## Appendix A: Model Estimation

Given the dynamic nature of the Hidden Markov Model, the multiple observations for one household are not independent over time and therefore the likelihood function must be computed over the whole sampling period of the household,

(A1) $$P\{(Y_o, Y_1, ..., Y_T); \lambda\} = \sum_{x_o, x_1, ..., x_T} P\{(Y_o, Y_1, ..., Y_T) \mid (x_o, x_1, ..., x_T); \lambda\} P\{(x_o, x_1, ..., x_T); \lambda\}.$$

Following the formulation of our model in (1)-(6), the likelihood function for a given household can be computed as

(A2) $$P\{(Y_o, Y_1, ..., Y_T); \lambda\} = \sum_X \left[ b_{x_o}(Y_o) b_{x_1}(Y_1) ... b_{x_T}(Y_T) \right] \left[ \pi_{x_o} a_{x_o x_1} a_{x_1 x_{12}} ... a_{x_{T-1} x_T} \right].$$

Since the latent class memberships in $X$ (the set of all the possible sequences of hidden states) are not observed, they are treated as missing data, using the E-M algorithm (Dempster, Laird and Rubin, 1977). We provide a brief description of the E-M algorithm below; readers interested in more details are referred to the excellent tutorial by Rabiner (1989).

E-STEP

In the E-step, we use the current parameter estimates to compute expectations for the missing data ($x_{nt}$). Because the missing data follow a (hidden) first order Markov process, we must use the forward-backward Baum-Welsh algorithm to compute the likelihood of the observed data sequence.

Forward

Let $\alpha_{nt}(i)$ represent the likelihood for the observed data sequence up to time $t$, conditional on household $n$ being in state $i$ at $t$. This likelihood can be computed recursively as,

(A3) $$\alpha_{no}(i) = \pi_i b_i(Y_{no})$$

(A4) $$\alpha_{n,t+1}(i) = \left[\sum_{j=1}^{M} \alpha_{nt}(j) a_{ji}\right] b_i(Y_{n,t+1}) \qquad t=0,1,2,\ldots,T-1\ ,$$

up to the last observation at time *T*. At that point the likelihood of the observed sequence can be computed as,

(A5) $$P\{(Y_{no}, Y_{n1}, \ldots, Y_{nT}); \lambda\} = \sum_{i=1}^{M} \alpha_{nT}(i)$$

<u>Backward</u>

In a similar way as in the forward sequence, we can compute $\beta_{nt}(i)$, the likelihood of the observed sequence backwards, starting from the last observation down to the current one at time *t*, conditional on household being at state *i* at the present time *t*, as

(A6) $$\beta_{nT}(i) = 1$$

(A7) $$\beta_{nt}(i) = \sum_{j=1}^{M} a_{ij} b_j(y_{n,t+1}) \beta_{n,t+1}(j) \qquad t=T-1, T-2, \ldots, 2, 1, 0$$

Once the forward and backward likelihoods are known, posterior probabilities for household *n* being at state *i* at time *t* can be computed as (Rabiner 1989, pg.263),

(A8) $$\gamma_{nt}(i) = \frac{\alpha_{nt}(i)\beta_{nt}(i)}{\sum_{i'=1}^{M} \alpha_{nt}(i')\beta_{nt}(i')},$$

and the posterior transition probabilities are obtained from

(A9) $$\zeta_{nt}(i,j) = \frac{\alpha_{nt}(i) a_{ij} b_j(Y_{n,t+1}) \beta_{n,t+1}(j)}{\sum_{i'=1}^{M}\sum_{j'=1}^{M} \alpha_{nt}(i') a_{i'j'} b_{j'}(Y_{n,t+1}) \beta_{n,t+1}(j')}$$

M-STEP

In the M-step we replace the missing data with their expected values and obtain new estimates for the parameters of the model via maximum likelihood estimation:

(A10) $$\hat{\pi} = \arg\max \sum_{n=1}^{N} \sum_{i=1}^{M} \gamma_{no}(i) \ln \pi_i$$

(A11) $\hat{B} = \arg\max \sum_{n=1}^{N} \sum_{T=0}^{T-1} \gamma_{nt}(i) \ln b_i(y_{nt})$

(A12) $\hat{A} = \arg\max \sum_{n=1}^{N} \sum_{t=0}^{T-1} \zeta_{to}(i, j) \ln a_{ij}$

This maximization step leads to the following estimates,

(A13) $\hat{\pi}_i = \sum_{n=1}^{N} \gamma_{no}(i) / N$

(A14a) $\hat{\theta}_{ik}^{(l)} = \frac{\sum_{n=1}^{N} \sum_{0 \le t \le T, y_{knt} = l} \gamma_{nt}(i)}{\sum_{n=1}^{N} \sum_{0 \le t \le T} \gamma_{nt}(i)}$ if item $k$ is multichotomous with $L$ categories

(A14b) $\hat{\theta}_{ik}^{(1)} = \frac{\sum_{n=1}^{N} \sum_{t=0}^{T} \gamma_{nt}(i) y_{nt}}{\sum_{n=1}^{N} \sum_{t=0}^{T} \gamma_{nt}(i)}$ and $\hat{\theta}_{ik}^{(2)} = \frac{\sum_{n=1}^{N} \sum_{t=0}^{T} \gamma_{nt}(i) y_{nt}^2}{\sum_{n=1}^{N} \sum_{t=0}^{T} \gamma_{nt}(i)} - (\theta_{ik}^{(1)})^2$ if item $k$ is continuous

(A15) $\hat{a}_{ij} = \frac{\sum_{n=1}^{N} \sum_{t=0}^{T-1} \zeta_{nt}(i, j)}{\sum_{n=1}^{N} \sum_{t=0}^{T-1} \gamma_{nt}(i)}$.