Online supplement to the paper

Semi-Parametric Independence Testing for Time Series of Counts and the Role of the Support

David Harris, Department of Economics, University of Melbourne

Brendan McCabe, Management School, University of Liverpool

This supplement provides Appendix B to the paper "Semi-Parametric Independence Testing for Time Series of Counts and the Role of the Support", which includes the proofs of Theorem A.1 and Lemmas A.1 and A.2 that were stated in Appendix A.

B Proofs of Theorem A.1, Lemmas A.1 and A.2

B.1 Proof of Theorem A.1

Using (2) the log likelihood ratio can be expressed

$$\log \frac{L_T(\theta_T h)}{L_T(\theta_0)} = \sum_{t=1}^{T} \left(\log \sum_{k=0}^{y_t \wedge y_{t-1}} \mathrm{Bi}\left(y_{t-1}, k; \beta_T(h_\beta)\right) \frac{\pi_{T,y_t-k}(h_\pi)}{\pi_{y_t}} \right).$$

Applying the mean value expansions

$$\begin{aligned} (1-\beta)^n &= 1 - n\beta + \binom{n}{2}\beta^2 - \binom{n}{3}\beta^3 (1-\beta^*)^{n-3}, && \text{(A.32)} \\ n\beta(1-\beta)^{n-1} &= n\beta - 2\binom{n}{2}\beta^2 + 3\binom{n}{3}\beta^3 (1-\beta^{**})^{n-2}, && \text{(A.33)} \\ \binom{n}{2}\beta^2(1-\beta)^{n-2} &= \binom{n}{2}\beta^2 - 3\binom{n}{3}\beta^3 (1-\beta^{***})^{n-1}, && \text{(A.34)} \end{aligned}$$

to the $k = 0, 1, 2$ terms of the inner summation and collecting powers of T gives

$$\sum_{k=0}^{y_t \wedge y_{t-1}} \mathrm{Bi}\left(y_{t-1}, k; \beta_T(h_\beta)\right) \frac{\pi_{T,y_t-k}(h_\pi)}{\pi_{y_t}} = 1 + T^{-1/2} a_{1,t} + T^{-1} a_{2,t} + T^{-3/2} a_{3,t},$$

where

$$\begin{aligned}
a_{1,t} &= h_\beta y_{t-1}\left(\frac{\pi_{y_t-1}}{\pi_{y_t}}-1\right)+\left(h_{\pi,y_t}-\sum_{j\in\mathcal{U}}\pi_j h_{\pi,j}\right),\\
a_{2,t} &= h_\beta^2\binom{y_{t-1}}{2}\left(1-2\frac{\pi_{y_t-1}}{\pi_{y_t}}+\frac{\pi_{y_t-2}}{\pi_{y_t}}\right)\\
&\quad -h_\beta y_{t-1}\left(\left(h_{\pi,y_t}-h_{\pi,y_t-1}\right)-\left(\frac{\pi_{y_t-1}}{\pi_{y_t}}-1\right)\left(h_{\pi,y_t-1}-\sum_{j\in\mathcal{U}}\pi_j h_{\pi,j}\right)\right),\\
a_{3,t} &= h_\beta^3\binom{y_{t-1}}{3}\left(-\left(1-\beta_T^*\right)^{y_{t-1}-3}\frac{\pi_{T,y_t}\left(h_\pi\right)}{\pi_{y_t}}+3\left(1-\beta_T^{**}\right)^{y_{t-1}-2}\frac{\pi_{T,y_t-1}\left(h_\pi\right)}{\pi_{y_t}}\right.\\
&\quad \left.-3\left(1-\beta_T^{***}\right)^{y_{t-1}-1}\frac{\pi_{T,y_t-2}\left(h_\pi\right)}{\pi_{y_t}}\right)+T^{3/2}\sum_{k=3}^{y_t\wedge y_{t-1}}\mathrm{Bi}\left(y_{t-1},k;\beta_T\left(h_\beta\right)\right)\frac{\pi_{T,y_t-k}\left(h_\pi\right)}{\pi_{y_t}}.
\end{aligned}$$

Applying the mean value expansion

$$\log\left(1+x\right)=x-\frac{x^2}{2}+\frac{x^3}{3}\frac{1}{\left(1+x^*\right)^3}$$

and collecting powers of T gives

$$\log\sum_{k=0}^{y_t\wedge y_{t-1}}\mathrm{Bi}\left(y_{t-1},k;\beta_T\left(h_\beta\right)\right)\frac{\pi_{T,y_t-k}\left(h_\pi\right)}{\pi_{y_t}}=T^{-1/2}a_{1,t}-\frac{1}{2}T^{-1}\left(a_{1,t}^2-2a_{2,t}\right)+r_{T,t}$$

where

$$\begin{aligned}
r_{T,t} &= T^{-3/2}\left(a_{3,t}-a_{1,t}a_{2,t}\right)-T^{-2}\left(\frac{1}{2}a_{2,t}^2+a_{1,t}a_{3,t}\right)-T^{-5/2}a_{2,t}a_{3,t}-\frac{1}{2}T^{-3}a_{3,t}^2\\
&\quad +\frac{1}{3}\left(T^{-1/2}a_{1,t}+T^{-1}a_{2,t}+T^{-3/2}a_{3,t}\right)^3\frac{1}{\left(1+a_{T,t}^*\right)^3}.
\end{aligned}$$

Thus the log-likelihood ratio has the representation

$$\log\frac{L_T\left(\theta_T h\right)}{L_T\left(\theta_0\right)}=T^{-1/2}\sum_{t=1}^T a_{1,t}-\frac{1}{2}T^{-1}\sum_{t=1}^T\left(a_{1,t}^2-2a_{2,t}\right)+\sum_{t=1}^T r_{T,t}. \tag{A.35}$$

(i) The first term of (A.35) is the linear operator

$$S_T h=T^{-1/2}\sum_{t=1}^T a_{1,t}=S_{T,\beta}h_\beta+S_{T,\pi}h_\pi$$

as defined in (A.4), (A.5). Under H_0, so that y_t is i.i.d., the process $a_{1,t}$ is a stationary, ergodic (1-dependent) finite-variance martingale difference for any h, and therefore obeys a central limit theorem

$$S_T h\rightsquigarrow N\left(0,\langle h,Vh\rangle\right)$$

where

$$\langle h,Vh\rangle=E\left[\left(y_{t-1}\left(\frac{\pi_{y_t-1}}{\pi_{y_t}}-1\right)h_\beta+\left(h_{\pi,y_t}-\sum_{j=1}^\infty\pi_j h_{\pi,j}\right)\right)^2\right].$$

This central limit theorem clearly also holds jointly for finite collections of h's, while asymptotic equicontinuity is provided by the empirical process formulation of Lemma 1(c) of Drost *et al.* (2009). Thus S_T converges to a tight Gaussian process with covariance operator V.

(ii) For any h, each term of $a_{2,t}$ is stationary and ergodic and therefore satisfies a WLLN. Moreover $a_{2,t}$ has mean zero under Assumption 2, as shown by

$$\begin{aligned} & E\left[\binom{y_{t-1}}{2}\left(1-2\frac{\pi_{y_t-1}}{\pi_{y_t}}+\frac{\pi_{y_t-2}}{\pi_{y_t}}\right)\right] \\ = \; & E\left[\binom{y_{t-1}}{2}\right]\left(1-2\sum_{k\in\mathcal{U}}\pi_{k-1}+\sum_{k\in\mathcal{U}}\pi_{k-2}\right) \\ = \; & 0, \end{aligned}$$

since $\sum_{k\in\mathcal{U}}\pi_{k-1}=\sum_{k\in\mathcal{U}}\pi_{k-2}=1$ under Assumption 2, and

$$\begin{aligned} & E\left[y_{t-1}\left(\left(\frac{\pi_{y_t-1}}{\pi_{y_t}}-1\right)\left(h_{\pi,y_t-1}-\sum_{j=1}^{\infty}\pi_j h_{\pi,j}\right)-\left(h_{\pi,y_t-1}-h_{\pi,y_t}\right)\right)\right] \\ = \; & E\left[y_{t-1}\right]\left(\sum_{k\in\mathcal{U}}\left(\pi_{k-1}-\pi_k\right)h_{\pi,k-1}-\sum_{k\in\mathcal{U}}\left(h_{\pi,k-1}-h_{\pi,k}\right)\pi_k\right) \\ = \; & E\left[y_{t-1}\right]\left(\sum_{k\in\mathcal{U}}\pi_k h_{\pi,k}-\sum_{k\in\mathcal{U}}\pi_{k-1}h_{\pi,k-1}\right) \\ = \; & 0. \end{aligned}$$

(These terms are not zero under Assumption 3.) Thus $T^{-1}\sum_{t=1}^{T}a_{2,t}\xrightarrow{p}0$.

Similarly $a_{1,t}^2$ satisfies a WLLN with limit expressed as

$$E\left(a_{1,t}^2\right)=\langle h,Vh\rangle=h_\beta^2V_{\beta\beta}+2h_\beta V_{\beta\pi}h_\pi+\langle h_\pi,V_{\pi\pi}h_\pi\rangle,$$

in which

$$\begin{aligned} V_{\beta\beta} &= E\left[\left(y_{t-1}\left(\frac{\pi_{y_t-1}}{\pi_{y_t}}-1\right)\right)^2\right] \\ &= E\left[y_{t-1}^2\right]\sum_{k\in\mathcal{U}}\left(\frac{\pi_{k-1}^2}{\pi_k^2}-2\frac{\pi_{k-1}}{\pi_k}+1\right)\pi_k \\ &= E\left[y_{t-1}^2\right]\left(\sum_{k\in\mathcal{U}_1}\frac{\pi_{k-1}^2}{\pi_k^2}-1\right), \end{aligned}$$

since $\sum_{k\in\mathcal{U}}\pi_{k-1}=1$ under Assumption 2,

$$V_{\beta\pi}h_\pi=E\left[y_{t-1}\left(\frac{\pi_{y_t-1}}{\pi_{y_t}}-1\right)\left(h_{\pi,y_t}-\sum_{j=1}^{\infty}\pi_j h_{\pi,j}\right)\right]=E\left[y_{t-1}\right]\sum_{k=1}^{\infty}\left(\pi_{k-1}-\pi_k\right)h_{\pi,k},$$

again using Assumption 2, and

$$\langle h_\pi, V_{\pi\pi} h_\pi \rangle = E\left[\left(h_{\pi,y_t} - \sum_{j=1}^{\infty} \pi_j h_{\pi,j}\right)^2\right] = \sum_{j=1}^{\infty} \pi_j h_{\pi,j}^2 - \left(\sum_{j=1}^{\infty} \pi_j h_{\pi,j}\right)^2.$$

(iii) The remainder $r_{T,t}$ contains terms of order $T^{-3/2}$ and below. Two of these terms will be considered, with the others following similarly. The first is

$$\begin{aligned}
&\left| T^{-3/2} \sum_{t=1}^{T} \binom{y_{t-1}}{3} (1-\beta_T^*)^{y_{t-1}-3} \frac{\pi_{T,y_t}(h_\pi)}{\pi_{y_t}} \right| \\
\leq\ & \sup_j \left| \frac{\pi_{T,j}(h_\pi)}{\pi_j} \right| T^{-3/2} \sum_{t=1}^{T} \binom{y_{t-1}}{3} \\
\leq\ & \left(1 + 2T^{-1/2} \sup_j |h_{\pi,j}|\right) T^{-3/2} \sum_{t=1}^{T} \binom{y_{t-1}}{3} \\
\xrightarrow{p}\ & 0,
\end{aligned}$$

since $h_\pi \in \ell^\infty$ and u_t is assumed to have finite third moment. The second uses (A.32)–(A.34) to obtain

$$\sum_{k=3}^{y_{t-1}} \mathrm{Bi}(y_{t-1}, k; \beta_T(h_\beta)) = 1 - \sum_{k=0}^{2} \mathrm{Bi}(y_{t-1}, k; \beta_T(h_\beta)) \leq 7 \binom{y_{t-1}}{3} \beta_T(h_\beta)^3$$

and hence

$$\begin{aligned}
E\left| \sum_{k=3}^{y_t \wedge y_{t-1}} \mathrm{Bi}(y_{t-1}, k; \beta_T(h_\beta)) \frac{\pi_{T,y_t-k}(h_\pi)}{\pi_{y_t}} \right| &\leq E\left(\sum_{k=3}^{y_{t-1}} \mathrm{Bi}(y_{t-1}, k; \beta_T(h_\beta))\right) E\left(\frac{\pi_{T,y_t-k}(h_\pi)}{\pi_{y_t}}\right) \\
&\leq 7T^{-3/2} h_\beta \mu_3,
\end{aligned}$$

so that

$$\sum_{t=1}^{T} \sum_{k=3}^{y_t \wedge y_{t-1}} \mathrm{Bi}(y_{t-1}, k; \beta_T(h_\beta)) \frac{\pi_{T,y_t-k}(h_\pi)}{\pi_{y_t}} \xrightarrow{p} 0.$$

This completes the proof. ■

B.2 Proof of Lemma A.1

Following CHS, define the effective score

$$S_{T,\beta}^* = S_{T,\beta} - S_{T,\pi} V_{\pi\pi}^{-1} V_{\pi\beta}.$$

From the expressions given in Theorem A.1, it can be seen that $V_{\pi\pi}^{-1}$ is defined by

$$\left\langle h_\pi, V_{\pi\pi}^{-1} h_\pi \right\rangle = \sum_{k=1}^{\infty} \frac{1}{\pi_k} h_{\pi,k}^2 + \frac{1}{\pi_0} \left(\sum_{k=1}^{\infty} h_{\pi,k}\right)^2.$$

It immediately follows that $S^*_{T,\beta}$ is given by (A.9). The weak convergence result for S_T reported in Theorem A.1(i) implies that

$$S^*_{T,\beta} \rightsquigarrow Z^* \sim N\left(0, \omega^2\right), \tag{A.36}$$

where the variance is

$$\omega^2 = V_{\beta\beta} - V_{\beta\pi} V_{\pi\pi}^{-1} V_{\pi\beta} = \sigma_u^2 \left(\sum_{k=1}^{\infty} \frac{\pi_{k-1}^2}{\pi_k} - 1 \right), \tag{A.37}$$

which verifies the asymptotic null distribution of ξ_T.

Theorem 1 of CHS states that an asymptotically uniformly most powerful test at a given Π is provided by the (infeasible) test that rejects H_0 for $\xi_T > z_\alpha$, where z_α is the upper α quantile of the standard normal distribution. The asymptotic local power of this test can be computed from the asymptotic distribution $\xi_T \rightsquigarrow N(\omega h_\beta, 1)$ under θ_T, which follows from Le Cam's third lemma and the LAN property in Theorem A.1. In particular, under θ_T,

$$\Pr(\xi_T > z_\alpha) \to 1 - \Phi(z_\alpha - \omega h_\beta)$$

which gives the local power. ■

B.3 Proof of Lemma A.2

Denote the INAR(1) transition probabilities under $\beta_T = T^{-1} h_\beta$ and any $\Pi = \{\pi_j\}_{j \in \mathcal{U}^{(0)}}$ as

$$p_{i|j} = \Pr(y_t = i | y_{t-1} = j) = \sum_{k=0}^{i \wedge j} \text{Bi}(j, k; \beta_T) \pi_{i-k}. \tag{A.38}$$

The identities

$$(1-\beta_T)^j = 1 - \beta_T \sum_{k=1}^{j} (1-\beta_T)^{k-1} \tag{A.39}$$

$$= 1 - \beta_T j + \beta_T^2 \sum_{k=1}^{j-1} (j-k)(1-\beta_T)^{k-1} \tag{A.40}$$

are applied to $\text{Bi}(j, 1; \beta_T)$ and $\text{Bi}(j, 0; \beta_T)$ respectively to give an expression for the first two terms in (A.38) as

$$\text{Bi}(j, 0; \beta_T) \pi_i + \text{Bi}(j, 1; \beta_T) \pi_{i-1} = q_{i|j} - \beta_T^2 \left(s_{0,j} \pi_i + s_{1,j} (\pi_{i-1} - \pi_i) \right)$$

where

$$s_{0,j} = \sum_{k=1}^{j-1} k (1-\beta_T)^{k-1}, \; s_{1,j} = j \sum_{k=1}^{j-1} (1-\beta_T)^{k-1},$$

and where it is convenient for subsequent analysis to represent $q_{i|j}$ as

$$q_{i|j} = \begin{pmatrix} \pi_i \\ \pi_{i-1} - \pi_i \end{pmatrix}' \begin{pmatrix} 1 \\ \beta_T j \end{pmatrix}. \tag{A.41}$$

Thus

$$p_{i|j} = q_{i|j} + r_{i|j},$$

where

$$r_{i|j} = -\beta_T^2 \left(s_{0,j}\pi_i + s_{1,j}\left(\pi_{i-1} - \pi_i\right)\right) + \sum_{k=2}^{i\wedge j} \mathrm{Bi}\left(j,k;\beta_T\right)\pi_{i-k}. \tag{A.42}$$

Then

$$\begin{aligned}
& \Pr\left(y_{t_1} = i_1, \ldots, y_{t_k} = i_k \text{ and } y_t \notin \mathcal{A} \text{ for all } s \neq t_1, \ldots, t_k\right) \\
= & \sum_{j_T \notin \mathcal{A}} \cdots \sum_{j_{t_k+1} \notin \mathcal{A}} \sum_{j_{t_k-1} \notin \mathcal{A}} \cdots \sum_{j_{t_1+1} \notin \mathcal{A}} \sum_{j_{t_1-1} \notin \mathcal{A}} \cdots \sum_{j_1 \notin \mathcal{A}} \sum_{j_0 \in \mathcal{U}^{(0)}} p_{j_T|j_{T-1}} p_{j_{T-1}|j_{T-2}} \cdots \\
& \quad p_{j_{t_k+1}|i_k} q_{i_k|j_{t_k-1}} \cdots p_{j_{t_1+1}|i_1} p_{i_1|j_{t_1-1}} \cdots p_{j_1|j_0} \pi_{j_0} \\
= & \sum_{j_T \notin \mathcal{A}} \cdots \sum_{j_{t_k+1} \notin \mathcal{A}} \sum_{j_{t_k-1} \notin \mathcal{A}} \cdots \sum_{j_{t_1+1} \notin \mathcal{A}} \sum_{j_{t_1-1} \notin \mathcal{A}} \cdots \sum_{j_1 \notin \mathcal{A}} \sum_{j_0 \in \mathcal{U}^{(0)}} q_{j_T|j_{T-1}} q_{j_{T-1}|j_{T-2}} \cdots \qquad \text{(A.43)} \\
& \quad q_{j_{t_k+1}|i_k} q_{i_k|j_{t_k-1}} \cdots q_{j_{t_1+1}|i_1} q_{i_1|j_{t_1-1}} \cdots q_{j_1|j_0} \pi_{j_0} \\
& + \sum_{j_T \notin \mathcal{A}} \cdots \sum_{j_{t_k+1} \notin \mathcal{A}} \sum_{j_{t_k-1} \notin \mathcal{A}} \cdots \sum_{j_{t_1+1} \notin \mathcal{A}} \sum_{j_{t_1-1} \notin \mathcal{A}} \cdots \sum_{j_1 \notin \mathcal{A}} \sum_{j_0 \in \mathcal{U}^{(0)}} r_{j_T|j_{T-1}} q_{j_{T-1}|j_{T-2}} \cdots \qquad \text{(A.44)} \\
& \quad q_{j_{t_k+1}|i_k} q_{i_k|j_{t_k-1}} \cdots q_{j_{t_1+1}|i_1} q_{i_1|j_{t_1-1}} \cdots q_{j_1|j_0} \pi_{j_0} \\
& + \sum_{j_T \notin \mathcal{A}} \cdots \sum_{j_{t_k+1} \notin \mathcal{A}} \sum_{j_{t_k-1} \notin \mathcal{A}} \cdots \sum_{j_{t_1+1} \notin \mathcal{A}} \sum_{j_{t_1-1} \notin \mathcal{A}} \cdots \sum_{j_1 \notin \mathcal{A}} \sum_{j_0 \in \mathcal{U}^{(0)}} p_{j_T|j_{T-1}} r_{j_{T-1}|j_{T-2}} \cdots \\
& \quad q_{j_{t_k+1}|i_k} q_{i_k|j_{t_k-1}} \cdots q_{j_{t_1+1}|i_1} q_{i_1|j_{t_1-1}} \cdots q_{j_1|j_0} \pi_{j_0} \\
& \vdots \\
& \sum_{j_T \notin \mathcal{A}} \cdots \sum_{j_{t_k+1} \notin \mathcal{A}} \sum_{j_{t_k-1} \notin \mathcal{A}} \cdots \sum_{j_{t_1+1} \notin \mathcal{A}} \sum_{j_{t_1-1} \notin \mathcal{A}} \cdots \sum_{j_1 \notin \mathcal{A}} \sum_{j_0 \in \mathcal{U}^{(0)}} p_{j_T|j_{T-1}} p_{j_{T-1}|j_{T-2}} \cdots \qquad \text{(A.45)} \\
& \quad p_{j_{t_k+1}|i_k} q_{i_k|j_{t_k-1}} \cdots p_{j_{t_1+1}|i_1} p_{i_1|j_{t_1-1}} \cdots r_{j_1|j_0} \pi_{j_0}
\end{aligned}$$

The main result is found from (A.43), with the sum of (A.44)–(A.45) being shown to be negligible below. Defining

$$\eta_0 = \sum_{j \notin \mathcal{A}} \left(\pi_{j-1} - \pi_j\right) = \sum_{j \in \mathcal{U}^{(0)} \cup \mathcal{U}^{(1)}} \pi_{j-1} - \sum_{j \in \mathcal{A}} \pi_{j-1} - \sum_{j \in \mathcal{U}^{(0)}} \pi_j = -\sum_{j \in \mathcal{A}} \pi_{j-1}$$

and

$$\begin{aligned}
\sum_{j \notin \mathcal{A}} j\left(\pi_{j-1} - \pi_j\right) &= \sum_{j \in \mathcal{U}^{(0)} \cup \mathcal{U}^{(1)}} j\pi_{j-1} - \sum_{j \in \mathcal{A}} j\pi_{j-1} - \sum_{j \in \mathcal{U}^{(0)}} j\pi_j \\
&= \sum_{j \in \mathcal{U}^{(0)}} (j+1)\,\pi_j - \sum_{j \in \mathcal{A}} j\pi_{j-1} - \sum_{j \in \mathcal{U}^{(0)}} j\pi_j \\
&= 1 - \sum_{j \in \mathcal{A}} j\pi_{j-1}
\end{aligned}$$

gives

$$\sum_{j\notin\mathcal{A}}\begin{pmatrix}1\\ \beta_T j\end{pmatrix}\begin{pmatrix}\pi_j\\ \pi_{j-1}-\pi_j\end{pmatrix}'=\begin{pmatrix}1 & -\eta_0\\ \beta_T\mu_u & -\beta_T\eta_1\end{pmatrix}.$$

This matrix can be diagonalised as $V\Lambda V^{-1}$ in which $\Lambda = \text{diag}(\lambda_2,\lambda_1)$, $\lambda_2>\lambda_1$

$$(\lambda_2,\lambda_1)=\frac{1}{2}\left(1-\beta_T\eta_1\pm\sqrt{\beta_T^2\eta_1-2\beta_T\left(2\eta_0\mu_u-\eta_1\right)+1}\right)$$

and

$$V=\begin{pmatrix}\lambda_2+\beta_T\eta_1 & \lambda_1+\beta_T\eta_1\\ \beta_T\mu_u & \beta_T\mu_u\end{pmatrix}.$$

Thus, using $\pi_i=0$ for all $i\in\mathcal{A}$,

$$\begin{aligned}&\Pr\left(y_{t_1}=i_1,\ldots,y_{t_k}=i_k \text{ and } y_t\notin\mathcal{A} \text{ for all } s\neq t_1,\ldots,t_k\right)\\ \approx&\begin{pmatrix}1\\ -\eta_0\end{pmatrix}'V\Lambda^{T-t_k-1}V^{-1}\begin{pmatrix}1\\ \beta_T i_k\end{pmatrix}\begin{pmatrix}0\\ \pi_{i_k-1}\end{pmatrix}'V\Lambda^{t_k-t_{k-1}-1}V^{-1}\\ &\times\begin{pmatrix}1\\ \beta_T i_{k-1}\end{pmatrix}\begin{pmatrix}0\\ \pi_{i_{k-1}-1}\end{pmatrix}'V\ldots V\Lambda^{t_1}V^{-1}\begin{pmatrix}1\\ 0\end{pmatrix}.\end{aligned}\tag{A.46}$$

Taking second order expansions of λ_2 and λ_1 around $\beta_T=0$ gives

$$\begin{aligned}\lambda_2&=1-\beta_T\mu_u\eta_0+\beta_T^2\frac{\left(\mu_u\eta_0-\eta_1\right)\mu_u\eta_0}{\left(\beta_T^{*2}\eta_1-2\beta_T^*\left(2\mu_u\eta_0-\eta_1\right)+1\right)^{3/2}}\\ \lambda_1&=\beta_T\left(\mu_u\eta_0-\eta_1\right)-\beta_T^2\frac{\left(\mu_u\eta_0-\eta_1\right)\mu_u\eta_0}{\left(\beta_T^{**2}\eta_1-2\beta_T^{**}\left(2\mu_u\eta_0-\eta_1\right)+1\right)^{3/2}}\end{aligned}$$

for $0\le\beta_T^*,\beta_T^{**}\le\beta_T$. These show that

$$\begin{aligned}\lambda_2&\to 1,\lambda_1\to 0,\\ \lambda_2^T&\to\exp\left(-h_\beta\mu_u\eta_0\right).\end{aligned}$$

It follows that

$$\begin{pmatrix}1\\ -\eta_0\end{pmatrix}'V\approx\begin{pmatrix}1\\ -\eta_0\end{pmatrix}'\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}=\begin{pmatrix}1\\ 0\end{pmatrix}',$$

and

$$V^{-1}\begin{pmatrix}1\\ 0\end{pmatrix}=\frac{1}{\lambda_2-\lambda_1}\begin{pmatrix}1\\ -1\end{pmatrix}\approx\begin{pmatrix}1\\ -1\end{pmatrix},$$

and

$$\begin{aligned}V^{-1}\begin{pmatrix}1\\ \beta_T i\end{pmatrix}\begin{pmatrix}0\\ \pi_{i-1}\end{pmatrix}'V&=\frac{\beta_T\mu_u\pi_{i-1}}{\left(\lambda_2-\lambda_1\right)}\begin{pmatrix}1-i\left(\lambda_1+\beta_T\eta_1\right)/\mu_u\\ -1+i\left(\lambda_2+\beta_T\eta_1\right)/\mu_u\end{pmatrix}\begin{pmatrix}1\\ 1\end{pmatrix}'\\ &\approx\beta_T\mu_u\pi_{i-1}\begin{pmatrix}1\\ -1+i/\mu_u\end{pmatrix}\begin{pmatrix}1\\ 1\end{pmatrix}'.\end{aligned}$$

Substituting these into (A.46) gives

$$
\begin{aligned}
& \Pr\left(y_{t_1}=i_1,\ldots,y_{t_k}=i_k \text{ and } y_t \notin \mathcal{A} \text{ for all } s \neq t_1,\ldots,t_k\right) \\
\approx\ & \left(\prod_{i\in\mathcal{A}} \beta_T \mu_u \pi_{i-1}\right) \begin{pmatrix} 1 \\ 0 \end{pmatrix}' \begin{pmatrix} \lambda_2^{T-t_k-1} & 0 \\ 0 & \lambda_1^{T-t_k-1} \end{pmatrix} \begin{pmatrix} 1 \\ -1+i_k/\mu_u \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix}' \\
& \begin{pmatrix} \lambda_2^{t_k-t_{k-1}-1} & 0 \\ 0 & \lambda_1^{t_k-t_{k-1}-1} \end{pmatrix} \begin{pmatrix} 1 \\ -1+i_{k-1}/\mu_u \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix}' \cdots \begin{pmatrix} \lambda_2^{t_1} & 0 \\ 0 & \lambda_1^{t_1} \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \\
=\ & \left(\prod_{i\in\mathcal{A}} \beta_T \mu_u \pi_{i-1}\right) \left(\lambda_2^{T-t_k-1}\left(\lambda_2^{t_k-t_{k-1}-1} - \lambda_1^{t_k-t_{k-1}-1}\left(-1+i_{k-1}/\mu_u\right)\right)\ldots\left(\lambda_2^{t_1}-\lambda_1^{t_1}\right)\right) \\
\approx\ & \left(\prod_{i\in\mathcal{A}} \beta_T \mu_u \pi_{i-1}\right) \lambda_2^{T-k} \\
\approx\ & \left(\prod_{i\in\mathcal{A}} \beta_T \mu_u \pi_{i-1}\right) \exp\left(-h_\beta \mu_u \eta_0\right), \qquad \text{(A.47)}
\end{aligned}
$$

as required.

Now consider the sum of (A.44)–(A.45). Write $r_{i|j} = r_{i|j}^{(1)} + r_{i|j}^{(2)}$, where $r_{i|j}^{(2)} = \sum_{k=2}^{i\wedge j} \mathrm{Bi}\left(j,k;\beta_T\right)\pi_{i-k}$ is the less obvious term that is explicitly derived here. Observe that

$$
\begin{aligned}
0 \ <\ & \sum_{i\notin\mathcal{A}}\sum_{j\notin\mathcal{A}} r_{i|j}^{(2)}\pi_j = \sum_{j\in\mathcal{U}^{(0)}} \pi_j \sum_{k=2}^{j} \mathrm{Bi}\left(j,k;\beta_T\right) \sum_{i\notin\mathcal{A}, i\geq k} \pi_{i-k} \\
\leq\ & \sum_{j\in\mathcal{U}^{(0)}} \pi_j \sum_{k=2}^{j} \mathrm{Bi}\left(j,k;\beta_T\right) \\
=\ & \sum_{j\in\mathcal{U}^{(0)}} \pi_j \left(1-\left(1-\beta_T\right)^j - \beta_T j \left(1-\beta_T\right)^{j-1}\right) \\
=\ & \sum_{j\in\mathcal{U}^{(0)}} \pi_j \left(1-\left(1-\beta_T j + \beta_T^2 \sum_{h=1}^{j-1}(j-h)\left(1-\beta_T\right)^{h-1}\right) - \beta_T j \left(1-\beta_T\sum_{k=1}^{j}\left(1-\beta_T\right)^{k-1}\right)\right) \\
=\ & \beta_T^2 \sum_{j\in\mathcal{U}^{(0)}} \pi_j \left(\sum_{h=1}^{j-1} h\left(1-\beta_T\right)^{h-1}\right) \\
\leq\ & \beta_T^2 \sum_{j\in\mathcal{U}^{(0)}} \pi_j j^2 \\
=\ & \beta_T^2 E\left[u_t^2\right],
\end{aligned}
$$

and similarly

$$
0 < \sum_{i\in\mathcal{U}^{(0)}}\sum_{j\in\mathcal{U}^{(0)}} r_{i|j}^{(2)} \pi_{j-1} \leq \beta_T^2\left(E\left[u_t^2\right] + E\left[u_t\right]\right).
$$

For any $n = 1,\ldots,T$, define h such that $t_h \leq n$ and $t_{h+1} > n$. Then the n^{th} term among

(A.44)–(A.45) satisfies

$$\begin{aligned}
&\sum_{j_n \notin \mathcal{A}} \cdots \sum_{j_{t_k+1} \notin \mathcal{A}} \sum_{j_{t_k-1} \notin \mathcal{A}} \cdots \sum_{j_n \notin \mathcal{A}} \cdots \sum_{j_{t_h+1} \notin \mathcal{A}} \sum_{j_{t_h-1} \notin \mathcal{A}} \cdots \sum_{j_{t_1+1} \notin \mathcal{A}} \sum_{j_{t_1-1} \notin \mathcal{A}} \cdots \sum_{j_1 \notin \mathcal{A}} \sum_{j_0 \in \mathcal{U}^{(0)}} p_{j_T|j_{T-1}} \cdots \\
&\quad p_{j_{t_k+1}|i_k} p_{i_k|j_{t_k-1}} \cdots p_{j_{n+1}|j_n} r_{j_n|j_{n-1}} q_{j_{n-1}|j_{n-2}} \cdots q_{j_{t_h+1}|i_h} q_{i_h|j_{t_h-1}} \cdots q_{j_{t_1+1}|i_1} q_{i_1|j_{t_1-1}} \cdots q_{j_1|j_0} \pi_{j_0} \\
\leq\ & \sum_{j_n \notin \mathcal{A}} \sum_{j_{n-1} \notin \mathcal{A}} \cdots \sum_{j_{t_h+1} \notin \mathcal{A}} \sum_{j_{t_h-1} \notin \mathcal{A}} \cdots \sum_{j_{t_1+1} \notin \mathcal{A}} \sum_{j_{t_1-1} \notin \mathcal{A}} \cdots \sum_{j_1 \notin \mathcal{A}} \sum_{j_0 \in \mathcal{U}^{(0)}} r^{(2)}_{j_n|j_{n-1}} q_{j_{n-1}|j_{n-2}} \cdots q_{j_{t_h+1}|i_h} q_{i_h|j_{t_h-1}} \cdots \\
&\quad q_{j_{t_1+1}|i_1} q_{i_1|j_{t_1-1}} \cdots q_{j_1|j_0} \pi_{j_0} \\
\approx\ & \left(\prod_{j=1}^{h} \beta_T \mu_u \pi_{i_j-1} \right) \sum_{j_n \notin \mathcal{A}} \sum_{j_{n-1} \notin \mathcal{A}} r^{(2)}_{j_n|j_{n-1}} \begin{pmatrix} \pi_{j_{n-1}} \\ \pi_{j_{n-1}-1} - \pi_{j_{n-1}} \end{pmatrix}' \begin{pmatrix} \exp\left(-h_\beta \mu_u \eta_0\right) \\ 0 \end{pmatrix} \\
=\ & \left(\prod_{j=1}^{h} \beta_T \mu_u \pi_{i_j-1} \right) \exp\left(-h_\beta \mu_u \sum_{j=1}^{h} \pi_{i_j-1} \right) \sum_{j_n \notin \mathcal{A}} \sum_{j_{n-1} \notin \mathcal{A}} r^{(2)}_{j_n|j_{n-1}} \pi_{j_{n-1}} \\
\leq\ & \beta_T^2 \left(E\left[u_t^2\right] + E\left[u_t\right] \right) \left(\prod_{j=1}^{h} \beta_T \mu_u \pi_{i_j-1} \right) \exp\left(-h_\beta \mu_u \sum_{j=1}^{h} \pi_{i_j-1} \right)
\end{aligned}$$

showing that, compared to (A.47), each term in (A.44)–(A.45) is bounded by an expression with an extra $\beta_T^2 = O\left(T^{-2}\right)$, and hence that the sum of T of these terms is asymptotically negligible relative to (A.47). The terms involving $r^{(1)}_{i|j} = -\beta_T^2 \left(s_{0,j}\pi_i + s_{1,j}\left(\pi_{i-1} - \pi_i\right)\right)$ will clearly also involve the extra β_T^2 factor and hence are negligible. ■