BAV Project

**Model estimation**

**Model from first-round paper**

**Market share model (MNL and MCI)**

*MNL*

1. The attraction Abt for brand b in period t and its associated market share sbt in an MNL model are given by

(1)

and

(2)

where αby is brand- and year-specific intercept (set to zero in one year for one brand), *Xbkt* is the matrix of explanatory variables (price, promotion, distribution, and ad-stock) with heterogenous response coefficients, Wbl is the matrix of brands’ attribute levels with homogenous response coefficients, εbt is the error term, and m is the number of all available brands b in time period t.

1. Linearizing equation (1) yields

(3)

1. Introduction of endogeneity

Gaussian copulas are used to model the correlation between the potentially endogenous regressors contained in X (Pricebt, Promotionbt, Distributionbt, and AdStockbt), and the error term . Following Park and Gupta (2012), we compute control variables, and add them to the matrix X. We apply the following transformation (see footnote 3, p. 572 in Park and Gupta 2012): X\*bkt=, where is the inverse distribution function of the standard normal, and H(·) is the empirical cumulative distribution function of Xbk.

1. We follow the base-brand approach (which yields the exact same coefficients as the log mean-centering approach advocated by Cooper and Nakanishi, 2010). We subtract an arbitrary base brand B (typically one that has maximum coverage in terms of number of observations), and transform

(4)

to:

(5)

or, equivalently

(6)

for all . Note that this sytem contains B-1 equations, and that is set to zero for all years. Estimation either via SUR, or Maximum Likelihood (better for model extensions, see below).

*MCI*

1. The attraction Abt for brand b in period t and its associated market share sbt in an MCI model are given by

(7)

and

(8)

1. Linearizing equation (7) yields

(9)

1. Introduction of endogeneity

Gaussian copulas are used to model the correlation between the potentially endogenous regressors contained in X (Pricebt, Promotionbt, Distributionbt, and AdStockbt), and the error term . Following Park and Gupta (2012), we compute control variables, and add them to equation (9). We apply the following transformation (see footnote 3, p. 572 in Park and Gupta 2012): X\*bkt=, where is the inverse distribution function of the standard normal, and H(·) is the empirical cumulative distribution function of Xbk.

1. Subtract an arbitrary base brand B, and transform

(10)

to:

(11)

or, equivalently

(12)

**Computation of SBBE for MNL and MCI model**

1. Sriram et al. (2007) ensures identification by setting αBy to zero for the store brand.
   1. “[The coefficient]” measures the incremental utility of a brand with respect to the store brand”
   2. Further, the authors make explicit that the coefficient need to be interpreted relative to the store brand’s BV in a given year, see on p. 64:



This confirms our intuition that the retrieved are to be interpreted as SBBE for the focal brand in year y, relative to “an identical variant offered with an identical marketing mix by the store brand in that [year]” (Sriram et al. 2007, p. 64)

* 1. Harald mentions that “a rank order of brands in terms of BV is not necessarily the same as a rank order of brands in terms of market share”; this statement is not to be interpreted as “market share” (as retrieved by exp(x)/(1+exp(x)), but as the “raw” market share in the data. The transformation, see above, is “monotonic”, so ranking in terms of BV or predicted market share must yield the same rank ordering.

1. Following Kamakura and Russel (1993), we can scale all N intercepts so that they sum to zero across all brands (see p. 17, note in Table 3).

**Extensions to include CBBE**

1. In a later stage, we can make and a function of CBBE instead of estimating dummy variables, e.g.

(7)

(9)

For identification, α0,b=1,y=1 and are set to zero.

1. Alternatively, a HLM model can be obtained by adding random components to and in (8) and (9)

(10)

(11)

where and