

Vertical Integration and the Theory of the Firm: A Review of Theoretical and Empirical Literature

Ashwin Iyengar (1521001)
ashwin.iyengar15@iimb.ernet.in

October 21, 2016

Abstract

This article presents a review of the theoretical foundations of the dominant theory of the firm used to explain Vertical Integration. We then review a few empirical articles on Vertical Integration and map them to the theories that have found support. We summarize with a status of the literature on vertical integration and suggest interesting open topics for further inquiry.

Keywords: Vertical Integration, Theory of the Firm, New Institutional Economics, Transaction Cost Economics

1 Introduction

"New Institutional Economics does not consist primarily of giving new answers to traditional questions of economics - resource allocation and degree of utilization. Rather, it consists of answering new questions, why economic institutions have emerged the way they did and not otherwise..."

Kenneth Arrow

In reviewing the academic literature on the theory of the firm, two aspects strike out as being salient. First, our understanding of nature of the firm has indeed come a long way from our conceptualization of the firm as a black box that takes capital and labor as inputs, that produces products as output, guided by the intention of wealth maximization and governed by the law

of returns. Second, the impact of New Institutional Economics, and particularly of Transaction Cost Economics in our understanding of firm boundaries and of firm scope is unparalleled by any other theory or set of theories in the literature. A review of the literature on the nature of Vertical Integration of firms as explained by theory and as observed empirically therefore is a topic that we are naturally drawn into¹.

2 Antecedents

Prior to Transaction Cost Economics, the common explanation for vertical integration was that it was due to technological factors. It was commonly believed that comprehensive integration, both backward into materials and forward into distribution was the means by which complex products and services were brought to the market. However Williamson (1985) argued that technology would fully determine economic organization only if there were a single technology that was decisively superior to all others, and if technology implied a unique organizational form. Williamson (1985) argued that these assumptions do not hold as long as alternate means of contracting hold despite employing the same technology. This gap in the existing understanding of firms allowed for the development of a theory of the firm based on governance costs of economic organization.

3 Theory of the Firm

Williamson (1988) suggests that the state of transaction cost economics in 1972 was largely where Coase had left it in 1937, and this might have been largely due to the failure to operationalize the concept. The operationalization has required, first that microanalytic factors that contribute to differences in transaction costs across transactions be identified, second that transaction costs be aligned with governance structures, and finally that the process aspects of economic organization be attended to (Williamson, 1988). Transaction Cost Economics has defined these microanalytic factors with respect to the following: 1. behavioral assumptions, 2. dimensionalizing transactions, and 3, process features (Williamson, 1988).

¹The author gratefully credits Prof. Murali Patibandla for having guided him into the fascinating realms of New Institutional Economics as part of the doctoral seminar course for which this article is being submitted

3.1 Behavioral Assumptions

Transaction Cost Economics employs two behavioral assumptions, first that human agents are boundedly rational, and the second that human agents are given to opportunism. The premise therefore is that economic activity is organized so as to economize on bounded rationality while simultaneously safeguarding the transactions in question against the hazards of opportunism. The implications as set out by Williamson (1988) are summarized in Table 1.

Behavioral Assumptions / Implications	Bounded Rationality	Opportunism
For contractual theory	Comprehensive contracting is infeasible	Contract as promise is naive
For economic organization	Exchange will be facilitated by modes that support adaptive, sequential decision making	Trading requires the support of spontaneous or crafted safeguards

Table 1: Implications of Behavioral Assumptions adapted from Williamson (1988)

3.2 Dimensionalizing

Transaction cost economics relies on the following principal dimensions for purposes of describing transactions: First, the frequency with which transactions occur, second, the degree and type of uncertainty to which they are subject, and finally to the condition of asset specificity, which is measure to which an asset can be redeployed to alternative uses by alternative users without a decline in the productive value of the asset.

3.3 Process Features

Williamson (1988) suggests that while the process aspects are underdeveloped in Transaction Cost Economics, and that it is widely resisted by economists, that process arguments play a prominent role. Specifically the question asked is if conditions of trade evolve over time because of transaction specific investments and incomplete contracts. Table 2 captures the contrast between the incentive and process views of integration.

	Incentive	Process
ownership	ex post decision rights only	assets and ex post decision rights
incentives	HPI	HPI gives way to LPI
controls	unchanged by integration	extended by integration
bureaucratic costs	unchanged by integration	greater under integration
adaptation costs	unchanged by integration	reduced by integration
who acquires whom	matters	negligible

Table 2: Implications of Integration adapted from Williamson (1988)

HPI - High Powered Incentives, LPI - Low Powered Incentives

4 Core Tenets

In their analysis of empirical findings of TCE, David and Han (2004) describe the core tenets of TCE as follows.

Tenet 1. As asset specificity increases, the transaction costs associated with market governance increase

Tenet 2. As asset specificity increases, hybrids and hierarchies become preferred over markets; at high levels of asset specificity, hierarchy becomes the preferred governance form

Tenet 3. When asset specificity is present to a nontrivial degree, uncertainty raises the transaction costs associated with market governance

Tenet 4. When asset specificity is present to a nontrivial degree, increasing uncertainty renders markets preferable to hybrids, and hierarchies preferable to both hybrids and markets

Tenet 5. When both asset specificity and uncertainty are high, hierarchy is the most cost-effective governance mode

Tenet 6. Governance modes that are aligned with transaction characteristics should display performance advantages over other modes; for example, when both asset specificity and uncertainty are high, hierarchy should display performance advantages over markets and hybrids

5 Relationship-Specific Investments

Klein et al. (1978) highlighted the importance of relationship-specific investments, combined with incomplete contracts as the critical factors that lead simple spot market transactions to be affected by transactional issue. They suggest that long-term contracts may be an attractive alternative to spot markets, but that incomplete long-term contracts can run into performance problems too. Therefore, they argue that coordination problems may be less severe if the transaction is internalized, and therefore that vertical integration is more likely to become the preferred governance structure the more important are specific investments by the buyer and the seller.

Williamson (1979) focuses on the structure of vertical relationships between firms independent of internal firm structure and internal labor markets. The contribution of the Williamson (1979) article is in mapping the degree of asset specificity, the importance of uncertainty and the frequency of transactions with the transaction cost minimizing governance structure for transactions. This study makes predictions about the conditions under which firms will respond by moving from spot market contracting to more complex long-term contractual arrangements, and eventually to internal organization (vertical integration).

Asset specificity implies that opportunism problems must be confronted during the execution stage. Since complete contracts are costly to write, monitor and enforce when uncertainty and complexity are higher, incomplete contracts must inevitably serve as an attractive alternative to spot markets. While internal organization also has its own costs, as costs of incomplete long-term contracts increase, internal control is likely to become more attractive.

6 The Vertical Integrating Decision

It is interesting to explore the circumstances under which firms choose to vertically integrate. Empirical literature on vertical integration decision making may be organized to fall within two categories: first, the decision whether to integrate forward into retailing, and second, the decision to backward integrate into input production (also known as the make or buy decision).

6.1 Forward Integration into Retailing

The literature has been concerned with distribution under exclusive dealing (e.g. franchising), rather than common agency (e.g. sale through department or grocery stores). The literature reveals systematic evidence that franchisors and manufacturers rely on independent retailers or franchisees to a greater extent when the effort of the franchisee is more important or when the

operations of the firm are more geographically dispersed. Risks and variability have been seen as being positively correlated with the use of franchising, while a higher value for the brand is seen as positively associated with the tendency to vertically integrate.

6.2 Backward Integration into Input Production

The literature here has been concerned with a manufacturer's decision to integrate with its suppliers. The theories here have built on property-rights as well as on TCE, though TCE has been the prominent theoretical lens used. The literature has established that backward integration is more likely for more complex inputs and when the environment is more uncertain.

7 Empirical Studies on Vertical Integration

We noted in the introduction that Coase's 1937 work had languished till 1972 because of the inability to operationalize the notion of transaction costs. Empirical work also requires several conditions to be satisfied before it can proceed to explore interesting questions. First, a theorization is required where a clear structural relationship is defined between a set of dependent variables and independent variables. Second, these variables need to be measured. Finally, the data should allow for sufficient variation so that theories may be falsified or supported. The direction of empirical work on vertical integration has captured the theme that spot markets fail because asset-specific investments transform the situation from several market participants bargaining ex ante, to a few who are able to negotiate agreements and contracts that may extract quasi-rents ex post. It therefore follows that the dependent variable in vertical integration studies should be the nature of organizational arrangement - spot market, long-term contracts or vertical integration. Since the theory suggests that asset specific investments drive the decision to organizational arrangement, some measure of asset specificity would ideally serve as the primary independent variable. This effect may be moderated by other factors such as internal organization costs, economies of scale, experience and regulation. These may therefore be the other typical independent variables in studies on vertical integration.

Williamson (1983) suggests four distinct types of asset-specific investments: First, site specificity where buyer and seller are geographically colocated so as to minimize inventory and transportation costs. Second, physical asset specificity where one or both parties make investments in equipment and machinery that have lower values in alternative uses. Third, human asset specificity, where human capital is developed by doing (experience). Fourth and final, dedicated assets where investments are made with the specific prospect of selling to a particular customer.

7.1 Early Empirical Work

Monteverde and Teece (1982) seems to have been the first systematic study to empirically examine the role of asset specificity in determining the decision to vertically integrate. Using 133 automotive components that are either produced internally or sourced through the markets by Ford and General Motors, they measure variations in choice of vertical integration as a dichotomous variable. Their hypothesis stated thus "The greater is the application engineering effort associated with the development of any given automobile component, the higher are the expected appropriable quasi rents, and therefore, the greater is the likelihood of vertical integration of production for that component" (Monteverde and Teece, 1982). This study seems to fall into the Williamson (1983) categorization of "human asset specificity". Monteverde and Teece (1982) supports the view that variations in asset specificity affect the choice between vertical integration and market procurement.

Another early work, Masten (1984) explores the make or buy decision by an aerospace firm for the components it had contracted to supply the US government. Masten (1984) found that variations in the importance of asset specificity affects the choice between vertical integration and market procurement. This study therefore provided further evidence that asset specificity considerations are important for understanding the structure of organizations.

Yet another early empirical work was Joskow (1985), where the aspect of site-specificity (Williamson, 1983) is explored in the context of vertical integration by electric utilities into coal production. Joskow (1985) found that vertical integration was more likely for mine-mount plants rather than other coal-burning plants, and that when vertical integration was not chosen, unusually long and detailed contracts were used to support the exchange.

7.2 Recent Empirical Work

While Transaction Cost Economics (TCE) has been and continues to be the predominant theory used to explain vertical integration in firms, recent research has tended to explore certain nuances to TCE in a particular empirical context.

Nickerson and Silverman (2003) are concerned with the question of why and how organizations change. They note that prior organization theorists have, on the one hand assumed an adaptation based view of organizational change where boundedly rational agents shape the organization toward most efficient organization form possible as explained by TCE theory. On the other hand, organization theorists have assumed selection-based approaches that contend that organizations rarely change inefficient governance mechanisms. The authors claim to bridge this divide with this article. With detailed data on large firms in the for-hire trucking industry, the

authors demonstrate that adjustment costs inhibit the extent or organizational adaptation toward the efficient boundary prescribed by TCE theory.

Consistent with TCE theory, Nickerson and Silverman (2003) recognize that costs vary with the characteristics of transactions. The for-hire trucking industry was a well selected empirical context to explore the question of how organizations change because the unexpected deregulation of the industry in 1980 threw up a variation in organizations' responses that helped in dealing with the problem of unobserved heterogeneity that might have otherwise been an issue. Additionally, the for-hire trucking industry also provides two additional advantages given the authors' emphasis on asset specificity in their theorization: First the highly asset specific nature of the LTL (Less-than-truckload) business makes the measurement and exposition of the relationship between asset specificity of assets owned and adaptability much easier to demonstrate. Second, the fact that the for-hire trucking industry had itself another line of business, TL (Truckload) which is not asset specific but similar in other respects helps in providing necessary variance in the data but controlling for other factors that are outside the scope of the particular research question considered.

Novak and Stern (2009) emphasize interdependence among various vertical integration decisions rather than seeing them as separate transactions. By drawing on a detailed data set of the luxury automobile market, Novak and Stern (2009) show support for complementarity in product development, suggesting therefore that contracting complementarity may be particularly important when coordination is important to achieve but difficult to monitor.

David and Han (2004) do an extensive assessment of empirical support for TCE and find that asset specificity found strong support across several studies. On the other hand however, David and Han (2004) find considerable disagreement on how to operationalize some of TCE's other central constructs and propositions, and relatively lower levels of support for uncertainty and performance. Figure 1 shows that asset specificity was quite successful at predicting both make-vs.-buy choice as well as predicting the degree of integration between independent buyers and sellers.

8 Contribution of Theories of the Firm

8.1 Predictive Ability

Williamson (1975) provided for the foundation for a theory of institutional choice, thus yielding clear causal relationships between transactional characteristics and institutional arrangements. As the results from David and Han (2004) indicate, asset specificity has been a strong predictor

	# Tests	# Supported	% Supported	# Counter	% Counter
<i>A. Independent variable</i>					
Asset specificity	107	64	60%	4	4%
Uncertainty	87	21	24%	14	16%
Transaction costs of the exchange	51	23	45%	9	18%
Asset specificity * Uncertainty	30	13	43%	4	16%
Frequency	13	9	69%	0	0%
Degree of integration between buyer and seller	12	8	67%	0	0%
Opportunism	7	5	71%	0	0%
Vertical integration	1	1	100%	0	0%
Total	308	144	47%	31	10%
<i>B. Dependent variable</i>					
Hierarchy vs. market	117	53	45%	9	8%
Degree of integration between buyer and seller	67	40	60%	5	7%
Transaction costs—market exchange	27	11	41%	1	4%
Hierarchy vs. hybrid	26	10	38%	11	42%
Hybrid vs. market	23	6	26%	3	13%
Performance of market	12	8	67%	0	0%
Opportunism	12	6	50%	0	0%
Performance of hierarchy	10	1	10%	2	20%
Performance of hybrid	8	5	63%	0	0%
Greater degree of integration within hierarchy	4	2	50%	0	0%
Transaction costs—hierarchy	2	2	100%	0	0%
Total	308	144	47%	31	10%

Figure 1: Empirical Studies in TCE adapted from David and Han (2004)

of both forward and backward integration decision in firms across time periods.

8.2 Explanatory Power

A salient feature of the theories of the firm discussed in this article is that despite it being a parsimonious theory, it has been successful at explaining firm level phenomena with just the few constructs and assumptions detailed above. This combination of explanatory power with predictive ability makes TCE and the theory of the firm amongst the best constructions to disseminate the learning to researchers and practitioners alike.

8.3 Modelling Assumptions

Williamson (1975) provides that there is wide range of institutional arrangements that can be used to govern transactions between economic agents. The theory suggests that specific institutional arrangements emerge in response to various transactional considerations in order to minimize the total cost of making transactions. While markets and hierarchies are presented as the two primary institutional mechanisms for allocating resources, firms can take on many

different organization structures depending upon market transactions taking on different forms ranging from simple spot market transactions to complex long-term contracts. TCE and the theory of the firm has therefore been a fertile ground to accommodate several different organizational forms not just those of the spot market and of the strict hierarchy.

8.4 Resilience in the presence of other theories

Building on the simplicity of the assumptions of the theory and the flexibility that the modeling has allowed, TCE has been used to explain firm decision making even in environments of high change and complexity. We note that while several studies have adopted a Resource based view of firm advantages or that of Dynamic capabilities to explain firm performance, TCE has remained applicable and robust despite the presence of these other lenses.

9 Further Inquiry

In his Nobel Prize lecture, Coase (1992) pointed out that "there is little doubt that a great deal more empirical work is needed" in TCE. David and Han (2004) suggest that two problems need to be addressed. First, some of the key propositions, including that relating to uncertainty have been loosely interpreted, and second that some key variables such as performance have received little scrutiny. David and Han (2004) also suggest that tests about the effects of governance forms on performance are likely to suffer from self-selection bias. They also suggest that tests of the relationship between asset specificity and governance form are tests of the largest surviving firms thereby indicating a survivor bias. David and Han (2004) also suggest that TCE has not been applied across time, in longitudinal studies. These open up interesting opportunities to further research at the intersection of vertical integration and TCE.

10 Conclusion

In this article we started out with the objective of reviewing the foundations of the theory of the firm and then applying it to vertical integration. We reviewed the theoretical assumptions and constructs of TCE and then reviewed how that was applied to explaining vertical integration. We reviewed selected empirical research on vertical integration in the early stages of the theory as well as more recently, and found that TCE as a theory has remained relevant and the strongest in explaining firm behavior with respect to vertical integration. Recent research has focussed more on the nuances and specifics of circumstances but support for asset specificity as an explanatory

for vertical integration continues to be strong. We conclude with empirical opportunities to further research on firm level phenomena building on the theory of the firm.

Acknowledgments

I owe significant debt to Prof. Murali Patibandla for having introduced and then guided me along the key principles, assumptions and implications of New Institutional Economics to applications in business and strategy.

References

- Coase, R. H. (1992). The institutional structure of production. *The American Economic Review*, 82(4):713–719.
- David, R. J. and Han, S.-K. (2004). A systematic assessment of the empirical support for transaction cost economics. *Strategic Management Journal*, 25(1):39–58.
- Joskow, P. L. (1985). Vertical integration and long-term contracts: The case of coal-burning electric generating plants. *Journal of Law, Economics, & Organization*, 1(1):33–80.
- Klein, B., Crawford, R. G., and Alchian, A. A. (1978). Vertical integration, appropriable rents, and the competitive contracting process. *The Journal of Law & Economics*, 21(2):297–326.
- Masten, S. E. (1984). The organization of production: Evidence from the aerospace industry. *The Journal of Law & Economics*, 27(2):403–417.
- Monteverde, K. and Teece, D. J. (1982). Supplier switching costs and vertical integration in the automobile industry. *The Bell Journal of Economics*, pages 206–213.
- Nickerson, J. A. and Silverman, B. S. (2003). Why firms want to organize efficiently and what keeps them from doing so: Inappropriate governance, performance, and adaptation in a deregulated industry. *Administrative Science Quarterly*, 48(3):433–465.
- Novak, S. and Stern, S. (2009). Complementarity among vertical integration decisions: Evidence from automobile product development. *Management Science*, 55(2):311–332.
- Williamson, O. E. (1975). Markets and hierarchies. *New York*, pages 26–30.
- Williamson, O. E. (1979). Transaction-cost economics: the governance of contractual relations. *The journal of law & economics*, 22(2):233–261.

- Williamson, O. E. (1983). Credible commitments: Using hostages to support exchange. *The American Economic Review*, 73(4):519–540.
- Williamson, O. E. (1985). *The economic institutions of capitalism*. The Free Press.
- Williamson, O. E. (1988). The logic of economic organization. *Journal of Law, Economics, and Organization*, 4(1):65–93.