Organization Science

Articles in Advance, pp. 1–18
ISSN 1047-7039 | EISSN 1526-5455

DOI 10.1287/orsc.1090.0457

Problems Deciding: How the Structure of Make-or-Buy Decisions Leads to Transaction Misalignment

Matthew Bidwell
The Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania 19104,
mbidwell@wharton.upenn.edu

This paper explores how the structure of decision making affects the way that firms manage their boundaries. Achieving transaction alignment requires firms to balance multiple goals. Drawing on the behavioral theory of the firm, I note that firms often assign different goals to different organizational units. As a consequence, simple problems about whether to make or buy can be affected by multiple decisions taken by multiple, locally rational units. I use a case study of the management of IT consultants in a financial services firm to explore how make-or-buy decisions are made. I find that senior managers at the firm focused on cost and organizational flexibility, whereas frontline managers concentrated on exploiting workers' existing knowledge. The narrow focus of these two groups interacted with the complex demands of transaction alignment to create three problems: separation of related decisions about internal capacity and project staffing, incomplete information when deciding on organizational capacity, and incentive misalignment in staffing consultants. These problems led the firm to become dependent on its consultants. I build on the case study to develop theoretical propositions about the characteristics of decisions and organizational structure that are most likely to lead boundary decisions to deviate from existing predictions.

Key words: transaction cost economics; knowledge-based view; firm boundaries; behavioral theory of the firm; organizational decision making; misalignment; contingent work
History: Published online in *Articles in Advance*.

A central question for firms is which activities to perform in-house and which activities to contract for externally. Existing theories generally propose that firms should manage such boundaries in ways that minimize transaction costs, specifically by achieving discriminating alignment between these "make-or-buy" decisions and transaction characteristics (Williamson 1988). Empirical studies confirm that firms often achieve the predicted alignment in their make-or-buy decisions (David and Han 2004), and that achieving such alignment improves transaction and organizational performance (e.g., Nickerson and Silverman 2003, Mayer and Nickerson 2005).

These accounts of transaction alignment usually assume that firms make a simple decision about whether to make or buy for each transaction. Argyres and Liebeskind (1999, 2002) have suggested that this focus on individual transactions may be too narrow; their work enriches transaction cost economics by showing how the governance of previous transactions can constrain the choices made about subsequent, related transactions. This paper seeks to extend our understanding of transaction alignment in a different direction by unpacking the structure of decisions that firms make as they manage their boundaries. I suggest that the alignment of a given transaction can sometimes be the result of multiple decisions made by different decision makers with different concerns. I examine the consequences of such complex decision structures for firms' ability to achieve transaction alignment.

Research on firm boundaries emphasizes that aligning a single transaction requires balancing multiple goals, such as allowing for adaptation as requirements change, creating effective incentives for participants, fostering the development of new knowledge, and making the best use of existing knowledge. Drawing on the behavioral theory of the firm, I suggest that this goal complexity can shape the structure of make-or-buy decisions. March and Simon (1958) noted that firms are structured into different units that are responsible for managing different goals. Such differentiation simplifies the complex task of running the organization; each unit focuses on solving a single component of the task, based on the goals it has been assigned. Yet this simplification comes at a cost; problems affecting multiple goals may be divided up among different units with different agendas. And the same local rationality that simplifies problem solving can prevent individual units from considering the full goals of the firm as they carry out their assigned tasks. Seemingly simple decisions can therefore be structured in complex ways, with important consequences for their eventual outcomes.

Using a case study of the use of information technology (IT) consultants in a large financial services institution (the Bank), I explore how such complex decision making can affect firms' ability to achieve

discriminating alignment. Achieving transaction alignment at the Bank involved both avoiding dependence on consultants and making effective use of consultants' and employees' existing knowledge. The multidimensional nature of this make-or-buy question interacted with the differentiated structure of the organization to lead to three related problems. First, related decisions were taken separately by different organizational units, specifically, by senior managers and frontline managers. This decision separation meant that no actor considered the full set of issues involved in achieving transaction alignment. It also created a potential for inconsistencies across the resulting decisions. Second, information asymmetries across the groups meant decisions were based on incomplete information. Third, incentive misalignments led decision makers to focus disproportionately on a subset of the organizations' goals. I demonstrate how these three problems led to misalignment in how consultants were staffed to work at the Bank.

I suggest that examining the structure of make-or-buy decisions contributes to our understanding of transaction alignment in a number of ways. Understanding how pressures for transaction alignment interact with organizational structure can potentially extend existing theories and provide more nuanced predictions about when firms will make versus buy. Perhaps most importantly, examining decision making can help us understand why misalignment appears common in organizations, despite the proven benefits of achieving discriminating alignment (Anderson 1988, Nickerson and Silverman 2003, Mayer and Nickerson 2005). Existing work has suggested that alignment may be inhibited by prior contractual commitments and bargaining power (Argyres and Liebeskind 1999, Nickerson and Silverman 2003), or has simply proposed that misalignment is a consequence of bounded rationality (Masten 1993; Argyres and Liebeskind 2002, p. 213), without explaining how this bounded rationality might manifest itself, or when it will be more or less severe. By providing some of the first research on how make-or-buy decisions are structured, I provide a more detailed explanation of why organizations may fail to achieve discriminating alignment, and make predictions about when such misalignment will be more versus less common. These explanations of misalignment provide new insights into how we can help firms better manage their boundaries. I also further our understanding of organizational decision making by articulating how firms' structures determine the kinds of problems that they find most difficult to solve.

Theory

Transaction cost economics argues that make or buy decisions should seek to minimize transactions costs, broadly defined as the costs of "planning, adapting, and monitoring task completion" (Williamson 1985, p. 2). A core tenet of transaction cost economics is that different instruments are available for governing activities within versus across firm boundaries. Outside the firm, formal contracts can be used to define how transacting parties will behave. Internally, asset ownership (Hart 1995) and legal doctrine (Masten 1988, Williamson 1991) allow firms to control activities through the exercise of authority. Such authority helps the firm to manage opportunism and adapt to changing circumstances, but may also carry costs of increased bureaucracy and reduced incentives (Williamson 1985, 1991).

Transaction cost economics focuses on transactions as its central unit of analysis, rather than individuals or decision premises (Williamson 1996, pp. 234–235). A central prediction is that "transactions, which differ in their attributes, are aligned with governance structures, which differ in their costs and competencies, in a discriminating (mainly, transaction-cost-economizing) way" (Williamson 1991, p. 277). Specifically, Williamson argues that transactions requiring large asset-specific investments should be internalized. Such specific investments increase the potential for opportunistic haggling between partners, haggling that can be minimized within the more flexible governance structure of the firm. Empirical research confirms that transaction alignment leads to increased performance, both at the organization level (Nickerson and Silverman 2003) and at the level of the individual transaction (Anderson 1988, Masten et al. 1991, Mayer and Nickerson 2005).

Whereas transaction cost economics focuses on firms' needs to manage the costs of governance, an alternative set of theories examines how firms' need to access and manage knowledge affects their make-or-buy decisions (Kogut and Zander 1992, Winter 1988). Managing firms requires complex knowledge that resides in organizational routines (Kogut and Zander 1992, Madhok 2002) and the heads of individual workers (Conner and Prahalad 1996, Grant 1996). These knowledge-based theories argue that this productive knowledge is not distributed evenly among firms. Causal ambiguity and path dependence in knowledge accumulation lead different firms to acquire different knowledge (Barney 1991, Hoetker 2005). As a result, firms may not always have the knowledge necessary to produce an input efficiently themselves.

Knowledge-based views of the firm have been used to modify predictions about discriminating alignment in two ways. First, scholars have explored how the problems of acquiring and protecting knowledge might affect the transaction costs of internal and external production (Macher 2006, Mayer 2006, Mayer and Salomon 2006, Nickerson and Zenger 2004, Schilling and Steensma 2002, Silverman 1999). Research in this tradition argues that transactions relating to the development or use of certain kinds of knowledge should remain inside the firm. Second, scholars have suggested that existing

stocks of knowledge may be an independent influence on firms' make-or-buy decisions: firms will choose to outsource activities if they lack the knowledge to do those activities themselves (e.g., Argyres 1996). This second set of accounts does not offer clear predictions about the kinds of transactions that should be internalized. Instead, they suggest that firms will make idiosyncratic choices based on differences between their knowledge and that of their suppliers.[1] Empirical studies confirm that these differences in firm-level capabilities also affect make-or-buy decisions (Jacobides and Hitt 2005, Hoetker 2005).

According to these theoretical perspectives, then, discriminating alignment requires economizing on both transaction costs and on the costs of developing and utilizing the idiosyncratic knowledge of internal resources and external suppliers for a given transaction. I draw on this definition of transaction alignment throughout this paper.

Firm Boundaries and Organizational Decision Making

Although research on firm boundaries identifies a complex array of factors that shape make-or-buy decisions, it provides a very simple view of how those decisions are made. The firm will choose the option that maximizes overall profitability. The behavioral theory of the firm, however, proposes a more sophisticated view of decision making. Research on the behavioral view of the firm draws on many of the same assumptions as transaction cost economics (Cyert and March 1963, Williamson 1985). Individuals are assumed to be boundedly rational and may have goals that differ from those of the organization. Where transaction cost economics explores how these assumptions affect the way transactions should be governed, the behavioral theory shows that these same assumptions have implications for how decisions are made within organizations.

A central principle of the behavioral view of the firm is that the task of coordinating activity within organizations is too complex to be solved by optimization. Instead, "a fundamental technique for simplifying the problem is to factor it into a number of nearly independent parts, so that each organizational unit handles one of these parts and can omit the others from its definition of the situation" (March and Simon 1958, p. 173). This strategy leads firms to adopt differentiated structures, where different units are responsible for addressing different kinds of problems. Most firms are differentiated horizontally into functional, geographic, customer, or product divisions. Firms also differentiate vertically into a "technical" level that is responsible for managing the organizational technology, and higher "managerial" levels of the firm (Thompson 1967, p. 10).

Factoring the complex task of managing the organization reduces complexity and allows the organization to devote considerable attention to multiple aspects of the task (Allison and Zelikow 1999, p. 167). These advantages come at a cost, however. First, all aspects of a given problem may not be considered simultaneously. Instead, different aspects of the problem may fall within the purview of different organizational units, leading a single problem to be addressed by multiple decisions taken by different units. Second, each of those organizational units will solve their parts of the problem in ways that are locally rational, rather than reflecting the goals of the overall organization. This local rationality stems in part from the different goals assigned to each unit (March and Simon 1958, p. 60; Lawrence and Lorsch 1967, pp. 36–37). Local rationality is also a result of differences in the information that organizational units encounter in their work (March and Simon 1958, p. 174), and the way that unit interests and organizational reporting channels shape attention (March and Simon 1958, p. 174; Ocasio 1997). This structural differentiation causes particularly difficulties when firms must balance the goals of multiple units. By its very nature, factored decision making prevents firms from seeking the most profitable trade-off among the goals of different organizational units. Instead, Cyert and March (1963, p. 164) argue that firms will seek simply to achieve "quasi-resolution of conflict" between goals, where "the procedures for 'resolving' such conflict do not reduce all goals to a common dimension or even make them obviously internally consistent."

These perspectives on organizational decision making have been used to examine how firms should best organize to process information and solve problems (e.g., Galbraith 1977, Tushman and Nadler 1978). They have also been applied to understand how the structure of the organization might shape major strategic decisions (e.g., Allison and Zelikow 1999, Fredrickson 1986). Less research has addressed how local rationality might affect the way that organizations deal with ongoing, day-to-day problems. Prior research has carefully characterized how make-or-buy decisions *should* be made; examining how these decisions are made in practice has the potential to add to our understanding of when firms struggle to solve problems effectively. In this paper, I explore how the structures of these decisions—and of the organizations taking them—might lead to transaction misalignment.

Research Setting and Methods

I investigated how decision processes shape transaction alignment using a detailed case study of how the IT department of a large financial services institution (the Bank) managed its independent consultants. Decision making within organizations involves multiple actors and is largely informal. It is therefore difficult to study with cross-firm survey research or archival work. An in-depth description of a single case may therefore

be more valuable than a less detailed comparison of multiple cases (Dyer and Wilkins 1991, Allison and Zelikow 1999).

The Bank was a highly respected, highly profitable company, and its IT department was seen in the industry as very professional and a premium employer. The work of the IT department involved maintaining the computer systems and fixing problems, as well as carrying out new development to support changing business practices and regulatory requirements. The department employed around 10,000 individuals and was divided into highly autonomous divisions that supported each of the Bank's different business units.

Independent Consultants at the Bank

The Bank made extensive use of independent consultants in managing its IT systems. These consultants comprised up to one-third of the IT workforce and were effectively employed as independent contractors.[2] The consultants did not possess substantially different skills from the regular employees, and were employed to carry out general programming rather than to provide specialist expertise: I surveyed 39 projects that used consultants; of these, only two were using consultants because of the specialized expertise that they were bringing. Nor were consultants brought in as a way of screening new hires (Houseman 2001). Although consultants did sometimes convert to regular employees, only once did a manager propose screening to me as a reason to hire consultants. Furthermore, the average tenure of consultants at the Bank was three years—far longer than any reasonable probation period. The Bank instead used independent consultants to increase employment flexibility. As we will see, firing employees was costly for the Bank. Using consultants allowed the Bank to adjust the size of its IT workforce in response to changes in demand. Such flexibility was important at the Bank, where IT spending was very volatile. From 2000 to 2001, for example, the IT budget of one of the largest divisions fell by 25%.

Although consultants were more flexible than employees, they were also more expensive to employ. The Bank calculated that the total cost of the average consultant was 25%–30% more than the equivalent employee, including benefit costs. Some of this differential came from the commissions paid to staffing agencies. The rest compensated consultants for taking on less secure work. This increased expense was the clearest cost to the Bank of using consultants, and was seen as a premium for their increased flexibility.

The Bank's relationships with its consultants were governed in very different ways to relationships with employees, corresponding closely to the dimensions that Williamson (1991, p. 281) used to describe how governance changes as transactions move from market to hierarchy. Williamson argued that hierarchies are characterized by greater use of administrative controls, reduced use of incentives, adaptation by cooperation rather than unilateral action, and more flexible contract law. Each of these dimensions applied to the differences between consultants and employees.

Relationships with employees involved a very high level of administrative controls. Like many large, white-collar employers, the Bank adopted employment practices that structured workers' promotion opportunities and increased their job security (Osterman 1987, Williamson 1985). Jobs were organized into different levels, which each had a defined pay band. Promotion through these levels was the main way for employees to increase their pay and benefits. Workers also had a great deal of job security. The Bank had strong norms against downsizing employees, and had historically performed very few layoffs. Perhaps most importantly, layoffs were very costly. When the Bank laid off an employee, it would pay compensation and benefits that were 60%–70% of the employee's annual cost. In contrast, relationships with consultants had very few controls. The consultants could be let go at any time for any reason. They were not part of any career ladder, nor was their pay determined by the firm's pay scales, although some efforts were made to tie their rates to a rate card.

The Bank's personnel policies also created stronger incentives for consultants than employees. It was very difficult to terminate nonperforming employees for cause. Managers had first to go through an elaborate process of putting the employee on probation, setting targets, and monitoring performance. When consultants failed to perform they were dealt with very rapidly. As a result, consultants had much stronger incentives to satisfy project managers.

The different controls applied to consultants and employees also prompted very different forms of adaptation. The Bank's personnel policies strongly encouraged the two parties to engage in adaptation by cooperation. It was costly for the Bank to dismiss existing employees and hire new ones as its skills needs evolved. Instead, the costs of layoffs encouraged the Bank to redeploy and retrain its employees. Although these structured employment practices may have made it more difficult for the Bank to dismiss underperforming workers, they helped to protect workers' investments in firm-specific skills. Structured promotion ladders and pay scales also discouraged employees from engaging in opportunistic bargaining by limiting their ability to obtain unique pay raises and tying their fate more closely to the firm's long-term interests (Williamson 1985, pp. 243–249). In contrast, market-based modes of adaptation were much more straightforward with consultants. Existing consultants could be easily dismissed, and new consultants brought in from the market. Differences in the legal status of employees and contractors reinforced these different modes of adaptation. The Bank's

formal employment relationship with its employees gave it far more authority to effect cooperative adaptation than did its commercial relationships with consultants (Masten 1988).

Although relationships between the Bank and its consultants were closer than arm's-length market relationships, all of these differences between employees and consultants made the Bank an effective setting for testing the predictions of transaction cost economics (see also Mayer and Nickerson 2005; Masters and Miles 2002; Williamson 1975, pp. 57–81).

Data Collection

I used a variety of methods to collect data on the use of consultants at the Bank. My data collection focused on a single business unit, which I refer to as Consumer. Consumer was one of the largest business units and had an IT workforce of roughly 2,000 at the time of the study. Focusing on Consumer allowed me to gain a deep understanding of the work, decisions, and outcomes in a single unit. I then conducted extensive interviews in other groups to compare their experiences.

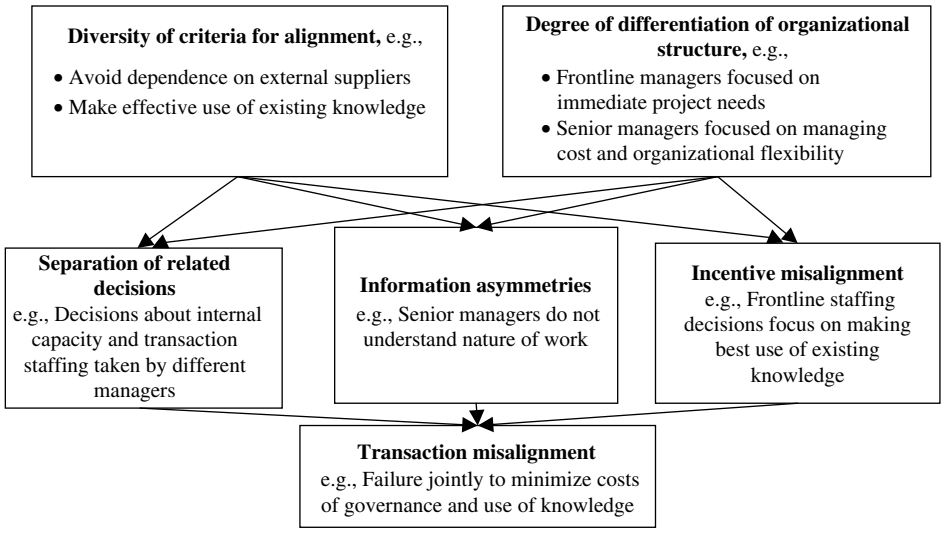
I collected data on how decisions about the use of consultants were made through semistructured interviews with 49 informants throughout the organization. My informants included 9 senior managers, 4 human resource managers, 7 sourcing managers, 22 frontline managers, and 7 developers (of whom 3 were consultants). I also collected systematic data on the work process and the use of consultants using a survey of 44 managers of projects in the Consumer division. This survey provided complete data on how 120 employees and 90 consultants had been used on those projects. Further details about this survey are presented in Bidwell (2009). I also asked the project managers a number of questions about their attitudes to consultants. Finally, I collected accounting data from the Consumer division that showed how the numbers of consultants and employees at the Bank changed over time. More information on how these data were collected and analyzed is presented in the appendix.

Decision Making and Transaction Misalignment

Figure 1 shows how the structure of decision making affected the way that the Bank used consultants, leading ultimately to transaction misalignment. As noted in the theory section, achieving transaction alignment requires simultaneous attention to multiple goals, including managing transaction costs and making productive use of knowledge. Yet these goals were primarily addressed in different decisions made by different parts of the organization. The interaction between the multidimensional nature of the make-or-buy problem and the differentiated structure of the organization led to three specific problems. First, decisions that affected transaction alignment were assigned separately to senior managers and frontline managers, reflecting the disparate goals that the use of consultants addressed. This decision separation meant that no actor considered the full set of issues involved in achieving transaction alignment; it also lead to inconsistencies across related decisions. Second, structural specialization meant that senior managers did not have access to information about how knowledge was used in the organization, information that was necessary for aligned decisions. Third, specialized incentives led frontline managers to focus disproportionately on the effective utilization of workers' knowledge, at the expense of managing transaction costs.

I start by defining transaction alignment in how the Bank used consultants, and go on to provide evidence that the Bank failed to achieve alignment. I then show

Figure 1 A Model of Decision Making and Transaction Misalignment



how the structure of decision making at the Bank led to this misalignment.

Defining Transaction Alignment in the Use of Consultants

Williamson (1985, p. 1) defines a transaction as the transfer of a good or service across a technologically separable interface. Within the Bank's IT department, an individual transaction would therefore be defined as a task that a worker carries out,[3] such as making changes to a module (a subunit of an application) as part of a project. Such a definition allows for workers to carry out multiple different transactions over time, as was usually the case at the Bank.

According to transaction cost economics, firms achieve discriminating alignment by internalizing those transactions that require greater investments in specific assets. The key source of such asset specificity at the Bank was learning about its proprietary applications. The vast majority of development work at the Bank consisted of incremental changes to existing applications, rather than developing entirely new software. These applications were generally not commercial applications, such as SAP or PeopleSoft. Instead, they were proprietary applications that had been written and rewritten over many years by the Bank's developers. These applications were only used at the Bank, often by very few people. They were highly complex and poorly documented. To be productive, workers needed to make a large, specific investment in learning how a given application worked.

The nature of application skills at the Bank created a risk that the Bank would become dependent on consultants' knowledge. Because the Bank used many different applications, only a few developers usually had experience in any given module. Because these applications were overwhelmingly proprietary, this experience could only be gained on the job at the Bank. Prior experience with the relevant application modules was therefore a central determinant of individuals' productivity on a project. One manager described for me how he staffed a project:

> All of the people on this project have prior [specific application] experience. They were what we most needed. With the experience of the system you can get away with one less resource for the project.

Evidence of the importance of this narrow, specific knowledge came from how managers staffed projects. I asked managers to rank the factors that had determined their decision about who to staff to a specific project. Of the 41 managers who responded to the survey, 31 said that workers' experience with the particular application under development was the first or second most important factor in their decision. Of all of the individuals working for a given manager, only a few would know enough about a particular application to be productive working on it. These would be the workers who would be staffed to the project.

Achieving transaction alignment therefore required consultants to be staffed to roles where they would not use or acquire unique valuable knowledge. Otherwise, mutual dependence would develop between the Bank and consultants, creating a potential for haggling over these rents and increasing transaction costs (Williamson 1985, pp. 175–179). Consultants' acquisition of unique valuable knowledge would also conflict with the Bank's strategy of using consultants to achieve flexibility. If the Bank became dependent on its consultants, it would be difficult to lay off those consultants in a downturn.

However, limiting dependence on consultants was not the only factor that managers needed to weigh up in achieving discriminating alignment. As highlighted by knowledge-based views of the firm, the Bank also needed to staff individuals in ways that made efficient use of the knowledge that they possessed. That said, there was no clear reason why managers could not have made productive use of workers' knowledge without creating dependence on consultants. On entering the Bank, both consultants and employees had very similar skills. It was only subsequently that workers acquired heterogeneous expertise, through the way that they were staffed to applications. If managers avoided staffing consultants to work that involved them acquiring unique valuable skills, it should have been possible to deploy the right knowledge sets without the Bank becoming dependent on its consultants.

We therefore define discriminating alignment at the Bank as having employees and consultants assigned to tasks in ways that minimize dependence on consultants while making effective use of existing knowledge. This alignment required segregating consultants from tasks where they would acquire unique and valuable knowledge that was not shared by other workers.

Evidence of Misalignment

I did not find that the Bank avoided staffing consultants to roles where they would acquire unique valuable knowledge, as required for discriminating alignment. Consultants and employees were not segregated in the projects that they worked on. Only five of the 44 projects I surveyed did not use consultants—and those projects did not appear systematically different from the others. Nor were there many differences in which projects and roles used more consultants. Table 1 provides basic descriptive statistics on the work done by consultants and employees, based on the project survey. I found that consultants were somewhat less likely than employees to be team leads and were also less likely to be analysts, who required greater knowledge of the business. Consultants were *more* likely to be staffed on projects that required more time spent on business issues rather

Table 1 Basic Summary Statistics for Project and Role Characteristics by Employment Status

	Employees		Consultants	
Project and role characteristics	Mean	Std. dev.	Mean	Std. dev.
Time spent modifying or enhancing already existing proprietary applications rather than developing or implementing totally new systems (%)	81.2	26.6	82.7	28.8
Time spent managing interfaces with other Bank systems (%)	13.9	17.1	16.3	16.5
Time spent resolving business issues rather than technical issues (%)	21.4	16.5	28.1*	18.2
How important do you think the outcome of this project was to senior management? (1–7)	5.57	1.39	5.22	1.68
Is the system business critical—that is, would temporary problems with the system be likely to lead to a significant loss of revenue for the Bank? (1–7)	4.63	1.92	4.58	2.02
How important was it to be technically innovative to implement an appropriate solution? (1–7)	5.05	1.49	4.62*	1.58
How high was the project's need for technical expertise? (1–7)	5.38	1.09	5.28	1.03
Project duration (months)	9.0	3.9	8.5	3.8
Team lead	0.18	0.39	0.06**	0.23
Analyst	0.11	0.31	0.03*	0.18
N	120		90	

Notes. Items measured as percentages were responses to the question, "What proportion of the project, measured as a percentage of the total hours spent, involved:…?" Items measured on a 1–7 scale were responses to the following: "I would like you to compare this project with other new development projects that go on at the Bank. I will ask a series of questions about this project, and would like you to rate it against each of these criteria, on a scale of 1 to 7, where 1 is not at all, 4 would be about an average amount for a project, and 7 is a great deal."

*Means are different at the 5% level; **means are different at the 1% level. All data are taken from the project survey answered by frontline managers. Data come from 36 different projects. All items were measured at project level except *team lead* and *analyst* variables.

than technical issues. Overall, though, the data demonstrates very little attempt to segregate consultants and employees. Time spent modifying existing applications is an indicator of highly specific skills; managing interfaces with other systems also required such specific skills. The table demonstrates that there was substantial variance in projects' need for such specific skills, yet consultants were no less likely than employees to be working on projects with a high need for specific skills.

A variety of qualitative evidence confirmed that managers staffed consultants and employees very similarly. I asked managers a series of open-ended questions about how they used consultants. Only 3 of the 44 managers surveyed said that they assigned consultants to different kinds of work to minimize dependence on them. Thirty-one of the managers said there was no difference in the way they managed consultants and employees. The survey therefore demonstrates that managers generally staffed consultants similarly to employees, and did not minimize their dependence on consultants by segregating them from work that required narrow, application-specific skills.

I also found abundant evidence that this failure to differentiate work done by employees and consultants led managers to become dependent on consultants' application knowledge. In fact, 12 of the 44 managers I surveyed gave unprompted accounts of their dependence on consultants. One manager told me:

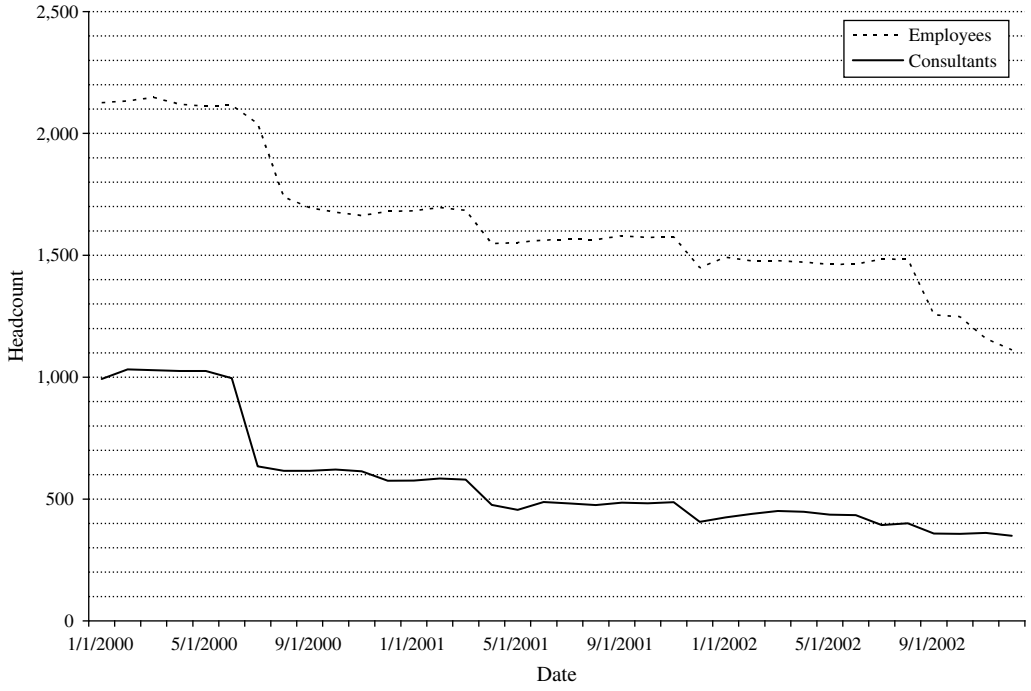
> I'm fortunate because my consultants are experts in this system. So the ones that I have left—I rely on them heavily for the internal knowledge of this product. It was a package that we've modified significantly. The consultants bring significant knowledge of the package and the business. I always shied away from relying on them but once we started downsizing they became more critical.

The economic downturn of 2000–2002 demonstrated the extent of this dependence. As we have seen, the Bank hired consultants because of their flexibility, and paid a significant premium to do so. In the event of a downturn, the use of consultants should provide a low-cost means of reducing the workforce. Yet when precisely such a downturn occurred, the Bank found it surprisingly difficult to lay the consultants off. Because of how they had been used, many consultants had acquired valuable application knowledge that was not shared by employees. As a result, some consultants were too important to let go. As one project manager told me:

> When you have a downturn the consultants are the sacrificial lambs at first—these are the lower level consultants. Then later you get to a stage where your consultants are very important to you. They're your best coder or they're key to making the business run. So although management is saying fire the consultants and protect the employees at all costs, you end up saying can we fire these two lower level employees with a package, in order to save this higher level consultant?

The Bank's inability to dismiss many of its consultants illustrates how dependent the Bank had become on the consultants. Figure 2 graphs the headcount changes for consultants and employees in Consumer's IT group from 2000 to 2002. Although many consultants were dismissed during this period, so were many employees. Despite the fact that consultants could be terminated

Figure 2 Change in Employee and Consultant Headcount Over Time in Consumer Division



without any compensation, the Bank retained many consultants while paying to lay off over 1,000 employees. The graph also overstates consultants' flexibility: some 97 consultants were converted to employees during this period, contributing to the difference between the numbers of consultants and employees.

The decision to retain some consultants while firing employees may have been ex post efficient. The Bank was able to retain consultants whose knowledge of the Bank's systems made them more productive than potential replacements. However, the decision also demonstrates ex ante transaction misalignment. The basic prediction of transaction cost economics is that firms should use internal governance when transactions involve high levels of specificity. Yet the Bank had frequently staffed external workers to work requiring highly specific knowledge. And the Bank incurred substantial costs from this misalignment. By becoming dependent on external workers, the Bank exposed itself to significant costs of opportunism and inefficient adaptation (Williamson 1975, 1985). For example, several managers complained that consultants were less committed to the organization and less hard working than employees. Becoming dependent on the consultants made such problems more difficult to deal with because the main incentive for consultants—the threat of dismissal—became harder to exercise in practice. Perhaps most concretely, the Bank had to engage in costly layoffs of many more employees than would have been necessary if it had not become dependent on its consultants. Had the Bank achieved transaction alignment, it could simply have laid off all of its consultants instead.

And achieving transaction alignment should have been possible. After all, the Bank's willingness to let go of so many employees demonstrates substantial variation in workers' acquisition of unique, valuable knowledge. The problem was that consultants were often staffed in ways that the Bank became *more* dependent on them than it was on some employees. Avoiding such dependence would have required frontline managers to exercise greater foresight in how they staffed consultants, but should have involved minimal disruption to the production process. Alternatively, the Bank could have reduced the additional costs of hiring expensive consultants if it had hired more individuals as employees ex ante. By hiring many consultants, but then using them in ways that made it dependent on them, the Bank achieved a costly transaction misalignment.

It is possible that some of the frontline managers' reluctance to lay off consultants stemmed not from transaction misalignment, but from general skills differences between employees and consultants. I did find that some managers believed that consultants had higher technical skills than employees. Overall, however, consultants and employees were remarkably similar. Many of the employees I interviewed had previously been consultants either at the Bank or elsewhere. Half of the consultants that I talked to were employed as consultants because of headcount freezes when they were hired, rather than because they had chosen to operate in a different segment of the labor market. Perhaps most importantly, when managers spoke to me about the difficulties involved in laying off consultants, almost all of their discussions centered on consultants' specific knowledge of

their systems. Over the years, these consultants had built up detailed expertise on the specific proprietary systems they managed. When those consultants left, other workers would struggle to work with those applications. One manager explained to me:

> Usually when I look at the project I try to use people who have the technical and business knowledge. I had to bring in new people when I lost the consultant . . . . Clearly I would have preferred it if my consultant didn't leave—I needed three people to replace one.

Perhaps the most persuasive evidence of misalignment came from senior managers' reactions to consultants' lack of flexibility. One of the most senior managers in the Consumer division commented at the end of my study, "We've been riffing [laying off] our employees, and we still have 350 consultants." When asked what they were getting for the 35% premium for flexibility, the manager made a zero with thumb and forefinger.

Senior managers did more than simply express their disappointment. Once the downturn revealed how dependent the Bank had become on many of its consultants, senior managers took corrective action to reduce misalignment. Controls on hiring consultants were increased. At the same time, many controls on hiring employees were actually eased so that consultants could be converted and positions could be filled with cheaper employees rather than consultants. Many groups encouraged managers to convert valuable consultants to employees. These conversions reduced the costs of continuing to employ these critical workers, and brought their governance back into alignment with the nature of the work they were doing. Yet the conversions did not render misalignment costless. The Bank had still incurred substantial costs from engaging these workers as consultants for a number of years when they could have been more cheaply hired as employees. Indeed, the need to make these conversions underscores the presence of misalignment by emphasizing how some supposedly flexible consultants had become critical to the Bank.

Misalignment and the Structure of Decision Making

We can understand how this transaction misalignment occurred by exploring the structure of decision making at the Bank. The Bank sought to achieve alignment in a single underlying set of transactions: how individual consultants and employees should be staffed to work. Yet this alignment affected multiple goals within the organization. Among the most important of these goals were balancing cost and flexibility through managing the numbers of employees and consultants, and maximizing productivity through how projects were staffed. Within the Bank, those goals were managed by different groups: senior managers on the one hand and frontline managers on the other. The way that the alignment decision was divided among these two groups, combined with the focus of each group on its own narrow goals, led directly to the transaction misalignment we observed.

Separation of Related Decisions. Rather than reflecting a single decision about whether a given project should be staffed internally or externally, transaction alignment was determined by two sets of related decisions taken by different groups of managers: how many consultants versus employees to hire, and how then to use those consultants and employees on specific projects. This separation of decisions reflected the way that the use of consultants intersected with the organizational allocation of responsibilities. But it greatly complicated the task of achieving discriminating alignment.

Decisions about how many consultants and employees to hire were largely taken by senior management. Each year, senior managers within the IT department would negotiate with frontline managers to set the numbers of regular employees in each of the different departments. These headcount numbers were based on existing employment, adjusted to reflect changes in the amount of work that the groups would perform. The headcount numbers set limits on how many employees could be hired, requiring managers to go through several layers of senior management approval if they wanted to hire additional employees during the year. At times, hiring freezes would make the hiring of new employees all but impossible.

The number of consultants hired was then determined by the difference between the allowed employee headcount, the amount of work to be done, and the budget available. Before the downturn revealed the difficulties of terminating consultants, senior managers did not feel a need to control the overall number of consultants as tightly as they did employees. After all, consultants were supposed to be flexible. Frontline managers therefore had to go through far fewer layers of approval to hire consultants. A frontline manager explained that:

> It is not a free market within the firm. Instead, there are a lot of HR policies that say—well, for the last few years every January there would be a hiring freeze for employees, but not for consultants, so we would hire in consultants.

Project managers repeatedly pointed to these headcount constraints in explaining why they used consultants. Similarly, four of the eight consultants I spoke to explained that they were taken on as consultants because of headcount constraints at the time they joined.

Although senior management set the headcount constraints that determined the balance between consultants and employees, it was frontline managers who decided how consultants were staffed. Subject to headcount constraints, frontline managers would take the final decision

to hire a consultant for a given project. Frontline managers also had exclusive responsibility for deciding how the consultants were staffed to subsequent projects: once hired, consultants would usually move from project to project, remaining at the Bank for an average of three years. Neither senior management nor any of the staff functions monitored the specific task that the consultants were hired into initially, or the work that they subsequently performed.

This separation of related decisions across managerial levels was a rational response to the complex demands of managing governance and knowledge. The firm needed to make long-term decisions about the balance between fixed employee capacity and the variable supply of consultants; given the expense of firing employees, the firm could not decide whether to hire employees or consultants for each individual project, particularly when projects often lasted just a few months. Furthermore, these capacity decisions needed to be resolved at the level of the overall organization where the firm's portfolio of activities could be considered. On the other hand, individual staffing decisions needed to be taken project by project. Effectively assigning individual transactions to internal versus external supply required an in-depth understanding of the specific details of the transaction, understanding that only frontline managers possessed. The nature of the decision problem therefore led decisions to be factored across the two levels of the organization.

Yet this same separation of related decisions across managerial levels fostered transaction misalignment in two ways. First, separation of decisions meant that no decision maker ever considered the full set of issues involved in making a transaction alignment decision. Senior managers could make capacity decisions without regard to which particular projects a given consultant or employee would work on, and frontline managers' staffing decisions were heavily constrained by existing limits on employee headcount. Because no decision maker ever took the full, unconstrained decision about how consultants would be used, no decision maker had the opportunity to properly manage the trade offs involved in achieving transaction alignment.

Second, separation of related decisions increased the risks of inconsistencies across the decisions of frontline and senior managers. The evidence presented above suggests that transaction misalignment stemmed from problems in how both senior managers and frontline managers took decisions. The headcount limits set by senior managers encouraged the hiring of large numbers of consultants; yet the critical importance of narrow, specific application knowledge meant that transaction alignment could only be achieved with limited numbers of consultants. Frontline managers then paid scant attention to issues of dependence in how they staffed consultants to work. Each of these factors contributed to transaction misalignment. Yet that misalignment was exacerbated by the inconsistency across those decisions. Senior managers' decisions to hire too many consultants would have been less damaging if frontline managers had done more to reduce dependence on the consultants hired. And frontline managers' inattention to dependence issues would have been less problematic had senior managers promoted the hiring of fewer consultants. Instead, by hiring a large number of consultants and using them indiscriminately, senior management and frontline managers magnified the problems created by each others' decisions.

Information Asymmetries and Allocation of Attention. In addition to the lack of coordination across separated decisions, misalignment also stemmed from local rationality in how both senior and frontline managers made their decisions. Among senior managers, this local rationality was a consequence of information asymmetries. These information asymmetries led senior managers to focus on the organization's need to manage cost and flexibility, and to largely ignore how the nature of the work affected the use of consultants.

Capacity decisions at the Bank required the consideration of diverse information about high-level questions of cost and flexibility, and low-level issues of how development takes place. Those individuals who were closely involved in shaping headcount numbers were either in senior management positions or in staff positions, far removed from the actual work. In addition, many of the senior and staff managers had no direct experience in managing technical projects at the Bank. Instead, they had backgrounds in consulting or in finance, legal, and human resource functions that provided little insight into the nature of development work within the Bank. The Bank's control processes also focused senior managers' attention on cost and governance issues. At the more senior and staff levels of the Bank, management processes focused on budgeting and resource constraints, favoring a focus on governance issues rather than the details of production and knowledge management. Few aspects of their day-to-day work forced these managers to focus on the nature of technical knowledge. Hence, although some senior managers did begin to understand how the work shaped the use of consultants, many of them failed to understand how they could become dependent on those consultants.

Senior and staff managers' descriptions of how consultants were staffed provided a clear example of their focus on governance rather than work. All of the work at the Bank was divided into projects so that expenses could be allocated to different tasks and business groups. To senior and staff managers, the language of projects implied that the work itself could be split into discrete pieces, with independent staffing choices made about each project. When talking about the use of consultants,

human resource managers would describe how consultants were brought in for specific short-term projects. For the project managers, though, these projects were often an accounting device rather than a way of defining truly discrete pieces of work. Consultants were really carrying out a stream of work on an application, drawing on a cumulative body of knowledge. Once hired, they moved from project to project, all the while carrying out similar work on the applications.

These misunderstandings were particularly pronounced among the human resource managers, who played a central role in facilitating headcount negotiations. Because of their role in shaping the employment system, human resource managers' view of consultants was shaped by the formal terms of their employment; consultants were not eligible for severance pay and were therefore flexible. One senior human resources manager explained that:

> Philosophically the difference between consultants and employees is the difference between fixed expense and variable expense. Employees get benefits, and when they are let go there is severance payments and the possibility of litigation expenses. You don't have these things with consultants. The assumption is that they are flexible.

As the study progressed and I probed deeper into how consultants were being used at the Bank, the human resource managers began to express bewilderment about the use of consultants. One manager responded to my preliminary findings with:

> Where I first worked we learned that a temp is a temp is a temp. You pay 40% to the agency, and then they stay with you for a short time and you get rid of them. Then, when I moved companies, there were temps who had been there three, four, five years. And it's the same here. And I still don't get it.

Without direct experience of the work process, it was hard for these managers to understand how consultants were being used in practice. Yet understanding the nature of the work that consultants and employees were doing should have been critical to establishing the appropriate numbers of employees and consultants. Without this understanding, senior and staff managers facilitated the hiring of far more consultants than was appropriate for discriminating alignment.

Incentive Misalignments. Although senior managers' failure to understand how consultants were used contributed to transaction misalignment, it was not the only cause. Frontline managers also paid too little attention to managing dependence problems when staffing consultants. Instead, frontline managers focused on the knowledge-based issues of making the best immediate use of workers' existing application experience. As described above, frontline managers rated existing application experience as by far the most important determinant of staffing decisions in my survey. Many managers also cited technical skills as important. By contrast, managers rated knowledge retention as the least important of the factors I listed: no manager said this was the most important factor in their decision; only one listed it as the second most important factor. Even if senior managers had introduced far fewer consultants at the Bank, the way that frontline managers skewed their decisions towards the short-term demands of using knowledge would have led to misalignment.

Frontline managers' focus on the demands of knowledge over governance was a product of the incentives that these managers faced. Frontline managers experienced strong pressure to keep their systems running and were held closely to account for the quality and timeliness of the work they carried out. By contrast, they were much less accountable for other consequences of their actions. Many of the problems of transaction misalignment manifested in higher staffing costs. As managers became dependent on consultants, they exposed the firm to being held up for higher pay rates. Their dependence also reduced the possibility of achieving market adaptation by dismissing the consultants during a downturn. Because consultants were more expensive to employ than regular employees, this dependence meant that the Bank was paying a premium for employment flexibility that it would not be able to exercise. But project managers were not responsible for these costs. Although some talked about wanting to minimize the costs, they had few incentives to do so. While they remained within the overall budget allocated to them, project managers were not assessed on the cost effectiveness of their development. In some parts of the organization, project managers did not even know what their project budget was. As two of the most senior managers commented on reviewing my findings:

> [First manager] Our managers still believe that they are being graded on delivery. Not for managing on budget and process.
> [Second manager] And they are, at that level.

These weak cost controls stemmed in part from the same information asymmetries that led the Bank to hire too many consultants. The optimal cost of a given project depended on the right mix of resources for the work. It was difficult, however, for outsiders to assess what the right mix should be, because this required a detailed understanding of the nature of the work. Without being able to set reasonable cost targets for a given group, staff managers found it difficult to improve project managers' cost incentives.

The consequence of these imbalanced incentives was that short-term pressures of getting the work done dominated longer-term concerns about managing dependence in staffing decisions. When discussing staffing decisions, managers repeatedly talked about what would allow them to get the work done most quickly. One of the

managers explicitly explained how short-term pressures compromised her efforts to spread knowledge within the group:

> [My staffing] was based on availability and also knowledge of the person in this area. I also try to do cross training so different people try different things. In reality, in order to go through rotation, it needs more time for the project. So this does not work 100%.

Project managers' focus on short-term speed over long-term cost and dependency problems is not indicative of a general lack of concern about cost in the organization. During the course of my study, pressures to reduce costs led the IT department to impose deep and painful headcount cuts, almost halving the size of the organization. The quest to reduce development costs also led managers to outsource some of their most critical applications to vendors based overseas. The demands of effectively deploying knowledge did not therefore dominate cost pressures for the overall organization, only for project managers.

Alternative Interpretations of the Case

The evidence presented here suggests that the Bank experienced transaction misalignment in its use of consultants because of the structure of its decision making. To establish the validity of this interpretation, though, it is important to consider alternative explanations.

One possibility is that misalignment was an inevitable consequence of the duration of the Bank's relationship with consultants. Where there is strong governance inseparability between transactions, decisions at the beginning of the relationship can affect transaction governance years later. Misalignment can then be caused by unanticipated shocks, which leave transaction characteristics very different from those first envisaged. Such exogenous shocks were absent from the Bank's employment system, though. The Bank did face a downturn in 2001–2002. Yet, although the timing of this downturn was unanticipated, the Bank had been well aware of the need to maintain flexibility in its headcount. It was precisely this need for flexibility that had led the Bank to hire consultants in the first place. The downturn did not, therefore, challenge the assumptions that had underpinned the Bank's use of consultants.

Could misalignment have been a consequence of particularly incompetent managers? There is no evidence that the Bank's managers were of low ability; in fact, the Bank was a prestigious employer with well educated, experienced, and hard-working staff. The managers appeared no less thoughtful than the average. Instead, the evidence supports Cyert and March's (1963) argument that behavior in any complex organization is necessarily boundedly rational.

It is also possible that the Bank never intended to achieve alignment in its use of consultants, and was not concerned about dependence on them. For example, some reports suggest that firms hire consultants to massage headcount numbers reported out to analysts (Barley and Kunda 2004). There might then be pressure to fire employees rather than consultants, in order to push headcount numbers down further. Yet senior managers' decision to increase the number of employees at the expense of consultants, through conversions and easing hiring restrictions, is not consistent with this interpretation of events.

A further alternative explanation for why frontline managers were reluctant to lay off consultants could be that they preferred the higher incentives that they could give consultants, and the more direct control they had over them. However, although some managers said that they liked consultants' raised incentives, many more expressed concern that consultants had less organizational commitment than employees. As we have seen, the weight of evidence suggests that managers were reluctant to dismiss consultants because they became dependent on their application knowledge, rather than because they preferred working with them.

Discussion

The case study presented above demonstrates how concepts from the behavioral theory of the firm can inform transaction cost and related knowledge-based theories of firm boundaries, and provide a better understanding of why firms often fail to achieve transaction alignment. The Bank's use of consultants deviated from the predictions of transaction cost economics and knowledge-based theories of the firm, not because those theories made the wrong predictions about what the Bank should do, but because the structure of decision making prevented the Bank from achieving the prescribed alignment. By providing a theoretical lens for understanding decision making, the behavioral theory of the firm allows us to build a deeper understanding of how firm boundaries will be set in practice.

In this section, I draw on the behavioral theory of the firm to develop the insights from the case study into more general propositions about when transaction misalignment is likely to occur. These propositions can help us to understand when firms will face problems in managing their boundaries, and when organizations are less likely to achieve the transaction alignment prescribed by our theories.

This paper's central argument is that the firm's ability to achieve alignment is shaped by the interaction between the structure of the organization and the nature of the decision problem. The behavioral theory of the firm notes that organizations are often divided into specialized units that have access to different information and are charged with different goals. At the Bank, different units focused on the problems of managing governance-related issues of cost and organizational

flexibility on the one hand, and managing knowledge and the development process on the other. Such differentiation allows problems to be simplified to a level where individual decision makers can manage them (March and Simon 1958). But structural differentiation creates difficulties in solving problems that have simultaneous implications for the goals of different units, or problems that require information that is dispersed among different parts of the organization—problems such as how to manage firm boundaries.

Although this case highlights problems in simultaneously managing transaction costs and knowledge, recent research has demonstrated how the make-or-buy decision can impact many other organizational goals. Within the broad rubric of transaction costs, firms often need to manage problems of expropriation of knowledge (Pisano 1990) and the appropriation of customer relationships (He and Nickerson 2006), as well as the classic problem of avoiding hold up. Issues of capability development and utilization (e.g., Argyres 1996, Leiblein and Miller 2003) and competitive foreclosure can also have a bearing on integration decisions (He and Nickerson 2006, Stuckey and White 1993). Often make-or-buy decisions have simultaneous implications for several of these goals (Mayer and Nickerson 2005, He and Nickerson 2006). As we begin to understand the diverse range of outcomes that can be shaped by boundary decisions, we should appreciate that this diversity creates real challenges for decision making. Boundary decisions require the reconciliation of multiple demands, yet specialized organizational units focus disproportionately on managing only one or a few demands. The case study demonstrates the specific problems that can result from this interaction of goal diversity and organizational differentiation.

First, the case study highlights how the interaction between multifaceted problems and differentiated organizational structures can lead related decisions to be assigned to different parts of the organization. One consequence of this factoring of decisions is that no actor has full control over transaction alignment, or fully considers all of the issues needed to achieve discriminating alignment. Related decisions taken by different actors may also be inconsistent with one another, exacerbating misalignment. At the Bank, factored decision making took the form of separating decisions about overall capacity and individual transactions. Such separation is likely to be common to many other firms. Several recent studies have noted the prevalence of "tapered integration" strategies, where firms both make and buy similar inputs (Parmigiani 2007, He and Nickerson 2006, Jacobides and Billinger 2006). Contractual commitments and specialized assets often mean that there is a high cost to such firms in switching between internal and external supply. Argyres and Liebeskind (1999) note how the resulting governance inseparability across transactions can constrain individual transactions away from the optimum. This paper highlights how such governance inseparability also affects the nature of decision making: long-term decisions about how much internal capacity to maintain become separated from ongoing decisions about how to manage individual transactions. Organizations then face the risk of conflicting decisions being taken.

We should not conclude, though, that organizations should avoid such factored decision making. There are many reasons why it makes sense to divide decisions across different groups of managers. Higher levels of management often lack the information necessary to make many of the basic decisions about how to coordinate frontline work (e.g., Aoki 1990, Nonaka 1994). Organizations therefore benefit from pushing many decisions down to the frontlines. Yet it is impossible to delegate all decisions to frontline managers, because these managers often lack the information and incentives to consider how their decisions affect the broader organization. At the Bank for example, frontline managers did not bear the costs of layoffs. Had they been responsible for deciding the balance of consultants and employees, they might have hired more employees than was optimal for the firm. Some form of factored decision making is therefore inevitable in managing complex organizations.

Similarly, factored decision making may not always lead to misalignment. Organizations may be able to use integrating mechanisms, such as cross-level committees and task forces, detailed policies, or tailored incentives, to help coordinate related decisions and ensure that the goals of multiple units are considered. Indeed, how organizations can foster higher quality decision making around firm boundaries is an important topic for future research. Nonetheless, managers need to be aware that factoring decisions increases demands on decision making. To the extent that these demands are not reflected in the way decisions are made, we would predict:

PROPOSITION 1. *Transaction misalignment is more likely to occur when decisions about the level of internal capacity are taken by different actors from those making decisions about how to align resources with specific transactions.*

A second consequence of organizational differentiation is that any individual unit only attends to a limited range of information. The resulting information asymmetries are particularly problematic when decisions require processing diverse, complex information that is held in different parts of the organization. Information asymmetries between senior managers and frontline managers are an important form of such problems, as highlighted by the case study.

Research in the behavioral view of the firm highlights several factors that influence information asymmetries. For example, March and Simon (1958, p. 174)

suggest that information flows more easily within organizational and professional groups. We might therefore predict that transaction misalignment is more likely when senior managers are drawn from different business units or occupations than frontline managers, as was often the case at the Bank. Information asymmetries can also result from deliberately biased reporting within the organization (Cyert and March 1963, p. 85). This suggests that we may see more misalignment when frontline managers face stronger incentives. Research on attention also emphasizes that the firm's reporting processes shape what information is attended to (Ocasio 1997). Information asymmetries might then be more of a problem when the firm's reporting channels neglect important aspects of transaction characteristics, perhaps because of a focus on financial information. Information asymmetry can also be related to the nature of the information itself. Some information, such as financial data, is readily quantifiable and easily communicated within organizations. Other information, such as production knowledge, is much more tacit. We expect misalignment to be greater when alignment requires decision makers to access largely tacit information.

At the Bank, information asymmetries between senior management and frontline managers led to a particular form of misalignment: underintegration. We can speculate that underintegration may be the most common response to information asymmetries. Williamson (1991, p. 279) notes that market-based governance should generally be preferred in the absence of asset-specific investments. Where senior managers generally have more information on governance issues and less information on the nature of the work, they might systematically underestimate requirements for asset-specific investments. Such underestimation would be particularly likely when specific investments are based on complex knowledge, making them hard to quantify. More broadly, we would predict the following:

Proposition 2. *The greater the information asymmetry between frontline managers and senior managers, the greater the probability of transaction misalignment when make-or-buy decisions require consideration of diverse information.*

A third consequence of organizational differentiation is incentive specialization, where managers' incentives emphasize some subset of the organization's goals. If the decision problem requires attention to a broad range of goals, misaligned decision making will result.

It is possible to identify a number of contingencies that lead units to adopt more specialized incentives. First, incentive specialization can reflect differentiation in units' goals. Within a functional organization, manufacturing departments receive cost and quality goals, and sales departments receive sales goals (Lawrence and Lorsch 1967, pp. 36–37). In such organizations, the prevalence of narrow goals such as manufacturing productivity (or project completion) will promote transaction misalignment. In contrast, where a manager has full profit-and-loss responsibility and expects to remain in position for a long period of time, transaction alignment is more likely.

The use of specialized incentives can also reflect the difficulties of measuring some outcomes (Holmstrom and Milgrom 1991). At the Bank, misalignment resulted in higher cost and reduced flexibility. However, it was difficult for senior managers to make frontline managers fully responsible for the costs of the resources that they used, because of the problems of assessing when higher-priced consultants should be used. Other forms of misalignment may affect frontline managers more directly. For example, Ang and Slaughter (2001) report a case where poor use of consultants damaged their productivity. We might expect frontline managers to be more likely to minimize such direct misalignments. Yet other forms of misalignment are more difficult to measure. For example, expropriation of knowledge by suppliers (Pisano 1990) may be very difficult for senior managers to detect and have little consequence for frontline managers. We would expect that organizations would be less likely to eliminate such hard-to-measure forms of misalignment.

Specialized incentives have straightforward implications for the form that transaction misalignment will take: we expect alignment decisions to take greater account of factors that relate more directly to decision makers' rewards. At the Bank, we saw that frontline managers overweighted the management of knowledge because of the way in which they were rewarded. Elsewhere, incentives may lead managers to consider very different factors. The goals of decision makers are therefore a variable that needs to be incorporated into our theories of firm boundaries. As regards misalignment, we predict the following:

Proposition 3. *The more specialized decision makers' incentives, the greater the probability of transaction misalignment when make-or-buy decisions must achieve multiple goals.*

Limitations and Future Research

The research presented in this paper describes a single firm, and caution must be taken in generalizing the findings to other settings. That said, there is some evidence that other firms have experienced similar problems in their use of IT consultants. For example, Barley and Kunda (2004) and McNeill (2004, p. 3) describe other examples of firms becoming dependent on consultants. A study by Mayer (2006) also reports managers paying surprisingly little attention to knowledge retention in how they staff external vendors to projects.

The theoretical model developed here is most likely to generalize to similar settings. We would expect to

see similar problems where the demands of knowledge are complex, and where decisions about individual transactions are taken by frontline managers. A number of industries are likely to fit these criteria, including the construction industry, where transactions are short relative to employment relationships, and high-tech firms in which knowledge is complex and important. A particular question is how the model translates to boundary decisions that extend beyond the use of individual workers, such as when to outsource particular projects or source specific components externally. I anticipate that the model will apply most closely to situations where it is costly to adjust internal capacity. Such adjustment costs could stem either from employment contracts, as at the Bank, or from the need to make hard-to-reverse investments in capacity. These adjustment costs are likely to promote the separation of decisions about capacity versus transaction alignment, leading firms to encounter similar problems to those of the Bank.

The results of this study will also generalize most readily to boundary decisions that are not subject to strong evolutionary pressures for alignment (Masten 1993). Some environments are more tolerant of poor performance than others; some boundary decisions are also likely to be more critical to the performance of the organization. When firms face very strong selection pressures from their environment, or when transaction alignment has profound consequences for organizational performance, then organizations with serious transaction misalignment may swiftly fail (Nickerson and Silverman 2003). Theories that predict transaction alignment should hold in such settings, regardless of how firms make decisions. In many situations, however, evolution creates more limited pressures. Although misalignment did impose important costs on the Bank, those costs were small enough in proportion to the overall organization that they did not threaten its survival. In such a weak selection environment, predictions of transaction alignment rely on organizations making efficient decisions through calculation. Whether they can make the right calculations will be shaped by the structure of decision making.

Indeed, how organizations respond to misalignment is an important question for future research. It could be argued that the misalignment observed at the Bank was a transient phenomenon. As we have seen, the Bank took a number of steps to improve transaction alignment: several consultants with unique valuable knowledge were converted to employees, and policies for hiring consultants and employees were changed to reduce such misalignments in future. Similarly, Nickerson and Silverman (2003) found evidence that trucking firms reduced misalignment over time as they adapted to the environment. However, such adaptation addresses the symptoms of misalignment rather than the structural problems that lead to inappropriate make-or-buy decisions. We may therefore see cycles of under- versus overintegration, because sequential attention to different problems leads firms to restrict the use of internal and then external resources. Even where organizations successfully identify criteria that they should focus on in managing their boundaries, without addressing the structure of decisions they can face different misalignment problems in the future. For example, by the end of my study, the Bank was attempting to outsource many IT activities to lower-cost, offshore developers. Senior managers argued that one attraction of outsourcing was that the vendors would absorb fluctuations in demand, buffering the Bank's workforce from downturns. In spite of their experiences with consultants, no attention was paid to how this goal might determine what work should be outsourced and how those decisions should be made.

Extensions of this research could also consider the impact of decision processes on *how* firms contract externally. Although this study has focused on a simple make-or-buy decision, firms often manage external transactions through complex contracts that seek to manage both governance and production problems (e.g., Ang and Beath 1993, Mayer and Argyres 2004). Understanding how these contracts are written is as important as understanding when firms choose to use external vendors.

Conclusion

This paper integrates concepts from the behavioral theory of the firm with transaction cost economics and related knowledge-based theories of the firm to build a better understanding of how firms manage their boundaries. Transaction cost economics and knowledge-based theories provide powerful explanations of the criteria that firms should use to manage their boundaries. Yet they do not address how these criteria are translated into actual decision outcomes; the prevailing assumption is that firms will make the profit-maximizing decision. The behavioral theory of the firm offers a more nuanced view of decision making: organizations are usually differentiated into many specialized units pursuing different goals, and those units make locally rational decisions, rather than considering the goals of the firm. By explaining how this differentiation affects decision making, the behavioral theory of the firm helps us to understand why, and when, organizations often fail to take profit-maximizing decisions.

I build on these foundations to predict when firms will struggle to achieve transaction alignment. I argue that alignment is less likely when it requires reconciling goals that are managed in different, specialized parts of the organization. In such situations, there can be a disconnect between the diverse criteria that each specialized unit *should* use to make decisions, and the actual criteria that it *does* use. Drawing on a case study of the use

of IT consultants, I highlight three forms that this disconnect can take: separation of related decisions among different organizational units, information asymmetries, and incentive misalignment. Incorporating these variables into our theories of firm boundaries can increase our understanding of how those boundaries are set: we should be able to predict when boundaries will be set according to knowledge-based and transaction cost logics, and when those logics will have less influence on firms' decisions.

This paper also contributes to research on the behavioral theory of the firm by showing how misaligned decisions can result from the interaction of organizational structure and the nature of the decision. Recent surveys of the behavioral theory of the firm emphasize that modern research has moved away from the work's original focus on internal decision-making structures (Argote and Greve 2007, p. 344; Gavetti et al. 2007, pp. 524, 528). Rather than examining the internal effects of specialized structures, scholars have focused more on the effects of learning and standard operating procedures. This paper seeks to return to Cyert and March's (1963) initial focus on examining the mechanics of decision structures. I show how ideas from the behavioral theory of the firm can be used to make clear predictions about the outcomes of organizational decisions. In particular, I specify variables at the level of the decision problem and the organization that will shape decision outcomes. I show how these interactions between structure and problem can lead to errors in making apparently simple decisions, such as when to use external consultants. Although I focus on how decision-making structures lead to misalignment in firm-boundary decisions, related problems may also occur in making other strategic decisions. As with the decisions studied at the Bank, decision making is likely to be problematic when a single strategic decision has important implications for the ability of different organizational units to achieve their specialized goals. Better understanding these interactions between organizational differentiation and the complex nature of individual decisions can help us to both improve the predictive power of our theories of firm behavior and offer better advice to firms about how decisions should be made.

Acknowledgments

This research was financed in part by the Cambridge-MIT Institute and the MIT Center for E-Business. The author would like to thank Diane Burton, Bob Gibbons, Javier Gimeno, Henrich Greve, Quy Huy, Sarah Kaplan, Thomas Kochan, Daniel Levinthal, Paul Osterman, Gabriel Szulanski, Peter Zemsky, Andrew von Nordenflycht, the slump-management research group, Bill McEvily, and three anonymous reviewers for their helpful comments and suggestions. Seminar audiences at MIT, INSEAD, Columbia, and Wharton also provided helpful feedback. The author is particularly grateful to the many individuals at the Bank who generously gave their time to take part in this research.

Appendix. Data Collection and Analysis

Interviews

Each interview was guided by an interview protocol that I tailored to the informant and the stage of the study, as I sought to build a comprehensive picture of the Bank's use of consultants. I used very broad questions with frontline managers in the early stage of the study to understand their use of external resources and the nature of the work in the organization. In later interviews with developers, I focused on understanding their experience of work and the differences between consultants and employees. Interviews with human resources managers contained questions about the technical details of employment policies at the Bank. Most of the interviews were not recorded (recording seemed to make the subjects uncomfortable on the occasions I attempted it), but instead I took copious notes throughout, using the respondents' own language where possible. All interviews were written up on the same day.

Project Survey

The survey was conducted in person to maximize the response rate. Each survey took between one and three hours to administer (four surveys were conducted by phone because the respondents were in remote sites; all of the other surveys were conducted face to face).

Survey respondents were identified from a sampling frame of large, new development projects that had taken place in the Consumer group the previous year. Almost all of these projects consisted of incremental modifications to existing systems, rather than development of entirely new applications or modules. Focusing on this kind of project allowed me to study the use of consultants on comparable projects, and focus on projects that were more salient and would therefore yield better data. Of the 70 project managers contacted, 49 agreed to be surveyed, giving a response rate of 70%. I had particular difficulty in identifying managers in a remote location, who were therefore underrepresented in the survey. Otherwise, t-tests showed no differences between the observable characteristics of respondents and nonrespondents. Five of the surveyed managers turned out to have little knowledge of the projects in question. An additional three managers had been involved in the project in question, but had not directly managed it. I do not include their responses in the quantitative analysis because of concerns about the comparability of their responses. I did however use their broader comments about managing the projects in the qualitative analysis. The qualitative analysis is therefore based on 44 surveys, whereas the quantitative analysis uses 36 surveys for which there was complete data.

Data Analysis

The first step of the data analysis used the survey data to build a clear description of the work process and the way consultants were used. This analysis raised the possibility of transaction misalignment, because I found that the way that the Bank was using consultants was inconsistent with both the predictions of transaction cost economics and the Bank's stated goal of using

consultants to achieve flexibility. I therefore conducted further interviews and collected additional headcount data to verify that the Bank was hiring consultants to achieve flexibility, and that transaction misalignment had indeed taken place.

Next, I carried out a detailed review of my field notes and survey data to analyze the decisions that led to this misalignment. I used the survey and the interviews to trace the sequence of actions by various participants that determined how consultants were hired into the Bank and how they were then used. Evidence from many informants was used to confirm the basic sequence of events. I also studied sections of the interviews and survey where managers were talking about hiring and staffing decisions, and coded the themes that they discussed (Miles and Huberman 1994). Through this process, I identified the sources of inconsistency in the different decisions that shape transaction alignment. I compared these insights to the existing literature to sharpen my insights (Eisenhardt 1989).

I validated the emerging model in several ways (Yin 2003). I presented my preliminary findings to managers at the Bank to check the key details of the account and gain further insights into the case. I looked for disconfirming evidence that would force me to refine my concepts. I also developed alternative explanations of my findings and tested these against the data. This process was aided by presenting my findings to a number of academic and practitioner audiences, and by responding to journal reviewers.

Endnotes

[1]Argyres and Zenger (2007) suggest that knowledge and transaction cost considerations may never be fully independent in the long run, because transaction costs determine firms' choices about what knowledge to develop. In the short term, however, the empirical evidence shows that knowledge considerations exert a strong influence on make-or-buy decisions, independent of transaction costs.

[2]Although the independent consultants at the Bank were formally employed by third party consulting firms, these consulting firms did not form close employment relationships with their consultants in the way that high-end consulting firms such as Ernst & Young or Accenture would. Instead, the relationships between the consultants and consulting firms would often last only as long as the consultants were engaged by the Bank. Nor did the consulting firms do any day-to-day management of the consultants. The only role that they played in the relationship was as a staffing firm bringing the consultants and the Bank together. In addition to using independent consultants, the Bank did make some use of high-end consulting firms that formed deeper relationships with their consultants. However, the employees of these high-end firms were used very differently from independent consultants by the Bank, and were excluded from this study.

[3]We might alternatively define a transaction as all of the work carried out by a particular worker over the course of their relationship with the Bank. I have defined the transaction at the smallest possible level, consistent with Williamson's (1985) emphasis on technological separability. Nonetheless, defining the transaction at this higher level would not materially alter the facts of transaction misalignment, or the factors that contributed to it.

References

Allison, G., P. Zelikow. 1999. *Essence of a Decision: Explaining the Cuban Missile Crisis*, 2nd ed. Longman, New York.

Anderson, E. 1988. Strategic implications of Darwinian economics for selling efficiency and choice of integrated or independent sales forces. *Management Sci.* **34** 599–618.

Ang, S., C. M. Beath. 1993. Hierarchical elements in software contracts. *J. Organ. Comp.* **3** 329–361.

Ang, S., S. Slaughter. 2001. Work outcomes and job design for contracts versus permanent information systems professionals on software development teams. *MIS Quart.* **25** 321–350.

Aoki, M. 1990. Toward an economic model of the Japanese firm. *J. Econom. Lit.* **28** 1–27.

Argote, L., H. R. Greve. 2007. A behavioral theory of the firm—40 years and counting: Introduction and impact. *Organ. Sci.* **18** 337–349.

Argyres, N. S. 1996. Evidence on the role of firm capabilities in vertical integration decisions. *Strategic Management J.* **17** 129–150.

Argyres, N. S., J. P. Liebeskind. 1999. Contractual commitments, bargaining power and governance inseparability: Incorporating history into transaction cost theory. *Acad. Management Rev.* **24** 49–63.

Argyres, N. S., J. P. Liebeskind. 2002. Governance inseparability and the evolution of US biotechnology industry. *J. Econom. Behav. Organ.* **47** 197–219.

Argyres, N. S., T. Zenger. 2007. Are capability-based theories of firm boundaries really distinct from transaction cost theory? Working paper, Washington University, St. Louis.

Barley, S., G. Kunda. 2004. *Gurus, Hired Guns, and Warm Bodies: Itinerant Experts in a Knowledge Economy*. Princeton University Press, Princeton, NJ.

Barney, J. 1991. Special theory forum the resource-based model of the firm—Origins, Implications, and Prospects. *J. Management* **17**(1) 97–98.

Bidwell, M. 2009. Do peripheral workers do peripheral work? Comparing the use of highly skilled contractors and regular employees. *Indust. Labor Relations Rev.* **62**(2) 200–225.

Conner, K. C., C. K. Prahalad. 1996. A resource-based theory of the firm: Knowledge versus opportunism. *Organ. Sci.* **7** 477–501.

Cyert, R. M., J. G. March. 1963. *A Behavioral Theory of the Firm*. Prentice Hall, Englewood Cliffs, NJ.

David, R. J., S. K. Han. 2004. A systematic assessment of the empirical support for transaction cost economics. *Strategic Management J.* **25**(1) 39–58.

Dyer, W. G., A. L. Wilkins. 1991. Better stories, not better constructs, to generate better theory: A rejoinder to Eisenhardt. *Acad. Management Rev.* **16** 613–619.

Eisenhardt, K. 1989. Building theories from case study research. *Acad. Management Rev.* **14** 532–550.

Fredrickson, J. W. 1986. The strategic decision process and organizational structure. *Acad. Management Rev.* **11** 280–297.

Galbraith, J. R. 1977. *Organization Design*. Addison Wesley, Reading, MA.

Gavetti, G., D. Levinthal, W. Ocasio. 2007. Neo-Carnegie: The Carnegie school's past, present and reconstructing for the future. *Organ. Sci.* **18** 523–536.

Grant, R. 1996. Toward a knowledge-based theory of the firm. *Strategic Management J.* **17** 109–122.

Hart, O. 1995. *Firms, Contracts and Financial Structure*. Oxford University Press, Oxford, UK.

He, D., J. Nickerson. 2006. Why do firms make and buy? Efficiency, appropriability and competition in the trucking industry. *Strategic Organ.* **4** 43–69.

Hoetker, G. 2005. How much you know versus how well I know you: Selecting a supplier for a technically innovative component. *Strategic Management J.* **26** 75–96.

Holmstrom, B., P. Milgrom. 1991. Multitask principal-agent analyses: Incentive contracts, asset ownership and job design. *J. Law, Econom., Organ.* **7** 24–52.

Houseman, S. N. 2001. Why employers use flexible staffing arrangements: Evidence from an establishment survey. *Indust. Labor Relations Rev.* **55**(1) 149–170.

Jacobides, M. G., S. Billinger. 2006. Designing the boundaries of the firm: From "make, buy or ally" to the dynamic benefits of vertical architecture. *Organ. Sci.* **17** 249–261.

Jacobides, M. G., L. M. Hitt. 2005. Losing sight of the forest for the trees? Productive capabilities and gains from trade as drivers of vertical scope. *Strategic Management J.* **16** 1209–1227.

Klein, B., R. Crawford, A. Alchian. 1978. Vertical integration, appropriable rents, and the competitive contracting process. *J. Law Econom.* **21** 297–326.

Kogut, B., U. Zander. 1992. Knowledge of the firm, combinative capabilities, and the replication of technology. *Organ. Sci.* **3** 383–397.

Lawrence, P. R., J. W. Lorsch. 1967. *Organization and Environment: Managing Integration and Differentiation*. Harvard Business School Press, Boston.

Leiblein, M. J., D. J. Miller. 2003. An empirical examination of transaction- and firm-level influences on the vertical boundaries of the firm. *Strategic Management J.* **24** 839–859.

Macher, J. T. 2006. Technological development and the boundaries of the firm: A knowledge-based examination in semiconductor manufacturing. *Management Sci.* **52**(6) 826–843.

Madhok, A. 2002. Reassessing the fundamentals and beyond: Ronald Coase, the transaction cost and resource-based theories of the firm and the institutional structure of production. *Strategic Management J.* **23** 535–550.

March, J. G., H. A. Simon. 1958. *Organizations*. John Wiley & Sons, New York. (Reprint, Blackwell, Cambridge, MA, 1993. Page references are to the 1993 edition.)

Masten, S. E. 1988. A legal basis for the firm. *J. Law, Econom., Organ.* **4** 181–198.

Masten, S. E. 1993. Transaction costs, mistakes, and performance: Assessing the importance of governance. *Managerial Decision Econom.* **14** 119–129.

Masten, S., J. Meehan, E. A. Snyder. 1991. The costs of organization. *J. Law Econom. Organ.* **7** 1–25.

Masters, J., G. Miles. 2002. Predicting the use of external labor arrangements: A test of the transaction costs perspective. *Acad. Management J.* **45** 431–442.

Mayer, K. J. 2006. Spillovers and governance: An analysis of knowledge and reputational spillovers in information technology. *Acad. Management J.* **49** 69–84.

Mayer, K. J., N. S. Argyres. 2004. Learning to contract: Evidence from the personal computer industry. *Organ. Sci.* **15** 394–410.

Mayer, K. J., J. Nickerson. 2005. Antecedents and performance implications of contracting for knowledge workers: Evidence from information technology services. *Organ. Sci.* **16** 225–242.

Mayer, K. J., R. M. Salomon. 2006. Capabilities, contractual hazards, and governance: Integrating resource-based and transaction cost perspectives. *Acad. Management J.* **49** 942–959.

McNeill, R. 2004. *IT Staffing Market Consolidates*. Forrester Research Inc., Cambridge, MA.

Miles, M. B., A. M. Huberman. 1994. *Qualitative Data Analysis: An Expanded Sourcebook*. Sage, Thousand Oaks, CA.

Nickerson, J., B. Silverman. 2003. Why firms want to organize efficiently and what keeps them from doing so. Inappropriate governance, performance and adaptation in a regulated industry. *Admin. Sci. Quart.* **48** 433–465.

Nickerson, J., T. Zenger. 2004. A knowledge-based theory of the firm—The problem-solving perspective. *Organ. Sci.* **15** 617–632.

Nonaka, I. 1994. A dynamic theory of organizational knowledge creation. *Organ. Sci.* **5** 14–37.

Ocasio, W. 1997. Towards an attention-based view of the firm. *Strategic Management J.* **18** 187–206.

Osterman, P. 1987. Turnover, employment security and the performance of the firm. M. M. Kleiner, R. N. Block, M. Roomkin, S. W. Salsburg, eds. *Human Resources and the Performance of the Firm*. Industrial Relations Research Association, Madison, WI, 275–318.

Parmigiani, A. 2007. Why do firms both make and buy? An investigation of concurrent sourcing. *Strategic Management J.* **28** 285–311.

Pisano, G. P. 1990. The R&D boundaries of the firm: An empirical analysis. *Admin. Sci. Quart.* **35** 153–176.

Schilling, M. A., H. K. Steensma. 2002. Disentangling the theories of firm boundaries: A path model and empirical test. *Organ. Sci.* **13** 387–401.

Silverman, B. S. 1999. Technological resources and the direction of corporate diversification: Toward an integration of the resource-based view and transaction cost economics. *Management Sci.* **45** 1109–1124.

Stuckey, J., D. White. 1993. When and when not to vertically integrate. *Sloan Management Rev.* **34** 71–83.

Thompson, J. D. 1967. *Organizations in Action*. McGraw-Hill, New York.

Tushman, M. L., D. A. Nadler. 1978. Information processing as an integrating concept in organizational design. *Acad. Management Rev.* **3** 613–624.

Williamson, O. E. 1975. *Markets and Hierarchies: Analysis and Antitrust Implications*. Free Press, New York.

Williamson, O. E. 1985. *The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting*. Free Press, New York.

Williamson, O. E. 1988. The logic of economic organization. *J. Law, Econom., Organ.* **4** 65–93.

Williamson, O. E. 1991. Comparative economic organization: The analysis of discrete structural alternatives. *Admin. Sci. Quart.* **36** 269–296.

Williamson, O. E. 1996. *The Mechanisms of Governance*. Oxford University Press, New York.

Winter, S. 1988. On coase, competence, and the corporation. *J. Law, Econom., Organ.* **4** 163–180.

Yin, R. K. 2003. *Case Study Research: Design and Methods*, 3rd ed. Sage Publications, Thousand Oaks, CA.