A quantile regression perspective on consumer heterogeneity

Cristina Davino
Federico II Univ. of Naples
Dep. Economics and Statistics
cristina.davino@unina.it

Rosaria Romano
Federico II Univ. of Naples
Dep. Economics and Statistics
rosaroma@unina.it

Domenico Vistocco
Federico II Univ. of Naples
Dep. Political Science
domenico.vistocco@unina.it

Table of contents

A case study on muscadine juices

Case study product set

Ten muscadine grape juices product[1].

ProductID	Overall	Appearance	Aroma	Color	Flavor
Black Beauty	6.92	6.70	6.60	6.48	6.65
Carlos	6.78	6.75	6.70	6.50	6.73
Granny Val	7.07	6.87	6.75	6.60	6.85
Ios	6.90	7.47	7.02	7.42	6.75
Nestitt	6.62	7.33	6.47	7.23	6.43
Post Red	6.25	7.35	6.40	7.28	6.02
Post White	6.27	6.43	6.07	6.10	5.93
Southern Home	6.88	6.93	7.02	6.53	6.92
Summit	6.95	7.02	6.90	6.93	6.90
Supreme	6.57	6.95	6.68	6.73	6.33

Hedonic means for 10 *muscadine grape juices.*

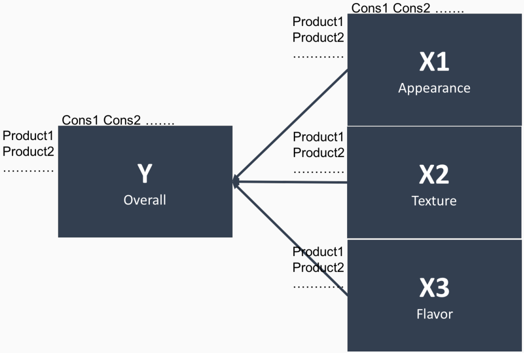
Consumer Testing

Sixty consumers were asked to evaluate the **appearance**, **flavor**, **texture**, **overall impression** of each sample on a 9-point hedonic scale with 9 being "like extremely" and 1 "dislike extremely".

[1] Meullenet, J. F., Xiong, R., Findlay, C. J. (2008). Multivariate and probabilistic analyses of sensory science problems (Vol. 25). John Wiley & Sons.

Case study objective

To provide the product developer with clear directions for improving products, it is necessary to identify sensory attributes that are important to product liking.



Analysing relations between specific and overall liking:

- on a **global level** (full panel);
- on an **individual** level to detect consumers' segments.

The toolkit: quantile regression for handling heterogeneity

Quantile Regression

Quantile Regression[2] (QR) can be considered an extension of conditional mean models to the whole conditional distribution of the response variable:

$$Q_\theta(\hat{\mathbf{y}}|\mathbf{X}) = \mathbf{X}\hat{\beta}(\theta)$$

- $Q_\theta(.|.)$ is the conditional quantile function for the θ-th quantile.
- θ is a given conditional quantile, with $0 < \theta < 1$.

Estimates in QR follows the classical interpretation:

Each $\hat{\beta}(\theta)$ coefficient represents the rate of change in the θ-th conditional quantile of the dependent variable per unit change in the value of the j-th regressor ($j = 1, \ldots, p$), holding the others constant.[a]

[a] Davino, C., Furno, M., Vistocco, D. (2013). *Quantile regression: theory and applications*. Wiley.

[2] Koenker, R., Bassett Jr, G. (1978). Regression quantiles. *Econometrica*, 46, 33-50.

Quantile Regression in consumer study

Previous studies

QR has been used in consumer studies for estimating the conditional quantiles of liking:

- Davino C., Romano R., Naes T. (2015).
  The use of quantile regression in consumer studies.
  Food Quality and Preference, **40**, 230-239:
  - Relating segments of consumers obtained based on their acceptance pattern to additional consumer characteristics.
- Davino C., Romano R., Vistocco D. (2018).
  Modelling drivers of consumer liking handling consumer and product effects.
  Italian Journal of Applied Statistics, **30**, 359–372:
  - Assessing heterogeneity across product similarities.
- Davino C., Naes T., Romano R., Vistocco D. (submitted).
  A quantile regression perspective on preference mapping.

Quantile Regression in consumer study

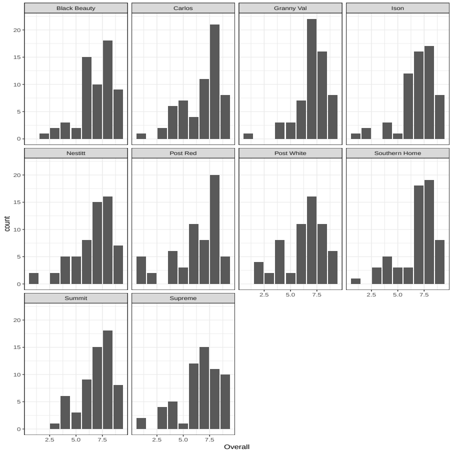
Current study

- QR to relate specific liking attributes to overall liking
  - **Global Model**
- QR-based strategy to detect **consumer effect** in such a relation[a]:
  1. Finding the best model for each consumer.
  2. Identification of consumers' segments.
  3. Assessing differences among the groups.

[a] Davino, C., Vistocco, M. (2018). Handling heterogeneity among units in Quantile Regression. Investigating the impact of students' features on university outcome, *Statistics and its interface*, 11, 541–556.

Main results

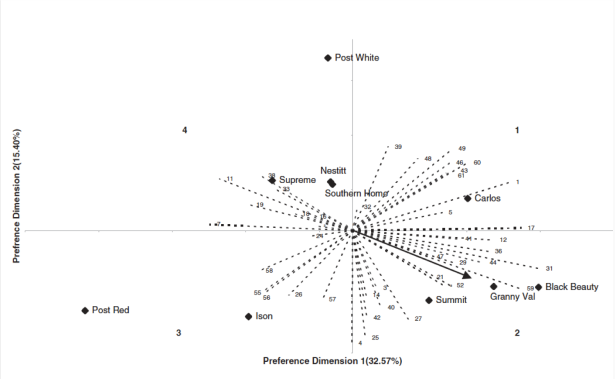
Overall liking: product differences



Univariate analysis of Liking

- All products were of high quality $\rightarrow$ left skewed distributions.
- Segments of consumers with varying preferences exist in this data.

Overall liking: individual differences



Multivariate analysis of Liking

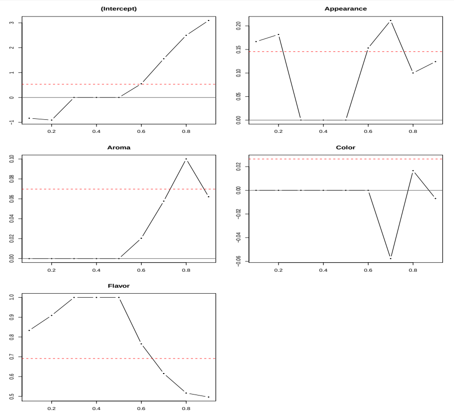
- The analysis clearly demonstrates that consumers showed preferences for different products.
- Consumer vectors are not concentrated in one direction of the map but are instead distributed in all four quadrants.

Quantile Regression in consumer study

Current study

- QR to relate specific liking attributes to overall liking
  - **Global Model**
- QR-based strategy to detect **consumer effect** in such a relation:
  1. Finding the best model for each consumer.
  2. Identification of consumers' segments.
  3. Assessing differences among the groups.

Relating specific attributes to overall liking - Global Model



Global model results:

- *Flavor* is the main driver of *overall liking*, followed by *appearance* and *aroma*.
- The impact of *flavor* decreases in the higher part of the distribution ($\theta \geq 0.50$), i.e. the less satisfied consumers are more influenced by this attribute.
- *Appearance* and *Flavor* show an opposite trend.

Quantile Regression in consumer study

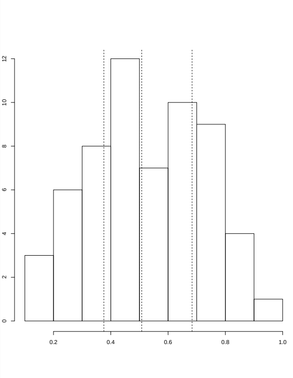
Current study

- QR to relate specific liking attributes to overall liking
  - **Global Model**
- QR-based strategy to detect **consumer effect** in such a relation:
  1. Finding the best model for each consumer.
  2. Identification of consumers' segments.
  3. Assessing differences among the groups.

Relating specific attributes to overall liking - Consumer effect

1. Find the best model for each consumer: consumer reference quantile

Average by consumer of the rank percentiles of each statistical unit with respect to the response variable: θ_g^{best} $g = 1, \ldots, m$



Best model for each consumer

$\theta_g^{best} = mean(rank\_perc(y_i))$
$i = 1, \ldots, n_g$ and $g = 1, \ldots, m$.

Relating specific attributes to overall liking - Consumer effect

1. Estimation of the best model for each consumer

- QR is carried out on the whole sample using the θ_g^{best} quantiles, namely estimating the 60 best models: $Q_\theta(\hat{\mathbf{y}}|\mathbf{X}) = \mathbf{X}\hat{\beta}({}_g\theta^{best})$

	Appearance	Aroma	Color	Flavor
C1				
C2				
...	...	...	...	
C60				

Relating specific attributes to overall liking - Consumer effect

2. Identification of consumers' segments

- Cluster Analysis on the matrix '*consumers* $\times$ *coefficients*'.

Clusters	θ min	θ max	θ^{best} mean	overall	appearance	aroma	color	flavor
1	0.10	0.21	0.15	4.30	5.45	4.62	5.30	4.05
2	0.23	0.51	0.38	6.04	6.38	6.11	6.12	5.91
3	0.54	0.62	0.58	7.19	7.14	6.80	6.96	7.10
4	0.64	0.69	0.66	7.53	7.80	7.24	7.54	7.33
5	0.71	0.73	0.72	7.79	7.87	7.56	7.81	7.40
6	0.75	0.88	0.81	8.17	8.20	8.18	8.15	8.05
7	0.98	0.98	0.98	8.90	8.00	8.50	7.80	8.90

Description of clusters by variables.

Relating specific attributes to overall liking - Comsumer effect

3. Assessing differences among groups

- QR is carried out on the whole sample using the representative quantiles assigned to each cluster.

Coefficient	$\theta = 0.15$	$\theta = 0.38$	$\theta = 0.58$	$\theta = 0.66$	$\theta = 0.72$	$\theta = 0.81$	$\theta = 0.98$
Intercept	-1.07*	0.00	0.41	1.09**	1.78**	2.54**	4.33**
Appearance	0.20	0.00	0.14	0.17 *	0.22**	0.11*	0.11
Aroma	0.07	0.00	0.04**	0.04	0.05**	0.11*	0.09*
Color	0.00	0.00	0.00	-0.03	-0.05	0.00	0.04
Flavor	0.80*	1.00	0.77*	0.68*	0.58*	0.51*	0.32*

('*' and '**' denotes coefficients significant respectively for $\alpha \leq 0.05$ and $0.05 < \alpha \leq 0.1$.)

Conclusion

Conclusions

Recap:

- Estimation of the group dependence structure
- Friendly interpretation of the final results
- Assessing differences among groups

Future avenues:

- Testing differences among groups
  - Separate testing on each slope coefficient.
  - Testing if all the slope coefficients of the groups are identical.
- Management of multi-group structures: external information.

THANK YOU FOR YOUR ATTENTION!

Cristina Davino
Federico II Univ. of Naples
cristina.davino@unina.it

Rosaria Romano
Federico II Univ. of Naples
rosaroma@unina.it

Domenico Vistocco
Federico II Univ. of Naples
domenico.vistocco@unina.it

Global Model