10.1287/msom.1060.0150

Store Manager Incentive Design and Retail Performance: An Exploratory Investigation.

Technical Appendix

Alternative Model Specifications and Robustness Checks

We estimated the parameters of the two models – one predicting shrink and one predicting sales-separately. Because all observations were drawn from the same firm, there is a possibility of correlation between the error terms in the two equations. The existence of such correlation suggests the two models should be treated as seemingly unrelated regressions (SUR). We found no correlation between the residuals of the fitted model predicting sales and the fitted model predicting shrinkage ($r=0.05$, $p=0.43$). Consequently, there was no need to estimate SUR parameters [See Greene (2000) for details on this methodology].

It may be argued that the impact of the incentive change on shrink is mediated by the observed change in sales. We therefore test for the existence of a mediation effect using the criteria proposed by Baron and Kenny (1986). Regressing SHRINK on LNSALES, TWTR, ADJ_INVEN while controlling for differences across stores, we find no relationship between LNSALES and SHRINK ($t=-0.86$, $p=0.3918$). Moreover, TWTR remains a statistically significant predictor ($t=5.94$, $p<0.0001$) of SHRINK despite the presence of LNSALES in the model. These findings suggest that that LNSALES does not mediate the effect of TWTR on SHRINK but rather there exists a strong direct relationship between the change in incentives and SHRINK. Moreover, results from the Sobel test (Sobel 1982) allow us to conclude that no statistically significant mediation effect exists (test statistic=-0.84, $p=0.40$).

Because advertising expenditure for some months had to be estimated by us in collaboration with BMS and Tweeter managers, we check the robustness of our findings using several alternative estimates of advertising expenditure. First, we simply replace the original estimates with values that are thirty percent greater (Model 1, **Table A1**). Second, we replace the original estimates with values that are thirty percent less (Model 2, **Table A1**). Comparing Models 1 and 2 (**Table A1**) to the model employing our original advertising estimates (Model 4, **Table A1**), we find no substantive differences in the parameter estimates or standard errors suggesting our results are robust to errors in managerial estimates.

We then employed an approach similar to multiple imputation where we simulate several plausible values for the missing advertising expenditure observations, fit our model using these

simulated values, and integrate our findings into a final result (Rubin 1996). Because the values of advertising that were missing were all prior to the change in incentive design, automated multiple imputation techniques were infeasible. These techniques require missing values to be missing at random (Ibrahim et al. 2005). We therefore treated the original managerial estimates as baseline values and added uncertainty around these values. For each managerial estimate we added a random number drawn from a normal distribution with a mean of zero and variance of 0.07. This variance represents the variance of the residual error from a model predicting advertising expenditure as a function of LNRETAIL, LNINVEN, and store dummies. We then fit our chosen model predicting sales using these new values for advertising expenditure. After executing these steps five different times, we had five sets of regression coefficients (e.g., parameter estimates and covariance matrices). The number of sets required for imputation is usually quite small, often 3 or 5 (Rubin 1987). We then pooled these fitted results and calculated the mean, variance, and p-value of this set of regression coefficients (Model 3, **Table A1**). Again, we did not find any substantive change in the coefficients or standard errors when comparing this model to the model using our original advertising expenditure estimates (Model 4, **Table A1**).

Table A1: Predicting the Natural Log of Sales by Incentive System and Select Independent Variables Using Alternative Estimates for Advertising Expenditure (n=264).

	Parameter Estimates (Standard Errors)			
Fixed Effects	**Model 1: Addition of 30%**	**Model 2: Reduction of 30%**	**Model 3: Multiple Imputation**	**Model 4: Original Estimates**
TWTR	0.11*** (0.03)	0.10* (0.04)	0.11** (0.03)	0.11*** (0.03)
LNRETAIL	1.48*** (0.13)	1.47*** (0.13)	1.48*** (0.13)	1.48*** (0.13)
LNINVEN	0.72*** (0.11)	0.72*** (0.11)	0.72*** (0.11)	0.71*** (0.11)
LNAD	0.001 (0.002)	0.003 (0.004)	0.002 (0.003)	0.002 (0.003)

~p<0.10 * p<0.05 **p<0.01 ***p<0.001

Store dummies were included in each of these models but coefficients and standard errors are not shown here.

References

Baron, R.M., D.A. Kenny. 1986. The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations. *Journal of Personality and Social Psychology* **51** 1173-1182.

Greene, W.H. 2000. *Econometric Analysis*. Prentice Hall, Upper Saddle River, NJ.

Ibrahim, J.G., M. Chen, S.R. Lipsitz, A.H. Herring. 2005. Missing-Data Methods for Generalized Linear Models: A Comparative Review. *Journal of the American Statistical Association* **100** 332-346.

Rubin, D.B. 1996. Multiple Imputation After 18+ Years. *Journal of the American Statistical Association* **91** 473-489.

Rubin, D.B. 1987. Multiple Imputation for Nonresponse in Surveys. John Wiley & Sons, New York, NY.

Sobel, M.E. 1982. Asymptotic Confidence Intervals for Indirect Effects in Structural Equation Models. In S. Leinhardt (Ed.) *Sociological Methodology* (p. 290-312). Washington DC: American Sociological Association.