Organization Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

# Selection at the Gate: Difficult Cases, Spillovers, and Organizational Learning

Mihaela Stan, Freek Vermeulen,

Organization Science
Vol. 24, No. 3, May–June 2013, pp. 796–812
ISSN 1047-7039 (print) | ISSN 1526-5455 (online)

informs
http://dx.doi.org/10.1287/orsc.1120.0763




# Selection at the Gate: Difficult Cases, Spillovers, and Organizational Learning

Mihaela Stan
Management Science and Innovation, University College London, London WC1E 6BT, United Kingdom, m.stan@ucl.ac.uk

Freek Vermeulen
London Business School, London NW1 4SA, United Kingdom, fvermeulen@london.edu

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 compared with those firms that admit difficult cases, because the latter develop steeper learning curves. That is, difficult cases enable greater learning from prior experience because they promote experimentation, communication among 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 with those of firms that do select at the gate.

*Key words*: organizational learning; longitudinal research; organizational capabilities
*History*: Published online in *Articles in Advance* July 17, 2012.

## Introduction

League tables and rankings publicizing the quality of firms' output are an important phenomenon in many industries. Institutions such as business schools, law firms, and medical clinics often publish their performance measures in the public domain. However, the quality of an organization's output is often substantially determined by the quality of its input. Consequently, organizations often 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 preselection 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 it 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 find jobs, and management consultants. For example, the business consulting firm Bain & Company 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 organization we examine in this paper, in vitro fertilization (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 (*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 suggest 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 experience from 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, over 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 rates through selection at the gate "shot themselves in the foot."

Our paper makes four contributions to the literature. First, we provide insights on and 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 et al. 1990, Darr et al. 1995, Dutton and Thomas 1984, Mihm et al. 2003). The analysis has been extended from research into 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 identification of the types of experience that 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 a firm's portfolio in terms of the complexity of its cases, to show why some firms learn more quickly than others—a topic that is relatively unexplored in this literature.

The third contribution of this paper lies in the broader implications of our findings. Firms face various pressures and incentives to optimize short-term results, actions that 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 performance. The choice of some IVF clinics to turn away difficult cases has similar ramifications; although some cases initially may seem less attractive and less profitable, their spillover benefits could make them profitable elements of the firm's portfolio. Our study shows that it is important to take account of these indirect, long-term benefits.

The fourth contribution 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, but they are attracting more challenging cases because they are getting better. For example, the cancer units in very reputable institutions often display higher mortality rates because patients with the most complex etiologies (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 using 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. 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 accept relatively few difficult cases—through three variables: (1) the proportion of women with a relatively poor prognosis because they have failed to conceive under 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 older (35 is the industry's standard cutoff rate). All these categories are 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 those firms that take on more difficult cases improve their success rates more quickly.

## 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 et al. 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 measure whether the success rate of IVF clinics increases with experience. "Success" in this business (IVF clinics) is very clearly defined: an IVF treatment that results in a birth is a success; otherwise, it is a failure. Hence, in our analysis, the measure of an organization's success rate is the proportion of live births among the patients receiving IVF treatment. Experience refers to the cumulative number of patients treated with IVF by the organization. Specifically, we examine whether the success rates of the organizations increase more or less quickly depending on the proportion of difficult cases that they treat.

## The Process of Learning

Studies examining the process of learning constitute another strand in the organizational learning literature. The learning process is often seen as consisting of multiple, interdependent stages representing the search for, choice of, 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 to the accumulation of new knowledge (Zollo and Winter 2002). Prencipe and Tell (2001) refer to this as a process of 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), through the process 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), which Prencipe and Tell (2002) referred to as learning by adapting, implementing, and replicating.

These processes, in combination, form a learning cycle (Gibson and Vermeulen 2003), which, we posit, is enhanced by dealing with difficult cases. Analyzing airlines' learning from accidents involving heterogeneous causes, Haunschild and Sullivan (2002) argued that such problems cause 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 the actors, which leads to a richer understanding of the problem. Similarly, we argue that dealing with nonstandard, difficult cases forces organizations and the people involved in these cases to analyze the underlying structure of a problem in more depth, thereby enhancing their understanding of the issue and changing how they deal with it and how they exploit 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 complex characteristics of the new case may mean that standardized approaches are not effective. It 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 new practices and allows those involved to apply and evaluate novel solutions that 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 cases will be viewed in a new, more comprehensive light and with a better understanding of what Haunschild and Sullivan (2002, p. 614) call "the underlying structure of the problem." Thus, when organizations employ less stringent selection at the gate, novel challenges and opportunities are more likely to emerge. As a result, the firm's general competence to address new cases increases with each unfamiliar case. Experience in dealing only with standard cases, in contrast, will not lead to richer understanding and might even be a disincentive to trying out new solutions.

Some of the clinics where we conducted in-depth interviews for this project seemed to deliberately use complex cases for training purposes (see Table 1 for quotes). These difficult cases were introduced to confront people with unfamiliar situations and new ways of doing things. Hence, dealing with complex cases enhances the general abilities and skill levels of the people involved and leads to a deeper understanding, which promotes the development of new solutions, which 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

**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 polycystic 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 aids 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 and 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 aids 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 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." |

**Table 1 (cont'd)**

| Subprocesses of learning | How they enhance the firm's success rate | Interview quotes |
|---|---|---|
| | Knowledge capture in procedures and technology | Interviewees gave numerous examples of spillovers from the more complex to the less problematic cases, and 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 [preimplantation genetic diagnosis] 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!'" |

> 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 overreliance on a standard set of operating procedures. It motivates the firm to experiment to try to enhance overall performance. Haunschild and Sullivan (2002) suggested that the attention of the organizational members involved in dealing with a complex problem forces situational analyses that 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 for the execution of simpler cases. Subsequent experiences are evaluated on the basis of the new cognitive model and set of solutions.

Difficult cases also stimulate learning by heightening the 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 with one another and to coordinate. In the context of standard cases, coordination is achieved through adherence to established routines and procedures. Nonstandard 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 process 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 cycle. Methods and procedures developed under difficult circumstances in trying to solve complex problems may improve performance in standard cases. For example, in an experiment, Schilling et al. (2003) 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 more quickly. 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,

> 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 higher proportions of difficult cases. Formally stated,

Hypothesis 1. *The relationship between experience and success will be more positive for organizations with higher proportions of difficult cases.*

This means that organizations that select tightly 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 will lead to less experimentation, lower levels of interaction, and less knowledge codification. 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 more stringently 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 IVF baby was born in 1978 as the result of the work of two British scientists, Edwards and 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 United Kingdom. The HFEA allowed us to trace back data on variables such as experience and success for the population of all fertility clinics based in the United Kingdom since 1991, one year prior to the introduction of IVF as an authorized treatment, up to 2006 (the final year made available). We also conducted 32 face-to-face interviews with people in the industry to supplement our quantitative analysis; 10 of these were conducted after completing the first draft of this paper.

More than two-thirds of UK IVF clinics are private, usually owned by the doctors operating them; the remainder are state owned (by the National Health Service). Among the former, there is, as one respondent put it, "fierce competition between clinics." Patients can choose from which clinic they want to buy treatment and, in the process, many consult the annual *Patients' Guide* published by the HFEA. This league table—as it is generally referred to in the industry—lists the clinical results for every UK IVF clinic and ensures that success rates are publicly available and easy to access. For example, in 2010, when 45,264 patients received IVF treatment in the United Kingdom, the *Patients' Guide* website recorded more than 600,000 hits.

Table 2 presents some interview quotes on the roles of the league table and selection at the gate. These extracts show that the existence of the league table puts pressure on clinics to present good success rates. It is a motivation to select patients with ex ante higher chances of success. As one respondent put it, "The best way to move yourself up the table is to treat prognostically the better group of patients." Interviewees indicated that all clinics do this to a greater or lesser extent; some are highly selective, whereas others are less restrictive.

For a variety of reasons, this setting and these data are ideal for testing the relationship between experience and success. First, this is so 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 a clear 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 various 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% of the clinics in the population and hence complete data on their prior experience. We reran the analyses excluding the left-censored observations, and the results were unchanged. Finally, because approximately 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, and the newest for just 1 year. The largest clinic had treated over 13,000 patients during the window of observation; the average number

**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 maximize your outcome and one of those, of course, is to select your patient." |
| | A nursing manager stated, "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." |
| | 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 one would recognize 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." |

of prior cases was approximately 4,000 patients. In total, these clinics had performed over 400,000 IVF cycles on approximately 300,000 women, who had delivered over 75,000 IVF babies by the end of 2006.

## Dependent Variable

*Success Rate.* The *Patients' Guide* shows success rates according to six patient age groups: under 35, 35–37, 38–39, 40–42, and over 43 years. Patients under 35—by far the largest group—are generally regarded in the industry as the "standard patient group" (Johnson et al. 2007), and they are used as the primary basis for comparisons between clinics because this "reduce[s] the impact of 'patient-mix' on the comparability of results between different centres" (Sharif and Afnan 2003, p. 484). Based on 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 35 and under 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 35 and under who underwent one or more fresh IVF cycles in the year of observation. We chose standard cases (i.e., women 35 and under) 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 presented in Table 4.

## Independent Variables

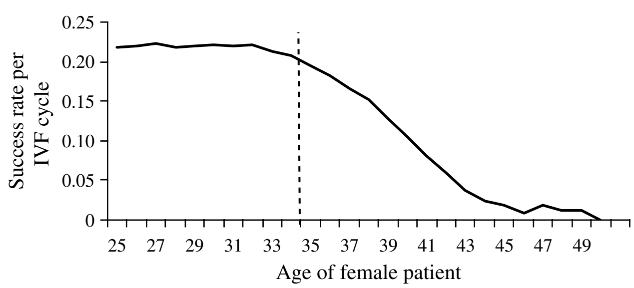
*Prior Experience.* Experience was measured by cumulating all the prior cases that involved one or more IVF cycles using the patient's own fresh eggs. The *Patients' Guide* does not provide information on the number of patients treated by each clinic in a given year; these data were obtained directly 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.[2]

To measure the level of difficulty in a clinic's patient mix, we identified three factors that IVF practitioners regard as relevant dimensions for assessing the complexity of a case: patients over 35 years old, patients that 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, because increasing age of the female reproductive system generally reduces the chance of successful pregnancy. Thirty-five years is used as the standard cutoff 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 35 and under are considered standard cases, and patients aged over 35 are seen as more difficult cases. One of our interviewees told us that "[a]ge 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.

*Patients with a History of Prior Failed Treatment.* The second variable used to measure relatively poor prognosis cases is the number of patients with unsuccessful previous treatment (*patients who failed before*). A history of prior IVF failure can signal a possible problematic underlying etiology. Patients who have had previous IVF treatment and failed to conceive may have health conditions that require further investigation and intervention, or they 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.

**Figure 1 IVF Success Rates**




*Note.* A long-term analysis of HFEA data, 1991–2005, is shown.

*Proportion of Patients with Low Egg Reserve.* The last indicator of a difficult case, *patients who produce few eggs*, builds on a patient's ovarian reserve, i.e., 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 treatment. Clinics can choose to measure this before commencing treatment. One respondent explained, "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 available to 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 2010).[5] The higher 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 are challenging in that there are more elements for clinics to address 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 ovaries and uterus change shape, 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 (customization of the hormonal doses to ensure ovarian response; performance of micromanipulation procedures 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 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 nonlinear (e.g., Haveman 1993). We use a dummy variable to control for whether the clinic has intracytoplasmic sperm injection (ICSI) 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 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 in 2005, for example, 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) and 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 United Kingdom up to the year of observation. We also reran 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. The results of both were almost identical. We present the models using total industry experience. Finally, all of 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 may not necessarily lie within this interval. To correct for this, Papke and Wooldridge (1996) proposed the fractional logit estimator, which they 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 $p < 0.10$. In the remainder of this 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 3 presents summary statistics for the variables included in our models. The correlations between the three indicators of difficult patients are the fairly low or even negative, and we conducted further analysis and interviews to try to understand the underlying reasons. One of the reasons for the relatively low correlations, for instance, between patients with previous unsuccessful treatment and patients who produce very few eggs (0.19) 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 (NHS clinics did not differentially select at the gate), the correlation is higher (0.23).

Furthermore, the variable *patients over the age of 35* differs from the other two variables because of clinic location. We took each of the three difficulty indicators as dependent variables in a simple regression model (with fixed effects), and we used gross domestic product (GDP) per capita for the specific area in which the clinic is located as the explanatory variable (representing 185 different areas within the United Kingdom). The results showed that local GDP per capita was significantly negatively associated with patients who had failed before ($-0.48$, $p < 0.01$) and patients who produce few eggs ($-1.51$, $p < 0.001$), but it was significantly positively associated with patients aged over 35 (2.76, $p < 0.001$). 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, e.g., because of higher incidences of sexually transmitted diseases. In contrast, in wealthier areas, patients tend

**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* (up to 35) | 0.254 | 0.114 | 0 | 0.667 | | | | | | | |
| 2. *Clinic size* | 457 | 434 | 1 | 2,271 | 0.203 | | | | | | |
| 3. *ICSI technology* | 0.655 | 0.475 | 0 | 1 | 0.501 | 0.077 | | | | | |
| 4. *Industry IVF experience* | 11.4 | 1.06 | 8.57 | 12.5 | 0.489 | 0.056 | 0.729 | | | | |
| 5. *Clinic IVF experience* | 6.22 | 2.27 | 0 | 9.48 | 0.417 | 0.588 | 0.473 | 0.589 | | | |
| 6. *Patients over the age of* 35 | 0.460 | 0.121 | 0 | 1 | 0.383 | 0.146 | 0.418 | 0.384 | 0.325 | | |
| 7. *Patients who failed before* | 1.16 | 0.080 | 1 | 1.49 | 0.118 | −0.098 | −0.067 | −0.088 | −0.099 | −0.103 | |
| 8. *Patients who produce few eggs* | 0.784 | 0.198 | 0 | 1 | 0.073 | 0.087 | −0.056 | −0.011 | 0.023 | 0.080 | 0.194 |

*Note.* $n = 1{,}004$ clinic-years.

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 aged over 35.

In line with the above findings, we ran an exploratory factor analysis (using principal components with varimax rotation), which showed that patients who failed before and patients with few eggs load on the same factor (with factor loadings 0.74 and 0.76, respectively). However, the second factor was composed solely of patients over 35 (0.89). Nevertheless, a combined measure of difficult patients, constructed by standardizing and averaging the three individual indicators, led to the 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 becomes negative only at +1.35 standard deviations above the mean, the estimations suggest that clinic size could have a negative influence on success rates 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 approximately 4%–5%. Overall industry experience—as a proxy for overall progress in the field—is also positive and significant. The results are very robust across the various model specifications, which, in combination with the low correlations between the explanatory variables, make us confident that multicollinearity is not an issue.

## 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 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 entailed, clinics that admit more difficult cases show a faster increase in success rates than clinics that have more selective admission procedures.

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 a clinic's experience and success when all three indicators are one standard deviation below the mean (i.e., a clinic dealing with a low proportion of difficult cases) versus when all three indicators are one standard deviation above the mean (i.e., a clinic accepting a high proportion of difficult cases). The former is labeled "low selection at the gate" and the latter "high selection at the gate." The results show clearly that treating difficult cases has an initially depressing effect on a clinic's success rate; the success rates for clinics that admit more difficult cases are as much as 10% lower than those for clinics that mainly select more promising cases. However, the graphs show that these clinics start to catch up rapidly. After some hundred cases, their success rates equal those of clinics that select heavily at the gate,[7] and subsequent success rates improve faster than those of their counterparts. After 400 cases, the overt success rate for clinics that admit more difficult cases is 3.3% higher than that of clinics that deal with small numbers of difficult patients.

## Alternative Explanations and Additional Analysis

*Reverse Causality.* We tested whether the results presented above potentially could be confounded by some

**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 1,000 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 over 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 over 35* | | 0.048*** | | | 0.047*** |
| × *IVF experience* | | (0.015) | | | (0.015) |
| *Patients who failed before* | | | 0.040** | | 0.031* |
| × *IVF experience* | | | (0.017) | | (0.017) |
| *Patients who produce few eggs* | | | | 0.012* | 0.011* |
| × *IVF experience* | | | | (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) | 1,004 | 1,004 | 1,004 | 1,004 | 1,004 |
| $N$ (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 |

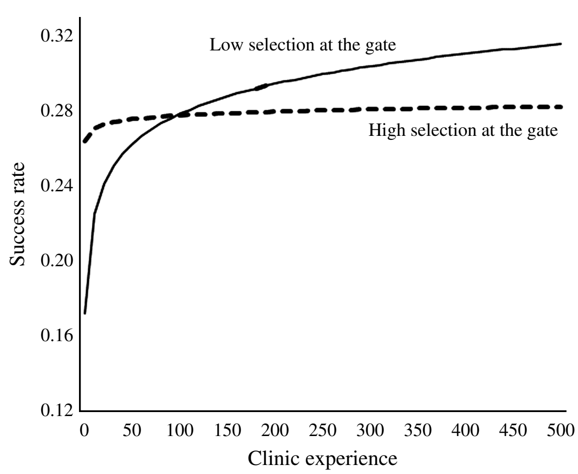
*Note.* Standard errors are in parentheses.
†$p < 0.10$; *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$ (one-tailed tests if hypothesized; two-tailed tests otherwise).

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 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 vice versa. To test directly for this possible effect, we estimate 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 referred 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 referred to a particular clinic based on their home address/post code. Therefore, because poor prognosis patients cannot choose clinics with higher success rates, there is no possibility of a reverse causality effect.

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



However, some NHS clinics deal with a substantially higher number of difficult cases due to chance but also because particular areas/post 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. Despite the smaller sample size, the coefficients of all three predicted interactions confirm our hypothesis (patients over $35 = 0.0444$, $p < 0.05$; prior failure $=$ $0.0371$, $p < 0.10$; poor egg reserve $= 0.0096$, $p < 0.05$); 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, employ more specialists to accommodate those patients. The presence of these additional 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 annual) inspections of all clinics, including collecting data on the number of specialist functions. Hence, we had access to data on the number of specialists for 606 (of a total of 1,004) year-clinic observations, representing 85 different clinics. Using the number of specialists as the 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, $p < 0.001$; which indicates that with progression in the field more specialist functions were developed), and ICSI technology drove the number of specialist down ($-1.56$, $p < 0.001$; 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. Hence, this does not support the alternative explanation for our findings.[8]

*More Technology.* Similarly, it could be conjectured that clinics that deal with more difficult patients might exploit the additional technology—in the form of equipment and specialist treatments—for standard cases, which might improve success rates and could potentially drive our results. Although it is impossible to measure all the technologies and specialist procedures, we can proxy them by ICSI treatment, and specifically to what extent clinics used this technology to treat standard patients. ICSI is an expensive procedure initially used only for difficult patients with a particular set of fertility problems; however, some clinics also started using it more widely. The HFEA provided data on how many ICSI treatments were provided by each clinics in a given year. Because we have information on how many difficult patients these clinics treated, we could add a control proxying ICSI treatment for standard patients. This measure was insignificant in all of the models; apparently, ICSI treatment does not improve success rates in standard patients. Importantly, all our results (reported in Table 4) were fully replicated.

*Better Learning Environments.* Another alternative explanation might be the better learning environment in some clinics, in the sense that some are more motivated to push the frontier of knowledge on IVF and therefore are more inclined to accept a larger proportion of difficult patients. A better learning environment might also attract better doctors, which in turn could lead to higher success rates. Although this explanation is quite well accounted for in our fixed effects models, we also tried to test and control for it more directly. The HFEA provided data on the total number of research projects each clinic engaged in each year. We added this variable as a control to our models, assuming that more research-oriented clinics would represent a more interesting learning environment for ambitious doctors.[9] The estimate of the variable, and its interaction with cumulative clinic experience, was insignificant. Importantly, it did not reduce the support for any of our predictions. Apparently, engagement in research is not a substitute for experience: it does not cause the clinic to learn more quickly.[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—eventually may disadvantage the firms that engage in this practice. The British IVF clinics in our study that admitted fewer difficult cases initially enjoyed higher success in the form of more live births per patient treated. However, this advantage disappeared quite rapidly because the clinics that treated relatively higher numbers of difficult cases learned more quickly from their prior IVF experiences, and the organizations that were more selective at the gate ultimately had significantly lower overt success rates than their more inclusive counterparts. What initially seemed an advantage became a disadvantage in the longer run; clinics that were more selective eventually found themselves "on the back foot" in terms of success rates. Although they were treating mainly patients with relatively simple etiologies, the number of live births per patient was lower over time 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 that are guided by protocols. The various medical professionals conducting the different stages are not required to meet and communicate to follow established procedures. However, being confronted with nonstandard cases forces these professionals to consider new solutions, establish new communication patterns, and alter their work methods. As Malterud (2001, p. 397), reflecting on knowledge development in medicine, put it, "Clinical knowledge consists of interpretive action and interaction—factors that involve communication, opinions, and experiences." 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 relationship between cumulative experience and success. Much progress has been made 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) and 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. 1995, 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] The present study provides additional insights into why certain strategic choices (e.g., the proportion of difficult products in a portfolio) enable some firms to learn more quickly than others.

The implications of our results are much broader, however. They are in line with other studies that theorize about how short-term pressures may 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, e.g., ISO 9000, intended to boost quality and efficiency, in the long run can 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, although research by Guthrie and Datta (2008) indicated that firms are usually worse off in the long run because of lower levels of commitment and increased turnover among remaining employees (e.g., Trevor and Nyberg 2008). Other studies document the unanticipated effects of discontinuing certain activities—for instance, as a result of 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 additional evidence that difficult activities, which seemingly are not attractive for firms to perform (e.g., because of low margins), can have indirect positive effects on an organization's performance, in the form of learning benefits, which can make their undertaking worthwhile.

## Additional Insights and Limitations

Our study shows that there are trade-offs related to firm choice. The firms in our sample had to choose between boosting success rates by selecting out difficult cases and forgoing the long-term learning benefits that the treatment of difficult cases entail. There are also 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 said that "they need more care, they need more counselling, they need more interventions, they need more monitoring." On the other hand, screening patients with the intention of selecting out the difficult ones is also costly. 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." We only studied organizational success in terms of live births, but we do not know how such improvements relate to other costs and benefits, and hence how they 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 not explicitly examined in this paper is whether firms can take on too many difficult cases. It would seem possible that if the learning benefits decrease, they will no longer outweigh the costs involved, or that undertaking high proportions of difficult cases may inhibit learning (see Haunschild and Sullivan 2002). To begin to explore this, we ran additional models including squared terms of our three indicators of difficult cases and interacted them with clinic experience. Only in the case of patients with unsuccessful previous treatment did this lead to statistically significant results, both for the interaction with the main term (0.064, $p < 0.001$) and the interaction with its square ($-0.074$, $p < 0.001$), suggesting that accepting too many patients with a history of failures starts to inhibit learning. It is possible that, for the other two indicators, the clinics in our sample had not reached saturation point, but future research with an explicit focus might shed more light on these issues.

This leads to another important nuance in our theory and empirical findings: that it is not simply that clinics learn from difficult cases—which would be a main effect—but that it is the combination of treating difficult cases and ample prior experience that enhances an organization's success rate. Some of the comments of our interviewees (summarized in Table 2) suggest, simply, that "you learn from difficult cases," but this is probably because, for people within the industry, it is difficult to disentangle the different effects cognitively. Our regression analyses show that it is the interaction that matters; clinics that deal with a relatively large proportion of difficult cases benefit more from their prior experience with standard cases than clinics that focus solely on the latter category.[12]

An important boundary condition of our study is that a "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. Purely routine work might even lead to boredom and demotivation to explore new solutions. However, in the case of standard IVF cases, there is much room for improvement because success rates are relatively low. Our findings and conclusions relate to situations where efficiency, error rates, and other indicators of success could potentially be improved for more standard cases. It would seem feasible that, in these contexts, some firms might deliberately opt to take on difficult cases in order to learn; for instance, project-based organizations might seek challenging client problems to build new skills and knowledge (Prencipe and Tell 2001). Future research could focus on cases of deliberate learning.

This leads us to speculate about whether clinics are aware of the learning benefits provided by difficult cases and, if so, to what extent this provides motivation to accept them. Although we do not have any quantitative evidence on this question, our interview data suggest that although some practitioners recognize the learning benefits (as per Table 1), this does not apply to all of them. Although a number of people recognized the learning advantages, the majority did not. Also, when we asked clinics why they operated a policy of less stringent admission criteria, they did not cite learning benefits as a reason; they largely 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 displayed by certain individuals did not necessarily extend to an awareness and the admission policy at the clinic level. However, the clinics that admitted difficult cases—largely, as they claim, for ethical reasons—received an unintended boost to their learning benefits, enabling higher success rates in the long run.

## Acknowledgments

The authors thank Phanish Puranam, senior editor Mary Benner, and two anonymous reviewers for their helpful comments on this study.

## 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 hence do not constitute a source of variation among firms.

[2]We checked for nonlinearity by estimating the effect of experience (not logged) and its square. It appeared, for the firms in our setting, that the effect of experience on success was immediately positive (cf. Haleblian and Finkelstein 1999) and did not turn negative at the highest levels (cf. Ingram and Baum 1997, Shepherd et al. 2003).

[3]The division by age in reporting the results is consistent across clinics. 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 up to 35 years old category (the standard group) and how many are 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 decreases their quality. Nevertheless, if a woman produces very few oocytes, it is generally considered a sign of a poor prognosis.

[6]The main negative effect of treating a higher proportion of older patients (>35 years) is somewhat surprising, because our dependent variable is related to the success rate among patients aged up to 35 years. Perhaps clinics that are more lenient about admitting older patients are also more lenient toward patients with lower ex ante success rates 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 approximately a year for the "low selection at the gate" to catch up with the "high selection at the gate" clinics.

[8]When we reran models 1–5 on the 606 observations for which we have data on specialists, the predictors for patients over 35 and patients who have failed before lost their statistical significance, but that for patients with few eggs remained significant for the interaction both with the main term ($-0.36$, $p < 0.05$) and with clinic experience ($0.06$, $p < 0.01$). When, in the model estimating the effect of patients with few eggs, we included the number of specialists as an additional control, the results were essentially identical ($-0.36$, $p < 0.05$ and $0.06$, $p < 0.01$, respectively). Likewise, adding the number of patients per specialist strengthened the conclusion that the potential alternative explanation centered around specialists is not driving our results. The loss of significance of the former two predictors is merely due to lost observations, not to the additional control.

[9]We also constructed a dummy variable for whether a clinic was affiliated with a university hospital, under the assumption that those that are will be more interested in developing better knowledge about IVF. Of the 37 university-affiliated clinics, 19 were private. Using our three indicators of difficult patients as dependent variables, random effects models showed that university-affiliated clinics treat more patients who have failed before ($0.281$, $p < 0.001$) and patients who produce few eggs ($0.065$, $p < 0.05$), but not more patients over 35. Because this variable appeared to be largely time invariant and therefore already controlled for in our fixed effects models, it was dropped as a main effect. An interaction with cumulative experience, added to the models in Table 4, remained insignificant, suggesting that they do not learn more quickly than others (over and above the effect of dealing with difficult cases). All estimates of our predictors were virtually unaffected by the inclusion of this variable, whether estimated through fixed or random effects models.

[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 more from the experience of others. Yet this appeared not to be the case; the interactions were all insignificant. This might be because general knowledge in the field also spreads effectively through other means, such as medical training, conferences, and the literature.

[11]For an exception, see Pisano et al. (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.

[12]It could be argued that our interactions could be interpreted as clinics that deal with a large proportion of difficult cases benefit more from experience. However, our results for women over 35 especially rule out this alternative explanation, because our dependent variable is success among women 35 and under, i.e., standard cases.

## References

Abdalla HI (2010) Interview with the clinical director of the Lister Fertility Clinic, London. Personal communication, September 28, London.

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

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

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

Argote L, Ingram P (2000) Knowledge transfer: A basis for competitive advantage in firms. *Organ. Behav. Human Decision Processes* 82(1):150–169.

Argote L, Ophir R (2002) Intraorganizational learning. Baum JAC, ed. *The Blackwell Companion to Organizations* (Blackwell, Malden, MA), 181–207.

Argote L, Beckman SL, Epple D (1990) The persistence and transfer of learning in industrial settings. *Management Sci.* 36(2):140–154.

Argyris C (1976) Single-loop and double-loop models in research on decision making. *Admin. Sci. Quart.* 21(3):363–376.

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

Bain & Company (2010) Our clients outperform the market 4 to 1: Helping make companies more valuable. Accessed December 6, 2011, http://www.bain.com.

Bannert M (2002) Managing cognitive load—Recent trends in cognitive load theory. *Learning Instruction* 12(1):139–146.

Barkema HG, Shenkar O, Vermeulen F, Bell JHJ (1997) Working abroad, working with others: How firms learn to operate international joint ventures. *Acad. Management J.* 40(2):426–442.

Baum JAC, Dahlin KB (2007) Aspiration performance and railroads' patterns of learning from train wrecks and crashes. *Organ. Sci.* 18(3):368–385.

Benner M, Tushman M (2002) Process management and technological innovation: A longitudinal study of the photography and paint industries. *Admin. Sci. Quart.* 47(4):676–706.

Benner M, Tushman M (2003) Exploitation, exploration, and process management: The productivity dilemma revisited. *Acad. Management Rev.* 28(2):238–256.

BBC (2007) IVF undercover: Transcript of interview with Mohamed Taranissi. Accessed June 9, 2012, http://news.bbc.co.uk/2/hi/programmes/panorama/6264997.stm.

Cohen WM, Levinthal D (1990) Absorptive capacity: A new perspective on learning and innovation. *Admin. Sci. Quart.* 35(1):128–152.

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

Dutton JM, Thomas A (1984) Treating progress functions as a managerial opportunity. *Acad. Management Rev.* 9(2):235–247.

Edmondson AC (1999) Psychological safety and learning behavior in work teams. *Admin. Sci. Quart.* 44(2):350–383.

Edmondson AC, Bohmer RM, Pisano G (2001) Disrupted routines: Effects of team learning on new technology adaptation. *Admin. Sci. Quart.* 46(4):685–716.

Edwards RG, Steptoe PC (1980) *A Matter of Life: The Story of a Medical Breakthrough* (Hutchinson, London).

Gibson C, Vermeulen F (2003) A healthy divide: Subgroups as a stimulus for team learning behavior. *Admin. Sci. Quart.* 48(2):202–239.

Guthrie JP, Datta DK (2008) Dumb and dumber: The impact of downsizing on firm performance as moderated by industry conditions. *Organ. Sci.* 19(1):108–123.

Haleblian J, Finkelstein S (1999) The influence of organizational acquisition experience on acquisition performance: A behavioral learning perspective. *Admin. Sci. Quart.* 44(1):29–31.

Hargadon A, Sutton RI (1997) Technology brokering and innovation in a product development firm. *Admin. Sci. Quart.* 42(4):716–749.

Haunschild P, Sullivan BNI (2002) Learning complexity: Effects of prior accidents and incidents on airlines' learning. *Admin. Sci. Quart.* 47(4):609–643.

Haveman HA (1993) Organizational size and change: Diversification in the savings and loan industry after deregulation. *Admin. Sci. Quart.* 38(1):20–50.

Hayward ML (2002) When do firms learn form their acquisition experience? Evidence from 1990–1995. *Strategic Management J.* 23(1):21–39.

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

Hoang HA, Rothaermel FT (2005) The effect of general and partner-specific alliance experience on joint R&D project performance. *Acad. Management J.* 48(2):332–345.

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

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

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

Kim J-Y, Miner AS (2000) Crash test without dummies: A longitudinal study of interorganizational learning from failure experience in the US commercial banking industry, 1984–1998. *Acad. Management Best Paper Proc., Toronto.*

Kolb DA (1984) *Experiential Learning* (Prentice-Hall, Englewood Cliffs, NJ).

*Lancet* (1999) Information versus choice in infertility treatment. 353(9168):1895.

Levitt B, March JG (1988) Organizational learning. *Annual Rev. Sociol.* 14(1):319–338.

Macher JT, Boerner CS (2006) Experience and scale and scope economies: Trade-offs and performance in development. *Strategic Management J.* 27(9):845–865.

Malterud K (2001) The art and science of clinical knowledge: Evidence beyond measures and numbers. *Lancet* 358(9279):397–400.

March JG (1991) Exploration and exploitation in organizational learning. *Organ. Sci.* 2(1):71–87.

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

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

Papke LE, Wooldridge JM (1996) Econometric methods for fractional response variables with an application to 401(K) plan participation rates. *J. Appl. Econometrics* 11(6):619–632.

Papke LE, Wooldridge JM (2008) Panel data methods for fractional response variables with an application to test pass rates. *J. Econometrics* 145(1–2):121–133.

Pisano GP, Bohmer RMJ, Edmondson AC (2001) Organizational differences in rates of learning: Evidence from the adoption of minimally invasive cardiac surgery. *Management Sci.* 47(6):752–768.

Prencipe A, Tell F (2001) Inter-project learning: Processes and outcomes of knowledge codification in project-based firms. *Res. Policy* 30(9):1373–1394.

Reagans R, Argote L, Brooks D (2005) Individual experience and experience working together: Predicting learning rates from knowing who knows what and knowing how to work together. *Management Sci.* 51(6):869–881.

Reitzig M, Wagner S (2010) The hidden costs of outsourcing: Evidence from patent data. *Strategic Management J.* 31(11):1183–1201.

Schilling MA, Vidal P, Ployhart RE, Marangoni A (2003) Learning by doing something else: Variation, relatedness, and the learning curve. *Management Sci.* 49(1):39–56.

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

Shepherd DA, Zacharakis A, Baron RA (2003) VCs' decision processes: Evidence suggesting more experience may not always be better. *J. Bus. Venturing* 18(3):381–401.

Sweller J (1989) Cognitive technology: Some procedures for facilitating learning and problem solving in mathematics and science. *J. Ed. Psych.* 81(4):457–466.

Szulanski G (1996) Exploring internal stickiness: Impediments to the transfer of best practice within the firm. *Strategic Management J.* 17(1):27–43.

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

Trevor CO, Nyberg AJ (2008) Keeping your headcount when all about you are losing theirs: Downsizing, voluntary turnover rates, and the moderating role of HR practices. *Acad. Management J.* 51(2):259–276.

Tyre MJ, von Hippel E (1997) The situated nature of adaptive learning in organizations. *Organ. Sci.* 8(1):71–83.

Vermeulen F, Barkema HG (2001) Learning through acquisitions. *Acad. Management J.* 44(3):457–476.

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

Weick KE, Sutcliffe KM, Obstfeld D (2005) Organizing and the process of sensemaking. *Organ. Sci.* 16(4):409–421.

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

Wiersma E (2007) Conditions that shape the learning curve: Factors that increase the ability and opportunity to learn. *Management Sci.* 53(12):1903–1915.

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

Zollo M, Winter SG (2002) Deliberate learning and the evolution of dynamic capabilities. *Organ. Sci.* 13(3):339–351.

**Mihaela Stan** is an assistant professor of management and organizations at the University College London. She received her Ph.D. in 2010 from the London Business School. Her main research interest focuses on organizational learning and innovation, with an emphasis on the healthcare sector.

**Freek Vermeulen** is an associate professor of strategy and entrepreneurship at the London Business School. He received a Ph.D. in strategic management in 1999 from Tilburg University and a Ph.D. in organization science in 2010 from Utrecht University, the Netherlands. He studies the limitations of organizations. His work has appeared in the *Academy of Management Journal*, *Administrative Science Quarterly*, the *Harvard Business Review*, and the *Strategic Management Journal*, among others.