A Simple Theory of Growth in a Complicated World

Steven Callander*
Stanford University

Niko Matouschek†
Northwestern University

PRELIMINARY AND INCOMPLETE

April 2013

Abstract

We examine a setting in which managers learn about the effectiveness of managerial practices by trial-and-error. We show that such a learning process generates persistent performance differences across firms. If learning is centralized, persistent performance differences arise if firms have different initial conditions. And if learning is decentralized and managerial practices are complementary, persistent performance differences can arise even if firms are identical. Moreover, if managerial practices are complementary, performance differences can persist even if managers can imitate each other. Finally, we explore the implications of our model for the growth and development of nations.

Keywords: learning, complementarities, complexity

JEL classifications: D21, D83, L25

*Graduate School of Business, Stanford University, sjc@GSB.Stanford.edu.
†Kellogg School of Management, Northwestern University, n-matouschek@kellogg.northwestern.edu.

1 Introduction

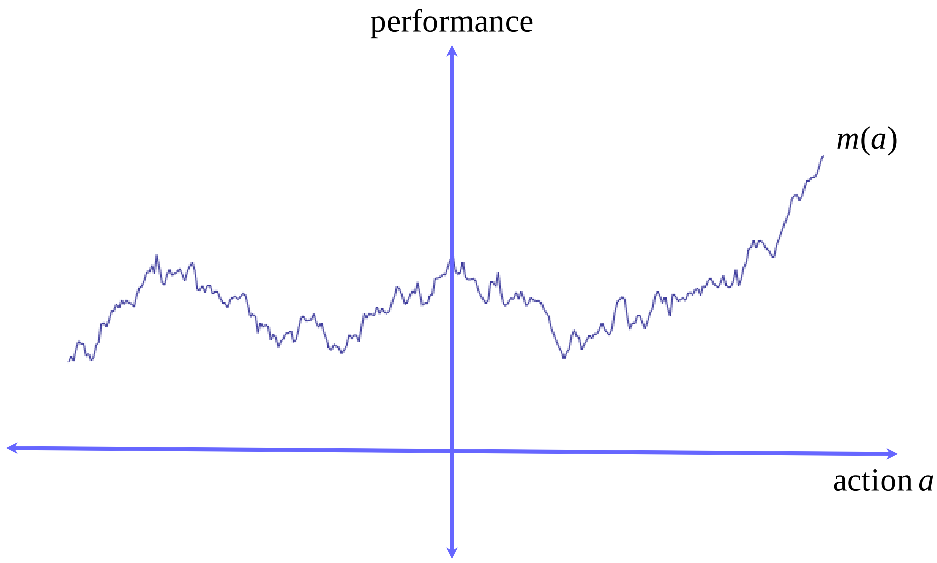
Firms that operate within the same industries, and even plants that operate within the same firms, often exhibit large and persistent performance differences. In recent years, researchers have made significant progress in exploring the implications of this fact for macroeconomics, industrial organization, labor markets, and trade (for a survey of the empirical evidence and implications see Syverson (2011)). An open question, however, is why performance differences exist in the first place. One view that has long been popular with managers and management scholars is that performance differences exist because different managers employ different managerial practices. In line with this view, a variety of recent studies have provided evidence for correlations and, in some cases, causal relationships between managerial practices and performance (see, for instance, Bloom et al. (2007, 2013) and, for a survey, Gibbons and Henderson (2012)).

If managerial practices affect performance, however, then why do different managers employ different practices? And why do they continue to employ different practices even when it has become apparent that some work better than others? After all, the managerial practices of successful firms, such as Toyota's just-in-time production system and Lincoln Electric's organizational design, are often well-known and are not protected by patents.

The aim of this paper is to shed light on these questions. Our premise is that the problem of choosing between different managerial practices is too complex to be solved analytically. Instead, managers tackle this problem by trying out different managerial practices and observing how they affect performance. In other words, managers learn about managerial practices by trial-and-error. We show that such a learning process generates persistent performance differences. Moreover, if there are complementarities between managerial practices, these performance differences can persist, even when managers can imitate each other.

Managerial Learning: To fix ideas, suppose that there is a single managerial practice that can take any value on the real line. There are infinitely many periods and in every period a manager has decide which specific practice or "action" to take. A production function maps the actions into the firm's performance. This function is increasing on average but also has many local peaks and troughs. The production function therefore looks like a rugged landscape that is tilted upwards, as illustrated in Figure 1.

Specifically, we follow Callander (2011) and model the production function as the realized path of a Brownian motion in which the action replaces time as the independent variable. The path is realized before the first period and then stays the same for all subsequent periods. The production function therefore does not change over time. We assume that the drift of the Brownian motion is positive, which ensures that the production function is increasing on average. And we interpret the variance of the Brownian motion as a measure of the complexity of the production process. In particular, the more complex the production process is, the larger is the variance; and the larger is the variance, the more rugged is the production function.



The Production Function m(a)

In the first period, the manager knows the performance generated by a status quo action. And he knows the basic characteristics of the production function, as captured by the drift and the variance of the Brownian motion. Before the manager decides what action to take, he uses this information to form beliefs about the performance generated by any action that is different from the status quo.

The assumption that the production function is generated by a Brownian motion ensures that these beliefs take a simple form. In particular, the manager's beliefs are normally distributed with a mean that is increasing in the action and a variance that is increasing in both, the complexity of the production process and the distance between the action and the status quo. The manager therefore knows that actions that are further to the right tend to generate better performance. And he knows more about an action, the closer it is to the status quo, or the less complex the production process is.

In this way the model captures the notion that managers know the general direction in which to search for better managerial practices. They may know, for instance, that they should reduce inventory or tie pay more closely to performance. While managers know the general direction in which to search for better managerial practices, however, they do not know the effect of a specific change on performance. Moreover, they are more uncertain about the effect of a specific change, the larger is the change, and the more complex is the managerial practice.

After the manager has formed his beliefs in the first period, he either takes the status quo action, in which case he learns no new information. Or he engages in search by taking a different action, in which case he learns the performance generated by that action. In the second, and any subsequent, period, the manager then updates his beliefs based on any new information he learned. And after he has updated his beliefs, he decides between taking one of the known actions and engaging in further search. The manager can therefore learn about the production function over time by trying out different actions. The question then is what actions the manager should try out and when, if ever, he should stop and settle for one of the known actions.

We show that for a broad class of utility functions, and under the assumption that the manager maximizes expected utility on a period-by-period basis, the optimal learning rule takes a simple form. In particular, if status quo performance is below a threshold, the manager takes the status quo action in the first period, and he continues to do so in all subsequent periods. If, however, status quo performance is above the threshold, the manager starts to search in the first period by taking an action that is strictly to the right of the status quo. In subsequent periods, he then continues to search by taking actions that are further and further from the status quo until he reaches a period in which performance falls below another threshold. Once that happens, the manager reverts to the best known action and then continues to take that action in all subsequent periods.

Persistent Performances Differences: A key implication of the optimal learning rule is that it generates persistent performance differences. To see this, consider the effect of an increase in status quo performance. We already noted that such an increase makes it more likely that the manager engages in search in the first period. Below we also show that if the manager does engage in search in the first period, an increase in status quo performance increases his action and makes it more likely that he will engage in further search in any subsequent period. An increase in status quo performance therefore unambiguously favors search. And because an increase in status quo performance favors search, it generates an even larger increase in expected performance in all subsequent periods. Differences in status quo performance therefore do not only persist, they actually grow larger.

This explanation for persistent performance differences relies on differences in status quo performances. Once we allow for multiple, complementary managerial practices, however, persistent performance differences can arise even if firms are identical. To see this, suppose that there is a second action. For simplicity, the production function for this second action is identical to the first. Each action is taken by a different manager who cares about the performance generated by his action minus the squared difference in the actions. The actions are therefore complementary.

In this setting, we can compare centralized learning–in which case the managers take their actions jointly–to decentralized learning–in which case they take them independently of each other. If learning is centralized, the optimal learning rule is the same as the one described above.

If learning is decentralized, however, there is a key difference. In particular, in any period other than the first one, the manager's actions are now determined by two thresholds. If performance is above the upper threshold, both managers engage search. And if performance is below the lower threshold, both managers revert to the best known action. If performance is between the two thresholds, however, there are multiple equilibria: either both managers engage in search or neither does. For intermediate levels of performance, the managers can therefore get stuck in a coordination failure. Moreover, the more complementary the actions, the more likely it is that coordination failures arise. Indeed, if the actions become very complementary, the only way to rule out coordination failures is for firm performance to improve continuously over time, without a single stumble or fall.

In summary, we show that persistent performances arise naturally if managers learn about managerial practices through trial-and-error. In particular, they arise if learning is centralized and firms face different initial conditions or if learning is decentralized and managerial practices are complementary.

Barriers to Managerial Imitation: So far we have abstracted from imitation. As we observed above, however, the managerial practices of successful firms are often well-known and not protected by patents. The question then is why under-performing firms don't simply imitate the top performing ones. The answer that we develop below is that imitation may simply be too risky.

To see why imitation may be risky, suppose that some managerial practices are industry-specific–in the sense that their production functions are the same for all firms in the industry–while others are firm-specific. The mapping from managerial actions into corporate culture, for example, may be firm specific. This would be the case, for instance, if an American car manufacturer such as General Motors has to take different actions to create a collaborative culture than a Japanese one such as Toyota.

If managerial practices are independent of each other, one would expect under-performers to imitate the top performers' industry-specific managerial practices. If, however, there are complementarities between industry- and firm-specific managerial practices, then imitating industry-specific managerial practices requires changes to firm-specific ones. Suppose, for instance, that the success of Toyota's just-in-time production system depends on its corporate culture and that, as we argued above, the production function that maps actions into corporate culture is firm-specific. To imitate Toyota's just-in-time production system, General Motors would then also need to change its corporate culture. But if managers learn about the production function for corporate culture by trial-and-error, then trying to change corporate culture is inherently risky.

Our final set of results characterizes conditions under which the risk associated with imitating industry-specific managerial practices is prohibitive. In particular, we show that if the top performers' industry-specific managerial practices are far from those employed by an under-performing firm, and the complementary, firm-specific managerial practices are sufficiently complex, then the under-performing firm is better off if it does not imitate the top performers.

We cast our model in the context of firms and industries. The main ingredients of our model–complexity, complementarities, and learning by trial-and-error–are also relevant in other settings. The vast literature on the growth and development of nations, in particular, touches on many of the issues that arise in our setting. We will discuss to this application of our model in the conclusion.

2 The Model

There is a single manager. At the beginning of every period $t = 1, 2, ...$, the manager takes an action that determines his income. After the manager has consumed his income, time moves on to the next period. Our aim is to characterize the manager's optimal actions given the technology, preferences, and information structure that we describe next.

Technology: At the beginning of period $t = 1, 2, ...$, the manager takes an action $a_t \in \mathbb{R}$ that determines his income level $m_t \in \mathbb{R}$. The mapping from actions into income is given by the production function $m(a_t)$, where $m : \mathbb{R} \to \mathbb{R}$. We follow Callander (2011) and model m as the realized path of a Brownian motion with drift $\mu > 0$ and variance $\sigma^2 > 0$. For reasons that will become apparent, we interpret the variance σ^2 as a measure of the complexity of the production process. Moreover, we refer to $a_0 = 0$ as the status quo action and denote status quo income by $m_0 = m(a_0)$. The realized path of the Brownian motion is determined by nature before the start of the game and does not change over time. Figure 1 shows one possible realization of the

Brownian motion.

Preferences: The manager's utility is given by $u(m)$, where m is his income. We assume that this function is four times continuously differentiable and satisfies

$$u'(m) > 0 \text{ and } u''(m) < 0 \text{ for all } m \in \mathbb{R}.$$

The first condition implies non-satiation and the second risk aversion.

To further characterize the manager's risk preferences, let

$$r(m) = -\frac{u''(m)}{u'(m)}$$

denote the coefficient of absolute risk aversion and let

$$p(m) = -\frac{u'''(m)}{u''(m)}$$

denote the coefficient of absolute prudence. We assume that absolute prudence is decreasing, that is, $p'(m) \leq 0$ for all $m \in \mathbb{R}$. Prudence was introduced by Kimball (1990) who shows that a risk averse agent engages in precautionary savings if and only if he is prudent, that is, $p(m) > 0$. Moreover, he shows that precautionary savings are decreasing in income if and only if absolute prudence is decreasing. Finally, he shows that on an unbounded domain, decreasing absolute prudence implies decreasing absolute risk aversion. In our setting, the manager's utility function therefore also satisfies $r'(m) \leq 0$ for all $m \in \mathbb{R}$, where $r(m)$ denotes the coefficient of absolute risk aversion defined above.

Finally, we assume that the coefficient of absolute risk aversion crosses $2\mu/\sigma^2$, where μ and σ^2 are the drift and the variance of the Brownian motion. We will see below that if $r(m)$ did not cross $2\mu/\sigma^2$, the solution to the manager's problem would either be trivial or would not exist. We denote the largest income level for which $r(m) = 2\mu/\sigma^2$ by $\widehat{m}$.

Information: In any period, the manager knows the income generated by the status quo action and by any action he took in any previous period. We refer to these actions as "known actions" and to all other actions as "unknown actions." In addition to the known actions, the manager knows that the production function was generated by a Brownian motion with drift μ and variance σ^2. The manager does not, however, know the realization of the Brownian motion. In any period t, the manager's information set is therefore given by $I_t = \left\{\mu, \sigma^2, (a_0, m_0), ... (a_{t-1}, m_{t-1})\right\}$.

Optimal Learning Rule: For simplicity we assume that the manager maximizes expected utility on a period-by-period basis. An optimal learning rule is therefore given by $(a_1^*, a_2^*, ...)$, where

$$a_t^* \in \arg\max_{a_t} \mathrm{E}\left[u\left(m_t\right)|I_t\right].$$

Our goal is to characterize the set of optimal learning rules.

3 Beliefs and Expected Utility

We start by examining the manager's beliefs about the income generated by any unknown action. For this purpose, consider any period t and let l_t and r_t denote the left-most and right-most known actions. Consider now an unknown action a_t that is to the right of r_t. For any such action, the manager believes that income $m\left(a_t\right)$ is drawn from a normal distribution with mean

$$\mathrm{E}\left[m\left(a_t\right)\right] = m\left(r_t\right) + \mu(a_t - r_t) \tag{1}$$

and variance

$$\mathrm{Var}\left(m\left(a_t\right)\right) = \left(a_t - r_t\right)\sigma^2, \tag{2}$$

where μ and σ^2 are the drift and the variance of the Brownian motion. The manager therefore expects an action to generate higher income, the further it is to the right of r_t. At the same time, however, the further an action is to the right of r_t, the more uncertain the manager is about the income generated by that action. The beliefs for actions to the left of the left-most action l_t are analogous.

Notice that the manager's beliefs about any action to the right of r_t depend only on r_t and that his beliefs about any action to the left of l_t depend only on l_t. Similarly, the manager's beliefs about any action between l_t and r_t depend only on the known actions closest to that action. To see this without having to introduce more notation, suppose that there are no known actions between l_t and r_t. For any action $a_t \in [l_t, r_t]$, the manager then believes that income $m(a_t)$ is normally distributed with mean

$$\mathrm{E}\left[m\left(a_t\right)\right] = \frac{a_t - l_t}{r_t - l_t} m\left(r_t\right) + \frac{r_t - a_t}{r_t - l_t} m\left(l_t\right) \tag{3}$$

and variance

$$\mathrm{Var}\left(m\left(a_t\right)\right) = \frac{\left(a_t - l_t\right)\left(r_t - a_t\right)}{r_t - l_t}\sigma^2. \tag{4}$$

The manager's expected income is therefore a convex combination of the income generated by l_t and r_t. Moreover, the further the action is from the closest known action, the more uncertain the manager is about the income generated by that action.

The assumption that the production function is generated by a Brownian motion therefore ensures that the manager's beliefs take a simple form that satisfies several intuitive properties. First, beliefs are normally distributed. Second, the manager knows more about an action, the closer the action is to a known action, and the less complex is the production process. Third, the manager engages in directed search, that is, he knows where he can expect better actions and he focuses his search in that direction. And finally, even if, over time, the manager learns the income generated by a very large number of actions, he can never infer the entire production function. In this sense, there is a limit to theoretical knowledge and thus a need to learn about the world by trial-and-error.

Now that we have examined the manager's beliefs, we can specify his expected utility. Suppose that the manager believes that income is normally distributed with mean M and variance V and let z denote a random variable that is drawn from a standard normal distribution. We can then state our first lemma, which is proven in Hilfsatz 4.3 in Schneeweiss (1966) and Theorem 1 in Chipman (1973).

LEMMA 1 (Schneeweiss 1966 and Chipman 1973). *Suppose that* $|u(m)| \leq A\exp(-Bm^2)$ *for some* $A > 0$ *and* $B > 0$. *Then the expected utility function*

$$W(M,V) = \mathrm{E}\left[u\left(M + \sqrt{V}z\right)\right]$$

exists for all $M \in (-\infty, \infty)$ *and* $V \in (0, 1/(2B))$. *Moreover, the expected utility function satisfies*

$$2\frac{\partial W(M,V)}{\partial V} = \frac{\partial^2 W(M,V)}{\partial M^2}.$$

The restriction in the lemma ensures that expected utility is integrable, and for the remainder of the paper we assume that it holds. Notice that since we are free to choose any positive parameters A and B, this restriction is mild. The partial differential equation in the lemma is known as the "heat equation" and we will use it in a number of proofs below.

The next lemma establishes another property of the expected utility function that is proven in Theorem 3 in Chipman (1973) and Theorem 2 in Lajeri and Nielsen (2000).

LEMMA 2 (Chipman 1973 and Lajeri and Nielsen 2000). *The expected utility function* $W(M,V)$ *is concave.*

The key property that ensures concavity of the expected utility function is decreasing absolute prudence. In the next section we will see that in the relevant range both expected income and its variance are linear in the manager's action. Decreasing absolute prudence therefore ensures that the manager's problem is concave.

4 Managerial Learning

We first focus on the optimal action in the first period and then turn to the second and all subsequent periods.

4.1 The First Period

In the first period, the manager can either take the status quo action, in which case he is certain to realize status quo income m_0. Or he can take some action $a_1 \neq a_0$, in which case he is uncertain about what income he will realize. In the previous section, we saw that for any action $a_1 < a_0$, expected income is strictly less than status quo income. The manager will therefore never take an action strictly to the left of the status quo.

Suppose then that $a_1 \geq a_0$ and let $\Delta_1 = a_1 - a_0 \geq 0$ denote the size of the step the manager takes in the first period. We then know from (1) and (2) that the manager's expected income is given by $m_0 + \mu\Delta_1$ and its variance is given by $\sigma^2\Delta_1$. We can therefore write the manager's problem as

$$\max_{\Delta_1 \geq 0} W\left(m_0 + \mu\Delta_1, \sigma^2\Delta_1\right),$$

where $W(\cdot)$ is the expected utility function defined in Lemma 1. As observed above, Lemma 2 ensures that this problem is concave.

Next, by differentiating $W(\cdot)$ with respect to Δ_1, and making use of the heat equation in Lemma 1, we obtain

$$\frac{\mathrm{d}W\left(m_0 + \mu\Delta_1, \sigma^2\Delta_1\right)}{\mathrm{d}\Delta_1} = \mathrm{E}\left[u'\left(m_0 + \mu\Delta_1 + \sigma\sqrt{\Delta_1}z\right)\right]\frac{\sigma^2}{2}\left(\frac{2\mu}{\sigma^2} - R\left(m_0, \Delta_1\right)\right), \qquad (5)$$

where

$$R\left(m_0, \Delta_1\right) \equiv -\frac{\mathrm{E}\left[u''\left(m_0 + \mu\Delta_1 + \sqrt{\Delta_1}\sigma z\right)\right]}{\mathrm{E}\left[u'\left(m_0 + \mu\Delta_1 + \sqrt{\Delta_1}\sigma z\right)\right]}.$$

Notice that $R(m_0, 0)$ is equal to the coefficient of absolute risk aversion $r(m_0)$. At $\Delta_1 = 0$, the slope of the expected utility function is therefore determined by the relative size of the coefficient of absolute risk aversion and the ratio $\left(2\mu/\sigma^2\right)$. In particular, we have

$$\frac{\mathrm{d}W\left(m_0, 0\right)}{\mathrm{d}\Delta_1}\begin{cases} > 0 & \text{if } m_0 > \widehat{m} \\ \leq 0 & \text{if } m_0 \leq \widehat{m}, \end{cases} \qquad (6)$$

where $\widehat{m}$ denotes the largest income level m for which $r(m) = 2\mu/\sigma^2$. Intuitively, the manager prefers an action just to the right of the status quo to the status quo itself if and only if he is sufficiently wealthy, in which case his coefficient of absolute risk aversion is sufficiently small.

We can now establish our first proposition which characterizes the manager's optimal action in the first period.

PROPOSITION 1. *The manager's optimal first period action is unique and given by*

$$a_1^* = \begin{cases} a_0 + \Delta^*(m_0) & \text{if } m_0 \geq \widehat{m} \\ a_0 & \text{if } m_0 < \widehat{m}, \end{cases}$$

where $\Delta^*(m_0)$ *is implicitly defined by*

$$R(m_0, \Delta^*(m_0)) = \frac{2\mu}{\sigma^2} \tag{7}$$

and satisfies $\Delta(\widehat{m}) = 0$ *and* $\Delta'(m_0) > 0$ *for all* $m_0 \geq \widehat{m}$.

If $m_0 < \widehat{m}$, it is therefore optimal for the manager to take the status quo action. This result follows immediately from the above observations that expected utility is concave and that for $m_0 < \widehat{m}$, the slope of expected utility is negative at $\Delta_1 = 0$. If, instead, $m_0 > \widehat{m}$, there exists a unique optimal action that is strictly larger than the status quo. Moreover, the step the manager takes in the first period is larger, the larger is his status quo income.

In the proof of the proposition we show that existence and uniqueness of an optimal action follows from the assumption that absolute risk aversion crosses $(2\mu/\sigma^2)$. If absolute risk aversion did not cross $(2\mu/\sigma^2)$, an optimal action might not exist. In particular, suppose that $r(m) < 2\mu/\sigma^2$ for all $m \in \mathbb{R}$. It then follows that from (5) that $W_{\Delta_1}(m_0, \Delta_1) > 0$ for all $m_0 \in \mathbb{R}$ and for all $\Delta_1 \in \mathbb{R}$. In this case the manager therefore always benefits from taking a larger and larger action.

Notice also that if it were the case that $r(m) \geq 2\mu/\sigma^2$ for all $m \in \mathbb{R}$, then $W_{\Delta_1}(m_0, 0) \leq 0$ for all $m_0 \in \mathbb{R}$. Since $W(m_0, \Delta_1)$ is concave, it is then immediate that $a_1^* = a_0$ would be an optimal action for all $m_0 \in \mathbb{R}$. Moreover, taking the status quo action would be uniquely optimal unless the manager's income level is given by an m_0 for which $r(m_0) = p(m_0) = 2\mu/\sigma^2$. The assumption that absolute risk aversion crosses $2\mu/\sigma^2$ therefore ensures the existence of a non-trivial solution to the manager's first period problem, as well as to the manager's problems in all subsequent periods.

4.2 The Second and Subsequent Periods

We just saw that if $m_0 \leq \widehat{m}$, the manager takes the status quo action in the first period. In any subsequent period, the manager then faces the same problem as in the first, and takes the status quo action again. For the remainder of this paper, we therefore assume that

$$m_0 > \widehat{m},$$

in which case the manager engages in search in the first period by taking an action a_1^* that is strictly larger than the status quo a_0.

Since the manager engages in search in the first period, there are two known actions in the second period: the status quo action a_0 and his first period action a_1^*. From (3) we know that for any action between a_0 and a_1^*, expected income is a convex combination of status quo income m_0 and first period income $m_1^* \equiv m(a_1^*)$. The manager therefore always prefers one of the two known actions to any action between those two actions. Moreover, just as in the first period, the manager always prefers the status quo to any action strictly to the left of the status quo. If an optimal action a_2^* exists, it therefore satisfies $a_1^* \in a_0 \cup [a_1^*, \infty)$.

To characterize the optimal second period action, we proceed in two steps. First, we assume that the manager is constrained to take an action that is larger than his first period action, that is, we assume $a_2 \geq a_1^*$. Second, once we know the constrained optimal action, we examine under what conditions the manager prefers that action to the status quo.

Suppose therefore that $a_2 \geq a_1^*$ and let $\Delta_2 = a_2 - a_1^* \geq 0$ denote the size of the step the manager takes in the second period. The manager's constrained problem is then given by

$$\max_{\Delta_2 \geq 0} W\left(m_1^* + \mu\Delta_2, \sigma^2\Delta_2\right),$$

where $W(\cdot)$ is the expected utility function defined in Lemma 1. Notice that this is the same problem the manager faced in the first period, except that the income level associated with a step of size zero is now given by m_1^* rather than m_0. It then follows from Proposition 1 that the solution to the manager's constrained problem depends on whether m_1^* is smaller than $\widehat{m}$. In particular, if $m_1^* \leq \widehat{m}$, the solution is given by a_1^*. And if $m_1^* > \widehat{m}$, the solution is given by $a_1^* + \Delta^*(m_1^*)$, where $\Delta^*(m_1^*) > 0$ is implicitly defined in (7).

Next we examine when the manager prefers the constrained optimal action to the status quo. For this purpose, consider the next lemma.

LEMMA 3. *There exists a unique income level* $\widetilde{m}(m_0) \in (\widehat{m}, m_0)$ *such that*

$$u(m_0) = W\left(\widetilde{m}(m_0) + \mu\Delta^*(\widetilde{m}(m_0)), \sigma^2\Delta^*(\widetilde{m}(m_0))\right), \tag{8}$$

where $\Delta^*(\widetilde{m}(m_0)) > 0$. *The derivative of this income level satisfies*

$$0 < \frac{\mathrm{d}\widetilde{m}(m_0)}{\mathrm{d}m_0} < 1.$$

If first period income m_1^* is equal to $\widetilde{m}(m_0)$, the manager is indifferent between the status quo and the constrained optimal action. It then follows that the manager strictly prefers the constrained optimal action to the status quo if $m_1^* > \widetilde{m}(m_0)$ and that he strictly prefers the status quo to the constrained optimal action if $m_1^* < \widetilde{m}(m_0)$. Notice that this implies that if m_1^* is strictly between $\widetilde{m}(m_0)$ and $\widehat{m}$, the manager reverts to the status quo even though the marginal return from engaging in further search is positive. The reason is that while in this case search is better than m_1^*, it is not better than the status quo $m_0 > m_1^*$.

The problem that the manager faces in any period $t > 2$ is very similar to the one he faces in period 2. The only difference is that in any period $t > 2$, the manager does not necessarily compare his expected utility from engaging in further search with his utility from the status quo, as he does in the second period. Instead, the manager compares his expected utility from engaging in further search with his utility from whatever known action generates the largest income level, which may be the status quo action or some other known action.

To state the proposition that characterizes the manager's optimal action in all periods $t \geq 2$, let $\overline{m}_t$ denote the largest known income level in period t, that is, let

$$\overline{m}_t = \max\{m_0, m_1^*, m_2^*, ..., m_{t-1}^*\}.$$

Also, let $\overline{a}_t$ denote the action that generates $\overline{m}_t$, that is, let

$$\overline{a}_t \in \{a_0, a_1^*, a_2^*, ..., a_{t-1}^*\} \text{ such that } m(\overline{a}_t) = \overline{m}_t.$$

And finally, recall that r_t denotes the largest known action in period t. We can then state our next proposition.

PROPOSITION 2. *The manager's optimal action in period $t \geq 2$ is unique and given by*

$$a_t^* = \begin{cases} r_t + \Delta^*(m(r_t)) & \text{if } m(r_t) > \widetilde{m}(\overline{m}_t) \\ a(\overline{m}_t) & \text{otherwise,} \end{cases}$$

where $\Delta^*(m(r_t)) > 0$ is defined in (7) *and $\widetilde{m}(\overline{m}_t)$ is defined in* (8).

In any period $t \geq 2$, the manager therefore engages in search if and only if the income level $m(r_t)$ associated with the largest previously taken action r_t is above a threshold $\widetilde{m}(\overline{m}_t)$, where the threshold is increasing in the largest known income level $\overline{m}_t$.

4.3 Discussion

The two propositions we just derived characterize the optimal learning rule and show that it takes a simple form. In particular, if status quo income is below a threshold, the manager takes the status quo action in the first period, and he continues to do so in all subsequent periods. If, however, status quo income is above the threshold, the manager starts to search in the first period by taking an action that is strictly larger than the status quo. In subsequent periods, he then continues to search by taking larger and larger actions until he reaches a period in which income falls below another threshold. Once that happens, the manager reverts to the best known action and then continues to take that action in all subsequent periods.

Figuratively, the manager's exploration of the rugged landscape depends on the height of his starting point. If that point is too low, he just stays put. If his starting point is sufficiently high, however, he starts exploring the rugged landscape by taking discrete steps towards the right. The manager continues his rightward march until he falls off a sufficiently large cliff. At that point, continuing his march is too risky for the manager, who instead returns to the highest peak he discovered during his exploration.

Notice that even though the manager eventually settles for an action that guarantees him some level of income, he is not satisficing. In particular, the manager does not stop searching for better actions once he found one that is "good enough." Instead, he stops searching for better actions after having taken a sufficiently bad one. The manager then reverts to the best known action because further search is too risky. Notice also that when the manager does settle for an action, that action is a local peak *given the manager's information*. Since the manager's information is coarse, however, his action will not, in general, be an actual local peak.

A key implication of the model is that it generates persistent performance differences. We discuss this implication in the next section.

5 Persistent Performance Differences

In this section we first show that persistent performance differences arise if firms have different status quo income levels. We then show that if there are multiple, complementary managerial practices and learning is decentralized, persistent performance differences can arise even if firms are identical.

5.1 Differences in Status Quo Income Levels

To understand the effect of an increase in status quo income, recall that the optimal first period action is increasing in status quo income. An increase in status quo income therefore generates higher expected income in the first period, both because the starting point from which the manager searches is higher, and because his first step is larger. A larger step, however, also generates more uncertain income. And higher status quo income raises the threshold below which the manager finds it optimal to revert to the status quo in the second period. It is therefore not immediately obvious whether an increase in status quo income makes it more likely that the manager will engage in further search in the second period. Specifically, it follows from Proposition 2 that the probability with which the manager engages in further search in the second period is given by

$$prob\left(m_1^* > \widetilde{m}\left(m_0\right)\right) = F\left(\frac{m_0 + \mu\Delta^*\left(m_0\right) - \widetilde{m}\left(m_0\right)}{\sqrt{\Delta^*\left(m_0\right)}\sigma}\right),$$

where $F\left(\cdot\right)$ is the cumulative density function of the standard normal distribution. Notice that an increase in m_0 increases both the numerator–the difference between expected first period income $m_0+\mu\Delta^*\left(m_0\right)$ and the threshold level of income $\widetilde{m}\left(m_0\right)$ above which the manager engages in further search–and the denominator–the standard deviation of first period income $\sqrt{\Delta^*\left(m_0\right)}\sigma$. The next lemma shows that the effect on the numerator dominates the effect on the denominator. And it shows that this is true for all subsequent periods.

LEMMA 4. *The probability that the manager engages in further search in period* $t \geq 2$ *is strictly increasing in status quo income, that is,*

$$\frac{\mathrm{d}prob\left(m_t^* > \widetilde{m}\left(\overline{m}_t\right)\right)}{\mathrm{d}m_0} > 0.$$

An increase in status quo income therefore favors search unambiguously: it makes it more likely that the manager engages in search in the first period, it increases the size of the first period action if the manager does engage in search, and it makes it more likely that the manager will engage in further search in any subsequent period. As a result, an increase in status quo income generates an even larger increase in expected income in any subsequent period, as shown in the next proposition.

PROPOSITION 3. *An increase in status quo income generates a disproportionate increase in expected income in any period* $t \geq 1$*, that is,*

$$\frac{\mathrm{dE}\left[m_t^*(m_0)\right]}{\mathrm{d}m_0} > 1.$$

To interpret the lemma, suppose there are two firms that differ in their status quo income. The two firms use the same production function. Their managers, however, cannot observe each other and thus learn about the production function independently. The lemma implies that the difference in expected income in any period $t \geq 1$ is strictly larger than the difference in status quo income. Differences in status quo income therefore do not only persist, they actually grow larger.

5.2 Complementarities and Decentralized Learning

The explanation for persistent performance differences that we just discussed relies on differences in status quo income. In this section we show that if there are multiple, complementary managerial practices, persistent performance differences can arise even if firms are identical.

For this purpose, suppose there are two managers, A and B. In any period $t = 1, 2, ...$ the managers make decisions $a_t^A \in \mathbb{R}$ and $a_t^B \in \mathbb{R}$ and then realize their incomes m_t^A and m_t^B. In particular, their incomes are given by

$$m_t^A = m\left(a_t^A\right) - \frac{1}{2}\delta\left(a_t^A - a_t^B\right)^2$$

and

$$m_t^B = m\left(a_t^B\right) - \frac{1}{2}\delta\left(a_t^A - a_t^B\right)^2,$$

where $\delta \geq 0$ is a parameter that measures the importance of coordination between the managers' decisions. The function $m(\cdot)$ is once again the realized path of a Brownian motion with drift μ and variance σ^2 and for which $m_0 = m(0)$. Notice that this function is the same for both managers. The managers are therefore learning about the same production function.

The managers know that the production function is generated by a Brownian motion with drift μ and variance σ^2 and that $m(0) = 0$. Moreover, in any period t, the managers know the actions that they took in previous periods and the income that these actions generated. Managers A and B therefore have the same information set

$$I_t = \left\{\mu, \sigma^2, (0, m_0), \left(a_1^A, m_1^A\right), \left(a_1^B, m_1^B\right), ..., \left(a_{t-1}^A, m_{t-1}^A\right), \left(a_{t-1}^B, m_{t-1}^B\right)\right\}.$$

And they have the same utility function $u(\cdot)$ that satisfies the same properties described above.

We distinguish between centralized learning–in which case the managers coordinate their actions–and decentralized learning–in which case the managers take their actions independently and simultaneously. If learning is centralized, the optimal learning rule is the same as the one we derived above. If learning is decentralized, however, there is a key difference, which we derive below. Since the formal analysis of this version of the model is similar to the single-action version we described above, we relegate the formal analysis to the appendix.

5.3 The First Period

The managers never find it optimal to take actions to the left of the status quo, just as in the single-action model. Suppose therefore that $a_1^A \geq a_0^A$ and $a_t^B \geq a_0^B$ and let $\Delta_1^A = a_1^A - a_0^A$ and

$\Delta_1^B = a_1^B - a_0^B$ denote the size of each manager's first period step. We can then write manager A's expected utilities as

$$W^A\left(m_0 + \mu\Delta_1^A - \frac{1}{2}\delta\left(\Delta_1^A - \Delta_1^B\right)^2, \sigma^2\Delta_1^A\right) = E\left[u\left(m_0 + \mu\Delta_1^A - \frac{1}{2}\delta\left(\Delta_1^A - \Delta_1^B\right)^2 + \sigma\sqrt{\Delta_1^A}z\right)\right],$$

where z is again a random variable drawn from the standard normal distribution. The definition of manager B's expected utility is analogous. The managers' first period problem is then given by

$$\max_{\Delta_1^A \geq 0} W^A\left(m_0 + \mu\Delta_1^A - \frac{1}{2}\delta\left(\Delta_1^A - \Delta_1^B\right)^2, \sigma^2\Delta_1^A\right)$$

and

$$\max_{\Delta_1^B \geq 0} W^B\left(m_0 + \mu\Delta_1^B - \frac{1}{2}\delta\left(\Delta_1^A - \Delta_1^B\right)^2, \sigma^2\Delta_1^B\right).$$

The solution to this problem gives the managers' optimal first period actions, which we characterize in the next proposition.

PROPOSITION 4. *The managers' optimal first period actions are unique and given by*

$$a_1^{A*} = a_1^{B*} = a_1^*$$

where a_1^* is the optimal action in the single-action model defined in Proposition 1*.*

The optimal first period actions are therefore the same as in the single-action model. Notice that this implies that there cannot be any coordination failures in the first period, even if the actions are very complementary, that is, even if δ is very large.

To understand why coordination failures cannot arise in the first period, suppose that δ is very large and that $\Delta_1^A = \Delta_1^B = 0$. Notice that if $\Delta_1^A = \Delta_1^B$, the effect of a marginal increase in the size of a manager's action on expected utility is second order. As long as $m_0 > \widehat{m}$, manager A has an incentive to unilaterally increase the size of his action by a small amount. But once A has increased his action, manager B benefits from increasing her action by the same amount. And once B has increased her action by the same amount, A benefits by unilaterally increasing her action by a further small amount. This process continues until each manager's action is equal to the optimal action in the single-action model, at which point neither manager has an incentive to unilaterally change his or her action.

5.4 The Second and Subsequent Periods

Consider first the second period. As in the single-action model, it is never optimal to take an action that is either strictly to the left of the status quo or strictly between the status quo and the optimal first period action a_1^*. And as in the first period of the two-action model, the two managers will always take the same action. In contrast to either setting, however, the managers may now get stuck in a coordination failure.

To understand why coordination failures can arise in the second period, suppose that m_1^* is just marginally larger than the threshold income level $\widetilde{m}(m_0)$. We saw above that if $m_1^* > \widetilde{m}(m_0)$, then in the single-action model the optimal second period action is given by $a_1^* + \Delta^*(m_1^*) > a_1^*$, where $\Delta^*(m_1^*)$ is defined in (7). Suppose now that in the second period manager A takes action $a_1^* + \Delta^*(m_1^*)$. Since the managers benefit from coordination, it is then optimal for manager B to take the same action. It is therefore an equilibrium for each manager to take action $a_1^* + \Delta^*(m_1^*)$. Suppose now, however, that manager A takes the status quo action a_0. If δ is large enough, it is then optimal for manager B to take the same action. It is therefore also an equilibrium for each manager to take the status quo action, in which case the managers are stuck in a coordination failure.

To understand the conditions under which coordination failures can arise in the second period, suppose that $m_1^* > \widetilde{m}(m_0)$ and let $\Delta^{k*}(m_1^*, 0)$ denote the solution to

$$\max_{\Delta_2^A \geq 0} W^A\left(m_0 + \mu\Delta_2^A - \frac{1}{2}\delta\left(\Delta_2^A\right)^2, \sigma^2\Delta_2^A\right)$$

and

$$\max_{\Delta_2^B \geq 0} W^B\left(m_0 + \mu\Delta_2^B - \frac{1}{2}\delta\left(\Delta_2^B\right)^2, \sigma^2\Delta_2^B\right)$$

In the appendix we show that there then exists a unique income level $\widetilde{m}(m_0, \delta)$ such that

$$u(m_0) = W^K\left(\widetilde{m}(m_0,\delta) + \mu\Delta^{k*}(\widetilde{m}(m_0,\delta)) - \frac{1}{2}\delta\left(\Delta^{k*}(\widetilde{m}(m_0,\delta))\right)^2, \sigma^2\Delta^{k*}(\widetilde{m}(m_0,\delta))\right) \quad (9)$$

for $k = A, B$. In words, there exists a unique income level such that if one manager takes the status quo action, the other is indifferent between also taking the status quo action and engaging in search. Notice that $\widetilde{m}(m_0, 0) = \widetilde{m}(m_0)$. Moreover, we show in the appendix that while $\widetilde{m}(m_0, \delta)$ is strictly increasing in δ, it is always strictly smaller than m_0. For any $\delta > 0$ there therefore exists a region $(\widetilde{m}(m_0, \delta), m_0)$ such that if m_1^* is within that region, there are two equilibria. In one of those equilibria, both managers take the status quo action. And in the other, both managers take action $a_1^* + \Delta^*(m_1^*)$, which is the optimal second period action in the single-action model.

As in the single-action model, the problem faced by the managers in periods $t \geq 2$ is very similar to the first period problem. To characterize the optimal actions in periods $t \geq 2$, notice first that it is never optimal for managers A and B to take different actions. We can therefore once again use $\overline{m}_t$ to denote the largest income level known in period t. Similarly, we can once again let $a(\overline{m}_t)$ denote the action associated with $\overline{m}_t$. And finally, we once again let r_t denote the right-most known action in period t. We then have the following proposition.

PROPOSITION 5. *The managers' optimal actions in period $t \geq 2$ are given by*

$$a_t^{A*} = a_t^{B*} = a_t^* = \begin{cases} r_t + \Delta^*(m(r_t)) & \text{if } m(r_t) > \widetilde{m}(\overline{m}_t, 0) \\ a(\overline{m}_t) & \text{if } m(r_t) \leq \widetilde{m}(\overline{m}_t, \delta) \end{cases}$$

where $\Delta^*(m(r_t)) > 0$ and $\widetilde{m}(\overline{m}_t, \delta)$ are defined in (7) and (9) and where $\widetilde{m}(\overline{m}_t, \delta)$ is strictly increasing in δ and satisfies $\widetilde{m}(\overline{m}_t, \delta) < \overline{m}_t$ for any $\delta \geq 0$.

In any period t, multiple equilibria therefore arise if the income level associated with the right-most action r_t is strictly between $\widetilde{m}(\overline{m}_t, 0)$ and $\widetilde{m}(\overline{m}_t, \delta)$. We already noted that the critical income level $\widetilde{m}(\overline{m}_t, \delta)$ is strictly increasing in δ. In the appendix we further show that as $\delta \to \infty$, $\widetilde{m}(\overline{m}_t, \delta) \to \overline{m}_t$. This implies that as coordination becomes very important, coordination failures become pervasive. In particular, the only way to avoid coordination failures for sure is for each period's income level to be larger than the previous period's income level.

6 Barriers to Imitation

So far we have abstracted from imitation. As we observed in the Introduction, however, the managerial practices of successful firms are often observable and are protected by patents. This raises the question of why under-performers do not imitate the managerial practices of top performers.

To address this question, suppose now that there is a single manager who has to take actions a_t^A and a_t^B in every period t. These actions are so complementary that the manager always needs to set them equal to each other, that is, $a_t^A = a_t^B = a_t$. The manager's income is then given by

$$m_t = m^A(a_t) + m^B(a_t),$$

where $m^A(a_t^A)$ and $m^B(a_t^B)$ are the production functions for the two actions. Each production function is an independently realized path of a Brownian motion with drift μ and variance σ^2. The status quo actions are given by $a_0^A = a_0^B = a_0 = 0$ and each status quo action generates status quo income $m^A(0) = m^B(0) = m_0/2$, where we divide by two for notational convenience. The rest of the model is as in the single-action model.

In the first period, the manager's expected utility is then given by

$$W\left(m_0 + 2\Delta_1\mu, 2\Delta_1\sigma^2\right) = E\left[u\left(m_0 + 2\mu\Delta_1 + \sigma\sqrt{2\Delta_1}z\right)\right],$$

where $\Delta_1 = a_1 - a_0$ denotes the size of the manager's first period step. Notice that this expression is equivalent to the manager's first period expected utility in the single-action model if the drift were given by 2μ and the variance were given by $2\sigma^2$. We can therefore use Proposition 1 to determine the manager's optimal decisions in the first period.

To focus on imitation, suppose that in the absence of any other information, the manager finds it optimal not to engage in search in the first period, that is, suppose that $m_0 < \widehat{m}$, where $\widehat{m}$ is defined in Section 4.1.

Suppose now that the manager learns about a competitor that uses the same production function for action A. The production function for this action is therefore industry-specific. In contrast, the production function for action B is firm-specific, in the sense that for each firm the production function is an independent realization of the Brownian motion with the same drift, variance, and status quo action described above. The manager knows that the actions of the competitor are given by $b^A = b^B = d > 0$, where d measures the distance in the action space. Moreover, the manager knows that $m^A(d) = D + m_0/2$, $D > 0$ measures the distance in performance.

We are interested in the conditions under which the manager prefers the status quo to imitation, that is, the conditions under which he prefers action $a_1 = 0$ to $a_1 = d$. To determine these conditions, notice that the expected utility from imitation is given by

$$W\left(m_0 + D + \mu d, \sigma^2 d\right).$$

The manager therefore prefers the status quo to imitation if and only if

$$u(m_0) \geq W\left(m_0 + D + \mu d, \sigma^2 d\right).$$

We can now use the results from the single-action model to characterize $W\left(m_0 + D + \mu d, \sigma^2 d\right)$. In particular, it follows from (6) that

$$\frac{\mathrm{d}W(m_0 + D, 0)}{\mathrm{d}d} > 0$$

if and only if $m_0 + D > \widehat{m}$. Moreover, it follows from Lemma 2 that $W\left(m_0 + D + \mu d, \sigma^2 d\right)$ is concave in d. And finally, it follows from Proposition 1 that if $m_0 + D > \widehat{m}$ there exists a unique $d > 0$ at which $W\left(m_0 + D + \mu d, \sigma^2 d\right)$ is maximized. Together these properties imply that there exists a critical distance $\widehat{d} > 0$ such that the manager prefers the status quo to imitation if and only if $d \geq \widehat{d}$. The manager therefore prefers the status quo to imitation if the competitor's managerial practices are too far its own, in which case imitation is simply too risky. The following proposition summarizes this result and shows how the critical distance varies with the economic environment.

PROPOSITION 6. *The manager prefers the status quo to imitation if and only if*

$$d \geq \widehat{d},$$

where $\widehat{d} > 0$ *is increasing in* D *and decreasing in* σ^2.

The manager therefore prefers the status quo to imitation if the competitor is too far away in terms of its actions, not in terms of its performance. Indeed, the further ahead the competitor is in terms of performance, that is, the larger D is, the more likely it is that the manager prefers imitation to the status quo. Notice also that the barrier to imitation is the complexity of the firm specific managerial practice. The less complex this managerial practice is, that is, the smaller σ^2 is, the more likely it is that the manager prefers imitation to the status quo. And if $\sigma^2 = 0$, the manager will always find it optimal to imitate that other firm.