities in different tasks might not only arrive randomly, but also vary with respect to the value or size of a given project.

[10] Similar qualitative results arise when the agent does not receive the full expected surplus, but any (fixed) fraction. This compensation structure might arise, for example, through a third party (a clerk in a barristers' chambers, a more senior partner in a firm, or an agency) setting prices and allocating work.

and precision h_A and is expected to exert effort $a_A^*(m, H)$. Then, if the agent undertakes an A task and an output $y = \eta_A + a_A + \widetilde{\varepsilon}_A$ is observed, the expressions for the posterior means and posterior precisions are well-known (see DeGroot, 1970). Specifically, we can define the innovation $z_A := y - a_A^*(m, H)$. The posterior is, then, given by

$$\widetilde{m}_A = \frac{h_A m_A + z_A}{h_A + 1}. \tag{3}$$

In particular, after t tasks have been undertaken, the posterior mean on task A is

$$\widetilde{m}_{A,t} = \frac{(h_A + t)m_{A,t-1} + z_{A,t-1}}{h_A + t + 1}, \tag{4}$$

where $m_{A,t-1}$ is the mean immediately before the last task was undertaken, and $z_{A,t-1}$ is the last innovation. Thus, ability is a martingale. The precision matrix evolves deterministically, following a new observation. Specifically, the precision on task A, h_A, evolves to $h_A + 1$, so that, after t tasks, the posterior precision is given by $h_A + t$. These characterizations are identical to those in Holmström (1982/99).

However, here, tasks arrive at random. In order to characterize the equilibrium effort, we first describe the problem over short, but positive time intervals of length ds. Let $V_S(m, m', h_A)$ denote the value of a currently idle agent, where the public belief is m with precision h_A, and her own beliefs and precision are given by m' and h_A. This value can be written as a discounted sum. The first component is the value of staying idle and maintaining the same continuation value (as in the first line of (5)). Next, there is a component that includes the probability that an A task arrives in the next instant (with probability "πds"), multiplied by the sum of the current payoff from the task—the outsiders' expectation of output, which is the sum of ability m and anticipated effort a^*—minus the cost of the actual effort exerted $c(a)$, and the discounted expected continuation value of being idle once the task is completed. This last value will incorporate updated beliefs and precision. Although it is possible that two tasks arrive in the same instant, this probability is of second order. So, the optimality equation can be written, to the second order, as

$$\begin{aligned} V_S(m_A, m'_A, H) &= (1 - \pi ds)\,(1 - r ds)\, V_S(m_A, m'_A, h_A) + \\ &\pi ds \max_a \left\{ m_A + a_A^*(m_A, m'_A, h_A) - c(a) + \delta \mathbb{E}[V(\widetilde{m}_A, \widetilde{m}'_A, h_A + 1)] \right\} + o(ds), \end{aligned} \tag{5}$$

where $\widetilde{m}$ denotes the posterior mean belief (which depends on the agent's action and the realization of the random variable ε_A), and $a^*(m_A, m'_A, h_A)$ denotes the market's expectations of the agent's effort when in state (m_A, m'_A, h_A). Note that although the agent knows the market's belief, the market only infers the agent's, and so its expectation of effort depends only on the part of the state that it can observe.

Rearranging and taking the limit as ds tends to zero, it follows that

$$V_S(m_A, m'_A, h_A) = \frac{\pi}{\pi + r} \max_a \left\{ m_A + a^*_A(m_A, m'_A, h_A) - c(a) + \delta \mathbb{E}[V(\widetilde{m}_A, \widetilde{m}'_A, h_A + 1)] \right\}. \tag{6}$$

(Appendix C provides a more formal argument). Therefore, in equilibrium,

$$a^*(m_A, m'_A, h_A) = \arg\max_a \left\{ \delta \mathbb{E}[V(\widetilde{m}, \widetilde{m}', H + 1)] - c(a) \right\}. \tag{7}$$

Note that the agent's incentives for effort arise from the desire to generate an output that is higher, to convince the market that he is of relatively high ability. The strength of this incentive does not vary with either his own or outsiders' current belief about his ability (since the extent to which output varies with effort and changes the posterior is independent of the prior mean). This allows us to note that $a^*(m_A, m'_A, h_A)$ is constant in m_A and m'_A. Note, also that in equilibrium, since private and social beliefs coincide initially and social expectations on actions are correct, social and private beliefs move in tandem along the equilibrium path. It is convenient, therefore, to write $S(m_A, h_A) := V_S(m_A, m_A, h_A)$.

We can then provide a relatively simple characterization of the agent's value function and optimal effort.

Theorem 1 *The value function in the case of specialization is equal to*

$$S(m_A, h_A) = \frac{\pi m_A}{\pi(1-\delta) + r} + \frac{\pi}{r + \pi} \sum_{t=0}^{\infty} \left(\frac{\pi \delta}{r + \pi} \right)^t [a_t(h_A) - c(a_t(h_A))], \tag{8}$$

where $a_t(h_A)$ is the unique solution to

$$c'(a_t) = \int_0^1 \frac{u^{h_A + t}}{\frac{\pi + r}{\pi \delta} - u} du.^{11} \tag{9}$$

[11] The reader might be surprised at the appearance. Given (8), one would expect optimal effort to be characterized by a series. As shown in Appendix A, the series can be re-expressed as an integral via the properties of hypergeometric functions.

Observe that $S(m_A, h_A)$ and $a_t(h_A)$ are continuously differentiable in all the parameters. Therefore, this observation, together with the characterization above, shows the following:

Corollary 1 *The first component of the value function, namely*

$$\frac{\pi m_A}{\pi(1-\delta)+r}, \tag{10}$$

is increasing in π, m_A and δ and decreasing in r.

Effort a_t is decreasing in h_A, t, and r and increasing in π and δ.

These comparative statics are intuitive: The part of the value that is independent of effort is higher if tasks arrive more frequently (π increases), if the future is worth more (r is lower), and if tasks are completed more quickly (δ is higher), allowing for less delay between tasks.

The comparative statics of effort are familiar from standard career-concern models. There is more effort if there is greater uncertainty about the agent's type (that is, h_A is lower), if the future is more valued, or if the next task arrives more rapidly (which will be the case when r is lower and π is higher).

It is not clear, in general, whether the function $a \mapsto a - c(a)$ is increasing or decreasing in a at $a = a_t$. At the efficient, first-best, level of effort a^{fb}, this should be constant—that is, $c'(a^{fb}) = 1$. However, at a_t, in general, $c'(a_t)$ can be either greater or less than 1. The intuition here is that if there is considerable uncertainty about the agent's type (which will be the case early in the career), then effort can shift beliefs about the agent's ability and has an effect that will diminish as the agent undertakes more tasks, but will do so slowly. After the agent has already undertaken many tasks, there is high precision about the agent's ability, and so effort has a rather muted effect and, in this case, is below the efficient level. More formally, note from (9) that $c'(a_t)$ is decreasing in t and tends to 0 as $t \to \infty$. It follows that $a_t - c(a_t)$ is decreasing in a_t if and only if $t < \bar{t}$ for some $\bar{t}$ (possibly 0).

Bringing together the results from Corollary 1 and the paragraph above, and then ignoring the first $\bar{t}$ periods, or assuming that h_A is large enough, the comparative statics for the value function follow, because comparative statics agree for both terms of the value function whenever $a - c(a)$ increases in a. In this case, therefore, the agent's value function is increasing in π and m_A and decreasing in h_A, r and l.

3.2 Generalist

We now consider an agent who undertakes whichever task first presents itself. We start by characterizing the extent to which a specialized agent is idle or employed on some task. Following the same steps as before, the optimality equation for a broad agent who is idle and in state $(m_A, m_B, m'_A, m'_B, h_A, h_B)$ is given by

$$\begin{aligned} & V_G(m_A, m_B, m'_A, m'_B, h_A, h_B) \\ = \quad & \frac{\pi}{r+2\pi} \max_{a_A} \left\{ m_A + a_A^* - c(a_A) + \delta \mathbb{E}[V_G(\widetilde{m}_A, m_B, \widetilde{m}'_A, m'_B, h_A + 1, h_B)] \right\} \\ & + \frac{\pi}{r+2\pi} \max_{a_B} \left\{ m_B + a_B^* - c(a_B) + \delta \mathbb{E}[V_G(m_A, \widetilde{m}_B, m'_A, \widetilde{m}'_B, h_A, h_B + 1)] \right\}, \end{aligned} \tag{11}$$

where a_A^* and a_B^* are the anticipated equilibrium efforts.

Following (3) and analogous to (4), we can write the posteriors, noting that following an A task, posteriors about ability in B are unchanged and vice versa. Note that, just as in the specialized case, following a task of type $i = A, B$, precision evolves deterministically to $h_i + 1$ and h_{-i}, and that efforts depend only on the precision and not on the level of beliefs. As above, it is convenient to define $G(m_A, m_B, h_A, h_B) := V_G(m_A, m_B, m_A, m_B, h_A, h_B)$.

Theorem 2 *The value function in the case of generalizing is equal to*

$$G(m_A, m_B, h_A, h_B) = \frac{\pi(m_A + m_B)}{r + \pi(2-\delta)} + \frac{\pi}{r + \pi(2-\delta)} \sum_{t=0}^{\infty} \left(\frac{\pi\delta}{r+2\pi} \right)^t [a_A(t) - c(a_A(t)) + a_B(t) - c(a_B(t))], \tag{12}$$

where $a_A(t)$ is the unique solution to

$$c'(a_A(t)) = \int_0^1 \frac{u^{h_A+t}}{\frac{\pi + (1-\delta)\pi + r}{\pi\delta} - u} du, \tag{13}$$

and analogously for $a_B(t)$.

Proof. The steps are analogous to the proof of Theorem 1 and, so, are omitted. ■

As for the specialist case, for t_A and t_B or h_A and h_B high enough, then $a_A - c(a_A(t_A))$ and $a_B - c(a_B(t_B))$ are increasing functions. And, in this case, the first component of V_G that is independent of efforts, and the second component that depends on effort will agree in their comparative statics. Thus, overall, and similar to the specialist case and with identical intuitions, the value to a generalist increases in perceived abilities, m_A and m_B

and the frequency with which tasks arrive, π, and decreases in precisions about abilities h_A and h_B, the interest rate r and the length of time it takes to complete a task l.

3.3 Comparing Specializing and Generalizing

We now turn to the central question posed in the introduction: When should an agent choose to be a specialist rather than a generalist?

If effort is irrelevant—e.g., if it is prohibitively costly or beliefs are fully precise ($h_A = h_B = \infty$)—this is simply determined. Even in this case, the choice is not trivial since the agent must trade off that, as a generalist, he will likely undertake fewer A tasks because sometimes these will arrive when he is engaged in a B task, against the greater number of tasks overall. This trade-off depends on the extent to which A tasks are more lucrative than B tasks, the frequency with which tasks arrive, the duration of a task, and the rate of time discounting.

By comparing the value from acting as a specialist with the value from acting as a generalist, we obtain the following condition that ensures that the agent prefers to specialize if and only if

$$\frac{m_A}{m_B} > 1 + \frac{r}{\pi(1-\delta)}. \tag{14}$$

Note that this condition is violated whenever $m_A \leq m_B$; in this case, the overriding effect is that, as a specialist, the agent undertakes fewer tasks overall. The higher the relative value of A tasks as compared to B tasks (the higher is $\frac{m_A}{m_B}$), the more attractive specialization is. Other comparative statics are also intuitive. A higher value of π suggests that tasks arrive more frequently, and so a specialist is seldom idle; similarly, an increase in the length of time to complete a task suggests that the specialist spends a large fraction of time occupied. Indeed, in the limit—that is, if $\pi \to \infty$ or $\delta \to 0$, so that tasks are constantly available or tasks last a lifetime—the agent will prefer to specialize unless he is equally adept in both tasks. This highlights a natural conclusion—that market or organizational scale or institutions that minimize idle time benefit specialization. Finally, notice that increasing r makes Condition (14) more difficult to satisfy; that is, greater time discounting makes generalization more attractive. The intuition here is that generalizing hastens the arrival of the first task since the agent undertakes it as soon as it arrives rather than waiting for a specific task, and this effect is more important the more the agent values current revenue over future revenue.

When compensation for effort accounts for a significant part of an agent's lifetime earnings, comparisons between specializing and acting as a generalist are more complicated.

It is immediate on comparing (9) and (13) that, controlling for the number of A tasks already undertaken, a specialist exerts more effort in the next A task than a generalist does. The intuition is clear: A generalist expects a longer delay before the next time (and subsequent occasions) that he will undertake an A task, and so the disciplinary pressure that boosting his reputation imposes on effort is weaker than it is for a specialist. However, in considering comparative statics and comparing the total value of acting as a specialist or as a generalist, it is not enough to simply compare the series of efforts, term by term, since these are discounted at different rates, and comparative statics affect both the individual efforts and the discount rates. In addition, only efforts in the A task are relevant for the specialist, while for a generalist, one must include efforts in both tasks. Moreover, the contributions to value from effort in individual tasks are complicated (though closed-form) expressions since effort has an effect not only on the output of the current task, but also on the reputation (and compensation) in all future tasks of that type. In order to build on the intuition from the comparative statics of effort, we simplify the environments that we consider. We analyze, first, a stationary version of the model, in which efforts are constant throughout, and then a two-period analog of the environment.

Since the effects of anticipated abilities m_A and m_B are easy to understand and independent of the value of effort, going forward, we simplify notation by setting $m_A = m_B = m$. Note that with $m_A = m_B$, absent contributions from the compensation for efforts, the agent unambiguously prefers to act as a generalist.

To summarize, while each version of the model (specializing and generalizing) displays intuitive comparative statics, the comparison between their payoffs is less clear-cut. The point, however, is that committing to be a specialist can dominate acting as a generalist in a variety of circumstances. We believe that our framework might provide a good starting point for more-general (perhaps numerical) analyses, allowing, for instance, the agent to optimally switch between being a specialist and a generalist. Definitive comparative statics require us to impose more structure. To make our dynamics more manageable, two routes are available. Following Holmström (1982/99), we first consider a stationary variant of our model. We then consider a two-period version, which allows us shed some light on the issue of switching between being a specialist and a generalist.

4 Stationary model

We follow Holmström (1982/99) and assume that ability fluctuates over the agent's experience, according to the process

$$\eta_{it+1} = \eta_{it} + \psi_t, \tag{15}$$

where the innovations ψ_t are independently and normally distributed with mean 0 and precision h.[12]

Note that (15) implies that ability fluctuates following performance of a task rather than evolving over time. In Holmström (1982/99), where the agent undertakes an identical task in each period, this distinction is irrelevant. Here, primarily to make the problem more tractable, we assume that innovations in skills arise in the course of performing tasks; however, this may be a realistic assumption, as skills improve (and bad habits or set ways of thinking develop) as a result of experience with tasks rather than simply through the passage of time.

Given this assumption, eventually, precision will no longer diverge with experience. Instead, if we introduce

$$\mu^* := 1 + \frac{1}{2h} - \sqrt{\frac{1}{4h^2} + \frac{1}{h}}, \tag{16}$$

then it is standard to show that the precision h_{it} of the belief about ability converges to

$$h^* = \frac{\mu^*}{1-\mu^*}. \tag{17}$$

Note that $\mu^* < 1$ and that it is increasing in h and, thus, so is h^*. This is intuitive: If there is less noise in innovations (that is, h is high), then beliefs about ability are more precise.

We focus on the stationary distribution and, following derivations similar to Theorems 1 and 2, we obtain the following characterization.

Theorem 3 *In the stationary case with innovation of precision h, and initial beliefs given by $m_A = m_B = m$, the values of specializing and acting as a generalist, evaluated at the arrival of a task, are, respectively,*

$$V_S = \frac{1}{\delta}\frac{\beta_S}{1-\beta_S}(m + a_S^* - c(a_S^*)), \text{ and } V_G = \frac{1}{\delta}\frac{2\beta_G}{1-2\beta_G}(m + a_G^* - c(a_G^*)), \tag{18}$$

[12] Note that supposing that innovations for the A and B tasks have the same distribution ensures that, in the stationary state, the precisions for the two abilities are identical. This simplifies notation but can easily be generalized.

where $c'(a_i^*) = \frac{\beta_i(1-\mu^*)}{1-\beta_i\mu^*}$, $i = S, G$, $\beta_S := \frac{\delta\pi}{r+\pi} < 1$, $\beta_G := \frac{\delta\pi}{r+2\pi} < 1/2$, *and* μ^* *is defined in (16).*

Note that $\beta_G < \beta_S < 1$, so that efforts for a generalist are lower than for a specialist, as in Section 3.3, and both are inefficiently low. Note that effort and value are increasing in β_i, and effort is decreasing in μ^*. Comparative statics for the specialist and the generalist are, then, immediate and similar to the non-stationary case.

It is immediate on inspecting β_S and β_G that values and efforts depend on π and r only through $\frac{r}{\pi}$. This suggests that all comparative statics with respect to π are exactly opposite the comparative statics with respect to r.

Corollary 2 *Effort and lifetime value, for both a specialist and a generalist, are increasing in* π *and decreasing in* h, l *and* r.

We now compare the values of acting as a specialist or as a generalist (V_S and V_G). Theorem 3 provides a full closed-form characterization, and so for any combination of parameters and cost function, this is easily done.

First, note that the components that are independent of effort are easily compared—in particular, when effort is prohibitively costly, then $\frac{V_S}{V_G} = \frac{1}{2}(1 + \frac{1}{\frac{r}{\pi(1-\delta)}+1})$. Trivially, then, the ability component of the value function for specializing is always less than the ability component of the value function for acting as a generalist (as we might expect). However, this difference is smaller the smaller is r or δ and the larger is π—that is, the more the agent discounts time, either directly or through delays from engaging in tasks. This is entirely consistent with our findings for the non-stationary model in Section 3.3.

Note, also, that increasing m favors the generalist: If the expected ability is higher, the cost of idleness increases. Logically, being a generalist is, then, more attractive. To put it differently, agents of higher expected ability should be expected to choose positions as generalists.

To consider the comparative statics of parameters on the component of the value function that depends on effort, we consider the case that $m = 0$ and, for tractability, we restrict attention to quadratic cost functions for efforts. We obtain the following comparative statics.

Proposition 1 *In the stationary case with innovation of precision* h, *initial beliefs given by* $m_A = m_B = 0$, *and the cost of effort given by* $c(a) = c \cdot a^2$, *the comparative statics are as follows:*

1. *Increasing the persistence of type, h, (or, equivalently, the precision on ability, μ^*, in the stationary distribution) can change the agent's preferred option from acting as a generalist to acting as a specialist, but not vice versa.*

2. *Increasing r (decreasing π)—that is, making the agent more impatient—can change the agent's preferred option from acting as a specialist to acting as a generalist, but not vice versa.*

3. *Increasing δ, or, equivalently, reducing the time that it takes to complete a task, l, can change the preferred option from acting as a specialist to acting as a generalist, but not vice versa.*

These results have clear intuitions. First, increasing the precision on ability will lead to lower efforts for both a specialist and a generalist. Note, however, that a specialist always exerts more effort than a generalist and, so, is exerting effort where the marginal cost of effort is relatively high and the cost of effort function is relatively steep. Thus, reducing the benefit of exerting effort (but increasing precision) has relatively little effect on the effort exerted.

Second, with regard to the agent's impatience, there is both a direct effect—a generalist gets more frequent payments and has less of a delay between tasks—and an indirect effect—less patience reduces efforts, in a way that might benefit a specialist, as in the paragraph above. Here, the direct effect always dominates.

Finally, reducing the time it takes to complete a task has several effects. One effect is to reduce the delay before the agent can engage in a new task, and, so, similar to the effects of impatience, through this channel, specializing can be more or less valuable. However, there is a further effect. As the time to complete a task diminishes, undertaking one task is less likely to preclude undertaking another task. When a task can be accomplished instantaneously ($\delta = 1$), taking on B tasks, as a generalist, does not interrupt the rate at which the agent receives A tasks at all, and so acting as a generalist must be the preferred option. Proposition 1 suggests that this effect—that increasing δ involves less sacrifice of A jobs when acting as a specialist—is dominant at the point where acting as a generalist is preferred.

5 Two-period model

In a two-period analog of this setup, analysis is somewhat simpler since there is no effort in the second (final) period, and, thus, it is only the first-period effort that must be considered. We must adapt the model a little to allow for the discrete time required by two periods. To present a two-period analog of the model, we imagine each period as a length of time in which the agent might engage in many jobs, but in which realizations and updating occur only at the end of a period.

To capture the cost of idle time in specializing, we suppose that when an agent specializes, he spends only a fraction λ of the period working (and scale the costs and the information that is generated to be consistent with this interpretation) but spends the rest of the period idle. We benchmark this against a generalist who is fully occupied—spending half the length of the period in each task—and again scale the costs and information generated to correspond with this interpretation.

Formally, we suppose that if an agent undertakes task i as a specialist, then the output produced in each period is given by

$$y_S = \lambda(\eta_i + a_i + \varepsilon_i), \tag{19}$$

where a_i is the effort that the agent undertakes, ε_i is a noise term that is normally distributed with mean 0 and precision λ, and λ is a parameter such that $1 \geq \lambda \geq \frac{1}{2}$. Instead, if the agent undertakes task i as a generalist, then the output produced in that task is given by

$$y_G = \frac{1}{2}(\eta_i + a_i + \varepsilon_i), \tag{20}$$

where the precision on ε_i in this case is $\frac{1}{2}$.[13]

Consistent with the interpretation of a period as a length of time and a reduced-form for undertaking several tasks, we assume that the cost of exerting effort a as a specialist is $\lambda c(a)$, and the cost of exerting effort a_A in task A and a_B in task B as a generalist is $\frac{c(a_A)+c(a_B)}{2}$. For the purpose of exposition, we further simplify by supposing that $c(a) = c{\cdot}a$, where $1 > c > 0$ and $a \in [0, 1]$, so that, in effect, the effort decision is discrete.

Note that, trivially, neither a specialist nor a generalist exerts effort in the second

[13] The different precisions on the error terms for the generalist and specialist are chosen to be consistent with the interpretation of spending a fraction λ rather than $\frac{1}{2}$ of the period engaged in task A, and so generating correspondingly more information.

period.

Finally, the analog to time discounting in the two-period model is to suppose that the agent values revenues in the second period at W times the first. In principle, we could allow for $W > 1$ to reflect the interpretation of the second period as a summary of many future periods; however, for ease of comparison with the stationary case, where effort is inefficiently low, and to reduce the number of cases considered, it is convenient to restrict that $W \leq 1$.

5.1 Comparing Specializing and Generalizing

Without loss of generality, we focus on an agent who specializes in task A and a generalist. Analogously to (4), we can write down how posteriors evolve given equilibrium expectations of effort, a_s^* and a_G^*, for the case where the agent acts as a specialist and as a generalist, respectively,[14] as

$$\widetilde{m}_S = \frac{hm + \lambda(\frac{y}{\lambda} - a_S^*)}{h + \lambda}, \text{ and } \widetilde{m}_{i,G} = \frac{hm + \frac{1}{2}(2y_i - a_G^*)}{h + \frac{1}{2}}. \tag{21}$$

In the second period, a specialist earns $\lambda \widetilde{m}_S$ and a generalist $\frac{1}{2}\widetilde{m}_{A,G} + \frac{1}{2}\widetilde{m}_{B,G}$. Since payoffs in the second period are weighted at W times the costs incurred in the first period, it follows that in equilibrium, efforts for a specialist and for a generalist a_S^* and a_G^*, are determined by

$$a_S^* = \left\{ \begin{array}{ll} 1 & \text{if } W\frac{\lambda}{h+\lambda} \geq c \\ 0 & \text{otherwise} \end{array} \right\} \text{ and } a_S^* = \left\{ \begin{array}{ll} 1 & \text{if } \frac{W}{2h+1} \geq c \\ 0 & \text{otherwise} \end{array} \right\}. \tag{22}$$

The following result is immediate on noting that in equilibrium, consumer expectations of efforts are accurate, and so $\mathbb{E}[\widetilde{m}_S] = m$ and $\mathbb{E}[\widetilde{m}_{A,G}] = \mathbb{E}[\widetilde{m}_{B,G}] = m$.

Proposition 2 *The equilibrium lifetime values of acting as a specialist and as a generalist, respectively, are given by*

$$\begin{aligned} S &= \lambda\left[m(1+W) + a_S^*(1-c)\right], \text{ and} \\ G &= m(1+W) + a_G^*(1-c), \end{aligned}$$

where a_S^* and a_G^* are determined as in (22).

[14] Note that, by symmetry, efforts in the different tasks that a generalist takes are identical.

Corollary 3 *The expected lifetime earnings of a specialist, S, increase in the fraction of time engaged in the task (λ), the initial reputation (m), and the value of the second period (W) and are non-decreasing in the uncertainty about the agent's type (h^{-1}). The expected lifetime earnings of a generalist increase in the initial reputation (m), and the value of the second period (W) and are non-decreasing in the uncertainty about the agent's type (h^{-1}).*

As is the case in Section 3, where $m_A = m_B = m$, and in Section 4, absent any effort concerns, the agent prefers to act as a generalist since he values both tasks equally, and, as a generalist, spends less time idle. However, when effort is a substantive concern, specialization may be preferred since it leads to greater incentives for effort. These incentives arise from the agent trying to influence market beliefs about his ability in that task. This has a larger payoff when the agent specializes since he will spend a greater proportion of his time engaged in task A as a specialist than as a generalist ($\lambda > \frac{1}{2}$). More generally, we show the following:

Proposition 3 *The agent is more likely to choose to act as a specialist, the lower is m and the higher is λ.*

An increase in h will make specialization more likely only if it leads a specialist, but not a generalist, to exert effort—that is, if $W\frac{\lambda}{h+\lambda}$ is just less than c, and an increase in h benefits generalizing when $\frac{W}{2h+1}$ is just under c. Since $\frac{1}{2h+1} > \frac{\lambda}{h+\lambda}$, increasing h first benefits specializing (albeit only for one value of h) over generalizing but then harms specialization (again, only at one value of h). Similar concerns arise with respect to W, but, in addition, there is the direct effect $\frac{\partial(G-S)}{\partial W} = (1-\lambda)m$ that acts in favor of generalizing. The intuition for this effect is that putting more weight on the future (when there are no efforts) raises the importance of the contribution of ability to lifetime income over the contribution arising from efforts, and, as for the comparative statics in m, this is a force in favor of generalizing.

5.2 Switching from Specialization to Generalization

In this relatively simple environment, we can examine the possibility that the agent is able to switch careers between the two periods. In fact, there are many possibilities that could be examined. For conciseness and to limit the number of cases, we suppose that an agent who starts out as a specialist can switch to being a generalist in the second period, but doing so comes at a cost K.[15] In effect, we examined above the extreme case that $K = \infty$. Here, we characterize an optimal switching cost.

The benefit of a lower switching cost is essentially to provide some partial insurance should the agent turn out to have very low ability in task A, making the restriction to being engaged only in that task relatively costly. In general, we expect a trade-off, in that a higher switching cost should lead to greater effort (or in this discrete effort example, a higher likelihood of effort) but also should increase the probability of ex-post allocative inefficiency.

Characterizing the optimal switching cost is complicated since the agent's effort has an impact only on his revenue for future A tasks. Thus, in order to determine whether or not to exert effort, he must calculate the likelihood that he will stay a specialist in the next period or switch to being a generalist. We obtain the following result:

Proposition 4 *The minimal (and optimal) switching cost that ensures that the agent exerts effort is given by*

$$K = m(1-\lambda) - \frac{2\lambda - 1}{2}\sqrt{\frac{2\lambda}{h(h+\lambda)}}\,\mathrm{erf}^{-1}\left(\frac{2\lambda+1}{2\lambda-1} - \frac{4c}{2\lambda-1}\frac{h+\lambda}{\lambda W}\sqrt{\frac{h+\lambda}{hl}}\right),$$

where erf^{-1} *denotes the inverse error function. At this switching cost, the probability that an agent switches from acting as a specialist to acting as a generalist in the second period is*

$$\frac{2\lambda}{2\lambda-1} - \frac{c}{W}\frac{2}{2\lambda-1}\frac{h+\lambda}{\lambda}\sqrt{\frac{h+\lambda}{h\lambda}}$$

as long as this takes a value in $(0,1)$, *which requires*

$$\frac{\lambda^2}{h+\lambda}\sqrt{\frac{h\lambda}{h+\lambda}} > \frac{c}{W} > \frac{1}{2}\frac{\lambda}{h+\lambda}\sqrt{\frac{h\lambda}{h+\lambda}}.$$

The following comparative statics are, then, immediate.

Corollary 4 *The optimal switching cost that ensures that a specialist exerts effort is increasing in the initial reputation m and in the cost of effort c, and it is decreasing in the importance of the second period W. The probability of switching, when the switching*

[15] In principle, one might also consider switching from specializing in one task to specializing in the other, or switching from being a generalist to being a specialist in one task or the other.

cost is optimally chosen, is independent of m, but is increasing in W and decreasing in c. Comparative statics with respect to other parameters—the length of time engaged in task A when specialized λ, and the precision of beliefs on abilities h—can go in either direction.

These results can be explained as follows. First, when initial reputations are high, an agent will be more likely to switch in order to spend more time in (relatively) lucrative tasks. Anticipating that he is more likely to generalize, the agent would have less incentive to exert effort. Note that since the initial reputation, m, has no direct effect on incentives for effort, the optimal switching probability that induces effort is independent of m. It follows that the optimal switching cost must rise to make switching less attractive and force effort to achieve this same switching probability. If effort is more expensive, or the future is less valued, then the incentives to exert effort are lower and must be boosted to ensure that the agent undertakes effort—such incentives are boosted by ensuring that the agent spends more time in the specialized task in the second period (which, in turn, requires that the cost of switching to act as a generalist increases). Finally, the remaining two parameters, λ and h, have several effects, and so the overall comparative statics with respect to these parameters are not straightforward.

A higher value of h, intuitively suggests that there is more precision on ability and, so, the output has relatively less impact on shifting priors. This implies that incentives for effort are weaker and, thus, should be boosted by reducing the probability of switching, but note that for a fixed threshold, higher precision in any case might reduce the probability of switching (depending on the value of λ). In the extreme case that the cost of switching is close to 0, there is some chance that the agent with $h \approx 0$ does not switch since the market may observe an extremely high first-period output, suggesting that ability in task A is much higher than anticipated (and higher than expected in task B). Instead, when $h \approx \infty$, an extremely high output is attributed to noise, and the agent is likely to switch in any case.

The comparative statics with respect to λ incorporate subtleties similar to those associated with h, but there is an additional effect: Increasing λ ensures that the when the agent is specialized in the second period, he spends more time on (and gets more payoffs from) the task, suggesting that there is more incentive to exert effort. However, overall, the comparative statics can go in either direction.

6 Conclusions

In this paper, we have adapted a standard career-concerns model to consider an agent's choice between working as a specialist or as a generalist. There is a purely technological trade-off that arises from the environment. An agent may prefer to spend more time in a more-lucrative task and, thus, not congest his opportunities by undertaking less-valuable tasks, leading to specialization; this concern is traded off against the fact that a specialist spends more time idle. More interesting, and more complex, is that specialization and acting as a generalist have different implications for the implicit incentives for effort that arise. We have illustrated that such incentive concerns may overwhelm the purely technological aspect. So, for example, even if an agent starts with identical reputations in both tasks, he might choose to specialize and, in effect, to commit to exert effort.

The model, therefore, makes several predictions that can, in principle, be tested. In common with models of specialization—discussed in the introduction and building on Smith (1776)—specialization is limited by the extent of the market: if there is sufficient volume of work such that a specialist is fully occupied ($\pi \to \infty$ in the infinite horizon model, or $\lambda \to 1$ in the two period model), the agent prefers acting as a specialist to acting as a generalist. Thus, there is a real trade-off to be considered only in thinner markets, in which work-flow is a substantive concern. A new prediction, however, concerns markets in which career concerns or reputation play a significant role in providing incentives—as in consulting, legal, medical and other professional services that motivate this study. In such cases, the more important is non-contractible effort, the more likely is an agent to specialize. This is consistent with a noted trend for greater complexity and increased specialization, and complements the explanation that more ex-ante training or knowledge is required (Jones, 2009)

Since we provide characterizations of the value of acting as a specialist and as a generalist, as well as a closed-form solution for efforts, it is easy to compare them for a given choice of parameters; however, general analytical comparative statics are complicated, in so far as parameters have effects in the future, and current actions are influenced by the sum of all future effects, which arrive at different rates for a specialist and a generalist. These complexities can be somewhat mitigated by simplifying the framework and the future values, by considering either a stationary model or a two-period model. We do this in Sections 4 and 5 and are able to obtain further results. In particular, in the two-period model, we analyze switching costs that effectively allow for a partial commitment to stay in the task. These give somewhat stronger incentives than would arise in the absence of any such commitment, while protecting the agent from severe misallocation if he finds himself incapable of performing the task. We demonstrate that while some parameters (initial reputation, cost of effort and discount rate between periods) allow for simple comparative statics, the other parameters (time engaged in the task and precision on ability) have subtle effects since they affect the agent's expectation that he will choose to switch careers. Such effects would be present, though the analysis would be much more analytically involved in allowing for switching costs in the infinite-horizon model; this, however, could be analyzed or approximated numerically and used to consider career dynamics.

Our analysis takes the definition of being a generalist or a specialist as essentially exogenous: Either our agent accepts both types of tasks or only one, and in (most of) the analysis, we suppose that the agent is committed indefinitely. In practice, however, there is considerable diversity in the scope of a professional's activity. Lawyers, doctors, and other professionals have varying breadths of competencies, and they can change the nature of their work, though often at some cost of retraining and accreditation. We believe that a finer analysis of the determinants of how specialized one should be could provide further insight.

References

[1] Abramovitz, M. and I. A. Stegun (1965). *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables*, Dover Publications, Inc., New York.

[2] Acemoglu, D. and D. Autor (forthcoming) "Skills, Tasks and Technologies: Implications for Employment and Earnings," in Handbook of Labor Economics Volume 4, Orley Ashenfelter and David E. Card (eds.), Amsterdam: Elsevier, forthcoming.

[3] Becker, G. S. and K. M. Murphy (1992). "The division of labor, coordination costs, and knowledge," *Quarterly Journal of Economics*, 107(4), 1137–60.

[4] Burguet, R., Caminal, R. and C. Matutes (2002). "Golden cages for showy birds: Optimal switching costs in labor contracts," *European Economic Review*, 46, 1153–1185.

[5] Casas-Arce, P. (2010). "Career Tournaments," *Journal of Economics and Management Strategy*, 19, 667–698.

[6] Chatain, O. and P. Zemsky (2007). "The Horizontal Scope of the Firm: Organizational Tradeoffs vs. Buyer-Supplier Relationships," *Management Science*, 53(4), 550–565

[7] DeGroot, M. H. (1970). *Optimal Statistical Decisions*, McGraw-Hill, Inc.

[8] Dewatripont, M., I. Jewitt and J. Tirole (1999). "The Economics of Career Concerns, Part I: Comparing Information Structures," *Review of Economic Studies*, 66.1 Special Issue on Contracts, 183–198.

[9] Dewatripont, M., I. Jewitt and J. Tirole (1999). "The Economics of Career Concerns, Part II: Application to Missions and Accountability of Government Agencies" *Review of Economic Studies*, 66.1 Special Issue on Contracts, 199–217.

[10] Fama, E. (1980). "Agency Problems and the Theory of the Firm," *Journal of Political Economy*, 88, 288–307.

[11] Friedman, M. and S. Kuznets (1945). *Income from Independent Professional Practice*, National Bureau of Economic Research, New York.

[12] Gagné, R. and P. T. Léger (2005). "Determinants of physicians' decisions to specialize," *Health Economics*, 14, 721–735.

[13] Garicano, L. and T. Hubbard (2009). "Specialization, Firms, and Markets: The Division of Labor Within and Between Law Firms," *Journal of Law, Economics, and Organization*, 25(2), 339-371.

[14] Gibbons, R., Waldman, M. (1999). "Careers in organizations: Theory and evidence," In: Ashenfelter, C., Card, D. (Eds.), *Handbook of Labor Economics*, Vol. 3B. North-Holland, Amsterdam, 2373–2437.

[15] Hart, O. (1983). "Optimal labour contracts under asymmetric information: An introduction," *Review of Economic Studies*, 50, 3–35.

[16] Heyman, D.P. and M.J. Sobel (1982). *Stochastic Models in Operations Research, Vol. I*, McGraw-Hill, New York.

[17] Holmström, B. (1999). "Managerial Incentive Problems—A Dynamic Perspective," *Review of Economic Studies*, 66.1 Special Issue on Contracts, 169–182.

[18] Holmström, B. and P. Milgrom (1991) "Multi-task Principal-Agent Analyses: Incentive Contracts, Asset Ownership and Job Design," *Journal of Law, Economics and Organization*, 7, 24-52.

[19] Jensen, A. (1953). "Markoff Chains as an Aid in the Study of Markoff Processes," *Skand. Aktuarietidskr.*, 36, 87–91.

[20] Jones, B. F. (2009). "The Burden of Knowledge and the "Death of the Renaissance Man": Is Innovation Getting Harder?" *Review of Economic Studies*, 76(1), 283–317.

[21] Kleiner, M. M. (2006). *Licensing Occupations: Ensuring Quality or Restricting Competition?* W. E. Upjohn Institute for Employment Research, Kalamazoo, Michigan

[22] Kovrijnykh, A. (2007). "Career Uncertainty and Dynamic Incentives," mimeo, University of Chicago.

[23] Law Society (2004). "Solicitors' Separate Business Code 1994," http://www.lawsociety.org.uk/documents/downloads/Profethics_SepBusiness.pdf

[24] Levin, J. and S. Tadelis (2005). "Profit Sharing and the Role of Professional Partnerships," *Quarterly Journal of Economics*, vol. 120(1), 131–171.

[25] Malcomson, J. (1999). "Individual employment contracts," In: Ashenfelter, C., Card, D. (Eds.), *Handbook of Labor Economics*, Vol. 3B. North-Holland, Amsterdam, 2291–2372.

[26] Martinez, L. (2009). "Reputation, Career Concerns, and Job Assignments," *B. E. Journal of Theoretical Economics*, 9(1): Art. 15.

[27] Nicholson, S. (2002). "Physician Specialty Choice Under Uncertainty," *Journal of Labor Economics*, 20, 816–47.

[28] Rosen, S. (1978). "Substitution and Division of Labour," *Economica,* 45, 235–250.

[29] Rosen, S. (1983). "Specialization and Human Capital," *Journal of Labor Economics*, 1, 43–49.

[30] Sloan, F. A. (1970). "Lifetime Earnings and Physicians' Choice of Specialty," *Industrial and Labor Relations Review*, 24(1), 47–56.

[31] Smith, A. (1776) *An Inquiry into the Nature and Causes of the Wealth of Nations*, Oxford University Press, Oxford.

A Omitted Proofs

Proof. Theorem 1.

We conjecture that in equilibrium, when $m' = m$, the value function $S(m_A, h_A)$ takes the form

$$V_S(m, h_A) = k_S m_A + g(h_A), \tag{23}$$

for some constant k_S. We verify below that this conjecture holds.

We begin by characterizing the agent's equilibrium behavior. In particular, it is useful to use Equation (6) to write

$$S(m_A, h_A) = \frac{\pi}{\pi + r}\mathbb{E}\left[\max_{\{a\}} \sum\nolimits_{t=0}^{\infty} (\frac{\pi\delta}{\pi + r})^t \left[\widetilde{m}_{A,t} + a_t - c(a_t)\right]\right]. \tag{24}$$

From Equation (4), we have

$$\widetilde{m}_{At+1} = \frac{h_A m_A + \sum_{j=0}^{t} z_{Aj}}{h_A + t + 1}. \tag{25}$$

Note that the maximization problem is concave, since effort affects future payoffs linearly. Taking first-order conditions in (24) (*note that this does not yet use our conjecture*), we get that

$$c'(a_t) = \sum_{s=t}^{\infty} \frac{(\frac{\pi\delta}{\pi+r})^{s+1-t}}{h_A + s + 1} = \frac{\pi\delta}{\pi + r} \frac{F(1, h_A + t + 1, h_A + t + 2; \frac{\pi\delta}{\pi+r})}{h_A + 1}, \tag{26}$$

where $F(a, b, c; z)$ denotes the Gauss hypergeometric function (see Abramovitz and Stegun, 1965, p. 556). Note, also, that as anticipated in the text, it depends only on h_A and not on m_A. Using standard properties of the hypergeometric functions, we can also write this as

$$c'(a_t) = \int_0^1 \frac{u^{h_A + t}}{\frac{\pi+r}{\pi\delta} - u} du. \tag{27}$$

Note that a_t does not depend on m_t, but only on h_A, so that we can write $a_t = a(h_A + t)$. Note, further, that, since $\mathbb{E}[m_{At+1}] = m_A$ on the equilibrium path, we obtain

$$S(m_A, h_A) = \frac{\pi}{\pi + r} \sum_{t=0}^{\infty} \left(\frac{\pi\delta}{\pi + r}\right)^t m_A + g(h_A) = k_S m_A + g(h_A), \tag{28}$$

with $k_S := \frac{\pi}{\pi(1-\delta)+r}$, validating our earlier claim about the structure of the value function. Finally, note that

$$g(H) := \frac{\pi}{\pi + r} \sum\nolimits_{t=0}^{\infty} \left(\frac{\pi\delta}{\pi + r}\right)^t [a_t - c(a_t)] \tag{29}$$

is finite since $a_t - c(a_t)$ is bounded above. ■

Proof. Theorem 3.

Skipping analogous steps to the characterization of the non-stationary case, we obtain that when the market expectation of effort is denoted a^*, the value of specialization is

$$V_s(m) = \frac{\pi}{r+\pi} \max_a \left\{ m + a^* - c(a) + \delta \mathbb{E}[V_S(\widetilde{m})] \right\}, \tag{30}$$

where

$$\widetilde{m} = \frac{h^* m + (m + \varepsilon + \psi + a - a^*)}{h^* + 1} \tag{31}$$

is the realized posterior belief. Again, the solution is affine in m with solution

$$\pi \frac{m + a_S^* - c(a_S^*)}{\pi(1-\delta) + r}, \tag{32}$$

where a_S^* denotes optimal effort and, as the notation suggests, this effort is independent of m. The optimal effort must solve

$$\begin{aligned} c'(a_S^*) &= \frac{1}{1+h^*} \delta \frac{\pi}{r+\pi} \sum_{t=0}^{\infty} (\delta \frac{\pi}{r+\pi} \frac{1}{1+h^*})^t \\ &= (1-\mu^*) \delta \frac{\pi}{r+\pi} \sum_{t=0}^{\infty} (\delta \frac{\pi}{r+\pi}(1-\mu^*))^t = \frac{\delta \pi (1-\mu^*)}{r + \pi(1-\delta\mu^*)}. \end{aligned} \tag{33}$$

Similarly, we can write

$$V_G(m_A, m_B) = \frac{\pi(m_A + a_G^* - c(a_G^*)) + \pi(m_B + a_G^* - c(a_G^*))}{r + 2\pi(1-\delta)}, \tag{34}$$

where

$$c'(a_G^*) = \frac{\delta \pi (1-\mu^*)}{r + 2\pi(1-\delta\mu^*)}. \tag{35}$$

■

Proof. Proposition 1. The strategy of the proof is the same for all three parameters of interest. The difference in payoff between acting as a specialist vs. a generalist is proportional to

$$\delta^2(1-\mu) \left(\frac{2(1+r) - \delta(1+\mu)}{(1-\delta+r)(1-\delta\mu+r)^2} - \frac{2(2(2+r) - \delta(1+\mu))}{(2(1-\delta)+r)(2-\delta\mu+r)^2} \right), \tag{36}$$

where $m = 0$ and π is normalized to 1. We aim to identify the sign of the derivative of this expression, evaluated at a point at which it is equal to 0. It is convenient to introduce the variables $x := 1 - \delta\mu + r$ and $y := \delta(1-\mu) - x$, which allows us to solve for the roots of the equation that sets (36) to 0, by solving for μ in terms of x and y.

1. Change with respect to h: Upon differentiation and substitution, we get

$$\frac{\partial(V^S - V^G)}{\partial \mu}|_{V^S=V^G} \propto \frac{(y(2-y)+x(2y-1))(x(1+x)^2+(x^3+x^2-3x-1)y+(1+2x-x^2)y^2)}{x^3(1+x)^3(x-y)(2+x-y)y},$$

and both numerator and denominator are positive: As h increases, we can only switch from acting as a generalist to acting as a specialist.

2. Change with respect to r: Upon differentiation and substitution, we get

$$\frac{\partial(V^S - V^G)}{\partial r}|_{V^S=V^G} \propto (x+y)\frac{M_1 M_2}{4x^4(1+x)^3(x-y)(2+x-y)y^2},$$

where

$$M_1 := x(1+x)^2 + (x^3+x^2-3x-1)y - (x^2-2x-1)y^2,$$

and

$$M_2 := x^2(x^3+3x^2+x-1) + 2x(1+x+x^2-x^3)y + (1+x)(x^2+6x-1)y^2 - 4xy^2.$$

The denominator is positive because it equals $(2(1+r)-\delta(1+\mu))(2(2+r)-\delta(1+\mu))$. It is not difficult to see that $M_1 \leq 0$ and $M_2 \geq 0$, so that the derivative is negative: As r increases, we can switch only from acting as a specialist to acting as a generalist.

3. Change with respect to δ: Upon differentiation, substitution, and simplification,

$$\frac{\partial(V^S - V^G)}{\partial r}|_{V^S=V^G} \propto -\frac{x^3(1+x)^3 - x^2(1+x)\alpha_1 y + x\alpha_2 y^2 - \alpha_3 y^3 - \alpha_4 y^4 + 4x(1+2x)y^5}{x^4(1+x)(2+x-y)y^2(x(1+x)^2+(x^3+x^2-3x-1)y-(x^2-2x-1)y^2)},$$

where

$$\alpha_1 = 5x^3+9x^2+9x+3, \alpha_2 = x^4+14x^3+18x^2+12x+3, \alpha_3 = 7x^4-2x^3-2x^2+4x+1,$$

and $\alpha_4 = 13x^3+23x^2+5x-1$. Both numerator and denominator are negative, but given the negative sign in front of the ratio, the sign of the derivative is negative: As δ increases, we can switch only from acting as a specialist to acting as a generalist.

■

Proof. Proposition 3.

We write

$$\Delta := S - G = -(1-\lambda)m(1+W) + \lambda a_S^*(1-c) - a_G^*(1-c). \tag{37}$$

Trivially, $-(1-\lambda)m(1+W)$ is decreasing in m and increasing in λ. Since a_S^* and a_G^* are independent of m, it follows that Δ is decreasing in m.

Note that a_G^* is independent of λ. Since $\frac{d}{d\lambda}(W\frac{\lambda}{h+\lambda}) = h\frac{W}{(h+\lambda)^2} > 0$, it follows that a_S^* is non-decreasing in λ; overall, therefore, Δ increases in λ. ■

Proof. Proposition 4. An agent will pay a cost K to switch from specializing to acting as a generalist when the market's posterior about his ability in task A is sufficiently low. In particular, he earns $\lambda\widetilde{m}_A$ if he stays as a specialist and $\frac{\widetilde{m}_A+m}{2}-K$ if switching, and so he switches whenever $\frac{\widetilde{m}_A+m}{2}-K>\lambda\widetilde{m}_A$ or, equivalently, if and only if

$$\frac{m-2K}{2\lambda-1}>\widetilde{m}_A=\frac{hm+\lambda(\eta_A+\varepsilon+a-a_S^*)}{h+\lambda}, \tag{38}$$

where a_S^* is the market's expectation of his effort in the first period and a is his actual effort; in equilibrium, of course, these are equal.

We define

$$H:=\frac{(h+\lambda)(m-2K)}{(2\lambda-1)\lambda}-\frac{hm}{\lambda} \tag{39}$$

and note that $\eta_A+\varepsilon$ is a sum of two normal distributions, and so is also normally distributed with mean m and precision $\frac{\lambda h}{h+\lambda}$. The threshold H is the level for the realization of the random variable $\eta_A+\varepsilon$ below which the agent prefers to switch tasks (in equilibrium when effort is correctly anticipated). This allows us to write the agent's value of starting out as a specialist when the cost to switching is given by K as

$$S(K)=\arg\max_a\left\{\lambda(m+a_S^*-ca)+W\left[\begin{array}{c}\lambda\int\limits_{H+a_S^*-a}^{\infty}(\frac{hm+\lambda(x+a-a_S^*)}{h+\lambda})\sqrt{\frac{\lambda h}{h+\lambda}}\phi(\sqrt{\frac{\lambda h}{h+\lambda}}(x-m))dx\\ +\frac{1}{2}\int\limits_{-\infty}^{H+a_S^*-a}(\frac{hm+\lambda(x+a-a_S^*)}{h+\lambda}+m-2K)\sqrt{\frac{\lambda h}{h+\lambda}}\phi(\sqrt{\frac{\lambda h}{h+\lambda}}(x-m))dx\end{array}\right]\right\}, \tag{40}$$

where $\phi(.)$ is the pdf of the standard Gaussian distribution (the cdf is denoted by Φ).

It is clear that by taking the derivative of the above expression with respect to a (and noting that in equilibrium, $a=a_S^*$) that the agent exerts maximal effort if and only if

$$c<W\lambda\frac{\lambda}{h+\lambda}\sqrt{\frac{\lambda h}{h+\lambda}}(1-\Phi\left(\sqrt{\frac{\lambda h}{h+\lambda}}(H-m)\right))+W\frac{1}{2}\frac{\lambda}{h+\lambda}\sqrt{\frac{\lambda h}{h+\lambda}}\Phi\left(\sqrt{\frac{\lambda h}{h+\lambda}}(H-m)\right). \tag{41}$$

It follows that the minimal switching cost K that ensures effort induces an H such that

$$c=W\frac{\lambda}{h+\lambda}\sqrt{\frac{\lambda h}{h+\lambda}}\left[\lambda-(\lambda-\frac{1}{2})\Phi\left(\sqrt{\frac{\lambda h}{h+\lambda}}(H-m)\right)\right]. \tag{42}$$

Trivially, a higher switching cost leads to more misallocation but does not lead to more effort, so that the minimal switching cost is also optimal.

We can use (42) to solve for the K as follows

$$c=W\frac{\lambda}{h+\lambda}\sqrt{\frac{\lambda h}{h+\lambda}}\left[\lambda-(\lambda-\frac{1}{2})\Phi\left(\sqrt{\frac{\lambda h}{h+\lambda}}(\frac{(h+\lambda)(m-2K)}{(2\lambda-1)\lambda}-\frac{hm}{\lambda}-m)\right)\right], \tag{43}$$

$$\Phi\left(\sqrt{\frac{\lambda h}{h+\lambda}}\left(\frac{(h+\lambda)(m-2K)}{(2\lambda-1)\lambda}-\frac{hm}{\lambda}-m\right)\right)=\frac{2\lambda}{2\lambda-1}-\frac{2c}{2\lambda-1}\frac{h+\lambda}{\lambda W}\sqrt{\frac{h+\lambda}{hl}}, \tag{44}$$

$$\begin{aligned}\sqrt{\frac{\lambda h}{h+\lambda}}\left(\frac{(h+\lambda)(m-2K)}{(2\lambda-1)\lambda}-\frac{hm}{\lambda}-m\right) &= \Phi^{-1}\left(\frac{2\lambda}{2\lambda-1}-\frac{2c}{2\lambda-1}\frac{h+\lambda}{\lambda W}\sqrt{\frac{h+\lambda}{hl}}\right) \quad (45)\\ &= \sqrt{2}\,\mathrm{erf}^{-1}\left(\frac{2\lambda+1}{2\lambda-1}-\frac{4c}{2\lambda-1}\frac{h+\lambda}{\lambda W}\sqrt{\frac{h+\lambda}{hl}}\right),\end{aligned}$$

and so

$$K=m(1-\lambda)-\frac{2\lambda-1}{2}\sqrt{\frac{2\lambda}{h(h+\lambda)}}\,\mathrm{erf}^{-1}\left(\frac{2\lambda+1}{2\lambda-1}-\frac{4c}{2\lambda-1}\frac{h+\lambda}{\lambda W}\sqrt{\frac{h+\lambda}{hl}}\right). \tag{46}$$

It is easy to solve for the probability of switching, $\Phi\left(\sqrt{\frac{\lambda h}{h+\lambda}}(H-m)\right)$, directly from (42) and to verify when this takes values in $[0,1]$. ■

B Correlated Abilities

In the central model in Section 2, we suppose that abilities are independent. Here, we show that correlation can be, relatively easily incorporated, and values and efforts for a specialized and generalist agent can be characterized as in Section 3.

We adapt the model of Section 2, and allow for correlation by supposing that the precision matrix is given by

$$H_0 = \begin{pmatrix} h_{0A} & \tau \\ \tau & h_{0B} \end{pmatrix}. \tag{47}$$

Note here that τ must be negative.

It is convenient to write the observation of output and the noise terms as vectors, where information is generated only on the task undertaken. Thus, if the agent engages in task A, we write the observed output as

$$y_t := \begin{pmatrix} \eta_A + a_{A,t} \\ \eta_B \end{pmatrix} + \varepsilon_t^A, \tag{48}$$

where $\varepsilon_t^A = \left(\varepsilon_{A,t}, \varepsilon_{B,t}^N\right)^\top$ is normally distributed with mean $(0,0)^\top$ and precision matrix

$$N_A := \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}. \tag{49}$$

Note, in particular, that since the precision on ε_B^N is 0, the observation $\eta_B + \varepsilon_B^N$ provides no information. We can, similarly, define ε_t^B and N_B for the case in which the agent takes on task B.

We consider, first, the case of specialization, and define the innovation

$$\begin{pmatrix} z_A \\ z_B \end{pmatrix} := y - \begin{pmatrix} a_A^*(m, H) \\ 0 \end{pmatrix}; \tag{50}$$

then, again following DeGroot (1970), it holds that

$$\begin{pmatrix} \widetilde{m}_A \\ \widetilde{m}_B \end{pmatrix} = (H + N_A)^{-1} \left(H \begin{pmatrix} m_A \\ m_B \end{pmatrix} + N_A \begin{pmatrix} z_A \\ z_B \end{pmatrix} \right). \tag{51}$$

In particular, after t tasks have been undertaken, the posterior mean on task A is

$$\widetilde{m}_{A,t} = \frac{(h_A + t - \lambda^2/h_B)m_{A,t-1} + z_{A,t-1}}{h_A + t - \lambda^2/h_B + 1}, \tag{52}$$

where $m_{A,t-1}$ is the mean immediately before the last task was undertaken, and $z_{A,t}$ is the last innovation. Ability remains a martingale and the precision matrix, again, evolves deterministically following each new observation. Specifically, the precision matrix H evolves to $H + N_A$, so that, after t tasks, the posterior precision matrix is given by

$$\begin{pmatrix} h_A + t & \tau \\ \tau & h_B \end{pmatrix}. \tag{53}$$

As in Section 3, we can write

$$V_S(m, m', H) = \frac{\pi}{\pi + r} \max_a \{m_A + a_A^*(m, H) - c(a) + \delta \mathbb{E}[V(\widetilde{m}, \widetilde{m}', H + N_A)]\}, \tag{54}$$

and, in equilibrium,

$$a^*(m, H) = \arg\max_a \delta \mathbb{E}[V(\widetilde{m}, \widetilde{m}', H + N_A)] - c(a). \tag{55}$$

Following similar steps to Theorem 1, we can obtain the following result.

Theorem 4 *The value function in the case of specialization is equal to*

$$V_S(m, H) = \frac{\pi}{\pi(1-\delta) + r} m_A + \frac{\pi}{r + \pi} \sum_{t=0}^{\infty} \left(\frac{\pi\delta}{r + \pi}\right)^t [a_t - c(a_t)], \tag{56}$$

where $a_t(H)$ is the unique solution to

$$c'(a_t) = \sum_{s=t}^{\infty} \frac{(\frac{\pi\delta}{\pi+r})^{s+1-t}}{h_A + s + 1 - \frac{\tau^2}{h_B}}. \tag{57}$$

Comparative statics, unsurprisingly, are similar to those in the case of independence.

Corollary 5 *Effort a_t is decreasing in h_A, h_B, t, r and l (through δ), and increasing in τ^2 and π. The first component of the value function, namely*

$$\frac{\pi m_A}{\pi(1-\delta) + r}, \tag{58}$$

is increasing in π, m_A and decreasing in both l (through δ) and r.

Note that the only new comparative static here relates to the extent of information about ability in task A that arises from knowing about task B which is the case if h_B is high or if τ^2 is low (that is there is low correlation); this is analogous to the effect of more-precise information on ability in task A in 3.1.

We now consider an agent who undertakes whichever task first presents itself. We start by characterizing the extent to which a specialized agent is idle or employed on some task. Following the same steps as before, we can write an optimality equation for a broad agent who is idle and in state $(m_A, m_B, m'_A, m'_B, H)$ as

$$\begin{aligned} & V_G(m_A, m_B, m'_A, m'_B, H) \\ = \; & \frac{\pi}{r + 2\pi} \max_{a_A} \left\{m_A + a_A^*(m, H) - c(a_A) + \delta \mathbb{E}[V_G(\widetilde{m}^A, \widetilde{m}^{A\prime}, H + N_A)]\right\} \\ & + \frac{\pi}{r + 2\pi} \max_{a_B} \left\{m_B + a_B^*(m, H) - c(a_B) + \delta \mathbb{E}[V_G(\widetilde{m}^B, \widetilde{m}^{B\prime}, H + N_B)]\right\}. \end{aligned} \tag{59}$$

Following (51) and analogous to (52), we can write the posteriors, following an A task, as

$$\widetilde{m}_A^A = \frac{(h_A - \tau^2/h_B)m_A + z_A}{h_A - \tau^2/h_B +}, \text{ and} \tag{60}$$

$$\widetilde{m}_B^A = \frac{(h_A + 1 - \tau^2/h_B)m_B + \tau m_A/h_B - \tau z_A/h_B}{h_A - \tau^2/h_B + 1}, \tag{61}$$

with similar expressions following a B task. Thus, the ability in each task is once again a martingale. Note that, just as in the specialized case, following a task of type $i = A, B$, the precision matrix evolves deterministically. Specifically, the precision matrix H evolves to $H + N_i$. From Equations (60) and (61) we have:

$$m_A^A(t_A + 1, t_B) = \frac{(h_A + t_A - \frac{\tau^2}{h_B + t_B})m_A^A(t_A, t_B) + z_A}{h_A + t_A + 1 - \tau^2/(h_B + t_B)}, \text{ and} \tag{62}$$

$$m_B^A(t_A + 1, t_B) = \frac{(h_A + t_A + 1 - \frac{\tau^2}{h_B + t_B})m_B^A(t_A, t_B) + \frac{\tau}{h_B + t_B}m_A^A(t_A + 1, t_B) - \frac{\tau z_A}{h_B + t_B}}{h_A + t_A + 1 - \tau^2/(h_B + t_B)}, \tag{63}$$

where $m_i^A(t_A, t_B)$ denotes the ability in the i task following a sequence that includes t tasks in A and s tasks in B, with the last task being an A task.

Note that in contrast to the analysis of Section 3.2, in which abilities are independent, here, efforts in a task A depend on the current precision not only about ability in task A, but also about the task B. Therefore, we write $a_A(t_A, t_B)$ to denote the optimal effort and, if interior, it is given by the first-order condition, namely

$$c'(a_A(t_A, t_B)) = \sum_{i \geq 0, j \geq 0} \left(\frac{\pi\delta}{r + 2\pi}\right)^{i+j+1} \begin{pmatrix} i+j \\ i \end{pmatrix} \frac{h_B + t_B + j - \tau}{(h_A + i + t_A + 1)(h_B + t_B + j) - \tau^2}. \tag{64}$$

This and similar steps to those taken above, therefore, show the following,

Theorem 5 *The value function in the case of generalizing is equal to*

$$V_G(m, H) = \frac{\pi(m_A + m_B)}{r + 2(1 - \delta)\pi} + f(H), \tag{65}$$

with

$$f(H) = \frac{\pi}{r + 2\pi} \sum_{k=0}^{\infty} \sum_{i=0}^{k} \left(\frac{\pi\delta}{r + 2\pi}\right)^k \left[a_A(k - i, i) - c(a_A(k - i, i)) + a_B(k - i, i) - c(a_B(k - i, i))\right], \tag{66}$$

and $a_A(t_A, t_B)$ *is the unique solution to*

$$c'(a_A(t_A, t_B)) = \sum_{i \geq 0, j \geq 0} \left(\frac{\pi\delta}{r + 2\pi}\right)^{i+j+1} \begin{pmatrix} i+j \\ i \end{pmatrix} \frac{h_B + t_B + j - \tau}{(h_A + i + t_A + 1)(h_B + t_B + j) - \tau^2}, \tag{67}$$

and analogously for $a_B(t_A, t_B)$.

C Note on the derivation of the optimality equations (6) and (11)

The argument used in the text for the derivation of the optimality equations (6) and (11) relied on a heuristic limiting argument. A rigorous derivation is immediate, using a standard *uniformization* argument, as first established by Jensen (1953). Let $\{X_t, t \geq 0\}$ be a Markov process on a finite state space S with generator Q (here, the two states are "idle" or "busy") such that the diagonal elements of Q are uniformly bounded (which is clearly the case, as the switching intensities are positive constants), and let $\Gamma = \sup_{i \in S} q_i$. Then, there exists a Markov chain (discrete time) $\{Y_n, n = 0, 1, \ldots\}$ on S with transition matrix $P = Q/\Gamma + I$ and a Poisson process $\{N(), t \geq 0\}$ with rate Γ, which are independent of each other, such that the processes $\{Y_{N(t)}, t \geq 0\}$ and $\{X(t), t \geq 0\}$ have the same finite dimensional distributions and, thus, are probabilistically identical. Therefore, discrete-time dynamic programming can be used, which immediately gives (6) and (11), where the coefficients (e.g., $\pi/(\pi + r)$) are the transition probabilities of Q. (see Heyman and Sobel, pp. 310–311, for example.)