MAKE, BUY, OR ALLY: A TRANSACTION COST THEORY
META-ANALYSIS

INGE GEYSKENS
Tilburg University

JAN-BENEDICT E. M. STEENKAMP
University of North Carolina at Chapel Hill

NIRMALYA KUMAR
London Business School

Since the publication of Williamson's Markets and Hierarchies, many empirical articles have investigated the tenets of transaction cost theory. Using meta-analytic techniques, we quantitatively synthesized and evaluated transaction cost-based empirical research on organizational boundary (make, buy, or ally) decisions. We found strong support for the theory for both make versus buy and ally versus buy decisions. However, we did not find evidence that asset specificity had stronger predictive power than uncertainty. Hierarchical and relational governance appropriately aligned with transaction dimensions both led to enhanced performance. On the basis of our meta-analysis, we provide directions for future research.

Transaction cost theory has become the predominant theoretical framework for explaining organizational boundary decisions. Like most influential theories, transaction cost theory was not fully developed at the outset. It has been and continues to be refined and reformulated, corrected and expanded, in response to new theoretical and empirical developments. The basic premise of transaction cost theory has its origins in Coase's (1937) classic article, The Nature of the Firm, in which he described markets and hierarchies as alternative governance structures. Coase argued that the choice between markets and hierarchies was determined principally by differences in transaction costs. However, the difficulty of directly measuring transaction costs resulted in Coase's 1937 article being "much cited and little used" (Coase, 1972: 67).

The operationalization problem of transaction cost theory was resolved by Williamson, who demonstrated that testable hypotheses could be developed by associating the relative efficiency of alternative governance structures with observable dimensions of transactions, namely asset specificity, uncertainty, and transaction frequency. Efforts to subject transaction cost theory to empirical testing began shortly after the publication of Williamson's (1975) seminal book Markets and Hierarchies and have continued unabated since then in a variety of disciplines. Now, about three decades later, few other organizational frameworks have been studied for a longer period of time or have been accorded as much scholarly attention as transaction cost theory.

Despite the large number of empirical studies conducted to test transaction cost theory—or perhaps because of this large number—insights from transaction cost applications have not always been cumulative. Two specific challenges hamper researchers who try to build on the existing transaction cost literature. First, there is the breadth of transaction cost research to contend with. Though the extant empirical research has led to important refinements of early versions of the transaction cost framework, many of these refinements have appeared in different disciplines, including economics, organization, law, sociology, marketing, finance, accounting, and operations management. Unfortunately, the lack of integration across these different disciplines has limited their overall impact on the development of transaction cost theory. Second, even within a given discipline, the depth of transaction cost theory has been a problem. The past 30 years have witnessed a veritable explosion of research efforts. In a recent narrative review of the transaction cost literature covering multiple so-
cial science disciplines, Boerner and Macher (2002) identified more than 600 articles that investigate some aspect of transaction cost theory. Transaction cost theory has grown so large in both breadth (number of disciplines) and depth (number of studies within disciplines) that it seems desirable to bring into sharper quantitative focus this diverse and sizable literature.

To date, several authors have presented narrative reviews of transaction cost theory (e.g., Boerner & Macher, 2002; Rindfleisch & Heide, 1997; Shenkar & Klein, 1995). Recently, a review using the vote-counting method, which consists of a tabulation of significant and nonsignificant findings, appeared (David & Han, 2004). Although bringing some degree of coalescence to a broad field, these reviews leave several issues unaddressed. Through this meta-analysis, we aim to contribute to the transaction cost literature in the following ways:

First, to the best of our knowledge, no study has quantitatively synthesized the empirical research over a wide variety of disciplines and studies, although such a synthesis would result in significant gains of inferential power over traditional narrative reviews and vote-counting methods.

Second, the lack of quantitative integration limits a full understanding of the relative importance of the transaction dimensions vis-à-vis governance choice. Researchers have rarely examined all transaction dimensions in the same study or attempted to compare the effects of each.

Third, doubts have been raised as to the adequacy of the explanation of relational governance modes provided by transaction cost theory. Zajac and Olsen (1993) contended that transaction cost theory’s exclusive focus on single-party cost minimization provides little insight into relational governance, which is not only about cost minimization but also about joint value maximization. In addition, Noorderhaven (1994) maintained that transaction cost theory pictures governance structures as the fruits of planned and intentional action, whereas relational governance modes can also have spontaneous origins in the accumulation of trust over time. As a result, the deliberative choice assumed by transaction cost theory may not always be implicated in the case of relational governance modes. This analysis has provoked the question of whether transaction cost theory can be "stretched" to explain relational governance modes.

Finally, the normative implications of transaction cost theory have been questioned (Dyer, 1997; Ghoshal & Moran, 1996), partly because empirical findings have been divergent with respect to the effect of governance choice on performance. The divergent and often inconsistent results may be a consequence of endogeneity bias (Masten, 1993), as the majority of studies have assessed the effect of governance choice on performance without accounting for the fact that firms purposely, as opposed to randomly, choose their governance modes (Hamilton & Nickerson, 2003). Through our meta-analysis, we provide more insight into the relation between governance choice and performance by explicitly incorporating the selection decision.

**TRANSACTION COST HYPOTHESES**

**Transaction Cost Theory and Hierarchical Governance**

The central question of transaction cost theory is whether a transaction is more efficiently performed within a firm (vertical integration) or outside it, by autonomous contractors (market governance). Transactors are assumed to be “boundedly rational” and “risk neutral,” and at least some actors are assumed to be “opportunistic.” The a priori transaction cost theory assumption is that market governance is more efficient than vertical integration owing to the benefits of competition. Transactions within integrated companies may be insulated from competitive pressure and subject to bureaucratic phenomena. However, certain dimensions of transactions raise transaction costs and combine to create “market failure,” making vertical integration more efficient than market governance. These dimensions are asset specificity, uncertainty, and transaction frequency (Williamson, 1975, 1985). According to transaction cost theory, economic organization is an effort to “align transactions, which differ in their attributes, with governance structures, which differ in their costs and competencies, in a discriminating (mainly, transaction cost economizing) way” (Williamson, 1991: 79).

Transaction-specific assets are assets that are tailored to a particular transaction and cannot be easily redeployed outside the relationship of the parties to the transaction. Their idiosyncratic nature gives rise to a safeguarding problem, because market competition will not restrain opportunistic exploitation. The solution to the safeguarding problem identified in transaction cost theory is vertical integration. In contrast to markets, the authority relationships and hierarchical control procedures available through vertical integration are assumed to embody greater safeguarding capabilities.

The second dimension, uncertainty, arises either when the relevant contingencies surrounding an exchange are too unpredictable to be specified ex ante in a contract (there is environmental uncertainty) or performance cannot be easily verified ex
post (there is behavioral uncertainty). The primary consequence of environmental uncertainty is an adaptation problem; that is, difficulties with adjusting agreements raise transaction costs, a problem that can be addressed through hierarchical governance. However, a number of authors have argued that high environmental uncertainty also encourages firms to maintain flexibility, which would argue against hierarchical governance. As Klein noted, “It appears that uncertainty is too broad a concept and that different facets of it lead to both a desire for flexibility and a motivation to reduce transaction costs” (1989: 256). Different ways of splitting the environmental uncertainty construct have been proposed (Klein, 1989; Walker & Weber, 1984). In this study, we adopted Walker and Weber's (1984) classification. They extended Williamson's basic framework by distinguishing between two types of environmental uncertainty: volume uncertainty and technological uncertainty. Their classification has become particularly influential, and our meta-analysis tested its relevance for understanding governance choice.

Volume uncertainty is the inability to accurately forecast the volume requirements in a relationship (Walker & Weber, 1984). When volume uncertainty is high, suppliers experience unexpected production costs or excess capacity, and buyers experience “stock-outs” or excess inventory. Since a firm should be able to coordinate variations occurring in a hierarchically organized production stream more efficiently than variations occurring with market suppliers, volume uncertainty should increase the likelihood of hierarchical over market governance. Technological uncertainty is the inability to accurately forecast the technical requirements in a relationship (Walker & Weber, 1984). Such uncertainty may follow from unpredictable changes in the standards or specifications of components or end product, or from general technological developments. Unlike volume uncertainty, which motivates hierarchical governance to facilitate adaptation, technological uncertainty is managed more efficiently through market governance. By using market governance, firms retain the flexibility to terminate relationships and switch to partners with more appropriate technological capabilities (Balakrishnan & Wernerfelt, 1986), and they avoid being locked into a technology that may become obsolete (Heide & John, 1990).

The effect of behavioral uncertainty is a performance evaluation problem—that is, difficulty in ascertaining ex post whether contractual compliance has taken place. According to transaction cost theory, the general response to the performance evaluation problem is vertical integration. The greater degree of control available through vertical integration is assumed to embody greater evaluation capabilities.

The complete transaction cost framework also includes transaction frequency, although this construct has received limited attention in the transaction cost literature. Transaction frequency refers to the extent to which transactions recur. Williamson (1985) argued that transaction frequency provides an incentive for firms to employ hierarchical governance because the overhead cost of hierarchical governance will be easier to recover for recurring transactions. We did not include transaction frequency in this meta-analysis because of the lack of research including this construct. Rindfleisch and Heide noted in 1997 that transaction frequency had received far less attention in the empirical literature than asset specificity and uncertainty, and this has not changed.

The conceptual and empirical research launched in the shadow of Williamson's (1975) book led to the proposition that asset specificity is the critical determinant of the choice between markets and hierarchies (Williamson, 1985). Indeed, transaction cost theory maintains that “asset specificity is the big locomotive to which transaction cost economics owes much of its predictive content” (Williamson, 1998: 36). Williamson's contention that asset specificity is the critical driver has been supported by much other research, including prior reviews (e.g., David & Han, 2004; Shelanski & Klein, 1995).

Subsequent theoretical extensions have shown that the benefits of vertical integration stem not from ownership or integration per se, but rather from the ability to exercise decision control (Heide, 1994). As Stinchcombe argued, the ability to govern by means of authority is not limited to intrafirm settings, but also can be achieved between firms by means of contractual provisions, which essentially “produce the effects of hierarchies” (Stinchcombe, 1985: 165). Therefore, following Stinchcombe (1985) and Heide (1994), we define hierarchical governance as explicitly based on enforcement by means of legitimate authority, either through an employment relation or a contractual arrangement that provides decision-making authority in certain areas.

Finally, it should be pointed out that the original transaction cost framework prediction is that uncertainty is only problematic in the presence of specific assets. Uncertainty coupled with transaction-specific assets demands hierarchical governance, since uncertainty allows for expropriation when a party’s investment is exposed. Uncertainty without transaction-specific assets favors the market. If asset specificity is absent and thus potential
transaction partners are numerous, a new transaction arrangement can easily be arranged if necessary (Williamson, 1975). The effect of uncertainty on governance choice thus is conditional on asset specificity being present to a nontrivial degree.

In summary, we set out to assess the generalizability of transaction cost theory’s classic framework over a wide variety of available studies and tested the following hypotheses:

**Hypothesis 1.** As asset specificity increases, hierarchical governance becomes preferred over market governance.

**Hypothesis 2.** As volume uncertainty increases, hierarchical governance becomes preferred over market governance.

**Hypothesis 3.** As technological uncertainty increases, market governance becomes preferred over hierarchical governance.

**Hypothesis 4.** As behavioral uncertainty increases, hierarchical governance becomes preferred over market governance.

**Hypothesis 5.** Asset specificity has a greater effect than uncertainty on choices between hierarchical and market governance.

**Hypothesis 6.** The transaction dimensions interact in shaping governance choice.

**Transaction Cost Theory and Relational Governance**

Originally, transaction cost theory focused on the dichotomy between market and hierarchical governance. However, researchers have argued that transaction cost theory overstates the desirability of integration and of explicit contractual safeguards to protect against transaction hazards (Poppo & Zenger, 2002). This view recognizes that in many industries managers engage in collaborative exchanges (Dyer, 1997). That is, relational governance (alliance) may be a viable alternative to hierarchy when the market fails.

Relational governance modes incorporate a large informal component and are therefore not easily legally enforceable. Instead, nonjuridical mechanisms such as mutual dependence, trust, parallel expectations, joint action, and procedural fairness sustain them (Bradach & Eccles, 1989). They are usually open-ended relationships, with no finite or foreseeable termination points (Heide, 1994). The mechanisms through which relational governance mitigates exchange hazards are both economic and sociological. Economists emphasize the rational, calculative origins of relational governance, emphasizing expectations of payoffs from future cooperative behavior that prompt present cooperation (Axelrod, 1984). In this view, if trust arises, it is carefully calculated. Sociologists emphasize shared values and affective feelings that emerge from a history of trustworthy interactions (Uzzi, 1997). Despite differences, both economists and sociologists have argued that repeated exchanges provide information about the cooperative behavior of exchange partners that may allow for informed choices of whom to trust or not trust (Poppo & Zenger, 2002). In addition, both economists and sociologists have argued that relational governance operates as a self-enforcing safeguard: the value of a future relationship is sufficiently large that neither party wishes to renege (Telser, 1980). Thus, there is considerable overlap in the arguments of sociologists and economists concerning relational governance.

Over the past two decades, relational governance modes have become more prominent. This development has motivated transaction cost analysts to incorporate them into transaction cost theory’s explanatory framework. High asset specificity should favor relational governance over market governance since the former is another way (in addition to hierarchical governance) to address the safeguarding problem posed by asset specificity. However, Williamson (1991) argued that relational governance addresses uncertainty less effectively than market governance, since relational adaptations cannot be made unilaterally, as can market adaptations, but require mutual consent. Building consent takes time, which may be in short supply in uncertain environments. Finally, we also tested whether Williamson’s proposition that asset specificity is more important than uncertainty in shaping governance is supported for relational versus market governance.

In sum, we tested the following hypotheses:

**Hypothesis 7.** As asset specificity increases, relational governance becomes preferred over market governance.

**Hypothesis 8.** As uncertainty (volume, technological, and behavioral) increases, market governance becomes preferred over relational governance.

---

1 Although interaction effects have been primarily discussed in the context of hierarchical governance, there is no a priori reason to assume that they are not applicable to the relational governance context. However, we do not state an explicit hypothesis for the relational governance case because we discovered only six interactions between the transaction cost constructs for relational governance (of which four were significant).
Hypothesis 9. Asset specificity has a greater effect than uncertainty on choices between relational and market governance.

Transaction Cost Theory and Performance

Although the previous hypotheses pertain to whether firms follow the prescriptions of transaction cost theory, transaction cost theory is explicitly normative. Firms that follow its prescriptions and align organizational form with transaction dimensions will economize on transaction costs, which in turn should translate into performing better than those who do not (Williamson, 1985). Note that this position is not without its critics. For example, Ghoshal and Moran (1996) expressed substantial skepticism about the normative implications of transaction cost theory because of its strongly self-fulfilling assumptions (for example, the assumption of opportunism can become a self-fulfilling prophecy in that opportunistic behavior increases when hierarchical controls are imposed). Dyer (1997) criticized transaction cost theory’s exclusive focus on minimizing transaction costs as an efficiency criterion as governance may also influence transaction value. In this article, we set out to assess transaction cost theory’s normative value by testing the following hypotheses:

Hypothesis 10. Hierarchical governance appropriately aligned with transaction dimensions leads to enhanced performance.

Hypothesis 11. Relational governance appropriately aligned with transaction dimensions leads to enhanced performance.

EXPLORATORY MODERATOR ANALYSES

In addition to testing the hypotheses described above, we conducted two exploratory moderator analyses. First, we examined differences in the operationalization of the variables in individual studies contributing to our meta-analysis. It has been observed that the various transaction cost constructs are rather broad (e.g., David & Han, 2004). Although broad constructs give a theory flexibility and relevance for a wider range of organizational contexts, they may be so broad as to subsume components whose relationships with other constructs are, although directionally similar, of differing magnitudes. We examined this issue by testing whether the focal relationships proposed in transaction cost theory are different for different components/operationalizations of a particular construct. More specifically, we tested for differences in focal relationships with respect to the following four transaction cost constructs:

1. **Asset specificity.** We compared four constructs, testing site specificity (idiosyncratic investments in facilities) versus physical asset specificity (idiosyncratic investments in equipment) versus human asset specificity (idiosyncratic investments in human capital) versus goodwill asset specificity (idiosyncratic investments in brand name capital) (Williamson, 1991).

2. **Hierarchical governance.** Vertical integration (the performance of a transaction within a firm) was contrasted with formal governance (governance modes characterized by explicit contractual provisions that simulate the effects of organizational hierarchies) (Heide, 1994).

3. **Relational governance.** We contrasted goodwill relational governance (measures emphasizing socially derived norms and social ties that have emerged from prior exchange) and calculative relational governance (measures emphasizing the rational, calculative origins of relational governance, particularly stressing expectations of future exchanges that prompt present cooperation) (Poppo & Zenger, 2002).

4. **Performance.** Cost-inclusive performance (measures that encompass the costs of generating performance, such as profitability) versus cost-exclusive performance (measures that do not directly encompass the costs of generating performance, such as sales performance) was our comparison (Poppo & Zenger, 2002).

Second, we tested the following ten study characteristics as potential moderators of all focal relationships: (1) year of publication, (2) year of sample, (3) publication outlet (top-tier journal versus other source), (4) data source (primary versus secondary data), (5) design (cross-sectional versus longitudinal/experimental), (6) direction of integration (vertical versus horizontal), (7) domain (purchasing versus distribution), (8) country (U.S. versus other), (9) industry type (products versus services), and (10) product/service type (consumer versus industrial).

Because we had no a priori reason to believe that specific operationalizations or study characteristics would lead to stronger relationships, we do not offer directional hypotheses for these moderators.

---

2 Williamson (1991) also identified dedicated assets and temporal specificity. The number of correlations involving these types of asset specificity was too small for them to be included in our moderator analysis.
METHODS

Rationale for Conducting a New Meta-Analysis

To date, six major narrative reviews of transaction cost theory (Boerner & Macher, 2002; Joskow, 1988; Lyons, 1996; Masten & Saussier, 2000; Rindfleisch & Heide, 1997; Shelanski & Klein, 1995) have been presented. Whereas narrative reviews can marshal and summarize work on a particular topic, they allow scope for subjective interpretation. Narrative reviews do not subject the studies they examine to statistical tests. Thus, they cannot estimate whether those who conducted the studies mistook chance results for meaningful ones (and thus reached falsely positive conclusions based on sampling error) or used samples so small that chance factors concealed important results (leading to falsely negative conclusions) (Hunt, 1997; Hunter & Schmidt, 1990). Recently, David and Han (2004) performed the first quantitative review of the transaction cost literature. Our meta-analysis builds on and extends their study in several ways. First, we quantitatively summarize 200 articles in various disciplines, a number that compares favorably with the 63 articles David and Han reviewed. Second, those authors relied on the vote-counting method, which consists of a tabulation of significant and nonsignificant findings. This technique is essentially a dichotomous accept-reject method. Our approach refines David and Han’s study by incorporating degree of support (for instance, collectively five p-values of .06 constitute much stronger evidence of a relationship than five p-values of .45); thus, we are able to provide information on the magnitude of effects. Third, David and Han’s meta-analysis treats all studies alike, taking their authors’ results to be as stated. In contrast, we corrected for the influence of statistical artifacts such as sampling error and measurement error. Fourth, we give extensive attention to the performance implications of governance choice, a topic David and Han (2004) only briefly discussed. A general overview of our meta-analytic procedures follows.

Literature Search

We used four phases of data collection for identifying studies (articles, book chapters, and unpublished reports) as input for our meta-analysis. First, we examined two computerized databases, ABI/INFORM Global and EconLit. Second, manual searches involved 19 major journals of widely accepted scholarly value from 1975 (the year in which Williamson’s Markets and Hierarchies appeared) through 2003.4 In the third phase, we consulted the reference sections of all the articles retrieved in the second phase as well as the reference sections of all major previous reviews of transaction cost theory to identify any studies that we might have overlooked. In the fourth and final phase, we contacted 113 researchers in the area in order to obtain unpublished studies.

We used several decision rules to determine the studies that would be retained for the meta-analysis. First, although transaction cost theory has been applied to explain a variety of problems of economic organization, ranging from marriage to corporate finance, we restricted this meta-analysis to make, buy, or ally decisions. Second, a study had to either (1) report on one or more relationships between these constructs: asset specificity, volume uncertainty, technological uncertainty, behavioral uncertainty, hierarchical (vs. market) governance, relational (vs. market) governance, and performance; or (2) report on one or more relationships between one of the constructs and one or more of these control variables: competitive intensity, environmental munificence, and firm size (see below). Table 1 summarizes our definitions of the constructs. As noted above, transaction frequency was not included, because the number of correlations involving it was too low (more specifically, no correlations were available between transaction frequency and volume and technological uncertainty). Third, a study had to report sample sizes and an additional substantive keyword and at least one methodological keyword (cf. David & Han, 2004). The additional substantive keywords included were “organization,” “governance,” “opportunism,” “rationality,” “integration,” “hierarch*,” “make-or-buy,” “merger*,” “relational,” “cooperation,” “alliance*,” “uncertainty,” “asset specificity,” “transaction-specific,” “performance,” and “Williamson.” The methodological keywords introduced were “data,” “empirical,” “test,” “statistical,” “finding*,” “result*,” and “evidence.”

TABLE 1
Definitions of the Transaction Cost Constructs and Representative Measures

<table>
<thead>
<tr>
<th>Transaction Cost Construct</th>
<th>Definition and Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asset specificity</td>
<td>Construct definition: The degree to which the assets that support a given transaction are tailored to it and cannot be redeployed easily outside of a particular exchange relationship. Includes investments in buildings, equipment, learning, and/or brand name capital that are specific to a particular relationship. Representative measures: Anderson (1985); Heide &amp; John (1990).</td>
</tr>
<tr>
<td>Behavioral uncertainty</td>
<td>Construct definition: The degree of difficulty in verifying whether compliance with established agreements has occurred. Representative measures: Anderson (1985); Poppo &amp; Zenger (2002).</td>
</tr>
<tr>
<td>Hierarchical governance</td>
<td>Construct definition: Governance modes that, by means of an authority structure, provide one exchange partner with the ability to develop rules, give instructions, and in effect impose decisions on the others. Includes vertical integration and formal governance modes that are characterized by explicit contractual provisions that simulate the effects of organizational hierarchies. Representative measures: Parkhe (1993); Walker &amp; Weber (1987).</td>
</tr>
<tr>
<td>Relational governance</td>
<td>Construct definition: Governance modes characterized by the parties to a transaction jointly developing policies directed toward the achievement of certain goals. Representative measures: Carson, Madhok, Varman, &amp; John (2003); Subramani &amp; Venkatraman (2003).</td>
</tr>
<tr>
<td>Performance</td>
<td>Construct definition: Includes “cost-inclusive” performance measures (measures that encompass the costs of generating performance, such as level and growth of profit and abnormal stock returns) and “cost-exclusive” performance measures (measures that do not directly encompass the costs of generating performance, such as level and growth of sales). Representative measures: Dyer (1996); Murray &amp; Kotabe (1999).</td>
</tr>
</tbody>
</table>

outcome statistic (e.g., $r$, univariate $F$, $t$, $\chi^2$) that allowed the computation of a correlation coefficient with the formulas provided by Hunter and Schmidt (1990: 272). We read each article in the final set and extracted data on the variables of interest, including outcome statistics, sample sizes, statistical artifacts, and study characteristics. The first author coded all the articles and categorized all harvested correlations on the basis of the construct operationalizations. We ensured investigator blindness in making determinations that could substantially affect our findings by keeping the methods sections separated from the results sections and by having the papers photocopied in such a way that their origins could not be determined (Orwin, 1994). Intercoder agreement in extracting information from primary studies is an important concern in meta-analysis; it is not a problem for calculation-based coding (e.g., coding effect sizes, sample sizes, and reliabilities) but may be a problem for judgment-based coding, such as categorizing operationalizations into transaction cost constructs (Heller, Watson, & Ilies, 2004). To address this issue, the second and the third authors rechecked those categorizations; interrater agreement was 98 percent, and we resolved remaining discrepancies via discussion and reaching consensus.

Data Set

Nonindependence. We used the following three criteria to ensure an acceptable level of independence among correlation coefficients in our database: (1) For studies with multiple independent samples, correlations from each sample were in-
Meta-Analytic Calculations

Our meta-analyses were conducted via Hunter and Schmidt’s (1990) approach, which allows for the correction of statistical artifacts and thus provides a relatively accurate estimate of the true average strength and variance of a relationship in the population of interest. We corrected the retrieved correlation coefficients (r’s) for the biasing influence of eight statistical artifacts: (1) sampling error, (2) measurement error in the dependent variable, (3) measurement error in the independent variable, (4) dichotomization of a continuous dependent variable, (5) dichotomization of a continuous independent variable, (6) range restriction in a dependent dichotomous variable, (7) range restriction in an independent dichotomous variable, and (8) downward bias in the retrieved correlation coefficient as a measure of the population correlation.

For each data point (i.e., for each r), we had information on artifacts 4–8. Following Hunter and Schmidt (1990), we first corrected each data point for these artifacts. Next, the partially corrected data points were meta-analyzed, a process that yielded a sample-size-weighted mean correlation and variance, both corrected for artifacts 4–8. In this step, the variance in correlations was also corrected for sampling error (artifact 1). Third, we corrected the mean and variance obtained in step 2 for measurement error (artifacts 2–3), using the method of artifact distributions (see Hunter and Schmidt [1990] for details), as this information was not available for all data points. These corrections yielded the following summary statistics for each bivariate analysis of interest to our study: the average corrected correlation (average rho; \( \bar{\rho} \)), the corresponding standard deviation of the corrected correlations in the population (s.d.,p), the 95% confidence interval around the average rho (CIp), and the true residual variance in the observed correlations (s.d.,res) after removal of variance due to artifacts 1–8.

We further examined whether the individual correlations on which the average correlations were based were drawn from the same population using two tests (Hunter & Schmidt, 1990): (1) the 75% rule-of-thumb, stating that looking for moderators is warranted if statistical artifacts explain less than 75 percent of the observed variance in correlations, and (2) Hunter and Schmidt’s chi-square test, a statistical significance test for whether the observed variation is greater than that expected by chance; a significant value suggests the presence of possible moderator variables. Finally, we tested for possible publication bias using the “trim and fill method” (Duval & Tweedie, 2000). Publication bias may occur if studies producing null results or effects opposite those envisaged are less likely to be accepted for publication (Hunter & Schmidt, 1990).

Testing Transaction Cost Hypotheses

Hypotheses 1–11 required that we simultaneously test a system of equations that explained in turn the impact of the transaction dimensions on a firm’s governance choice and the relationship between governance choice and performance. Testing these equations independently would introduce biased estimates due to the endogeneity of governance choice in the performance equation (Hamilton & Nickerson, 2003). For hierarchical (vs. market) governance, we estimated the following system of equations:

Hierarchical vs. market governance

\[
= \beta_1 \text{asset specificity} + \beta_2 \text{volume uncertainty} + \beta_3 \text{technological uncertainty} + \beta_4 \text{behavioral uncertainty} + \epsilon, \quad (1)
\]
**Performance**

\[ \text{Performance} = \beta_h \text{ hierarchical vs. market governance} + \beta_s \text{ size} + \beta_c \text{ competitive intensity} + \beta_e \text{ environmental munificence} + \epsilon. \] (2)

In these equations, we included two industry variables and one firm variable (competitive intensity, environmental munificence, and firm size, respectively) to control for other key factors that may affect performance (Leiblein, Reuer, & Dalsace, 2002). A similar system of equations was estimated for examining how the transaction dimensions affected relational governance choice and how relational governance influenced performance.

Our meta-analytic path analysis required that in addition to estimating the correlations between the predictors and the criterion variable, we also estimate the correlations between all predictors (cf. Geyskens, Steenkamp, & Kumar, 1999). Thus, we also conducted meta-analyses relating all predictors to one another using data from all primary studies in our database providing this information. We corrected correlations between the predictors for statistical artifacts and removed outliers using the procedures described in the previous section.

Our analyses were conducted using the full information maximum likelihood method and LISREL software on the meta-analytically derived correlation matrix. This estimation method explicitly accounted for potential simultaneity bias in the choice of governance mode and performance (Greene, 2000). Given the variability in sample sizes associated with each correlation coefficient in the meta-analytic correlation matrix, we based this analysis on the harmonic mean of the sample sizes comprising each entry in the meta-analytic correlation matrix (Viswesvaran & Ones, 1995) (harmonic mean \( N = 1,265 \) [1.075] for hierarchical [relational] governance).

Since our meta-analytic path analysis was based on a pooled correlation, we could only test relationships for which an adequate number of correlations appeared in the existing literature. Since correlations involving interactions between transaction dimensions were seldom available, we were not able to include interaction effects (Hypothesis 6) between the transaction dimensions in our path model. As an alternative, we employed a vote-counting procedure whereby prior studies are categorized on the basis of the direction and significance of the findings (Bushman, 1994). A nonparametric sign test, which tests the hypothesis that the interaction effects between the transaction dimensions from a collection of \( k \) independent studies are all zero (H\(_0\): \( \pi = .5 \), where \( \pi \) is the proportion of significant interaction effects in the population), was employed.

**Exploratory Moderator Analyses**

For continuous moderators (e.g., year of publication), we correlated the moderators with the corrected correlations (\( \rho \)'s). For categorical moderators (e.g., data source), using the Z-statistic we divided the data into subsets based on the level of the moderator in question and determined whether the average rho statistically differed between subsets (Hunter & Schmidt, 1990).

**RESULTS**

**Bivariate Correlations**

In Table 2, for the three dependent variables examined in this article (hierarchical versus market governance, relational versus market governance, and performance), for each predictor we report: the number of data points (\( k \)), total sample size (\( N \)), the average rho (\( \bar{\rho} \)), the standard deviation of rho in the population (s.d.\( _p \)), and the 95% confidence interval around the mean rho (\( CI_\alpha \)). We also report the true residual variance in the observed correlations (s.d.\( _{res} \)) with variance due to artifacts 1–8 removed, the percentage of observed variance accounted for by artifacts, and Hunter and Schmidt’s (1990) chi-square test for heterogeneity (\( Q \)).

Heterogeneity for the various relations was significant. Absence of heterogeneity would have been preferable for testing Hypotheses 1–11, but did the degree of heterogeneity reported in Table 2 preclude quantitative meta-analyses? Drawing on recent work by Cortina (2003), we argue that this is not the case. Cortina analyzed 1,647 meta-analyses and reported a mean residual standard deviation (s.d.\( _{res} \)) of .122. Our mean residual standard deviation is .129. More importantly, he developed a cut-off value for an acceptable standard deviation of rho in a population (s.d.\( _p \)). According to Cortina (2003), if s.d.\( _p \) is less than .05, then there is absolutely no reason to be concerned about interpretation of effect sizes. On the other hand, if s.d.\( _p \) is larger than .265 (a number Cortina labeled “the point of no return”), then the average correlation must be regarded as uninterpretable because it is a mean of sample values that are no less discrepant than would be values taken from \( k \) populations. The area in between .05 and .265 is a grey zone: analyses can proceed, though a note of caution is in place. The s.d.\( _p \)'s of our focal relationships varied between .10 and
TABLE 2
Meta-Analytic Results for the Focal Relationships

<table>
<thead>
<tr>
<th>Predictor</th>
<th>k</th>
<th>N</th>
<th>( \bar{\rho} )</th>
<th>s.d.( \rho )</th>
<th>Cl( \rho ) 5%</th>
<th>Cl( \rho ) 95%</th>
<th>s.d.( \rho_{res} )</th>
<th>% Variance Accounted For</th>
<th>Q</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predictors of hierarchical</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>governance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asset specificity</td>
<td>75</td>
<td>40,444</td>
<td>.16</td>
<td>.16</td>
<td>.09</td>
<td>.16</td>
<td>.13</td>
<td>10.7</td>
<td>756.6*</td>
</tr>
<tr>
<td>Volume uncertainty</td>
<td>16</td>
<td>3,315</td>
<td>.04</td>
<td>.14</td>
<td>-.03</td>
<td>.09</td>
<td>.10</td>
<td>31.7</td>
<td>50.8*</td>
</tr>
<tr>
<td>Technological uncertainty</td>
<td>7</td>
<td>2,991</td>
<td>-.08</td>
<td>.11</td>
<td>-.13</td>
<td>.01</td>
<td>.08</td>
<td>25.8</td>
<td>27.6*</td>
</tr>
<tr>
<td>Behavioral uncertainty</td>
<td>20</td>
<td>7,238</td>
<td>.13</td>
<td>.15</td>
<td>.04</td>
<td>.15</td>
<td>.11</td>
<td>19.0</td>
<td>109.6*</td>
</tr>
<tr>
<td>Predictors of relational</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>governance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asset specificity</td>
<td>51</td>
<td>11,681</td>
<td>.28</td>
<td>.21</td>
<td>.17</td>
<td>.27</td>
<td>.17</td>
<td>12.6</td>
<td>432.6*</td>
</tr>
<tr>
<td>Volume uncertainty</td>
<td>6</td>
<td>1,430</td>
<td>-.27</td>
<td>.19</td>
<td>-.34</td>
<td>-.08</td>
<td>.15</td>
<td>15.4</td>
<td>41.2*</td>
</tr>
<tr>
<td>Technological uncertainty</td>
<td>7</td>
<td>2,137</td>
<td>-.07</td>
<td>.17</td>
<td>-.17</td>
<td>.05</td>
<td>.14</td>
<td>14.9</td>
<td>47.3*</td>
</tr>
<tr>
<td>Behavioral uncertainty</td>
<td>9</td>
<td>2,121</td>
<td>-.03</td>
<td>.20</td>
<td>-.13</td>
<td>.09</td>
<td>.15</td>
<td>15.1</td>
<td>59.7*</td>
</tr>
<tr>
<td>Predictors of performance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hierarchical governance</td>
<td>22</td>
<td>8,246</td>
<td>.09</td>
<td>.13</td>
<td>.03</td>
<td>.12</td>
<td>.10</td>
<td>20.0</td>
<td>112.0*</td>
</tr>
<tr>
<td>Cost-inclusive performance</td>
<td>16</td>
<td>7,293</td>
<td>.07</td>
<td>.10</td>
<td>.01</td>
<td>.11</td>
<td>.10</td>
<td>25.6</td>
<td>63.3*</td>
</tr>
<tr>
<td>Cost-exclusive performance</td>
<td>10</td>
<td>4,172</td>
<td>.20</td>
<td>.21</td>
<td>.04</td>
<td>.27</td>
<td>.17</td>
<td>19.5</td>
<td>53.2*</td>
</tr>
<tr>
<td>Relational governance</td>
<td>27</td>
<td>5,140</td>
<td>.44</td>
<td>.21</td>
<td>.31</td>
<td>.45</td>
<td>.18</td>
<td>11.8</td>
<td>259.5*</td>
</tr>
<tr>
<td>Cost-inclusive performance</td>
<td>15</td>
<td>3,279</td>
<td>.40</td>
<td>.22</td>
<td>.24</td>
<td>.45</td>
<td>.19</td>
<td>9.4</td>
<td>176.2*</td>
</tr>
<tr>
<td>Cost-exclusive performance</td>
<td>10</td>
<td>1,630</td>
<td>.57</td>
<td>.12</td>
<td>.40</td>
<td>.55</td>
<td>.10</td>
<td>32.2</td>
<td>49.9*</td>
</tr>
</tbody>
</table>

\* \( k \) = number of data points; \( N \) = total sample size; \( \bar{\rho} \) = estimate of corrected population correlation; \( s.d.\( \rho \) \) = estimated standard deviation of corrected correlations in population; \( Cl.\( \rho \) 5\% \) = lower bound of confidence interval for \( \bar{\rho} \); \( Cl.\( \rho \) 95\% \) = upper bound of confidence interval for \( \bar{\rho} \); \( s.d.\( \rho_{res} \) \) = residual standard deviation; \% variance accounted for = percentage of observed variance accounted for by statistical artifacts; \( Q \) = Hunter and Schmidt’s chi-square test for heterogeneity.

\* \( p < .05 \)

.21. Nevertheless, concern about heterogeneity was legitimate, and the Q-statistics for the meta-analyses that populated the pooled correlation matrix indicated that most relationships were moderated by unknown variables.

Assessing Publication Bias
For each focal relationship, we applied the trim and fill method to assess publication bias (Duval & Tweedie, 2000). For 11 out of 14 relationships reported in Table 2, we estimated the number of studies missing because of publication bias to be 0 (using the estimator \( R_0 \); see Duval and Tweedie [2000] for details). For the remaining three relationships (i.e., asset specificity, volume uncertainty, and technological uncertainty with hierarchical governance), the number of studies missing because of publication bias was estimated to be 2, 1, and 3, respectively. Only the last number was significant (\( p < .05 \)). We then imputed the missing values to derive effect-size estimates adjusted for the missing studies, finding the adjusted (original; cf. Table 2) estimates to be .16 (.16), .03 (.04), and -.04 (-.08), respectively. The extent of bias is very modest and did not affect any of our research conclusions reported below. In sum, publication bias did not appear to be a serious issue.

Meta-Analytic Correlation Matrix
Table 3 presents the meta-analytic correlation matrix. Each cell in this matrix represents one individual meta-analysis. Thus, the table is the culmination of 44 individual meta-analyses. Each entry in the matrix contains a sample-size-weighted, average correlation coefficient corrected for artifacts (\( \bar{\rho} \)), the standard deviation of rho (\( s.d.\( \rho \) \)), the total sample size for each correlation (\( N \)), and the number of samples included in each weighted average (\( k \)). Standard deviations marked with a dagger indicate relationships in which moderator variables are present. With the exception of one relationship involving a control variable, all relationships included data from at least two samples (\( N_s = 194-40,444 \)). All focal relationships included data from at least three samples (\( N_s = 576-40,444 \)).5 Note that no sample included all variables

Some relationships in our meta-analytic correlation matrix were based on rather small numbers of samples and sample sizes. Thus, the magnitudes of those relationships should be interpreted with caution. However, it should be noted that the problem of small numbers of samples and small sample sizes is less likely to seriously affect the estimates of average correlations (\( \bar{\rho} \)) than it is to influence estimates of the standard deviation (\( s.d.\( \rho \) \)) of the correlations (Hunter & Schmidt, 1990).
<table>
<thead>
<tr>
<th>Construct</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. Relational governance</td>
<td>5.149</td>
<td>11.661</td>
<td>1.430</td>
<td>2.137</td>
<td>2.121</td>
<td>1.323</td>
<td>378</td>
<td>938</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Performance</td>
<td>0.9 [0.13]</td>
<td>0.44 [0.21]</td>
<td>0.05 [0.17]</td>
<td>2.062</td>
<td>5.645</td>
<td>4.630</td>
<td>16.648</td>
<td>4.571</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Asset specificity</td>
<td>0.16 [0.16]</td>
<td>0.28 [0.21]</td>
<td>0.5 [0.17]</td>
<td>2.062</td>
<td>5.645</td>
<td>4.630</td>
<td>16.648</td>
<td>4.571</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Volume uncertainty</td>
<td>0.04 [0.14]</td>
<td>-0.27 [0.19]</td>
<td>-0.17 [0.05]</td>
<td>-0.09 [0.23]</td>
<td>972</td>
<td>1.521</td>
<td>913</td>
<td>393</td>
<td>1,617</td>
<td></td>
</tr>
<tr>
<td>6. Technological uncertainty</td>
<td>-0.08 [0.11]</td>
<td>-0.07 [0.17]</td>
<td>0.23 [0.24]</td>
<td>0.25 [0.10]</td>
<td>0.02 [0.51]</td>
<td>2.232</td>
<td>5.057</td>
<td>2.666</td>
<td>2.833</td>
<td></td>
</tr>
<tr>
<td>7. Behavioral uncertainty</td>
<td>0.13 [0.15]</td>
<td>-0.03 [0.20]</td>
<td>0.10 [0.20]</td>
<td>0.06 [0.20]</td>
<td>-0.05 [0.24]</td>
<td>0.06 [0.00]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Firm size</td>
<td>0.17 [0.21]</td>
<td>0.07 [0.00]</td>
<td>0.00 [0.17]</td>
<td>-0.01 [0.16]</td>
<td>-0.09 [0.00]</td>
<td>0.05 [0.09]</td>
<td>0.05 [0.00]</td>
<td>194</td>
<td>4,403</td>
<td></td>
</tr>
<tr>
<td>9. Competitive intensity</td>
<td>0.06 [0.10]</td>
<td>0.07 [0.00]</td>
<td>-0.01 [0.23]</td>
<td>0.00 [0.07]</td>
<td>0.01 [0.00]</td>
<td>-0.01 [0.03]</td>
<td>-0.13 [0.00]</td>
<td>0.03 [0.00]</td>
<td>3.258</td>
<td></td>
</tr>
<tr>
<td>10. Environmental munificence</td>
<td>0.01 [0.07]</td>
<td>0.15 [0.06]</td>
<td>-0.13 [0.09]</td>
<td>0.01 [0.11]</td>
<td>-0.29 [0.14]</td>
<td>0.05 [0.00]</td>
<td>0.06 [0.00]</td>
<td>0.00 [0.16]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* The triangle to the lower left of the diagonal contains the average corrected correlations ($\hat{r}$s), and the standard deviations (s.d.'s) of the corrected correlations are in parentheses. The triangle to the upper right of the diagonal shows the total sample sizes (Ns) and the number of samples (k's, in parentheses) from which the means were derived. The dagger symbol (\textdagger) indicates cases where the chi-square test was significant, suggesting the existence of moderators.
of interest.Thus, the number of samples contributing to each meta-analytic correlation is far fewer than the total number of samples.

Transaction Cost Theory and Hierarchical Governance

Table 4 presents the results for the path analysis. Since path analysis controls for redundancy in measures of independent variables, our analysis provides a more precise test of our hypotheses than could be obtained by bivariate analyses. Asset specificity ($\beta = .19, p < .01$), volume uncertainty ($\beta = .07, p < .01$), and behavioral uncertainty ($\beta = .13, p < .01$) led to a choice of hierarchical governance over market governance, supporting Hypotheses 1, 2, and 4. As predicted by Hypothesis 3, technological uncertainty increased the likelihood of market governance ($\beta = -.14, p < .01$).

Whereas volume uncertainty had a weaker impact on governance choice than asset specificity ($\Delta \chi^2[1] = 10.96, p < .01$), the effects of technological uncertainty ($\Delta \chi^2[1] = 2.88, p = .09$) and behavioral uncertainty ($\Delta \chi^2[1] = 2.77, p = .10$) were not significantly different in magnitude from the effect of asset specificity. Moreover, the joint effect of uncertainty (volume, technological, and behavioral) on governance choice was larger than that of asset specificity ($\Delta \chi^2[1] = 12.56, p < .01$). Thus, Hypothesis 5 was not supported.

Ten out of 21 studies in our meta-analytical database that reported interaction results found significant interaction effects between the transaction dimensions in their models of hierarchical governance choice. The sign test described above yielded an estimate of $\pi$ (of .48 (10/21), which corresponded to a cumulative probability of .5 from a binomial table. Thus, we could not reject the null hypothesis that the interaction effects between the transaction dimensions were all zero. In addition, it is plausible that 10 out of 21 was the upper limit for the number of studies finding interaction effects, as some of the studies in our meta-analytical database might have tested for interaction effects but not reported nonsignificant findings (the so-called file drawer problem). The preponderance of empirical evidence therefore suggested that the transaction dimensions worked primarily through main effects and did not support Hypothesis 6.

Transaction Cost Theory and Relational Governance

Supporting transaction cost theory’s predictions, asset specificity ($\beta = .29, p < .01$) motivated relational governance over market governance, but vol-

TABLE 4

Results of Analyses of Transaction Cost Hypotheses\(^a\)

<table>
<thead>
<tr>
<th>Predictors</th>
<th>Hierarchical Governance</th>
<th>Performance for Hierarchical Governance</th>
<th>Relational Governance</th>
<th>Performance for Relational Governance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transaction dimensions</td>
<td>$\beta$</td>
<td>$p$</td>
<td>$\beta$</td>
<td>$p$</td>
</tr>
<tr>
<td>Asset specificity</td>
<td>.19</td>
<td>&lt; .01</td>
<td>.29</td>
<td>&lt; .01</td>
</tr>
<tr>
<td>Volume uncertainty</td>
<td>.07</td>
<td>&lt; .01</td>
<td>-.24</td>
<td>&lt; .01</td>
</tr>
<tr>
<td>Technological uncertainty</td>
<td>-.14</td>
<td>&lt; .01</td>
<td>-.14</td>
<td>&lt; .01</td>
</tr>
<tr>
<td>Behavioral uncertainty</td>
<td>.13</td>
<td>&lt; .01</td>
<td>-.05</td>
<td>.01</td>
</tr>
<tr>
<td>Governance choice</td>
<td></td>
<td></td>
<td>.10</td>
<td>&lt; .01</td>
</tr>
<tr>
<td>Hierarchical governance</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relational governance</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control variables</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Competitive intensity</td>
<td>-.02</td>
<td>&gt; .10</td>
<td>-.04</td>
<td>&gt; .10</td>
</tr>
<tr>
<td>Environmental munificence</td>
<td>.15</td>
<td>&lt; .01</td>
<td>.04</td>
<td>&gt; .10</td>
</tr>
<tr>
<td>Firm size</td>
<td>-.02</td>
<td>&gt; .10</td>
<td>-.03</td>
<td>&gt; .10</td>
</tr>
<tr>
<td>Harmonic mean N</td>
<td>1,265</td>
<td></td>
<td>1,075</td>
<td></td>
</tr>
<tr>
<td>$\chi^2$ (7)</td>
<td>162.65</td>
<td></td>
<td>278.07</td>
<td></td>
</tr>
<tr>
<td>GFI</td>
<td>.97</td>
<td></td>
<td>.95</td>
<td></td>
</tr>
<tr>
<td>RMSR</td>
<td>.05</td>
<td></td>
<td>.06</td>
<td></td>
</tr>
</tbody>
</table>

\(^a\) A positive beta coefficient ($\beta$) signifies that the higher the corresponding variable, the more likely it is that hierarchical or relational governance will be chosen. Negative coefficients indicate that higher values of the corresponding variable encourage market governance.
Transaction Cost Theory and Performance

The empirical results provide support for Hypotheses 10 and 11. The coefficient estimates for the governance choice–performance relationship were positive and highly significant, for both hierarchical (β = .10, p < .01) and relational governance (β = .44, p < .01), which indicated that choosing hierarchical or relational governance in response to transaction hazards increases performance.

Exploratory Moderator Analyses

We tested whether the focal relationships proposed by transaction cost theory were different for different operationalizations of four transaction cost constructs and ten study characteristics, conducting a total of 84 analyses. With respect to hierarchical governance, with one exception no significant differences were found for vertical integration as opposed to formal governance, the exception being the relationship with volume uncertainty. However, for both vertical integration and formal governance, the correlation with volume uncertainty was positive (ρ = .02 for vertical integration versus .15 for formal governance). The average corrected correlations (ρ’s) were significantly (p < .05) different for the relationships between governance choice and cost-exclusive versus cost-inclusive performance, suggesting that the how performance is measured moderates the strength of relations (hierarchical governance: ρ = .07 for cost-inclusive performance versus .20 for cost-exclusive performance; relational governance: ρ = .40 for cost-inclusive performance versus .57 for cost-exclusive performance). We therefore reestimated our system of equations, estimating Equation 2 separately for cost-inclusive and cost-exclusive performance. We found that the effect of governance mode (hierarchical or relational) on performance did not differ significantly (p < .05) by type of performance. The moderators pertaining to different operationalizations of asset specificity and relational governance were not significant.

With respect to the study characteristics, 13 out of 71 effects were significant (p < .05). We discovered, however, no systematic pattern in the significant effects, either in terms of the moderator being tested, or in terms of the relationship being analyzed. Thus, the study characteristics we examined could not adequately explain the heterogeneity in effect sizes.

DISCUSSION AND DIRECTIONS FOR FUTURE EMPIRICAL RESEARCH

Despite what almost 30 years ago may have appeared to be insurmountable obstacles to acquiring the relevant data, today transaction cost theory stands on a remarkably broad empirical foundation. Through this meta-analysis, we have contributed to the transaction cost literature in the following ways: First, by psychometrically meta-analyzing the results from 200 empirical papers on transaction cost theory—papers that all concerned the make, buy, or ally decision and that represented 209 independent samples and 557 correlations—we were able to test a comprehensive model that no individual study has tested. Second, we found no evidence supporting the superior predictive power of asset specificity over uncertainty, in contrast to previous reviews that have supported Williamson’s contention that asset specificity is the critical driver of governance choice. Third, we showed that the transaction cost approach can be used to explain relational governance modes, a result that differs from David and Han’s (2004) conclusion that transaction cost theory is less effective in predicting the choice of relational governance over market governance than in predicting the choice of hierarchy over markets. Fourth, we rebutted critics and found strong support for the normative implications of transaction cost theory.

Although we endeavored to overcome depth and
breadth limitations of previous reviews of the transaction cost literature, limitations to our coverage remain. First, we only included studies related to make, buy, or ally decisions. Second, the databases used (ABI/Inform and EconLit) did not contain all relevant studies. Third, we used only studies whose results could be converted to correlation coefficients. Fourth, we did not specifically attempt to include dissertations on transaction costs. We did find, however, that a number of the published articles were based on dissertations. Despite these limitations, we have performed the largest quantitative review of transaction cost theory as it applies to interorganizational relationships in order to take stock of what is known, answer some persisting questions, and point out directions for future research.

**Hierarchical Governance**

Williamson originally intended transaction cost theory to predict the choice between markets and hierarchies. Our meta-analysis shows that the theory performs well with respect to main effects, provided refinements in conceptualizing uncertainty are incorporated. We found that asset specificity, volume uncertainty, and behavioral uncertainty promote a choice of hierarchical governance over market governance. In contrast, in the face of technological uncertainty, market governance is preferred over hierarchical governance.

In two areas, however, Williamson’s claims were not supported. First, we were unable to demonstrate superior predictive power for asset specificity. Thus, Williamson's contention that asset specificity is the “locomotive” was unsubstantiated. A future area of empirical investigation might be to assess the conditions under which asset specificity may have greater impact than uncertainty.

Second, Williamson argued for the interactive effects of the transaction dimensions rather than the main effects. However, empirical testing has primarily concentrated on the main effects. Only 21 studies reported testing for interactions, and 10 of these reported significant effects. Thus, we were unable to reject the null hypothesis. These results must be considered in light of the reasonable assumption that some (other) studies probably did test for interactions but did not report insignificant results. There are at least two possible reasons for why interaction effects were not found. First, the search for empirical evidence of interaction effects may have been disappointing because of the low power of interaction effects (McClelland & Judd, 1993). Second, the transaction dimensions may operate primarily through main effects. Future research should examine which is true by, for example, using experiments in which asset specificity and uncertainty are manipulated by using role-playing scenarios.

Enormous empirical attention has been paid to the asset specificity construct. Empirical investigations of asset specificity in the future will fruitfully deepen researchers’ understanding only if they go beyond simply demonstrating a main effect into the more complex unresolved questions such as distinguishing between transaction cost reasoning and alternative arguments. Different theories sometimes make the same predictions but specify quite different underlying motivations. For example, does asset specificity motivate hierarchical governance because it renders contract formation troublesome and potentially exposes firms to opportunistic behavior, or do firms integrate because such internalization facilitates the coordination of asset-specific activities in their quest for sustainable competitive advantage (as in knowledge-based theories)? In a similar vein, do firms integrate in the presence of transaction-specific assets to economize on the costs resulting from planning and adapting contracts because of a continuing trading relationship, or do they integrate to reduce resource dependence? Given the congruity of several antecedents of governance choice over competing theories, in future empirical work researchers should attempt to gain insight into the underlying mechanisms driving governance decisions, by measuring the managerial motivations mediating the relationships between transaction dimensions and governance mode chosen. Schilling and Steensma (2002) have taken a step in this direction.

According to Williamson, transaction frequency makes hierarchical governance more likely because the overhead costs of hierarchical governance should be easier to recover for recurring transactions. Unfortunately, the available correlations relating to transaction frequency were too few to include this construct in our meta-analysis. Transaction frequency is deserving of greater empirical attention.

**Relational Governance**

Our meta-analysis showed strong support for the application of transaction cost theory to relational governance. In fact, the variance explained was higher than for hierarchical versus market governance. This higher variance is probably seen because relational governance is usually measured as a perceptual variable, while the choice of hierarchical versus market governance is usually measured with a secondary data indicant. Thus, there is sub-
stantially less shared method variance between the independent and dependent variables when researchers examine hierarchical (vs. market) governance.

Our test examined relational versus market governance choices. We were unable to model relational versus hierarchy choices because there were simply too few correlations between relational governance versus hierarchical governance and the constructs in our model. We encourage future research to further pursue this area.

Following Williamson, empirical researchers have considered market, hierarchical, and relational governance as discrete and mutually exclusive governance modes rather than as complements. To a large extent, the validity of this approach depends on what one considers a transaction. Often, when a firm outsources a process, the decision is made that while a third party will perform part of the contract, the firm will keep a small percentage in-house to retain learning. This decision allows for a more informed (as one knows the true costs of doing the activity and the challenges) and balanced (as a threat, one can take the process in-house) negotiation when the contract is up for renegotiation. The exploration of governance modes as complements is a fertile area for empirical work.

Although we found that each of the three dimensions of uncertainty rendered relational governance a less preferred option, this distinction may be conditional upon the network in which a particular dyadic exchange is embedded. More specifically, the negative effects of uncertainty on relational governance attest to relational governance’s potential dark side in that it may lock firms into unproductive relationships or preclude partnering with other viable firms. By forging multiple alliances, each of which is viewed as a “real option,” firms can isolate themselves from such lock-in effects (Gulati, Nohria, & Zaheer, 2000). They may maintain their flexibility by using alliances as a portfolio of options. As such, relational governance may become a more suitable alternative than market governance in the face of high volume or technological uncertainty if it is embedded in a network that allows a firm to flexibly use different specific kinds of expertise and production facilities from the various firms in the network (Hage & Alter, 1997).

Behavioral uncertainty is also expected to have a differential effect on relational governance in the context of a network. More specifically, one would expect a positive effect of behavioral uncertainty on relational governance, since the social mechanisms of reputation and collective sanctions safeguard specific exchanges in the network by dispersing information about behavior and increasing the costs of malfeasance (Jones, Hesterly, & Borgatti, 1997). Collective sanctions (i.e., group members punishing other members who violate group norms, values, or goals) reduce behavioral uncertainty by increasing the costs of opportunism and by decreasing the costs of monitoring to any individual party. Similarly, reputations—which have economic consequences for participants in networks—reduce behavioral uncertainty by providing information about the reliability and goodwill of others.

Future empirical research should move transaction cost theory beyond a dyadic focus and apply a network lens. Although the transaction cost perspective stresses the efficacy benefits gained from reducing the governance cost of a transaction, a network approach allows consideration of the benefits from optimizing not just a single relationship but a firm’s entire network of relationships.

Transaction Cost Theory and Performance

We found support for the performance implications of transaction cost theory. Choosing hierarchical or relational governance in response to transaction hazards increases performance. We further found that the effect of relational governance on performance was substantively larger than that of hierarchical governance. This result may be attributable to at least three not mutually exclusive causes. First, there is the fact that relational governance measures, as opposed to hierarchical governance measures, often confound a governance decision per se with the quality of that decision (e.g., achieving a collaborative exchange). Another plausible explanation is the uniqueness of relational governance as a governance mechanism because it not only minimizes transaction costs but also creates value in the exchange relationship through superior information sharing (Dyer & Chu, 2003). A third explanation is possible bias due to shared method variance. In general, the relational governance studies in our review employed multi-item perceptual measures to tap both relational governance and performance. In contrast, common methods were less of a problem for hierarchical governance. The studies in our review measured hierarchical governance using “objective” dichotomous measures (e.g., a dummy variable representing whether hierarchical or market governance had been chosen), whereas they measured performance using either perceptual measures or secondary data indicants. Future research should investigate which cause or causes give rise to the larger effect of relational governance. To disentangle the effect
of shared method variance from that of substantive causes, longitudinal studies using recent advances in structural equation modeling are needed (Baumgartner & Steenkamp, 2006).

Although our meta-analysis supports the assertion that transaction cost considerations are important drivers of governance choice, most of the studies included in our meta-analysis were limited to documenting whether firms followed transaction cost descriptions rather than to examining how firm decisions impacted performance. The studies that did examine performance differences between governance modes usually failed to appropriately correct for the selective adoption of governance modes. Consequently, with a few exceptions, this research ends up answering the question “How does the performance of firms that adopt a particular governance arrangement compare to that of firms that adopt alternatives to that arrangement?” when the correct question, from a governance-choice perspective is “How does the performance of a firm that adopted a particular arrangement compare with how that same firm would have performed had it adopted an alternative?” (Masten, 1993: 124). Greater effort to understand the influence of governance choice on performance is needed.

Related to the above performance question, our study looked at the performance implications of alignment and showed that hierarchical and relational governance appropriately aligned with the transaction dimensions led to enhanced performance. What we did not study was the cost of misalignment. Some promising research emerging in this area has suggested that the performance implications from deviations of optimal alignment may be asymmetric (Leiblein et al., 2002; Sampson, 2004).

Also, a more complete understanding of governance choice requires a greater sensitivity to the interdependence of production and exchange relations. If hierarchies are superior to markets for reasons of efficiency, this superiority may well be due not just to transaction cost reductions but also to productivity-enhancing factors tied to superior skills and knowledge. Production cost measurements should always be comparisons of in-house and supplier costs. Unfortunately, of the studies in our meta-analysis, only a few have measured production costs in this way. As a result, the number of correlations involving the appropriate production cost measure and the other constructs in our models was not sufficiently large to warrant its inclusion.

Methodological Findings

Our examination of the measures used in transaction costs research led to several observations. Often, the empirical studies were limited by single-item survey measurement and proxy measures. The use of proxy measures was most apparent in the studies that investigated transaction frequency. Some studies have treated frequency as synonymous with business size, while others have measured frequency as an indicator of production costs. However, neither of these captures the meaning Williamson afforded the construct.

The literature is further complicated by the presentation of measures without a description of psychometric properties, rendering it difficult to assess the quality of measures and/or their relationships to other variables. In addition, the majority of studies analyzed were correlational, and causal interpretations should therefore be made cautiously. Notable exceptions include the study by Ohanian (1994), who estimated changes in the pattern of vertical integration over time. Other notable exceptions are Silverman, Nickerson, and Freeman (1997) and Nickerson and Silverman (2003), studies that respectively explore the organizational failure and adaptive actions over time of inappropriately aligned organizations; both studies take advantage of the natural experiment associated with a deregulatory shock (which ensures unusually high environmental change and consequently a need for organizational change). To remedy these problems, future research should use multi-item scales, provide adequate data on reliability, and examine longitudinal data.

In sum, transaction cost theory is well established and empirically corroborated. Yet, for all its depth and scope, transaction cost theory has only begun to explore the variety and complexity of organizational forms. There is still much to learn. It is hoped that further research will explore some of the above issues in more detail.

REFERENCES*


---

*a References marked with an asterisk indicate studies included in the meta-analyses.


*Gulati, R., & Nickerson, J. A. 2003. *Interorganizational trust: the choice of make, buy, or ally; and the performance of interorganizational relationships in the U.S. auto industry*. Working paper, Kellogg School of Management, Northwestern University, Evanston, IL.


Lyons, B. R. 1996. Empirical relevance of efficient con-


Williamson, O. E. 1991. Comparative economic organization: The analysis of discrete structural alterna-


Inge Geyskens (i.geyskens@uvt.nl) is an associate professor of marketing at Tilburg University. She received her Ph.D. from the Catholic University of Leuven. Her research interests include interorganizational relationships and meta-analysis.

Jan-Benedict E. M. Steenkamp (jbs@unc.edu) is the C. Knox Massey Professor of Marketing and Marketing Area Chair at the Kenan-Flagler Business School, University of North Carolina at Chapel Hill. He received his Ph.D. from Wageningen University. His current research focuses on interorganizational relationships, global marketing, marketing research methods and techniques, and the effectiveness of marketing instruments.

Nirmalya Kumar (nkumar@london.edu) is a professor of marketing, the director of Centre for Marketing, and a codirector of the Aditya V. Birla India Centre at London Business School. He received his Ph.D. from Northwestern University. His current research interests include marketing strategy, interorganizational relationships, and branding.