How Does Outsourcing Affect Performance Over the Product Lifecycle? Evidence from the Auto Industry

Sharon Novak
Kellogg School of Management
Northwestern University
snovak@mit.edu

Scott Stern
Kellogg School of Management, Northwestern University and NBER
s-stern2@northwestern.edu

Preliminary and Incomplete!

Draft: January, 2005

This paper examines the impact of vertical integration on the dynamics of performance in the context of automobile product development. Building on recent work in contract theory (Bajari and Tadelis, 2001), we first examine a number of detailed case studies to evaluate the relationship between vertical integration and different performance margins. On the one hand, outsourcing facilitates access to cutting-edge technology and the use of high-powered performance contracts. On the other hand, vertical integration allows firms to respond to adapt to unforeseen contingencies and customer feedback, maintain more balanced incentives over the product lifecycle, and develop firm-specific capabilities over time. Together, these effects suggest that outsourcing will be associated with higher levels of *initial* performance, while vertical integration will be associated with a higher rate of performance improvement over the product lifecycle. We test these ideas with detailed data from the luxury segment of the global automobile industry. The data combine detailed performance measures over time with nuanced measures of the extent of vertical integration, as well as measures of the contracting and technology environment. Using both OLS and an instrumental variables estimator, we establish four key results. First, initial performance is declining in the level of vertical integration. Second, the level of performance improvement is significantly increasing in the level of vertical integration. Moreover, even after controlling for other factors impacting performance, the magnitude of these two effects are roughly identical – there is no relationship between vertical integration and "overall" performance. Finally, taking advantage of outsourcing during the early part of the product lifecycle and internal development during the latter years of the lifecycle depends on the institutional and strategic environment. For example, the long-term benefits to vertical integration are erased for those firms with a strong union presence. Overall, the empirical findings highlight that vertical integration is associated with both costs *and* benefits and that different performance margins will reflect the tradeoffs associated alternative contracting modes.

I. Introduction

The modern theory of the firm has made considerable recent progress explaining the determinants of vertical integration and firm boundaries (Williamson, 1985; Hart and Moore, 1990; Holmstrom and Milgrom, 1994; Baker, Gibbons and Murphy, 2002). Research in strategy and economics has placed increasing empirical content on the key tradeoffs that drive outsourcing choices, and identified the impact of shifts in the environment on make-versus-buy decisions (Whinston, 2002; Baker and Hubbard, 2004). However, while an increasing body of research has focused on the choice between vertical integration versus outsourcing, relatively little research builds on the modern theory of the firms to explore the performance *consequences* of vertical integration choices (Boerner and Macher, 2004; Klein, 2004).

The link between vertical integration choices and performance is subtle, since alternative contracting modes are *endogenous* to the economic, strategic, and organizational environment. Since the likelihood of being vertically integrated into a particular activity (e.g., the design and manufacture of a specific component or system for a product) will tend to be higher for those firms that have chosen to be vertically integrated, a simple comparison of the overall performance results between firms who have chosen vertical integration versus outsourcing are likely to be misleading (or ambiguous). For example, if the returns to vertical integration are quite significant for those firms that have adopted an integrated structure, and the returns to outsourcing are equally high for those firms who have adopted outsourcing, a cross-sectional performance comparison need not find any performance consequence to vertical integration, even though each decision maker faced a clear performance tradeoff. This challenge has motivated a small but growing literature on the "costs" of "transactional misalignment" – the performance loss associated with adopting (or inheriting) an organizational form which is "inappropriate" in a given economic or strategic environment (Masten, et al, 1991; Poppo and Zenger, 1998; Nickerson and Silverman, 2004; Saussier, 2000; Sampson, 2000).

This paper proposes an alternative approach to evaluate the relationship between performance and vertical integration choices. Specifically, our analysis builds on the insight that a single vertical integration choice affects multiple performance dimensions. Just as the modern theory of the firm is premised on the idea that any organizational choice has *both* costs and benefits, our empirical approach traces out the costs and benefits of both vertical integration and outsourcing. In other words, we draw on research in both economics and strategy to develop specific hypotheses about precise margins which are impacted by vertical integration (either positively or negatively), and then we test for an impact of vertical integration on those margins using a detailed dataset that includes margin-specific performance measures, detailed vertical integration choice measures, and a set of instrumental variables which influence vertical integration choice but are independent of the margin-specific performance measures.

We apply this insight and methodology to evaluate the relationship between vertical integration and performance over the product lifecycle in the global luxury automobile sector.[1] Our theoretical analysis draws on recent research in organizational economics (Bajari and Tadelis, 2001), as well as strategy research on the role of firm boundaries in learning, knowledge accumulation and capability development (Kogut and Zander, 1992; 1996; Nickerson and Zenger, 2003). To take advantage of these theoretical insights in our empirical context, we exploit the process of procurement and product development contracting which accompanies a "major" model change. First, each "major" model change for an automobile model provides an opportunity to significantly alter product positioning, technologies, and contracting choices for that automobile model. Typically, there are approximately five years between major model changes, with a process that takes three to five years between initiation and automobile launch. Moreover, while broad positioning choices are made by a coordinated internal team, a given major model change involves hundreds of individual contracting and governance choices, impacting each of several distinct "systems" within an automobile, such as the brakes, engine, body, etc.. Though historical factors shape contracting choices during each of major change, individual managers are able to choose a governance mode for the duration of the major model change (changes in governance mode during the product lifecycle is rare).

This setting allows us to explore how vertical integration is likely to impact performance over the product lifecycle. During the initial product development and sourcing stage, outsourcing will facilitate contracting on a global basis for cutting-edge technology, and the use of high-powered incentive contracting. In other words, the major model change provides an opportunity to take advantage of external innovations and to impose detailed performance contracts on an external supplier. Conversely, internal development limits the ability to take

[1] Macher (2004) is the only other paper we are aware of which evaluates the impact of organizational form on multiple non-subjective performance dimensions. Macher provides a complementary analysis, in several ways, and we discuss the relationship between the two papers more carefully in Section II.

advantage of frontier technology, and internal wage contracts will offer only muted incentives to reach specific initial performance targets. However, the process of product development will eventually necessitate a period of ex-post adaptation, as a particular component or system is tested and evaluated over time in the market. While market and consumer feedback provides concrete guidance for potential improvements over the remainder of the major model, external suppliers may have very limited incentives to contribute to such improvements. Specifically, because the precise needs for ex-post adaptation cannot be anticipated (and the manufacturer cannot guarantee a precise volume of "work," since the extent of needs is also uncertain), such terms will not be included in the initial performance contract. Moreover, in many cases, the capabilities required for improvement rely on detailed model-specific knowledge (e.g., effective improvement may require coordinating with other units in the firm, or may rely on idiosyncratic knowledge about the precise technical characteristics of a given automobile). Consequently, relative to outsourcing, vertical integration allows firms to adapt to unforeseen contingencies more effectively and maintain effort over the product lifecycle for those systems that require improvement.

Overall, our theoretical analysis suggests that while outsourcing may yield a higher level of "initial" performance, vertical integration will facilitate performance improvement over the product lifecycle. Moreover, there is not a "general" advantage to outsourcing; instead, the benefits (and costs) of outsourcing are realized across different performance margins realized over the product lifecycle.

Our empirical analysis proceeds in two phases. First, we report a series of short case studies illustrating the costs and benefits of outsourcing over the product lifecycle. Our qualitative findings suggest that external sourcing allows firms to access state-of-the-art technology but leaves them open to hold-up and low effort supply after the initial terms of the contract are satisfied, and that internal development is associated with inferior technology development and high costs for an initial model-year, but there are much greater opportunities for improvements over time.

We turn to a more systematic empirical analysis in Section V. Our analysis exploits an original and detailed dataset covering luxury automobile models over a fifteen year period. For each model, we observe both the degree of vertical integration and the contracting environment for seven distinct automobile systems (e.g., the brake system, the seat system, etc.). Moreover, we link these measures of system-specific vertical integration to system-specific performance measured at different points over the product lifecycle. Specifically, the analysis draws upon the annual system-specific automobile quality ratings reported in *Consumer Reports*. Given our particular empirical setting in luxury automobiles where competitive advantage is closely tied to realized quality, these measures are a useful proxy for overall vehicle commercial success (relative to alternatives for that automobile model). Finally, for each different model-system, we observe a similar set of system-specific vertical integration drivers. For example, for each system, we observe whether the firm has existing in-house sunk investments in plant and equipment. Together, these data allow for a detailed examination of the relationship between vertical integration and performance over the product lifecycle, focusing on the benefits and costs of vertical integration for different performance margins.

Though we are cautious in our interpretation, the basic empirical patterns are striking. First, systems with a low level of vertical integration are associated with a much higher level of initial performance (as measured by ratings during the first two years after the "major"), but outsourced systems experience almost no ratings improvement during the latter years of the product lifecycle. In contrast, systems that are more vertically integrated have much lower initial scores, but a very rapid rate of improvement over the lifecycle. Both of these basic patterns in the data are robust to the inclusion of alternative control structures in the performance equations, the use of system fixed effects, time fixed effects, and various company controls. As well, the qualitative results are consistent using both OLS and an instrumental variables estimator, where the instruments for vertical integration are factors that are likely independent of realized quality levels (but relate to historical circumstances that affect the adjustment costs associated with a particular governance choice). Moreover, the results are robust to alternative performance definitions and methods for capturing the difference between "early" performance measures and methods for capturing "improvements over time." Finally, the results suggest that there is no "overall" benefit to outsourcing or vertical integration within our sample. The benefits of outsourcing during the early stage of the lifecycle are equivalent to the incremental benefits received over the lifecycle associated with vertical integration.

The remainder of the paper is organized as follows. The next section provides a brief overview of the costs and benefits of outsourcing, and motivates the underlying theoretical framework. We then turn to a series of case studies to evaluate the salience of each of the key effects. After developing our empirical hypotheses more precisely, we present the dataset, and report the key empirical findings. A final section concludes.

II. The Costs and Benefits of Outsourcing[2]

Our theoretical analysis draws on recent research in organizational economics (Bajari and Tadelis, 2001), as well as strategy research on the role of firm boundaries in learning, knowledge accumulation and capability development (Kogut and Zander, 1992; 1996; Nickerson and Zenger, 2003). To develop our key predictions, we first draw on a detailed understanding of the process of automobile product development. This evaluation provides a foundation for drawing out the potential implications of vertical integration on different performance margins.

Contracting in Automobile Product Development

At the heart of our analysis is the product lifecycle for automobile models. While automobiles receive incremental upgrades annually, an automobile model undergoes a "major" model change approximately every five years. A "major" model change provides an opportunity to significantly alter product positioning, technologies, and contracting choices for an automobile model. Of course, even for major model changes, a manufacturer is constrained by the history of the vehicle, sunk investments, etc. However, the process underlying a major model change is substantial, and there is typically a 3-5 year period between initiation of the product development process for a major change and the launch of a new automobile model.

Product development of a new vehicle or a major model change begins with a "vehicle integrity" team which chooses broad vehicle performance and positioning (i.e. "The Ultimate Driving Machine"). Work is decomposed into key system technology requirements (e.g., Engine Horsepower) and further decomposed into sub-systems and then individual components. Once the key positioning and technology choices have been made, sourcing and procurement take place at the component level. The purchasing decision determines the extent of external product development contracting. Although purchasing decisions are made at the component level, there are significant technological interdependencies at the system level. For example, the energy absorbing device is a seemingly simple sheet metal piece that functions as part of the steering

[2] This section is PRELIMINARY, including the omission of key references supporting the argument.

system. By its appearance ("simple" design, readily available materials and processes), it looks as if its production should be outsourced, but every automobile manufacturer produces it in-house because of the important role it plays and of the complex interactions it has with virtually every other component of the steering system. These interactions require it to be developed from a system level perspective and not a component level, as any changes to the energy absorbing device must be carefully coordinated with all other parts, as they can drive changes to any or all of them in product development. The key technology and contracting choices made for the "major" model change can significantly constrain contract choice for the life of the major. Firms lack flexibility to transition from in-house production to outsourcing because it is extremely costly to contract for external suppliers if the project has been maintained internally in its initial stages. The difficulty of finding external suppliers for a "short" contract is compounded by the significant penalties external suppliers impose for supplier switching during contract life if they meet observable performance requirements. However, though the decisions are fixed in the "medium-term," the underlying contracts combine detailed specifications with a large degree of contractual incompleteness.

Contracts contain detailed provisions governing initial contract performance requirements for external contracts, including the ability to pass key safety and production thresholds, commitments to satisfy specific technical requirements, etc. Although the contract language includes requirements for continued involvement and updating in response to customer feedback, and incremental model improvement, there are very few mechanisms to enforce these contract provisions.

Figure 1 outlines the timeline underlying product development contracting over the product lifecycle. During the earliest stages, the manufacturer has latitude to access any supplier and set detailed requirements, most of which relate to a large number of foreseeable contingencies. Whereas external suppliers are offered high-powered performance requirements contracts, internal suppliers are provided more muted incentives, often yielding a higher level of "coordination" with other components and/or systems (Novak and Stern, 2004). Initial performance realization motivates incremental innovation and improvement over the life of the major. Whereas external suppliers have few incentives for further effort, internal teams are provided a constant level of incentives, are learning during the project, and can be directed through authority relationships. Total profits depend on the success of the vehicle across the entire product lifecycle.

These aspects of the product development process lead to a linkage between specific performance margins and the extent of vertical integration. Overall, our theoretical analysis suggests that while outsourcing may yield a higher level of "initial" performance, vertical integration will facilitate performance improvement over the product lifecycle. Moreover, there is not a "general" advantage to outsourcing; instead, the benefits (and costs) of outsourcing are realized across different performance margins realized over the product lifecycle.

The "Early" Years in the Lifecycle

There are two key benefits to outsourcing realized during the early years of the product lifecycle, the ability to access frontier global technology and the ability to write and enforce detailed procurement contracts with high-powered incentives. First, when one chooses an external procurement mode, one is able to access the "best" in global technology and capabilities through a competitive bidding process. In contrast to the *ex ante* capability levels of internal teams, each bidder is themselves a specialist, vying with each other to achieve "best in class" and are able to take advantage of their learning from multiple projects within a given system. This type of access to frontier technology is particularly important in quality-sensitive segments such as the luxury segment, as we discuss below.

In addition, there may be significant incentive effects associated with outsourcing that would favor the short-term performance margin. Similar to the analysis of Bajari and Tadelis (2001), one of the key potential benefits of outsourcing is the ability to (endogenously) induce contractibility on a set of observable performance measures. These detailed contract provisions induce high-powered incentives for meeting specific performance requirements relating to the achievement of technical specifications and cost objectives before the initial date of product launch. In contrast, internal development teams are governed by wage contracts and authority relationships, and there are only modest performance penalties in place for a given failure. While subjective incentive schemes and promotion can provide reasonable incentives, it is very unlikely (and, more importantly, not even optimal, given the other unobservable dimensions of effort you would like to encourage effort towards) to employ internal development teams with the same high-powered incentive schemes as one would an external supplier.

The "Late" Years in the Lifecycle

A very different set of capability and incentive effects characterize the later years in the lifecycle. In particular, while access to global technology was important in the early part of the lifecycle, performance improvement after product launch requires detailed model-specific knowledge. While it is likely that an external supplier is the "best in class" from a global perspective, internal teams will have capabilities and knowledge that is idiosyncratic to a particular model, thus facilitating performance improvement over the lifecycle.

At the same time, incentive contracting also tilts towards vertical integration. While outsourcing involves detailed contracts, the enforceable terms of the contract (in terms of technology specification and quality) are largely satisfied by the time that the initial major model is introduced. While prospects for future contracts and general reputation provide some incentives for continuing effort, the lack of a direct authority relationship or reliance on subjective incentive schemes reduces incentives for ongoing quality improvements. This effect is reinforced by the organization of supplier activities, in that at the end of a project, supplier employees are immediately allocated to new projects. A post-contract change may require ten engineers who are already staffed elsewhere. The inability to access key personnel is reinforced by legal contracts surrounding future projects. To maintain secrecy and avoid expropriation (separate staff, facilities), contracts limit the extent to which employees can be pulled from other projects to return to an earlier one. In contrast, the balanced incentives and authority relationships that characterize internal development allow these teams to be able to provide significant effort in response to the need for ex-post adaptation (Bajari and Tadelis, 2001). In particular, it is both feasible to provide incentives to internal teams to undertake activities which are "non-contractible" (e.g., through promotion incentives), and, perhaps as importantly, the firm can use its authority relationships to adjust the level of effort and the composition of personnel optimally to respond to the specifics of consumer feedback. As a result, the incentive effects of vertical integration may be most salient during the latter years of the product lifecycle.

Overall, we have simply drawn out the benefits and costs of vertical integration in the context of procurement, and suggested that, while there may be no relationship between vertical integration and overall performance, there may be close linkages between vertical integration and specific performance margins. To deepen this intuition, we draw on detailed knowledge about the process of product development through a series of case examples, which we turn to in the next section.

III. Case Studies from Automobile Product Development

In this section we present four cases drawn from our data collection to illustrate the key potential benefits and costs resulting from internal development versus outsourcing of systems. In particular, we highlight differences in capabilities and in performance incentives.

The Volvo 850

The 850 model was Volvo's first US front-wheel drive vehicle. In order to balance weight in a front-wheel drive vehicle, the engine is aligned east to west, rather than north to south, as in rear-wheel drive vehicles. Such an east-west, or "transverse axis" engine design required an extremely narrow, "short" gearbox for the automatic transmission. With no experience producing automatic transmissions, Volvo lacked the internal expertise to produce such a complex and unusual design in time to meet its product launch deadlines. Instead, Volvo contracted the gearbox design to an outside supplier, Aisin Seiki. As part of the contract, Aisin Seiki maintained resident staff at Volvo to meet daily with body and engine designers as changes were being made. Not only was initial product performance successful, but, over time, Volvo was able to take over key functions through learning, such as maintaining software, etc.

Volvo's outcome with Aisin Seiki is an example of using outsourcing to access the best in global technology. Any given company will only have a limited number of sources of internal expertise that will allow a luxury automaker to differentiate in terms of design and technology. Of course, commitment to an outside supplier will lock the manufacturer in to a given technology and vendor, raising the potential for hold-up. However, if there are few unforeseen contingencies in initial product development, as in the Volvo-Aisin Seiki example, the manufacturer can realize a technological "leap" along with the opportunity for learning going forward through outsourcing.

The 1992 Cadillac Seville ABS Brake System

In contrast, internal development may lead to poor initial performance, and the need for costly revisions. Until the early 1990s, the dominant antilock brake system (ABS) design was mechanically-based, that is, the designs were based on motors. By 1990 Cadillac had invested $100 million in mechanically-based ABS technology. As a result of that investment, Cadillac's internal brake division, AC Delco, was heavily focused on the specifics of the mechanically-based technology. However, in 1989, Bosch, a leading global supplier, dramatically redesigned an electronically-based system, based on solenoids, a completely different approach to ABS, which offered better performance at half the price of the mechanically-based systems. Cadillac was already in development of the 1992 Seville, and attempted to respond to the Bosch innovation by implementing a "hybrid" solution that combined Bosch and AC Delco parts. The hybrid system had significantly higher cost and worse performance than the Bosch alternative.

The Cadillac outcome demonstrates one of the key risks of internal development, as firms rely on a narrower range of in-house capabilities when sourcing internally. As internal suppliers are not as close to global technology developments, they are at a greater risk of being leapfrogged in terms of initial product development due to being more constrained by internal capability limitations.

The Toyota Lexus LS400

On the other hand, internal development can offer the potential for significant improvements over time. During the late 1980s, Toyota created Lexus as a new brand and division in order to position Toyota to enter the luxury auto segment. Toyota had no experience in producing such high-end vehicles, with their related increase in quality requirements, and indeed the initial design was extremely expensive for Toyota to produce, particularly to maintain a quality level that would not suffer too strongly in terms of quality ratings. This design relied on use of a large number of separate parts for reinforcement, and there were are few key outside suppliers who could have completed the job for Toyota at a lower cost. Toyota maintained a single project manager over the first two Lexus product development projects, and over time, Toyota realized significant reductions in cost alongside continuing quality improvements in terms of design and part simplicity, eventually rising to the #1 position in quality ratings. The Lexus project manager stated, "Even when we did the two piece body outer, we thought about doing a one piece. We started to plan for this during the first program." Employees were rewarded for their ability to contribute to this goal. This quote reflects one of the key benefits of internal development, that project management can manage beyond the "life" of an individual project and invest in objectives and criteria that are difficult to write down and enforce in contracts. Toyota used its high personnel stability along with internal promotion incentives to design a plan for gradually building the key skills needed to deliver high quality products. In this way, internal development reinforces the potential to exploit authority relationships and subjective incentive compensation.

The Saab 9000

Finally, significant quality problems can remain unresolved with outsourcing because external suppliers lack the ongoing incentive to improve product design after meeting the initial contract terms. SAAB, for example, went through two near bankruptcies and merger attempts in the 1980s/1990s. The original 9000 was planned as a sister vehicle to the Fiat Lancia, but finally ended up as part of the General Motors Opel Vectra platform in the 1990s. SAAB had initially contracted with Hella for brakes prior to the merger with GM, but adapting the 9000 to the Vectra platform after the merger necessitated a great deal of changes to the brake design. The resulting brake performance was fraught with problems that required costly changes. According to a VP at Saab, "Savings can disappear overnight due to changes." The original contract with Hella did not include language for changes driven by the merger, and it was very costly to enforce those changes, as Hella had already met the terms specified by the original agreement. The SAAB story exemplifies a key issue with outsourcing, in that external suppliers face too few incentives, and, after putting forth initial effort to satisfy the terms of the contract, have few incentives to invest in the specific problems and challenges that arise in the context of an individual manufacturer.

Figure 2 summarizes the findings from these case examples, In particular, Figure 2 complements our theoretical discussion and motivates the specific hypotheses we examine in our empirical analysis. In particular, the case examples offer two testable implications of the impact of vertical integration of different performance margins.

Hypothesis 1: The level of initial performance is lower for higher levels of vertical integration.

Hypothesis 2: The level of performance improvement is higher for higher levels of vertical integration.

Of course, if the size of one of these two performance margins was significantly larger than the other, there would be a net positive return to a particular contracting model. For example, if the performance improvement benefits from integration were much larger than the costs associated with the initial performance penalty, then, for most firms, the optimal contracting mode would be through integration. However, when we observe significant variation in contracting practices (as we do in the automobile industry), our framework offers a third potentially testable implication:

Hypothesis 3: There will be no systematic relationship between vertical integration and a measure of "overall" performance over the full product lifecycle.

The remainder of this paper focuses on testing these three hypotheses. To do so, we must exploit a dataset that links performance data over the product lifecycle with measures of vertical integration, and measures of the overall contracting and technology environment. The next section describes these data in detail.

IV. Data

Sample and Methods

This study combines a proprietary and original dataset based of contracting choices and the contracting environment in the global auto industry with system-specific ratings drawn from *Consumer Reports.* As discussed more fully in Novak and Eppinger (2001), the data on contracting choices, product architecture, and the contracting environment was constructed from a multi-year study of the global automobile industry. The dataset consists of observations from the luxury performance car segment (defined by *Consumer Reports* as vehicles priced above $30,000 in 1995) and the companies included in the sample are drawn from Europe, the U.S. and Japan, accounting for roughly 90% of revenues in the global luxury performance market. As flagship vehicles developed in different environments over time, wide variation in contracting practices (and the contracting environment) was expected, and competitive advantage in this segment is likely highly sensitive to quality (relative to price tradeoffs). By focusing on a single vehicle segment, we limit the measurement problems that arise from combining information from different vehicle types.

The unit of analysis is an automotive system for a specific "major" for a given automobile model. As discussed earlier, "major" model changes," which are typically implemented at approximately five-year intervals, provide an opportunity to significantly alter product positioning, technologies, and contracting choices for an automobile model. Overall, the dataset includes comprehensive information about seven systems for 19 automobile "major" model versions between 1980 and 1995.[3] The data were collected through on-site interviews with over 1000 people, including CEOs, chief engineers, project managers and the system engineers involved in the development of each model-year. All participants were assured that only aggregate data would be presented, and confidentiality agreements were signed with each company.

Data collection proceeded in several stages. After signing an agreement with each firm, a letter was sent requesting interviews with relevant project managers, system engineers, design engineers, purchasing managers and manufacturing engineers for each vehicle for each time period. The relevant parties were identified by the corporate liaison for each company, and on-site meetings were arranged. To ensure data accuracy, interviewees were given an overview of the research project and definitions for key terms. Subjects were given a list of questions pertaining to the design and sourcing of components within their respective systems. The questions focused on principally objective information (e.g. number of parts in the body side) so as to minimize the likelihood of response bias. The interviews were conducted on-site at each company, in time intervals ranging from three days to three months. All interviewees were given the option of being interviewed in their native languages. US and European interviews were conducted in English and Japanese interviews were conducted in Japanese.[4]

We combine this contracting choice data for a given "major" model change with system-

[3] More precisely, the overall dataset includes information about 8 distinct car models, many of which are observed at (roughly) five-year intervals, with 19 total "model-years" for which complete data were available.
[4] All interviews were conducted by one of the authors. Professor Kentaro Nobeoka, a scholar with extensive experience in the Japanese auto industry, provided Japanese interview interpretation.

specific performance ratings published by *Consumer Reports.* These performance ratings are used to construct the key dependent variables throughout the analysis. The *Consumer Reports* rating varies from 1-5, with 5 as the "highest rating") for each system for each model-year. For each of the 19 "major" model changes that we examine, we gather data for the first four years after the introduction of that "major" change. We measure short-term performance in terms of the ratings supplied for the first two model-years of the major, and long-term performance in terms of the ratings supplied for the third and fourth model-years of the major. When performance measures were available from both of the two years, we simply took an average; when only one year is available (within the relevant two-year span), we take that as our performance measure.

It is important to note that, for each model-year, there are potentially multiple years of performance data, since the performance ratings for a given model-year will be updated even after the year of initial introduction. In our main analysis, we examine the model-year performance rating associated with the *first* rating provided for that vehicle model-year; as well, our key results are robust to alternative formulations of the year-by-year performance measures.

The original sample consists of 133 model-year systems, drawn from nineteen distinct "major" model changes (associated with seven different automobile models) and across seven distinct systems for each model: engine, transmission, body, electrical, suspension, steering, and brakes. From this initial dataset of 19 models, each of which includes seven distinct systems, 2 models were excluded from the analysis for inadequate data, leaving 119 observations. As well, a small number of system-specific performance measures were unavailable for individual years and systems (largely because of "inadequate data").[5]

The final dataset consists of 112 observations of system-specific contracting choice, the contracting environment, and performance. Table 1 provides all variable names, definitions, as well as summary statistics.

System-specific performance measures

The key dependent variables throughout the analysis are a series of performance measures drawn from *Consumer Reports.* For each system i on model j in year t after a major

[5] We actually have complete data for 114 observations for our regression related to short-term performance and 112 observations for regressions related to the long-term performance measure. The results do not change if the 2 observations for which no data on long-term performance are included or excluded from the analysis.

model change, PERFORMANCE RATING$_{ijt}$ is the *Consumer Reports* quality rating for that system, ranging from 1-5, with 5 as the "highest rating.").[6] The mean of PERFORMANCE RATING is 3.54, with a standard deviation of just under 1. Overall, while relatively few vehicles receive a rating of 1, there is significant variation in the performance ratings, across systems, automobiles, and time. We use PERFORMANCE RATING$_{ijt}$ to calculate our measures of short-term performance, long-term performance, performance change, and overall performance.

First, SHORT-TERM PERFORMANCE is the average of PERFORMANCE RATING$_{ij0}$ and PERFORMANCE RATING$_{ij1}$, as available. When only one *Consumer Reports* rating is available, only one is used. In other words, SHORT TERM PERFORMANCE is a measure of the performance measure during the first two years of the product lifecycle (inclusive of the introduction year). Similarly, LONG-TERM PERFORMANCE is the average of PERFORMANCE RATING$_{ij2}$ and PERFORMANCE RATING$_{ij3}$ as available. When only one *Consumer Reports* rating is available, only one is used. We use these two measures ton construct a measure of PERFORMANCE CHANGE, which is set equal to the difference between the LONG TERM PERFORMANCE and SHORT TERM PERFORMANCE variables. There is significant difference in the mean levels of SHORT TERM PERFORMANCE and LONG TERM PERFORMANCE (mean = 3.43 versus 3.70). In other words, across the sample, there is a modest upward trend in vehicle ratings over the product lifecycle. Finally, OVERALL PERFORMANCE is the average of the LONG TERM PERFORMANCE and SHORT TERM PERFORMANCE variables, and provides a measure of the overall performance of the vehicle over the duration of the major change.[7]

Contracting Variables

The key contracting measure throughout the analysis is VERTICAL INTEGRATION, the percentage of the system produced in-house, with 1 indicating in-house production of all components within that system.[8] For each component, system, vehicle model, and time period,

---

[6] As mentioned earlier, for each model-year, the rating that is chosen for a given model-year is the "first" *Consumer Reports* rating available for that model-year.

[7] It is useful to note that, for any given major-mode, OVERALL PERFORMANCE is not simply the average of PERFORMANCE RATING. Instead, we require only one observation in the first two years after the major model introduction, and one observation within the third and fourth year of the major model. As such, a simple average will confound differences in the number of reported quality ratings with differences in the levels of those ratings.

[8] Masten et al (1989) use a similar measure at the component level.

we have collected data on the make / buy decision outcome. The vertical integration measure at the system level is calculated as the average across the individual components for that system, and each component is weighted equally. Parts supplied to firms by wholly-owned subsidiaries, such as the Delphi division of General Motors, are treated as in-house. Parts produced by partially owned suppliers, such as Nippondenso (Toyota group), were treated as outside suppliers.

VERTICAL INTEGRATION exhibits substantial variation across the sample, ranging from 0 (fully outsourced) to 1 (in-house production), with a mean of .51 and a standard deviation of .32. Moreover, it should be emphasized that much of the variation in VERTICAL INTEGRATION is "model-specific." For example, in an OLS regression of VERTICAL INTEGRATION on individual model-year dummies, R^2 = 0.58, most of the individual model-year effects are individually significant, and the overall F-test statistic is highly significant at 8.74. In other words, vertical integration is "clustered" according to model-year. We will exploit this correlation in VERTICAL INTEGRATION across systems within a given model to construct instrumental variables for VERTICAL INTEGRATION in the context of a performance regression. To do so, we calculate VERTICAL INTEGRATION$_{-i}$, which is the sum of VERTICAL INTEGRATION across all *other* systems within that model. Consistent with the framework and evidence in Novak and Stern (2004), we will assume that VERTICAL INTEGRATION$_{-i}$ may impact VERTICAL INTEGRATION, but that it will have no direct impact on system-specific performance measures (since the correlation between VERTICAL INTEGRATION among systems accrues to performance associated with the interaction and coordination across systems, a separate performance margin which we do not directly observe in the current paper).

System-Specific Contracting and Performance Drivers

Our analysis also includes a set of system-specific contracting and performance drivers. Overall, these measures are included to control for model-specific performance drivers that may be themselves correlated with VERTICAL INTEGRATION. In other words, in our discussion of each of these measures, we discuss both their relationship to VERTICAL INTEGRATION (and, in some cases, highlight their potential as instrumental variables), and their relationship with different performance margins.

First, SUNK COST is a dummy variable indicating whether there is pre-existing in-house sunk investments for each system (mean = 0.14). Specifically, managers were asked whether or not existing plant equipment directly affected their design choices for the system, as systems are often designed around plant-specific process equipment investments. On the one hand, the existence of pre-existing in-house capital investment will tend to favor a positive relationship between VERTICAL INTEGRATION and SUNK COST at the system level; as such, we employ SUNK COST_{-i} as an instrumental variable for VERTICAL INTEGRATION in the IV analysis. As well, when SUNK COST = 1, this may *either* indicate that a company has significant experience and capabilities in a given system (favoring a positive relationship with performance), or that they may face relatively high adjustment costs in adopting frontier technology (perhaps leading to a negative relationship with performance, particularly in the earliest parts of the product lifecycle).

LOW CAPACITY is a dummy variable indicating that, prior to contracting, the level of in-house capacity is insufficient to manufacture the system in-house (mean = 0.17). If a certain system, like a one-piece body side, exceeds the capacity of current plant equipment, this will necessitate new physical investment. As with SUNK COST, the impact on performance is ambiguous. Specifically, LOW CAPACITY may indicate a lack of capabilities in a given system (favoring a negative relationship with performance), or perhaps suggest an increased propensity to adopt frontier technology (perhaps leading to a positive relationship with performance, particularly in the earliest parts of the product lifecycle). However, it is useful to note that LOW CAPACITY is likely to enhance the relative returns to outsourcing, and so we predict a negative relationship between VERTICAL INTEGRATION and LOW CAPACITY.

Turning to factors related to system-specific design and technology choice, PLATFORM is a dummy variable equal to one for models with platform requirements where the component was designed to be used by more than one vehicle. Overall, this measure may have a complicated impact on performance over the product lifecycle. In the short-term, platform requirements may enhance or detract from initial performance, depending on a combination of the level of investment, innovation and capabilities underlying the platform development process. However, platform requirements are predicted to have a positive impact on PERFORMANCE CHANGE (as the firm is likely developing relevant competencies, and also has higher incentives to improve in response to feedback). Most importantly, PLATFORM may enhance the potential positive impacts of VERTICAL INTEGRATION over the latter stages of the lifecycle. Specifically, precisely to the extent that platform requirements will be associated with the development of specific capabilities and higher intrinsic incentives for improvement over time, PLATFORM may enhance the boost to performance over time associated with VERTICAL INTEGRATION. Moreover, PLATFORM may have a direct effect on VERTICAL INTEGRATION itself. In terms of the impact on contracting, platform requirements could support in-house production through economies of scale achieved through parts sharing. For this reason, we expect a positive relationship between PLATFORM and VERTICAL INTEGRATION.

As well, the degree of system-specific complexity may impact realized performance and the level of VERTICAL INTEGRATION. As developed in Novak and Eppinger (2001), the degree of system-level complexity will impact the need for coordination across component elements of the system, encouraging in-house contracting. Our measure of system complexity draws on several measures, based on detailed system design and manufacturing data. For each system, we estimate product complexity on a scale from 0 to 1 (no complex system interactions to high product complexity) based on an unweighted average of characteristics of design complexity.[9] For some systems, measures include characteristics such as "newness" - the degree to which a design configuration has been used in the company and in the vehicle. For example, product complexity in the suspension system is calculated as an unweighted average of three (0-1) measures: newness of the design, number of moving parts in the suspension and whether the suspension is active or passive.[10] The measure used in our analysis, COMPLEXITY (mean = .39), is the result of applying this procedure for each component within each system.

As mentioned earlier, we will use these system-specific measures to construct instrumental variables for VERTICAL INTEGRATION. Specifically, while, for any system, *i*, we include the direct effects associated with these measures (e.g. ,SUNK COST_i), we will use the drivers of VERTICAL INTEGRATION on other systems (e.g., SUNK COST_{-i}) as instruments for the level of VERTICAL INTEGRATION associated with system i. Interestingly, within this data, there is substantial variation in these measures across systems within each model-year. Of the 19 model-years in our sample, 13 exhibit variation across

[9] The system-specific complexity measure is based on system engineering principles (Novak and Eppinger, 2001).
[10] *See* Novak and Eppiinger (2001).

systems in SUNK COST and 17 exhibit variation in LOW CAPACITY (and all models have variation in at least one of these measures). As well, all models display variation in the PLATFORM and COMPLEXITY measures. In other words, in this dataset, there is substantial variation in exogenous measures of system-specific variation, which we will exploit to identify vertical integration in the context of assessing the impact of vertical integration on different performance margins.

PERFORMANCE is a dummy variable equal to 1 if an individual system is associated with "high" system-specific performance goals. The importance of performance goals were provided by vehicle product managers, on a 0-10 scale, with 0 indicating no importance for product performance goals and 10 indicating that the vehicle competes based on high performance. While PERFORMANCE is reflecting the ex ante performance expectations for the system, PERFORMANCE is predicted to have a positive impact on each of the performance measures. However, the relationship with VERTICAL INTEGRATION may be subtle. Certain performance goals necessitate more complex product designs, such as more integrated architectures (Ulrich, 1995). The need for such integration enhances the returns to vertical integration. However, as discussed earlier, accessing global frontier technology may necessitate outsourcing. As such, there is an ambiguous relationship between PERFORMANCE and vertical integration.

SKILL SHORTAGE (mean = .15) is a dummy variable equal to 1 if key system-specific worker skills are absent within current plant locations. For example, it is much more costly to produce a body design featuring many complex manual welds in an area where workers are not trained in advanced welding. Vehicle product managers were asked whether the absence of worker skills played a role in design considerations for each system. SKILL SHORTAGE may reduce performance across the product lifecycle, though the potential to alleviate a skill shortage through learning and investment over time suggests that SKILL SHORTAGE may be associated with a higher level of PERFORMANCE CHANGE.

Finally, we observe one measure at the model (rather than model-system) level, UNION. UNION is a dummy variable which is equal to 1 if *any* component is produced in house and covered under a union agreement. While the role of unions in initial performance is unclear (e.g., UNIONS may be associated with higher or lower *ex ante* capability levels), a high UNION presence may reduce the potential for ex-post adaptation. As such, we expect that UNION will be negatively related to PERFORMANCE CHANGE, and, moreover, that the interaction effect between PERFORMANCE CHANGE and UNION will also be negative. However, through much of the analysis, we will introduce company fixed effects; because there is insufficient variation across model-years within a company, we cannot separately identify UNION after the inclusion of company fixed effects.

System, Year, and Company Fixed Effects

We also calculate fixed effects for each of the seven automobile systems (SEATS are the excluded category), and also introduce various time controls (including a linear time trend, and annual fixed effects). As well, in several specifications, we include company dummies. The empirical analysis explores each of these control structures to identify the precise source of variation in the dataset driving our key findings and to highlight the robustness of key results to focusing on alternative sources of variation.

V. The Empirical Framework

The empirical framework is straightforward. Essentially, we examine three different performance measures (SHORT-TERM PERFORMANCE, PERFORMANCE CHANGE, and OVERALL PERFORMANCE), and examine each using both OLS and instrumental variables estimation. In other words, the objective of the empirical analysis is to estimate the following three equations:

$$SHORT\ PERFORMANCE_{i,j} = \beta_{STP} + \lambda_{STP} VI_{i,j} + \beta_{STP,Z_i} Z_{i,j} + \chi_{i,j}^{STPERF}$$
$$PERFCHANGE_{i,j} = \beta_{PC} + \lambda_{PC} VI_{i,j} + \beta_{PC,Z_i} Z_{i,j} + \chi_{i,j}^{PC}$$
$$OVERALL_{i,j} = \beta_{OVERALL} + \lambda_{OVERALL} VI_{i,j} + \beta_{OVERALL,Z_i} Z_{i,j} + \chi_{i,j}^{OVERALL}$$

According to our key hypotheses, $\lambda_{SP} < 0$, $\lambda_{PC} > 0$, and $\lambda_{OVERALL} = 0$

As suggested above, we will report both OLS and instrumental variables estimates for each of these equations. However, it should be emphasized that while the extent of vertical integration is certainly an endogenous choice variable, an OLS estimator can still offer a consistent estimate of the impact of vertical integration on the specific performance margins that we observe. In particular, in our OLS analysis, we simply assume that, conditional on observable, each firm receives a mean-zero relative cost shock that affects the costs of vertical integration (relative to outsourcing) for each system-model-year. As long as this sunk cost shock (one might think of this as an adjustment cost, or a cost to effective coordination across systems (Novak and Stern, 2004), the OLS estimates in the performance equation will provide a consistent estimate of the impact of vertical integration on these specific performance margins.

Of course, it is possible that the error in the vertical integration equation is in fact correlated with one or more of the above performance equations. In particular, those firms who choose a high level of vertical integration for a particular model-system are likely facing relatively high returns to vertical integration, even in terms of the performance benefits of vertical integration. As such, we compare our estimates for each procedure with a set of instrumental variables which are uncorrelated with the system-specific performance margins. To do so, we draw on our own prior work which identified the potential for interdependencies among vertical integration choices for different systems across an automobile (Novak and Stern, 2004). In that paper, we were able to present evidence that the level of vertical integration on a given system is influenced by the level of vertical integration on other systems for the same automobile model. Moreover, this complementarity across systems is driven by the returns to coordination across systems, which is independent of system-specific performance. As such, we are able to construct a set of instrumental variables for VERTICAL INTEGRATION: the level of vertical integration on other systems (VERTICAL INTEGRATION$_{-i}$), as well as measures associated with the drivers of vertical integration on other systems (e.g., LOW CAPACITY$_{-}$i, SUNK COSTS$_{-i}$). Consistent with the preliminary nature of the analysis, we present both types of estimates, which are broadly consistent with each other.

VI. Empirical Results

The empirical analysis proceeds in several stages. First, we simply present descriptive evidence about the patterns of performance over the product lifecycle, according to the extent of vertical integration. Second, we evaluate the impact of vertical integration on initial performance. Third, we examine the impact of vertical integration on the rate of performance improvement. We then evaluate the impact of vertical integration on an overall performance measure. Our final set of empirical results concerns the impact of interaction effects between vertical integration and institutional and organizational factors that might impact the returns to vertical integration on specific performance measures.

Performance Dynamics Over the Lifecycle

Our analysis begins with Figure 3, where we plot the mean of PERFORMANCE RATING, by the years since the introduction of the "major," divided according to whether the system-model is above or below the median level of vertical integration. In the initial model-year, there is a quite pronounced difference in the performance level (3.62 versus 3.09). However, over the product lifecycle, there is convergence in the "raw" performance levels by the fourth year after product introduction. A similar pattern is observed in Table 2, where we divided out the average of SHORT TERM PERFORMANCE and LONG TERM PERFORMANCE by the extent of vertical integration. As before, there is a significant difference in SHORT TERM PERFORMANCE, but convergence in the raw levels for LONG TERM PERFORMANCE. Of course, it is important to recognize that these patterns could simply reflect correlation with other potential drivers of performance, and so it is important that we evaluate these patterns more systematically in a regression model.

Short-Term Performance

Table 3 reports the findings associated with SHORT TERM PERFORMANCE. First, reflecting the patterns in Table 2, a simple regression with VERTICAL INTEGRATION has a large and statistically significant negative relationship with SHORT TERM PERFORMANCE. As well, this pattern is robust to the inclusion of a linear time trend, as well as a set of six system-level dummy variables. It is useful to note that there is a significant upward time trend in the level of SHORT TERM PERFORMANCE, but that this is essentially independent of the relationship between VERTICAL INTEGRATION and SHORT TERM PERFORMANCE. In (3-3), we include the full set of control variables; interestingly, while there is only a modest decline in the level of the VERTICAL INTEGRATION coefficient, no single regressor (except for the time trend) is also significant. Indeed, across a wide range of OLS specifications, there is a strong relationship between VERTICAL INTEGRATION and SHORT TERM PERFORMANCE, and, except for the robust presence of a positive time trend, there is no other measure which has a consistent relationship with SHORT TERM PERFORMANCE.

In Table 3B, we turn to the instrumental variables approach described earlier. Specifically, the instrumental variables are drawn from prior work emphasizing the potential for interdependencies among the vertical integration choices of different systems across an automobile. In Novak and Stern (2004), we emphasized that complementarity across system should *not* impact system-specific performance, but does affect overall vehicle performance through coordination and competitive advantage through secrecy concerns. This finding motivates the use of VERTINT-i,j , as well as SUNK COST-i,j and LOW CAPACITY-i,j . Instrumental variables point estimates are even larger (in absolute value) than OLS point estimates (though one cannot reject equality of the estimates). Moreover, the results are robust to the inclusion of a full set of control variables, including firm-level fixed effects. As in the OLS specification, there is no robust driver of SHORT TERM PERFORMANCE except for VERTICAL INTEGRATION and the time trend,. Overall, whether one employs OLS or instrumental variables, there is a quantitatively and statistically significant relationship between vertical integration and SHORT TERM PERFORMANCE.

Performance Change

Table 4 reports a similar set of regression specifications for PERFORMANCE CHANGE. In (4-1), we simply include VERTICAL INTEGRATION by itself, and it has a positive and quite large impact on the predicted level of PERFORMANCE CHANGE (recall that PERFORMANCE CHANGE has a mean of 0.28 and a standard deviation of 0.83). We then include a set of control variables. In particular, in addition to the measures we included earlier, we also include SHORT TERM PERFORMANCE. While it is feasible to shift both up and down in the ratings, a higher initial rating will tend to "constrain" the potential for performance improvement, and so we control for that effect directly in the analysis. Interestingly, the magnitude of the VERTICAL INTEGRATION coefficient increases, as it does with the inclusion of system-level dummy variables, and a linear time trend. As well, the data suggest that there is indeed a "mean reversion" effect, so that models receiving higher levels of initial performance ratings tend to experience a lower level of PERFORMANCE CHANGE (equal in magnitude approximately to the size of the VERTICAL INTEGRATION effect that is the focus of the analysis). Finally, in contrast to the SHORT TERM PERFORMANCE results, we do observe two significant coefficients within the control variables. Both UNION and COMPLEXITY have a large potential impact on PERFORMANCE CHANGE; while UNION is associated with a lower rate of PERFORMANCE CHANGE, COMPLEXITY is associated with a higher rate.

Table 4B extends the analysis with the instrumental variables approach. The instruments are the same as used in Table 3B (VERTINT-i,j, SUNK COST-i,j and LOW CAPACITY-i,j). While the estimates are a bit noisier, the overall magnitude and pattern of results is the same. VERTICAL INTEGRATION is associated with a large and significant improvement in PERFORMANCE CHANGE. While the pairwise correlation (4-5) is not significant, this is essentially driven by the fact that we have not included a linear time trend. Across a wide variety of specifications, there is a robust relationship, after inclusion of the linear time trend.

Overall, whether one employs OLS or instrumental variables, there is a quantitatively and statistically significant relationship between vertical integration and the level of performance change.

Overall Performance Results

Table 5 summarizes our key findings relating to the OVERALL PERFORMANCDE measure. Essentially, OVERALL PERFORMANCE has no robust relationship with vertical integration. In the absence of any controls, there is a small negative correlation (very weakly significant). However, with the inclusion of the time trend variable, the effect becomes insignificant. Indeed, when one controls for the economics and strategic environment, as in (5-3), the estimated coefficient on VERTICAL INTEGRATION is positive. Interestingly, when one implements instrumental variables in this context, one gets results similar to the "transactional misalignment" literature (Nickerson and Silverman, 2004). Specifically, when one focuses on that part of VERTICAL INTEGRATION which is driven by exogenous shocks to the adjustment costs associated with vertical integration or the returns to cross-system coordination, there is a negative relationship between OVERALL PERFORMANCE and VERTICAL INTEGRATION. In other words, when firms are "constrained" to be vertically integrated, the relationship between VERTICAL INTEGRATION and performance is negative. However, thinking about the overall effect estimated to be zero, the results in Figure 5 are consistent with an industry equilibrium in which supplier entry and internal investment are set so that the (expected) marginal returns to outsourcing equal the (expected) marginal returns to internal development.

Interaction Effects Summary

Motivated by the case studies, we examined several specific potential interaction effects, focusing both on SHORT TERM PERFORMANCE and PERFORMANCE CHANGE. UNION involvement at the manufacturer has no effect on initial performance, but completely eliminates the "benefits" of vertical integration as realized by PERFORMANCE CHANGE. The gains from vertical integration in terms of PERFORMANCE CHANGE are higher for systems which are incorporated into company-wide platform efforts. Firms with substantial sunk investments in a given system face no short-term performance loss from vertical integration and vertical integration is additionally associated with a much higher rate of performance improvement for firms with high sunk system-specific investments.

Summary of Findings

Under both OLS and instrumental variables treatments., SHORT TERM PERFORMANCE is declining in the level of VERTICAL INTEGRATION, and PERFORMANCE CHANGE is increasing in the level of VERTICAL INTEGRATION. There is no robust pattern of correlation between VERTICAL INTEGRATION and OVERALL PERFORMANCE. Moreover, the costs of vertical integration in terms of short-term performance are lower for those firms with substantial system-specific sunk investments. The benefits of vertical integration in terms of performance change is lower for those companies constrained by union activity, higher for systems which are incorporated into company-wide platforms, and higher for firms with substantial system-specific sunk investments. Moreover, it should be emphasized that these core findings are robust to alternative definitions of the underlying ratings and alternative definitions for SHORT TERM PERFORMANCE and PERFORMANCE CHANGE.

VII. Concluding Remarks

This paper examined the impact of vertical integration on the dynamics of performance in the context of automobile product development. Our key insight is that vertical integration will have differential impacts on different performance margins. Specifically, we looked at how vertical integration is related to the performance profile over the product lifecycle. On the one hand, outsourcing facilitates access to cutting-edge technology and the use of high-powered performance contracts. On the other hand, vertical integration allows firms to respond to adapt to unforeseen contingencies and customer feedback, maintain more balanced incentives over the product lifecycle, and develop firm-specific capabilities over time. Together, these effects suggest that outsourcing will be associated with higher levels of *initial* performance, while vertical integration will be associated with a higher rate of performance improvement over the product lifecycle.

A series of case examples reinforces each of these insights. While outsourcing yields benefits at the time of initial product introduction, vertical integration is associated with performance improvement over time. We then subjected these ideas to a more systematic empirical analysis. Our data combine detailed performance measures over time with nuanced measures of the extent of vertical integration, as well as measures of the contracting and technology environment. Using both OLS and an instrumental variables estimator, we establish four key results. First, initial performance is declining in the level of vertical integration. Second, the level of performance improvement is significantly increasing in the level of vertical integration. Moreover, even after controlling for other factors impacting performance, the magnitude of these two effects are roughly identical – there is no relationship between vertical integration and "overall" performance. Finally, taking advantage of outsourcing during the early part of the product lifecycle and internal development during the latter years of the lifecycle depends on the institutional and strategic environment. For example, the long-term benefits to vertical integration are erased for those firms with a strong union presence. Overall, the empirical findings highlight that vertical integration is associated with both costs *and* benefits and that different performance margins will reflect the tradeoffs associated alternative contracting modes

While the current analysis focuses on two specific but still not comprehensive set of performance margins, a more complete analysis should yield additional insight into the economic and strategic logic underlying the observed empirical patterns. Do the differences between the lifecycle patterns of outsourcing versus internal development reflect changes in competencies over time or changes in the contractibility and incentive environment over time? What is the role played by union involvement, platform requirements, and sunk investments in shaping the performance tradeoffs over the lifecycle? What are the implications of the analysis for dynamic theories of vertical integration and contracting over specific investments?

REFERENCES

Alexander, C. (1964), Notes on the Synthesis of Form, Harvard University Press, Cambridge.

Argyres, N.S. and J. Porter Liebeskind (1999), "Contractual Commitments, Bargaining Power, and Governance Inseparability: Incorporating History into Transaction Cost Theory", *Academy of Management Review*, v. 24: 49-63.

Bajari, Patrick and Tadelis, Steven (2001) "Incentives Versus Transaction Costs: A Theory of Procurement Contracts." RAND Journal of Economics, Autumn 32(3) pp. 387-407.

Baker, G. and T. Hubbard (2004), "Contractibility and Asset Ownership: On-Board Computers and Governance in U.S. Trucking," *Quarterly Journal of Economics*, forthcoming.

Baker, G., R. Gibbons and K. Murphy (2002), "Relational Contracts and the Theory of the Firm," *Quarterly Journal of Economics.*

Baldwin, C., and K.B. Clark (2000), Design Rules: Volume I, The Power of Modularity, MIT Press.

Bigelow, L. (2003). "Make-or-Buy Revisited: A Longitudinal, Population-wide Test of Transaction Cost Alignment," mimeo, Olin School of Management, Washington University.

Boerner, C., and J. Macher (2001), "Transaction Cost Economics: An Empirical Assessment and Review," working paper.

Clark, K. and T. Fujimoto, (1991). Product Development Performance, Strategy, Organization, and Management in the World Auto Industry, Harvard Business School Press, Boston.

Corts, Kenneth (2002) "Fixed-Price vs. Cost-Plus: The Determinants of Contractual Form in Offshore Drilling," mimeo, Harvard Business School

Corts, Kenneth and Jasjit Singh (2003) "The Effect of Relationships on Contract Choice: Evidence from Offshore Drilling," mimeo, Harvard Business School

Crocker, K. and Reynolds, K. (1993), "The Efficiency of Incomplete Contracts: An Empirical Analysis of Air Force Engine Procurement," *Rand Journal of Economics*, Vol. 24, pp. 126-46.

Fine, C. (1998). *Clockspeed: Winning Industry Control in the Age of Temporary Advantage*, Perseus Books, Reading, Massachusetts.

Fine, C. and D. Whitney (1996). "Is the Make versus Buy Decision A Core Competence?", MIT International Motor Vehicle Program Working Paper.

Fujimoto, T. (1997). "The Japanese Automobile Supplier System: Framework, Facts and Reinterpretation", Working Paper, Tokyo University.

Grossman, S. and O. Hart, 1986, "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration", *Journal of Political Economy*, 94: 691-719.

Hart, O. and J. Moore, 1990, "Property Rights and the Nature of the Firm", *Journal of Political Economy*, 98: 1119-1158.

Helper, S. (1995), "Supplier Relations and the Adoption of New Technology: Results of Survey Research in the U.S. Auto Industry," NBER Working Paper 5278.

Helper, Susan (1997), "Complementarity and Cost Reduction: Evidence from the Automobile Supply Industry," NBER Working Paper 6033.

Holmstrom, B. and P. Milgrom (1994), "The Firm as an Incentive System," *American Economic Review*, 84: 972-991.

Ichniowski, C., K. Shaw, and G. Prennushi (1997), "The Effects of Human Resource Management Practices on Productivity," *American Economic Review*, 87(3): 291-313.

Joskow, Paul, (1988). "Asset Specificity and the Structure of Vertical Relationships: Empirical Evidence", *Journal of Law, Economics and Organization*, vol. 4, pp. 95-117.

Klein, B., R. Crawford, and A. Alchian (1978). "Vertical Integration, Appropriable Rents and the Competitive Contracting Process," *Journal of Law and Economics* 21: 297-326.

Klein, Peter G. (2004) , “The Make-or-Buy Decision: Lessons from Empirical Studies,” Contracting and Organizations Research Institute Working Paper No. 2004-07.

Kogut, B., U. Zander (1992), “Knowledge of the Firm, Combinative Capabilities, and the Replication of Technology”. *Organization Science* 3(3) 383-397.

Kogut, B., U. Zander (1996), “What Firms Do? Coordination, Identity and Learning.” *Organization Science* 7(5) 502-518.

Leiblein, Michael J., J J. Reuer and F. E Dalsace (2002), “Do Make Or Buy Decisions Matter? The Influence Of Organizational Governance On Technological Performance” *Strategic Management Journal* 23: 817–833.

MacDuffie, J. (1995), “Human Resource Bundles and Manufacturing Performance: Organizational Logic and Flexible Production Systems in the World Auto Industry,” *Industrial and Labor Relations Review*, 48 (2).

Macher Jeffrey T. (2004). Technological Development and the Boundaries of the Firm: A Knowledge-Based Examination in Semiconductor Manufacturing. Georgetown University working paper.

Masten, S. (1984).“The Organization of Production: Evidence from the Aerospace Industry,” *Journal of Law and Economics*, 27: 403-417.

Masten, S., J. Meehan and E. Snyder (1989), “Vertical Integration in the U.S. Auto Industry,” *Journal of Economic Behavior and Organization*, 12: 265-273.

Masten, S., J. Meehan and E. Snyder (1991) “The Costs of Organization”, *Journal of Economic Behavior and Organization*, 7: 1-25.

McAfee, R.P. and McMillan, J. (1986), “Bidding for Contracts: A Principal Agent Analysis” Rand Journal of Economics, Vol. 17, pp. 326-38.

Monteverde, K. and D. Teece (1982). “Supplier Switching Costs and Vertical Integration in the Automobile Industry", *Bell Journal of Economics*, 13: 206-213.

Nickerson, J. (1997). *Toward an Economizing Theory of Strategy: The Choice of Strategic Position, Assets, and Organizational Form*, Dissertation, UC-Berkeley.

Nickerson, J. A., B. S. Silverman. (2004), “Why Firms Want to Organize Efficiently and What Keeps Them from Doing So: Inappropriate Governance, Performance, and Adaptation in a Deregulated Industry” Administration Science Quarterly.(forthcoming) 1-49.

Nickerson, J. and T. Zenger (2005) "A Knowledge-based Theory of Governance Choice," Organization Science, forthcoming

Novak, S. and S. Eppinger (2001), “Sourcing by Design: Product Complexity and the Supply Chain,” *Management Science*

Poppo, L., T. Zenger (1998), “Testing Alternative Theories of the Firm: Transaction Cost, Knowledge-Based, and Measurement Explanations for Make-or-Buy Decisions in Information Services. *Strategic Management Journal* 19 853-877.

Sampson, R. C. (2000). “R&D Alliances & Firm Performance: The Impact of Technological Diversity and Alliance Organization on Innovation.” *Stern School of Business Working Paper*.

Saussier, S. (2000). “Transaction Costs and Contractual Incompleteness: The Case of Electricite de France.” *Journal of Economic Behavior and Organization* **42**: 189-206.

Suh, N. (1990). *The Principles of Design*, Oxford University Press, New York.

Suh, N. (1999). “A Theory of Complexity, Periodicity, and the Design Axioms”, *Research in Engineering Design*, vol. 11, pp. 116-131

Ulrich, K. and D. Ellison (1998). “Beyond Make-Buy: Internalization and Integration of Design and Production", Wharton School Working Paper.

Ulrich, K. and D. Ellison (2003), “Customer Requirements and the Design-Select Decision", *Management Science*.

Ulrich, K., and S. D. Eppinger (1995). *Product Design and Development,* McGraw-Hill, New York.

Ulrich, K., et al (1993). "Including the Value of Time in Design-for-Manufacturing Decision Making", *Management Science*, 39(4): 429-447.

Whinston, M. D., (1997). "On the Transaction Cost Determinants of Vertical Integration", Columbia Law School Make Versus Buy Conference.

Williamson, O. E. (1979), "Transaction-Cost Economics: The Governance of Contract Relations", *Journal of Law, Economics and Organization*, 22: 3- 61.

Williamson, O.E. (1975) Markets and Hierarchies, New York: Free Press.

Williamson, O.E., (1985) The Economic Institutions of Capitalism, New York: Free Press.

FIGURE 1
Timing of Procurement and Ex-Post Adaptation

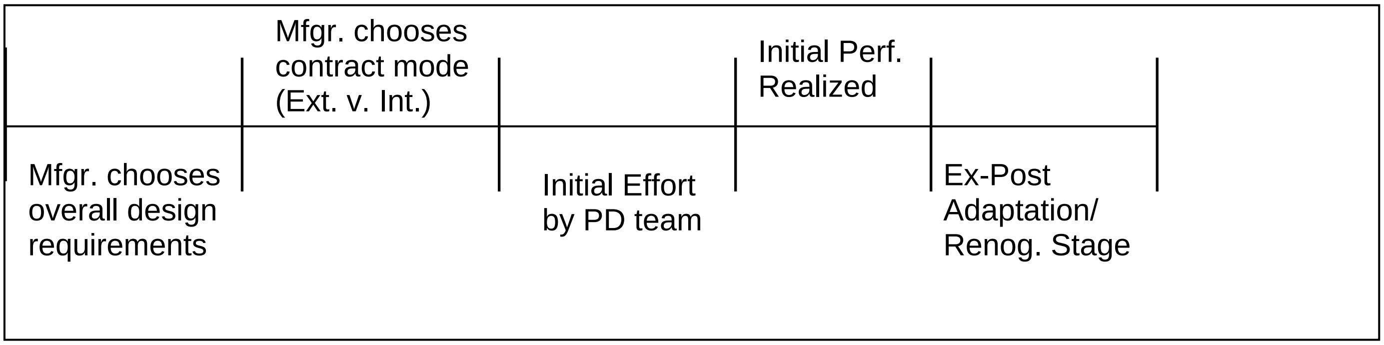
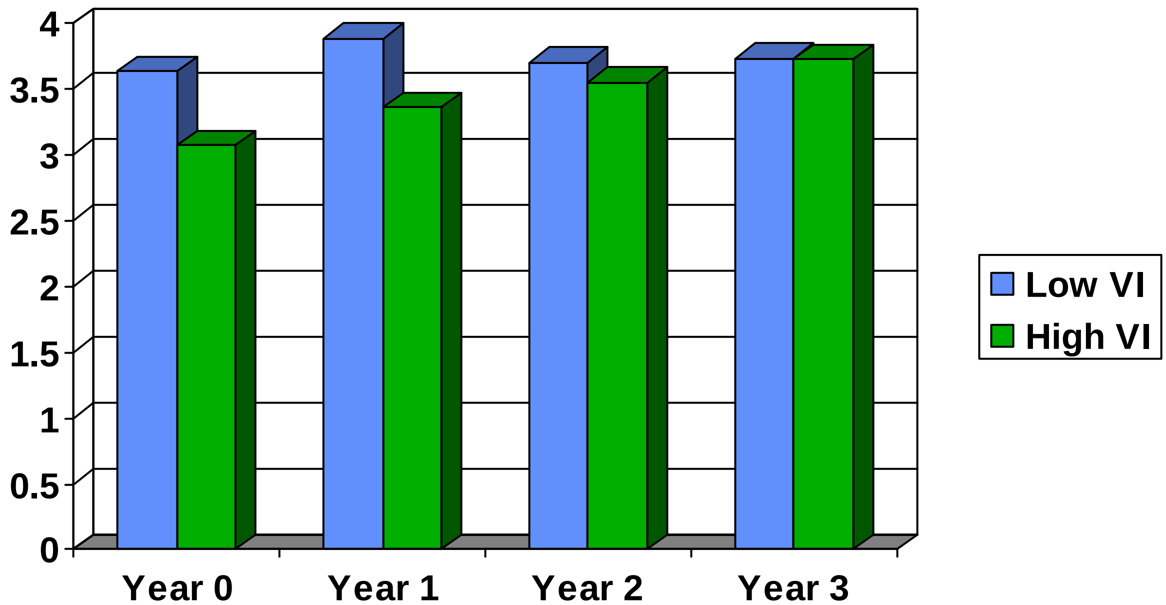


FIGURE 2
The Costs and Benefits of Outsourcing:
Synthesizing the Case Evidence

	Vertical Integration	External Sourcing
Ex Ante Contracting Opportunities	• Deep vehicle- specific knowledge base • Less knowledge of system-specific technology • Difficult to enforce specific performance criterion	•Global supply opportunities •Opportunity for well-defined performance contracts
Ex Post Renegotiation Outcomes	• Continuing authority relationship allows for redirection • Potential for learning	• Hard to enforce contracts after key requirements have been met Fewer continuing relationships

FIGURE 3
Average Performance Rating By High or Low Vertical Integration, by Years Since "Major" Model Introduction



Years Since "Major" Model Introduction

TABLE 1
Variables & Definitions

VARIABLE	DEFINITION	MEAN	STD. DEV.
	PERFORMANCE MEASURES		
PERFORMANCE RATING	*Consumer Reports* rating (from 1-5, w/ 5 as the "highest rating") for system *i* on model *j* in year *t* after a major model change. The rating that is chosen for a given model-year is the "first" *CR* rating available for that model-year	3.541	.970
SHORT TERM PERFORMANCE	Average of PERFORMANCE RATING$_{ij0}$ and PERFORMANCE RATING$_{ij1}$, as available. When only one *CR* rating is available, only one is used.	3.432	.968
LONG TERM PERFORMANCE	Average of PERFORMANCE RATING$_{ij2}$ and PERFORMANCE RATING$_{ij3}$, as available. When only one *CR* rating is available, only one is used.	3.705	.967
PERFORMANCE CHANGE	LONG TERM PERFORMANCE – SHORT TERM PERFORMANCE	0.286	.832
OVERALL PERFORMANCE$_i$	Average of SHORT TERM PERFORMANCE and LONG TERM PERFORMANCE	3.560	.969
	CONTRACTING MEASURES		
VERTICAL INTEGRATION	Percentage of the system produced in house between 0 and 1 (1 indicates all in-house production)	.513	.318
VERTICAL INTEGRATION$_{-i}$	Sum of VERTICAL INTEGRATION for all systems excepting *i* on model j	3.076	1.468
	SYSTEM-SPECIFIC CONTRACTING ENVIRONMENT MEASURES		
SUNK COST	Dummy = 1 if pre-existing in-house sunk costs and/or plant investment for system *i*	.143	.351
LOW CAPACITY	Dummy = 1 if plant has insufficient capacity to manufacture system design in-house	.170	.377
PLATFORM	Dummy = 1 the component was designed to be used for more than one vehicle model	.527	.502
COMPLEXITY	Degree of System Complexity, ranging from 0 to 1 (See Novak and Eppinger, 2001).	.392	.275
PERFORMANCE	Measure for desired performance goals at the system level, ranging from 0 (low) to 1 (high)	.457	.311
SKILL SHORTAGE	Dummy = 1 if key worker skills are missing in existing plant locations	.161	.369
	MODEL-YEAR MEASURES		
UNION	Dummy = 1if a component has been produced in-house and is covered under union agreement	.464	.501

TABLE 2
Performance Rating Margins
By High or Low Vertical Integration

VERTICAL INTEGRATION	SHORT TERM PERFORMANCE	LONG TERM PERFORMANCE	PERFORMANCE CHANGE
“Below” Median	3.71	3.73	0.02
“Above” Median	3.24	3.68	0.44

Median VERTICAL INTEGRATION = 0.50

TABLE 3A
Short-Term Performance:
OLS Estimation

DEPENDENT VARIABLE = SHORT TERM PERFORMANCE N = 114			
	(3-1)	**(3-2)**	**(3-3)**
VERTICAL INTEGRATION	-0.850*** (0.279)	-0.715*** (0.259)	-0.510*** (0.296)
UNION			-.308 (.219)
COMPLEXITY			.302 (.362)
PERFORMANCE			-.569 (.417)
PLATFORM			.018 (.171)
SKILL SHORTAGE			-.052 (.426)
SUNK COST			-.281 (.297)
LOW CAPACITY			-.016 (.287)
SYSTEM DUMMY VARIABLES		Insignificant	Insignificant
YEAR		0.088*** (.014)	0.075*** (.016)
R^2	.079	.416	.447

Table 3B
Short-Term Performance:
IV Results

DEPENDENT VARIABLE = SHORT TERM PERFORMANCE N = 114			
	(3-4)	**(3-5)**	**(3-6)**
VERTICAL INTEGRATION	-1.961***	-0.922***	-2.507**
	(0.443)	(0.332)	(1.101)
COMPLEXITY			.532
			(.427)
PERFORMANCE			-.613
			(.475)
PLATFORM			.187
			(.213)
SKILL SHORTAGE			.021
			(.498)
SUNK COST			.030
			(.357)
LOW CAPACITY			-.364
			(.471)
SYSTEM DUMMY VARIABLES		**Significant**	**Significant**
COMPANY DUMMY VARIABLES			Insignificant
YEAR		0.085***	0.075***
		(.015)	(.016)

*Instrumental Variables for VERTICAL INTEGRATION = VERTICAL INTEGRATION-i, SUNK COST-i,, LOW CAPACITY-i

TABLE 4A
Performance Change:
OLS Results

Dependent Variable : PERFORMANCE CHANGE			
	(4-1)	**(4-2)**	**(4-3)**
VERTICAL INTEGRATION	**0.611*****	**0.708*****	**0.828*****
	(0.246)	**(0.274)**	**(0.294)**
UNION		-.663***	-.608***
		(.176)	(.154)
COMPLEXITY		.862***	.412*
		(.229)	(.238)
PERFORMANCE		-.050	-.090
		(.287)	(.315)
PLATFORM		.064	-.099
		(.120)	(.120)
SKILL SHORTAGE		-.326	-.593
		(.238)	(.267)
SUNK COST		.152	.392
		(.209)	(.245)
LOW CAPACITY		.002	.100
		(.174)	(.179)
SHORT TERM PERFORMANCE		-.546***	-.697***
		(.072)	(.073)
SYSTEM DUMMY VARIABLES			Insignificant
YEAR			.078***
			(.015)
R^2	.055	.424	.583

TABLE 4B
Performance Change:
IV Results

	(4-5)	(4-6)	(4-7)
VERTICAL INTEGRATION	0.569 (0.397)	**0.790*** **(0.418)**	**0.979***** **(0.390)**
UNION		-.597*** (.184)	-.645*** (.181)
COMPLEXITY		.404** (.215)	.400* (.241)
PERFORMANCE		-.068 (.234)	-.067 (.312)
PLATFORM		-.086 (.104)	-.089 (.126)
SKILL SHORTAGE		-.465** (.214)	-.393 (.264)
SUNK COST		.345* (.206)	.393 (.245)
LOW CAPACITY		.080 (.158)	.124 (.177)
SHORT TERM PERFORMANCE		-.714*** (.067)	-.690*** (.072)
SYSTEM DUMMY VARIABLES			Insignificant
YEAR		0.081*** (.015)	0.081*** (.015)

*Instrumental Variables for VERTICAL INTEGRATION = VERTICAL INTEGRATION-i, SUNK COST-i, LOW CAPACITY-i

TABLE 5
Overall Performance:
OLS Results

Dependent Variable : OVERALL PERFORMANCE			
	(5-1)	(5-2)	(5-3)
VERTICAL INTEGRATION	-0.488*	-0.256	0.116
	(0.27`)	(0.238)	(0.264)
UNION			-.514***
			(.161)
COMPLEXITY			.386
			(.266)
PERFORMANCE			-.404
			(.261)
PLATFORM			-.050
			(.129)
SKILL SHORTAGE			-.332
			(.329)
SUNK COST			-.021
			(.255)
LOW CAPACITY			.050
			(.229)
SYSTEM DUMMY VARIABLES		**Significant**	Insignificant
YEAR		0.108***	0.090***
		(.012)	(.014)
R^2	.032	.543	.612

TABLE 6
INTERACTION EFFECTS:
UNION

	SHORT TERM PERF	PERFORMANCE CHANGE
VERTICAL INTEGRATION	-0 802**	1.438***
	(0.413)	(0.366)
VI *UNION	0.621	-1.245
	(0.513)	(0.431)
UNION	-0.644*	0.070
	(0.348)	(0.303)
System-Specific Controls	Included	Included
System Fixed Effects	Included	Included
Time Trend	Included	Included

TABLE 7
INTERACTION EFFECTS: PLATFORM

	SHORT TERM PERF	PERFORMANCE CHANGE
VERTICAL INTEGRATION	-0 590*	0.569**
	(0.311)	(0.286)
VI *PLATFORM	0.208	0.637*
	(0.495)	(0.378)
PLATFORM	-0.117	-0.406*
	(0.304)	(0.231)
System-Specific Controls	Included	Included
System Fixed Effects	Included	Included
Time Trend	Included	Included

TABLE 8
INTERACTION EFFECTS:
SUNK COSTS

	SHORT TERM PERF	PERFORMANCE CHANGE
VERTICAL INTEGRATION	-0 557*	0.775***
	(0.297)	(0.304)
VI *SUNK COSTS	1.453***	1.250*
	(0.637)	(0.703)
SUNK COSTS	-1.004*	-0.234
	(0.516)	(0.426)
System-Specific Controls	Included	Included
System Fixed Effects	Included	Included
Time Trend	Included	Included