**Online Appendix**

**Investigating Emergent Nested Geographic Structure in Consumer Purchases: A Bayesian Dynamic Multiscale Spatiotemporal Modeling Approach**

**Part I: Spatiotemporal clustering Algorithm (Ferreira et al. 2011)**

The Ferreira et al. 2011algorithm includes five steps to cluster regions at a given multi-scale level.

Step 1: Estimate the discrepancy measurement model (DMM). Assume the observations at time in two different sub-regions at multiscale level is denoted as and , respectively, for The DMM is defined as follows in the current study

  **(1)**

 **(2)**

where , , , , and are the observational variances in observations and , and are the systematic variances in the two systematic parameters and . In equation (1), it shows that is the shared mean process with weights and between the two times series of observations, while is the discrepancy between series of observations from the two subregions. The variance components and can be estimated from MLE method for the above dynamic linear model (equation (1) and (2)).

Step 2: The relative discrepancy between the two sub-regions are defined as the ratio between the two variance components and , that is, /. The smaller the ratio, the more similar the two sub-regions are.

Step 3: For each sub-region in the multiscale level , calculate the relative discrepancy between and all its neighboring sub-regions. Build a link between subregion and its neighboring sub-region that has the smallest relative discrepancy (highest similarity).

Step 4: If two sub-regions at the multi-scale level are connected through a path (one link or a sequence of links), these two subregions are defined as in the same cluster. If there is no path between two sub-regions, these two sub-regions are defined as in different clusters. Thus the clustering structure is defined at the multi-scale level .

Step 5: Start Step 1- Step 4 at the finest level and recursively repeated for the immediately coarser level of resolution till

**Part II: MCMC algorithm for the multiscale spatiotemporal model**

In the following discussion, we will present detailed derivation of the MCMC algorithm for parameters in the multi-scale model. The standard Bayesian theorem applies and by the conjugacy of the prior distribution, full conditional posterior distributions can be obtained and Gibbs sampling can be easily implemented.

1. The signal-to-noise coefficients for the latent mean process at the coarsest level .

The full conditional posterior distribution of is

,

where and

with mean equal to and variance equal to . The hyper-parameter = 3 and . By conjugacy, the conditional posterior distribution of (is also inverse gamma distribution with

1. The signal-to-noise ratio of the multi-scale coefficients . The full conditional posterior distribution is as the form

With ) and with mean equal and variance equal to , the conditional posterior distribution of is an inverse gamma distribution with

where is the generalized inverse of .

The posterior estimation on is obtained by the standard forward filtering, backward smoothing methods as specified in (West and Harrison 1995 page 570). The details are specified as follows. Under the multi-scale structure, two sets of dynamic linear models (DLM) are constructed to model the temporal evolutions. The first DLM is on the coarsest multi-scale level for and the latent mean process as

where and The parameters ’s, , have the implication of signal-to-noise ratios.

The second DLM is on the multi-scale coefficients with the following observational equation on and the system equation on as

with **,** where serves as a signal-to-noise ratio, and .

1. First, run the standard analysis forward from to as in the following steps (1)-(4). At each stage , compute and save the quantities .
2. Posterior at , ;
3. Prior at , , where and ;
4. One-step forecast: where and ;
5. Posterior at : where and , .
6. At time a vector of is sampled, then sequence backward from to compute the element and at each stage and generating a value of by where , , and .
7. The posterior estimation of the multi-scale parameters Since is distributed as a singular Gaussian distribution, the usual FFBS algorithm as in part 3 cannot be directly applied. FHB (2011) and FBH(2010) developed a revised version of FFBS which they term Singular Forward Filter Backward Sampler (SFFBS). We include the details from FBH (2010) and theorem 2 in FHB(2011) for reference.
8. Posterior at , ;
9. Prior at , , where and ;
10. Posterior at , , where and ;
11. Through step 1-3, a vector of is obtained by the usual forward filtering from to Then sequence from sampling from the posterior at and then sequence backward from to , sampling according to where

;

.

**Table 1 Three level clusters and the counties included**

|  |  |  |
| --- | --- | --- |
| **Level 1** | **Level 2** | **Level 3 (county)** |
| **1** | 1 | massachusetts,nantucket |
|  |  |  |
| **2** | 2 | new york,clinton |
|  |  |  |
| **3** | 3 | new york,new york |
|  |  |  |
| **4** | 34 | new york,kings ; new york,nassau ; new york,queens ; new york,suffolk |
|  |  |  |
| **5** | 4 | connecticut,fairfield ; connecticut,litchfield ; massachusetts,hampden ; massachusetts,hampshire |
| 5 | connecticut,hartford ; connecticut,middlesex ; connecticut,new haven ; connecticut,new london ; connecticut,tolland ; connecticut,windham ; massachusetts,worcester |
| 16 | massachusetts,barnstable ; massachusetts,norfolk ; massachusetts,plymouth |
| 18 | massachusetts,bristol ; rhode island,kent ; rhode island,newport ; rhode island,washington |
| 19 | massachusetts,essex ; massachusetts,middlesex ; massachusetts,suffolk |
| 49 | rhode island,bristol ; rhode island,providence |
| **6** | 6 | delaware,kent ; maryland,kent |
| 7 | delaware,new castle ; maryland,cecil ; maryland,harford |
| 8 | delaware,sussex ; maryland,somerset ; maryland,wicomico ; maryland,worcester |
| 14 | maryland,queen annes ; maryland,talbot |
| **7** | 9 | maryland,allegany ; pennsylvania,bedford ; pennsylvania,blair ; pennsylvania,cambria ; pennsylvania,huntingdon |
| 40 | pennsylvania,allegheny ; pennsylvania,armstrong ; pennsylvania,beaver ; pennsylvania,butler ; pennsylvania,clarion ; pennsylvania,crawford ; pennsylvania,erie ; pennsylvania,indiana ; pennsylvania,lawrence ; pennsylvania,mercer |
| 46 | pennsylvania,elk ; pennsylvania,warren |
| 47 | pennsylvania,fayette ; pennsylvania,greene ; pennsylvania,somerset ; pennsylvania,washington ; pennsylvania,westmoreland |
| **8** | 10 | maryland,anne arundel ; maryland,baltimore ; maryland,baltimore city ; maryland,calvert ; pennsylvania,york |
| 11 | maryland,carroll ; maryland,frederick |
| 12 | maryland,charles ; maryland,st marys |
| 13 | maryland,howard ; maryland,montgomery ; maryland,prince georges |
| 15 | maryland,washington ; pennsylvania,adams ; pennsylvania,franklin |
| 42 | pennsylvania,centre ; pennsylvania,clinton ; pennsylvania,columbia ; pennsylvania,lycoming |
| 44 | pennsylvania,cumberland ; pennsylvania,perry |
| 45 | pennsylvania,dauphin ; pennsylvania,lebanon ; pennsylvania,northumberland |
| 48 | pennsylvania,lackawanna ; pennsylvania,luzerne ; pennsylvania,monroe ; pennsylvania,pike ; pennsylvania,wayne ; pennsylvania,wyoming |
| **9** | 17 | massachusetts,berkshire ; massachusetts,franklin ; new york,columbia ; new york,rensselaer ; new york,washington |
| 28 | new york,albany ; new york,fulton ; new york,montgomery ; new york,saratoga ; new york,warren |
| 32 | new york,dutchess ; new york,greene ; new york,ulster |
| 38 | new york,schenectady ; new york,schoharie |
| **10** | 20 | new jersey,atlantic ; new jersey,cape may ; new jersey,cumberland |
| 21 | new jersey,bergen ; new jersey,hudson ; new york,bronx |
| 22 | new jersey,burlington ; new jersey,mercer ; new jersey,somerset |
| 23 | new jersey,camden ; new jersey,gloucester ; new jersey,hunterdon ; new jersey,salem ; pennsylvania,bucks ; pennsylvania,delaware ; pennsylvania,philadelphia |
| 24 | new jersey,essex ; new jersey,passaic |
| 25 | new jersey,middlesex ; new jersey,monmouth ; new jersey,ocean |
| 26 | new jersey,morris ; new jersey,sussex ; new jersey,warren ; pennsylvania,lehigh ; pennsylvania,montgomery ; pennsylvania,northampton |
| 27 | new jersey,union ; new york,richmond |
| 36 | new york,orange ; new york,putnam ; new york,sullivan |
| 37 | new york,rockland ; new york,westchester |
| 41 | pennsylvania,berks ; pennsylvania,carbon ; pennsylvania,schuylkill |
| 43 | pennsylvania,chester ; pennsylvania,lancaster |
| **11** | 29 | new york,broome ; new york,chenango ; new york,cortland ; new york,madison ; new york,onondaga |
| 33 | new york,herkimer ; new york,jefferson ; new york,oneida ; new york,oswego |
| **12** | 30 | new york,cattaraugus ; new york,chautauqua ; new york,erie ; new york,niagara ; pennsylvania,mckean |
| 31 | new york,chemung ; new york,genesee ; new york,livingston ; new york,monroe ; new york,steuben ; new york,wayne ; pennsylvania,bradford |
| 35 | new york,ontario ; new york,seneca ; new york,yates |
| 39 | new york,tioga ; new york,tompkins |



Figure 1 The study area. There are a total of 204 counties. 31 counties (grey) are excluded because either is no order (10 counties) or there are less than three years with order history between 1997-2001. The areas with blue are those with no neighbors (not connected by land within the study area.

**Figure 2 (legend at the end)**



1. (b) (c)



 (d) (e) (f)



 (g) (h) (i)



 (j) (k) (l)



 (m) (n) (o)



 (p) (q) (r)



 (s) (t) (u)



 (v) (w) (x)

Figure 2: The sum of log order values (y-axix) at level 1 clusters in the study area in Year 1997-2001 (x-axis). Each row represents one cluster from Level 1 from cluster 5-12. The first 4 clusters are not shown here as there are clustered because they do not have geographical neighborhood. Column one shows Level 1 total (a, d, g, j, m, p, s, v); Level one total disaggregated by Level Two clusters as shown in the second column (b, e, h, k, n, q, t,w); the empirical multiscale coefficients are shown in the third column (c, f, I, l, o, r, u,x).