How much to copy?
Determinants of effective imitation breadth

Felipe A. Csaszar
3026 Steinberg Hall-Dietrich Hall
Wharton School
Philadelphia, PA 19104
V: (215) 746-3112
F: (215) 898-0401
fcsaszar@wharton.upenn.edu

Nicolaj Siggelkow
2211 Steinberg Hall-Dietrich Hall
Wharton School
Philadelphia, PA 19104
V: (215) 573-7137
F: (215) 898-0401
siggelkow@wharton.upenn.edu

The authors would like to thank Rich Burton and three anonymous referees for helpful comments. We are grateful to the Mack Center for Technological Innovation at the University of Pennsylvania for generous funding.

How much to copy?
Determinants of effective imitation breadth

Abstract: It is a common and frequently implicit assumption in the literature on knowledge transfer and organizational learning that imitating practices from high-performing firms has a positive impact on the imitating firm. While a large body of research has identified obstacles to successful imitation, not much is known about what breadth of imitation is most effective. In this paper, we use a simulation model to explore how interdependence among practices, context- and firm-similarity, and time horizon interact in non-trivial ways to determine the payoffs that arise from different breadths of imitation. The results of the model allow us to qualify and refine predictions of the extant literature on imitation. In particular, the results shed light on the conditions under which increases in imitation breadth, and hence investments that facilitate the faithful copying of more practices, are valuable. In addition, the results of the model highlight that imitation can serve two different functions – mimicking high performers, and generating search by dislodging a firm from its current set of practices – each one requiring different organizational routines for its successful implementation.

I. Introduction

Copying or transferring ideas and practices from other, usually high-performing organizations is a mechanism widely used by firms to improve their performance. Indeed, imitation is a key means of organizational learning (Argote, 1999) and has found substantial attention in the academic and the practitioner-oriented literatures. Imitative practices come in many forms: among others, firms expend substantial resources to identify and transfer best practices; firms hire consultants and experts to gain access to good ideas and practices that have worked in other firms; firms invest in trade associations to share information; young firms join business incubators and seek access to well-connected venture capitalists in part with the hope to gain access to good practices used by others. Likewise, in the intra-firm domain, firms spend considerable resources to create knowledge repositories and to transfer practices from one plant to another plant, or from one franchise to another franchise (Darr, Argote, and Epple, 1995). On the prescriptive side, firms are exhorted to invest in capabilities that allow them to more quickly and extensively copy others, to implement "best practices," and to invest in absorptive capacity. The underlying rationale for these activities and prescriptions is that firms benefit from imitation.

While a large literature on imitation and knowledge transfer has focused on identifying barriers to imitation and on increasing the probability of successfully transferring knowledge, the question of how much to copy, or what is an effective breadth of imitation, has received only modest attention. Fairly little is known about the conditions under which it is advisable to engage in small, intermediate, or very broad imitation attempts; or, in other words, when it is valuable to invest in activities that prior research has found to enable broader knowledge transfer, such as codifying knowledge (Zander and Kogut, 1995), building trust relationships (Szulanski, Cappetta, and Jensen, 2004), or increasing absorptive capacity (Cohen and Levinthal, 1990).

Our paper attempts to shed light onto this problem. In particular, we focus on three factors that have been highlighted in the existing literature on imitation, but whose impact on effective imitation breadth has not been investigated before: (a) the degree of interaction between the practices of a firm, (b) the time horizon over which a firm seeks to achieve improvements, and (c) the similarity between the source and

the recipient of the copied practices. With our analysis we address questions such as, "In the presence of many interactions among practices, should a firm attempt to copy very few, some, or many practices of a high-performing firm that operates in a similar business context?" and "How would the answer to this question change as the time horizon shrinks, or as the similarity in business context decreases?"

Methodologically, we use a simulation model as a disciplined way to develop a more nuanced theory of imitation. Simulation forces the researcher to be transparent about assumptions and each modeled construct, and takes care of correctly combining them to arrive at testable hypotheses. Simulation does not replace empirical testing – but allows stating a clear set of predictions, moving theory a step closer to empirical investigation.

The paper is structured as follows: We begin by summarizing the extant literature on the benefits of imitation as a set of propositions that depend on three contingencies central to studies of imitation. Then, we describe our agent-based simulation of imitation which includes these three contingencies as parameters. Next, we present the results of the simulation. Finally, we contrast the propositions derived from the extant literature with the predictions of the model, and discuss further implications of our results.

II. Theoretical motivation

Imitation is pervasive both as a concept in the academic literature and in practice. As Lieberman and Asaba (2006:366) have noted, "imitation of superior products, processes, and managerial systems is widely recognized as a fundamental part of the competitive process." Supporting this notion, Nutt (1998) empirically found that importing knowledge from other firms is one of the most frequently used tactics to solve problems. In a similar vein, Szulanski (1996:27) observed that "the identification and transfer of best practices is emerging as one of the most important and widespread practical management issues of the latter half of the 1990s," while Argote (1999:146) argued that "the popularity of 'benchmarking' and programs to transfer 'best practices' or 'lessons learned' from one organization to another reflects the usefulness of acquiring knowledge from other firms."

The abundant references to "best practices" and "benchmarking" in the practitioner literature provide further evidence of the prevalence of willful imitation. For many managers, imitation is an important fact of organizational life – a colorful illustration of this is provided by the president of Rexhall, a manufacturer of premium motorhomes, who gleefully revealed that, "In this industry, we call it R&C: research and copy" (Schnaars, 1994:3).

The literature on imitation is intrinsically connected to the broader literature on knowledge transfer. This literature has studied knowledge transfer between different firms ("across-organizational knowledge transfer") or between different units within the same firm, say different plants or franchises ("within-organizational knowledge transfer") (Darr, Argote, and Epple, 1995; Baum and Ingram, 1998). In this paper, we focus on the subset of the knowledge transfer literature that is concerned with the copying of particular practices. At the same time, our notion of imitation is broad. The model we develop below is agnostic as to whether practices are copied across or within organizations (although certain parameter values may be more representative of one context than the other.) Thus, while in the following we will talk about one "firm" copying from another "firm," it should be understood that both of these "firms" may be units of the same organization.

To gain clarity around prior work, and to show how earlier studies relate to our analysis, it is helpful to introduce some notation. Assume a firm is attempting to copy one practice from another firm. Let θ be the probability that the firm is successful in copying that practice. Hence, if the firm attempts to copy E practices, on average the firm will successfully imitate $\beta = \theta E$ practices of the target firm. We call β the effective *breadth* of an imitation, θ the *probability* of successfully imitating one practice, and E the *extent* of the desired imitation.

A large part of the existing literature on imitation and knowledge transfer has been concerned with three broad categories of factors that affect θ. First, θ can be affected by properties of the practice or knowledge that is attempted to be transferred. For instance, the degree to which a practice is codifiable (Zander and Kogut, 1995) or tacit influences the probability of successful copying.

Second, a large number of properties of the receiver and the sender have been found to influence θ. These include, among others, the absorptive capacity of the receiver (Cohen and Levinthal, 1990); the degree of receiver reluctance to accept knowledge from the outside ("not-invented-here syndrome")(Katz and Allen, 1982); the trustworthiness of the sender (Szulanski, Cappetta, and Jensen, 2004); and whether the receiver has access to a template (Winter, 1995). One might call these factors the "imitation capabilities" of both the recipient and the sender.

Third, a number of properties of the relationship between the receiver and the sender have been recognized to affect θ. For instance, knowledge flow between individuals is facilitated when the individuals are embedded in a dense web of third-party connections (Reagans and McEvily, 2003). Similarly, the arduousness of the relationship between source and recipient (Szulanski, 1996) as well as the length of the relationship between sender and receiver (Williams, 2007) have been shown to be important determinants of whether knowledge transfer is successful. Likewise, it is has been found that knowledge is more likely to transfer within organizations, e.g., across stores owned by the same franchisee or between establishments that are owned by the same parent organization, than across independent organizations (Darr, Argote, and Epple, 1995; Baum and Ingram, 1998).

The general tenor of the imitation and knowledge transfer literature is that imitating practices from high-performing firms has a positive impact on the imitating firm. Two caveats have been pointed out, however. First, as firms copy more practices from each other, strategic convergence can ensue, decreasing differentiation among firms, increasing competition, and consequently leading to profit erosion (Porter, 1996). Second, the copied practices, even if generally considered "key success factors" in the industry, may actually not be economically valuable. For example, Staw and Epstein (2000) showed that companies associated with popular management techniques did not have higher performance, but nevertheless were more admired, perceived to be more innovative, and offered a higher pay to their chief executives than firms not associated with these techniques. To understand more clearly the benefits of different breadths of imitation, it is useful, however, to focus on practices that are actually valuable and to leave possible competitive effects outside the scope of the analysis.

While the prior literature has focused primarily on factors that influence θ, our paper focuses on factors that determine an effective breadth of imitation, β. In particular, we are interested in the conditions under which an increase in β is performance enhancing. Since an increase in β tends to be costly, our results shed light onto the circumstances in which investments in the many factors that the prior literature has identified as allowing a firm to increase E or θ might be beneficial.

II.1. Propositions concerning the value of increasing imitation breadth (β)

To understand under which conditions it is beneficial to increase the breadth of imitation, prior work on imitation and knowledge transfer points to three important contingencies: complexity, similarity, and time frame. These three variables have been prominently featured in the work on imitation. While effective imitation breadths have not been studied explicitly before, one can infer a number of relationships between these contextual variables and the value of increasing the breadth of imitation from the prior literature, allowing us to state a number of propositions. The aim of these propositions is twofold: they summarize the conventional wisdom and they serve as reference points for the results we obtain from our simulation analysis.

Complexity. Complexity is often seen as a key impediment to successful imitation. For instance, Inkpen (2008:451), in his study of GM's efforts to transfer knowledge from its joint venture with Toyota to its other plants noted that "GM struggled to appreciate the *systemic* nature of the TPS [Toyota Production System]" (emphasis added). Similarly, Szulanski (1996) highlighted Intel's "copy exactly" maxim: The only way Intel found it could transfer its complex knowledge from one plant to another plant was to copy exactly every aspect of a plant. In a similar vein, Milgrom and Roberts (1995) suggested that the complexity of Lincoln Electric's set of activities created a barrier for imitation.

While complexity can be conceptualized in a variety of ways, a number of scholars have built on the work of Simon (1962), who described complexity as arising from interactions among practices. This conceptualization has led scholars to study the transfer of "subnetworks" of interconnected activities

(Argote and Ingram, 2000), and of practices that differ in their degree of "discreteness" and "self-containment" (Williams, 2007).

Using a simulation model, Rivkin (2000) showed more formally how interactions among a firm's activities can create formidable barriers to imitation. In particular, Rivkin (2000) modeled firms that attempt to copy all practices of another, high-performing firm (i.e., E is maximal), with each practice successfully copied with probability $\theta = 0.9$. Rivkin (2000) found that as complexity increased, the performance difference between the imitator and the target firm grew. Thus, this paper highlighted the importance of complexity in the context of imitation but was silent on the value of different breadths of imitation, i.e., of changes in θ (and thereby implicitly, in β). In a related paper, Rivkin (2001) analyzed the advantages that a firm which replicates its practices might have over a firm that tries to imitate these practices. Rivkin (2001) showed that a fixed difference in the probability of correctly imitating practices creates particularly high benefits for intermediate values of complexity. (For instance, he modeled the cases of $\theta = 0.8$ for the replicator vs. $\theta = 0.2$ for the imitator). Thus, this paper illustrated a particular instance in which a higher θ (and, since E was fixed, implicitly a broader imitation breadth β) had a positive value.

Generally, then, prior studies imply that in the presence of many interdependencies, the payoff from imitation increases with β:

Proposition 1: In the presence of high interdependency, an increase in imitation breadth is valuable.

In the absence of many interdependencies, firms are likely to find good sets of practices without imitating, using local search alone (Rivkin 2000). While local search can eventually lead to high performance, it may take longer, however, than just copying high-performing practices from other firms. This suggests that the value of increasing imitation breadth, in the absence of many interdependencies, may be contingent on the time horizon that is relevant for the firm. Consequently, we turn to the time horizon as the next contingency factor.

Time Horizon. The length of the time horizon that managers do or should take into considerations when deciding on a course of action has been a perennial topic of organization and strategy research. For instance, the debate on "short-termism," i.e., whether managers have time horizons that are too short, or not, has continued for the last twenty years (Laverty, 1996). Regardless of whether managers have the "right" time horizon, heterogeneity exists with respect to the time horizon that managers employ. For instance, as Das (1987) found, even within the same large and established firm (a commercial bank) corporate executives differed in their time perspectives, which in turn affected the planning horizons that they considered. Likewise, managers, operating in firms at different life stages, are likely to have different time horizons. For example, an entrepreneur seeking to achieve a particular milestone to ensure the next round of financing, is more likely to choose actions that generate a quicker payoff than managers in firms with higher levels of slack, less fierce competition, or a focus on long-term goals. Similarly, given the significant "liability of newness" (Stinchcombe, 1965), managers in young firms often do not have the luxury to consider long-term horizons but need to be concerned with the short-term payoffs of their actions.

Similarly, in fast-paced environments, in which firms face "time-based competition" (Stalk and Hout, 1990), time horizons tend to be short. In these environments, actions that lead to quick performance improvements for a given environment are more valuable than actions that might lead to a higher performance in the long run, since the long run may never occur in this particular environment (i.e., the environment will have changed by then). As Lieberman and Asaba (2006:373) have noted, it is in these environments "where quick action is necessary [that] imitating others becomes an attractive decision rule."

More concretely, if different imitation breadths lead to different speeds at which firms can improve their performance, the value of increasing the breadth of imitation for a given firm is likely to depend on the time horizon that is relevant for the firm. How might the time horizon influence the value of increasing the breadth of imitation? If quick performance improvements are wanted, broadly imitating the set of practices of a high-performing firm would appear to have advantages, regardless of the degree of interdependency among the practices. The underlying logic is that if many interdependencies exist,

improvements can only be expected if broad imitation takes place (from Proposition 1). In contrast, if few interdependencies exist, the copied practices are likely to fit into the firm's existing set of practices and consequently are immediately beneficial for the imitator. As a result:

Proposition 2: For firms with short time horizons, an increase in imitation breadth is beneficial for all levels of complexity.

Alternatively, if the time horizon is long and few interdependencies exist, firms are likely to find good sets of practices themselves, with or without imitation (Rivkin 2000). Therefore, we would expect much less benefit from increasing β. This suggests:

Proposition 3: For firms with long time horizons, an increase in imitation breadth will generate little benefit if few interdependencies among the practices exist. (If many interdependencies exist, an increase in imitation breadth is beneficial, as per Proposition 1.)

Similarity. A third important factor for imitation is the degree of similarity between the firm that imitates and the firm that is being imitated. In this literature, it is valuable to distinguish between two types of similarity: context and firm similarity. Context similarity refers to similarity of factors that lie outside the control of a firm, such as industry maturity, but that affect the relationship between activities and performance. In particular, we define context similarity as the degree to which the same practices in one firm lead to a similar performance in a different firm. Thus, the more similar the performance that two firms achieve with the same set of activities, the more context-similar are the firms.

In general, the prior literature cautions against imitation given a lack of context similarity. For instance, Azoulay and Shane (2001) reported on the misguided imitation of McDonald's policy of non-exclusivity by a new franchisor. As Azoulay and Shane pointed out, the new franchisor failed to recognize that McDonald's, a mature franchise system, operated in a substantially different context (i.e., derived different benefits from the same actions) than the new franchisor, and hence the imitated practice was unlikely to be optimal. (Indeed, McDonald's actually had a different policy in its early stage of development.) Similarly, Argote and Ingram (2000) have noted that if new practices do not fit well a new

context, it will be even more problematic to copy a large set of choices than individual practices. Using a simulation model, Gavetti, Levinthal, and Rivkin (2005) obtained similar results: the value of analogies (transferring ideas from one context to another) decreased as the similarity between the contexts declined. In sum, these arguments imply:

Proposition 4: In the presence of low context similarity, increasing imitation breadth is harmful.

The notion of similarity has been used in the literature in a second, distinct way, which we call "firm similarity." While context similarity describes factors that lie outside the control of a firm, firm similarity is related to the similarity of the activities that firms have chosen to implement. This type of similarity is most meaningful if context similarity is already given; otherwise, the fact that two firms are similar with respect to certain practices would not be performance relevant. For instance, Baum, Li, and Usher (2000), in a study of imitation of location choices of nursing homes, used similarity in firm size as a measure of firm similarity, because in their sample, firms of a similar size competed using similar operating and strategic resources. Thus, in their study, firm size proxied for similarity in the set of activities that firms had employed.

The likely relationship between firm similarity and effective imitation breadth has not been well developed. Moreover, as we explain in more detail in Section IV.3, analyzing the effect of firm similarity on imitation breadth is challenging, particularly for long time horizons. As a result, we treat our analysis of this issue as purely exploratory.

Our simulation model will allow us to engage in a systematic analysis to probe into the validity of the prior propositions, to discover potential boundary conditions, and to help clarify underlying mechanisms that generate the relationships between breadth of imitation and performance under different conditions.

III. Model

Our goal is to assess the benefits of different breadths of imitation over different time horizons and between firms that differ in context- (or firm-) similarity, in the presence of varying degrees of

interactions among the practices that are copied. The simulation model we develop gives us direct control over these contingencies. In the following sections, we provide detail on the elements of the model.

III.1. Interactions among practices

The starting point of our model is the conceptualization of firms as systems of interdependent practices (Milgrom and Roberts, 1990; Porter, 1996; Siggelkow, 2002). Each modeled firm is assumed to make N binary decisions about how to configure its practices. N reflects the fact that a real firm must make numerous decisions, for instance, whether or not to increase the price for its product, or whether or not to locate a new store at a particular location. We represent a firm's configuration of practices as a string $\mathbf{d} = d_1 d_2 ... d_N$ with each d_i either 0 or 1. To account for the possibility of interactions among the practices of a firm, the payoff or contribution of a practice is modeled as a function of the practice itself and the resolution of K other, randomly chosen practices. K is thus a measure of interdependence.

The contribution function for a particular practice is a mapping of all relevant practice configurations to payoff values. For instance, in the case of two practices ($N = 2$), with $K = 1$, the contribution function for practice 1 would have four values: $c_1(00)$, $c_1(01)$, $c_1(10)$, and $c_1(11)$, denoting the payoff to practice 1 if both practices are configured as "0"; practice 1 is configured as "0" and practice 2 configured as "1"; etc. Likewise for $c_2(\cdot)$. The firm's overall performance P is then given by the average across the contributions. For instance, if the firm had chosen the practice configuration 01, then $P = \frac{1}{2}[c_1(01) + c_2(01)]$. Lastly, all entries of $c_i(\cdot)$ for each contribution function are drawn independently from a uniform distribution over the unit interval. In order to compare the performances of firms that live on different landscapes, performance is linearly scaled to the range 0-1, where 1 is the global peak of that landscape, and 0 the global minimum of that landscape.

This procedure for generating payoff functions is adapted from Kauffman's (1993) NK model, a model originally developed in the context of evolutionary biology. Numerous management scholars have used this procedure in recent years to generate payoff functions that can be employed to examine organizational search (for a review see Porter and Siggelkow, 2008). It is common to interpret such

payoff functions in terms of high-dimensional performance landscapes. Each of the N decisions constitutes a "horizontal" axis in a high-dimensional space, and each decision offers different options. Resulting from each combination of practices is a payoff for the firm, which is plotted on the vertical axis. The goal of organizational search is to find and occupy a high spot on this landscape, i.e., to select a combination of practices that, together, are highly successful. Interactions among practices cause the landscape to become rugged and multi-peaked, making the search for a high peak eminently more difficult (Kauffman, 1993; Levinthal, 1997).

III.2. Context and firm similarity

As discussed above, context similarity reflects the degree to which practices that work in one context are likely to work in another context. Consequently, we operationalize context similarity as the extent to which an identical set of practices implemented by different firms leads to similar levels of performance. To model this, we developed a technique to create correlated performance landscapes, with each firm "living" on its own landscape. When context similarity is maximal, all landscapes are identical, and hence, firms implementing the same set of practices achieve exactly the same level of performance. As context similarity decreases, the correlation between the performances generated by identical sets of practices decreases. At the extreme, with no context similarity, the correlation becomes zero.

In the model, this is implemented as follows. Context similarity, S, is modeled as a number between 0 and N that describes how many contribution functions are shared between firms. Hence, if $S = 0$ the landscapes are completely unrelated; if $S = N$, landscapes are identical. For intermediate values, S contribution functions are randomly picked and shared by all firms.

Firm similarity, in turn, has a very direct correlate in the model. Since firm similarity captures the degree to which firms have made similar performance-relevant decisions, the degree of firm similarity, *H*, is given by the number of shared elements between Firm 1's decision vector $\mathbf{d_1}$ and Firm 2's decision vector $\mathbf{d_2}$. For instance, if $\mathbf{d_1}$ = 0000 and $\mathbf{d_2}$ = 0100, then the firm similarity between the two firms would be $H = 3$.

III.3. Local and imitative search

The firms in our model attempt to find combinations of practices that yield high performance. To this end, they can engage in local or imitative search. Building on the work of Simon on bounded rationality (Simon, 1947), search is often seen as occurring via evaluation of alternatives that are close to the current status quo. In particular, we model local search as testing all local alternatives, i.e., alternative practice configurations that differ in only one practice, and choosing the one with the highest performance. For instance, if $N = 2$, and the status quo choices are 00, then the firm would evaluate alternatives 01 and 10 (but not choice 11), and choose the configuration with the highest performance. For instance, if $P(01) > P(00) > P(10)$, the firm would chose 01. This configuration would then be the starting point for further search in the next period. If the firm does not find any alternative with higher performance, it retains the status quo practices. An intrinsic problem of local search is that it can get stuck on a low, local peak before ever reaching the higher peaks of the landscape (Levinthal, 1997).

To escape the bounds of local search, firms can engage in imitating other organizations. We model two different types of imitative search: discriminant and non-discriminant imitation. Both types of imitation involve copying practices from a template firm to a recipient firm, "moving" the recipient firm to a new position on its landscape. Given that the leading firms in an industry often serve as templates for imitation, the best performing firm among the M simulated firms is chosen as the template firm.

Discriminant and non-discriminant imitation differ in how much firms know about the context similarity between the template and the recipient firm. Discriminant imitation assumes that the recipient firm knows the identity of the practices for which the two firms have the same contribution functions. In this case, if the recipient firm copies β practices, it will randomly select β practices from the S shared dimensions and copy them. With non-discriminant imitation, the recipient firm does not know which dimensions it shares with the template firm. In this case, it randomly chooses β practices from the N practices of the template firm and copies them. The minimum value that β can take is zero (nothing is copied). The maximum value for β under discriminant imitation is S (copy everything that is shared),

and N under non-discriminant imitation (completely copy the other firm's practices). (We would expect that extremely high values of β are more likely to arise in an intra-firm setting.) In the model, either way of imitating takes one time period to happen.

In the simulation, firms perform a local search in each time period until they cannot improve their performance, i.e., get stuck. Once stuck, a firm will perform an imitative search in the following period, that is, the firm will copy β practices from the highest performing firm. After this imitation, the firm continues to perform local searches until stuck again, at which point it would imitate again, etc. The imitation can be seen as a long jump in the landscape, allowing firms to explore strategies that are quite distant from the current one. The subsequent local searches can be interpreted as more fine-grained adjustments to the new set of routines to further improve performance. Thus, our model of imitative search contains the two mechanisms of knowledge transfer described by Williams (2007), replication and adaptation: First, a firm replicates precisely β practices of a target firm; subsequently, the firm adapts (at least locally) its entire set of practices, which may be necessary due to the newly copied practices.

In sum, one should note that in our model imitation differs from local search in two important aspects. First, imitation allows the firm to change more than one practice at the same time. Second, imitation can lead to the implementation of practices that turn out to be detrimental to performance. (Recall, with local search firms assess alternatives before implementing them, thereby avoiding any new practices that might reduce performance.)

III.4. Implementation Details

At the beginning of each simulation run, new landscapes are created and firms are placed on random locations, i.e., are endowed with random practice configurations. Since we are interested in studying short-run and long-run performances generated by various breadths of imitation, we measure the performance of firms at four different time periods: after 4, 10, 100, and 1,000 periods. In our model, by period 100 practically all firms achieve a steady state (they are either stuck or they engage in a repetitive cycling behavior). As we show, performances achieved by period 1,000 are essentially identical to

performances achieved by period 100. In other words, the results that are observed at $T = 100$ can be understood as a long-term equilibrium situation. For environments that are stable enough and with firms that have long time horizons these are the relevant results.

In contrast, at $T = 4$ and at $T = 10$ some performance-improving adaptation has occurred but performance is still far from achieving a plateau. By $T = 4$, firms have engaged in (at most) one imitative attempt; by $T = 10$, many firms have had time to engage in two or three imitation attempts. Consequently, we interpret results at these time points as short-run scenarios. These short-run results also illustrate what happens in a turbulent environment. In a turbulent environment, the relationship between practices and performance frequently resets. Thus, a firm would face (somewhat) different landscapes every, for instance, 10 periods. As a result, a firm would only be interested in what performance it could achieve by period 10 in the "first" landscape it faces, and then again in period 10 in the "second" landscape, etc.

Even though the model is relatively parsimonious (it depends on only five parameters) it would be unwieldy to describe the results of all combinations and gradations of these five parameters. Hence, we resort to describing representative cases that we identified after carefully exploring the space of possibilities and after testing for the robustness of the findings (robustness is further discussed in Section IV.6). Specifically, we keep the number of decisions (N) fixed at 12; the number of firms (M) fixed at 5; and landscape complexity (K) and context similarity (S) take a low and a high value (2 and 10 for K, and 4 and 12 for S). The parameter of central interest, the extent of imitation (β), is varied over its entire range in every model that we analyze.

In order to report results that are caused by the nature of the model and not by any specific randomly generated landscape, the results for each set of parameters were computed by averaging 1,000 simulation runs. All reported results are significant at the 5% level.

IV. Results

In the first set of simulations, the focus is on firms that are context similar, while in the second set, we study firms that operate in different contexts. Within each set, we first model the case of low complexity,

followed by high complexity.

IV.1. Imitating firms operating in similar, low complexity contexts

We start the analysis by studying the effects of imitation in a "simple" environment: firms operate in similar contexts, and each firm's practices are not very interdependent. In particular, firms face identical landscapes ($S = 12$, i.e., they share all 12 contribution functions that define performance) and the complexity of the landscapes is low ($K = 2$).

Figure 1 shows the effect on performance of varying the breadth of imitation, with β going from 0 to 12, i.e., from not imitating at all to completely imitating the best performer. The effect is measured for the short ($T = 4$ and 10) and long term ($T = 100$ and 1,000). We will first discuss the short term curves of Figure 1. In the very short-run, at $T = 4$, imitation breadth does not have a large impact on performance. By this time, almost half of all firms have only engaged in local search; this implies that their performance is equal to the performance of firms that do not engage in imitation ($\beta = 0$). Unless firms copy all (or almost all practices), their performance does not improve significantly. However, once firms had a few opportunities to copy and to adjust their practices, at $T = 10$, the relationship between imitation breadth and performance has a clear upward slope: the more practices are copied, the higher the performance of a firm. One should note that at very high levels of imitation (when $\beta > 8$) the performance improvement due to imitation is slight. This happens because, due to the low complexity, even if valuable practices are not copied, they can be discovered by the imitator by performing local searches after the imitation. These results are consistent with the conventional wisdom summarized in Proposition 2: for firms with short time horizons, an increase in β appears to be beneficial.

< Insert Figure 1 about here >

The long term curves of Figure 1 tell a somewhat different story. In contrast to the short-term curves, the long-term curves ascend very rapidly, yet then level off. The marginal payoff of going from not imitating at all to imitating one or two practices is very high, since given enough time, repeatedly imitating just one or two practices every time the firm is stuck suffices to dislodge the firm from its

current local peak enabling it to find a higher peak. A real-world example would be a firm that has ceased to improve its profitability and decides to imitate a simple customer-service process from a similar competitor. Sometimes copying such process will be a mistake; but because the extent of the imitation is small, it is easy to reverse the changes. Other times the copied process will be equivalent to the process already in practice, so there will be no change in performance. And in other cases, the imitated process, either immediately or after a few local tweaks, will improve performance. Thus, on average, in these circumstances (high context similarity and low complexity) even copying few practices will improve performance.

These results offer a slight adjustment to Proposition 3. Even with few interdependencies, it pays off for firms to increase their imitation breadth from "not copying at all" to "copying a few practices, repeatedly." However, as long as firms imitate a few practices, so that they are able to dislodge themselves from their current set of activities (and are not immediately "sucked" back into their old set of practices after imitation), performance is relatively unaffected by imitation breadth: the long-run performance of copying a few or very many activities is very similar, consistent with Proposition 3.

IV.2. Imitating firms operating in similar, high complexity contexts

We now turn our attention to situations with rich interactions among the practices. Figure 2 is analogous to Figure 1 but under a highly complex environment ($K = 10$). With a high degree of interdependence among a firm's practices, very different results arise. Considering short-run performance first, we can observe a marked non-monotonicity: a moderate level of imitation is worse than both no imitation at all and a high breadth of imitation. This effect arises because given the large number of interdependencies among the practices, the newly imported practices, if unmodified, are likely to be incompatible with the rest of the firm's practices. In the short run, the firm does not have enough time to engage in all necessary adaptations and its performance suffers. Only when fairly large "chunks" of practices are imported does the short-run effect on performance turn positive, because many of the interactions of the imported practices will fall within that chunk. Figure 3 illustrates this effect by

measuring the percentage of imitations that end up decreasing performance in the short run. The measure includes not only the performance change due to the imitation, but also due to the subsequent local searches (until the next imitation). In other words, we measure the effect of the imitation plus all the subsequent local refinements. As can be seen in Figure 3, the relationship between the percentage of imitations that end up decreasing performance and the breadth of imitation (β) has an inverted U shape. Indeed, Figure 3 mirrors well the performance line for $T = 4$ in Figure 2.[1]

These results offer some important limitations to Propositions 1 and 2. In the presence of high interdependency, an increase in the breadth of imitation (β), and equivalently, an increase in the probability that any individual practice is faithfully copied (θ), is not always valuable since intermediate-sized imitations are not beneficial. In the short run, firms are better off engaging only in local search, or, if possible and not too costly, in engaging in large-scale imitations that capture many of the interdependencies among the copied elements. Thus, in the short-run, an increase in imitation breadth is only valuable at higher levels of β . As an interesting corollary one might note that if high values of β are more easily achieved by knowledge transfer within an organization than across organizations, firms that are mainly concerned with short-run performance or operate in turbulent environments may have an incentive to move the boundary of the firm such that the knowledge source lies inside the organization.

< Insert Figures 2 and 3 about here >

The long-run curves in Figure 2 also show a distinct non-monotonicity. Interestingly, what is worst for short time horizons (an intermediate breadth of imitation) is what is best for long time horizons. What explains the high long-run performance generated by intermediate-sized imitations? In the presence of many interdependencies, the performance landscapes become very rugged, offering many peaks on which firms may get stuck. This ruggedness increases the value of exploration. At one extreme, if a firm engages in only small imitations, a firm may not be able to dislodge itself from its current peak, that is, from its

[1] In contrast to the low complexity case, with high complexity almost all firms have completed at least one imitation attempt by $T = 4$.

current set of practices. (More formally, it is unable to leave the basin of attraction of its current local peak.) At the other extreme, as a firm copies a very large number of practices, its overall exploration of the landscape declines as well. This happens because by copying large numbers of practices from the current best performer, the firm is limiting its possibilities of finding a better peak than the one occupied by the current best performer. In the extreme, if the whole industry is always imitating what the best performer does, there will be many unexplored opportunities because no one will dedicate resources to innovating, potentially leading to pathological scenarios (Strang and Macy, 2001). In the intra-firm context, if a corporation has decided on its "best practices" that every unit (e.g., franchise outlet) has to copy exactly, no unit will be likely to improve on the identified "best practices." Figure 4 confirms this intuition. Here we plot the average number of configurations of practices (or, equivalently, locations on the landscape) that firms experience over their 1,000-period life span. As shown, exploration is maximized in the long run by intermediate sized imitations – large enough to allow a firm to escape from its current configuration, but small enough to avoid locking in the firm too quickly on the currently best configuration of the industry (which may not be the best achievable). These results provide a further limitation to the conventional wisdom summarized in Proposition 1. In the presence of many interdependencies, increasing imitation breadth beyond an intermediate level may not be beneficial in the long-run, and actually be harmful to performance.

< Insert Figure 4 about here >

As noted above, imitative search allows the firm to engage in a long jump, creating the possibility for the firm to escape from its current set of practices. While imitation is one way to create such long jumps, other mechanisms do exist. For instance, a firm could allow managers to experiment with, more or less radical, new sets of practices. To start assessing whether such long jumps have a similar effect to imitation, we modified our model slightly. Firms still search locally until they are stuck; then, they change R randomly picked practices; then they again search locally, etc. As the results reported in Figure 5 show, a strategy of random long jumps generally does not lead to short- nor long-run performance improvements.

The only exception is $R = 1$ in the long-run. In this case, the firm is at least sometimes able to dislodge itself from its current set of practices.[2] In all other cases, random jumps per se are not sufficient.

<Insert Figure 5 about here >

In sum, when many interdependencies among practices exist, our results identify a number of boundary conditions to the conventional wisdom that broader imitation is generally valuable. In the short-run, imitation is only valuable when it is very broad, i.e., when β is close to the maximum. If firms are unable to copy very broadly, it appears to be more beneficial for firms to rely solely on local search rather than risking performance losses that arise when firms copy only few practices that have high interdependencies with existing practices. In the short-run, firms do not have enough time to readjust their existing practices and to undo potentially poor imitation attempts. In the long-run, however, intermediate-sized imitation, which was so harmful in the short-run, turns out to be the most valuable imitation breadth, because it creates the greatest amount of search.

IV.3. The effect of firm similarity

As noted in Section II, the concept of "similarity" has been used in the literature in two different ways, denoting both context similarity ("To what degree are the landscapes in which two firms operate similar?") and firm similarity ("To what degree are the sets of practices that two firms employ similar?"). This subsection covers the case of different degrees of firm similarity, while the next covers the case of different degrees of context similarity.

In short, the question we try to address is, "How does the value of different levels of imitation breadth change as the imitation target becomes more similar to the imitator?" To address this question, we keep track of the firm-similarity measure between the imitator and the target firm. In this analysis, once stuck, a firm imitates β of the practices that are different between itself and the highest performing firm. For

[2] Changing one practice leads for sure to an immediate performance decline, since the firm was sitting on a local peak. In most cases, the firm will simply move back to the local peak on which it was located. However, if a higher-performing configuration is located two steps away from the current local peak, this configuration can now be reached. The random jump forces the firm into the "valley" between the local peak and the higher-performing configuration; a move the firm would not have made under purely local search.

instance, if the imitator and the highest-performing firm differ in four practices, the imitator could copy up to four practices. One problem arises in this analysis. Many combinations of imitation breadth and similarity are infeasible in the long run. For instance, the strategy of "copy many practices of best-performers that are very different" can not be applied many times in a row. After the firm has once copied many practices, and has engaged subsequently in local search, it is very unlikely to be stuck again with a set of practices that is very different from that of the best performer, i.e., it could not follow again its strategy of "copy many practices of best-performers that are very different." To keep the analysis tractable, so that we do not have to consider, for instance, different sequences of imitations of best-performers that have different similarities, we restrict our attention to the very short-run in which at most one imitation has happened, i.e., $T = 4$.

Figure 6 reports the results of the two cases of low and high complexity. Consider the line labeled $H = 9$. This line represents the performance at $T = 4$ of firms that, when faced with a best-performer that had nine similar and three ($= 12 - 9$) different practices, copied β = 0, 1, 2, or 3 of the different practices of the best-performer. (For ease of illustration, the figures include only the cases of H = 9, 8, 7, 5, and 3.) As first observation, we may note that in all cases full imitation yields the highest performance. This is very much consistent with the short-run results we saw in the previous subsection: if time is of essence, it pays to completely copy a high-performer's set of activities. Second, we observe a marked non-monotonicity for all levels of firm similarity: increasing β can lead to a performance decline, except when the fraction of copied practices is very high.

Lastly, a comparison of the two panels shows that firm similarity creates pronounced performance differences over a larger range of β when few interdependencies exist. In the presence of few interdependencies, the performance penalty created by intermediate-sized imitations is particularly strong when firms are copied that are very dissimilar. (The bottom of the U-shaped relationship between β and performance is much lower for low levels of *H* than for high levels of *H*.) In contrast, in the very short run, firm similarity plays much less of a role in complex environments. With many interdependencies,

even small differences in practices can have large performance impacts. Consequently, in this case, imitation, unless it is very broad, is equally detrimental regardless of firm similarity. (For longer time horizons, we again find that the performance penalty created by intermediate-sized imitations is particularly strong when firms are copied that are very dissimilar; results available from the authors.)

In sum, our results offer two insights into the relationship between firm similarity and effective imitation breadth. First, in the very short run, regardless of firm similarity, increases in imitation breadth can lead to a performance decline. Second, the more dissimilar the imitation target, the more detrimental it is to copy only an intermediate number of practices, i.e., the more pronounced the U-shape.

< Insert Figure 6 about here >

IV.4. Imitating firms operating in different contexts

In the previous sections, imitation was performed among firms that operated on the same performance landscape. In this section, we study firms that operate in different contexts, i.e., that live on different performance landscapes. An understanding of the effects of imitating firms that face different performance landscapes might be important, because in many circumstances firms may not have directly comparable peers. For example, an early startup in an innovative industry (e.g., iRobot in the household robots industry), an organization facing a unique task (e.g., the Justice Prisoner and Alien Transportation System in the US), and many oligopolistic firms serving unique markets (e.g., Hindustan Lever selling a broad range of consumer goods in India) do not have peers facing exactly the same landscape, but firms with varying degrees of context similarity which may be candidates for imitation.

We explore the effects of imitation among firms that face performance landscapes that are neither completely identical nor completely different, but have an intermediate degree of context similarity. In the following analyses, this is modeled by reducing the number of shared contribution functions (S) from 12 to 4. Our analysis follows the same scheme as in the previous section: we explore the effects on performance of varying the extent of imitation (β) in the short and the long run, on both low and high complexity landscapes (Figure 7). It is important to note that the range of values that imitation (β) can

take is constrained by the degree of similarity (S), since discriminant imitation is only defined for $\beta \leq S$; if firms' contexts are only similar for S practices (i.e., they share S contribution functions) they can at most imitate S practices. (In the next section, we will study non-discriminant imitation, which does not have this constraint.)

< Insert Figure 7 about here >

The most striking feature of these new figures is that for each one of the curves shown, the best performance is generally achieved at $\beta = 0$. In other words, when context similarity is low, the best decision is to avoid imitating others and to focus on local search. This happens because even if a practice has only few interactions, most of these interactions will be with practices that are not shared. In other words, the imported practices will not be compatible with the unique characteristics of the recipient. This is very much akin to the short-run results we saw for imitation among context-similar firms in complex environments.

Why does intermediate-sized imitation not work even in the long run, as it did for context-similar firms? With firms operating in different contexts, the problem is that the range of practices that can be copied is narrower. With firms in similar contexts, over time up to 12 practices of a leading firm can be copied (in chunks say of $\beta = 4$). With firms in different contexts, a total of only S (in our case 4) practices can be copied. If the best performer remains the same, an imitating firm with $\beta = 4$ would always copy the same set of practices from the best performer. In this case, imitation does not only fail to import valuable practices, but it also fails to create variety of new ideas. In sum, for discriminant imitation, our results are consistent with the conventional wisdom captured in Proposition 4. When a firm copies practices from a firm that operates in a very different context, imitation is likely to lead to performance penalties. Moreover, this effect tends to be larger, the larger the breadth of imitation.

One should note, however, that our results do not imply that imitation among firms that operate in different contexts is always a bad idea. Apart from lucky imitations, there are at least two cases in which the previous results do not hold: in very low complexity landscapes (e.g., $K = 0$) and under modular

interactions. If K is very low, the imported practices are completely self-contained, so a valuable imitated practice continues to be valuable in the receiving firm. (Indeed, when we run our simulation with $K = 0$, all firms regardless of β reach the global peak. Firms with higher values of β reach the global peak faster, though. Results are available from the authors.)

The second case not covered in the previous figures is imitation of shared "modules." Modularity means that all (or more generally, almost all) interactions of a set of practices fall within that set of practices. Or conversely, modularity means that there are no (or only few) interactions between different modules. Hence, an imitation of a shared module can be thought of as an imitation between smaller firms (the size of the module) that share all their practices. One example of imitating a shared module could be the Justice Prisoner and Alien Transportation System adopting the engine maintenance practices of American Airlines. Clearly both organizations are different, but probably their engine maintenance practices can be imitated because they are shared and modular. Hence, the results of the previous section concerning imitation between context-similar firms are applicable to imitation between shared modules of otherwise context-dissimilar firms.

IV.5. Non-discriminant imitation

Until now, imitation has been modeled as discriminant imitation which assumed that the imitating firm has a clear understanding of which practice dimensions are shared with the target firm. What happens if the imitating firm does not clearly know what is shared? Can imitation still add value under these circumstances?

When firms operate in similar contexts, i.e., $S = N$, non-discriminant imitation renders into discriminant imitation by definition (both copy β out of N practices). Thus in this case, we obtain the same results as shown in Figures 1 and 2. The results of non-discriminant imitation between firms operating in different contexts are shown in Figure 8. Given that the identity of the shared practices is not known, firms might copy up to 12 practices from the best performer. Thus, β can lie between 0 and 12. The curves for short time horizons lead to the same conclusion as under discriminant imitation: for firms

operating in different contexts, no imitation dominates any amount of imitation. For short time horizons, the newly adopted practices are likely to create misfits with the existing practices and lead to performance deterioration. In the case of high interdependency, we can observe that increases in β can be beneficial at high levels (see the right panel), yet the achieved performance is still below that of pure local search.

For long time horizons, however, results differ significantly. Even with low context similarity, small breadths of imitation and enough time allow the firm to achieve a higher performance level than that achieved by pure local search. This result holds for both the low- and the high interdependency case. At first, this result may seem surprising. Why is it that imitating small parts of firms that operate in markedly different contexts can be beneficial if firms do not constrain themselves to copying only from the set of shared practices? The explanation is that in loosening the constraint of only copying shared practices, valuable variety is generated. The argument is analogous to the explanation of why under high context similarity and high complexity intermediate-sized imitations were useful: importing an intermediate number of new practices is valuable in the long run because the new practices can dislodge the firm from a local peak, increase its exploration, and enable it to find a new, higher-performing configuration of practices. Figure 9, showing the average number of locations visited by firms supports this intuition; intermediate-sized imitation (β around 3) leads to the largest extent of exploration.

This last finding leads us to a speculative observation. The academic literature frequently tends to question the value of "airport-store business books," yet there seems to be a large market for them. Are managers just naïve, or do these books actually provide value? Our findings suggest that as long as only intermediate-sized changes are implemented, these changes can still provide value in the long run even if the new ideas may not be valuable per se. The effect of spurring the firm to change its sets of practices can renew the firm's search effort, and given enough time, allow the firm to reach a higher-performing set of practices.

Overall, our results regarding non-discriminant imitation qualify Proposition 4. When operating in different contexts, increasing imitation breadth to intermediate levels can actually be beneficial in the

long run. For the short-run, and for broad imitation breadths in the long run, we do find, however, consistent with Proposition 4, that increases in imitation breadth are harmful for firms operating in dissimilar contexts.

< Insert Figures 8 and 9 about here >

IV.6. Robustness

We studied the robustness of our findings regarding the value of different breadths of imitation (β) with respect to the four other parameters of the model (complexity, K ; number of firms, M ; dimensionality of the landscape, N ; and similarity, S). In addition, we relaxed the assumption that managers can only perform local searches, and we explicitly modeled turbulence in the environment to show that our short-horizon results are indeed representative of turbulent environments. In general, our results are robust to changes in these assumptions. The full set of robustness results are available in an on-line appendix.

V. Discussion

We used our simulation model to systematically explore the value of different imitation breadths. In particular, we studied how three contingency factors – complexity, similarity, and time horizon – influence the benefit of increasing imitation breadth β . Our results show that the value of increasing β is dependent on these three contextual factors in non-trivial ways. In particular, we find a number of non-linearities in the relationship between β and performance which were not predicted by the extant literature on imitation. These non-linear relationships imply that the value of increasing imitation breadth is also a function of its own level: In those cases in which we found an inverse U-shape relationship, the benefit of increasing β is only positive when β starts at a low level; once β reaches a high level, further increases of β lead to a performance decrease. The converse holds true for U-shaped relationships. Figure 10 summarizes our results, using the three contingency factors and the starting level of β . The Figure shows under which conditions it is valuable to increase imitation breadth (the grey cells that include a plus

sign), and those conditions in which it is detrimental to increase imitation breadth (the black cells including a minus sign). For low context similarity, we report results for discriminant and non-discriminant imitation separately; for high context similarity, results for both types of imitation are identical.[3]

Testable hypotheses and relationship to the propositions summarizing the extant literature. Each of the cells of Figure 10 can be interpreted as a hypothesis suggested by the results of our model. Rather than stating a large number of hypotheses, we summarize the results as follows:

H1: When context similarity is high, an increase in imitation breadth is generally beneficial, with two important exceptions:

(a) in the presence of high complexity and with a short time horizon, increases in imitation breadth from very low levels to intermediate levels are harmful, and

(b) in the presence of high complexity and with a long time horizon, increases in imitation breadth at very high levels are harmful.

H2: When context similarity is low, an increase in imitation breadth is generally harmful, with two important exceptions:

(a) regardless of complexity, with a long time horizon, increases in imitation breadth from very low levels to intermediate levels are beneficial if imitation is non-discriminant, and

(b) in the presence of high complexity and a short time horizon, increases in imitation breadth from very low levels are beneficial if imitation is non-discriminant. (Yet performance may still be worse than that achieved by local search alone.)

One should also note that the four exceptions are the key areas in which our results deviate from the conventional wisdom summarized in Propositions 1–4.

With respect to firm similarity, our results point to one further hypothesis:

H3: With short time horizons, the more dissimilar the copied firm, the more harmful is a strategy of intermediate-sized imitation attempts.

< Insert Figure 10 about here >

Mimicking versus dislodging. At the conceptual level, our paper unearths two different roles imitation can play: imitation as a way to quickly *mimic* high-performers, and imitation as a way to rekindle search

[3] In one condition, the relationship between increases in β and performance is flat, which is denoted by a zero. In one other condition, the relationship is positive, yet the benefit is below that achieved by local search. We denoted this case with a plus sign in parentheses.

by *dislodging* a firm from its current set of practices. The following examples clarify this distinction. A firm in a low complexity environment with a high-performer operating in a similar context should opt to copy, if possible, a substantial number of practices from the high-performer, that is, to mimic and to lock in on the activities of a high-performing firm. On the other hand, a firm on a complex landscape with enough time to experiment should opt to import only small chunks of practices from others, that is, to dislodge itself repeatedly from its current set of practices.

Even if mimicking and dislodging are both achieved by imitating others, they are conceptually and pragmatically two distinct strategies. To start, they both require radically different organizational commitments. Mimicking requires imitation of large chunks of practices with high fidelity; hence it requires mechanisms to assure the quality of extensive imitation attempts. As we have seen in Figure 5, random, unguided long-jumps tend not to be effective. On the other hand, dislodging does not require high fidelity nor copying large chunks of practices, but an organizational commitment to repeatedly adapt to small changes, i.e., to remain plastic. Dislodging requires imitating small chunks and then adapting subsequently in order to reach a better configuration of practices. It also requires alertness, because if the firm cannot discover a better configuration subsequent to the imitation, it has to rapidly undo the changes caused by the imitation. Consequently, the organizational commitments of both imitation strategies are opposed: one requires strict enforcement of rules while the other is more an exercise in creativity. This implies that if a firm tries to mimic but is in a "dislodging mode," the imitation might likely fail, and vice versa. A further difference is that both mechanisms react differently to practices that are difficult to copy. For example, practices that involve tacit knowledge, which are difficult to copy with fidelity, would seriously undermine the success of mimicking but not of dislodging.

Implications for exploitation and exploration. A comparison of local search, mimicking, and dislodging casts new light on the classic discussion of exploration and exploitation (March, 1991). "Dislodging" has characteristic features of exploration: First, it allows a firm to try a large range of new combinations of practices. As Figures 4 and 9 showed, intermediately sized imitation attempts led to the broadest search across the performance landscapes. In other words, dislodging creates variance. Second,

dislodging can lead a firm to escape from its current set of practices (the "basin of attraction" of its current local peak), and yield extensive change over time. In contrast, firms that are only engaged in local search, i.e., that do not imitate at all ($\beta = 0$) would correspond to firms that are exclusively performing an exploitation of their current set of activities; these firms improve upon their starting set of practices, but never veer off very far.

"Mimicking," however, defies an easy classification in terms of exploration and exploitation. On one hand, it shares characteristics of exploration, as it allows the firm to escape from its current local peak and to change radically. On the other hand, mimicking – in contrast to exploration – does not generate broad search, as the right hand sides of Figures 4 and 9 attest. A mimicking firm tries to quickly lock in on a set of new practices without broad search. In a sense, a firm that mimics tries to reduce variance and to "exploit" the practices found by a high-performing firm. Thus, mimicking appears to form a hybrid search strategy which cannot be cleanly mapped to either exploration or exploitation. We believe that the field may benefit from further discussion aimed to elucidate whether exploration and exploitation are the extremes of a continuum or whether further search strategies exist, such as mimic, which may fall outside this uni-dimensional model.

Extensions. As every model, our simulation set-up has a number of limitations which point toward further extensions. For instance, our results imply that imitating context similar firms can lead to much higher improvements than imitating context dissimilar firms (as seen when comparing the peaks achieved in Figure 8 to those in Figures 1 and 2). This is important, because in the real world, firms may not know which other firms are context similar to them since context similarity is costly to assess (e.g., the choice set may be large, be geographically disperse, or context similarity may not be evident unless a deep analysis is carried out). Hence, it may be profitable for firms to strive to find context similar peers before starting any imitative effort. A possible extension of our model could include a cost of assessing how context similar the focal firm is to a potential target.

VI. Conclusion

Our present work offers four main findings that extend the prior literature on imitation. First, imitation can serve two different functions. It can either be used to attempt mimicking the exact configuration of practices of other high-performing firms, or it can be used to dislodge a firm from its current set of practices, generating broader exploration. Second, an increase in imitation breadth, and consequently investments in actions that increase the probability of faithfully copying individual practices, is not always valuable. Third, in the presence of rich interactions among practices, the most effective imitation breadth differs dramatically depending on time horizon. Imitation breadths that are effective for long time horizons yield very poor performance for short horizons, while effective breadths for short horizons are not particularly valuable for long horizons. Fourth, imitation from context dissimilar firms can be harmful, especially for short time horizons. For long time horizons, small doses of imitation can, however, be helpful. In this case, it is not the direct imitation of a practice that is valuable, but the dislodging and subsequent search that can yield performance improvements in the long run.

Imitation – both revered and despised by practitioners – is a widespread phenomenon with important competitive implications. This paper has aimed to contribute to the discussion of the value of imitation by putting into place an analytical framework to study imitation in a rigorous manner, and using it as a stepping stone to create, in a disciplined way, new theoretical insights. This paper has tried to move toward a contingency theory of imitation, emphasizing that imitation is not a context-independent yes/no choice, but rather that imitation falls along a scale, with the value of increasing the breadth of imitation being contingent on complexity, similarity, and time frame.

VI. References

Argote, L. 1999. *Organizational learning: Creating, retaining, and transferring knowledge*. Kluwer Academic, Boston, MA.

Argote, L. and Ingram, P. 2000. Knowledge transfer: A basis for competitive advantage in firms. *Organizational Behavior and Human Decision Processes* **82**: 150-169.

Azoulay, P. and Shane, S. 2001. Entrepreneurs, contracts, and the failure of young firms. *Management Science* **47**(3): 337-358.

Baum, J. A. C. and Ingram, P. 1998. Survival-enhancing learning in the Manhattan hotel industry, 1898-1980. *Management Science* **44**: 996-1016.

Baum, J. A. C., Li, S. X., and Usher, J. M. 2000. Making the next move: How experiential and vicarious learning shape the locations of chains' acquisitions. *Administrative Science Quarterly* **45**: 766-801.

Cohen, W. M. and Levinthal, D. A. 1990. Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly* **35**: 128-152.

Darr, E. D., Argote, L., and Epple, D. 1995. The acquisition, transfer, and depreciation of knowledge in service organizations: Productivity in franchises. *Management Science* **41**: 1750-1762.

Das, T. K. 1987. Strategic-planning and individual temporal orientation. *Strategic Management Journal* **8**: 203-209.

Gavetti, G., Levinthal, D. A., and Rivkin, J. W. 2005. Strategy making in novel and complex worlds: The power of analogy. *Strategic Management Journal* **26**: 691-712.

Inkpen, A. C. 2008. Knowledge transfer and international joint ventures: The case of nummi and general motors. *Strategic Management Journal* **29**: 447-453.

Katz, R. and Allen, T. J. 1982. Investigating the Not Invented Here (NIH) Syndrome: A look at the performance, tenure, and communication patterns of 50 R&D project groups. *R&D Management* **12**(1): 7-19.

Kauffman, S. A. 1993. *The Origins of Order: Self-Organization and Selection in Evolution*. Oxford University Press, New York.

Laverty, K. J. 1996. Economic "short-termism": The debate, the unresolved issues, and the implications for management practice and research. *Academy of Management Review* **21**: 825-860.

Levinthal, D. A. 1997. Adaptation on rugged landscapes. *Management Science* **43**: 934-950.

Lieberman, M. B. and Asaba, S. 2006. Why do firms imitate each other? *Academy of Management Review* **31**: 366-385.

March, J. G. 1991. Exploration and exploitation in organizational learning. *Organization Science* **2**: 71-87.

Milgrom, P. and Roberts, J. 1995. Complementarities and fit: Strategy, structure, and organizational change in manufacturing. *Journal of Accounting & Economics* **19**: 179-208.

Milgrom, P. R. and Roberts, J. 1990. The economics of modern manufacturing: Technology, strategy, and organization. *American Economic Review* **80**: 511-528.

Nutt, P. C. 1998. How decision makers evalutate alternatives and the influence of complexity. *Management Science* **44**: 1148-1166.

Porter, M. E. 1996. What is strategy? *Harvard Business Review* **74**(6): 61-78.

Porter, M. E. and Siggelkow, N. 2008. Contextual interactions within activity systems and sustainability of competitive advantage. *Academy of Management Perspectives* **22**(2): 34-56.

Reagans, R. and McEvily, B. 2003. Network structure and knowledge transfer: The effects of cohesion and range. *Administrative Science Quarterly* **48**: 240-267.

Rivkin, J. W. 2000. Imitation of complex strategies. *Management Science* **46**: 824-844.

Rivkin, J. W. 2001. Reproducing knowledge: Replication without imitation at moderate complexity. *Organization Science* **12**: 274-293.

Schnaars, S. P. 1994. *Managing Imitation Strategies: How Later Entrants Seize Markets From Pioneers*. Free Press, New York.

Siggelkow, N. 2002. Evolution toward fit. *Administrative Science Quarterly* **47**: 125-159.

Simon, H. A. 1947. *Administrative Behavior*. MacMillan, New York.

Simon, H. A. 1962. The architecture of complexity. *Proceedings of the American Philosophical Society* **106**: 467-482.

Stalk, G. and Hout, T. M. 1990. *Competing Against Time: How Time-Based Competition is Reshaping Global Markets*. Free Press, New York, NY.

Staw, B. M. and Epstein, L. D. 2000. What bandwagons bring: Effects of popular management techniques on corporate performance, reputation, and CEO pay. *Administrative Science Quarterly* **45**: 523-556.

Stinchcombe, A. L. 1965. Organizations and social structure. In J. G. March (Ed.), *Handbook of Organizations*: 142-193. Rand McNally, Chicago.

Strang, D. and Macy, M. W. 2001. In search of excellence: Fads, success stories, and adaptive emulation. *American Journal of Sociology* **107**: 147-182.

Szulanski, G. 1996. Exploring internal stickiness: Impediments to the transfer of best practice within the firm. *Strategic Management Journal* **17**(Special issue: Knowledge and the firm): 27-43.

Szulanski, G., Cappetta, R., and Jensen, R. J. 2004. When and how trustworthiness matters: Knowledge transfer and the moderating effect of causal ambiguity. *Organization Science* **15**: 600-613.

Williams, C. 2007. Transfer in context: Replication and adaptation in knowledge transfer relationships. *Strategic Management Journal* **28**: 867-889.

Winter, S. G. 1995. Four Rs of profitability: Rents, resources, routines and replication. In C. Montgomery (Ed.), *Resource based and evolutionary theories of the firm: Towards a synthesis*: 147-178. Kluwer Academic, Boston, MA.

Zander, U. and Kogut, B. 1995. Knowledge and the Speed of the Transfer and Imitation of Organizational Capabilities - an Empirical-Test. *Organization Science* **6**: 76-92.

Figure 1: Short-run and long-run performance for context similar firms and low complexity

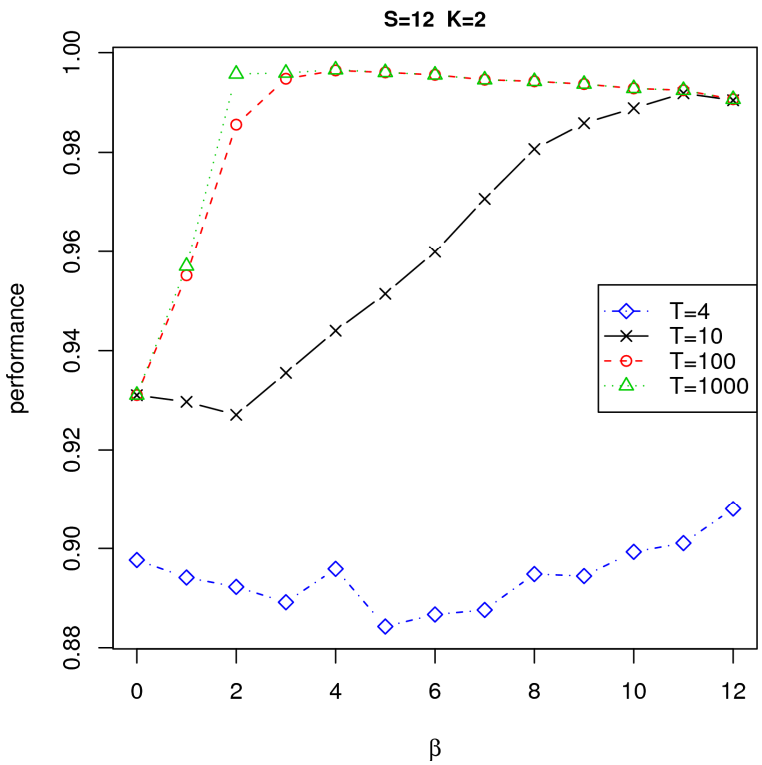


Figure 2: Short-run and long-run performance for context similar firms and high complexity

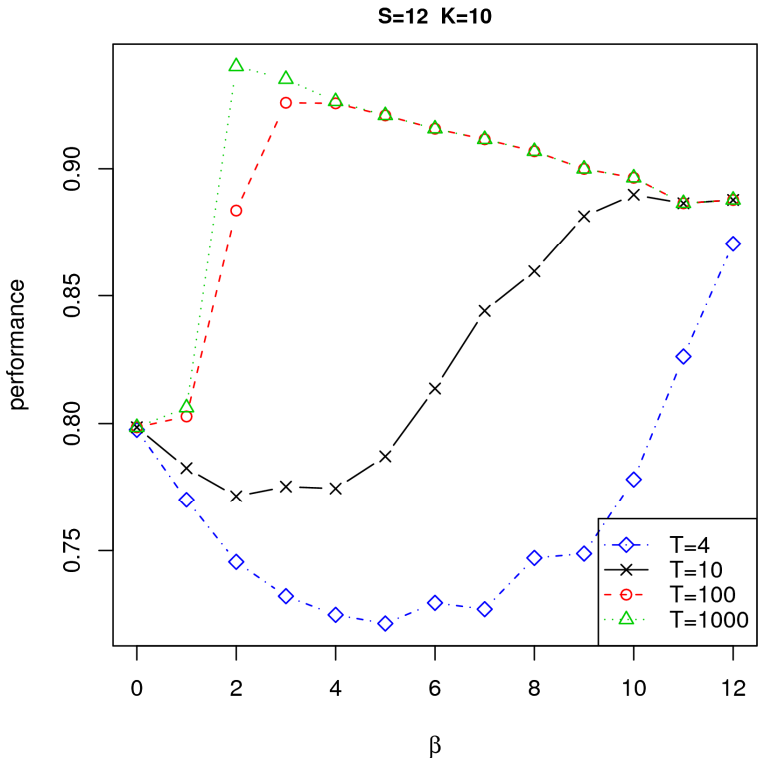
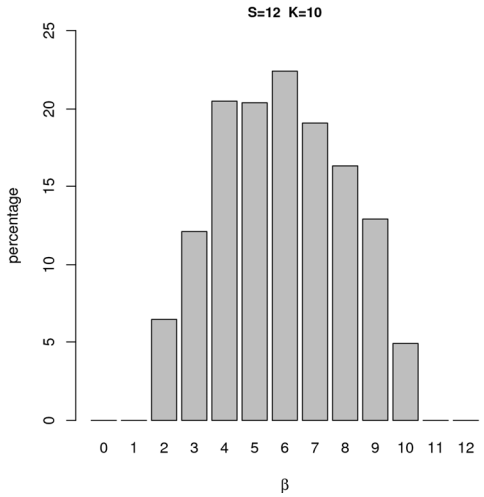


Figure 3: Percentage of first imitations that end up decreasing performance



Note: Performance is measured from just before the first imitation to just before the second imitation

Figure 4: Average number of visited locations

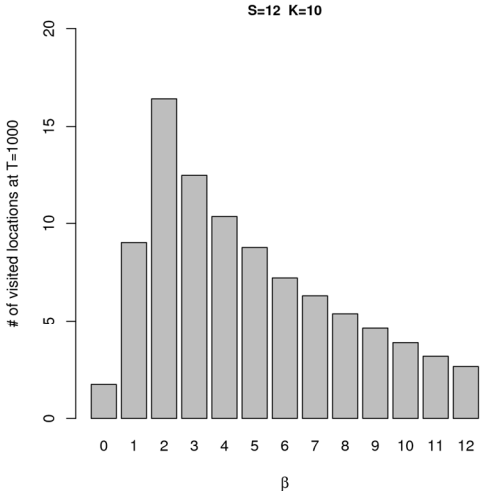


Figure 5: Performance achieved with local search and random long jumps

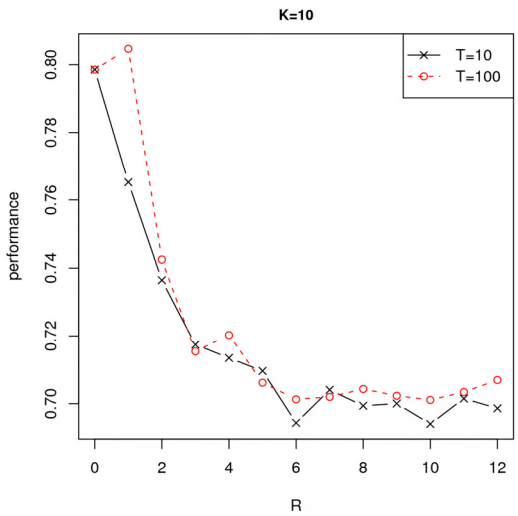


Figure 6: Short-run performance of firms copying firms of different similarity

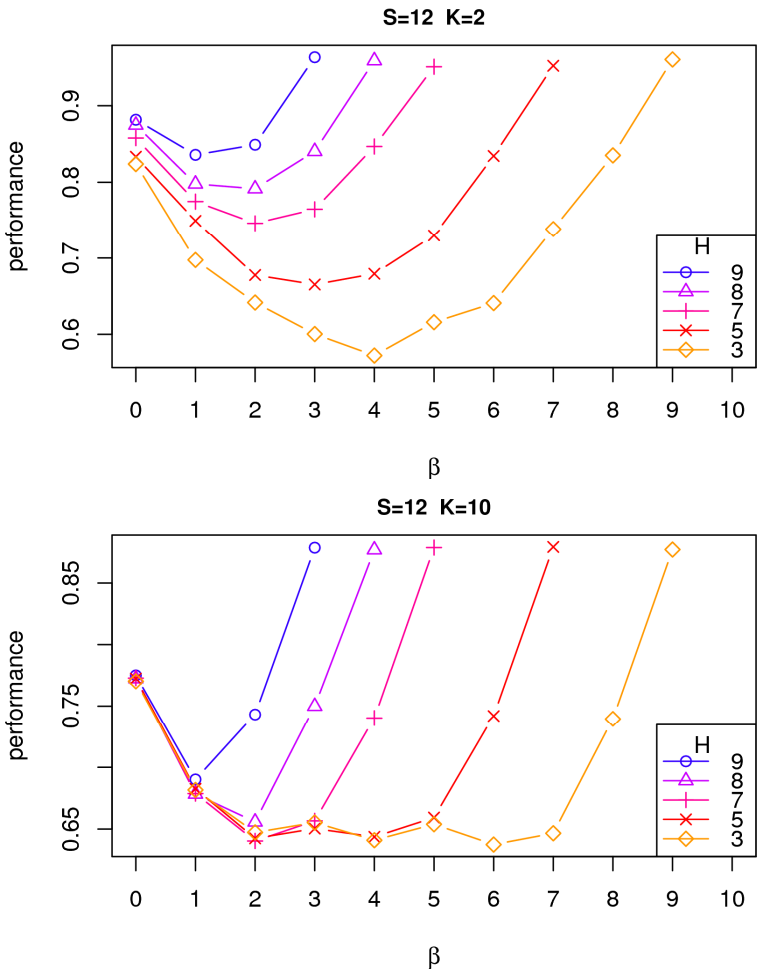


Figure 7: Short-run and long-run performance for firms operating in different contexts

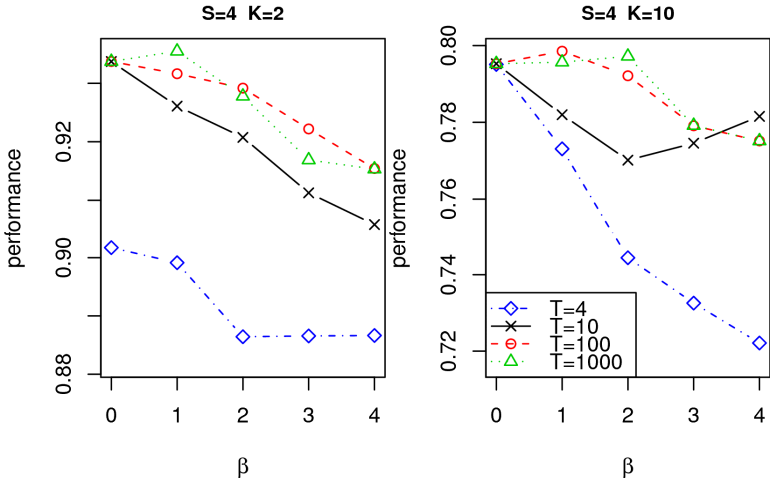


Figure 8: Results for firms using non-discriminant imitation

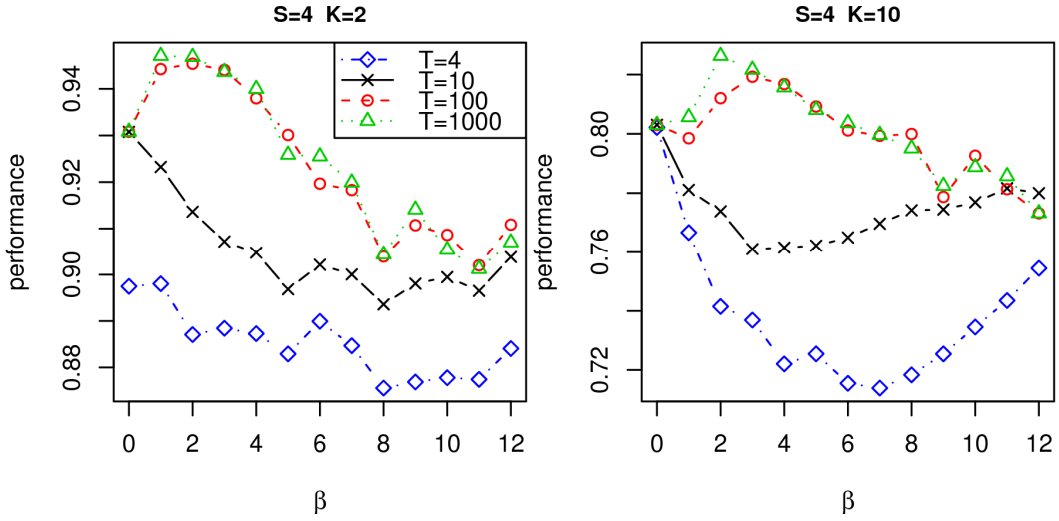


Figure 9: Average number of visited locations for firms operating in different contexts using non-discriminant imitation

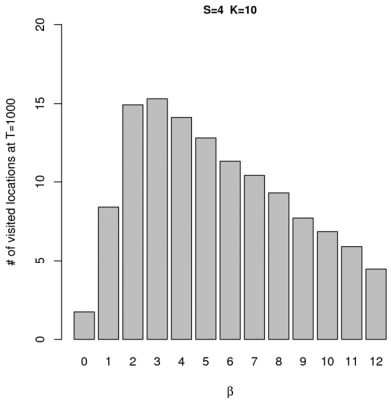


Figure 10: The value of increasing imitation breadth

	High Context Similarity				Low Context Similarity			
	Low Complexity		High Complexity		Low Complexity		High Complexity	
Starting β	Low β	High β	Low β	High β	Low β	High β	Low β	High β
Short Horizon	+	+	–	+	– / –	– / –	– / –	(+) / –
Long Horizon	+	0	+	–	+ / –	– / –	+ / –	– / –

Note**:** A plus sign denotes that performance increases when imitation breadth is increased. A negative sign denotes that performance declines when imitation breadth is increased. A zero denotes that the relationship between imitation breadth and performance is flat. In one condition, the relationship is positive, yet the benefit is below that achieved by local search. We denoted this case with a plus sign in parentheses. For low context similarity, results for non-discriminant imitation are in the upper left of each cell, results for discriminant imitation are in the lower right. For high context similarity, results for both types of imitation are identical.