Comments

- On Page 19, the v_t vector should be $[v_{11}^G, v_{11}^S, v_{11}^Y, v_{11}^A, \ldots, v_{n1}^G, v_{n1}^S, v_{n1}^Y, v_{n1}^A]$. The superscripts for the components should not be $[v_{11}^{NI}, v_{11}^I, v_{11}^C, v_{11}^A, \ldots, v_{n1}^{NI}, v_{n1}^I, v_{n1}^C, v_{n1}^A]$. I think it is a typo.
- Most of the tedious work of representing in the form of a standard DLM is already given in the paper on page 19. So that's a relief.
- Paper assumes a basic prior on β's. No fancy hierarchical priors specified. This means the most of the analysis can be done via EM algo estimation.
- The authors use FFBS algo for sampling. This is already coded in `dlm` package, so the whole model can be done via `dlm`
- Given the dimensionality of the problem, the simulation might take a long time to converge. This is where a alternative simulation smoother could be of help.If the car segments and number of cars analyzed is far less than what is done here (21 cars and 4 categories), the estimates might converge quickly and my comment is superfluous.
- Prior distributions are mentioned in Web Appendix A. So, one can use these prior distribution as the initial state vector distribution in EM estimation procedure. Obviously one can ignore about using these priors and use EM to estimate even the initial state parameters. It is all up to the modeler who might base it on the context and data.
- What are the key aspects of analysis ?
  - Short term elasticities : Straightforward inference from coefficients
  - Overall conversion rate for each car : Straightforward inference from coefficients
  - Long term elasticities - Authors talk about some simulation method. I think they are talking about some kind of impulse response modeling. I doubt its utility though!.
  - Correlation analysis between ad spending and baseline convertibility is a very interesting visual.
  - Endogenity : Straightforward inference from coefficients
  - Environment variable effects : Again a straightforward inference from coefficients
  - Use of discriminant validity to show that Google trends is a good proxy.
- Analysis has been only based on coefficients of the equations. No analysis has been done on the error variance estimates. Surprisingly, they are not even mentioned in the appendix. Very surprising to me as the first thing I want to look at is the signal to noise ratio. I can only guess that may be they were not alarming and authors did not find it worth mentioning.
- The whole analysis is based on multivariate normal error variance. The structure is not made complicated by assuming some fancy hierarchical priors across car segments. This means the model estimation is straight forward and can be done via MARSS. If you have a decent number of data points, you can take some diffuse prior for the initial state vector distribution and run it through MARSS set up.
- The standard form DLM is already given on page 19. So, one can directly take that as an input, and use the default MARSS model to get the estimates.
- I have not gone through MARSS and have not worked with it. So, I am not sure how the correlation analysis between spending and baseline convertibility is obtained. In a Bayes setting it is straightforward as you compute the correlation between samples generated for each of the coefficients. In a MARSS setting, one might have to dig in to the output and see how to obtain the correlation between various coefficients. There must be an object that throws out covariance matrix of the parameters based on which one can do the relevant analysis. However since I have not used the package, I do not know what exactly are the names of the list objects that give these estimates.

Proposed Model

Variables

- Observation variables
  - Sales data : Y_{jt}
  - Vehicle-shopping related searches : S_{jt}
  - Shopping and non-shopping related searches for vehicle : G_{jt}
  - Current Ad spend : A_{jt}
- State variables :
  - Consumer prepurchase information interest : I_{jt} ($I_{jt} = \log Q_{jt}$)
  - Non Shopping interest parametrized : NI_{jt}
  - Baseline convertibility of consumer interest : C_{jt}
  - Interest conversion : R_{jt}

Observation Equations

$$\begin{aligned}
\log S_{jt} &= I_{jt} + v^S_{jt}, \quad v^S_{jt} \sim N(K^S_j, V^S_j) \\
\log G_{jt} &= I_{jt} + NI_{jt} + v^G_{jt}, \quad v^G_{jt} \sim N(K^G_j, V^G_j) \\
\log Y_{jt} &= C_{jt} + (1+\phi_{jt})I_{jt} + v^Y_{jt} \\
\log R_{jt} &= C_{jt} + \phi_{jt} I_{jt} + v^Y_{jt}, \quad v^Y_{jt} \sim N(0, V^Y_j) \ \text{: This is an Implicit equation}
\end{aligned}$$

State Equations

$$\begin{aligned}
I_{jt} &= \alpha^I_{jt} + \beta^I_j X_{jt} \\
\alpha^I_{jt} &= \delta^I_{j,1}\alpha^I_{j,t-1} + \delta^I_{j2} \sum_{j'=1, j'\neq j}^{n} \alpha^I_{j',t-1} + \delta^I_{j3}\log A_{jt} + \delta^I_{j4}\log \tilde{A_{jt}} + w^I_{jt}, \quad w^I_{jt} \sim N(0, W^I_j) \\
NI_{jt} &= \alpha^{NI}_{jt} + \beta^{NI}_j X_{jt} \\
\alpha^{NI}_{jt} &= \delta^{NI}_{j,1}\alpha^{NI}_{j,t-1} + \delta^{NI}_{j2} \sum_{j'=1, j'\neq j}^{n} \alpha^{NI}_{j',t-1} + \delta^{NI}_{j3}\log A_{jt} + \delta^{NI}_{j4}\log \tilde{A_{jt}} + w^{NI}_{jt}, \quad w^{NI}_{jt} \sim N(0, W^{NI}_j) \\
C_{jt} &= \alpha^C_{jt} + \beta^C_j X_{jt} \\
\alpha^C_{jt} &= \delta^C_{j,1}\alpha^C_{j,t-1} + \delta j_2{}^C \sum_{j'=1, j'\neq j}^{n} \alpha^C_{j',t-1} + \delta^C_{j3}\log A_{jt} + \delta^C_{j4}\log \tilde{A_{jt}} + w^C_{jt}, \quad w^C_{jt} \sim N(0, W^C_j)
\end{aligned}$$

Controlling Endogenity

$$\begin{aligned}
\log A_{jt} &= \alpha^A_{jt} + \beta^A_j X_{jt} + v^A_{jt} \\
\alpha^A_{jt} &= \delta^A_{j,1}\alpha^A_{j,t-1} + \delta^A_{j2}\log Y_{j,t-1} + \delta^A_{j3}\log\log \tilde{A}_{j,t-1} + w^A_{jt}, \quad w^A_{jt} \sim N(0, W^A_j)
\end{aligned}$$