ONLINE APPENDICES TO

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.

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

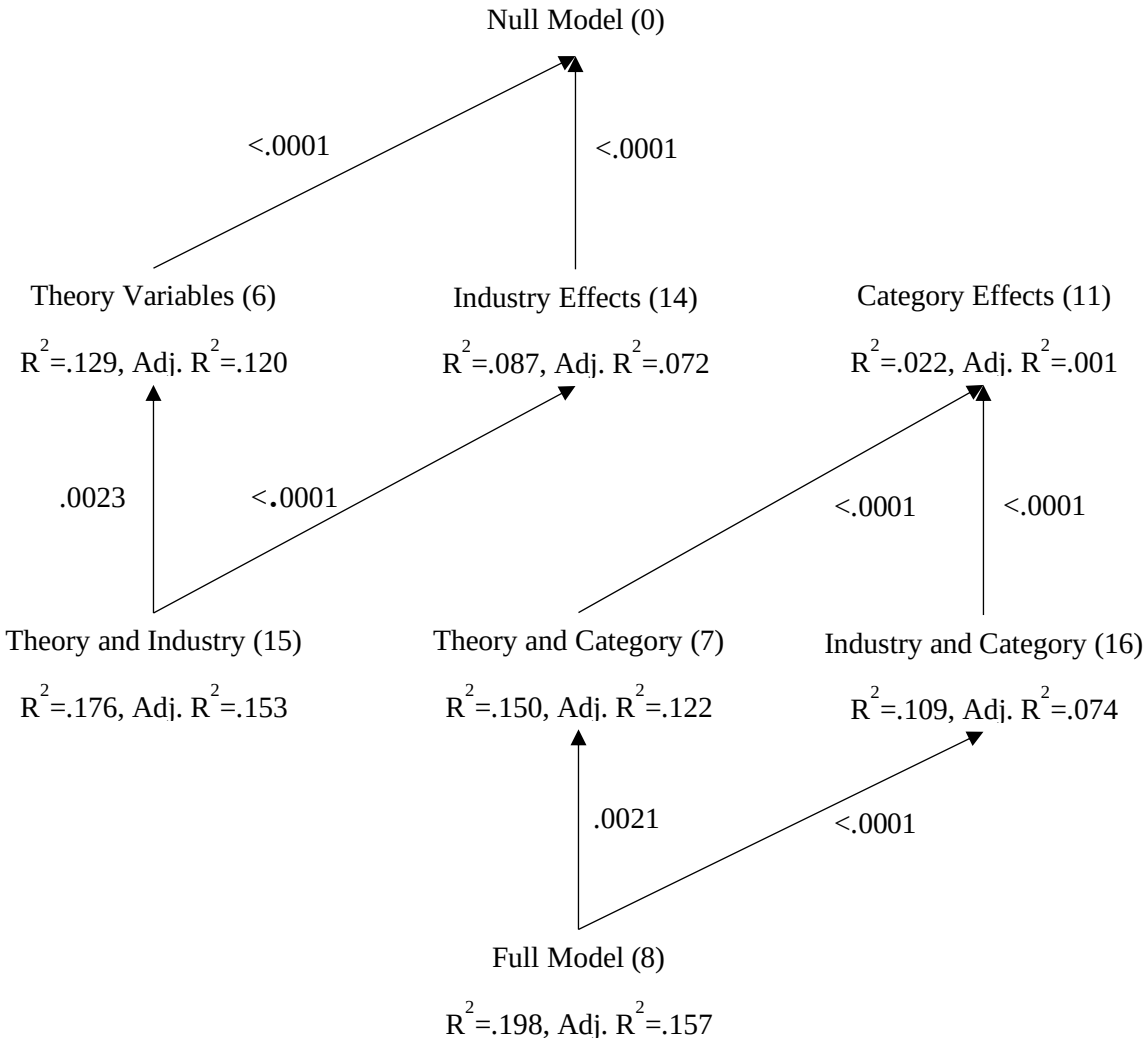
Table A5: Additional Regression Models for Figure 1 and Figure A1

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.