Better marketing for people.

MANAGEMENT SUMMARY

Mixed-Methods MMM

The Utilization of Ridge Regression
in Marketing Mix Modeling

June 2024

Introduction to Mixed-Methods MMM

1 Ongoing developments in media fragmentation and ad tracking challenges (e.g. Apple's ATT and Google's end of third-party cookies) prompt advertisers to adopt signal-resilient solutions like dynamic Marketing Mix Modeling (MMM).

2 MMM options include partnering with agencies, vendors, or developing in-house MMM solutions with open-source codes such as Meta Robyn and Google Meridian.

3 Automated MMM methods, driven by mainly machine learning methods can handle an increasing number of variables by prioritizing high-level modeling selection output results to address issues such as model overfitting.
However, these models tend to fail to generate statistically robust models at the individual variable level.

4 Bynd Consulting proposes an innovative two-stage modeling approach by combining conventional MMM methods with contemporary ML-related/Ridge regression techniques to produce statistically robust model results with respect to individual input variables (including online media channels and platforms).

Challenge

1 Generally, MMM faces challenges as media fragmentation increases. This can limit the inclusion of smaller impact factors (e.g. online media channels) in traditional models due to data constraints and multicollinearity issues.

2 To address this issue, a novel approach combining contemporary machine-learning techniques with conventional statistics is proposed.

3 This approach involves a two-stage modeling process, where online variables are quantified separately in a second modeling stage (auxiliary model) while non-digital media variables serve as control factors in the first modeling stage (reference model).

4 To demonstrate the proposed method, the effectiveness of this method will be applied on actual sales data (FMCG advertiser, Germany), and various media channels (including TV and digital platforms) functioning as input variables.

5 Finally, this study will compare the produced output results obtained through the proposed mixed-methods approach vs more traditional and contemporary methods.

Proposed Method: Two-stage Mixed-Methods MMM (1)

1 Historically, Marketing Mix Modeling (MMM) has evolved over time, ranging from traditional OLS methods to more contemporary regularization techniques in order to account for fragmented advertising channels.

2 The traditional OLS method provides variable-specific statistical significance results but struggles with smaller digital channels, while regularization techniques lack consensus on model creation.

3 The proposed modeling approach combines OLS with ridge regression techniques, leveraging the strengths of both methods.

4 Two-stage modeling involves applying OLS methods in the first stage to prioritize robustness and transparency, followed by applying ridge regression techniques to effectively quantify (smaller) digital channels in the process.

5 This mixed-method approach ensures the use of statistical significance while extending the number of potential variables in the model.

Proposed Method: Two-Stage Mixed-Methods MMM (2)

Stage 1

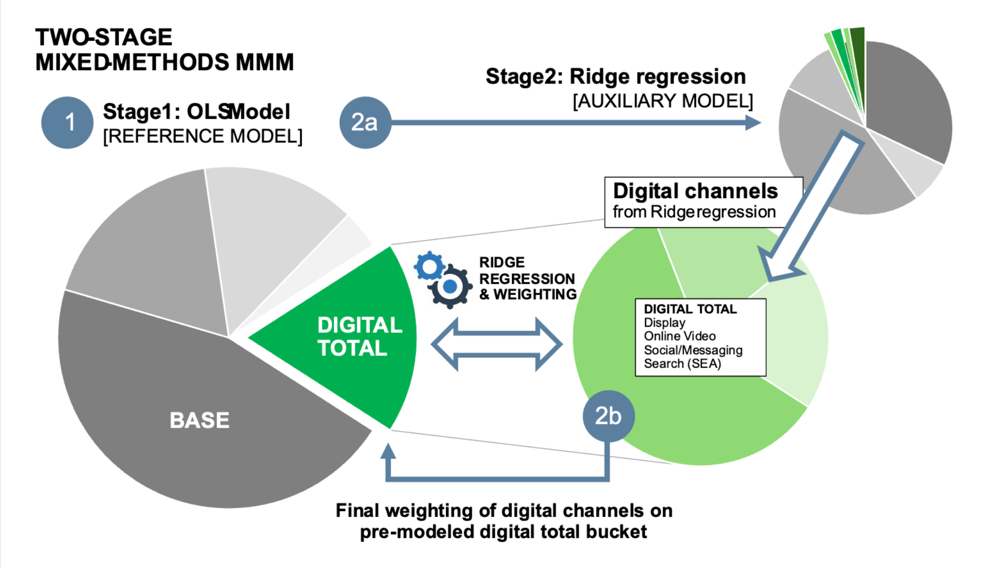
- Creation of a traditional OLS model by using a "total digital bucket" for grouping digital channels, serving as the reference model.

Stage 2a

- Development of an independent ridge-regression model to break out individual digital channels while maintaining statistical significance as demonstrated in 2(b).

Stage 2b

- Combining both model outputs by weighting the modeled digital variables against the pre-modeled digital total variable. This approach ensures statistical significance of modeled variables and addresses issues such as creating model bias and variance associated with regularization methods.



The proposed two-stage mixed-methods modeling approach allows for quantifying the impact of digital channels within a statistically robust modeling framework

Excursus: Introduction to Ridge Regression

Ridge regression is a recent modification of classical linear regression.

In linear regression, optimal weight factors are determined for a linear combination of independent variables to predict a dependent variable using least squares methods.

Practical application constraints include having fewer variables than data points for matrix invertibility.

Ridge regression introduces a regulating factor λ to overcome this constraint.

Lambda [λ] allows the model to absorb more variables, with coefficients often becoming positive as λ increases.

Ridge regression provides ”authentic" regression results despite potential deterioration in model fit (e.g., R^2) with increasing λ.

Suitable λ value must be determined through experimentation.

Strength of coefficients decreases as λ increases, but coefficient ratios remain relatively stable above a certain λ value.

Variables used in the Ridge regressions approach may not be statistically significant, and significance tests have limited interpretability.

A Sequential Modeling Approach is Recommended:

Consolidate "granular" variables into a collective sum variable using classical multiple linear regression.

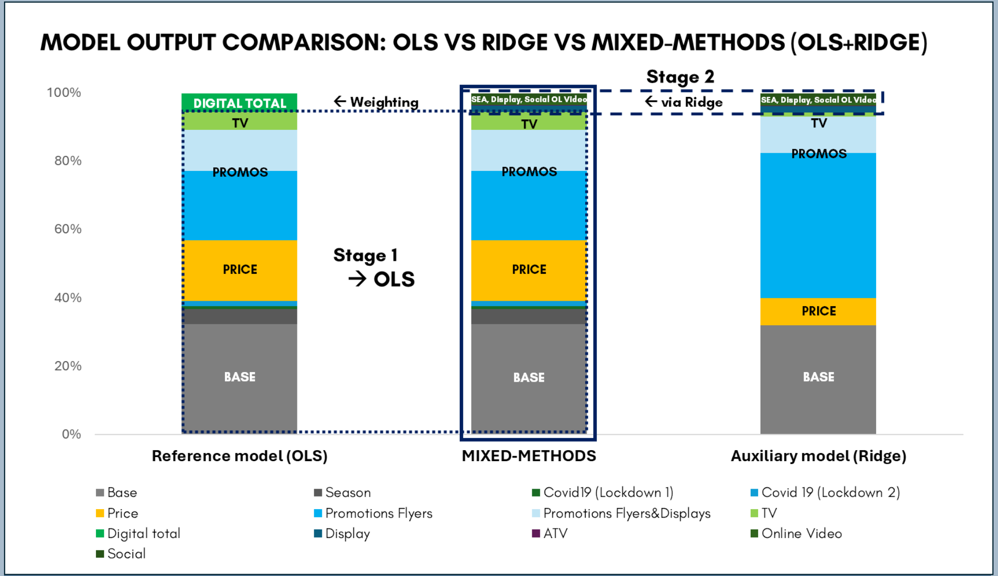
Compute a separate ridge regression model with an appropriate correction factor to address the effects of the "granular" variable.

Partition sum variables based on model coefficients determined through ridge regression techniques.

Advantages include establishing statistical significance for sum variables, stable variable relationships across λ variations, and unchanged effects of other significant variables.

Two-Stage Modeling Approach: Sales Decomposition

- **Example:** Produced sales decompositions via the proposed two-stage modeling method
- After the final weighting process for digital channels, respective sales decompositions for the final model can be calculated using the two-stage modeling approach.
- The illustration shows the final sales decomposition (→ Mixed-Methods) based on the OLS and ridge-only models as conducted in the two-stage modeling process.
- This methodology ensures a comprehensive understanding of the impact of digital channels on sales while maintaining statistical significance and stability in variable relationships.



The final model and sales decomposition can be determined by combining OLS and Ridge regression output results

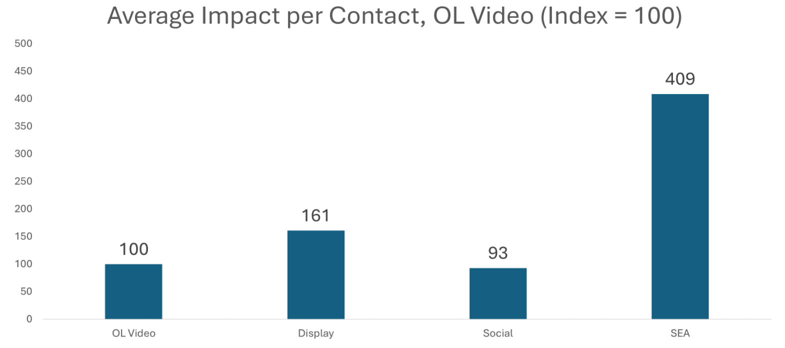
Cross-Case Analysis of Digital Channel Effects*

- Applying ridge regression techniques in the "Mixed-Methods MMM" approach yields stable results across cases.
- Average impact per contact is determined and normalized against OL Video coefficients. SEA/Search exhibits the strongest effect per contact due to end-of-journey searches.
- Social/messaging shows lower impact per contact but also lower costs per contact.
- Display advertising demonstrates the greatest variation in effect per contact, likely due to specific promotional offers. ROAS is indexed across cases, with OL Video ROAS set to 100.
- SEA shows the highest efficiency in ROAS, followed by social/messaging.
- ROAS is influenced by factors such as effect per contact, advertising pressure, and cost per contact.

Ridge regression techniques as applied in the two-stage mixed-methods approach provides stable insights into digital channel effects, with SEA (followed by social media) emerging as highly impactful in both impact per contact and ROAS.

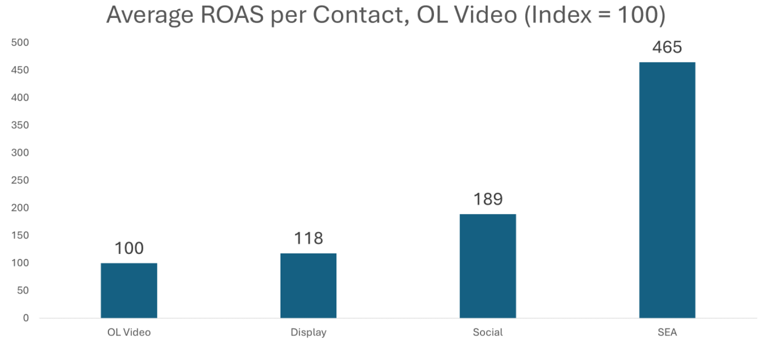
*Data base analysis: Cases range across 4 industry sectors (beverages, pharmaceuticals, FMCG, retail), with annual gross total media spend ranging from 2.8m to 109m in the German market.

Data base analysis (Impact)



Average Impact per Contact, n= 33

Data base analysis (ROAS)



Average ROAS per Contact, n = 17

General Recommendations for Ridge Regression

1 Selecting the correct lambda factor: Use the smallest lambda value where coefficients for media variables have positive signs and remain stable to each other.

2 Maintaining carry-over factors: Coefficients in ridge regression respond more strongly to changes in carryover. Use equal carryover rates for all variables unless specific indications suggest otherwise.

3 Variable set: Ensure maximum comparability by keeping the model's variable set consistent, except for variables involved in the split process.

4 Same scale or unit of variables: Ensure variables involved in ridge regression are measured in the same unit to avoid discrepancies in results.

5 → Following these recommendations ensures the reliability and consistency of results when employing ridge regression in MMM analyses.

Summary & Conclusion

Summary

- Ridge regression applied in a two-step approach generates accurate quantified results for "granular” media measures.
- This approach avoids overestimation and underestimation of effects.
- Initial model assessment determines the significance of such measures.

Conclusion and Future Research Directions

- While the proposed two-step modeling approach performs adequately, future research could enhance its effectiveness.
- Tasks for future research includes:
  - Specifying error probability for ridge regression coefficients.
  - Estimating optimal lambda based on data structures.
  - Defining requirements for variables to be partitioned.
  - Calculating maximum number of variables for partitioning.

Technical White Paper

White Paper available for download:

[bynd Consulting]

[Snapchat for Business]

Two-stage Mixed-Methods MMM:
The Utilization of Ridge Regression in Marketing Mix Modeling

MIXED - METHODS MMM

The Utilization of Ridge Regression in Marketing Mix Modeling

Prof. Dr. Alexander Preuss[1]   Antonio Prada[2]

Abstract

Since the introduction of Marketing Mix Modeling (MMM) in the advertising industry, it has undergone a significant evolution, transitioning into a more intricate measurement solution over time. Initially grounded in established econometric methodologies originating from academia, MMM has incorporated specific advertising research techniques, including variable transformation and the modeling of saturation and carryover effects, which are now integral components. Presently, the landscape of MMM confronts ongoing challenges such as the emergence of diverse digital platforms and the management of extensive datasets, coupled with current issues related to signal resiliency and reduced tracking possibilities. As a result, the domain of MMM is increasingly influenced by data science, featuring the integration of automation and machine learning (ML) techniques. Many practitioners have shifted from creating basic Ordinary Least Squares (OLS) models to more advanced and multifaceted models. Within the realm of frequentist modeling, the utilization of contemporary statistical and ML-related methods, involving techniques such as regularization methods (L1 - LASSO and L2 - Ridge), has gained traction to address issues with multicollinearity due to media fragmentation. However, using these methods alone may not consistently produce robust model results concerning crucial variables represented in a model. In this paper, we propose a mixed-methods approach to MMM, combining traditional OLS methods with contemporary statistical and ML-related elements, such as ridge regression, through a two-stage modeling process.

Authors

**Prof. Dr.
Alexander Preuss**

Professor of Quantitative Methodology, Fresenius University of Applied Sciences, Hamburg, Germany and Managing Partner Marketing Science (bynd GmbH, the strategy and technology consultancy of Omnicom Media Group Germany

alexander.preuss@bynd.consulting

Antonio Prada

Group Manager Marketing Science (Germany), Snap Inc.

aprada@snap.com