-

SELECTION AT THE GATE: DIFFICULT CASES, SPILLOVERS, AND ORGANIZATIONAL LEARNING

Abstract:

We analyze longitudinal data on British fertility clinics to examine the impact of "selection at the gate", i.e. the attempts of organizations to improve the success rate of their output by selecting promising cases as input. In contrast to what might be expected, we argue that more stringent input selection is likely to lead to lower overt performance in comparison to those firms that admit difficult cases, because the latter develop steeper learning curves. That is, difficult cases enable more learning from prior experience, because they promote experimentation, communication amongst various actors, and the codification of new knowledge. Our results confirm this prediction and provide clear evidence that organizations with more difficult cases in their portfolios gradually begin to display performance figures that compare favorably to those firms that do select at the gate.

-

INTRODUCTION

League tables and rankings publicizing the quality of firms' output are an important phenomenon in many industries. Business schools, law firms, medical clinics, etc. often publish their performance measurement in the public domain. However, the quality of an organization's output is often substantially determined by the quality of its input. Consequently, often organizations can influence (or perhaps manipulate) their output by carefully selecting what sort of cases to admit as input. For example, business schools usually explicitly publicize the job offers and the career progress of their MBA graduates but they also use very stringent pre-selection criteria in terms of GMAT scores, interview performance, and work experience. Hence, the quality of their output may be the result of the quality of their educational program but equally might be determined by the quality of their selection program. This applies also to law firms that only take on winnable cases, workforce reintegration firms that help unemployed people to find jobs, and management consultants. For example, the business consulting firm Bain habitually and prominently publishes the market performance of its clients, claiming that they (Bain) "make companies more valuable" (Bain & Company, 2010), yet it usually only takes on companies whose financial performance already exceeds that of their peers. The same applies to the type of organizations we examine in this paper, IVF clinics. In various countries, the success rate of these fertility clinics, as percentages of pregnancies among admitted patients, is published annually. However, it is an open secret within the business that many clinics find ways to refuse patients with poor prognoses (The Lancet, 1999; BBC, 2007; Sharif and Afnan, 2003) – we refer to this as "selection at the gate". Hence, IVF clinics can influence their output scores by carefully selecting their patients.

So far, we know very little about the impact of these selection practices on the firms that use them. In this paper, we argue that the benefits of selection – in terms of heightened success rate for firms – might be short-lived. Building on organizational learning theory (Argote and Epple, 1990; Argote, 1999; Haunschild and Sullivan, 2002) we argue that complex, "poor prognosis"

-

cases offer valuable opportunities for learning. Building on insights from extensive field work, we theorize that the knowledge and practices that organizations develop as a consequence of dealing with difficult cases, spill over and have a positive impact on the learning experiences of all cases. As a result, firms that deal with a relatively larger number of difficult cases develop steeper learning curves. In the case of the firms in our sample, their learning curves were so much steeper that, in time, those companies willing to take on difficult cases "caught up" with the firms that operated more stringent selection. Consequently, the overt success rates of the relatively inclusive firms ended up being higher than the rates achieved by those firms that had selected out the more difficult cases. Put differently, organizations that sought to increase their success rate through selection at the gate "shot themselves in the foot".

Our paper makes four contributions to the literature. First, we provide insights on and a systematic analysis of the important societal phenomenon of selection at the gate. Second, we contribute to theory on organizational learning. Research on organizational learning has documented learning curves for various industries (e.g. Argote, Beckman, Epple, 1990; Darr, Argote, Epple, 1995; Dutton and Thomas, 1984; Mihm et al., 2003), it has transferred this analysis from studying production processes to more complex experiences such as acquisitions (Haleblian and Finkelstein, 1999; Vermeulen and Barkema, 2001; Hayward, 2002) and alliances (Barkema et al., 1997; Hoang and Rothaermel, 2005), and it has begun to study what types of experiences are most beneficial for firms (Haunschild and Sullivan, 2002; Schilling et al., 2003). We contribute to this line of research by explicitly examining an important moderator at firm level, i.e. the composition of the firm's portfolio in terms of the complexity of its cases. Hence, we reveal one reason why some firms learn quicker than others – a topic relatively unexplored in this literature.

The third contribution of this paper is that our findings on selection at the gate have broader implications. Firms face various pressures and temptations to optimize short-term results, yet,

-

these actions may sometimes come at the expense of long-term benefits. For example, various management practices and choices, varying from the adoption of a process management system (Benner and Tushman, 2002; 2003), to downsizing programs (Guthrie and Datta, 2008), or outsourcing (Reitzig and Wagner, 2010), may boost short-term performance but have longer-term negative consequences for firm's performance. The choice of some IVF clinics to turn away difficult cases has similar ramifications; although at first sight some cases may seem less attractive and less profitable, their spill-over benefits could still make them profitable elements in the firm's portfolio. Our study shows that it is important to take account of these indirect, long-term benefits.

The final contribution of this paper is an empirical one. One of the problems related to the type of topic addressed in this paper is the potentially confounding effect of reverse causality; some organizations (e.g. IVF clinics) might not be improving *because* they are dealing with many difficult cases; instead, because they are better, they are attracting more challenging cases. For example, the cancer units in very reputable institutions often display higher mortality rates because patients with the most complex etiology (and hence the lowest ex-ante probability of survival) find their way to or are referred to the best hospitals. We are fortunate that the IVF data we use for this study allow us to test directly for reverse causality, due to the presence of a control group. About half of the IVF clinics in the sample are National Health Service (NHS) clinics: i.e. government hospitals. These clinics do not select at the gate;[1] their patients are assigned randomly (i.e. by postal code). Nevertheless, this assignment process results in some NHS clinics receiving very few poor prognosis patients and others receiving relatively many.[2] This enables us to test for – and rule out – the presence of reverse causality.

We test our prediction – that the learning curve is less steep for firms that perform relatively few difficult cases – through three variables: 1) the proportion of women with a relatively poor

-

prognosis because they have failed to conceive through previous IVF treatment; 2) the proportion of women with a relatively poor prognosis because they produce very few eggs; and 3) women who are aged 35 years or over (35 is the industry's standard cut-off rate), all categories known to have significantly lower chances of conceiving. All three variables support the study's prediction: their presence moderates the organization's learning curve so that firms that took on more difficult cases improved their success rate quicker.

THEORY

Experience and Learning

The literature on organizational learning includes several strands (Argote and Ingram, 2000). One tradition – pertaining to the study of learning curves – examines the relationship between cumulative experience and performance. It interprets a positive association as evidence that learning has taken place, but it treats the process and content of knowledge accumulation as a black box. In these studies, organizational learning refers to gradual improvements, for instance in the form of efficiency increases, as the firm gains more experience with a particular process. Learning is seen as occurring iteratively as firms engage repeatedly in an activity, draw inferences from their experience, and store and then retrieve the learning through future engagement in the activity (Argote and Ophir, 2002; Levitt and March, 1988). Learning from experience may result in reduced production inputs (Dutton and Thomas, 1984; Mihm, Loch, and Huchzermeier, 2003), reduced unit costs (Argote and Epple, 1990), improved completion times (Edmondson et al., 2001; Pisano et al., 2001; Reagans et al., 2005), acquisition efficiencies (Haleblian and Finkelstein, 1999; Hayward, 2002), and higher survival rates (Barkema et al., 1997; Ingram and Baum, 1997; Kim and Miner, 2000; Vermeulen and Barkema, 2001).

Our study follows this tradition. We will measure whether the success rate of IVF clinics increases with experience. "Success" in this line of business (IVF clinics) is very clearly defined:

-

an IVF treatment that results in a birth is a success; and otherwise it is a failure. Hence, in our analysis, the measure of an organization's success rates is the proportion of live births in the patients receiving IVF treatment. Experience concerns the cumulative number of IVF treatments performed by the organizatio. Specifically, we will examine whether the success rate of these organizations increases more or less quickly dependent on the proportion of difficult cases that they treat.

The Process of Learning

Studies examining the process of learning represent another strand in the organizational learning literature. The learning process is often seen as consisting of multiple, interdependent stages representing the search for, choice and implementation of solutions. This notion has led several authors to describe it as a cycle of activities engaged in by an organization to process knowledge that allows it to adapt and improve (Argyris and Schön, 1978; Kolb, 1984; Edmondson, 1999; Gibson and Vermeulen, 2003). For example, in the context of team learning, Gibson and Vermeulen (2003) describe the process as a cycle of experimentation, reflective communication, and knowledge codification. Similarly, we see the process of organizational learning as a cycle consisting of three subprocesses: experimentation, reflection and coordination, and capturing the newly acquired knowledge in the form of routines, technologies, and procedures.

First, individuals within the organization need to generate ideas about how to improve work through exploration or experimentation (Argyris, 1976; Levitt and March, 1988; March, 1991). They need to try out new things, which leads them to accumulate new knowledge (Zollo and Winter, 2002). This concerns the process that Prencipe and Tell (2001) refer to as learning by using and doing. Second, a common understanding about a proposed solution must be developed. Organizational members' engagement in experimentation can result in different mental schemas related to the experience. To reach a common understanding of what the experience or

-

information means, members transfer and combine their insights through a process of reflection and coordination (Jelinek, 1979; Walsh et al., 1988). This leads to the articulation of knowledge (Zollo and Winter, 2002), which is described by Prencipe and Tell (2001) as learning by reflecting, thinking, discussing, and confronting. Finally, the knowledge needs to be translated and codified (Zollo and Winter, 2002) into tangible, generalized concepts, decisions, or work methods (Argyris and Schön, 1978; Kolb, 1984); what Prencipe and Tell (2002) refer to as "learning by adapting, implementing, and replicating".

In combination, these processes form the learning cycle (Gibson and Vermeulen, 2003). We posit that dealing with difficult cases enhances this cycle. Analyzing airlines' learning from accidents involving heterogeneous causes, Haunschild and Sullivan (2002) argued that it caused organizations to look for connections between different elements of a problem to deal with a complexity of multiple underlying causes (rather than attributing it to one relatively simple cause), and generate debate among actors, which leads to a richer understanding of the problem. Similarly, we argue that dealing with non-standard, difficult cases will force organizations and the people involved in the cases to conduct a deeper analysis of the underlying structure of a problem, enhancing their understanding of the issue and changing how they deal with and respond to prior experience.

Hypothesis Development

When a firm, such as a medical clinic, admits and begins to work with a relatively difficult case, it will often need to depart from existing routines and protocols. The standardized approaches may not work given the complex characteristics of the new case. This will necessitate and trigger experimentation, from which may emerge new solutions, deeper insights, and different ways of working across the organization. Experimentation can increase the skill levels of the individuals involved. Working with challenging cases enables additional practice, and allows those involved

-

to apply and evaluate novel solutions which otherwise might not have been considered. Evidence from human information processing suggests that task complexity requires the learner to generate a more elaborate mental model (Wilson and Rutherford, 1989) and enhances the ability to carry an increased cognitive load (Bannert, 2002; Sweller, 1989). Subsequent experience with more standard cases will be viewed in a new, more comprehensive light, with a better understanding of what Haunschild and Sullivan called "the underlying structure of the problem" (2002, p. 614). Thus, when organizations employ less stringent selection at the gate, novel challenges and opportunities are more likely to be encountered. As a consequence, the firm's general competence to address new cases will increase with each unfamiliar case. In contrast, dealing with standard cases only will not lead to a richer understanding, and might even demotivate actors to try out new solutions.

Some of the clinics where we conducted in-depth interviews for this project seemed deliberately to use complex cases for training purposes – see Table 1 for quotes. These difficult cases were introduced to entice people to experience new situations and new ways of doing things. Hence, dealing with complex cases enhances the general ability and skill levels of the people involved and leads to a deeper understanding, which promotes the development of new solutions. These new solutions may also improve success rates among more simple cases. For example, one physician commented:

> What you see in the textbook or in the code of practice are treatment coordinates for standard cases, the typical patients showing up for consultation, young couples under 35, with good egg reserve, good sperm and good health … We have a lot of experience with them and they're easy cases where we rarely deviate from the standard procedure. And that's all fine, but when you get a difficult case, with complex pathology, and the standard procedure simply doesn't fit, what do you do? You change the practice, you start tinkering with the parameters, adding new things, adjusting doses and sequences so that it fits. And

-

is that all? No, it isn't; you start tinkering with the procedure for the easy case as well; you take what you've learned from that difficult case to the easy case.

As this and other interview extracts (see Table 1) suggest, finding solutions to difficult cases prevents the firm from inertia and over-reliance on a standard set of operating procedures. It motivates them to experiment to try to enhance overall performance. Haunschild and Sullivan (2002) suggest that the attention of the organizational members involved in dealing with a complex problem forces situational analyses which may go beyond simple responses, and leads to deeper analysis of the situation at hand. This is in line with Ocasio's (1997) principle of situated attention, which emphasizes that ventures that require greater attention represent greater cognitive resources and higher levels of concern and participation from the organizational members involved. The experimentation triggered by experience of complex cases leads to learning effects that are beneficial to the execution of simpler cases. Subsequent experiences are evaluated on the basis of the new cognitive model and set of solutions.

----- Insert Table 1 about here -----

Difficult cases also stimulate learning by heightening interaction among the various parties in an organization that contribute to the problem solving – the second part of the learning cycle. Because such cases necessitate a departure from standard procedures, this requires the organizational members to communicate and coordinate with one another. In the context of standard cases, coordination is achieved through adherence to established routines and procedures. Non-standard cases force people to interact directly. Increased levels of coordination will uncover problems in standard cases, fine-tune solutions, and aid the transfer of knowledge and routines (see also Hargadon and Sutton, 1997; Henderson and Cockburn, 1996). One physician summarized it thus:

-

> Clinics which admit more difficult cases tend to be more disciplined and more thorough in their work. And of course, even for us, the effort of treating [difficult patients] intensifies the interaction among our doctors, embryologists, and nurses. And we tend to take that with us, and to do it for the next patient which enters our office.

More direct communication also aids the processes of sensemaking (Weick et al., 2005), from which standard cases also benefit. It triggers collective reflection on how to deal with the problem at hand.

Finally, the newly developed insights and solutions need to be translated into formal and informal processes, in the form of procedures, technologies, and routines – the third stage in the learning process. Methods and procedures developed under difficult circumstances trying to solve complex problems, may improve performance standard cases. In an experiment, Schilling and colleagues (2003), for example, found that experience gained in one setting benefited learning in a related context. Similarly, Wiersma (2007) found that performing a diversity of related tasks enabled organizations to learn quicker. Both studies concluded that variation in related tasks leads to a deeper cognitive understanding of the underlying processes, which allows for easier transfer of solutions. Our interviewees contended similarly that protocols, checklists, technological devices, and informal procedures developed when dealing with complex cases were often transferred to and applied to standard cases. For example, one executive – a quality manager – commented that:

> Doctors have checklists; the more difficult their cases, the longer the checklists. And I am interested in their checklists because I want to revise mine and bring the system up to date. Are we getting cases with a new bullet point? Then I want to know about it, the other doctors want to know about it, the nurses as well.

To conclude, experience of difficult cases enhances learning in the organization. First, the experience triggers experimentation. Because complex cases imply a departure from established

-

procedures and routines, they force people to communicate and reflect on the newly developed

solutions. Finally, the experience and learning lead to the codification of new knowledge in terms

of newly adapted processes, routines, and technologies. Consequently, we predict that the learning

curves observed in the firms in our sample will be steeper among organizations with relatively

high proportions of difficult cases. Formally:

Hypothesis: *The relation between experience and success will be more positive for*

organizations with a higher proportion of difficult cases.

This means that organizations that select heavily at the gate, and therefore deal with fewer

difficult cases, miss out on learning benefits. They are less likely to depart from standard

procedures, which leads to less experimentation, lower levels of interaction, and less new

knowledge codification. As a result, these firms learn more slowly than organizations that apply

less stringent selection criteria. Firms that admit a relatively high proportion of complex cases,

initially will likely show lower success rates. However, if, as we predict, their rate of learning is

indeed substantially higher than that of firms that select heavily at the gate, the learning curves

could eventually cross, and firms with a higher share of difficult cases will begin to show

comparatively higher success rates. In our models, we test for this possibility directly.

METHOD

Research Setting and Data

The first in-vitro fertilization (IVF) baby was born in 1978 as the result of the work of two British

scientists, Robert Edwards and Patrick Steptoe (1980). Yet, it was not until 1992 that the

regulators in the United Kingdom permitted fertility centers to offer IVF treatment. Since then,

data on all UK IVF centers have been collected and published by the Human Fertilisation and

Embryology Authority (HFEA), which is the independent regulator that oversees the use of

gametes and embryos in infertility treatment and research in the UK. With the HFEA's

-

permission, we were able to trace back data on variables such as experience and success for the population of all fertility clinics based in the UK since 1991; the year prior to the introduction of IVF as an authorized treatment and up to 2006 (the final year made available). In addition, we had a total of 32 face-to-face interviews with people in the industry to supplement our quantitative analysis; 10 of which were conducted after completing the first draft of this manuscript.

Of all IVF clinics in the UK, more than two thirds are private, usually owned by the doctors operating them, with the remainder being state-owned (i.e. National Health Service). For the former, as one respondent put it, "there is fierce competition between clinics". Patients can choose at which clinic they want to buy treatment and, when doing so, many of them pay careful attention to the *Patients' Guide* published by the HFEA (2002). This annual guide lists the clinical results for each IVF clinic in the country. This league table – as it is generally referred to in the industry – assures that all clinic success rates are publicly available and easy to access. For example, in 2010, during which 45,264 patients received IVF treatment in the UK, the relevant website received more than 600,000 hits.

Table 2 comprises a number of interview quotes regarding the roles of the league table and selection at the gate. The quotes illustrate that the existence of the league table puts pressure on the clinics to present good success rates. It stimulates them to select patients with ex-ante higher chances of success (The Lancet, 1999; BBC, 2007; Sharif and Afnan, 2003). As one respondent put it, "The best way to move yourself up the table is to treat prognostically the better group of patients". Interviewees also indicated that clinics do this to a more or lesser extent; some are very selective where others are much less restrictive.

----- please insert Table 2 about here -----

-

For a variety of reseason, this setting and these data are ideal for testing the relationship between experience and success. First, because the measure of success is unambiguous: whether a cycle of IVF treatment results in a live birth or not is a clear goal and outcome. Furthermore, because each patient and treatment cycle is distinct, cumulative experience is clearly measurable and simple to operationalize, namely as the prior number of treatments performed. Moreover, there are some clear indicators as to whether a patient should be classed as a "difficult case" or not, based on poor prognosis, prior failure to conceive, or age. The database we constructed is very comprehensive. There were only 11 left-censored observations, which gives us complete longitudinal data on 90 percent of the clinics in the population, and hence complete data on their prior experience. We re-ran the analyses excluding the left-censored observations and this did not change any of the results. Finally, because around one third of the clinics in the sample are government hospitals (which do not select at the gate) we have a valuable control group to rule out some possible alternative explanations (e.g. reverse causality).

During the period 1991-2006, a total of 116 IVF clinics were set up; by the end of our sample period, 100 of these had more than 2 years of data, with the average 9.3 years of observation per clinic. The oldest clinics had been offering IVF for 15 years, the newest for just 1 year. The largest clinic had treated over 13,000 patients during the window of observation, while the average number of prior cases was approximately 4,000 patients. In total, these clinics have performed over 400,000 IVF cycles on around 300,000 women, who had delivered over 75,000 IVF babies by the end of 2006.

Dependent Variable

Success rate. In the *Patients' Guide*, success rates are presented separately for patients in six age groups: under 35, 35-37, 38-39, 40-42, and over 43 years. Patients under 35 - and by far the largest group - are generally regarded in the industry to be the "standard patient group" (Johnson

-

et al., 2007), and used as the primary basis for comparisons between clinics because it "reduce[s] the impact of 'patient-mix' on the comparability of results between different centres" (Sharif and Afnan, 2003: 484). In light of these field observations, our primary dependent variable is the success rate in standard cases, defined as the live-birth rate in the IVF patient group aged under 35, involving the use of the patient's own fresh eggs. It is calculated as the number of live-birth events per number of female patients under 35 who underwent one or more fresh IVF cycles in the year of observation. We chose standard cases (i.e. women < 35) as the dependent variable, rather than the success rate among women of all ages, because this is the information that is made public and hence informs (potential) patients when choosing a clinic. Repeating the analysis using the success rate across all cases (i.e. women of all ages) led to basically the same results as those displayed in Table 4.

Independent Variables

Prior experience. Experience was measured by cumulating all the prior cases that involved one or more IVF cycles using the patient's own fresh eggs. The number of patients treated by each clinic in a given year is not published in the *Patients' Guide*; it was provided directly from the HFEA. In line with prior research (e.g. Argote, 1999), we computed the natural log of experience. This assumes, for example, that the experiential difference between 100 and 200 cases is more influential than the difference between 10,000 and 10,100 cases.

To measure the level of difficulty in a clinic's patient mix, we identified three factors which IVF practitioners regard as relevant dimensions for assessing the complexity of a case: patients over 35 years old, patients that have failed to conceive with previous IVF treatment, and patients who produce very few eggs. All three categories are considered to represent more difficult cases with relatively poor prognoses.

-

Patients aged over 35. In the field of IVF, the variable patients over the age of 35 is generally considered the main predictor of success, since increasing age of the female reproductive system generally reduces the chance of successful pregnancy. Thirty-five years is used as the standard cut-off rate in the field because at around that age female fertility shows a rapid decline. Figure 1 shows that the success rate for IVF cycles is relatively stable up to the age of 35, but shows a sharp decline thereafter. So women aged under 35 are considered standard cases, and patients aged over 35 are seen as more difficult cases. One of our interviewees told us that: "Age is the most important parameter, [and] the only thing that we know and that is true beyond any scientific doubt is that increasing age equals more difficult cases and poorer outcomes".[3] We computed the proportion of older patients as the ratio of the number of IVF patients over the age of 35, to the total number of IVF patients treated each year by each clinic.

----- Insert Figure 1 about here -----

Patients with a history of prior failed treatment. The second variable used to measure relatively poor prognosis cases is number of patients where previous treatment has failed. A history of prior IVF failure can signal a possible problematic underlying etiology. Patients who have had a previous IVF treatment and failed to conceive may have health conditions that require further investigation and intervention, or may have a tolerance to the standard drugs. As one respondent put it, "[a] factor that affects IVF success rate [is], for example, how many times has the patient attempted IVF in the past". Of course, not all patients who fail to conceive following an IVF treatment will have a particularly complex or problematic etiology; in some cases, the failure may be sheer chance. And vice versa, some patients with poor prognoses may get pregnant after the first treatment. However, if we measure all of a clinic's patients experiencing previous treatment failure and compare with successful conceptions at first treatment, on average, the former group can be expected to include a significantly higher proportion of complex cases. We build on this information in the analysis. We proxy the proportion of patients who have received more than one

-

IVF cycle by dividing the total number of IVF cycles by the total number of IVF patients treated at a clinic, in the same year of observation.

Proportion of patients with low egg reserve. The last indicator of a difficult case builds on a patient's ovarian reserve; the number of oocytes (eggs) that can be retrieved from a woman's ovaries and fertilized, as an early indication of the likely outcome of the treatment. Clinics can choose to measure this before commencing treatment, as explained by one of our respondents: "They'll usually run a panel of hormones to get an idea of what they call ovarian reserve, in other words, does the ovary still have plenty of eggs". Practitioners tend to be more optimistic about patients with higher egg and embryo counts because a larger number of retrieved oocytes increases the probability of achieving valid embryos, from which to select the best candidates for transfer into the patient's womb. For these patients, the excess eggs or embryos are cryopreserved in case subsequent frozen cycles are necessary.[4] If there are no excess embryos that can be frozen, this is seen as indicative of an underlying, problematic etiology. Therefore, the number of frozen cycles relative to the number of fresh IVF cycles performed at a clinic is treated by sector analysts as indicative of the proportion of "good prognosis patients" in the clinic's patient mix (Abdalla, 2008).[5] The more the number of frozen cycles performed at a clinic relative to the number of fresh IVF cycles, the higher the incidence of good prognosis cases among its patients. To make the interpretation of this ratio consistent with the hypothesis in our study – which refers to the effort expended by clinics to handle more demanding cases – we reverse-coded this ratio to obtain a proxy for the number of poor prognosis patients.

High proportions of each of these three patient groups pose challenges to clinics in the sense that there are more elements that must be addressed at each stage in the treatment. For example, with increasing age, the woman's reproductive system generally exhibits more constraints to successful pregnancy: the pituitary and thyroid functions deteriorate; the shapes of the ovaries and uterus

-

change which affects their function; the body's neurotransmitters become less responsive; the morphology of the human gametes changes, etc. These elements and their complex relation to each other require additional procedures in the treatment cycle (e.g., customization of the hormonal doses in order to insure ovarian response; micro-manipulation procedures performed in the lab to ensure fertilization and/or selection of morphologically normal sperm, eggs, or embryos, etc.). Similarly, for patients who have experienced a failed previous cycle and for patients with lower egg counts, treatment cycles tend to involve more task elements in order to address the medical complications of having a recent failure, or a low egg count. For these patients, clinics cannot rely on the standard treatment protocol only.

Control variables. We control for clinic size using a commonly accepted measure of clinic capacity, i.e. the count of all licensed treatments (including IVF cycles, donor insemination, frozen cycles, and cycles involving donated eggs and embryos) performed at each clinic in the previous year. We include the square term of clinic size because prior research suggests that the effect might be non-linear (e.g. Haveman, 1993). We use a dummy variable to control for whether the clinic has ICSI (intra-cytoplasmic sperm injection) technology available. ICSI is an innovation that was introduced in the IVF industry during our period of observation. ICSI enables embryologists to address the problem of low sperm count or poor sperm mobility, by injecting a single sperm into the ovum. This makes a significant difference even for couples where male infertility is not an issue because it leaves less to chance (Takeuchi et al., 2000). At the beginning of our observation period there were no clinics with ICSI; at the end of the observation period all clinics were using ICSI, but not all clinics gained access to this technology at the same time. We control for this using a time-variant variable.

We control also for the industry experience in the field of IVF. The average success rate of IVF across all clinics has increased over the years. Hence, a clinic established, for example, 2005 is

-

likely to enjoy a higher immediate success rate than a clinic established at the first year of entry in 1995, purely because the field as a whole has progressed. Firms learn from the experience of others (Argote et al., 1990; Ingram and Baum, 1997), through medical training, employee mobility, conferences, and so forth. Therefore, we control for total industry experience, measured as the natural logarithm of the count of all IVF cycles performed in the UK up to the year of observation. As an alternative, we re-ran the models controlling for the highest clinic-level success rate achieved for the standard patient group in each given year, to indicate the "state-of-the-art" in the field. Both measures led to identical results. We present the models with total industry experience. Finally, all our models include fixed-effects (i.e. they include clinic dummies), representing a shift in the intercept of a firm's learning curve, to control for any remaining unobserved clinic-specific characteristics that may affect clinical performance.

Analysis

We ran various estimators to check the robustness of our models. Below, we present the results of the ordinary least squares (OLS) estimator with fixed effects (within and between). Models using a random-effects estimator produced nearly identical results. Also, because our dependent variable is a proportion (proportion of patients giving birth), it is bounded between 0 and 1, whereas the predicted values of an OLS model cannot be guaranteed to lie within this interval. To correct for this, Papke and Wooldridge (1996) proposed the fractional logit estimator, which they later expanded for use with panel data (Papke and Wooldridge, 2008). This estimator produces very similar results to those displayed below, although our second test (patients with previous failed treatment) only supports our hypothesis at $p < .10$. In the remainder of the paper, we present and discuss the results of the fixed effects OLS regression because it is easier to interpret the size of the coefficients for this estimator.

RESULTS

-

Table 2 presents summary statistics for the variables included in our models. One thing to note is the fairly low or even negative correlations between the three indicators of difficult patients. We conducted some further analysis and interviews to understand the background to this. One reason for the relatively low correlations, specifically between patients who failed before and patients who produce few eggs (.19), is simply that not all clinics deliberately select at the gate. Indeed, for the subsample of private clinics, which at least have the possibility to select their patients (the NHS clinics did not differentially select at the gate) the correlation is already higher (.23).

Furthermore, the variable patients above the age of 35 differs from the other two due to clinic location. We took each of the three difficulty indicators as dependent variables in a simple regression model (with fixed effects) with GDP per capita for the specific area in which the clinic is located as the explanatory variable (representing 185 different areas within the UK). The results showed that local GDP per capita was significantly negatively associated with patients who failed before (-.48; $p<.01$) and patients who produce few eggs (-1.51; $p<.001$), but significantly positively with patients above 35 (2.76; $p<.001$). The likely reason for this is that in relatively socially deprived areas (i.e. with lower GDP per capita) there are more patients with a difficult underlying etiology, for instance due to a higher incidence of sexually transmitted diseases. In contrast, in more wealthy areas, patients tend to be older when trying to conceive. When assessed across clinics, this leads to a low or even negative correlation between the first two measures and patients above 35.

----- Insert Table 3 about here -----

In conformity with the abovementioned findings, we also ran an exploratory factor analysis (using principal components with varimax rotation), which showed that patients who have failed before and patients with few eggs load on the same factor (with factor loadings .74 and .76 respectively). Yet, the second factor was comprised solely of patients above 35 (.89). Nevertheless, a combined

-

measure of difficult patients, constructed by standardizing and averaging the three individual indicators, led to the exact same results and conclusions as discussed below.

Table 4 displays the results of the regression analyses. The first column refers to the model with control variables only. The results show that, in general, if clinics grow larger their success rate increases. The quadratic effect of size is negative but because the overall relationship between size and success only turns negative at +1.35 standard deviations above the mean, the estimations suggest that clinic size might have a negative influence on the success rate only when clinics become very large. The effect of size is modest, but significant. For example, if the size of the average clinic increases by one standard deviation (i.e. from 457 to 891 treatments per year), the probability of a patient becoming pregnant at that clinic increases by 2%. The other two control variables show the expected effects. The use of the innovative ICSI technology increases success rates by about 4-5%. Overall industry experience – as a proxy for overall progress in the field – is also positive and significant.

----- Insert Table 4 about here -----

Testing the Hypothesis

Models 2-5 present the estimates of our three measures of difficult cases (age, prior failure, patients producing very few eggs) and their interaction with clinic experience. In all three cases the interaction is positive and significant. This indicates that the positive relationship between clinic experience and success is stronger for clinics that deal with the more difficult cases. Hence, all three tests strongly support our hypothesis: clinics that deal with a larger proportion of poor prognosis patients show steeper learning curves.

Note that the three main effects concerning difficult patients in these models are negative, as expected. This means that, by itself, taking on a higher proportion of complex cases depresses a

-

clinic's success rate, because dealing with many poor prognosis patients initially decreases the number of pregnancies resulting from the treatment.[6] However, because of the learning effects they entail clinics that admit more difficult cases show an increase in their success rates faster than clinics that are more selective over their admissions.

Figure 2 displays the estimated relationship between clinic experience and success rate. Using the results from model 5, keeping all other variables at their mean, we display the relationship between experience and success for a clinic where all three indicators are at one standard deviation below their mean (i.e. a clinic dealing with a low proportion of difficult cases), versus when all three indicators are at one standard deviation above their mean (i.e. one with a high proportion of difficult cases). Accordingly, the former is labeled "low selection at the gate", where the other is labeled "high selection at the gate". The results show clearly that treating difficult cases has an initially depressing effect on a clinic's success rate; a clinic that admits more difficult cases has a success rate as much as 10% lower than clinics that mainly select more promising cases. However, as the graphs show, these clinics start to catch up rapidly. After about a 100 cases their success rate is already equal to that of the clinics that do select heavily at the gate.[7] Subsequently, they their success rates improve at a faster rate than those of their counterparts; after 400 cases, their overt success rate is 3.3% higher than the clinics that deal with few difficult patients.

----- Insert Figure 2 about here -----

Alternative Explanations and Additional Analysis

Reverse causality. We tested whether the results presented above could potentially be confounded by some reverse causality. This is pertinent because it seems possible that (prospective) patients with a relatively poor prognosis, who can access and observe the various clinics performance indicators (which are published), might be more inclined to select and visit a clinic with a higher

-

overall success rate. Or, similarly, that the clinics that are improving rapidly will start to attract more difficult cases. Hence, it might be that clinics with higher (ex-ante) success rates attract more difficult cases, rather than the other way around. To test for this possible effect directly, we estimated our models on the subsample of NHS clinics only. These state-owned clinics cannot select at the gate; they are obliged to admit all the patients that are assigned to them. As one doctor working in an NHS center explained:

> I can't pick and choose here at all. I have to see everyone who comes to the door… I can't turn down someone who has bad endometriosis or poor egg reserve, I can't say 'sorry I can't treat you' because it's their right, they're NHS. In a private clinic I can, oh yes. I can tell them 'look, I don't want to include you in my statistics.' […] There is much more independence as to what some doctors can do in other clinics."

Patients also cannot select among NHS clinics; they are assigned to a particular clinic based on their home address/postal code. Therefore, because poor prognosis patients cannot choose clinics with higher success rates, there is no possibility of a reverse causality effect.

However, some NHS clinics deal with a substantially higher number of difficult cases than others, due to chance but also because particular areas/postal codes are associated with more complex pathologies (e.g. inner-city areas). For example, one of the doctors working in an inner-city clinic, who previously had worked in a clinic in a wealthy provincial area, commented:

> There I had mostly Caucasian patients, with less health issues than our patients here in [East London], where I see much more Africans than I ever saw in [Yorkshire]… Also patients from the young urban population here are more likely to have pelvic infections not using condoms, having unprotected sex, are more likely to get Chlamydia and all those things with tubal problems… So I see more pathologies, more problems and so on….

-

Hence, although not caused by selection at the gate, some NHS clinics deal with larger proportions of more difficult cases, but there is no chance of reverse causality. If our organizational learning theory holds, these clinics should also display steeper learning curves than their NHS counterparts that deal with relatively few difficult cases. This enables us to check whether our findings might be confounded by reverse causality issues.

Our longitudinal data include 39 NHS clinics, together accounting for 406 clinic-year observations. We estimated the effect of our three predictors on the success rates of these clinics using the models displayed in Table 3. In spite of the smaller sample size, the coefficients of all three predicted interactions confirm our hypothesis (patients above 35 = .0444, p<.05; prior failure = .0371, p<.10; poor egg reserve = .0096, p<.05); the results are very close to the results based on the full sample. To conclude, reverse causality does not explain our findings, which still fully support the prediction in this paper.

Additional specialists. Another alternative explanation could be that clinics that treat a larger proportion of difficult patients, over time, add more specialists to their personnel to accommodate those patients. The presence of these specialists might potentially also start to benefit standard cases, which could explain our results. To test for this possibility, we contacted the HFEA, who do regular (although not yearly) inspections of all clinics, among others collecting data on the number of specialist roles. Hence, we were able to collect data on the number of specialists for 606 year-clinic observations (out of a total of 1004), representing 85 different clinics. Using the number of specialists as a dependent variable, including the same independent variables as in all other models, we tested directly whether clinics that treat more difficult patients create more specialist positions. Although industry experience significantly predicted the number of specialists in a clinic (6.26, p<.001; which indicates that with progression in the field more specialist roles were developed), and ICSI technology drove the number of specialist down (-1.56, p<.001;

-

probably because ICSI replaced several other specialist roles), our three indicators of difficult cases, whether measured as a proportion or in interaction with cumulative experience, remained wholly insignificant in all models. This shows that clinics that deal with a relatively large number of difficult cases do not add more specialists to their team than others. Hence, this does not support the alternative explanation for our findings.[8]

More technology. Similar to the previous point, one could conjecture that clinics that deal with more difficult patients also start to use the additional technology – in the form of equipment and specialist treatments – for standard cases, which could also benefit their success rate. This too could potentially drive our results. Although it is impossible to measure all technologies and specialist procedures, we were able to find a proxy in the form of ICSI treatment, specifically to what extent a clinic also started using ICSI for standard patients. ICSI is an expensive procedure that usually is only used for difficult patients with a particular set of fertility problems, yet some clinics have started to use it more widely. The HFEA provided us with data about how many ICSI treatments each of the clinics provided any given year. Since we already had information about how many difficult patients they treated, we were able to add a control to our models proxy'ing ICSI for standard patients. This measure remained completely insignificant in all models; apparently treating standard patients with ICSI does not improve their success rate. Importantly, all our results, as reported in Table 4, were fully replicated.

Better learning environments. Another alternative explanation could be that some clinics form better learning environments, in the sense that they are motivated to push the frontier of knowledge about IVF, and are therefore inclined to accept a larger proportion of difficult patients. Yet, these attractive learning environments could also attract better doctors, which could in turn lead to a higher success rate. Although this alternative rationale should to a considerable extent be accounted for through our fixed effects models, we endeavored to test and control for it more

-

directly. First, we collected a dummy variable whether a clinic was affiliated with a university hospital, under the assumption that these are more interested to push knowledge development about IVF. This variable, however, appeared to be largely time-invariant, so that it was already controlled for – and thus dropped out – in our fixed effects models.[9] Next, via the HFEA, we managed to collect data on the total number of research projects that each clinic was engaged in for every year, assuming that the more research-oriented clinics would represent a more interesting learning environment for ambitious doctors. The estimate of the variable itself, as well as of its interaction with cumulative clinic experience, was insignificant. Importantly, it did not reduce any of the support for our predictions. Apparently, engagement in research projects is not a substitute for experience, as it does not cause a clinic to learn quicker.[10]

DISCUSSION

We have shown that selection at the gate – organizations that try to enhance their explicit success rate by selecting promising cases as input – may eventually disadvantage the firms that engage in this practice. The British IVF clinics in our study that admitted few difficult cases initially enjoyed higher success, in the form of more live births per patient treated. However, this advantage disappeared quite quickly; because the clinics that treated relatively higher numbers of difficult cases learned quicker from their prior IVF experiences, the organizations that selected heavily at the gate ended up having significantly lower overt success rates than their more lenient counterparts. What initially seems an advantage, turned against the firm in the longer run; the clinics that selected heavily, eventually found themselves "on the back foot" in terms of their success rates. Although they were treating mainly patients with relatively simple etiologies, the number of live births per patient ended up being lower than in clinics treating more difficult cases.

Our findings suggest that admitting and dealing with more difficult cases enhances learning because it requires organizations to depart from routine processes. Standard IVF treatment

-

consists of a series of routinized, sequential stages, which are guided by protocols. The various medical professionals conducting the different stages do not need to meet and communicate when they are following the established procedures. However, dealing with non-standard cases forces them to consider new solutions, establish new communication patterns, and alter their work methods. Our findings suggest that these new solutions and communication patterns influence the treatment of standard cases, which enhances their success rate.

Implications

This finding has implications for the literature on organizational learning, especially the tradition that examines the relation between cumulative experience and success. There has been much progress in terms of determining learning curves for various industries and processes (see, Argote, 1999; Argote and Ingram, 2000). The studies in this tradition have begun to disentangle what underlies the transfer of knowledge (Argote and Ingram, 2000; Almeida and Kogut, 1999; Szulanski, 1996), to examine interorganizational learning (Argote and Ophir, 2002; Ingram and Baum, 1997) and the diversity and type of experiences that lead to learning (Tyre and Von Hippel, 1997; Haunschild and Sullivan, 2002; Darr et al., 2005; Hoang and Rothearmel, 2005; Baum and Dahlin, 2007; Wiersma, 2007). However, relatively little is known about learning curve moderators, especially those under the control of the firm's management.[11] Hence, we provide greater insight into why certain strategic choices (e.g., the proportion of difficult products in a portfolio) enable some firms to learn quicker than others.

The implications of our results, however, are much broader. They tie in with other studies that have theorized about how short term pressures may sometime tempt organizations to adopt choices that lead to suboptimal results in the long run. Benner and Tushman (2002; 2003), for example, showed how the adoption of a process management system, such as ISO9000, intended to boost quality and efficiency, may in the long run lead to a decline in firm innovation, which

-

could potentially offset these short-term benefits. Similarly, downsizing programs may cut costs and boost a firm's short-term profits, but research by Guthrie and Datta (2008) indicated that firms are usually worse off in the long run, for instance due to lower commitment and increased turnover rates amongst the remaining employees (e.g. Trevor and Nyberg, 2008). Others studies document the unanticipated effects of discontinuing certain activities – for instance due to outsourcing – on the development of an organization's capabilities (Cohen and Levinthal, 1990; Macher and Boerner, 2006; Weigelt, 2009; Reitzig and Wagner, 2010). Our study provides another example that difficult activities, which seemingly are not attractive for firms to perform (e.g. due to low margins), can have indirect positive effects on an organization's performance in the form of learning benefits, which can make their undertaking worthwhile.

Limitations

Our study points to trade-offs. The firms in our sample had to choose between boosting their success rate by selecting out difficult cases, and foregoing on the long-term learning benefits that they entail. However, there were several other costs that we did not observe. For example, we did not examine the full financial implications for the organizations in our sample of engaging in selection at the gate. Dealing with difficult cases requires resources. As one interviewee put it "complex cases are very time consuming". Another one said, "they need more care, they need more counselling, they need more interventions, they need more monitoring". On the other hand, prescreening patients, with the intention of selecting out the difficult ones, is also costly. As one interviewee, whose clinic did not select at the gate, explained: "They do more tests to exclude patients. We do very basic investigations. Other centers, they have heaps of tests done". How these various costs and benefits compare could be a valuable subject of future research. We only studied organizational success in terms of live births but we do not know how such improvements influence overall profitability. Examining the trade-off between the various costs and learning

-

opportunities, and thus disentangling the different aspects of success, would lead to a more comprehensive understanding of the issue examined in this paper.

Another question we did not explicitly examine in this paper is whether firms can also do too many difficult cases. It seems possible that if the learning benefits decrease, they no longer outweigh the costs involved. Or perhaps doing very large proportions of difficult cases even inhibits learning (cf. Haunschild and Sullivan, 2002). To get to this point to some extent, we ran additional models including squared terms of our three indicators of difficult cases, and also interacted them with clinic experience. Only in the case of patients who failed before did this lead to statistically significant results, both on the interaction with the main term (.064, p<.001) and the the interaction with its square (-.074, p<.001), suggesting seeing too many patients with prior failures starts to inhibit learning. Perhaps on the other two indicators, our clinics simply did not reach this point of saturation, but future research focusing explicitly on this topic might shed more light on the issue. It did again highlight an important nuance of our theory and empirical findings, namely that it is not simply that clinics learn from difficult cases – that would be a main effect – it is the combination of doing difficult cases and ample prior experience with standard cases that enhances an organization's success rate. Put differently, clinics that dealt with a relatively large proportion of difficult cases, benefited more from their prior experience with standard cases than clinics that focused solely on this category.

An important boundary condition of our study is that the "standard case" should still imply a reasonable level of difficulty. Settings where standard cases are very simple, with near perfect success rates, might not benefit from the learning opportunities offered by difficult cases, simply because there is not much left to learn. In contrast, even for standard IVF cases, for example, there is still ample room for improvement, since success rates are relatively low. Our findings and conclusions relate to such situations, where the more standard cases can potentially be improved

-

in terms of efficiency, error rates, or other indicators of success. It seems possible that, in such situations, some firms might opt to deliberately take on difficult cases to learn; for instance project-based organizations might seek challenging client problems in order to build new skills and knowledge (Prencipe and Tell, 2001). Future research might focus on examining such cases of deliberate learning.

Following from the above, one could wonder whether clinics might be aware of the learning benefits difficult cases entail and, if so, to what extent they act upon it. Although we do not have any quantitative evidence on the issue, our interview data clearly suggested that although some practitioners recognize the learning benefits (as per Table 1), certainly not all of them do. Where there were a fair number of people who saw learning advantages, the majority of them did not. When we asked clinics with lenient selection criteria for the reason behind their policy, they never quoted learning benefits as a reason, but invariably focused on ethical considerations. As one interviewee put it, "There's a commercial pressure and there's an ethical commitment that struggle to always balance together". The awareness of potential learning benefits that particular individuals displayed did not necessarily translate into awareness and admission policy at the clinic level. However, the clinics that admitted difficult cases – largely, as they claim, for ethical reasons – unintendedly boosted their learning benefits, making their success rate higher in the long run.

-

References

Abdalla, H. 2008. The use of frozen IVF cycles in assisted conception in the U.K. Working paper.

Almeida, P. and B. Kogut. 1999. Localization of knowledge and the mobility of engineers in regional networks. *Management Science*, **45**(7) 905-917.

Argote, L. and D. Epple. 1990. Learning curves in manufacturing. *Science*, **247**(4945) 920-924.

Argote, L. 1999. *Organizational Learning: Creating, Retaining and Transferring Knowledge*. Kluwer Academic Publishers, Norwell, MA.

Argote, L. and P. Ingram. 2000. Knowledge transfer: A basis for competitive advantage in firms. *Organizational Behavior and Human Decision Processes*. **82**(1) 150-169.

Argote, L. and R. Ophir. 2002. Intraorganizational learning. *The Blackwell Companion to Organizations*. 181-207.

Argyris, C. and D. Schön. 1978. *Organizational Learning*. Addison-Wesley, London.

Argyris, C. 1976. Single-loop and double-loop models in research on decision making. *Administrative Science Quarterly*. **21**(3) 363-376.

Bain & Company. 2010. Our clients outperform the market 4 to 1: Helping make companies more valuable. Available at www.bain.com, last accessed 06.12.11.

Bannert, M. 2002. Managing cognitive load—recent trends in cognitive load theory. *Learning and Instruction*. **12**(1) 139-146.

Barkema, H. G., O. Shenkar, F. Vermeulen, J. H. J. Bell. 1997. Working abroad, working with others: How firms learn to operate international joint ventures. *Academy of Management Journal*. **40**(2) 426-442.

Barkema, H. G. and F. Vermeulen. 1998. International expansion through start-up or acquisition: A learning perspective. *Academy of Management Journal*. **41**(1) 7-26.

Baum, J. A. C. and K. B. Dahlin. 2007. Aspiration performance and railroads' patterns of learning from train wrecks and crashes. *Organization Science* **18**(3) 368.

Ingram, P., and J. A. C. Baum. 1997. Opportunity and constraint: Organizations' learning from the operating and competitive experience of industries. *Strategic Management Journal*. **18**(7) 75-98.

British Broadcasting Corporation (BBC). 2007. IVF undercover: Transcript of interview with Mohamed Taranissi. *Panorama*. Available at http://news.bbc.co.uk/2/hi/programmes/panorama/6264997.stm, last accessed 05.13.11.

Birkinshaw, J. 1997. Entrepreneurship in multinational corporations: The characteristics of subsidiary initiatives. *Strategic Management Journal*. **18**(3) 207-229.

Cohen, W. M. and D. Levinthal. 1990. Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly*. **35**(1) 128-152.

-

Darr, E., L. Argote, D. Epple. 1995. The acquisition, transfer and depreciation of knowledge in service organizations: Productivity in franchises. *Management Science*. **41**(11) 1750-1762.

Dutton, J. M., A. Thomas. 1984. Treating progress functions as a managerial opportunity. *Academy of Management Review*, **9**(2) 235-247.

Edmondson, A. C. 1999. Psychological safety and learning behavior in work teams. *Administrative Science Quarterly*. **44**(2) 350-383.

Edmondson, A. C., R. M. Bohmer, G. Pisano. 2001. Disrupted routines: Effects of team learning on new technology adaptation. *Administrative Science Quarterly*. **46**(4) 685-716.

Edwards, R. G. and P. C. Steptoe. 1980. *A Matter of Life: The Story of a Medical Breakthrough*. Hutchinson, London, UK.

Gibson, C. and F. Vermeulen. 2003. A healthy divide: Subgroups as a stimulus for team learning behavior. *Administrative Science Quarterly*. **48**(2) 202-239.

Haleblian, J. and S. Finkelstein. 1999. The influence of organizational acquisition experience on acquisition performance: A behavioral learning perspective. *Administrative Science Quarterly*. **44**(1) 29-31.

Hargadon, A. and R. I. Sutton. 1997. Technology brokering and innovation in a product development firm. *Administrative Science Quarterly*. **42**(4) 716-749.

Haunschild, P., B. N. I. Sullivan. 2002. Learning complexity: Effects of prior accidents and incidents on airlines' learning. *Administrative Science Quarterly*. **47**(4) 609-643.

Hayward, M. L. 2002. When do firms learn form their acquisition experience? Evidence from 1990-1995. *Strategic Management Journal*. **23**(1) 21-39.

Haveman, H. A. 1993. Organizational size and change: Diversification in the savings and loan industry after deregulation. *Administrative Science Quarterly*. **38**(1) 20-50.

Henderson, R., I. Cockburn. 1996. Scale, scope, and spillovers: The determinants of research productivity in drug discovery. *The RAND Journal of Economics*. **27**(1) 32-59.

Human Fertilisation and Embryology Authority (HFEA). 2002. *The Patients' Guide to IVF Clinics*. HFEA, London, UK.

Hoang, H. A., F. T. Rothaermel. 2005. The effect of general and partner-specific alliance experience on joint R&D project performance. *The Academy of Management Journal*. **48**(2) 332-345.

Ingram, P., J. A. C. Baum. 1997. Opportunity and constraint: Organizations learning from the operating and competitive experience of industries. *Strategic Management Journal*. **18**(1) 75-98.

Jelinek, M. 1979. *Institutionalizing Innovation: A Study of Organizational Learning Systems*. Praeger Publishers, New York.

Johnson, A., T. El-Toukhy, S. K. Sunkara, M. Khairy, A. Coomarasamy, C. Ross, S. Bora, Y. Khalaf, P. Braude. 2007. Validity of the in vitro fertilisation league tables: Influence of patients' characteristics. *BJOG*, **114**(12) 1569-1574.

Kim, J., A. S. Miner. 2000. Crash test without dummies: A longitudinal study of interorganizational learning from failure experience in the US commercial banking industry, 1984–1998. *Academy of Management Best Paper Proceedings*.

Kolb, D. A. 1984. *Experiential Learning*. Prentice-Hall: Englewood Cliffs, N. J.

Levitt, B., J. G. March. 1988. Organizational learning. *Annual Review of Sociology*. **14**(1) 319-338.

Lim, A. S. and M. F. Tsakok. 1997. Age-related decline in fertility: A link to degenerative oocytes? *Fertility and Sterility*. **68**(2) 265-271.

Macher, J. T., C. S. Boerner. 2006. Experience and scale and scope economies: trade-offs and performance in development. *Strategic Management Journal*. **27**(9) 845-865.

March, J. G. 1991. Exploration and exploitation in organizational learning. *Organization Science*. **2**(1/Special Issue) 71-87.

Mihm, J., C. Loch, A. Huchzermeier. 2003. Problem-solving oscillations in complex engineering projects. *Management Science*. **49**(6) 733-750.

Ocasio, W. 1997. Towards an attention-based theory of the firm. *Strategic Management Journal*. **18**(S1) 187-206.

Papke, L. E., J. M. Wooldridge. 1996. Econometric methods for fractional response variables with an application to 401 (K) plan participation rates. *Journal of Applied Econometrics*. **11**(6) 619-632.

Papke, L. E., J. M. Wooldridge. 2008. Panel data methods for fractional response variables with an application to test pass rates. *Journal of Econometrics*. **145**(1-2) 121-133.

Pelinck, M. J., A. Hoek, A. H. M. Simons, M. J. Heineman. 2002. Efficacy of natural cycle IVF: a review of the literature. *Human Reproduction Update*. **8**(2) 129-139.

Pisano, G. P., R. M. J. Bohmer, A. C. Edmondson. 2001. Organizational differences in rates of learning: Evidence from the adoption of minimally invasive cardiac surgery. *Management Science* **47**(6) 752-768.

Reagans, R., L. Argote, D. Brooks. 2005. Individual experience and experience working together: Predicting learning rates from knowing who knows what and knowing how to work together. *Management Science*. **51**(6): 869.

Schilling, M. A., P. Vidal, R. E. Ployhart, A. Marangoni. 2003. Learning by doing something else: Variation, relatedness, and the learning curve. *Management Science*. **49**(1) 39-56.

Sharif, K., M. Afnan. 2003. The IVF league tables: Time for a reality check. *Human Reproduction* **18**(3) 483-485.

Sorenson, O. 2003. Interdependence and adaptability: Organizational learning and the long-term effect of integration. *Management Science*. **49**(4) 446–463.

Sweller, J. 1989. Cognitive technology: Some procedures for facilitating learning and problem solving in mathematics and science. *Journal of Educational Psychology*. **81**(4) 457-466.

-

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

Takeuchi, S., H. Minoura, T. Shibahara, X. Shen, N. Futamura, N. Toyoda. 2000. In vitro fertilization and ontracytoplasmic sperm injection for couples with unexplained infertility after failed direct intraperitoneal insemination. *Journal of Assisted Reproduction and Genetics.* **17**(9) 515-520.

Tyre, M. J., E. von Hippel. 1997. The situated nature of adaptive learning in organizations. *Organization Science.* **8**(1) 71-83.

Vermeulen, F., H. G. Barkema. 2001. Learning through acquisitions. *Academy of Management Journal.* **44**(3) 457-476.

Walsh, J., C. Henderson, J. Deighton. 1988. Negotiated belief structures and decision performance: An empirical investigation. *Organizational Behavior and Human Decision Processes.* **42**(2) 194-216.

Weick, K. E., K. M. Sutcliffe, D. Obstfeld. 2005. Organizing and the process of sensemaking. *Organization Science* **16**(4) 409-421.

Weigelt, C. 2009. The impact of outsourcing new technologies on integrative capabilities and performance. *Strategic Management Journal.* **30**(6) 595-616.

Wilson, J. R., A. Rutherford. 1989. Mental models: Theory and application in human factors. *Human Factors.* **31**(6) 617-634.

-

Table 1: Qualitative Evidence of how Difficult Cases Enhance the Process of Learning

Subprocesses of learning	How they enhance the firm's success rate	Interview quotes
Exploration / experimentation	Leads to improved skill levels in individuals	One interviewee described that her clinic: *"has a system in place for nurses and junior doctors, and the way this has been done seems to work better than giving them only easy jobs… For example, Gina has now received a case with poly-cystic ovarian syndrome, which is one of the toughest diagnostics to work with; all her previous patients were young, straightforward cases which responded well [to drugs]. She's got the hang of it from those simple cases, but she needs the challenge to perfect her skill, to understand the various nuances and the subtleties of this job."*
	What is learned from difficult cases aides the relatively simple cases	A doctor explained: *"What you see in the textbook or in the code of practice are treatment coordinates for standard cases, the typical patients showing up for consultation, young couples under 35, with good egg reserve, good sperm and good health… We have a lot of experience with them and they're easy cases where we rarely deviate from the standard procedure. And that's all fine, but when you get a difficult case, with complex pathology, and the standard procedure simply doesn't fit, what do you do? You change the practice, you start tinkering with the parameters, adding new things, adjusting doses and sequences so that it fits. And is that all? No, it isn't; you start tinkering with the procedure for the easy case as well; you take what you've learned from that difficult case to the easy case."*
	Difficult cases deepen understanding	One IVF consultant stated: *"I think those difficult cases teach us much better how to do our job, how to understand the real depth of infertility as a medical condition, how to acknowledge our ignorance in order to overcome it. If you don't let the bad cases in, to teach you failure, to teach you pressure, you'll oversimplify, you'll miss many of the underlying causes."*
	Focuses attention and solution finding	A doctor in a clinic that sees a high proportion of patients of advanced maternal age and thus lower chances of successful treatment, stated: *"Treating older women who constantly remind their doctors that they are running out of time builds a feeling of urgency, a feeling of purpose in all those that enter in contact with them; they need the treatment fast and they need it done well!"*
Reflection & coordination	Difficult cases lead to new solutions and improve coordination	*"Clinics which admit more older women tend to be more experimental in the therapies that they offer. The effort of treating such patients –and patients with poor prognosis in general–intensifies the interaction among our doctors, embryologists and nurses."*

	Improved coordination also aides the more simple cases	*"Clinics which admit more difficult cases tend to be more disciplined and more thorough in their work. And of course, even for us, the effort of treating [difficult patients] intensifies the interaction among our doctors, embryologists and nurses. And we tend to take that with us, and to do it for the next patient which enters our office."*
	Reflection and interaction leads to enhanced understanding	One of the doctors working at a clinic in London contrasted the its patient base to that of a clinic in which he had worked that was in a provincial area: *"There I had mostly Caucasian patients, with less health issues than our patients here in [East London], where I see much more Africans than I ever saw in [Yorkshire]… also patients from the young urban population here are more likely to have pelvic infections, not using condoms, having unprotected sex, are more likely to get Chlamydia and all those things with tubal problems… So I see more pathologies, more problems and so on… but all these problematic cases add to our experience as doctors, it makes us talk to embryologists, to pharmacists; it matures us, it keeps us understand things, the physiology of different races and diseases, how drugs work for them, what their medical predispositions are".*
	They stimulate collective reflection and learning	*"If we have unusual cases or adverse outcomes, then we have regular clinical meetings, look at the cases, pull them to pieces and everybody tries to learn from those"*
The translation of what has been learned into processes and procedures	Knowledge capture	*"It's hard but treating severe cases comes with its rewards. I'm not talking only about the thrill of cracking a difficult case, I'm talking about the careful checklists that you put together and the resilience that you develop as you do that. Baby or no baby, the checklists and the ideas you tried stay with you, you'll try them again for less complicated cases again and again. Anything that leaves less to chance is worth trying again".*
	Knowledge capture and transfer	Several interviewees referred to the updating of clinical protocols and departmental interfaces based on experience with difficult cases. A quality control manager described it thus: *"Doctors have checklists; the more difficult their cases, the longer the checklists. And I am interested in their checklists because I want to revise mine and bring the system up to date. Are we getting cases with a new bullet point? Then I want to know about it, the other doctors want to know about it, the nurses as well."*
	Knowledge capture in procedures and technology	Interviewees gave numerous examples of spillovers from the more complex to the less problematic cases; most emphasized knowledge transfer and use of diagnostic tools and equipment. One

-

experienced doctor described how a catheter typically employed in difficult embryo transfers had become the tool of choice for most doctors in his clinic, regardless of the difficulty of the case: *"Because we have experienced many two-stage transfers in older age groups, our medical director has authorized the purchase of several Wallace Pro-Ultra catheters. These catheters made the two-stage procedure so much easier under ultrasound, that most of us are using it now for the ordinary single-stage transfers."*

Knowledge transfer to relatively simple cases	Another example of knowledge transfer from difficult to more straightforward cases was described by an embryologist related to involvement in a case with a history of treatment failure: *"She had beautiful embryos, symmetrical, with equally sized blastomeres, perfect for textbook illustrations… we couldn't understand why they didn't implant... but then we did the PGD test and discovered chromosomal abnormalities in three of them... Now when I see perfectly symmetrical embryos I don't get as excited as I used to. I always think 'they should do a PGD on those!'"*

\-

Table 2: Qualitative Evidence Regarding League Tables, Selection, and Difficult Patients

Concept	Interview quotes
League tables	Many interviewees commented that competition between clinics is fierce, and that the main variable of comparison is their success rate: *"The pressure is mainly to ensure that the success rates are comparable with the natural success rates. That's the most important thing. ... showing that you've got good success rates. ... Patients look at your success rate as your top-line".* *"If one IVF clinic has a success rate 5% or 10% higher than another, patients will notice this and that will have a big commercial effect on that clinic. ... To what extent they're truly comparable from one clinic to another is debatable, but, certainly, patients treat them as a league table".* *"If you don't treat poor prognosis patients, then your pregnancy rate per age group will be better. So, therefore, if you look at the league table publication by the HFEA, you will look better than a clinic which is just as good but who may treat poorer prognosis patients. You will appear higher in the league table, and therefore the interpretation of patients may be that [your clinic] is better".* *"The best way to move yourself up the table is to treat prognostically the better group of patients. If you have a group of patients that are poor prognosis, if where you come in the league table is important to your practice, you will not give those patients a chance. So, patient selection is critical".*
Selection at the gate	When asked why some clinics seem to select out difficult patients, one doctor, heading an Academic Unit of Reproductive and Developmental Medicine, commented: *"The driver [of selection] is success rates".* A senior doctor in a private clinic added: *"One of the competitive elements is success rate, inevitably, and you've got to do all that you possibly can to maximise your outcome and one of those, of course, is to select your patient".* *The only reason why any private clinic would refuse a private patient is purely for the success rates. They don't want these patients to dilute their success rates. ... What they do is they choose patients just so they are at the top.* (A nursing manager) Another senior doctor said: *"Some clinics, yes, do look at a patient's history and decline to treat certain patients knowing that they're unlikely to become pregnant. Other clinics, for sure, accept any patient, as long as they think there's some chance of getting a pregnancy".* *"It does seem to be the case that some clinics will turn down patients who don't fit certain parameters".*

Difficult patients	When asked what would constitute a difficult patient, and how you would recognisze her, interviewees generally brought up the factors represented in our measures: *"What you might say is a difficult patient, is a patient who, once you've looked at their history, you suspect that they're not going to have as high a chance of pregnancy as another patient who on the surface of it may seem similar"*.
	"I know some clinics who refuse treatment based on patient age or other prognostic factors".
	"So the most obvious factor is maternal age; that's the single biggest factor that affects IVF success rate. But there are other factors, too, like, for example, how many times has the patient attempted IVF in the past".
	"How to turn an average clinic result into a super clinic result: You simply take out the women who have a low ovarian reserve. That's been picked up [ ] by a number of clinics, so they just don't treat women who've got a low ovarian reserve".

Table 3: Descriptive Statistics and Correlations

| Variable	Mean	S.D.	Min	Max	1	2	3	4	5	6	7
1. Success rate for standard patient group (under 35s)	.254	.114	0	0.667							
2. Clinic size	457	434	1	2271	.203						
3. ICSI technology	.655	.475	0	1	.501	.077					
4. Industry IVF experience	11.4	1.06	8.57	12.5	.489	.056	.729				
5. Clinic IVF experience	6.22	2.27	0	9.48	.417	.588	.473	.589			
6. Patients above the age of 35	.460	.121	0	1	.383	.146	.418	.384	.325		
7. Patients who failed before	1.16	.080	1	1.49	.118	-.098	-.067	-.088	-.099	-.103	
8. Patients who produce few eggs	.784	.198	-1.83	1	.073	.087	-.056	-.011	.023	.080	.194

n = 1004 clinic/years

Table 4: Regressions of Clinic Success Rate on the Proportion of Challenging Cases in the Patient Mix

| Variable	model 1	model 2	model 3	model 4	model 5
Clinic size (per 1000 patients)	0.110***	0.110***	0.103***	0.102***	0.106***
	(0.030)	(0.030)	(0.031)	(0.031)	(0.031)
Clinic size – quadratic	-5.06e-05***	-5.54e-05***	-5.02e-05***	-4.71e-05***	-5.29e-05***
	(1.47e-05)	(1.47e-05)	(1.49e-05)	(1.47e-05)	(1.59e-08)
ICSI technology	0.049***	0.040***	0.050***	0.049***	0.041***
	(0.006)	(0.007)	(0.006)	(0.006)	(0.007)
Industry IVF experience	0.011*	0.009†	0.009†	0.012*	0.009†
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
IVF experience	0.011**	-0.009	-0.034†	0.001	-0.053**
	(0.004)	(0.007)	(0.019)	(0.007)	(0.020)
Patients above 35		-0.221*			-0.224*
		(0.103)			(0.103)
Patients who failed before			-0.226*		-0.192†
			(0.107)		(0.107)
Patients who produce few eggs				-0.024	-0.023
				(0.037)	(0.036)
Patients above 35 X IVF experience		0.048***			0.047***
		(0.015)			(0.015)
Patients who failed before X IVF experience			0.040**		0.031*
			(0.017)		(0.017)
Patients who produce few eggs X IVF experience				0.012*	0.011*
				(0.007)	(0.007)
Constant	0.005	0.117	0.279*	0.009	0.355*
	(0.049)	(0.061)	(0.135)	(0.048)	(0.144)
N_t (total clinic-years)	1004	1004	1004	1004	1004
N (total clinics)	105	105	105	105	105
F statistic	67.86***	89.94***	65.40***	66.01***	44.72***
Clinic fixed effects	Yes	Yes	Yes	Yes	Yes

† $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$ (one-tailed tests if hypothesized; two-tailed tests otherwise); standard errors in parentheses.

Figure 1: IVF Success Rates
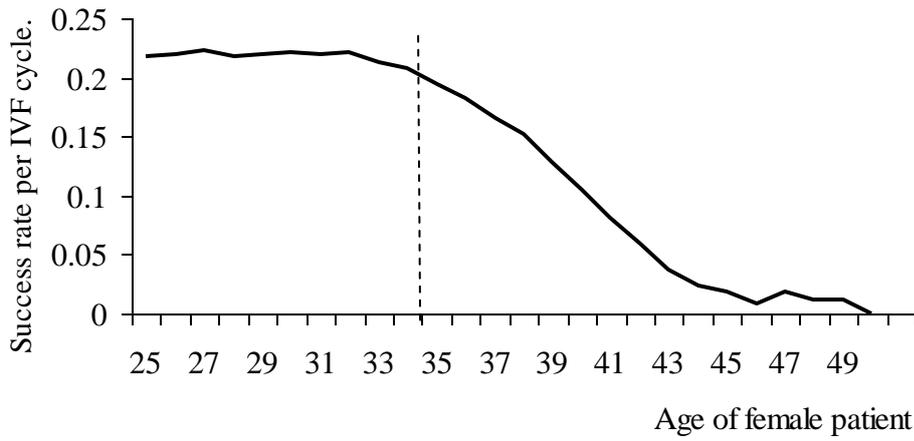
* Note: HFEA Long Term Data Analysis, 1991-2005

Figure 2: The Influence of Treating Difficult Cases on a Clinic's Success Rate

-
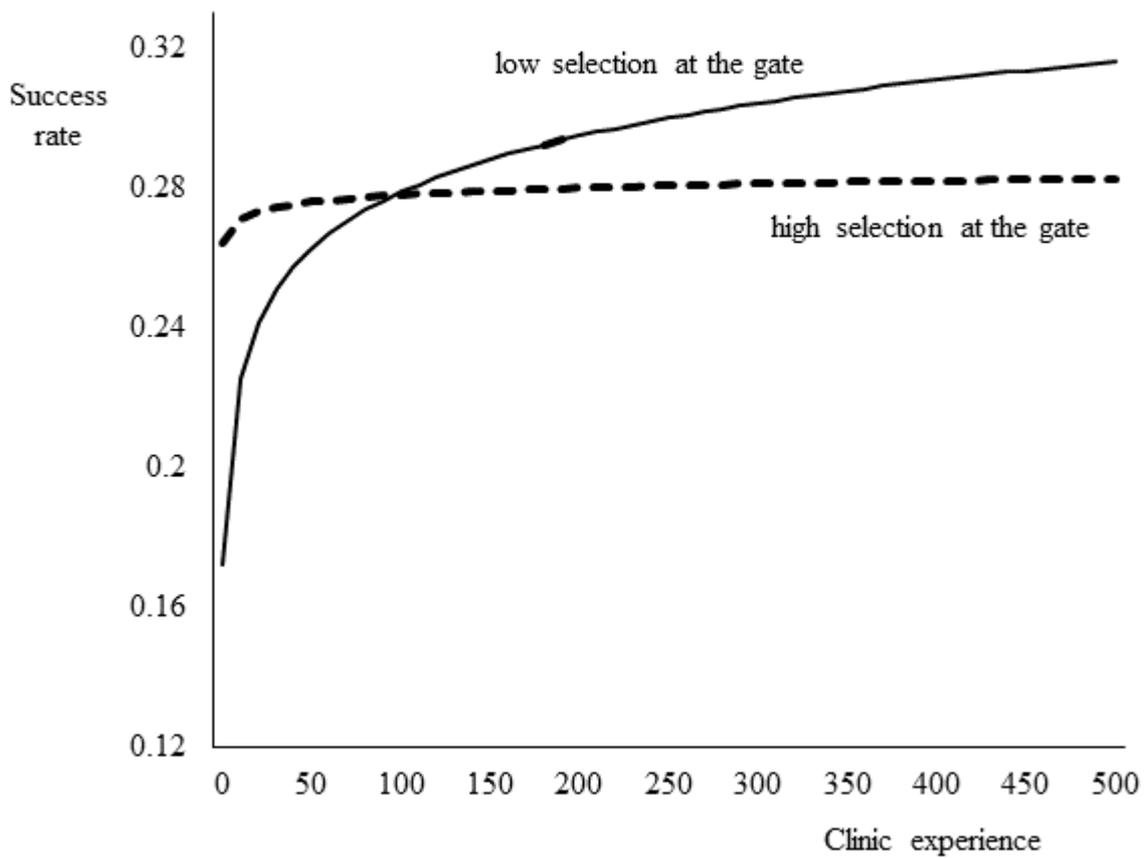
Endnotes

[1] To qualify for treatment in an NHS clinic women must be under 39 years of age and be childless. These selection criteria apply to all NHS IVF clinics and do not constitute a source of variation among firms.

[2] If input selection, through the learning curve, has an impact on clinic success rates rather than the other way around, the positively moderating impact of taking on complex cases should be visible also for these NHS clinics.

[3] The division by age in reporting the results is consistent across clinics. The exact ages of all the patients treated are unknown, so we cannot compute the mean or median, but we know how many of a clinic's patients are in the below 35 year old category (the standard group) and how many aged over 35 years (considered more difficult cases).

[4] A minority of clinics did not offer the option of using frozen cycles throughout the entire period of observation. Excluding those observations (90 clinic-year observations) from the analysis, did not alter any of the results or models in this paper.

[5] In the field, this is considered a more controversial measure than the other two, because it is not impossible that obtaining a larger number of oocytes also potentially decreases their quality. Nevertheless, if a woman produces very few oocytes, it is generally considered a bad sign in terms of prognosis.

[6] The main negative effect of treating a higher proportion of older patients (> 35 years) is somewhat surprising, since our dependent variable is related to the success rate among patients aged under 35 years. Perhaps clinics that are lenient about admitting older patients are also more lenient toward patients with a lower ex-ante success rate for other reasons (unobserved in our models), which would explain the overall negative effect.

[7] Note that, on average, clinics treat 116 cases in their first year of operation. Hence, it takes around a year for the "low selection at the gate" clinics to catch up with the "high selection at the gate" organizations.

[8] When we reran models 1-5 on the 606 observations for which we have data on specialists, with the inclusion of the number of specialists as an additional control variable, the predictors patients above 35 and patients who have failed before lost their statistical significance, yet, patients with few eggs remained significant, both on the main term ($-.36$, $p<.05$) and in interaction with clinic experience ($.06$, $p<.01$). When we included the number of specialists as an additional control, the results were essentially identical ($-.36$, $p<.05$; $.06$, $p<.01$). This again strengthens the conclusion that the potential alternative explanation centered around specialists is not driving our results, offering further support for our theory.

[9] Running a random effects model showed that university-affiliated clinics have slightly lower success rates (-2.7%, $p<.05$), perhaps because they accept more difficult cases. An interaction with cumulative experience remained insignificant, suggesting that they do not learn quicker than others (over and above the effect of dealing with difficult cases). All estimates of our predictors were virtually unaffected by the inclusion of these variables.

[10] We also replicated models 2-5, including an interaction between our predictors and industry experience, to test whether dealing with difficult cases also enables clinics to absorb and benefit from the experience of others better. Yet, this appeared not to be the case; the interactions always were insignificant. This might be because general knowledge in the field also spreads effectively through other means, such as medical training, conferences, and literature.

[11] For an exception, see Pisano, Bohmer and Edmondson (2001) who, in a study of two hospitals, found that the use of various management processes (formal procedures, cross-functional communication, feedback activities, team stability) enhanced learning.