

NBER WORKING PAPER SERIES

HOW DOES OUTSOURCING AFFECT PERFORMANCE DYNAMICS? EVIDENCE
FROM THE AUTOMOBILE INDUSTRY

Sharon Novak
Scott Stern

Working Paper 13235
<http://www.nber.org/papers/w13235>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2007

Mercedes Delgado provided outstanding research assistance. We thank Joshua Gans, Tom Hubbard, Mara Lederman, two anonymous referees, and seminar participants at various universities and conferences for helpful comments and suggestions. This research was partially funded by MIT Center for Innovation and Product Development and the International Motor Vehicle Program at MIT. All errors or omissions remain our own. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

© 2007 by Sharon Novak and Scott Stern. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

How Does Outsourcing Affect Performance Dynamics? Evidence from the Automobile Industry
Sharon Novak and Scott Stern
NBER Working Paper No. 13235
July 2007
JEL No. L24,L25,L62,O32

ABSTRACT

This paper examines the impact of vertical integration on the dynamics of performance over the automobile product development lifecycle. Building on recent work in organizational economics and strategy, we evaluate the relationship between vertical integration and different performance margins. Outsourcing facilitates access to cutting-edge technology and the use of high-powered performance contracts. Vertical integration allows firms to adapt to unforeseen contingencies and customer feedback, maintain more balanced incentives over the lifecycle, and develop firm-specific capabilities over time. Together, these effects highlight a crucial tradeoff: while outsourcing is associated with higher levels of initial performance, vertical integration will be associated with performance improvement over the product lifecycle. We test these ideas using detailed data from the luxury automobile segment, establishing three 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. Finally, the impact of vertical integration on alternative performance margins is mediated by the level of pre-existing capabilities, by the salience of opportunities to access external technology leaders, and by the scope for learning over the product lifecycle. Together, the findings highlight a strategic governance tradeoff between short-term performance and the evolution of firm capabilities.

Sharon Novak
Policy Department
D505
UCLA Anderson School of Management
Box 951481
Los Angeles, CA 90095
snovak@anderson.ucla.edu

Scott Stern
Management and Strategy Department
Kellogg School of Management
2001 Sheridan Road
Evanston, IL 60208
and NBER
s-stern2@northwestern.edu

I. Introduction

Recent work in strategy and economics has made considerable theoretical and empirical progress in explaining the determinants of vertical integration and firm boundaries. From a theoretical perspective, transaction cost theory (building on Williamson), the property rights approach (Hart and Moore, 1990; Baker, Gibbons and Murphy, 2002) and the knowledge-based view of the firm (Kogut and Zander, 1992, 1996; Grant, 1996; Conner and Prahalad, 1996; Nickerson and Zenger, 2004) have each developed rapidly over the last decade. Partly in response to this emerging theoretical maturity, recent attention has turned towards evaluating the empirical content of individual theories, and synthesizing relationships among them (among others, *see* Boerner and Macher, 2005). Empirical research has focused attention on the tradeoffs driving outsourcing choices, and the impact of environmental shifts on make-versus-buy decisions (Whinston, 2001; Baker and Hubbard, 2004; Nickerson and Silverman, 2003).

While a large body of research focuses on the choice between vertical integration versus outsourcing, only a small body of research explores the performance *consequences* of vertical integration (Masten, et al, 1991; Masten and Saussier, 2002; Nickerson and Silverman, 2003; Boerner and Macher, 2005; Yvrande-Billon and Saussier, 2005). From the perspective of strategy research, understanding the performance implications of vertical integration is crucial. Firm boundaries are often a long-lived strategic commitment, and a boundary choice made at one point in time will have performance and strategy consequences over time. For example, outsourcing a function constrains a firm's ability to adapt to unforeseen contingencies or benefit from learning and the experience associated with that function.

This paper proposes and implements a novel approach for evaluating the relationship between performance and vertical integration. Our analysis builds on the insight that a single vertical integration choice affects multiple performance *margins*. Theoretical work in economics, strategy, and organizations suggests that a key difference between vertical integration and outsourcing arises from the difference between what is achievable through specific formal contracts and what is achievable through maintaining activities within the boundaries of a single organization. As emphasized by Williamson (1985) or, more recently, Bajari and Tadelis (2001), the benefits from outsourcing depend on the ability to specify detailed formal contracts and offer contract-based incentives, while the benefit to vertical

integration results from the ability to adapt to unforeseen contingencies through a hierarchy. In the product development context, contract-based incentives provide high-powered incentives during the early stages of the procurement process, while the advantages to hierarchy will be realized as unforeseen developments arise over time. Similarly, the knowledge-based view (KBV) of governance choice emphasizes that, while outsourcing allows firms to rapidly access capabilities that are not currently maintained within the firm, vertical integration is a prerequisite for internal capability and knowledge development over time (Kogut and Zander, 1992, 1996; Grant, 1996; Conner and Prahalad, 1996, Nickerson and Zenger, 2004). Though distinct, both of these theoretical perspectives imply the key dynamic tradeoff at the heart of this paper: while product development outsourcing facilitates high-powered incentive contracts and the ability to access cutting-edge suppliers, vertical integration enhances the ability to achieve a higher level of flexibility and learning.

Evaluating the linkage between vertical integration choices and performance is subtle, since alternative contracting modes are endogenous. 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 adopting an integrated structure, and the returns to outsourcing are equally high for those firms adopting outsourcing, cross-sectional performance comparisons need not find any performance consequence to integration, even though decision makers face a clear performance tradeoff.

This challenge has motivated a small, growing literature on the performance implications of vertical integration. These studies focus on “transactional misalignment” –performance losses associated with adopting (or inheriting) an organizational form which is “inappropriate” for a given economic or strategic environment (Masten, et al, 1991; and Yvrande-Billon and Saussier, 2005 for an assessment). As well, building on work such as Poppo and Zenger (1998), Macher (2006) provides evidence from the semiconductor industry that the conditions supporting high performance differ depending on whether a firm is vertically specialized or vertically integrated. In this paper, we focus on the difference between the static and dynamic performance implications of vertical integration. As the modern theory of the firm

is premised on the idea that organizational choices have *both* costs and benefits, our analysis traces out the (static and dynamic) costs and benefits associated with vertical integration versus outsourcing.

We exploit the process of procurement and product development contracting accompanying a “major” model change in the automobile industry. 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. While broad positioning choices involve high levels of coordination, a major model change involves hundreds of individual governance choices, impacting several different “systems” within an automobile (e.g., brake, engine, body, etc). Individual managers choose a governance mode for the duration of the major model change, and changes in governance mode during the product lifecycle are rare.

This setting allows us to explore how vertical integration impacts performance over the lifecycle. During the initial product development and sourcing stage, outsourcing facilitates the specification of detailed high-powered contracts, and the ability to access and contract on a global basis for frontier technology. Conversely, internal development limits the ability to take advantage of frontier technology, and internal wage contracts will offer only muted incentives to reach specific initial performance targets. After a product is launched, however, product development involves 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, external suppliers may have limited incentives to innovate. Because the precise nature of adaptation cannot be anticipated (and the manufacturer cannot guarantee a precise volume of “work,” since the extent of adaptation is also uncertain), these terms cannot be precisely articulated in the initial performance contract. Moreover, in many cases, the capabilities required for improvement rely on detailed firm-specific knowledge (e.g., effective improvement may require extensive coordination), or may rely on idiosyncratic knowledge about the precise technical characteristics of a given automobile. 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. Consequently, while outsourcing

facilitates high performance in the early stages of the product lifecycle, vertical integration will accrue a dynamic benefit as the firm's performance increases over the product lifecycle.

Our empirical analysis exploits a 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.). 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*. Since competitive advantage is closely tied to quality in the luxury automobile segment, these measures are a useful proxy for overall vehicle commercial success (relative to alternatives for that automobile model). Finally, for each model-system, we observe a set of system-specific vertical integration drivers. For example, we observe whether the firm has existing in-house sunk investments in plant and equipment for each system. Together, these data allow for a detailed examination of the relationship between vertical integration and performance over the product lifecycle.

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, 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. These basic patterns in the data are robust across a wide range of specifications, including instrumental variables estimates that account for the endogeneity of the degree of vertical integration. Further, we are able to examine whether the impact of vertical integration on performance is mediated by the level of pre-existing capabilities, by the salience of opportunities to access external technology leaders, and by the scope for learning over the product lifecycle. Among other findings, we find that both the benefits and costs of vertical integration were particularly high for Japanese manufacturers, who were entrants into the luxury automobile segment.

II. The Implications of Vertical Organization for Performance over the Product Lifecycle

While a voluminous literature focuses on the drivers of vertical integration and contracting terms, a small literature focuses on the relationship between vertical integration and performance (Masten, et al,

1991; Masten, 1993; Silverman, et al, 1997; Nickerson and Silverman, 2003; Boerner and Macher, 2005; Yvrande-Billon and Saussier, 2005).¹ Though performance studies are central to evaluating the role of vertical integration for strategy, such studies must address a fundamental inference problem. Since governance is chosen in response to a firm's economic, strategic and organizational environment, differences among firms in their governance choices likely reflect idiosyncratic differences in the benefits and costs of vertical integration (Masten and Saussier, 2002; Yvrande-Billon and Saussier, 2005). Estimating the overall "impact" of vertical integration by comparing the performance of firms that choose "make" rather than "buy" confounds the impact of vertical integration with differences in the (expected) returns to vertical integration among organizations that choose one governance mode over another.²

While cross-sectional comparisons of vertical integration and performance may be misleading, individual decision-makers nonetheless face significant performance tradeoffs when choosing a governance mode. Within a given setting, alternative governance structures will be associated with particular benefits and cost profiles, and the decision to vertically integrate (or not) depends on comparing the *net* benefits and costs after accounting for multiple performance margins. As a result, the drivers of high performance can differ significantly under vertical integration and outsourcing (Poppo and Zenger, 1998; Macher, 2006). While the decision to vertically integrate can be summarized in terms of net benefits, the performance implications of vertical integration will be realized along specific margins.

Studying the relationship between governance choice and alternative performance "margins" is only useful to the extent that the performance dimensions under study reflect specific and crucial strategic tradeoffs. We focus on the impact of vertical integration on short-term versus long-term performance over the product lifecycle. From a theoretical perspective, the tradeoff between short-term and long-term performance is rooted in leading theoretical approaches to vertical integration, such as contract theory and

¹ Literature reviews include Shelanski and Klein (1995), Masten and Saussier (2002), Boerner and Macher (2005), Klein (2004), and Yvrande-Billon and Saussier, (2005). The approach here complements studies of the relationship between contractual incompleteness and contract design (Joskow, 1988; Crocker and Masten, 1991; Crocker and Reynolds, 1993; Saussier, 2000) as well as recent studies of the impact of performance margins (Poppo and Zenger, 1998; Macher, 2006; Forbes and Lederman, 2006).

² The most well-developed approach to overcoming this bias is to focus on the impact of transactional "misalignment," documenting the performance consequences of choosing an organizational form inconsistent with observable aspects of the environment (Masten, et al, 1991; Masten, 1993; Silverman, et al, 1997; Mayer, 2000; Leiblein, et al, 2002; Menard and Saussier, 2002; Bigelow, 2003; Nickerson and Silverman, 2003; Yvrande-Billon and Menard, 2003; Sampson, 2004; and Mayer and Nickerson, 2005).

transaction cost economics (TCE) as well as the knowledge-based view of the firm (KBV). As such, we consider these theoretical perspectives in detail and examine whether the key assumptions underlying these arguments are likely satisfied in our empirical setting of luxury automobile product development.

Consider the dynamic performance implications of transaction cost economics. As pioneered by Williamson (1985), transaction cost economics (and related work in contract theory) emphasizes three crucial adaptation advantages to vertical integration: authority relationships through hierarchy, the ability to coordinate multiple activities across an organization, and the ability to provide promotion-based incentives for non-contractible dimensions of performance. As “hierarchy is its own court of ultimate appeal” (Williamson, 1991, p. 274), vertical integration allows firms to internalize the benefits and costs of coordination across different actors, and offer long-term incentives (such as promotion or the allocation of discretionary budgets) that are based on non-contractible (and often purely subjective) evaluations of the degree of compliance to the demands of hierarchy or non-contractible contributions to firm-specific performance. Employees bound by employment relationships extending beyond the life of any single project can be provided with incentives that extend beyond the contract governing that project. Conversely, transaction cost economics emphasizes that markets offer the potential to implement high-powered incentives for “contractible” dimensions of performance, and allow vertically disintegrated units to adapt autonomously to changes in the overall economic and strategic environment.³

Building on these insights, Bajari and Tadelis (2001, hereafter BT) develop a model of procurement contracts in which market-based relationships are governed by the details of (costly) contract design, while vertical integration allows an organization to maintain control through authority relationships (or relational mechanisms) with its employees. BT integrate insights from the transaction costs approach and contract theory by drawing out the implications of the fact that firm boundaries matter because it is impossible (or too costly) to contractually specify all potential contingencies.⁴ A key contribution of BT is to highlight how endogenizing the degree of “contractual incompleteness” (i.e., the degree of vertical integration) impacts contract design and specific dimensions of organizational

³ This discussion does not explicitly distinguish the specific advantages and potential costs of hybrid forms such as alliances at a theoretical level; we address the role of hybrid forms as we adapt our theoretical discussion to the context of automobile product development, and in our empirical work.

⁴ Building on Dye (1985), a recent theoretical literature focuses on the implications of contracting costs on contract design (see, among others, Anderlini and Felli, 1999; Battigalli and Maggi, 2004).

performance. Firms face a tradeoff between high-powered performance incentives in order to guarantee performance over the limited and upfront contingencies covered by formal contracts, versus the benefits that come from being able to exercise authority to adapt to unforeseen or unspecified contingencies as they arise. This tradeoff implies a fundamental difference between outsourcing and vertical integration in terms of performance dynamics: outsourced projects are likely to be associated with a high level of initial performance (in order to satisfy the terms of the original contract), while internal projects are likely to be associated with increasing relative performance over time (since these organizational forms allow for adaptation in the face of unforeseen contingencies). Thus, a single vertical integration (or procurement) decision impacts both a short-term and long-term performance margin, and the choice between vertical integration and outsourcing depends on the relative importance of these distinct dimensions.

The KBV provides a second theoretical perspective on the impact of firm boundaries on firm performance (Kogut and Zander, 1992, 1996; Grant, 1996; Conner and Prahalad, 1996, Nickerson and Zenger, 2004).⁵ While research within the KBV offers multiple (and sometime competing) perspectives on the relationship between firm boundaries and learning (*see* Nickerson and Zenger, 2004, for a thoughtful synthesis), this literature offers several insights into the relationship between vertical integration and performance over the product lifecycle. First, the type of knowledge an organization accumulates over time depends on both the knowledge generated in the context of operational problem-solving and the mechanisms employed to apply knowledge over time. To the extent that organizations differ in terms of their capacity to generate and apply knowledge, vertical integration choices will impact the nature of organizational learning and knowledge accumulation. Whereas vertical integration encourages the application of firm-specific knowledge, outsourcing facilitates the exploitation of application-specific knowledge (Grant, 1996). Thus, the knowledge and capabilities brought to bear on a given problem will differ depending on whether the locus of problem-solving activity is internal or external to the firm. Governance also impacts the evolution of knowledge and the nature of learning. While vertical integration encourages the development and accumulation of knowledge to be used in

⁵ We do not attempt to summarize the disparate strands of this vast literature here, and focus on the core ideas motivating hypotheses about the impact of governance on specific performance margins.

future firm-specific activities, outsourcing encourages the development and accumulation of knowledge with application in follow-on contracts with *other* downstream firms (Nickerson and Zenger, 2004).

This reasoning has implications for the impact of alternative governance modes on performance over the product lifecycle. In the early stages of the product development process, initial problem-solving requirements will benefit from the ability to access external knowledge sources and will involve modular problem-solving tasks. When choosing 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 initial capability levels of internal teams, each bidder is a specialist, vying with the others to achieve “best in class” by taking advantage of learning across multiple prior projects. From a dynamic perspective, however, project-specific learning is likely higher under vertical integration. Internal employees have higher incentives for project-specific learning, and are more likely to make that knowledge available for others to use throughout the organization. Moreover, as a project evolves, problem-solving is likely more interdependent. Internal problem-solving efforts may facilitate coordination and management of complexity (Nickerson and Zenger, 2004). Thus, similar to TCE, the KBV posits a dynamic tradeoff: outsourcing allows firms to rapidly access knowledge that the firm lacks from an *ex ante* perspective, but vertical integration is a prerequisite for the development of internal capabilities and learning over time.

These insights motivate our main hypotheses. For product development contracting (where governance choices involves contractual incompleteness and the potential for learning), both economic and organizational theory suggest a dynamic tradeoff between vertical integration and performance:⁶

Hypothesis 1: Initial performance is lower for higher levels of vertical integration.

Hypothesis 2: Performance improvement is higher for higher levels of vertical integration.

The Impact of Governance on Performance over the Automobile Product Development Lifecycle

The remainder of this paper evaluates these ideas in the context of governance choices in luxury automobile product development (Figure 1 and Appendix A present an overview of this process). While automobiles receive annual incremental upgrades, an automobile model undergoes a “major” model

⁶ If the size of one of these performance margins was significantly larger than the other, there might be a net positive return to a particular contracting model. However, when we observe significant variation in contracting practices, as in the automobile industry, we do not expect a systematic relationship between governance choices and measures of “overall” performance over the entire product lifecycle.

change approximately every five years. This “major” change provides an opportunity to significantly alter product positioning, technologies, and contracting choices for an automobile model, and initiates the beginning of the “lifecycle” for a major change. While a manufacturer is constrained by the history of the vehicle (sunk investments, etc.), the process underlying a major change is substantial, and allows us to evaluate the impact of vertical integration choices at different stages of the product development lifecycle.

The “Early” Years in the Lifecycle. During the earliest stages of automobile product development, 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, 2007). Two benefits seem to arise from outsourcing 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, within the luxury automobile segment, external contracting is most often realized through a competitive bidding process, with an explicit emphasis (and bidding criteria) related to the ability of suppliers to provide “globally innovative” technology or capabilities. In most cases, internal teams are less experienced and have lower capabilities than a “best in class” supplier; this is not surprising as external suppliers take advantage of learning across multiple projects within their area of expertise. Second, the incentive systems associated with outsourcing may favor short-term performance margins. Automobile contracts (endogenously) specify detailed contract provisions and milestone payments, resulting in high-powered incentives for specific performance requirements relating to objective technical specifications and cost objectives prior to 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 the potential for promotion do provide incentives, it is extremely rare that internal teams face the type of high-powered incentive schemes that characterize an outsourcing relationship. Indeed, such schemes are unlikely to be optimal, given the learning and coordination that firms encourage in internal development.

The “Late” Years in the Lifecycle. The realization of initial performance (associated with the introductory model year) motivates incremental innovation and improvement over the lifecycle of the

major. Consistent with our earlier analysis, both incentive and learning effects suggest advantages to vertical integration during this process. While internal teams are provided a constant level of incentives, are learning during the project, and can be directed through authority relationships, external suppliers have relatively few incentives for further effort. By and large, the enforceable terms of the contract (in terms of technology specification and quality) are satisfied by the time that the initial major model is introduced. While prospects for future contracts and general reputation do provide important incentives for continuing effort, the lack of a direct authority relationship or reliance on subjective internal incentive schemes mutes formal incentives for ongoing quality improvements. This effect is reinforced by the organization of supplier activities: while a post-contract change may require ten engineers who worked on the original project, supplier employees are allocated to new projects shortly after the project satisfies specified formal requirements. The inability to access personnel is reinforced by the contracting provisions surrounding ex post adaptation: to maintain secrecy and avoid expropriation, contracts limit the extent to which employees can be pulled from current projects to return to an earlier one (e.g., most contracts include provisions restricting the freedom of suppliers to use shared physical facilities across projects). In contrast, incentives can be provided internally to undertake activities which are formally “non-contractible” (e.g., through promotion incentives), and the firm can simply use its authority relationship to adjust the level of effort and the composition of personnel to address unforeseen contingencies, coordinate interactions across multiple systems, and respond to the specifics of consumer feedback. Learning and knowledge considerations also tilt towards vertical integration. In contrast to the performance drivers in the early stages, a very different type of knowledge characterizes the ability to improve performance over the lifecycle. While access to global technology was important in the initial development stage, performance improvement after product launch requires detailed model-specific knowledge. While an external supplier may be “best in class” using objective criteria, internal teams will have idiosyncratic capabilities and knowledge relating to the needs of a particular automobile model (and will have strong incentives to develop and systematize that knowledge). Furthermore, suppliers may have too few incentives to share information with the downstream firm in case this causes them to lose their advantage in contracting. This generation and application of firm-specific knowledge facilitates performance improvements over the lifecycle.

The Limits to Hierarchy, the Limits to Contracting.

Hypothesis 1 and 2 suggest that since total profits depend on the success of the vehicle across the entire product lifecycle, luxury automakers face a dynamic governance tradeoff: while outsourcing may yield a higher level of “initial” performance, vertical integration facilitates improvements over the product lifecycle. Of course, this tradeoff only arises to the extent that hierarchy and vertical integration cannot replicate the high-powered contract-based incentives associated with outsourcing, and contracting relationships cannot replicate the adaptation and coordination advantages associated with hierarchy (Hart, 1988). We examine the limits to hierarchy and the limits to contracting in this empirical setting, and highlight potential mechanisms that may mitigate the impact of governance choices on performance.

The Limits to Contracting on Performance. Outsourcing contracts primarily focus on technical and cost objectives that are observable prior to market introduction, rather than incentive payments related to specific performance margins over the product lifecycle.⁷ This is not particularly surprising, since the specification and enforcement of performance-based contracts is quite costly. While technical milestones can be tailored to individual development contracts (e.g., specifying a particular tolerance for a particular part), the performance of an overall system depends on the interaction and coordination of multiple components, and so it is difficult to assign responsibility for performance to individual contractors (or internal employees, for that matter).⁸ Moreover, performance benchmarks are inherently noisy (and subjective); risk-averse product development personnel will seek to avoid a contract heavily dependent on such a measure. Thus, a key limit on contracting in our empirical setting is the inability to specify a contract based primarily on observable short-term and long-term performance measures.

The Limits to Relational Contracting. The power of hierarchy arises in part from the ability to implement relational contracts within an organization, and offer incentives (such as promotion) based on subjective evaluation criteria (Baker, Gibbons and Murphy, 1994). At the same time, relational contracts also operate *across firms*, and mediate the impact of contracts (Poppo and Zenger, 2002). If relational

⁷ Suppliers benefit from increased parts volume for successful vehicles; however, as a given supplier has a negligible impact on overall sales, sales-based incentives have a very limited impact on each supplier.

⁸ Some contracts do specify penalties if certain performance thresholds are not met; however, these provisions are very rarely enforced and are, according to managers, essentially unenforceable. Over the past decade, manufacturers have begun to use high-powered performance contracts for entire automobile systems (assigning all responsibility for an individual system to an individual supplier); we discuss this trend in light of our findings in the Conclusion.

contracts were the primary source of incentives and operated independently of governance mode, then the impact of governance choices on alternative performance margins would be muted. However, the nature of relational contracts differs within firms versus across firms (Baker, Gibbons and Murphy, 2002): while within-firm relational contracts are used to support the power of hierarchy (e.g., by encouraging hard-to-monitor firm-specific investments), across-firm relational contracts support the objectives of contracting (e.g., by enhancing the incentives to comply with contract specifications). These distinctions are likely to be important in the context of automobile product development, where internal teams are primarily long-term employees of a given manufacturer, while external teams work with multiple manufacturers over time and have only a more limited relationship with any one downstream firm. For example, whereas internal teams may be provided subjective incentives based on their perceived contribution to overall product performance, implicit incentives provided to external contractors are primarily related to their ability to contribute to the (potentially subjective) performance of the component or system covered by contract. In other words, though relational mechanisms operate both within and between firms, the nature of the relational contract varies by governance mode, and procurers are limited in their ability to use implicit contracts to exercise authority (beyond that which is covered by contract) across firm boundaries.

The Limits to Internal and External Learning and Experience. As discussed earlier, the KBV suggests that external contracting will favor early performance while internal development favors long-term performance. This assumes that the impact of learning and knowledge accumulation operate over the course of a single product lifecycle in isolation (i.e., from the initiation of one major model change to the next). However, if learning and experience operate across multiple projects and over many model generations, the impact of governance mode on short-term versus long-term performance may be muted. A few important cases of this mediating effect stand out. First, while the short-term performance penalty (and the opportunities for improvements over the course of the product lifecycle) will be particularly high for a novel manufacturer (i.e., a manufacturer who is at the earliest stages of learning and capability development), the effect of learning and knowledge on different performance margins may be more muted for manufacturers with substantial *ex ante* experience. Second, the short-term advantages to external contracting will be particularly high when one is able to outsource to a globally innovative technological leader; the short-term penalty to vertical integration may be lower when external

contractors possess more generic capabilities. Last, the opportunities for learning over the product lifecycle varies across components and systems; while it is possible to “tweak” modular systems such as the brakes, it is more challenging to achieve significant improvements in complex systems such as the engine over the product lifecycle. Taken together, these considerations motivate our final set of hypotheses, on the mediating impact of experience and learning on vertical integration and performance:

Hypothesis 3a: The impact of vertical integration on initial performance will be muted for firms with a higher level of ex ante experience and/or knowledge.

Hypothesis 3b: The impact of outsourcing on initial performance will be muted in the absence of a globally innovative external contractor.

Hypothesis 3c: The impact of vertical integration on performance improvement will be higher in environments where opportunities for learning are higher.

These interaction effects can be tested by taking advantage of differences across manufacturers in their experience level (most notably, we observe the “entry” of novel luxury models by Japanese manufacturers), variation in the availability of globally innovative external contractors, and in the potential for learning within particular systems. The remainder of the paper evaluates the hypotheses using a detailed dataset that combines information about the level and drivers of vertical integration at the automobile system level with system-level performance measures over the product lifecycle.

III. Data and Methods

Sample and Methods. This paper combines a proprietary, original dataset of contracting choices, product architecture, and the contracting environment in the global auto industry with system-specific ratings drawn from *Consumer Reports* (*CR* hereafter). Focusing on the luxury performance car segment (defined by *CR* as vehicles priced above \$30,000 in 1995), the sample includes European, US, and Japanese companies; accounting for roughly 90% of global luxury market revenues (Appendix B describes the sample and interview process). We focus on the luxury segment for three related reasons. First, as flagship vehicles developed in different environments over time, wide variation in contracting practices and the contracting environment was expected. Second, we not only limit the measurement problems arising from combining information across vehicle types, but also focus on a segment where competitive advantage is highly sensitive to product quality (relative to price and cost tradeoffs). Finally,

the data are based on more than 1000 detailed interviews with managers and engineers in this segment, and it would be difficult (if not impossible) to construct a similarly detailed and nuanced dataset for other industry segments. We combine our proprietary contracting choice data with system-specific *CR* performance ratings, which we use as the basis for the dependent variables throughout the analysis. The final dataset consists of 112 observations of system-specific contracting choice, the contracting environment, and performance. Table 1 provides all variable names, definitions and summary statistics.

System-specific performance measures. The dependent variables throughout the analysis are calculated using system-specific *CR* performance measures. For each of the 19 “major” model changes and seven automobile systems (brakes, transmission, etc) in our dataset, we gather performance data for the first four years after the introduction of each “major.” For each system i on model j in year t after a major model change, $Performance\ Rating_{ijt}$ is the first *CR* quality rating for that system (ranging from 1-5, with 5 as the “highest rating.”).⁹ 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 both short-term and long-term performance measures over the product lifecycle. First, *Short-Term Performance* is the average of $Performance\ Rating_{ij0}$ and $Performance\ Rating_{ij1}$, as available. 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 *CR* rating is available, only one is used. There is a significant difference in the mean levels of *Short Term Performance* and *Long Term Performance* (mean = 3.43 versus 3.70). In other words, over the product lifecycle, there is a modest upward performance trend. To measure the level of change, we construct *Performance Change* ($Long\ Term\ Performance - Short\ Term$

⁹ Each model-year generates multiple years of potential performance data, since the ratings for a given model-year are updated annually after the year of initial introduction. While our measures use the *first* rating for each vehicle model-year, the main findings are robust to alternative formulations using subsequent performance evaluation years.

Performance). Finally, *Overall Performance*, the average of *Short Term Performance* and *Long Term Performance*, measures performance over the lifecycle.¹⁰

Contracting Variables. The 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. For each component, system, vehicle model, and time period, we have collected data on the make / buy decision outcome. *Vertical Integration* at the system level is calculated as the average across the individual components for that system, with each component weighted equally.^{11,12} *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, the variation in *Vertical Integration* is “model-specific”: in an OLS regression of *Vertical Integration* on individual model-year dummies, $R^2 = 0.58$, most of the model-year dummies are individually significant, and the overall F-test is highly significant at 8.74. In other words, vertical integration is “clustered” according to model-year. As described below, we exploit this correlation in *Vertical Integration* across systems within model-year to construct instrumental variables for *Vertical Integration* in the context of a performance regression.

Contracting and Performance Drivers. Our analysis also includes system-specific contracting and performance drivers (Table 1 includes definitions, and Appendix C discusses these measures in detail, including their effect on performance and their relationship to *Vertical Integration*). The system-specific measures include *Sunk Cost*, *Low Capacity*, *Platform Complexity*, and *Design Goal*. As well, we include a measure of the geographic origin of each model (*Japan OEM*),¹³ and the timing of the major model introduction (*Year*). Together, these measures (a) control for system-specific and model-specific

¹⁰ *Overall Performance* is not simply the average of *Performance Rating*. Since we require only one observation in each two-year period after the major model introduction to calculate *Short Term Performance* and *Long Term Performance*, an average confounds differences in the number of ratings with the timing of those ratings.

¹¹ 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), were treated as outsourcing. For more information on this measure, see Novak and Eppinger (2001).

¹² Because our performance measures (and measures of the drivers of vertical integration) are at the *system* level, we are unable to take advantage of the limited supplier information we have at the *component* level (e.g., the length and scope of bilateral relationships, or the degree to which competitors share similar suppliers at the component level). Integrating such information requires aggregating different supply conditions across different supply relationships while accounting for the interdependencies among those supply relationships (Novak and Wernerfelt, 2007).

¹³ We also experimented with dummies to distinguish US and European manufacturers, with no separate impact on the results. We focus specifically on the *Japan OEM* dummy, since Japanese manufacturers were new entrants into the luxury segment, and so may have had high opportunities for learning over the lifecycle.

performance drivers that may be correlated with *Vertical Integration*, and (b) account for differences in the governance environment which may influence the impact of vertical integration on specific performance margins. For example, a platform system (where key components are shared among multiple models) may not only impact performance directly, but may enhance the impact of vertical integration and long-term performance (e.g., since platform requirements will be associated with high learning opportunities within vertically integrated systems). We also include system-specific dummy variables (Brakes, Engine, etc) to control for differences across rated systems. Finally, in line with Hypothesis 3, we construct system-specific measures of the opportunity to access globally innovative suppliers (*Innovative Supplier System*).

The Empirical Framework. Our empirical objective is to exploit this small but nuanced dataset to examine short-term and long-term performance for each of seven individual systems within a model. To focus on the core empirical patterns, we adopt a simple linear specification:¹⁴

$$SHORT\ TERM\ PERFORMANCE_{ij} = \beta_0^{STP} + \beta_{VI}^{STP} VERTICAL\ INTEGRATION_{ij} + \beta_X^{STP} X_{ij} + \mu_{ij}^{STP} \quad (1)$$

$$PERFORMANCE\ CHANGE_{ij} = \beta_0^{PERF\Delta} + \beta_{VI}^{PERF\Delta} VERTICAL\ INTEGRATION_{ij} + \beta_X^{PERF\Delta} X_{ij} + \beta_{STP}^{PERF\Delta} STP_{ij} + \mu_{ij}^{PERF\Delta} \quad (2)$$

where X_{ij} includes model-system measures (*Sunk Cost*, *Low Capacity*, *Platform Complexity*, and *Design Goal*) and model-level measures (*Japan OEM*, *Year*), and STP_{ij} abbreviates *Short Term Performance*.¹⁵ We are interested in evaluating whether $\beta_{VI}^{STP} < 0$ and $\beta_{VI}^{PERF\Delta} > 0$: the impact of governance mode on these performance margins is consistent with the idea that firm boundaries matter (e.g., one cannot “contract around” the problem), and that governance choice induces a particular performance profile. We also consider interaction effects between *Vertical Integration* and measures of the level of *ex ante* internal capabilities (*Sunk Costs*, *Japan OEM*), the potential for short-term gains from outsourcing (*Innovative Supplier System*), and the potential for performance improvement through internal development

¹⁴ This formulation assumes that both performance measures contribute separately to firm value; our qualitative evidence suggests that luxury segment managers believe that perceptions of high quality impact sales and market power for models over the entire lifecycle. As well, we assume value is continuous in the performance measures. While ordered probit allow for “threshold effects,” these nonlinear models yielded little additional insight (the qualitative results are identical and remain significant).

¹⁵ We can re-estimate (2) with *Long Term Performance* as the dependent variable. The results are identical (except for $\beta_{STP}^{PERF\Delta}$). As well, Table 4 explores non-linear forms for *Short Term Performance* in (2), and our results are also robust to a non-linear treatment of the time trend (e.g., by allowing for up to five different time periods). Finally, we experimented with alternative control structures, and the key findings are robust across these specifications.

(*Platform*). These interaction effects identify the mediating influence of knowledge and capabilities on the governance-performance relationship.

An important challenge in interpreting our results is that contract choice is endogenous. Firms that are vertically integrated into a particular activity (e.g., the design and manufacture of a specific system) will tend to have higher overall returns to vertical integration for that activity, relative to the returns from outsourcing. Comparing overall performance results between firms can therefore be misleading. For example, even if every firm faces a clear performance tradeoff, a cross-sectional comparison need not find any performance consequence to vertical integration. While readily acknowledging that an ideal experiment would allow us to separately identify the selection process for vertical integration independent of the impact of vertical integration choices on different selection margins, it is important to emphasize that the impact of selection may be limited in the current context. Specifically, we are not examining *total profits*, but are focusing on the implications of vertical integration for specific performance margins. Just as the modern theory of the firm is premised on the idea that organizational choices have *both* costs and benefits, our analysis traces out the (static and dynamic) costs and benefits associated with vertical integration versus outsourcing. An OLS estimator can therefore provide a consistent estimate of the impact of vertical integration on the observed performance margins to the extent that the level of vertical integration is independent of μ . Assuming that μ is random (e.g., determined by random noise (at least in expectation) in the determination of performance ratings), we will observe variation in *Vertical Integration* under the following two scenarios: (a) conditional on observables, each firm receives a mean-zero relative cost shock that affects the costs of vertical integration (relative to outsourcing) for each system-model-year (Masten and Saussier, 2002), or (b) conditional on observables, each firm receives an independent draw that determines the relative weight placed on short-term versus long-term performance.¹⁶ If variation in *Vertical Integration* comes from differences in the (unobserved) costs of integration or from differences across models in the relative

¹⁶ In other words, suppose that each firm receives a random draw, λ_i , on the relative importance of short-term versus long-term performance and the firm maximizes solves $\text{Max}_{VI_i} \lambda_i(\text{Short-Term}_i(VI_i)) + (1 - \lambda_i)(\text{Long-Term}_i(VI_i))$. While the relationship between each performance measure and *Vertical Integration* is fixed, firms will choose different levels of *Vertical Integration* depending on the realized level of λ .

valuations placed on alternate performance margins, we can consistently estimate the impact of *Vertical Integration* on each of the performance measures.

Of course, it is possible that those firms who choose a high level of vertical integration for a particular model-system may expect relatively high (conditional) returns to vertical integration. As such, we also experiment with an instrumental variables approach.¹⁷ To account for the potential endogeneity of *Vertical Integration*, we employ instruments correlated with *Vertical Integration* for a given system but exogenous to the performance of that system. Building on Novak and Stern (2007), which highlights the potential for complementarity across vertical integration decisions within an automobile model, we calculate instruments based on the drivers of vertical integration for *other* systems within a given automobile model (and control directly for system-specific drivers of vertical integration within each system). In other words, for system i in model j , we control directly for $Sunk\ Cost_{i,j}$, and construct $Sunk\ Cost_{-i,j}$, which is equal to the average of $Sunk\ Cost$ over other systems for model j . Using an analogous procedure for each system-specific driver, we calculate the excluded IV vector:¹⁸

$$Z_{ij} = \{SUNK\ COST_{-i,j}, LOW\ CAPACITY_{-i,j}, PLATFORM_{-i,j}, COMPLEXITY_{-i,j}, DESIGN\ GOAL_{-i,j}\}$$

In addition to a single-equation IV estimator, we also experiment with a 3SLS procedure, to account for correlation across the two performance equations, and improve efficiency. Finally, we have experimented extensively with alternative control structures, including the use of company fixed effects. While the pattern of results remains the same, the results are not as robust across all alternative specifications. Rather than overstate our results, we limit our presentation to specifications which exploit variation across companies, with standard errors clustered by company, and comment where necessary about the robustness (or not) of individual results to the use of company fixed effects.

IV. Empirical Results

The empirical analysis proceeds in several stages. First, we present descriptive evidence about performance patterns over the product lifecycle, according to the extent of vertical integration. Second,

¹⁷ *Vertical Integration* is a continuous rather than dichotomous variable, and so we cannot simply implement a standard selection correction (as in Masten, et al, 1991; Poppo and Zenger, 1998; and Nickerson and Silverman, 2003). Instead, following Wooldridge (1997), we implement a standard instrumental variables estimator to estimate an average treatment effect for a continuous endogenous variable which interacts with unobserved heterogeneity.

¹⁸ We have also experimented with using $Vertical\ Integration_{-i,j}$ directly. While this measure is itself endogenous, it is useful to note that the results remain qualitatively identical.

we evaluate the impact of vertical integration on initial performance and performance improvement. We then focus on interactions between vertical integration and factors that might impact the returns to outsourcing for specific performance margins.

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 pronounced difference in the performance level (3.59 versus 3.07). However, by the fourth year after product introduction, there is convergence in “raw” performance levels. A similar pattern is observed in Table 2, where we divide the average of *Short Term Performance* and *Long Term Performance* by the extent of vertical integration. There is a statistically and quantitatively significant difference in *Short Term Performance*, but convergence in the raw levels for *Long Term Performance*. Of course, these patterns could simply reflect spurious correlation, and so we turn to a more systematic regression framework.

Short-Term Performance. Table 3 reports the *Short Term Performance* regressions. First, a simple regression with *Vertical Integration* has a large and statistically significant negative relationship with *Short Term Performance*. This pattern is robust to the inclusion of a linear time trend, a set of six system-level dummy variables, and clustering the standard errors by company (which we do for the remainder of the specifications). While there is a significant upward time trend in the level of *Short Term Performance*, this trend is essentially independent of the relationship between *Vertical Integration* and *Short Term Performance*. When we include the complete set of control variables (3-3), the *Vertical Integration* coefficient is almost identical. This pattern holds across a wide range of specifications – while no other measure (except the time trend) has a consistent relationship with *Short Term Performance*, the *Vertical Integration* coefficient is positive, and quantitatively and statistically significant.¹⁹ The final column of Table 3 turns to an instrumental variables approach. To account for the potential endogeneity of *Vertical Integration*, we utilize the instrumental variables described in

¹⁹ These findings for *Vertical Integration* are statistically significant across a wide range of alternative specifications, including explorations of alternative control structures (including company fixed effects), non-linear time trends, and alternative definitions of the short-term performance measure. The only caveat is that the IV estimate with company fixed effects, though similar in magnitude, is noisy and just below the 10% significance level.

Section 3, based on the drivers of vertical integration for *other* systems within a given automobile model (controlling directly for system-specific drivers of performance within each system). The instrumental variables estimate on *Vertical Integration* is consistent with (and actually larger in absolute magnitude than) than the OLS coefficient. As in the OLS specification, there is no other robust driver of initial model performance except for *Vertical Integration* and the time trend.²⁰ Overall, using several measures for initial performance, employing OLS or instrumental variables, and controlling (or not) for other potential performance drivers, the results point to a significant relationship between *Vertical Integration* and *Short Term Performance*.

Performance Change. Table 4 reports an analogous set of specifications for *Performance Change*. In (4-1), we include *Vertical Integration*, which has a positive and 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). Compared with Table 3, the only difference is that, in the final two columns, we control for *Short Term Performance*: (4-2) includes a time trend and system dummies, (4-3) includes the complete set of controls (including a nonlinear treatment of *Short Term Performance*), and (4-4) implements an IV estimator. We control for *Short Term Performance* because while it is possible to shift up or down in the performance ratings (*Performance Change* ranges from -2 to 2), a high (or low) initial rating tends to “constrain” the potential for performance improvement (or erosion). Since the errors in the short-term and performance change equation are likely correlated, *Short Term Performance* may be endogenous, and so in (4-4), we instrument for both *Vertical Integration* and *Short Term Performance*. Finally, in Table 5, we present a 3SLS estimator, which accounts for the endogeneity of *Vertical Integration* and *Short Term Performance* in the context of an efficient linear estimator.

Across all of these specifications, there is a positive and significant relationship between *Vertical Integration* and *Performance Change*. The estimates suggest that a shift of *Vertical Integration* from 0 to 1 is associated with a shift in *Performance Change* ranging from 0.45 (in (4-3)) to more than 0.92 (in the 3SLS estimator).²¹ Interestingly, the 3SLS estimator offers a more precise estimate of the impact of

²⁰ Not surprisingly given the close relationship between the OLS and instrumental variables estimates, a Hausman test cannot reject the exogeneity of *Vertical Integration*, and we cannot reject the exogeneity of the instrumental variables vector in an overidentification test.

²¹ The IV vector remains the same as in (3-4). A Hausman test cannot reject the exogeneity of *Vertical Integration*, and we cannot reject the overidentifying restrictions.

Vertical Integration (and is quite robust across alternative specifications). Further, the data suggest a “mean reversion” effect: models receiving higher initial performance ratings tend to experience a lower level of *Performance Change* (the coefficient on *Short Term Performance* in the *Performance Change* equation is consistently negative). Finally, while *Low Capacity* is associated with a lower level of performance improvement (perhaps suggesting constraints on production or the adoption of new technologies), *Japan OEM* and *Sunk Cost* are associated with a positive boost in the change in performance (consistent with our theoretical discussion). Together, Tables 3-5 provide consistent evidence for a negative relationship between vertical integration and initial performance, but a positive relationship between vertical integration and performance improvement.²²

Do the Environment for Learning and Access to Capabilities Matter? Our final empirical exercise examines whether our findings are mediated by the knowledge environment. Table 6 examines four different interaction effects: *Japan OEM*, *Sunk Cost*, *Innovative Supplier Availability*, and *Platform*. We are interested in whether variation in *ex ante* capabilities (identified by *Japan OEM* and *Sunk Cost*), external expertise (*Innovative Supplier Availability*), or opportunities for learning (*Platform*) impact the role of *Vertical Integration* along short-term and long-term performance margins. Each of the regressions in Table 6 is identical to the specifications in (3-3) and (4-3), except for the inclusion of measures to identify the impact of the interaction between *Vertical Integration* and the dummy variables described above. We begin in Panel 6A with the *Japan OEM* measure. While both Japanese and non-Japanese manufacturers face a short-term performance penalty for vertical integration, the level of *Performance Change* arising from integration is more than twice as high for Japanese manufacturers. This is consistent with our hypothesis that, as novel producers in the luxury segment with a high level of internal flexibility (and a minimum of union restrictions), Japanese producers may be able to realize particularly significant performance improvement over the product lifecycle. Our evaluation of *Sunk Cost* interactions provides complementary evidence. The KBV suggests that firms with a deeper knowledge base may be able to apply their prior knowledge in order to mitigate some of the early-stage costs of vertical integration (in

²² Appendix D reports on the relationship between *Vertical Integration* and *Overall Performance*. Though the coefficient on *Vertical Integration* is negative in Appendix D, the estimates are small and insignificant; moreover, this coefficient is positive in alternative specifications (available upon request). This is consistent with the idea that while there is no systematic relationship between governance mode and overall performance (particularly if firms vary in their governance choices), there is a relationship along distinct performance margins.

terms of reduced short-term performance), but may be constrained in opportunities for additional learning over the lifecycle (Grant, 1996). In Panel 6B, for firms with *Sunk Cost* = 1, there is no early-stage loss from *Vertical Integration*, and there is only a noisy relationship between *Vertical Integration* and *Performance Change*. In contrast, for firms that lack sunk investments, vertical integration is associated with large *Short Term Performance* penalty and significant performance improvement after product introduction.

We next turn attention towards the impact of cutting-edge global suppliers on the tradeoffs arising from contracting versus integration. To the extent that one of the key benefits of outsourcing is the ability to gain access to frontier technology and best-in-class capabilities, the short-term costs of vertical integration will be particularly high when a globally innovative supplier is available. Among other factors, it is likely that a competitor will be able to access this leading technology, resulting in a relative performance loss. Panel 6C provides evidence consistent with this logic. While the coefficient on *Vertical Integration* in the *Short Term Performance* equation is more than three times larger when *Innovative Supplier* = 1, there is no significant difference in the impact of *Vertical Integration* on performance improvement by the presence or absence of a global technology leader.

Finally, we explore the interaction between *Vertical Integration* and *Platform*. Since there may be economies of scope in learning and adaptation, firms may invest more significantly over the product lifecycle when using a platform approach (i.e., the returns to a strong knowledge base are realized across a wider range of automobile models). For example, Toyota used its deep knowledge base in design-for-manufacturing techniques to rapidly improve product design in the Lexus luxury vehicle. Despite having only entered the luxury segment in 1989, Toyota used a platform approach to achieve global leadership by 1995. Panel 6D offers evidence consistent with this supposition. Though the *Short-Term Performance* penalty is similar regardless of whether one uses a platform approach or not, the performance improvement benefits arising from vertical integration are concentrated in model-systems employing a platform approach. The findings from Table 6 suggest that not only does *Vertical Integration* influence specific performance margins (in line with theory), but that the salience of these effects is grounded in the environment for knowledge development and application. As such, this evidence provides quantitative support for key hypotheses in the KBV.

VI. Concluding Remarks

This paper examines the impact of vertical integration on the dynamics of performance in the context of automobile product development. Our insight is that vertical integration will have differential impacts on different performance margins that are realized over the product lifecycle. On 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 adapt to unforeseen contingencies and customer feedback, to maintain more balanced incentives, and to develop firm-specific capabilities. These effects suggest that outsourcing will be associated with higher levels of initial performance, while vertical integration will be associated with performance improvement over the product lifecycle.

To test these ideas, we combine detailed performance measures over time with nuanced measures of the extent of vertical integration, and measures of the contracting and technology environment. We are cautious in interpreting our results: the size of our dataset is modest, and, while we account for endogeneity using instrumental variables, we are not using an unambiguous natural experiment. With these caveats in mind, we establish several robust findings. First, initial performance is declining in the level of vertical integration, and, second, the level of performance improvement is significantly increasing in the level of vertical integration. Finally, the impact of vertical integration on alternative performance margins depends on the environment for contracting and learning. For example, both the benefits and costs of vertical integration were particularly high for Japanese manufacturers, who were entrants with opportunities for learning in the luxury automobile segment. Overall, the findings highlight a strategic governance tradeoff between short-term performance and the evolution of firm capabilities.

These findings have implications for practice. Given the sharp tradeoffs associated with governance choice along specific performance margins, the results suggest that manufacturers may have incentives to contract directly on a measure of post-launch product performance. As discussed earlier, such contracts were essentially unenforceable during our sample period, as performance is realized at the system (or even vehicle level), while hundreds of individual contracts are signed at the component level. A potential solution for addressing the governance tradeoff is to aggregate individual contracting decisions; for example, one could specify a single outsourcing contract at the “system” level assigning responsibility for the system over the product lifecycle (with high-powered incentives based on

observable performance measures). Indeed, this is precisely what has been observed in practice after our sample period: manufacturers now specify system-level (or even multi-system) contracts (e.g., an “interior complete” contract assigned to a single Tier I supplier), where the contractor is subsequently responsible for all component-level sourcing within that system. In future work, it would be interesting to investigate the impact of these system-level contracts on observed performance dynamics.

Finally, our approach seems to offer a novel alternative for evaluating the relationship between governance and performance. Rather than focusing on the ambiguous relationship between vertical integration and overall performance, our analysis suggests that governance mode induces performance along specific performance margins, and the nature of these margins can be grounded in specific theoretical frameworks (e.g., transaction cost economics, contract theory, or the KBV). As such, similar to recent work by Forbes and Lederman (2006), our empirical approach may prove useful in testing the empirical implications of specific theories of firm boundaries and organizational design.

REFERENCES

- Anderlini, L. and L. Felli (1999), "Incomplete Contracts and Complexity Costs," *Theory and Decision* 46(1):23-50.
- Argyres, N. (1996), "Evidence on the Role of Firm Capabilities in Vertical Integration Decisions." *Strategic Management Journal*, 17: 129-150.
- Bajari, P. and S. Tadelis (2001) "Incentives Versus Transaction Costs: A Theory of Procurement Contracts." *RAND Journal of Economics*, 32(3): 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*, 1443-1480.
- Baker, G., R. Gibbons and K. Murphy (1994), "Subjective Performance Measures in Optimal Incentive Contracts," *Quarterly Journal of Economics* 109: 1125-56.
- Baker, G., R. Gibbons and K. Murphy (2002). "Relational Contracts and the Theory of the Firm," *Quarterly Journal of Economics*.
- Battigalli, P., and G. Maggi (2004). "Costly Contracting in a Long-Term Relationship," Working Paper EC-04-33, Stern School of Business, N.Y.U.
- 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 (2005). "Transaction Cost Economics: An Assessment of Empirical Research in the Social Sciences," mimeo, Georgetown University.
- Conner, K.R. and C.K. Prahalad (1996). A Resource-based Theory of the Firm: Knowledge versus Opportunism, *Organization Science*, 7: 477-501.
- Crocker, K.J. & S.E. Masten (1991). "Pretia ex Machina? Prices and Process in Long-Term Contracts," *Journal of Law & Economics*, 34(1): 69-99.
- Crocker, K. and K. Reynolds (1993). "The Efficiency of Incomplete Contracts: An Empirical Analysis of Air Force Engine Procurement," *Rand Journal of Economics*, 24: 126-46.
- Dye, R.A., (1985). "Costly Contract Contingencies," *International Economic Review*, 26(1): 233-250.
- Forbes, S.J. and M. Lederman (2005). "Control Rights, Network Structure and Vertical Integration:

- Evidence from Regional Airlines,” mimeo, UCSD.
- Grant, R.M. (1996). “Toward a Knowledge-Based Theory of the Firm,” *Strategic Management Journal*, 17:109-122.
- Hart, O. (1988). “Incomplete Contracts and the Theory of the Firm,” *Journal of Law, Economics and Organization*, 4(1): 119-139.
- Hart, O. and J. Moore (1990). “Property Rights and the Nature of the Firm,” *Journal of Political Economy*, 98: 1119-1158.
- Joskow, P., (1988). “Asset Specificity and the Structure of Vertical Relationships: Empirical Evidence,” *Journal of Law, Economics and Organization*, 4: 95-117.
- Klein, P.G. (2004). “The Make-or-Buy Decision: Lessons from Empirical Studies,” *Contracting and Organizations Research Institute Working Paper No. 2004-07*.
- Kogut, B., and U. Zander (1992). “Knowledge of the Firm, Combinative Capabilities, and the Replication of Technology,” *Organization Science*, 3(3): 383-397.
- Kogut, B., and U. Zander (1996). “What Firms Do? Coordination, Identity and Learning,” *Organization Science*, 7(5): 502-518.
- Leiblein, M.J., Reuer, J J., 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.
- Macher J.T. (2006). “Technological Development and the Boundaries of the Firm: A Knowledge-Based Examination in Semiconductor Manufacturing.” *Management Science*, forthcoming.
- 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.
- Masten, S. (1993). “Transaction Costs, Mistakes, and Performance: Assessing the Importance of Governance,” *Managerial and Decision Economics*, 14:119-129.
- Masten, S. and S. Saussier (2002). “The Econometrics of Contracts: An Assessment of Developments in

- the Empirical Literature on Contracting,” in *The Economics of Contracts: Theory and Application*, E. Brousseau and J-M. Glachant, ed., Cambridge.
- Mayer, K.J. (2000). "Transactional Alignment and Project Performance: Evidence from Information Technology", Working Paper.
- Mayer, K.J. and J. Nickerson (2005). “Antecedents and Performance Implications of Contracting for Knowledge Workers: Evidence from Information Technology Services, *Organization Science* 16.
- Ménard, C. and S. Saussier (2002). “Contractual Choice and Performance: the Case of Water Supply in France”, in *The Economics of Contracts: Theories and Applications*, Brousseau E. and Glachant J-M (eds.), Cambridge University Press: 440-462.
- Nickerson, J. A. and B. S. Silverman. (2003). “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*, 48(3): 433-465.
- Nickerson, J.A. and T. Zenger (2004). “A Knowledge-based Theory of Governance Choice: The Problem Solving Approach,” *Organization Science*, 15(6): 617-632.
- Novak, S. and S. Eppinger (2001). “Sourcing by Design: Product Complexity and the Supply Chain,” *Management Science*, 47(1): 189-204.
- Novak, S. and S. Stern (2007). “Complementarity Among Vertical Integration Decisions: Evidence from Automobile Product Development,” mimeo, Northwestern University.
- Novak, S. and B. Wernerfelt (2007). “The Design of Industry,” mimeo, MIT.
- Poppo, L. and 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.
- Poppo, L. and T. Zenger (2002). “Do formal contracts and relational governance function as substitutes or complements?” *Strategic Management Journal*, 23(8):707-725.
- Sampson, R.C. (2004). “The Cost of Misaligned Governance in R&D Alliances,” *J Law Econ Organ* 20: 484-526
- Saussier, S. (2000). “Transaction Costs and Contractual Incompleteness: The Case of Electricite de

- France.” *Journal of Economic Behavior and Organization* 42: 189-206.
- Shelanski, H., and P. Klein (1995). “Empirical Research in Transaction Cost Economics: A Review and Assessment,” *Journal of Law, Economics, and Organization*, 11: 335-361.
- Silverman, B.S., Nickerson, J.A., and J. Freeman (1997). “Profitability, Transactional Alignment, and Organizational Mortality in the U.S. Trucking Industry,” *Strategic Management Journal* 18:31-52.
- Whinston, M. D. (2001), “Assessing the Property Rights and Transaction-Cost Theories of Firm Scope,” *American Economic Review (Papers and Proceedings)* 91: 184-8.
- Williamson, O.E. (1985), *The Economic Institutions of Capitalism*, New York: Free Press.
- Williamson, O.E. (1991), “Comparative Economic Organization: The Analysis of Discrete Structural Alternatives,” *Administrative Science Quarterly*, 36(2):269-296.
- Wooldridge, J.M. (1997), “On two stage least squares estimation of the average treatment effect in a random coefficient model,” *Economics Letters*, 56(2):129-133.
- Yvrande-Billon, A. and C. Ménard (2003). “Institutional Constraints and Organizational Changes: The Case of the British Rail Reform,” mimeo, ATOM, N°2003-19.
- Yvrande-Billon, A., and S. Saussier (2005). “Do Organization Choices Matter? Assessing the Importance of Governance through Performance Comparisons,” in *New Ideas in Contracting and Organizational Researches*, H. James (ed.), Nova Science Publishers.

FIGURE 1
Timing of Procurement and Ex-Post Adaptation

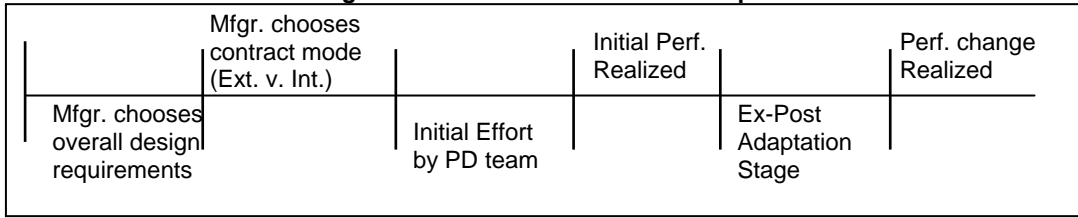


FIGURE 2
Average Performance Rating By High or Low Vertical Integration, by Years Since "Major" Model Introduction

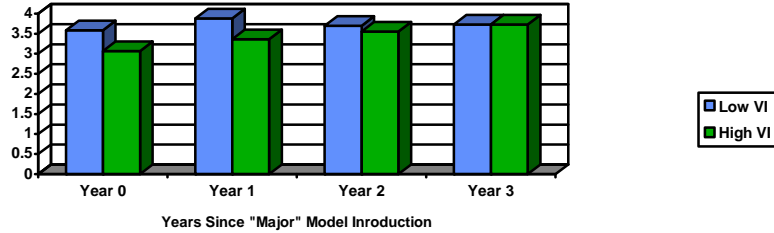


TABLE 1
Variables & Definitions

VARIABLE	DEFINITION	MEAN	STD. DEV.
PERFORMANCE MEASURES			
PERFORMANCE RATING	<i>Consumer Reports</i> rating (from 1-5, w/ 5 as the "highest rating") for system <i>i</i> on model <i>j</i> in year <i>t</i> after a major model change. The rating for a given model-year is the "first" <i>CR</i> 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 1 rating is available, only 1 is used).	3.420	.965
LONG TERM PERFORMANCE	Average of PERFORMANCE RATING _{ij2} and PERFORMANCE RATING _{ij3} , as available (when only 1 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	3.563	.871
CONTRACTING MEASURES			
VERTICAL INTEGRATION	Percentage of the system produced in house between 0 and 1 (1 indicates all in-house production)	.513	.318
SYSTEM-SPECIFIC CONTRACTING AND PERFORMANCE DRIVERS			
SUNK COST	Dummy = 1 if pre-existing in-house sunk costs and/or plant investment for system <i>i</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	.392	.275
DESIGN GOAL	Measure for desired performance goals at the system level, ranging from 0 (low) to 1 (high)	.457	.311
MODEL-YEAR MEASURES			
JAPAN OEM	Dummy = 1 if company headquarters are located in Japan	.366	.484
YEAR	Year of Product Introduction for Major Model change	1990.080	5.309
SYSTEM GROUPINGS			
INNOVATIVE SUPPLIER SYSTEM	Dummy = 1 if system is Transmission, Electrical, or Brakes. Based on availability (during the sample period) of globally available innovative suppliers	.447	.499

TABLE 2
Performance Rating Margins, By High or Low Vertical Integration

VERTICAL INTEGRATION	SHORT TERM PERFORMANCE	LONG TERM PERFORMANCE	PERFORMANCE CHANGE
"Below" Median	3.70	3.73	0.03
"Above" Median	3.25	3.69	0.44

Median VERTICAL INTEGRATION = 0.50

TABLE 3
Short-Term Performance

Dependent Variable : SHORT TERM PERFORMANCE (N=112)										
	(3-1)	(3-2)			(3-3)			(3-4)		
	Ordinary Least Squares						Instrumental Variables			
VERTICAL INTEGRATION	-0.794** (0.279)	-0.679** (0.249)			-0.681** (0.274)			-0.907** (0.245)		
SUNK COST					-0.373 (0.422)			-0.346 (0.420)		
LOW CAPACITY					0.043 (0.375)			-0.028 (0.365)		
PLATFORM					-0.023 (0.119)			-0.038 (0.123)		
COMPLEXITY					0.311 (0.357)			0.310 (0.342)		
DESIGN GOAL					-0.445 (0.523)			-0.508 (0.507)		
JAPAN OEM					0.074 (0.207)			0.081 (0.203)		
YEAR		0.090 (0.024)			0.081** (0.028)			0.078** (0.029)		
CONSTANT	3.827 (0.168)	3.949 (0.268)			4.031 (0.350)			4.235 (0.328)		
<i>Parametric Rest.</i>		#Restr F-stat p-value			#Restr F-stat p-value			#Restr F-stat p-value		
SYSTEM DUMMIES		6 4.84 .037			6 3.52 .075			6 5.28 .031		
<i>R-Squared</i>	0.068	0.407			0.424					
<i>RHS Endogenous Variables</i>								VERTICAL INTEGRATION		
<i>Instrumental Variables</i>								For system <i>i</i> of model <i>j</i> , sums of each model-specific measure for all systems but system <i>i</i> . This is defined as: $Z_{-i,j} = \left(\sum_{l=1, \dots, 7} Z_{lj} - Z_{ij} \right)$ $Z = \left\{ \begin{array}{l} \text{SUNK COST} \\ \text{LOW CAPACITY} \\ \text{PLATFORM} \\ \text{COMPLEXITY} \\ \text{DESIGN GOAL} \end{array} \right\}$		

Notes: (1) Stars denote statistical significance at 5% (**) and 10% (*), respectively.

(2) Standard errors are presented in parentheses. Except for (3-1), SE are clustered by company.

TABLE 4
Performance Change

Dependent Variable : PERFORMANCE CHANGE (N=112)										
	(4-1)	(4-2)			(4-3)			(4-4)		
	Ordinary Least Squares						Instrumental Variables			
VERTICAL INTEGRATION	0.611** (0.243)	0.845** (0.392)			0.456** (0.209)			0.648* (0.362)		
SUNK COST					0.502** (0.137)			0.419* (0.209)		
LOW CAPACITY					-0.212** (0.075)			-0.154 (0.169)		
PLATFORM					-0.137 (0.105)			-0.096 (0.090)		
COMPLEXITY					0.180 (0.321)			0.132 (0.279)		
DESIGN GOAL					0.112 (0.419)			0.104 (0.272)		
JAPAN OEM					0.795** (0.152)			0.670** (0.148)		
YEAR		0.036* (0.020)			0.047** (0.034)			0.051 (0.042)		
SHORT TERM PERFORMANCE (LOW RATING)					1.699** (0.309)					
SHORT TERM PERFORMANCE (MEDIUM RATING)					0.675** (0.292)					
SHORT TERM PERFORMANCE								-0.364 (0.308)		
CONSTANT	-0.028 (0.146)	-0.217 (0.282)			-0.877 (0.290)			0.915 (1.160)		
<i>Parametric Rest.</i>		#Restr	F-stat	p-value	#Restr	F-stat	p-value	#Restr	F-stat	p-value
SYSTEM DUMMIES		6	444.18	.000	6	70.84	.000	6	701.94	.000
<i>R-Squared</i>	0.055	0.126			0.520					
<i>RHS Endogenous Variables</i>								VERTICAL INTEGRATION SHORT TERM PERFORMANCE		
<i>Instrumental Variables</i>								For system i of model j , sums of each model-specific measure for all systems but system i . This is defined as: $Z_{-i,j} = \left(\sum_{l=1, \dots, 7} Z_{lj} - Z_{ij} \right)$ $Z = \left\{ \begin{array}{l} \text{SUNK COST} \\ \text{LOW CAPACITY} \\ \text{PLATFORM} \\ \text{COMPLEXITY} \\ \text{DESIGN GOAL} \end{array} \right\}$		

Notes: (1) Stars denote statistical significance at 5% (**) and 10% (*), respectively.
 (2) Standard errors, clustered by company, are presented in parentheses.
 (3) SHORT TERM PERFORMANCE (LOW RATING) = 1 if SHORT TERM PERFORMANCE < 3, 0 else
 (4) SHORT TERM PERFORMANCE (MEDIUM RATING) = 1 if SHORT TERM PERF > 3 and < 4.5, 0 else.

TABLE 5
SHORT-TERM PERFORMANCE AND PERFORMANCE CHANGE:
THREE STAGE LEAST SQUARES ESTIMATION

Dependent Variable	Three Stage Least Squares	
	(5A) SHORT TERM PERFORMANCE	(5B) PERFORMANCE CHANGE
VERTICAL INTEGRATION	-0.907** (0.406)	0.921** (0.336)
SUNK COST	-0.346 (0.287)	0.523** (0.222)
LOW CAPACITY	-0.028 (0.261)	-0.146 (0.197)
PLATFORM	-0.038 (0.164)	-0.084 (0.123)
COMPLEXITY	0.310 (0.325)	0.039 (0.248)
DESIGN GOAL	-0.508 (0.395)	0.256 (0.307)
JAPAN OEM	0.081 (0.167)	0.675** (0.125)
YEAR	0.078** (0.017)	0.028 (0.017)
SHORT TERM PERFORMANCE		-0.063 (0.156)
CONSTANT	4.235 (0.328)	-0.063 (0.745)
SYSTEM DUMMIES	Included	Included
<i>RHS Endogenous Variables</i>	VERTICAL INTEGRATION	VERTICAL INTEGRATION SHORT TERM PERFORMANCE
<i>Instrumental Variables</i>	<p>For system i of model j, sums of each model-specific measure for all systems but system i. This is defined as:</p> $Z_{-i,j} = \left(\sum_{l=1, \dots, 7} Z_{lj} - Z_{ij} \right)$ $Z = \left\{ \begin{array}{l} \text{SUNK COST} \\ \text{LOW CAPACITY} \\ \text{PLATFORM} \\ \text{COMPLEXITY} \\ \text{DESIGN GOAL} \end{array} \right\}$	

Notes: (1) Stars denote statistical significance at 5% (**) and 10% (*), respectively.

**TABLE 6
INTERACTION EFFECTS**

Dependent Variable	SHORT TERM PERFORMANCE	PERFORMANCE CHANGE
Ordinary Least Squares Regressions		
6A. Japan OEM		
VI * (1 – JAPAN OEM)	-0.655* (0.316)	0.362 (0.284)
VI * JAPAN OEM	-0.804** (0.322)	0.896** (0.350)
JAPAN OEM	0.147 (0.279)	0.530** (0.251)
<i>R-squared</i>	0.424	0.626
6B. Sunk Cost		
VI * (1 – SUNK COST)	-0.731** (0.262)	0.431** (0.198)
VI * SUNK COST	0.591 (0.913)	0.932 (0.807)
SUNK COST	-1.053 (0.772)	0.244 (0.406)
<i>R-squared</i>	0.435	0.625
6C. Innovative Supplier Availability		
VI * (1 – INNOVATIVE SUPPLIER)	-0.319 (0.553)	0.304 (0.276)
VI * INNOVATIVE SUPPLIER	-0.950** (0.401)	0.596 (0.413)
INNOVATIVE SUPPLIER	-.0356 (0.477)	-0.301 (0.265)
<i>R-squared</i>	0.432	0.622
6D. Platform		
VI * (1 – PLATFORM)	-0.704** (0.309)	0.256 (0.222)
VI * PLATFORM	-0.653 (0.425)	0.715** (0.292)
PLATFORM	-0.048 (0.269)	-0.360** (0.189)
<i>R-squared</i>	0.424	0.626

- Notes: (1) Stars denote statistical significance at 5% (**) and 10% (*), respectively.
(2) Standard errors, clustered by company, are presented in parentheses.
(3) VI stands for VERTICAL INTEGRATION in labels on interaction terms.
(4) All regressions include System-specific controls (SUNK COST, LOW CAPACITY, PLATFORM, COMPLEXITY, DESIGN GOAL, and JAPAN OEM), system fixed effects and year trend.

Appendix A

Contracting in Automobile Product Development

Our analysis focuses on the product lifecycle for automobile models. While autos are incrementally upgraded annually, an automobile model undergoes a “major” model change approximately every five years. A “major” model change is an opportunity to significantly alter product positioning, technologies, and contracting choices for an automobile model. While a manufacturer is constrained by the history of the vehicle, sunk investments, etc., the process underlying a major model change is substantial, and allows for significant changes in the design and organization of the 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 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.

Though the governance mode is relatively fixed over the product lifecycle, manufacturers devote considerable internal and external resources and attention to improving products as the lifecycle evolves. While establishing a perception of high quality in the initial introduction year is quite important (the reputational effect will spill over to future model-years), companies attempt to address the myriad technical issues arise as the result of large-scale use and respond to detailed (and often voluminous) customer feedback over the life of the major. Many consumers (particularly in the luxury segment) seek out models that are known to have resolved any engineering issues, and this important customer segment will be particularly sensitive to reports of improvement and changes over the product lifecycle. A key difference between internal development and external sourcing is the availability of personnel who had been assigned to product development efforts prior to product launch. While internal governance usually allows for reassignment to initiate improvements after the initial model-year, it is difficult to maintain the integrity of external teams (from the perspective of a downstream procurer). Indeed, many contracts are specified to limit the extent to which employees can be pulled from current projects to return to an earlier one (these provisions reflect trade secrecy concerns).

More generally, 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. In large part, the inability to enforce performance-oriented contracts is a consequence of the underlying production structure: the relationship between observed performance and individual contracts is very noisy and dependent on the actions of other contractors and internal development teams. While procurement takes place at the level of components, overall performance is the level of systems (and overall vehicle performance and vehicles sales depends on the interaction among these systems, combined with other factors such as marketing, distribution issues, etc). Because of the problem of assigning responsibility for failure (or success), the only “contractual” approach would involve outsourcing large segments of the

automobile (e.g., an entire system, or even a combination of systems). Indeed, after our sample period, manufacturers began to use such arrangements (e.g., an “interior complete” contract with a Tier I supplier).

Appendix B
Data Collection

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.²³

The unit of analysis is an automotive system for a specific “major” for a given automobile model. “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 (see Novak and Eppinger (2001) for further details). The data were collected through on-site interviews with over 1000 people, including CEOs, chief engineers, project managers and system engineers involved in development for each model-year.

The unit of analysis is the model-year-system. 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,

²³ 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.

electrical, suspension, steering, and brakes. From this initial dataset of 19 models, each of which includes seven distinct systems, 2 overall models and five individual model-systems were excluded due to inadequate data. While governance choices are at the component level, the performance measures and the contracting environment measures are at the model-system level. Consequently, we are unable to exploit the (limited) information we have about the nature of individual bilateral contracts (e.g., the duration of individual relationships, or the scope of activity covered by an individual contract). The final dataset consists of 112 observations at the model-system-year level of system-specific contracting choice, the contracting environment, and performance.

Appendix C

System-specific Contracting and Performance Drivers

Our analysis also includes a set of system-specific contracting and performance drivers, included to control for model-specific performance drivers that may be correlated with *Vertical Integration*, and also serve as a source of instrumental variables for the level of *Vertical Integration* on other systems within the same automobile model. There are six key measures.

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. When *Sunk Cost* = 1, this likely indicates that a company has significant experience and capabilities in a given system, which may be associated with a higher level of performance over the product lifecycle.²⁴ Moreover, we expect that the short-term performance penalty associated with vertical integration will be muted when *Sunk Cost* = 1, and also that there may be fewer opportunities for new learning and performance improvement within an individual 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. The relationship to 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).

Platform is a dummy variable equal to one for models with platform requirements where the

²⁴ It is also possible that *Sunk Cost* will be associated with high barriers to adopting frontier technology and production methods, perhaps limiting performance (particularly early in the lifecycle). Table 6 explores the interaction between *Vertical Integration* and *Sunk Costs*.

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*. *Platform* is likely itself correlated with *Vertical Integration*. Platform requirements could support in-house production through economies of scope achieved through parts sharing, and so we control for *Platform* in assessing the relationship between *Vertical Integration* and different performance margins.

The degree of system-specific complexity may also impact realized performance (as well as be correlated with *Vertical Integration*). 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. 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. *Complexity* (mean = .39), is the result of applying this procedure for each component within each system.

A separate measure of the design requirements is *Design Goal*, a 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 *Design Goal* reflects the ex ante objectives of the design process for each system, *Design Goal* is predicted to have a positive impact on each of the performance measures. We include it in our analysis as *Design Goal* may itself be correlated with *Vertical Integration* (and also with performance margins). However, the relationship with *Vertical Integration* may be subtle. Certain performance objectives necessitate more complex product designs, such as more integrated architectures, enhancing the returns to vertical integration. However, accessing global frontier technology may necessitate outsourcing. As such, while theory suggests an ambiguous relationship between *Design Goal* and vertical integration, we control for this measure directly in order to avoid conflating the impact of *Vertical Integration* from *Design Goal* on individual performance margins.

Model-Year Measures

Japan OEM (mean = .366) is a dummy variable equal to 1 if the model originates from a firm with company headquarters in Japan. This measure is useful in several ways: First, Japanese companies in this sample were new entrants to the luxury automobile market, and this measure allows us to isolate those firms with less ex-ante model-specific skills, and thus with more opportunities to learn over the product lifecycle. Additionally, Japanese firms are well known to invest heavily in continuous product improvement, and are able to achieve a much higher level of internal flexibility due to the absence of union restrictions. Consequently, the interaction between *Japan OEM* and *Vertical Integration* is predicted to have a negative relationship with *Short Term Performance* and a positive relationship with *Performance Change*.

We also calculate fixed effects for each of the seven automobile systems (*Seats* are the excluded category), and also introduce an overall (de-measured) time trend (*Year*). The average observation is from a 1990 major model change, with a range from 1980 to 1996. We have experimented extensively with alternative time trends, and company fixed effects.

System Groupings

Innovative Supplier System (mean = .447) is a dummy = 1 if the system in question varies in performance based on availability (during the sample period) of globally available innovative suppliers. Interviewees were asked whether suppliers varied in their innovative capacity at the system level and to evaluate the extent to which access to such suppliers was thought to be a direct performance determinant.

Such systems were identified to be Transmission, Electrical, and Brakes. As discussed in Section II, the availability of an innovative external supplier raises the returns to outsourcing, particularly in terms of *Short Term Performance*. As a result, we expect a negative interaction effect between *Vertical Integration* and *Innovative Supplier System* in the *Short Term Performance* equation.²⁵

²⁵ We also experimented with a measure of the potential for learning within each system, based on engineering principles. While the results were consistent with the remainder of the analysis, the interaction effects between vertical integration and this measure of *Adaptive System Potential* were imprecisely estimated; in the interest of space constraints, we drop this measure from our main analysis.

Appendix D
Overall Performance Regressions

Dependent Variable : OVERALL PERFORMANCE (N=112)						
	(C-1)	(C-2)			(C-3)	
	Ordinary Least Squares					
VERTICAL INTEGRATION	-0.488* (0.257)	-0.256 (0.321)			-0.248 (0.265)	
SUNK COST					-0.094 (0.275)	
LOW CAPACITY					0.046 (0.259)	
PLATFORM					-0.068 (0.093)	
COMPLEXITY					0.320 (0.337)	
DESIGN GOAL					-0.316 (0.354)	
JAPAN OEM					0.411** (0.151)	
YEAR		0.108** (0.018)			0.092** (0.018)	
CONSTANT	3.813 (0.155)	3.871 (0.269)			3.769 (0.303)	
<i>Parametric Rest.</i>		#Restr	F-stat	p-value	#Restr	F-stat p-value
SYSTEM DUMMIES		6	70.84	.000	6	21.18 .009
<i>R-Squared</i>	0.032	0.543			0.588	

Notes: (1) Stars denote statistical significance at 5% (**) and 10% (*), respectively.
(2) Standard errors, clustered by company, are presented in parentheses.