INTERNALIZATION OF ADVERTISING SERVICES: TESTING A THEORY OF THE FIRM

By

Birger Wernerfelt[1], Alvin Silk[2] and Shuyi Yu[3]

November 19, 2020

[1] MIT Sloan School of Management, bwerner@mit.edu
[2] Harvard Graduate School of Business Administration, asilk@hbs.edu
[3] MIT Sloan School of Management, shuyiyu@mit.edu
Support from the MIT Sloan School and the Division of Research at Harvard Business School is gratefully acknowledged. We thank Marta Stiglin for assistance in developing and pretesting our questionnaire. We benefitted from comments by the Editor, an Associate Editor, two referees, Gustavo Manso, and John van Reenen, as well as audiences at the 2019 SICS, the 2019 Marketing Science Conference, Yale, the MIT Organizational Economics Lunch, and the MIT Marketing Seminar.

Abstract

In 1956, a group of trade associations representing publishers and independent advertising agencies signed a consent decree aimed at ending a set of trade practices that for half a century effectively precluded advertisers from owning and operating in-house agencies. Since then, large firms have internalized more and more of the services formerly performed by external agencies, perhaps as many as half. We use this phenomenon to test a theory of the firm, thereby simultaneously offering an explanation for it. The theory suggests that firms should internalize activities for which their competitive position implies (1) that it is more important for human capital to be firm-specific as opposed to function-specific and (2) that frequent modifications are desirable. It also predicts (3) that these two effects reinforce each other. This is the first paper to report on a test of the specialization hypothesis and we find that it is robustly significant in a cross-sectional dataset covering nine different agency activities in 79 firms. In addition to the cross-sectional test, we informally present some time-series data suggesting that both specialization and frequency have grown over time along with the level of internalization.

I. INTRODUCTION

Theories of the firm aim to tell us which functions firms out-source and which they perform in-house. It is notoriously hard to test these theories because most functions are either universally internalized or universally outsourced. Because advertising services in the past several years have increasingly become partially internalized by many large firms[4], they now offer a promising setting in which to test theories of the firm. The setting is even more attractive because the management of these services is of independent interest to marketing scholars.

The theory (Wernerfelt 2015, 2016) trades off the relative advantages of specialization in the function versus firm domain against the costs of more or less frequent adaptation. [5] Specifically, two forces bear on internalization: (1) Employees can specialize in a single firm, while workers in external agencies typically work on campaigns for several firms but can focus on a narrower function, and (2) it is comparatively less costly to implement changes inside the firm.

The first force is stressed by Chris Grams, author and founder of New Kind: "In my experience the most valuable aspect of having an internal creative/branding team is how deeply connected the team can be to the company, its mission, values, and culture." And: "Probably the second most important benefit of having an internal creative team working on your brand is the depth of understanding the internal team may have about the company" (New Kind, 2018). The second of these forces is remarkably similar to that highlighted by Bill Duggan, Group Executive Vice President in the Association of National Advertisers (ANA): "[In-house agencies] are closer to their internal brand groups and internal decision groups and

[4] The magnitude of the phenomenon varies both over time and between different types of firms (Horsky, Michael, and Silk 2012). Here we investigated a sample of US firms with more than 1000 employees and found that every single one of them internalized to some degree, averaging 58% of all services. So advertising agencies have lost a lot of business.

[5] The earliest incarnation of the theory was Wernerfelt (1997), so it was not invented to fit the phenomena studied here.

have those conversations in the hallway rather than setting it up via phone and email“ (ANA 2012, p. 21). Many other articles explicitly stress the same two forces[6], while others highlight phenomena that could interpreted as indicators of them.[7]

We confront the theory with two types of data. We first test the theory on a cross-sectional data set obtained from a survey of a convenience sample of large US consumer marketing firms. We exploit the differences between advertisers and the effect of these differences on nine activities typically performed in developing and implementing an advertising campaign. Consistent with the theory, the data robustly suggests that more of an activity is internalized if the advantages of firm specialization are large relative to those of functional specialization. Since the specialization effect has not been tested before, this finding is the main contribution of the paper. Our findings about needs for more frequent change and the interaction between that and the specialization effects are more fragile in the sense that they are significant in regressions without firm fixed effects, but not in regressions with them.[8] In the second part of the empirical investigation, we use archival time-series data from several sources to suggest that the extent of integration has risen recently in concert with the relative value of firm versus function specialization and the importance of fast adaptation. Because these series are very short and chosen ex post, we do not draw any conclusions from them, but simply present the patterns as stylized facts about the changes underway in the vertical structure of the U.S. advertising and marketing services industry.

While we are unable to establish a causal test of our theory, we provide suggestive evidence in light of correlational data. Our hope is to provoke future empirical work that more deeply explores the ideas suggested here.

---

[6] For example, Anheuser-Busch InBev's US CMO, Marcel Marcondes, has recently explained the motivation behind their in-house agency as "a desire to … connect consumer data to the creative work and have more agility…" (Liffreing, 2019). This could presumably apply to other complementary factors as well.

[7] One such example is a recent survey in which IAB found that activities with real-time optimization are likely to be internalized, while those requiring large specialized investments are likely to be outsourced. (www.iab.com/2019-european-programmatic-in-housing)

[8] Since the frequency variable has been highly significant in other studies, it is quite surprising that is does not have a significant within firm effect in this dataset.

I.1 Theories of the Firm and Tests in Economics and Marketing

Theories of the firm compare the relative efficiency of firm versus arms-length exchange and most fall in one of two categories: Theories based on "ex ante investment distortions", notably the "Property Rights Theory" (PRT) due to Grossman and Hart (1986) and Hart and Moore (1990) look at the indirect costs that are incurred when fear of later hold-up cause the parties to shrink from making appropriate partner-specific investments. This theory looks for the optimal allocation of property rights to productive assets and defines an employee as someone who works with assets owned by another player. It suggests that the assets should be owned by the advertiser if her specific investment is more important than the agency's. It is not a direct mapping, but PRT arguably predicts internalization when it is more important that human capital be firm-specific rather than task-specific, just like our theory. However, it does not predict that this effect should be stronger when changes are more frequent nor that frequency by itself should lead to more internalization.

The theory has faced critiques on several fronts: There are examples in which employees own some of the assets they work with (e. g. craftsmen), it seems unnatural to have the same theory predict both asset ownership and employment, and the role of hold up seems overblown. (That is, while employment is ubiquitous, hold up is relatively rare.)[9] There is also a theoretical argument due to Maskin and Tirole (1999), but it is clear that many members of the profession, while they accept the theoretical point, are willing to look past it.

Theories relying on differences in "ex post adaptation costs" (e. g. Coase, 1937; Wernerfelt, 1997, 2015, 2016; Bajari and Tadelis, 2001; Hart and Moore, 2008; Dessein, Galeotti, and Santos, 2016) focus on the direct and indirect costs of adapting the services/products that are traded, and the prices of them, to

[9] See Hart and Moore (2008) for further critiques.

changing circumstances. Theories in this category use the more conventional definition of an employee as someone who takes orders and define a firm by an employer/employee dyad.[10]

The predictions tested here only come into play when adaptations are repeated. (1) Because employment involves the firm getting repeated services from the same worker, it is more efficient when firm-specific specialization has a bigger productivity effect. (Without firm specialization a sequence of spot market transactions would be better.) (2) Since the variable costs of adaptation under employment are lower than in other contracting regimes, more frequent adaptations make employment more attractive. Almost all other models in this literature, including Bajari-Tadelis (2001) and Hart-Moore (2008), look at single period adaptations, implying that these effects do not appear in them. While this practice is motivated by a desire to keep the models as simple as possible, it is worth noting that all actual incidents of employment involve multiple periods/adaptations. So it stands to reason that multi-period effects play a large role in their functioning.

Different versions of "Transactions Cost Economics" (TCE) due to Williamson (1979, 1985) use arguments from both the ex ante and ex post categories and invoke several non-standard assumptions. Since the theory, in its most general form, has not been formalized, it is not clear which assumptions are necessary for which results to hold. (In fact, the formal theories discussed above are sometimes referred to as "formalizations of aspects of TCE".) TCE makes a number of predictions about the use of firms, for example when productive assets are specific, when complete contracts are hard to write, when the future is uncertain, and when transactions are frequent. The latter prediction is obviously relevant to the current study, although it typically is part of the argument that if someone has to make a specific investment that cannot be protected with a contract, then that party should have power (be the party that integrates). Of course, the firms do not make any investments in our model and we make no attempt to measure any.

[10] Matouschek (2004) is an exception.

Theories of the firm have been tested in many contexts, and many of these tests, particularly those related to TCE have been outside economics, in areas like organization studies, strategy, marketing, etc. Unfortunately, the vast majority of these tests have looked at the same prediction: that more specific investments correlate with more integration. (e.g. Monteverde and Teece, 1982 and Anderson and Schmittlein, 1984). The specific asset hypothesis is consistent with both PRT and theories based on ex post adaptation costs. Indeed, since agents within firms trade with each other on a repeated basis, any theory about firms needs an exogenous or endogenous lock-in mechanism. (It is endogenous in our theory, but often treated as exogenous in tests of TCE.)

More recent empirical work in economics has been put in the context of PRT rather than TCE (e. g. Woodruff, 2002, Antras, 2003**,** and Baker and Hubbard, 2004**)**. However, as is apparent from the survey by Lafontaine and Slade (2007), most of the tests continue to look at specific investments and could still be interpreted from a TCE perspective. The adaptation cost theory has been tested by Forbes and Lederman **(**2009); Costinot, Oldenski, and Rauch (2011); and Novak and Wernerfelt (2012), among others.

While the theory tested here is about backward integration, the bulk of empirical work in marketing has sensibly focused on the forward integration. This literature is surveyed by Rindfleisch and Heide (1997), who find that almost all the studies look at specific assets, though some also look at uncertainty (a proxy for the need to adapt, e. g. Anderson, 1985), and a few ask whether economies of scale enjoyed by outside producers are so large that internal production is inefficient.[11] An exception is Horsky (2006), who models the cost differentials among alternative sourcing policies as a function of an ad hoc set of explanatory variables culled from the extensive trade literature relating to the advantages and disadvantages of the set of alternative "architectures".

[11] Anderson and Weitz (1986), John and Weitz (1988), Dutta, Bergen, Heide, and John (1995), and Leiblein and Miller (2003) are four representative papers in this very large literature. Geyskens, Steenkamp and Noorderhaven (2006) perform a meta-analysis of it.

The effects suggested by the theory tested in the present paper are different from, but related to, the effects tested in the marketing literature. The relative advantages of business vs. service specialization could be seen as a comparison between two different kinds of specific assets and the frequency of change is obviously related to uncertainty. So while somewhat similar tests have been performed, ours is the first paper, in any field, to test the specialization effect. Table 1 summarizes the results from empirical studies that provide supporting evidence for different theories of the firm.[12]

Table 1: Empirical Studies on Factors Affecting Firm Internalization

Author	Year	Industry	Data/Technique	Variable Examined	Effect
Adaptation Cost Theory					
Novak and Wernerfelt	2012	Automobiles	Cross Section; Structural Estimation	Adaptation Frequency	+
Costinot et al.	2011	Multiple	Cross Section; OLS	Adaptation Frequency (Task Routineness)	+
Forbes and Lederman	2009	Airline	Cross Section; Nested Logit	Adaptation Frequency (Schedule Adjustments)	+
Property Rights Theory (PRT)					
Baker and Hubbard	2004	Trucking	Panel; First Dif., IV	Contractibility (Monitoring Technology)	+
Woodruff	2002	Footwear	Cross Section; Probit	Importance of Worker's Investment	-
Transactions Cost Economics (TCE) tests of specific asset hypothesis					
Anderson	1985	Electronic	Cross Section;	Transaction-Specificity of	+

[12] Only seminal papers that are most relevant to our study are listed here for PRT and TCE. See Lafontaine and Slade (2007) for an extended survey.

		Components	Logit	Knowledge, and its Interaction with Uncertainty	+
Anderson and Schmittlein	1984	Electronic Components	Cross Section; Logit	Transaction-Specificity of Knowledge	+
Monteverde and Teece	1982	Automobiles	Cross Section; Probit	Amount of Firm-Specific Knowledge	+

I.2 Research on the Sourcing of Advertising Services

In spite of their widespread use, the recent literature on internal advertising agencies is very sparse. One of the few systematic studies is due to Horsky (2006). Dividing advertising services into "creative" and "media", she compares five formats: Both can be out-sourced to a single agency, they can be out-sourced to two different agencies, both can be internalized, or one can be performed in-house while the other is out-sourced. Horsky compares these in terms of their expected cost for the advertiser, taking into account a very large set of factors including market power, double marginalization, economies of scale and scope, coordination costs, advantages of specialization, etc. The model is estimated on data for 1,746 firms listed in the 2000 edition of the *Advertiser and Agency Redbook* with 1999 advertising budgets of $1 million or more. The present study differs in two main ways: Our predictions rely on a theory of the firm and our empirical test breaks advertising services into nine categories and focuses on sets of particular tasks or activities rather than bundling issues.[13]

In a more recent study, Silk and Stiglin (2016 and the references cited therein) use information about members of the In-House Agency Forum (IHAF), to argue that the core competencies underlying internal agencies are efficiency, adaptability, and corporate influence. Our study tests for, and finds evidence of, the first two of these. On the efficiency point, Silk and Stiglin suggest that today's internal agencies are

[13] Specifically, we consider three tasks (creative, production, and media buying) for digital, video, and print media.

able to compete more successfully for top talent than in the past, when they were widely presumed to be seriously disadvantaged in attracting and retaining personnel required to match the "expertise, specialization, and objectivity" available in external, full service agencies.[14] We build on this line of reasoning by noting that the ability to hire reflects the relative productivity of talent in external versus internal environments.

I.3 Plan of the Paper

The theory is presented in Section II, followed by the cross-sectional test in Section III and a historical account in Section IV. We conclude with a discussion in Section V.

II. THEORY

II.1 Intuition

Consider the relationship between a firm and its advertising agency (AA). The staff in an internal AA will be very familiar with the firm, its brands, and its strategy. The flip side of this is that external AAs, while they typically work for several clients, can specialize more narrowly, and will have access to all the most expensive specialized technologies and know how. So the decision to internalize also depends on the relative importance of firm-specific versus the function-specific expertise.

Another concern is flexibility. The ideal campaign/media plan changes all the time: Competition does something different, a news story draws positive or negative attention to various market trends, new data emerges, etc. However, change is not free, and depending on how costly it is, only some adjustments are made. The difficulty of change depends on whether the AA is external or internal. If it is external, changes require negotiation and re-contracting; but if the head of the AA is an employee of the firm, he or

[14] A similar phenomenon has recently been observed in the legal profession where many top lawyers leave professional partnerships to join corporate law departments.

she can simply be told what to do. So the dollar and time cost per change is smaller if the AA is internal and more changes will be implemented in such agencies. On the other hand, there are non-trivial fixed costs involved in setting up an internal AA and these will only be justified if frequent changes are sufficiently important.

The model has many equilibria but depending on parameter values, one of three is most efficient: A "Market" in which workers specialize in individual services (e. g. plumbing) and in each period match with a different business in need of that service, "Sequential Contracting" in which workers specialize in individual businesses and in each period negotiate a new contract with that business agreeing to perform a specific service the business needs in the period, and "Employment" in which workers also specialize in individual businesses but in the first period negotiate a blanket contract agreeing to follow orders and perform any service the business asks for with no new negotiation. The gains from specialization are the same in Employment and Sequential Contracting, and the former can only be more efficient if there are economies of scale in contracting such that the big blanket contract is cheaper to negotiate than a steady stream of one-service agreements.

There are undoubtedly many sources of ex post adaptation costs, including time, uncertainty, the dis-utility of bargaining, legal fees, etc., and no single micro-foundation can capture all of them. In Wernerfelt (2015) players are initially bargaining under two-sided incomplete information but can, and will want to, expend resources to learn their opponent's valuation. We do not offer any such micro-foundation here, but simply represent bargaining costs by a couple of parameters.

II.2 Simple Model

A formalization of the above intuition was introduced in Wernerfelt (1997) and further developed as the "Adaptation Cost Theory" in Wernerfelt (2015, and Ch. 5 of 2016). We here give a reduced form version of the model in Wernerfelt (2015) and with the purposes at hand in mind, state the key result.

Businesses, each operated by a different entrepreneur, produce by using workers to perform services.[15] Each of the B businesses, $b \in \boldsymbol{B}$, needs one of the S services, $s^t_b \in \boldsymbol{S}$, in each period t, and if a needed service s is performed by a worker w for business b, it results in q_{wbs} units of output. Any of the W workers can costlessly perform any service, but only for one per period and output cannot be expanded by using more than one worker or by performing an unneeded service. There are two periods, and $\delta > 0$ is both the weight on second period payoffs and the rate at which needs change.[16] All players are risk-neutral, financially unconstrained, and at all times fully informed about the history of the game.

The vector of needs in the period *1*, s^1, is commonly known ex ante, but needs in period *2*, s^2, are only revealed at the start of that period. For convenience, we assume that the distribution of s^2 is such that the total number of businesses needing each service is the same as that in s^1 and that no business needs the same service in both periods.

After observing s^1, workers choose human capital profiles by costlessly acquiring more or less narrow business and service skills.[17] In the service domain, human capital may be focused on an individual service or services in general, and similarly in the business domain, human capital may be focused on an individual business or businesses in general. So w's human capital is summarized in his profile $(h_{wB}, h_{wS}) \in \{\{b\}_{b \in \boldsymbol{B}}, \boldsymbol{B}\} \times \{\{s\}_{s \in \boldsymbol{S}}, \boldsymbol{S}\}$.

Human capital is reflected in productivity in the natural way. A worker is more productive if he is narrowly invested and works in his area of specialization, but lower if he works outside his area of expertise. The advantages of specialization are given by $q_{wbs} = q_B > 1$ if $(h_{wB}, h_{wS}) = (b, \boldsymbol{S})$, $q_{wbs} = q_S > 1$ if $(h_{wB}, h_{wS}) = (\boldsymbol{B}, s)$, and $q_{wbs} = Max\{q_B, q_s\}$ if $(h_{wB}, h_{wS}) = (b, s)$ versus $q_{wbs} = 1$ if $(h_{wB}, h_{wS}) = (\boldsymbol{B}, \boldsymbol{S})$. If w works outside his area of specialization, such that $h_{wB} = b' \neq b$ or $h_{wS} = s' \neq s$, then $q_{wbs} = \underline{q} < 1$. To

[15] To keep things simple, we assume that labor is the only factor of production. For physical assets, the analog of employment is ownership. If the advertiser owns an asset, say some equipment used to produce ads, she can use it any way she wants without having to negotiate with the agency.
[16] Changes happen once per period and if periods are long, the discount factor δ will be small.
[17] We do not model how this happens, it could be by education or experience.

eliminate equilibria in which the worker simply maximizes first period productivity, we assume that the importance of the second period and a worker's loss from working outside his area of expertise are sufficiently large. Formally:

Assumption 1: $\delta > (Max\{q_B, q_S\} - 1)/(1 - \underline{q})$

Trades are governed by mechanisms. If we think of markets as examples, the idea is that a mechanism provides a forum in which workers and businesses meet and arrive at a price for labor. More formally, each worker can enter one mechanism per period, and each business can be entered, by the entrepreneur who operates it, in one mechanism per period. A mechanism in period t specifies two sets: $(m^t_B, m^t_S) \subseteq \boldsymbol{B} \times \boldsymbol{S}$, and by entering, a worker agrees to perform any $s \in m^t_S$ for any $b \in m^t_B$ in exchange for a price determined by the number of workers and businesses who enter the mechanism. Similarly, the entrepreneurs agree to choose one $s \in m^t_S$ and pay the price to any worker who performs it. Before trading, businesses and workers are matched randomly within each mechanism. To rule out complete contracting, we assume that a mechanism can produce agreement on exactly one price per period, though prices agreed to in period *1* bind the parties for period *2* as well. That is, two parties to a period *1* agreement do not need to enter a period *2* mechanism as long as they both want to trade for a service/business pairing that is covered by the original agreement.

A mechanism clears iff it is entered by the same number of workers and businesses, and in that case the price is such that all participants (workers and entrepreneurs) get positive payoffs, independent of the workers' human capital. If a mechanism does not clear, all players on the long side get negative payoffs, and since they can guarantee themselves zero payoffs by not entering any mechanism, all mechanisms clear in all equilibria.

Reflecting the idea that larger markets are more efficient, participation in mechanisms is assumed to be costless, except when only one entrepreneur enters. In such cases, we make the non-standard, but

reasonable, assumption that the parties face costs of bilateral bargaining.[18] We assume that both sides incur some bargaining costs and that these total $K(\left| m^t_S \right|)$ per worker with $K(0) = 0$. The key assumption is that $K(\,)$ is sub-additive.[19] In fact, to allow us to work with a two-period model, we will assume that $K(1) < K(2) = K(S) < 2K(1)$. This assumption, which is weaker in models with more periods, means that it is cheaper to negotiate once over a single price for any of S services than to negotiate twice for prices on two different services.[20] Note, however, that the assumption leaves open the possibility that $K(2) > (1 + \delta)K(1)$, in which case it is cheaper to negotiate new one-service contact once per period than making a broader agreement up front. To eliminate equilibria in which players take on extra production costs in order to save on bargaining costs, we assume that the latter costs are "small", such that

Assumption 2: $Min\{K(S), (1 + \delta)K(1)\} < (1 + \delta)(1 - \underline{q})$.

The notation is summarized in Table 2 below.

Table 2: Summary of Model Notation

Symbol	Meaning
$b \in \boldsymbol{B}, B$	A generic business, the set of businesses, the number of businesses
$s, s^t_b \in \boldsymbol{S}, S$	A generic service, the service needed by b in period t , the set and number of services
s^1, s^2	The vectors of needs in periods *1* and *2*
w, W	A generic worker, the number of workers

[18] Coase (1937), Williamson (1985), and Hart and Moore (2008) make similar assumptions.
[19] While this is an unusual premise, it is not unreasonable: Most people prefer not to bargain, but if they have to, would rather bargain once over a $300 pie than 30 times over $10 pies. Consistent with this, Maciejovsky and Wernerfelt (2011) report on a laboratory experiment in which bargaining costs are found to be positive and sub-additive.
[20] If we generalize the model to allow complete contingent claims contracts, those would have bargaining costs $SK(1)$, and the sub-additivity would guarantee that $K(S) < SK(1)$.

δ	The discount factor
(h_{wB}, h_{wS}).	The business and service specializations of w
q_{wbs}, q_B,, q_S,, $\underline{q}$	Production if w performs s for b, production if w is specialized in a business (service) and works for (delivers) it, production if w works outside his area of specialization
(m^t_B, m^t_S)	Set of businesses and services covered by a mechanism in period t.
$K(\lvert m^t_S \rvert)$	Cost of a bilateral mechanism delivering a price for $\lvert m^t_S \rvert$ services

We are now ready to present the extensive form. The sequence of events is as follows:

0. Business needs for period *1* are realized and publicly observed.
1. Workers choose their publicly observed human capital profiles
2. Workers and businesses select mechanisms for period *1* and are randomly matched within each. Workers perform the service needed by the business with which they are matched.
3. Business needs for period *2*, s^2, are realized and publicly observed.
4. If necessary, workers and businesses select mechanisms for period *2* and are randomly matched within each. Workers perform the service needed by the business with which they are matched.
5. All payoffs, net of any bargaining costs, are distributed.

We will say that a worker is an employee if he negotiates with a single entrepreneur in the first period and the contract gives the entrepreneur the right to choose one of several second period assignments for the worker (give an order) with no additional negotiation, such that $(m^2_B, m^2_S) = (\emptyset, \emptyset)$. If a worker enters $(m^t_B, m^t_S) = (b, s^t_b)$ in both periods, he is a contractor, and he is a market worker if $(m^t_B, m^t_S) = (\mathbf{B}, s)$. Given these definitions and assumptions, we can prove:

PROPOSITION: There are many subgame perfect equilibria in the region defined by *Assumptions 1* and *2*,[21] but if we follow the literature and selects the most efficient one, it is always one of three:

Employment: $(h_{wB}, h_{wS}) = (b, \mathbf{S})$, $(m^1_B, m^1_S) = (b, \mathbf{S})$, and $(m^2_B, m^2_S) = (\emptyset, \emptyset)$, with (net two period) production $(1 + \delta)q_B - K(S)$,

Sequential Contracting: $(h_{wB}, h_{wS}) = (b, \mathbf{S})$, $(m^1_B, m^1_S) = (b, s^1_b)$ and $(m^2_B, m^2_S) = (b, s^2_b)$, with production $(1 + \delta)q_B - (1 + \delta)K(1)$, and

Market: $(h_{wB}, h_{wS}) = (m^t_B, m^t_S) = (\mathbf{B}, s)$, with production $(1 + \delta)q_S$.

Proof: We will proceed in two stages, first proving that Employment, Sequential Contracting, and the Market all are subgame perfect equilibria, and then that no other equilibria can be more efficient than the best of the three.

To the first point, define the most efficient of the three as a focal equilibrium. When players decide which mechanisms to select, any deviation would put the deviator on the long side of the chosen mechanism and thus not pay off. If a worker deviates and chooses a different human capital profile, the best he can hope for is that an entrepreneur makes a corresponding change at the mechanism choice stage. However, it will not pay off to become doubly specialized since this at most can yield $Max\{q_B, q_S\} + \delta\underline{q}$ which, by *Assumptions 1* and *2* is less than what he can get in the focal equilibrium. Furthermore, any other deviation, even if matched by an entrepreneur, will yield less than the focal equilibrium (by the definition of the latter).

To the second point, there are many other equilibria, but all of them involve mixtures of workers and entrepreneurs playing as in Employment, Sequential Contracting, and the Market. (To see this, note that any other strategies will involve human capital profiles in which workers are double specialists or double

[21] Intuitively, the multiplicity is a consequence of the assumption that the long side in mechanisms that do not clear get very low payoffs.

generalists and thus get lower payoffs.) Any mixture of the three equilibria will have weakly lower payoffs than one of the three.

Q.E.D

By comparing the payoffs, we immediately get:

KEY RESULT: Employment is comparatively better when the advantages of business specialization are larger than those of service specialization ($q_B - q_S$ is larger), and when changes are more frequent (δ is larger), and the two effects amplify each other:

$$(1 + \delta)q_B - K(S) > (1 + \delta)q_B - (1 + \delta)K(1) \Leftrightarrow (1 + \delta)K(1) > K(S)$$

$$(1 + \delta)q_B - K(S) > (1 + \delta)q_S \Leftrightarrow (1 + \delta)(q_B - q_S) > K(S)$$

Intuitively, market solutions work comparatively better when firms are similar, and services are different. Similarly, changes are implemented either by orders (cheap) or by negotiations (less cheap), so more of them make internalization more attractive. The interaction effect reflects that fact that productivity advantages are reaped every period, while the costs of establishing an employment relationship are incurred only once. The second prediction was highly significant in the tests of Klein (1989), Forbes and Lederman (2009), Costinot et al (2011), and Novak and Wernerfelt (2012), but the first and last have not been tested before.

III. CROSS-SECTIONAL EVIDENCE

III.1 Challenges to Data Collection

The advertising industry offers a near-ideal setting in which to test the theory: Production is based almost entirely on human capital and we observe wide variations in the degree to which firms integrate different types of services. However, to gather the relevant data poses a major challenge. There are no data bases publicly available with the detailed intra-firm data needed to capture the complexity and variety of sourcing policies marketing organizations employ, let alone the frequency of change and the relative importance of business- vs service-specific human capital. Moreover, it is very difficult to collect relevant primary data via a survey design. The preferred organizational informants tend to be high level managers who have many other demands on their time and may harbor concerns about preserving the proprietary nature of the information sought.[22] We made several efforts to overcome these barriers through unsolicited emails or third party representatives, but the results were uniformly disappointing. We therefore decided to use paid respondents recruited by the online market research aggregator Qualtrics. While this seems to be the only feasible means of gaining access to a convenience sample of qualified informants, it is obviously not ideal, in particular because it prevents us from finding multiple informants from each firm.

III.2. Operationalizing the Key Theoretical Constructs in the Context of Advertising Services

The theory predicts that an activity is more likely to be internalized if it is affected by more frequent changes. These changes can be due to actions taken by competitors, media, or customers, to technological developments, or simply to changes in the firm's own strategy and tactics. In the context of a survey, the frequency of change is observable by a suitably qualified organizational informant and we measured that construct by asking a direct question to capture that experience.

The second prediction, that an activity is more likely to be internalized if the advantages of business specialization are greater than those of service specialization, is more difficult to operationalize by direct

[22] For a discussion of the use of informants in research on marketing and other organizations, see Phillips (1981) and Phillips and Bagozzi (1986) and Bagozzi, Yi, and Philips (1991).

questions. There are at least two problems. First, because we would presumably not observe the counterfactual, and second, because the cost-adjusted productivity of small groups of employees is not a directly observable quantity. However, a good indirect measure of relative productivity is to ask about a manager's ability to hire employees of the same quality as those working in external agencies. To see why this works, make the standard assumptions that workers are more likely to take the highest offer and that both internal and external agencies make offers equaling marginal productivities. In such a labor market, internal agencies will find it easier to hire as workers become more productive when employed by them. This implies that managers' perceptions of their ability to hire, which are quite accessible, is an indirect measure of the relative advantages of business versus service specialization. So the often-voiced complaint "that it is hard to hire good people for internal agencies", reflects the historical situation in which such workers were much more productive in external agencies. On the other hand, if managers of internal agencies now feel that they can hire more or less the same people as external agencies, it implies that the productivity is catching up.

A possible problem with the above measure is that hiring difficulties could be more pronounced in small cities. To take this into account, we therefore control for the population of the city in which firms' headquarters are located.

III.3. Data

As mentioned above, data were collected from a sample of paid informants recruited by Qualtrics from several actively managed market research panels. No single panel had enough qualified respondents, but the questionnaire was administered to panel members judged as "most likely" to meet our targeting criteria. Each of these was then asked to answer a number of screening questions before administering the main questionnaire.

Qualtrics checks every IP address and uses a sophisticated digital fingerprinting technology to certify that panel members are authentic human subjects. Detailed psycho-demographic profiles of panel members are available, and each panelist enters their information during initial registration and subsequently can update that information when signing-in to participate in a study. To ensure that profiles are consistently updated, panels set an expiration date for each profiling question. Each panel has its own confirmation procedures including, but not limited to: TrueSample, Verity, SmartSample, USPS verification, and digital fingerprinting.[23] Panel partners verify respondent's mailing address, demographic information, and email address.

As mentioned above, the questionnaire begins with four qualification questions. Subjects had to be (1) senior managers employed in the (2) marketing organizations of companies that (3) advertise to consumers and (4) have more than 1000 employees. We paid for 100 respondents who met all four criteria and received the data in several small batches over a few weeks. The first few respondents caused us to make some revisions to the questionnaire and could therefore not be used with those we got later. For example, we originally used a three-point scale to measure the amount of integration but since we found that it produced too little variance, we changed to a continuous scale. The analysis that follows is thus based on 79 responses, many from well-known companies.[24]

After pretesting in personal interviews, we divided agency activities into nine categories denoting nine different "*Services Mixes*;" three "functions" (creative, production, and media buying) for each of three "media" (digital, print, video). Our survey instrument asked the following three questions about these nine categories: (1) "Approximately what percentage of the work do you do in-house?" Informants responded on a continuous scale from 0-100. (2) "How frequently do circumstances change such that you find it desirable to change what is done?" Informants responded on a five-point scale from "very

[23] Each of these branded services employ a series of techniques to verify respondents self-reported data and ensure that each respondent is unique and consistent over time.

[24] The last change involved revising the third qualification question from "sells to consumers" to "advertises to consumers". Completed questionnaires were returned by four pairs of respondents who worked for the same four companies. We dropped a randomly chosen member of each pair.

infrequently" to "very frequently". (3) "How do the corporate capabilities available to an internal agency compare to those available to external agencies?" Informants responded on a five-point scale from "very different" to "very similar".[25] In principle, the latter question does not distinguish between cases in which external agencies can hire better talent and cases in which internal agencies can do so. However, in all the interviews and trade press articles we went through, no one ever suggested that internal agencies had strictly superior talent. So we will interpret lower scores (less similar talent) as indicating that labor productivity is lower in internal agencies, and thus the advantages of firm specialization are higher than when scores are higher (more similar talent). Our pretests and interviews suggested that the specific formulation in (3) was easy to understand and reply to.

We also asked, but did not require, subjects to identify their employer, and in the case of 42 informants, we were able to determine the city in which company headquarters are located and thus its population (from the 2016 Census). Finally, we attempted to collect data on sales, advertising, and R&D, but were not able to obtain such information those for more than 23, 10, and 3 of our companies, respectively.

III.4 Descriptive Statistics

The descriptive statistics are given in Tables 3 and 4 below.

Table 3: Means (Standard Deviations), N = 79

	Share In-House (in Percentages)	*Adjustment Frequency*	*Similarity of Available Capabilities*
Service Mix Category			
Digital Creative	61.5 (28.9)	3.68 (1.26)	3.27 (1.27)
Digital Production	55.5 (31.8)	3.34 (1.14)	3.24 (1.22)
Digital Media Buy	51.9 (34.8)	3.32 (1.30)	3.13 (1.26)
Print Creative	61.1 (32.4)	3.38 (1.10)	3.46 (1.29)

[25] It would obviously be desirable to have multiple questions for each construct, but since the informants were senior managers we were warned against making the questionnaire too long.

Print Production	57.1 (31.5)	3.19 (1.28)	3.48 (1.18)
Print Media Buy	55.5 (33.7)	3.10 (1.39)	3.43 (1.12)
Video Creative	62.7 (28.0)	3.49 (1.27)	3.25 (1.22)
Video Production	58.3 (30.7)	3.37 (1.08)	3.15 (1.26)
Video Media Buy	54.5 (36.0)	3.39 (1.24)	3.30 (1.30)

Notes: Share In-House is on a continuous scale from 0-100. Adjustment Frequency is on a five-point scale from 1 ("very infrequently") to 5 ("very frequently"). Similarity of Available Capabilities is on a five-point scale from 1 ("very different") to 5 ("very similar"). Standard deviations are in parentheses.

The means reveal several unexpected phenomena. First, the average level of integration is very high; over 50% in all nine categories. Second, the mean Shares In-House and Adjustment Frequencies for digital are not significantly larger than those for print and video, suggesting that the recent explosion in internalization is driven by more than just the growth in digital advertising. Third, the creative functions are more integrated than production (paired t-test=3.75, p=.000)[26] and media buy (paired t-test=3.75, p=.000). Consistent with this, the creative functions require more frequent adjustment than both production (paired t-test=4.15, p=.000) and media buying (paired t-test=3.86, p=.000). Fourth, advertisers find it easier to compete for comparable talent in print media than in both digital (paired t-test=3.59, p=.000) and video (paired t-test=2.76, p=.003). Finally, all firms internalize something and only about 12% of the Shares In-House are 0 or 1. Hybrid solutions are much more common.

Table 4: Pearson Correlation Coefficients, N = 79

Panel A: Correlations for *Share In-House (in Percentages)*

	1	2	3	4	5	6	7	8	9
1. Digital Creative	1.00								
2. Digital Production	0.77	1.00							
3. Digital Media Buy	0.55	0.60	1.00						
4. Print Creative	0.66	0.61	0.49	1.00					
5. Print Production	0.76	0.74	0.59	0.77	1.00				
6. Print Media Buy	0.52	0.58	0.89	0.51	0.65	1.00			
7. Video Creative	0.63	0.63	0.58	0.46	0.62	0.57	1.00		
8. Video Production	0.58	0.72	0.57	0.43	0.64	0.58	0.85	1.00	
9. Video Media Buy	0.56	0.59	0.86	0.46	0.68	0.90	0.63	0.65	1.00

[26] One-tailed two-sample paired t-tests are performed in this section because we compare two correlated samples from the same population.

Panel B: Correlations for *Adjustment Frequency*

		1	2	3	4	5	6	7	8	9
1.	Digital Creative	1.00								
2.	Digital Production	0.75	1.00							
3.	Digital Media Buy	0.75	0.67	1.00						
4.	Print Creative	0.57	0.59	0.56	1.00					
5.	Print Production	0.58	0.64	0.64	0.79	1.00				
6.	Print Media Buy	0.52	0.59	0.70	0.69	0.75	1.00			
7.	Video Creative	0.66	0.67	0.66	0.61	0.60	0.61	1.00		
8.	Video Production	0.54	0.65	0.63	0.58	0.64	0.57	0.79	1.00	
9.	Video Media Buy	0.61	0.62	0.77	0.58	0.60	0.69	0.66	0.77	1.00

Panel C: Correlations for *Availability of Capabilities*

		1	2	3	4	5	6	7	8	9
1.	Digital Creative	1.00								
2.	Digital Production	0.79	1.00							
3.	Digital Media Buy	0.69	0.70	1.00						
4.	Print Creative	0.53	0.56	0.40	1.00					
5.	Print Production	0.51	0.63	0.52	0.67	1.00				
6.	Print Media Buy	0.55	0.54	0.75	0.43	0.55	1.00			
7.	Video Creative	0.51	0.51	0.37	0.42	0.48	0.38	1.00		
8.	Video Production	0.65	0.65	0.53	0.46	0.48	0.45	0.76	1.00	
9.	Video Media Buy	0.59	0.59	0.70	0.32	0.43	0.63	0.57	0.74	1.00

Four interesting relationships emerged from an inspection of the correlation matrix. First, firms tend to do a lot or very little of all their media buying in-house as indicated by the high correlations (.89, .86, and .90, Panel A). Second, the creative and production functions for each of the three media involve either frequent or infrequent adjustment (.75, .79, and .79, Panel B). Third, in case of digital media, the three types of functional capabilities all tend to be either easy or difficult to obtain by internal agencies (.79, .69, and .70, Panel C). Fourth, the correlations between the Adjustment Frequencies are very high, averaging .65. This suggests that some firms pursue strategies that require frequent changes in all aspects of advertising, while other firms have fewer changes across the board. In the clothing industry, purveyors of teenage fashions and men's underwear might be two examples. About 2/3 of the variance in Adjustment Frequencies is between -, rather than within, firms.

III.5. Main Results

We look at a number of linear regression analyses in Table 5 below. In the main model, (6), we estimate the *Share In-House* for firm *i* of *Service Mix* category *j* as:

$$\textit{Share In-House}_{ij} = \alpha + \beta_{AC}\, \textit{Similarity of Available Capabilities}_{ij}^{27} + \beta_{AF}\, \textit{Adjustment Frequency}_{ij} + \beta_{IN}\, \textit{Adj Freq}_{ij}\, x\, \textit{Sim of Cap}_{ij} + \beta_{CS}\, \textit{Log (City Size)}_{i} + \varepsilon_{ij}.$$

where β_{AC} captures the effect of *Similarity of Available Capabilities* on the percentage of work done in-house, β_{AF} captures the effect of *Adjustment Frequency*, β_{IN} captures the interaction between those two effects, and β_{CS} shows the effect of *City Size*. Our theory predicts that $\beta_{AC}>0$, $\beta_{AF}>0$, and $\beta_{IN}>0$.

Table 5: Linear Regressions Models

Independent Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Adjustment Frequency	3.89***		0.90	1.36	3.11**	3.86***	3.52**	2.90*	0.48
	(0.96)		(1.03)	(1.02)	(1.42)	(1.46)	(1.49)	(1.49)	(1.52)
Similarity of Available		7.57***	7.20***	7.35***	7.40***	7.68***	7.89***	7.09***	3.79***
Capabilities		(0.93)	(1.03)	(1.01)	(1.40)	(1.40)	(1.40)	(1.42)	(1.40)
Adj Freq x Sim of Cap				2.94***		1.92**	2.16**	2.07**	-0.75
				(0.68)		(0.97)	(0.98)	(0.98)	(1.01)
Log *(City Size)*					1.24	1.42*	1.41*	1.05	-2.42
					(0.85)	(0.85)	(0.85)	(0.96)	(2.41)
Category Fixed Effects	No	No	No	No	No	No	Yes	Yes	Yes
Industry Fixed Effects	No	No	No	No	No	No	No	Yes	No
Firm Fixed Effects	No	No	No	No	No	No	No	No	Yes
Observations	711	711	711	711	378	378	378	378	378
Adjusted R^2	0.02	0.08	0.08	0.11	0.11	0.12	0.12	0.16	0.59

[27] To enhance the interpretability of coefficients and reduce numerical instability caused by the multicollinearity in interaction models (Afshartous and Preston, 2011), we use mean-centered transformations for *Adjustment Frequency* and *Similarity of Available Capabilities.* So β_{AC} captures the effect of *Similarity of Available Capabilities* when *Adjustment Frequency* is at its average level (3.68) and β_{AF} captures the effect of *Adjustment Frequency* on the percentage of work done in-house when *Similarity of Available Capabilities* is at its average level (3.27), while β_{IN} is the same as the one estimated with untransformed variables (reported in Table A1 in Online Appendix A).

Notes: Standard errors are in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. The number of observations reflects 9 categories of data for 79 firms, i.e. 9 x 79 = 711 in Column (1)-(4) and the number of observations reflects data for 42 firms in Column (5)-(9).

The coefficient of *Similarity of Available Capabilities* is positive and highly significant throughout the Table. This is consistent with the theoretical prediction that employment is comparatively better when it is more important that human capital be firm-specific as opposed to function-specific. To get a sense of magnitudes, the main regression in column (6) shows that 9.69% more tasks will be done in-house if the measure of *Similarity of Available Capabilities* increases by one standard deviation (1.26 points) when *Adjustment Frequency* is at its average level.. The coefficient of *Adjustment Frequency* is always positive and most of the times significant. This is consistent with the theoretical prediction that employment is comparatively better when changes are more frequent. Based on Column (6), 4.83% more tasks will be done in-house if the measure of *Adjustment Frequency* increases by one standard deviation (1.25 points) when *Similarity of Available Capabilities* is at its average level. The coefficient on the interaction between *Similarity of Available Capabilities* and *Adjustment Frequency* is positive and significant in four of five regressions. From column (6), we estimate that the effect of *Adjustment Frequency* will increase by 2.42 if the measure of *Similarity of Available Capabilities* increases by one standard deviation from the average level. Finally, still using the numbers form column (6), it is estimated that a one standard deviation increase in log (*City Size*) (0.87) will lead to a 1.24 increase in *Share In-House*.

To test the robustness of our results, we introduce fixed effects in columns (7) – (9). In column (7) we add a vector of *Service Mix* category indicators, which control the baseline differences among 9 service mix categories. A vector of *Industry* indicators, which control the baseline differences among the environments of the 79 firms is added into the model in Column (8).[28] The estimation results in (7) and (8) are similar to those reported in Column (6). However, when we put in firm fixed effects in Column (9), *Adjustment Frequency* and *Adjustment Frequency x Similarity of Available Capabilities* are no longer

[28] The firms are classified into eight industries based on the first digit of the SIC code.

significant. This is quite a surprising finding, especially in light of previous tests of the *Adjustment Frequency* variable. However, given the small sample size, we have not been able to gain any real understanding of it.

To check whether the results depend on the functional form, we report fractional logit regressions in Table A2 in Online Appendix A. Another concern about our results is that the effects we want to capture may vary among functions or/and media. To address this issue, we estimated sets of linear and fractional logit regressions that included interaction terms between function and media indicators[29] along with the three basic predictors. These results are reported in Tables A3 and A4 in Online Appendix A. In all three cases, the main effects remain almost the same after controlling for heterogeneity and the differences among categories are not significant in most specifications, which suggests that the effects of heterogeneity is not a serious threat to the validity of our model. The results are very close to what we get from the linear model in terms of relative significance and magnitude.

As a final check, Table 6 presents the correlation matrix for the three explanatory theory variables in the model. None of those correlation coefficients has an absolute value higher than 0.33, indicating that multicollinearity is not a likely threat to the parameter estimation.

Table 6: Correlations Matrix for Explanatory Variables, N=378

	1	2	3
1. *Adjustment Frequency*	1.00		
2. *Similarity of Available Capabilities*	0.33	1.00	
3. Log *(City Size)*	-0.10	0.14	1.00

III. 6 Relative Significances

To throw more light on the relationship between firm fixed effects and the theory, we use the procedure from Schmalensee (1985) and Wernerfelt and Montgomery (1988). These results are given in Figure 1 below. The second row in the Figure gives the R^2 and adjusted R^2s of linear regression models with no

[29] All of the indicators are also mean centered to allow easy interpretation of the interaction terms.

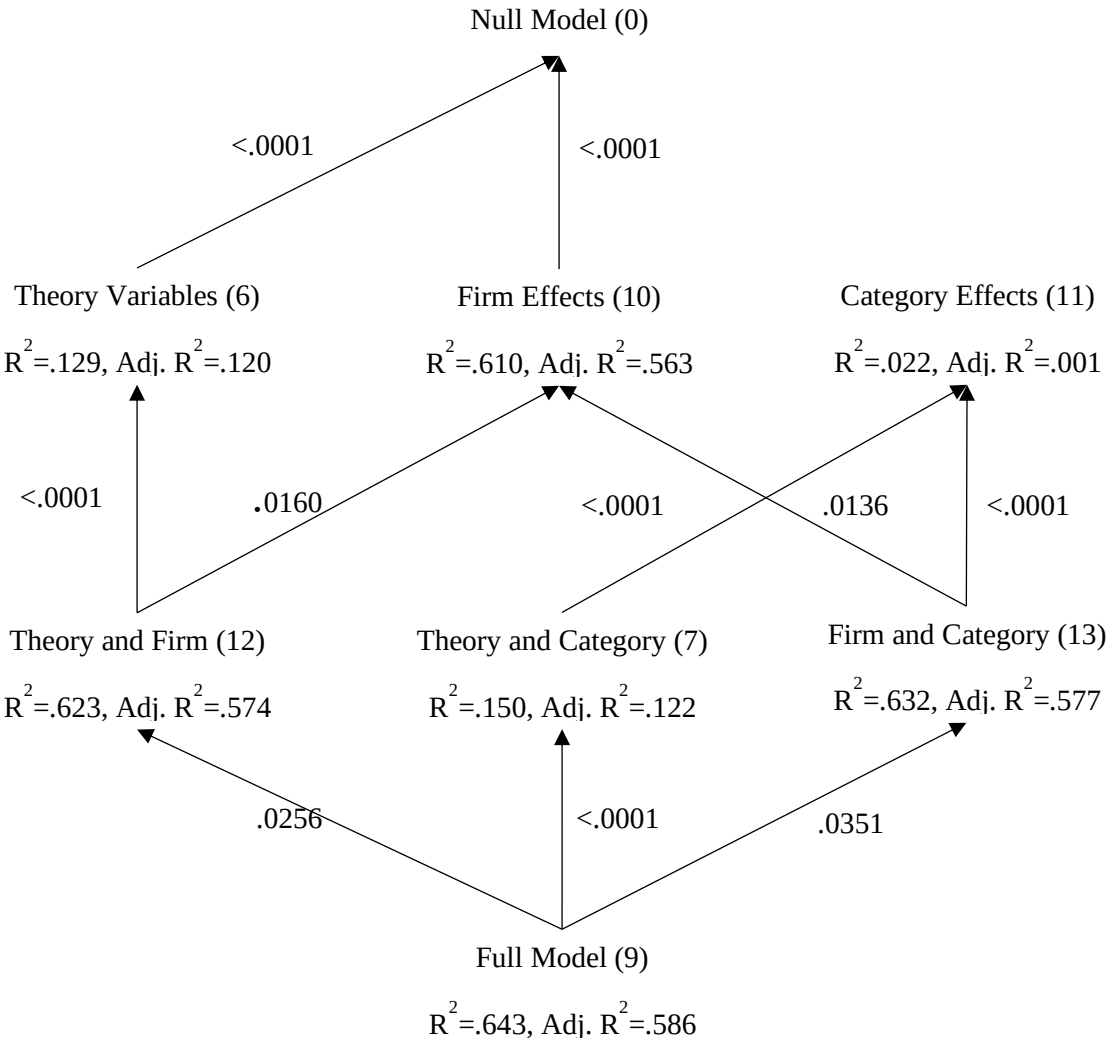
fixed effects ((6) in Table 5), with only firm fixed effects, and with only category fixed effects. Later rows combine those. The estimation results for the underlying regressions are reported in Table A5 in Online Appendix A.

To interpret this Figure, look at the line between "Theory Variables" and "Null Model": The .12 is the adjusted R^2 of (6) in Table 5 and <.0001 is the *p*-value of the F-test[30] that tells us whether the theory variables provide enough extra explanatory power as a group to justify their addition to the null model. The strong significance here indicates that the theory variables play a crucial role in explaining the variance in the level of internalization. Similarly, the missing line between "Theory and Category" and "Theory Variables" indicates that the difference between the model with both the theory variables and category dummies and the restricted model with only theory variables is insignificant. As can be seen, the theory variables are significant throughout the Figure. In order to understand the strength of the firm fixed effects, the most important number is the .0160 indicating that a model with the theory variables and firm fixed effects performs significantly better than a model with firm fixed effects alone. When we replace firm fixed effects with industry fixed effects and replicate the figure (Figure A1 in Online Appendix A), the theory variables remain extremely significant throughout.

[30] $F = \frac{(R^2_{Unrestricted} - R^2_{Restricted})/(p_{Unrestricted} - p_{Restricted})}{(1 - R^2_{Unrestricted})/(N - p_{Unrestricted} - 1)}$ where p is the number of predictors. The degrees of freedom are $p_{Unrestricted} - p_{Restricted}$ and $N - p_{Unrestricted} - 1$.

Figure 1: Tests for the Existence of Theory, Firm, and Category Effects – Linear Model

(All significant effects indicated by arrows)



In addition to the above, we have some suggestive, but interesting, time-series data.

IV. INTERNALIZATION OVER TIME

In this Section, we will first explain why the original level of internalization was low and a how regulatory change paved the way for more. We then show data suggestive of three recent, and

simultaneous, trends: The actual level of internalization has gone up, the value of fast adaptation has increased, and service specialization has become less important.

IV.1 The Recognition System and the 1956 Consent Decree

When "space brokers" originated in the middle of the nineteenth century, they did what we now call "media buy" and were compensated by a commission on the costs. The modern full-service advertising agencies were born when advertisers began to demand additional services needed to create and produce advertising and promotional programs. Paradoxically, despite the evolution from space broker to independent full-service provider, the method of agency compensation remained essentially unchanged in that it continued to be determined by the commission earned on media bookings. Reliance on the commission system of agency persisted for decades as a result of what Pope (1983, p. 153) has labelled as an "alliance of convenience" between publishers and independent agencies. That relationship came to known as "the recognition system" and consisted of "various policies, standards, and procedures used to qualify advertising agencies to do business with, and to be entitled to receive credit and agency commission" (U.S. District Court for the Southern District of New York, Complaint filed by the U.S. Department of Justice, 1955).

The recognition system initially served to alleviate a moral hazard problem. It is hard to write contracts on the quality of work going into the creation and production of ads, but if the advertiser chooses to run good ads more than bad ads, the qualities of creation and production are indirectly measured by media billings. However, the recognition system also restricted price competition among agencies by standardizing commission rates (15 percent) and functioned to discourage, if not preclude, advertisers from performing any of the functions themselves.

The recognition system resulted in an investigation by federal antitrust authorities on grounds that it constituted a conspiracy to restrain trade and thus violated the Sherman Act. The first complaint was dismissed in 1930, but the second led to the signing of a consent decree in 1956 by five trade associations

representing publishers of magazines and newspapers plus the independent agencies, the American Association of Advertising Agencies (4A)'s—i.e., the set of organizations that had been participated in the administration of the recognition system. While the consent decree was credited with effectively dismantling the administrative system that resided in the aforementioned trade associations (Holland, 1981), individual media vehicles remained free to grant commissions to independent agencies and withhold it from in-house agencies and advertisers (Klaw, 1956). Thus, the consent decree had little immediate effect on agency compensation (Wood, 1958, pp.471-472) and media commissions persisted as the dominant mode of agency compensation for several decades.

For our purposes, the key implication of the recognition system is that by compensating agencies for the entire bundle of services required to conceive of, produce, and place an ad, it removed all incentives for partial internalization.[31] The practice could only take off after the 1956 consent decree.

IV.2 Internalization Has Grown Over Time

We will here look at the growth of internal agencies through several indicators. While the data are incomplete and partially indirect, it seems clear that the use of internal agencies has grown over time.[32] As mentioned above, the Consent Decree did not result in the Recognition System being abandoned overnight. Tracking studies by the ANA showed that as late as 1982, 71 percent of the largest U.S. national advertisers utilized commissions-based compensation. That share fell to 61 percent in 1994 and continues to decline, dropping to only 16 percent in 2006 (Beals, 2007).

The current state of affairs has evolved not only through the entry of new firms specializing in narrow types of advertising (collectively referred to as "digital"), but also as a result of the disintermediation of established (i.e., traditional independent) advertising agencies and providers of related services.

[31] See Arzaghi et al. (2012) for an analysis of the unbundling of agency services over the period 1982-2007.
[32] Using data from the *Standard Director of Advertisers*, Horsky et al., 2012) found that across 69 2 digit SIC industries, the share of advertisers who reported operating an in-house agency increased from the 43.3% (of 9,527 advertisers) in 1990 to 53.4% (of 15,548 advertisers) in 1999.

Contemporary indicators of the growth of internal agencies over time are in short supply and all are limited in the sense that the time series are short or interrupted and tend to cover only a limited range of the total size distributions of interest. We will now discuss three indicators relevant to the purposes at hand and publicly available.

First, (very) intermittent tracking studies by the Association of National Advertisers (ANA) looked at the percentage of members (large U.S. advertisers) that "use an internal agency". These numbers are given in Table 7 below.

Table 7: ANA members using an internal agency[33]

Year	1976	2008	2016	2018
Percentage	11	42	58	78

Source: ANA (2018)

A second measure, presented in Table 8, is the number of members in the In-House Agency Forum (IHAF), a membership organization founded in 2005.[34] Since this series is more complete and the numbers are larger, this is what we will compare to measures of our independent variables.

Table 8: Members of IHAF

Year	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
Members	77	80	89	125	154	179	253	295	306	295

Source: IHAF

A third and less direct way to judge the extent and importance of internalization is to look at the economic fortunes of external agencies. To this end, we compared the cumulative return of an investment of one

[33] IAB (www.iab.com/2019-european-programmatic-in-housing), reports that 86% of a sample of large European advertisers had internalized all or part of their programmatic media buying. As reported in Section IV below, we find partial in-housing in 100% of a 2018 sample consisting of US-based advertisers.

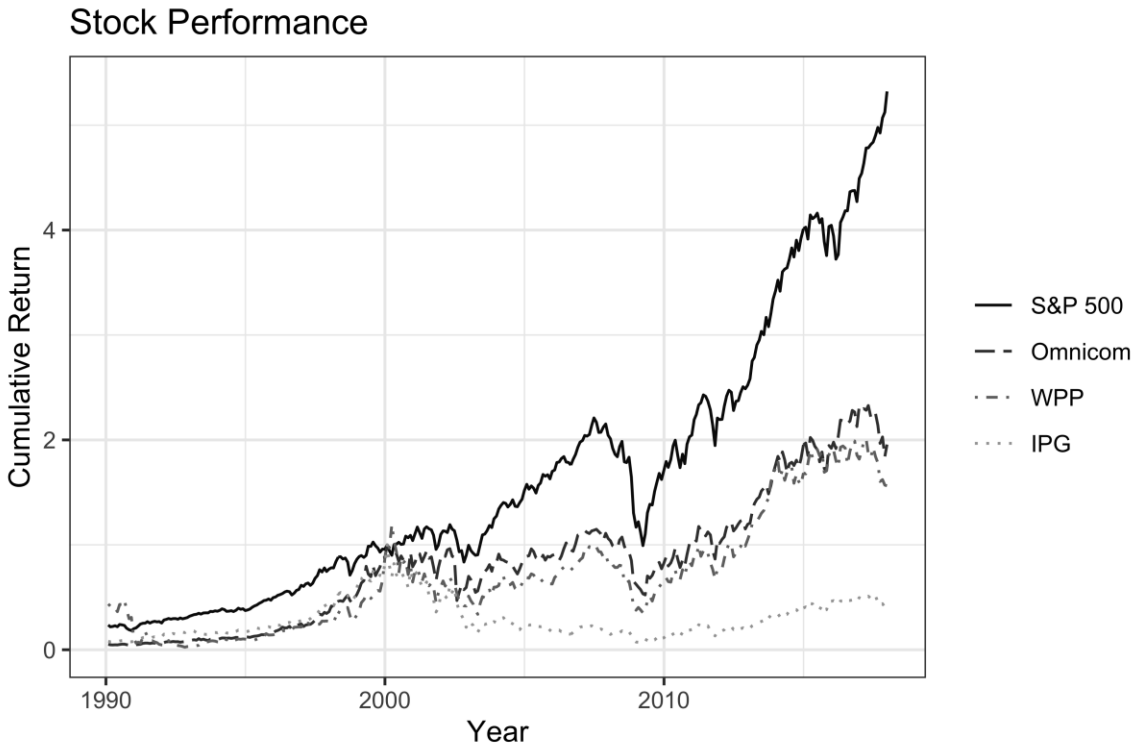
[34] The membership obviously does not include all firms that have partially of fully internalized, but the idea is that it includes a more or less constant fraction of them.

dollar in an S&P 500 index fund in 2000 to the cumulative return of an investment of one dollar in three of the largest suppliers of advertising and marketing services (A&MS) in the same year.[35] Figure 2 shows the stock performance for three of the "Big Four" holding companies whose stocks are publicly traded in the US: Omnicom (NYSE: OMC), Interpublic Group (NYSE: IPG), and WPP (NYSE: WPP). These three conglomerates operate globally and each own/control a portfolio of both full and specialized A&MS firms. Together with Publicis (the other "Big Four" holding company which is traded in France), they have about a quarter of the US A&MS market (Silk and King, 2013). It is clear that this trio of holding companies have underperformed the market since the beginning of the 21st century: Starting in 2010, the S&P 500 index fund has had at least twice the return than any of the three aforementioned major A&MS holding companies.[36] A possible explanation for this pattern is that internal agencies started to grow rapidly after 2010, (though there are many other possibilities including the growing competition from consulting companies).

Figure 2: Stock performance of advertising and marketing services holding companies 1990-2018

[35] 2000 is chosen as the benchmark year because Google launched AdWords in that year, thereby making it much easier to purchase advertising space.

[36] Given the notorious cyclicality of the advertising industry, one would have expected it to over-perform the broader market as the economy was coming out of the great recession of 2008-9 (by NBER's dating).



IV.3. Decreasing Value of Service Specialization

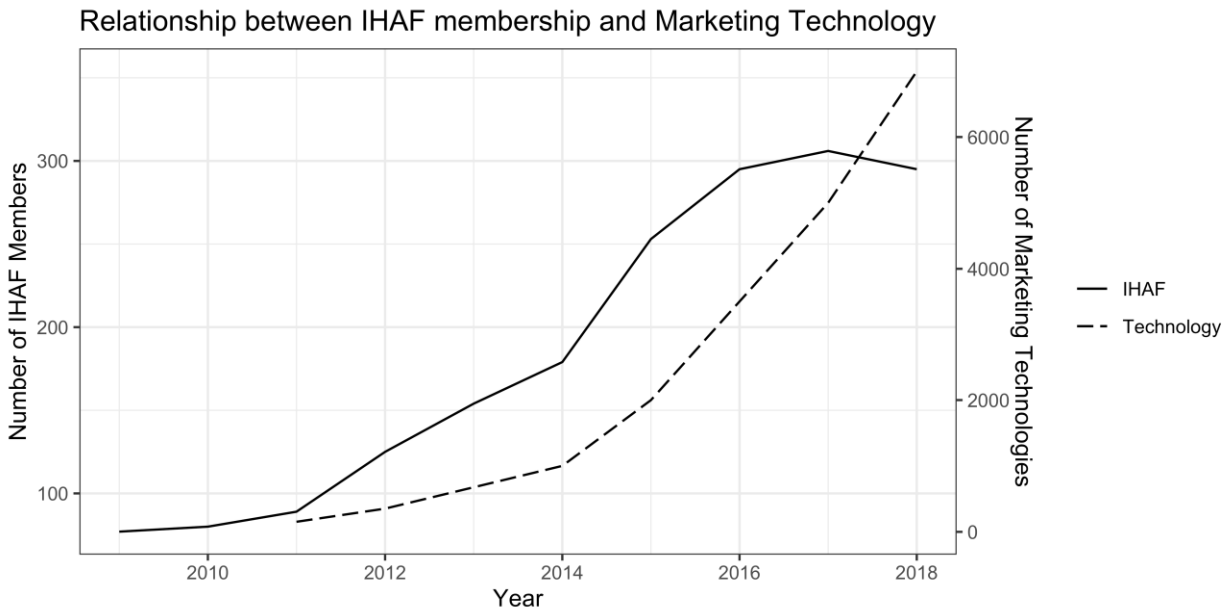
The relative values of firm – vs. service specialization are harder to gauge. However, one possibly relevant force is the growth in marketing technologies offered by third parties. At least up to the turn of the century, many marketing technologies (market research techniques, media planning programs, etc.) were owned by individual agencies. Since then, there has been an explosion of independently owned technologies (tools for marketing automation, business intelligence, analytics, SEO, CRM, targeting tools offered by Facebook and Google, etc.) that are available to all interested parties such that the productivity advantages of the external agencies have decreased.[37] The explosive rise of programmatic advertising, which typically utilizes several widely available tools, is a particularly important factor in this shift.[38]

[37] An anonymous referee has pointed out that it could also be argued that the explosion in marketing technologies makes it harder for individual firms to take advantage of all of them and therefore decreases, rather than increases, the value of firm-specialization. We agree that this is possible but consider our interpretation more likely - in particular when the number of third-party mar-techs was very small.

[38] It was recently reported that programmatic advertising accounts for 80% of display advertising on a revenue basis (IAB 2019).

Scott Brinker of the Chiefmartec.com blog estimates the number of marketing technologies as illustrated in Figure 3.

Figure 3: Relationship between IHAF membership and Marketing Technologies (2009-2018)



As can be seen from the Figure, there has, over the last ten years, been an exponential growth the in number of such technologies. The correlation between IHAF membership and the number of marketing technologies between 2011 and 2018 is .86 (N = 8, p = .013). So also this is consistent with our theory. [39]

IV.4 The Value of Fast Adaptation Has Increased

It is generally regarded as a truism that the "speed of business" is growing all the time. This alone, whatever its reasons, will increase the value of fast adaptation.

One obvious, and often mentioned, factor pointing in the same direction, is the growth of digital advertising. Advances in communication and information technologies have improved the efficiency and

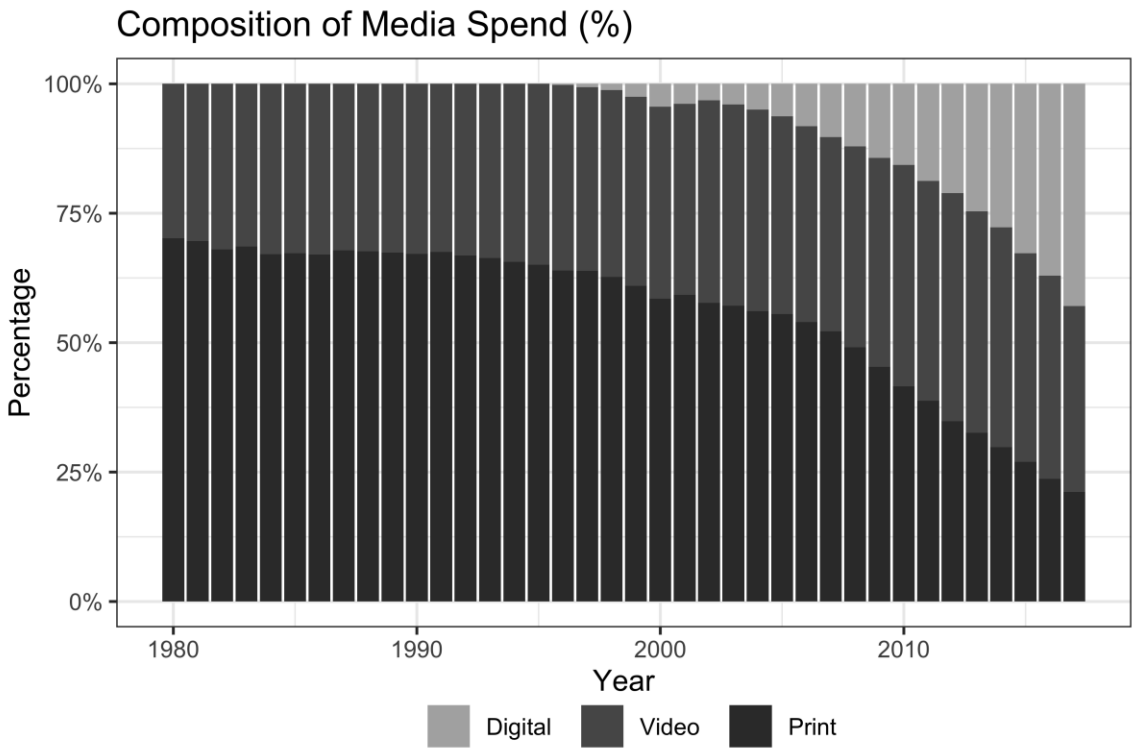
[39] It should be clear that the reported correlations do not constitute a test of the theory. While digital ad spend and the number of marketing technologies were the first two time series we looked at, we could, in principle**,** have picked them out from a large number of others. The relationships are interesting because they suggest (and just suggest) that the trend in in-housing reflects a gradual change in the equilibrium level, instead of a slow adjustment to a new equilibrium.

reduced the cost of digital advertising[40]. The resulting growth has transformed not only the media habits (Kivijarvi, 2018) and purchasing behavior of consumers (Goldfarb and Tucker, 2017), but also the distribution and advertising strategies sellers pursue and how those activities are organized and managed (Burton, 2009; Evans, 2008, 2009). One indication of the magnitude of this transformation may be gleaned from the striking shifts of advertising expenditures that has occurred across major media since the launch of digital advertising media in the U.S. As captured by data on the advertising receipts of media suppliers, the share of total advertising expenditures accounted for by digital media grew from 0.21 percent in 1996, to 35.6 percent in 2016. In that year, digital displaced television (33.0 percent) as the share leader among eight major categories of advertising media (Magna Global, 2015; Letang and Leszega, 2017). Concurrently, the shares of other traditional media declined dramatically, especially for print media, such as direct mail, magazines and newspapers. These trends are illustrated in Figure 4, where the market share of digital media spend is plotted over time.[41] As can be seen in the graph, the role of digital advertising has grown rapidly.

Figure 4: Share of Total Media Spend due to Digital Advertising

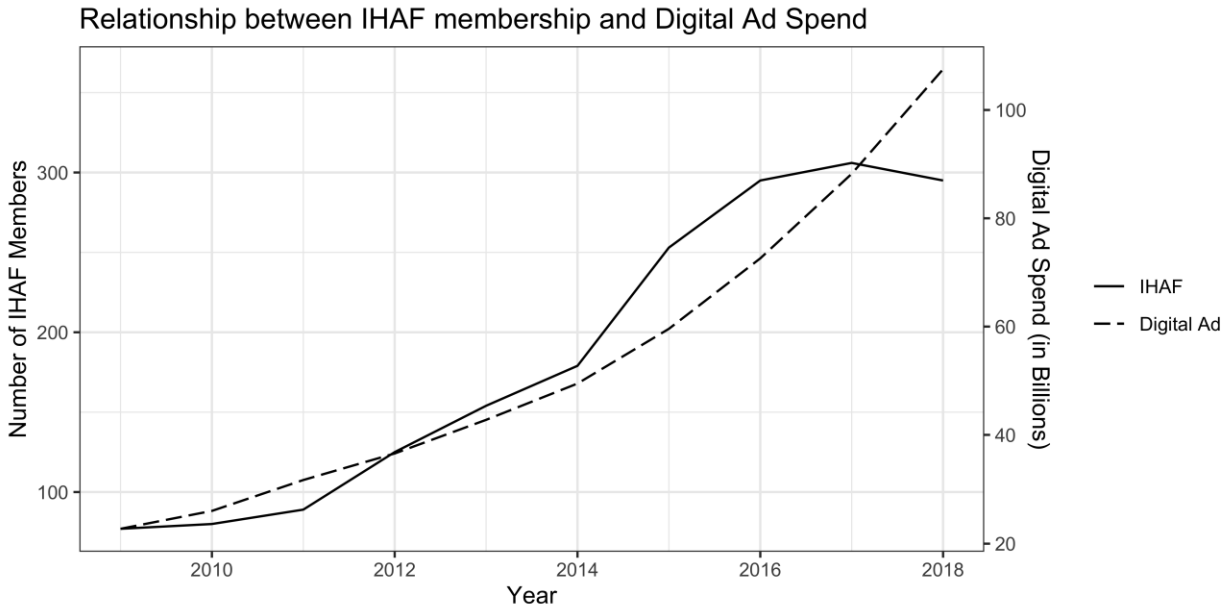
[40] The Bureau of Labor Statistics (BLS) price index for Internet Advertising (excluding that sold by print publishers) has fallen approximately 33% between 2012 (when BLS initiated the series) through 2018. In contrast, over that same period, the price indices for the "traditional" media have either risen (in the case of "Periodicals"), or declined only moderately (2-6%) in the cases of Radio, Television, and Directories and Mailing Lists. https://www.bls.gov/ppi/data.htm.

[41] Source: Magna Global U.S. Media Forecast (2018). The eight major categories of advertising media are grouped into three larger categories to make the bar chart more readable: Radio and television are grouped into video media; Directory, direct mail, magazine, newspaper, and out-of-home media are grouped into print media.



The relationship between digital advertising and in-house agency activity (measured by IHAF membership), is plotted in Figure 5. As can be seen, the relationship is extremely strong. (With N = 10, the correlation is .94 and $p < .0001$). However, we conjecture that it would be significantly weaker in a well-identified model. To maintain that digital advertising alone has driven the transition to internal agencies, one would effectively be dismissing an important chapter of well documented advertising history. On the other hand, the graph serves to highlight that digital technology has been an important factor in the recent expansion of in-house advertising services.

Figure 5: Relationship between IHAF membership and Digital Advertising Spend (2009-2018)



V. DISCUSSION AND CONCLUSION

V. 1 Alternative Theories

In this section, we consider plausible rival hypotheses that might account for the patterns observed in the data presented here. As mentioned in the Introduction, the correlation between integration and "*Similarity of Available Capabilities*", could reflect the specific asset correlation predicted by PRT (and TCE). Since that theory is based on relationship-specific investments, the idea would have to be that the differences in integration reflect differences in the relative importance of investments made by the advertiser and those made by the agency. One might argue that both agency and advertiser make investments in their employees and it seems reasonable to conjecture that these investments relate to the relative importance of firm vs. service specialization. So both PRT and the version of TCE that is concerned with ex ante investment distortions could predict the positive correlation. However, neither theory predicts that this effect should be stronger when changes are more frequent nor that frequency by itself should lead to more internalization. More fundamentally, it is hard to explain any role of "*Adjustment Frequency*" without relying on some sort of bargaining or contracting frictions combined with multi-period effects. So at least the formalized versions of other theories cannot capture it.

Beyond those suggested by alternative theories, there may be other factors, such as concerns over data sharing, behind the growing internalization of advertising agencies.

V.2 Improving Our Measures

We next discuss some post hoc analyses we undertook to assess the possible presence of certain threats to the construct validity of our measures and suggest some directions that might improve upon the research design and instruments employed here.

The half-matrices of the 36 pairwise associations among the 9 ratings for each of the "frequency" and "similarity" variables presented in Table 4 were all positive, varying in magnitude from 0.52 to 0.79 and 0.32 to 0.79 for the "frequency" and "similarity" ratings, respectively. Elevated levels of pairwise associations among "similar" measures (5 point ratings scales) of different constructs (9 activities consisting of combinations defined by medium and function) might be symptomatic of some component of systematic measurement error such as "yeasaying-naysaying response set" (Couch and Kenniston 1960) that can adversely affect the discriminant validity of measures collected via self-administered questionnaires in surveys (cf., Wells, 1961 and Silk, 1971). In Online Appendix B, we analyze correlation coefficients using ideas suggested by Campbell and Fiske (1959) and point out the possible presence of such a threat to construct validity of the aforementioned variables. While that analysis did not indicate that informants' ratings of "frequency" and "similarity" differed with respect to discriminant validity, that conclusion should be regarded as tentative given that the present study involved a mono-method design, precluding the assessment of both convergent and discriminant validity. It should, however, be noted that we found no indications that the amount of time informants spent on responding to the self-administered questionnaire was related to the within-informant variance for either the "Frequency" or "Similarity" ratings they provided. The correlations were 0.00 and -0.08, respectively.[42] Clearly, measure validation is an issue that warrants further investigation using a research design that involves collection of data from

[42] Kahneman (2011) and others have used fast response time as an indicator of less reflective answers.

two or more informants (Algesheirmer, Bagozzi, and Dholakia, 2018; Bagozzi, Yi, and Phillips, 1991; and Seidler, 1974).[43]

Our experience also suggests that reliance on "senior managers" as informants for measures of the organizational constructs related to the internalization of advertising and marketing services could be reconsidered. First, the term "senior manager" is open to alternative interpretations; especially given the heterogeneity in the design and staffing of marketing organizations across industries and/or firms. It seems likely that "middle managers" (e.g., brand manages) or specialized staff personnel (advertising department or procurement managers) could satisfy the criteria for "key informants" emphasized in the literature (e.g., "specialized knowledge, accessibility") dating back to Seidler (1974) and Phillips (1981). Unfortunately, we were unable to obtain any detailed information about the organizational positions, roles, or responsibilities of our informant sample. Further guidelines for identifying and recruiting informants could be culled by scrutinizing the literature on methods that were found to workable in informant research conducted in other organizational contexts (Seidler, 1974; Silk and Kalwani, 1982; and Parmigiana, 2007). Overall, our experience suggests that overcoming the challenges surrounding the design and implementation of future research on this topic calls for a two-step process whereby the initial phase would focus on diagnosing the structure and decision processes of a firm's marketing organization and collecting archival data on a the firm's advertising spending and agency relationships. The second stage would recruit informants and collect data from those identified in the first phase as satisfying the desiderata for a "key informant." in the target organization. These are clearly very difficult procedures to implement and serve to highlight the limitations of the field's current knowledge about the design and advertising planning and budgeting practices of contemporary marketing organizations (Moorman and Day 2016, Kolsari et al. 2020).

[43] Amabile and Kramer (2011) use data from even more informants and get very rich information

In conclusion, like all empirical studies, our cross-sectional study could be improved. However, it is very challenging to collect get data on the kinds of variables we seek to measure and the realistic alternative to coping with fallible measures is arguably preferable to ignoring altogether a problem of importance to both theory and practice.

V.3 Summary and Directions for Future Research

We have tested a theory of the firm in the context of the recent internalization of advertising services. The theory predicts that more internalization if (1) it is more important for human capital to be firm-specific as opposed to function-specific and (2) if frequent modification is more desirable. It also predicts that (3) these two effects reinforce each other. The specialization hypothesis has not been tested before and the fact that it is robustly significant in the cross-sectional test, constitutes the main contribution of the paper. In contrast, the frequency effect ceases to be significant once we introduce firm fixed effects, indicating that there is little within firm effect in this dataset.

Advertising services provide an attractive setting for the study of integration and invite extensions in at least three directions. First, and most importantly, the paper cries out for a causal test. We hope that the suggestive nature of our correlational results will inspire future work in that direction. Second, the industry has experienced numerous changes over the last ten years and it would be interesting to look at panel data to take a more formal look at the difference between gradual changes in the equilibrium level of integration versus a slow transition from one equilibrium to another. Third, as is clear from the data on our dependent variable (*Share In-House*), advertisers engage in a significant amount of dual sourcing, and this present both a challenge and an opportunity to learn about a currently poorly understood strategy.

REFERENCES

Afshartous, D. and R. A. Preston (2011). Key results of interaction models with centering. Journal of Statistics Education, 19 (3), pp. 1-24.

Algesheirmer, R., R. P. Bagozzi, and U. M. Dholakia (2018), "Key Informant Models for Measuring Group-level Variables in Small Groups: Application to Plural Subject Theory," *Sociological Methods & Research,* 47, pp. 277-313.

Amabile, Theresa, and Steven Kramer, *The Progress Principle: Using Small Wins to Ignite Joy, Engagement, and Creativity at Work,* Boston, MA: Harvard Business School Press, 2011.

Anderson, E. (1985), "The salesperson as outside agent or employee: A transaction cost analysis," *Marketing Science*, 4 (Summer), pp. 234-254.

Anderson, E. & D. C. Schmittlein (1984), "Integration of the sales force: An empirical examination," *RAND Journal of Economics*, 15 (Autumn), pp. 385-395.

Anderson, E., and B. A. Weitz (1986), "Make-or-Buy Decisions: Vertical Integration and Marketing Productivity," *Sloan Management Review,* 27 (Spring), pp. 3-19.

Antras, Pol, "Firms, Contracts, and Trade Structure", *Quarterly Journal of Economics,* 118, pp. 1375-1418, 2003.

Araghi, M., E.R. Berndt, J.C. Davis, and A.J. Silk (2012), "The Unbundling of Advertising Agency Services," *Review of Marketing Science,* 10, Issue 1, Article 4.

Association of National Advertisers (2012), *Industry Groups Urge Marketers to Take Affirmative Steps to Address Online Piracy and Counterfeiting,* New York, May.

Association of National Advertisers (2018), *The Continued Rise of the In-House Agency.* New York, October.

Bagozzi, R. P., and L. Phillips (1982), "Representing and Testing Organizational Theories: A Holistic Construal," *Administrative Science Quarterly*, 32, pp.459-489.

Bagozzi, R. P., Y. Yi, and L. W. Philips (1991), "Assessing Construct Validity in Organization Research," *Administrative Science Quarterly,* 36, pp. 421-458

Bajari, P., and S. Tadelis (2001), "Incentives Versus Transaction Costs: A theory of Procurement Contracts", *Rand Journal of Economics,* 32, pp. 387-407.

Baker, G. P. & T. N. Hubbard (2004), "Contractibility and asset ownership: On-board computers and governance in US trucking," *Quarterly Journal of Economics,* 119, pp. 1443-1479.

Beals, D. (2007), *Trends in Agency Compensation,* 14th Ed. New York: Association of National Advertisers'.

Burton, J. (2009), *A Marketer's Guide to Understanding the Economics of Digital Compared to Traditional Advertising and Media Services*. New York: American Association of Advertising Agencies.

Campbell, D. T. and D. K. Fiske (1959), "Convergent and Discriminant Validation via the Multitrait-Multimethod Matrix," *Psychological Bulletin,* 56, pp. 81-105.

Coase, R. H. (1937), "The Nature of the Firm", *Economica n. s. 4,* pp. 386-405.

Costinot, A., L. Oldenski, & J. Rauch (2011), "Adaptation and the boundary of multinational firms," *Review of Economics and Statistics,* 93, pp. 298-308.

Couch, A. and K. Keniston (1960), "Yeasayers and Naysayers: Agreement Response Set as a Personality Variable," *Journal of Abnormal and Social Psychology,* 60, pp. 151-174.

Dessein, W., A. Galeotti, and T. Santos, "Rational Inattention and Organizational Focus", *American Economic Review,* 106, pp. 1522-36, 2016.

Dutta, S., M. Bergen, J. Heide, and G. John (1995), "Understanding Dual Distribution: The Case of Reps and House Accounts", *Journal of Law, Economics, and Organization,* 11, pp. 189-204.

Evans, D. S. (2008), "The Economics of the Online Advertising Industry," *Review of Network Economics,* 7 (September), pp. 351-391.

_____ (2009), "The Online Advertising Industry: Economics, Evolution, and Privacy," *Journal of Economic Perspectives,* 23 (September), pp. 37-60.

Forbes, Silke J., and Mara Lederman,(2009), "Adaptation and Vertical Integration in the Airline Industry", *American Economic Review,* 99, pp. 1831-49.

Geyskens, I., Steenkamp, J., and N. Noorderhaven (2006), "Make, Buy, or Ally: A Transactions Cost Theory Meta-Analysis", *Academy of Management Journal,* 49, pp. 519-43.

Goldfarb, A. and C. E. Tucker (2017), "Digital Economics," NBER Working Paper 23684. Cambridge, MA: National Bureau of Economic Research, August.

Goodman, L. A., and W. H. Kruskal (1954), "Measures of Association for Cross Classifications", *Journal of the American Statistical Association,* 49, pp. 732-64.

Grams, C. (2018): "In-House vs. Agency Branding: 5 Benefits of Each Approach", https://newkind.com/in-house-vs-agency-branding-benefits/.

Grossman, S., and O. D. Hart (1986), "The Costs and Benefits of Ownership", *Journal of Political Economy,* 94, pp. 691-719.

Hart, O. D., and J. Moore (1990), "Property Rights and the Nature of the Firm", *Journal of Political Economy,* 98, pp. 1119-58.

_____ (2008), "Contracts as Reference Points", *Quarterly Journal of Economics,* 123, pp. 1-48.

Holland, D. R. (1981), "The Great Tradeoff," *Advertising Age*, 52 (July 6), pp. 41-42.

Horsky, S. (2006), "The Changing Architecture of Advertising Agencies", *Marketing Science,* 25, pp. 367-83.

Horsky, S., S. C. Michael, and A. J. Silk (2012), "The Internalization of Advertising Services: An Inter-Industry Analysis," *Review of Marketing Science,* 10 (Article 2), pp. 1-37.

Interactive Advertising Bureau (2019), (www.iab.com/2019-European-programmatic-in-housing).

Interactive Advertising Bureau (2019), *IAB Internet Advertising Revenue Report, 2018 Full year results.*

John, G., and B. A. Weitz (1988), "Forward Integration into Distribution: An Empirical Test of Transaction Cost Analysis", *Journal of Law, Economics, and Organization,* 4, pp. 337-55.

Kahneman, D. (2011), *Thinking Fast and Slow*, New York, NY: Farrar, Straus, and Giroux.

Kendall, M. and B. Babington Smith (1939), "The Problem of m Rankings," *Annals of Mathematical Statistics,* (10), 275-287.

Kivijarvi, L. (2018), "Time Spent on Advertising Daily," *Digital News Daily,* April 2. Available at: http://www.mediapost.com/publications/article/316932/time-spent-with-advertising-daily.

Klaw, S. (1956), "Advertising: The Battle of Fifteen Percent," *Fortune,* 54 (October), 142.

Klein, S. (1989), "A Transaction Cost Explanation of Vertical Control in International Markets", *Journal of the Academy of Marketing Science,* 17, pp. 253-60.

Kolsario, C., V. Vakratss, and P. Naik (2020), "The Anatomy of the Advertising Budget Decisions: How Analytics and Heuristics Drive Sales Performance," *Journal of Marketing,* 57(3), pp. 468-488.

Lafontaine, Francine, and Margaret Slade (2007), "Vertical Integration and Firm Boundaries: The Evidence", *Journal of Economic Literature,* 45, 3, pp. 629-85.

Leiblein, M. J. and D. J. Miller (2003), "An Empirical Examination of Transaction- and Firm – Level Influences on the Vertical Boundaries of the Firm", *Strategic Management Journal,* 24, pp. 839-59.

Letang, V. and M. Leszega (2017), "U.S. Advertising Forecast," March. New York: Magna Global.
Macaulay, S. (1963), "Non-Contractual Relations in Business", *American Sociological Review,* 28, pp. 55-70.

Liffreing, Ilyse (2019), "Anheuser-Busch InBev Launches In-House Agency Draft line", *AdAge,* May 16.

Macaulay, S. (1963), "Non-Contractual Relations in Business", *American Sociological Review,* 28, pp. 55-70.

Magna Global (2015), "US Media Forecasts: Methodology and Definitions," August. Available from vincent.letang@magnglobal.com.

Maciejovsky, B., and B. Wernerfelt (2011), "Costs of Implementation: Bargaining Costs versus Allocative Efficiency", *Journal of Economic Behavior and Organization,* 77, pp. 318-25.

Marozzi, M. (2014), "Testing for Concordance between Several Criteria," *Journal of Statistical Computing and Simulation,* (September), 1843-1850. Available at: http://www/researchgate.net/publication/26258626.

Maskin, Eric, and Jean Tirole (1999), "Unforeseen Contingencies and Incomplete Contracts", *Review of Economic Studies,* 66, pp. 83-114.

Masten, S. E. (1984), "The organization of production: Evidence from the aerospace industry," *Journal of Law and Economics,* 27, pp. 403-417.

Matouschek, N. (2004), “Ex Post Inefficiencies in a Property Rights Theory of the Firm”, *Journal of Law, Economics, and Organization,* 20, pp. 125-47.

Monteverde, K. & D. J. Teece (1982), “Supplier switching costs and vertical integration in the automobile industry.” *Bell Journal of Economics,*13, pp. 206-213.

Moorman C. and G.Day (2016), “Organizing or Marketing Excellence,” *Journal of Marketing,* 80 (November), pp. 6-35.

Novak, S. and B. Wernerfelt (2012), “On the Grouping of Tasks into Firms: Make-or-Buy with Interdependent Parts,” *Journal of Economics& Management Strategy,”* 21 (Spring), pp. 53-77. Pp.

Parmigiana, A. (2007), “Why Do Firms Both Make and Buy? An Investigation of Concurrent Sourcing,” *Strategic Management Journal*, 28, pp. 285-311.

Phillips, L. W. (1981), “Assessing Measurement Error in Key Informant Reports: A Methodological Note on Organizational Analysis in Marketing,” *Journal of Marketing Research,* 18, pp. 395-415.

Phillips, L. and R. Bagozzi (1986), “On Measuring Organizational Properties of Distribution Channels: Methodological Issues in Use. Of Key Informants,” *Research in Marketing,* Vol. 8, pp. 313-369.

Pope, D. (1983), *The Making of Modern Advertising*. New York: Harper.

Rindfleisch, A., and J. B. Heide (1997), “Transaction Cost Analysis: Past, Present, and Future Applications”, *Journal of Marketing,* 61, pp. 30-54.

Schmalensee, R. (1985), “Do Markets Differ Much?” *American Economic Review,* 75 (June), pp. 341-51.

Siedler, J. (1974), “On Using Informants for Collecting Quantitative Data and Controlling for Measurement Error in Organization Analysis,” *American Sociological Review,* 39, pp. 816-831.

Silk, A. J. (1971), “Response Set and the Measurement of Self-Designated Opinion Leadership,” *Public Opinion Quarterly,* Vol. 35, pp. 383-397.

_____and M. Kalwani (1982), “Measuring Influence in Organizational Purchase Decisions,” *Journal of Marketing Research*, 19, pp. 165-181.

_____and C. King III (2013), “How Concentrated Is the U.S. Advertising and Marketing Services Industry? Myth vs. Reality,” *Journal of Current Issues and Research in Advertising*, 34:1, pp. 166-93.

_____and M. S. Stiglin (2016), “Build It, Buy It, or Both? Rethinking the Sourcing of Advertising

Sources," *International Journal of Marketing Studies*, Vol 8, No. 1 (February), pp. 1-13.

Teles, J. (2012), "Concordance Coefficients to Measure the Agreement among Several Set of Ranks," *Journal of Applied Statistics*, 39(August), 1749-1764.

Wells, W. D. (1961), "The Influence of Yeasaying Response Style," *Journal of Advertising Research*, 1, pp. 1-12.

Wernerfelt, B. (1997), "On the Nature and Scope of the Firm: An Adjustment-Cost Theory", *Journal of Business,* 70, pp. 489-514.

_____ (2015), "The Comparative Advantages of Firms, Markets, and Contracts: A Unified Theory", *Economica,* 82, pp. 350-67.

_____ (2016), *Adaptation, Specialization, and the Theory of the Firm: Foundations of the Resource-Based View*. Cambridge, UK: Cambridge University Press.

_____, and C. A. Montgomery (1988), "Tobin's q and the Importance of Focus in Firm Performance", *American Economic Review,* 78 (March), pp. 246-50.

Williamson, O. E. (1979), "Transaction-Cost Economics: The governance of Contractual Relations", *Journal of Law and Economics,* 22, pp. 233-61.

_____(1985), *The Economic Institutions of Capitalism,* New York, NY: The Free Press.

Wood, J. P. (1958), *The Story of Advertising*. New York: Ronald Press.

Woodruff, C. (2002), "Non-contractible investments and vertical integration in the Mexican footwear industry." *International Journal of Industrial Organization,* 20, pp. 1197-1224.

ONLINE APPENDIX A: Additional Analyses

Table A1:Linear Regressions with Untransformed Variables

Independent Variable	(1')	(2')	(3')	(4')	(5')	(6')	(7')	(8')	(9')
Adjustment Frequency	3.89***		0.90	-8.34***	3.11**	-2.48	-3.62	-3.94	2.95
	(0.96)		(1.03)	(2.36)	(1.42)	(3.16)	(3.19)	(3.18)	(3.45)
Similarity of Available		7.57***	7.20***	-2.53	7.40***	1.23	0.62	0.13	6.31*
Capabilities		(0.93)	(1.03)	(2.47)	(1.40)	(3.41)	(3.46)	(3.51)	(3.39)
Adj Freq x Sim of Cap				2.94***		1.92**	2.16**	2.07**	-0.75
				(0.68)		(0.97)	(0.98)	(0.98)	(1.01)
Log *(City Size)*					1.24	1.42*	1.41*	1.05	-2.42
					(0.85)	(0.85)	(0.85)	(0.96)	(2.41)
Category Fixed Effects	No	No	No	No	No	No	Yes	Yes	Yes
Industry Fixed Effects	No	No	No	No	No	No	No	Yes	No
Firm Fixed Effects	No	No	No	No	No	No	No	No	Yes
Observations	711	711	711	711	378	378	378	378	378
Adjusted R^2	0.02	0.08	0.08	0.11	0.11	0.12	0.12	0.16	0.59

Table A2: Fractional Logistic Regressions

Independent Variable	(1f)	(2f)	(3f)	(4f)	(5f)	(6f)	(7f)	(8f)	(9f)
Adjustment Frequency	1.17***		1.04	1.07	1.14**	1.18***	1.17**	1.14**	1.02
	(0.05)		(0.05)	(0.05)	(0.07)	(0.07)	(0.07)	(0.07)	(0.08)
Similarity of Available		1.37***	1.35***	1.37***	1.36***	1.38***	1.40***	1.36***	1.22**
Capabilities		(0.06)	(0.06)	(0.06)	(0.08)	(0.09)	(0.09)	(0.09)	(0.11)
Adj Freq x Sim of Cap				1.14***		1.09**	1.10**	1.10**	0.97
				(0.04)		(0.05)	(0.05)	(0.05)	(0.06)
Log *(City Size)*					1.05	1.06*	1.06*	1.04	1.52
					(0.04)	(0.04)	(0.04)	(0.04)	(0.87)
Category Fixed Effect	No	No	No	No	No	No	Yes	Yes	Yes
Industry Fixed Effects	No	No	No	No	No	No	No	Yes	No
Firm Fixed Effects	No	No	No	No	No	No	No	No	Yes
Observations	711	711	711	711	378	378	378	378	378
Adjusted *Pseudo* R^2	0.00	0.02	0.02	0.03	0.03	0.03	0.00	-0.00	-0.10

Table A3: Linear Regressions with Interactions

Independent Variable	(17)	(18)	(19)	(20)	(21)
Adjustment Frequency	3.52**	3.31**	3.55**	3.37**	3.77**
	(1.49)	(1.50)	(1.50)	(1.51)	(1.52)
Adjustment Frequency		-1.84		-2.03	-3.88
** I (Creative)*		(3.54)		(3.61)	(4.01)
Adjustment Frequency		0.24		-0.29	0.38
** I (Media)*		(3.59)		(3.63)	(4.10)
Adjustment Frequency			2.39	3.01	3.88
** I (Digital)*			(3.49)	(3.55)	(3.93)
Adjustment Frequency			-4.08	-3.62	0.22
** I (Video)*			(3.49)	(3.54)	(3.98)
Similarity of Available	7.89***	7.86***	7.80***	7.75***	7.83***
Capabilities	(1.40)	(1.41)	(1.41)	(1.42)	(1.42)
Similarity of Available		-1.28		-1.20	-2.14
*Capabilities * I (Creative)*		(3.33)		(3.35)	(4.33)
Similarity of Available		3.90		4.39	7.27
*Capabilities * I (Media)*		(3.46)		(3.49)	(4.57)
Similarity of Available			-3.29	-4.05	-3.36
*Capabilities * I (Digital)*			(3.39)	(3.43)	(4.53)
Similarity of Available			1.60	0.80	6.66
*Capabilities * I (Video)*			(3.31)	(3.34)	(4.45)
Adj Freq x Sim of Cap	2.16**	2.33**	2.04**	2.24**	1.92*
	(0.98)	(0.99)	(0.99)	(0.99)	(1.00)
Adj Freq x Sim of Cap					1.16
** I (Creative)*					(1.75)
Adj Freq x Sim of Cap					-1.16
** I (Media)*					(1.65)
Adj Freq x Sim of Cap					-0.69
** I (Digital)*					(1.70)
Adj Freq x Sim of Cap					-3.50**

** I (Video)*					(1.68)
Log (*City Size*)	1.41*	1.42*	1.36	1.38	1.56*
	(0.85)	(0.85)	(0.85)	(0.85)	(0.85)
Category Fixed Effects	Yes	Yes	Yes	Yes	Yes
Observations	378	378	378	378	378
Adjusted R^2	0.12	0.12	0.12	0.12	0.13

Table A4: Fractional Logit Regressions with Interactions

Independent Variable	(17f)	(18f)	(19f)	(20f)	(21f)
Adjustment Frequency	1.17**	1.16**	1.17**	1.16**	1.18***
	(0.07)	(0.07)	(0.07)	(0.08)	(0.07)
Adjustment Frequency		0.93		0.92	0.85
** I (Creative)*		(0.13)		(0.13)	(0.13)
Adjustment Frequency		1.02		0.99	1.02
** I (Media)*		(0.17)		(0.17)	(0.19)
Adjustment Frequency			1.10	1.13	1.18
** I (Digital)*			(0.16)	(0.17)	(0.19)
Adjustment Frequency			0.84	0.85	1.01
** I (Video)*			(0.12)	(0.13)	(0.17)
Similarity of Available	1.40***	1.40***	1.40***	1.40***	1.41***
Capabilities	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)
Similarity of Available		0.95		0.96	0.92
*Capabilities * I (Creative)*		(0.13)		(0.13)	(0.17)
Similarity of Available		1.21		1.23	1.42*
*Capabilities * I (Media)*		(0.20)		(0.20)	(0.30)
Similarity of Available			0.87	0.84	0.87
*Capabilities * I (Digital)*			(0.14)	(0.13)	(0.18)
Similarity of Available			1.07	1.03	1.37*
*Capabilities * I (Video)*			(0.16)	(0.15)	(0.26)
Adj Freq x Sim of Cap	1.10**	1.11**	1.10**	1.11**	1.09**
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)

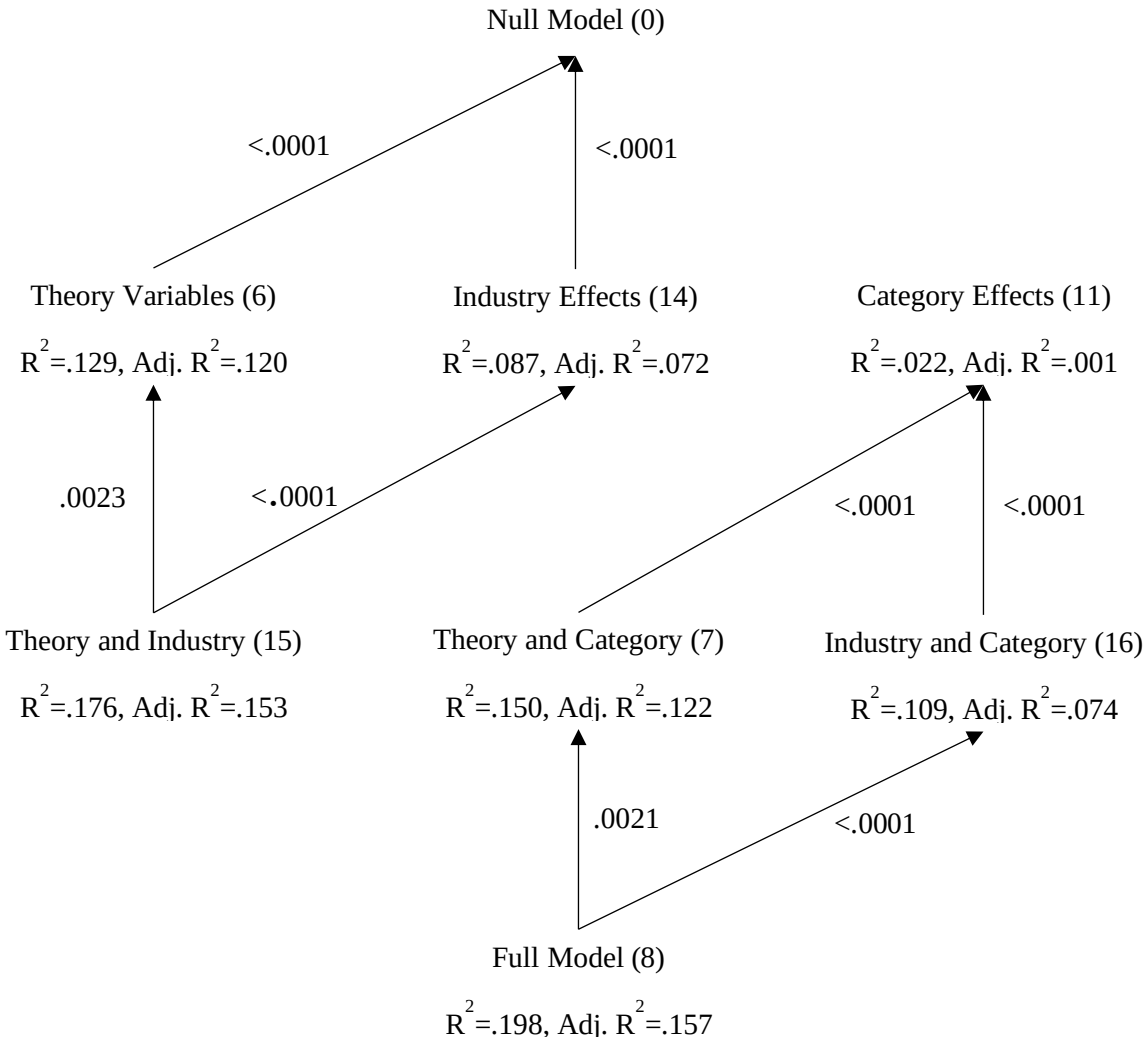
Adj Freq x Sim of Cap					1.05
** I (Creative)*					(0.08)
Adj Freq x Sim of Cap					0.95
** I (Media)*					(0.07)
Adj Freq x Sim of Cap					0.97
** I (Digital)*					(0.08)
Adj Freq x Sim of Cap					0.85**
** I (Video)*					(0.06)
Log (*City Size*)	1.06*	1.06*	1.06	1.06*	1.07*
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Category Fixed Effects	Yes	Yes	Yes	Yes	Yes
Observations	378	378	378	378	378
Adjusted *Pseudo* R^2	0.02	-0.01	-0.01	-0.02	-0.03

Table A5: Additional Regression Models for Figure 1 and Figure A1

Independent Variable	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Adjustment Frequency			1.49			3.26**	
			(1.48)			===(1.47)	
Similarity of Available			3.45**			6.90***	
Capabilities			(1.40)			(1.41)	
Adj Freq x Sim of Cap			-1.23			1.82*	
			(1.00)			(0.97)	
Log *(City Size)*			-2.50			1.06	
			(2.44)			(0.96)	
Category Fixed Effects	No	Yes	No	Yes	No	No	Yes
Firm Fixed Effects	Yes	No	Yes	Yes	No	No	No
Industry Fixed Effects	No	No	No	No	Yes	Yes	Yes
Observations	378	378	378	378	378	378	378
Adjusted R^2	0.56	0.00	0.57	0.58	0.07	0.15	0.07

Figure A1: Tests for the Existence of Theory, Industry, and Category Effects – Linear Model

(All significant effects indicated by arrows)



ONLINE APPENDIX B: Analysis of Gamma Coefficients

The task survey informants were asked to perform solicited assessments of a pair of different constructs ("Frequency of Adjustment" and "Similarity of Capabilities," hereafter denoted as FA and SC, respectively), for a set of 9 different conditions, representing all possible pairs of 3 types of media and 3 functions) on the same 5 point rating scale. The question naturally arises: In reporting ratings of both FA and SC pertaining to their organizations, did informants discriminate among the 9 conditions? Previous research suggests that such ratings may be afflicted by systematic errors such as those arising from informants' organizational positions (Silk and Kalwani 1982) and/or acquiescent response style (Couch and Keniston 1960, Wells1961). To assess the possible presence of such unwanted factors, we undertook some further analyses of the FA and SC ratings suggested by Campbell and Fiske's (1959) multitrait-multimethod approach to assessing the convergent and discriminant validity of measures. Campbell and Fiske observe that whereas constructs are typically validated by evidence of "convergence" or agreement among different measures of the same construct, they also stress that a measures can be invalidated by correlating "too highly" with measures of other constructs from which the focal construct was intended to differ—evidence of a lack of "discriminant" validity.

A full-blown application of the analyses Campbell and Fiske advocate calls for several traits to be measured by each of several methods in the same study. The present survey employed one method (ratings elicited from by a single informant/firm) to measure a given construct (FA or SC), under each of 9 different conditions defined by combinations of media and function. For the purposes at hand, the latter set of conditions can be viewed as multiple "traits."

Table B1 depicts the structure of the half matrix of 36 (9x8/2) pairwise associations among the ratings of 9 unique media/function combinations with respect to the same construct that each informant reported for his/her firm (n=79 firms).The analyses that follows treats the half-matrix of 36 pairwise associations as a set of **single method/multi-trait** inter-correlations and partitions it so as to identify 3 sub-groupings of the pairwise correlations that delineate differences in the composition of the media/function pairs rated, represented as triangles or diagonals in Table B1.

- ***Mono -Media/Hetero- Function Triangles***: TD (Digital), TP (Print), & TV (Video)
  (3 triangles x 3 pairwise coefficients/triangle) = 9 coefficients
- ***Hetero- Media/Mono- Function Diagonals:*** XC (Creative), XP (Production), XB (Media Buying)
  (3 diagonals x 3 pairwise coefficients/diagonal)= 9 coefficients
- ***Hetero-Media/Hetero-Function Pairs Triangles***:
  (6 triangles x 3 pairwise coefficients/triangle)= 18 coefficients.

Table B1: Partitioning of Half Matrix of 36 Pairwise Gammas Among 9 Raying of the Same Construct

	DC	DP	DB		PC	PP	PB		VC	VP	VB
DC	1										
DP	TD	1									
DB	TD	TD	1								
PC	XC	?	?		1						
PP	?	XP	?		TP	1					
PB	?	?	XB		TP	TP	1				
VC	XC	?	?		XC	?	?		1		
VP	?	XP	?		?	XP	?		TV	1	
VB	?	?	XB		?	?	XB		TV	TV	1

The ensuing analysis involves comparisons of the magnitudes of bivariate correlations. Utilization of Pearson coefficients for such purpose is problematical since the range of values a Person correlation can assume is affected by skewness in the marginal distributions of the underlying variables. Such is the case with our rating scale measures of FA and SC and rather than relying on the Person correlations presented in Table 4, we use Goodman and Kruskal's (1954) "Gamma" measure of association for cross classifications. Gamma indicates "how much more probable it is to get like than unlike orders in two classifications when two individuals are drawn at random from a population" (p. 749). Gamma is zero in the case of independence and +l, if the population is concentrated along the lower left to upper right diagonal or -1, if concentrated on the upper left to lower right diagonal of a two-way cross classification table. The half matrices of gamma coefficients for the ratings of FA and SC are presented in Panels A and B of Table B2, respectively.

Table B2: Goodman and Kruskal's (1954) Gammas, N = 79

Panel A: Gammas for *Adjustment Frequency (FA)*

	1	2	3	4	5	6	7	8	9
1. Digital Creative	1.000								
2. Digital Production	0.798	1.000							
3. Digital Media Buy	0.808	0.722	1.000						
4. Print Creative	0.624	0.663	0.619	1.000					
5. Print Production	0.614	0.679	0.677	0.867	1.000				
6. Print Media Buy	0.555	0.611	0.734	0.731	0.794	1.000			
7. Video Creative	0.670	0.661	0.683	0.667	0.626	0.641	1.000		
8. Video Production	0.624	0.719	0.715	0.649	0.699	0.638	0.863	1.000	
9. Video Media Buy	0.708	0.668	0.826	0.627	0.648	0.732	0.680	0.849	1.000

Panel B: Gammas for *Similarity of Available Capabilities (SC)*

		1	2	3	4	5	6	7	8	9
1.	Digital Creative	1.000								
2.	Digital Production	0.842	1.000							
3.	Digital Media Buy	0.728	0.726	1.000						
4.	Print Creative	0.564	0.591	0.401	1.000					
5.	Print Production	0.566	0.679	0.527	0.732	1.000				
6.	Print Media Buy	0.587	0.553	0.789	0.473	0.584	1.000			
7.	Video Creative	0.550	0.566	0.383	0.437	0.512	0.386	1.000		
8.	Video Production	0.675	0.691	0.549	0.488	0.521	0.471	0.823	1.000	
9.	Video Media Buy	0.639	0.607	0.749	0.367	0.447	0.691	0.580	0.757	1.000

Table B3: Summary of Analyses of Pairwise Associations among Ratings of "Frequency of Adjustment" and "Similarity of Capabilities".*

	Construct	Ratings
	Frequency of Adjustment	Similarity of Available Capabilities
All 36 Pairwise G's	.678 Range = .312 = (.867 - .555)	.573 Range = .475 = (.842 - .367)
Mono-Media/Hetero-Function ʌ's (3 ʌ's x 3 G's/ʌ = 9 G's)	.798 Range = .187 = (867 - .680) W = .333 (p-value=.528)	.728 Range = .369 = (.842 - .473) W =.778 (p-value =.194)
Hetero-Media/Mono-Function Diagonals (3 Diag. x 3 G's/Diag.= 9 G's)	.699 Range = .202= (.826 - .624) W = .778 (p-value =.194)	.690 Range = .352 = (.789 - .437) W = .778 (p-value =.194)
Hetero-Media/Hetero-Function Pairs (6 ʌ's x 3 G's /ʌ = 18 G's)	.615 Range = .160 = (.715 - .555)	.538 Range = .308 = (.675 - .367)

Notes: Cell entries are: Median Gamma (G)/Range= (Max G-Min G), W= Coefficient of Concordance, the p-values reported here are associated with the one-sided test where the null hypothesis is the independence of the rankings.

Table B3 summarizes the observed values of the set of gamma coefficients falling within each of the 3 groups defined above. For *case A*, ***(Mono-Media/Hetero-Function Triangles)***, each triangle contains 3 pairwise associations between ratings for each of the 3 **different functions (**Creative, Production, and Media Buying) and a ***single medium***, while each **triangle** relates to a different medium (digital, print, or video). Similarly for *case B*, (***Hetero-Media/Mono-Function Diagonals)***, *each* **diagonal** contain 3 pairwise associations between each of the 3 **different media** and a **single function**; with the single reference function (creative, production, and media buying) varying across the 3 diagonals. Hence, whereas *case A* relates to *discrimination across media*, *case B* relates to *discrimination across functions*.

Inspection of the values of the median and range of the gamma coefficients across the cells corresponding to the tripartite classification explained above reveals two discernible patterns of differences.

1. The **median values of the gammas for FA are greater than those for SC** not only for the total set of 36 coefficients but also for each of the three sub-categories (***Mono-Media/Hetero-Function, Hetero-Media/Mono-Function,*** and ***Hetero-Media/Hetero-Function).*** However, with respect to **dispersion** (as reflected by the corresponding values of **range)**, the reverse ordering holds. **The range or spread between the maximum and minimum gammas for the SC ratings exceeds than for FA ratings, overall and for each of the 3 sub-categories.**

   Such a pattern of difference in the levels and dispersions of the pairwise association would arise if informants were more discriminating in the making assessments of SC than of FA. To investigate this possibility, we analyzed the extent to which the rankings of the pairwise coefficients varied across the 3 **mono-media/hetero-functions triangles** (*case A*).The gamma coefficients within each triangle were for the same set of 3 different pairs of the same 3 functions while the media condition was different for each triangle. A similar analysis was made of the consistency of the ranking for the 3 **hetero-media/mono-media diagonals** comprising *case B*.

   In Table B3, we report the value of Kendall coefficient of concordance (W), a non-parametric measure of the degree of agreement among m (m>2) sets of rankings of k objects that varies from 0 to +1 (Kendall and Babington Smith 1939, Teles 2012). W=1 when there is perfect agreement among the m sets of rankings, if there are no ties. W can take the minimum value of zero when there is no agreement, only if m(k+1) is even (Teles 2012, p.751). The latter condition is satisfied here since k=3=m for both the mono-method/hetero-functions triangles (*case A*) and the hetero-media/mono-media diagonals **(***case B***).** Following Marozzi's (2014) recommendation for small sized samples, we test the null hypothesis that rankings are independent of one another and report p-value derived from the exact distribution of the test statistic (W) by permuting the k ranks in all possible ways, as discussed in Kendall and Babington Smith (1937, pp.277-278.) In the present case where m=k=3, this test is equivalent to testing that the population value of W=0.

   Referring to the p values shown in Table B3, we find that for both the FA and SA ratings, the null hypothesis of no agreement (W=0) cannot be rejected at conventional levels of significance in either case A or case B. In all four cases, the associated p values (one tail test) exceeded levels that might suggest even marginal significance; the smallest p value being 0.194. Thus, these results do not indicate that informants' ratings of FA and SC differed with respect to discriminant validity.

2. The levels of the median values for these ***Hetero-Media/ Hetero-Function*** Gammas shown in the Table B3 are .615 and .538 for the "Frequency" and "Similarity" ratings, respectively. What accounts for this substantial level of association across this set of seemingly unrelated if not independent pairings of media/function conditions? Variations across firms attributable to differences among firms with respect to the latent construct (i.e., FA or SC), or to differences among informants with respect to some source of systematic or random measurement error?

   We are unable to assess this issue directly with the data available from the study we conducted, which represented a single method/multi-construct design. However, the problem could be addressed by a study that employed a full-scale multi-method/multi construct design that would allow both convergent and discriminant validity to be assessed and enable comparisons to be made between the influence on pairwise associations of systematic and/or random measurement error components present in the ratings vs. that attributable to covariation between the "true" values of the underlying constructs.