Online Supplement

Appendix A. Algorithm used to compute the probability distribution of $\boldsymbol{D}(\mathrm{s}, \mathrm{a})$.

Figure A.1: Pseudo-code of the algorithm used to compute the probability distribution of $D(\mathbf{s}, \mathbf{a}) = D_{1K(\mathbf{s},\mathbf{a})}$.

```
for $b = 1$ to $K - 1$ do
    for $k = \sum_{t=1}^{(b+1)} \underline{d}_b$ to $\sum_{t=1}^{(b+1)} \bar{d}_b$ do
        for $m = \underline{d}_{b+1}$ to $\bar{d}_{b+1}$ do
            $\Pr\{D_{1(b+1)} = k\} = \Pr\{D_{1(b+1)} = k\} + \Pr\{D_{1b} = k - m\} \times \Pr\{d_{b+1} = m\}$
        end for
    end for
end for
```

Appendix B. Simulation results for initial states generated using the M policy.

Table B.1: Summary of the simulation results for Problem Setting 1 (Base Case). The bold font indicates the policy (policies) that provides (provide) the best mean performance for each metric in a statistical sense ($\alpha = 0.05$).

Metric	**Patient Class**	**Policy**			
		AOPD	**AOPS**	**FAS**	**M**
DC	–	32352 ± 536	**32055 ± 559**	50616 ± 2506	43213 ± 1525
AC	–	313.17 ± 1.77	**310.83 ± 1.71**	578.74 ± 14.14	435.71 ± 7.92
ACU	–	18.03 ± 0.05	18.03 ± 0.05	17.98 ± 0.05	**18.01 ± 0.05**
ATFAS	1	**1.25 ± 0.01**	1.27 ± 0.01	2.26 ± 0.02	2.19 ± 0.02
	2	**1.29 ± 0.01**	1.32 ± 0.01	2.80 ± 0.03	2.65 ± 0.03
	3	**1.34 ± 0.01**	1.37 ± 0.01	3.57 ± 0.05	3.27 ± 0.04
AWT	1	**2.01 ± 0.02**	2.17 ± 0.02	8.56 ± 0.20	6.23 ± 0.11
	2	**5.50 ± 0.04**	**5.49 ± 0.04**	8.71 ± 0.18	6.96 ± 0.14
	3	9.47 ± 0.05	9.63 ± 0.05	8.56 ± 0.15	**7.39 ± 0.14**
SL	1	**99.98 ± 0.01**	99.94 ± 0.02	10.85 ± 1.81	23.89 ± 1.87
	2	**100.00 ± 0.00**	**100.00 ± 0.00**	38.60 ± 2.86	68.41 ± 1.87
	3	**100.00 ± 0.00**	**100.00 ± 0.00**	**100.00 ± 0.00**	**100.00 ± 0.00**

Table B.2: Percent deviation from the lowest mean discounted cost (DC*) for Problem Setting 1 (Base Case) assuming different discrete service time distributions. The bold font indicates the policy that provides the best value for each probability distribution.

Policy	**Probability Distributions**			
	Geometric	**Neg. Binomial**	**Poisson**	**Uniform**
AOPD	1.0%	0.9%	0.9%	1.5%
AOPS	**0.0%**	**0.0%**	**0.0%**	**0.0%**
FAS	48.4%	57.9%	63.3%	61.3%
M	30.0%	34.8%	37.3%	36.5%
DC*	41190 ± 704	32055 ± 559	27994 ± 510	29088 ± 531

Table B.3: Summary of the simulation results for Problem Setting 2. The bold font indicates the policy (policies) that provides (provide) the best mean performance for each metric in a statistical sense ($\alpha = 0.05$).

Metric	**Patient Class**	**Policy**			
		AOPD	**AOPS**	**FAS**	**M**
DC	–	33907 ± 634	**33636 ± 651**	37577 ± 1410	34579 ± 984
AC	–	331.25 ± 1.81	**327.40 ± 1.79**	385.34 ± 5.44	341.16 ± 3.24
ACU	–	**13.99 ± 0.04**	**13.99 ± 0.04**	13.97 ± 0.04	13.98 ± 0.04
ATFAS	1	**1.34 ± 0.01**	1.43 ± 0.01	1.98 ± 0.02	1.90 ± 0.02
	2	**1.56 ± 0.01**	1.70 ± 0.02	3.37 ± 0.06	3.05 ± 0.04
AWT	1	**2.22 ± 0.02**	2.44 ± 0.02	5.73 ± 0.12	4.43 ± 0.07
	2	**5.76 ± 0.03**	6.26 ± 0.03	7.51 ± 0.12	6.64 ± 0.09
	3	8.79 ± 0.05	8.88 ± 0.05	4.47 ± 0.12	**3.80 ± 0.08**
SL	1	**99.76 ± 0.05**	99.40 ± 0.08	41.57 ± 1.55	53.68 ± 1.22
	2	99.95 ± 0.01	**100.00 ± 0.00**	56.16 ± 1.77	71.30 ± 1.28
	3	**100.00 ± 0.00**	**100.00 ± 0.00**	**100.00 ± 0.00**	**100.00 ± 0.00**

Table B.4: Summary of the simulation results for Problem Setting 3. The bold font indicates the policy (policies) that provides (provide) the best mean performance for each metric in a statistical sense ($\alpha = 0.05$).

Metric	**Patient Class**	**Policy**			
		AOPD	**AOPS**	**FAS**	**M**
DC	–	31293 ± 696	**30364 ± 707**	34006 ± 1844	31131 ± 1261
AC	–	302.11 ± 1.71	**294.29 ± 1.71**	357.80 ± 8.21	308.73 ± 4.39
ACU	–	**15.99 ± 0.04**	**15.99 ± 0.04**	15.98 ± 0.04	**15.99 ± 0.04**
ATFAS	1 & 2	**1.49 ± 0.01**	**1.49 ± 0.01**	3.64 ± 0.08	3.32 ± 0.06
AWT	1	**2.29 ± 0.02**	2.50 ± 0.02	6.29 ± 0.21	4.89 ± 0.12
	2	5.83 ± 0.03	5.67 ± 0.04	6.53 ± 0.20	**5.49 ± 0.13**
	3	9.94 ± 0.05	8.80 ± 0.06	6.61 ± 0.19	**5.73 ± 0.14**
SL	1	**99.89 ± 0.04**	99.82 ± 0.05	34.83 ± 2.66	45.76 ± 2.19
	2	99.99 ± 0.00	**100.00 ± 0.00**	67.10 ± 2.42	83.21 ± 1.34
	3	**100.00 ± 0.00**	**100.00 ± 0.00**	**100.00 ± 0.00**	**100.00 ± 0.00**

Table B.5: Percent difference between the mean discounted cost (DC) associated with the AOPD policy and that associated with the AOPS policy for Problem Setting 3 assuming different idle time and overtime cost values.

Overtime Cost/Idle Time Cost				
\$100/\$0	**\$100/\$25**	**\$100/\$50**	**\$100/\$75**	**\$100/\$100**
1.36%	2.25%	3.06%	2.84%	3.13%

Table B.6: Summary of the simulation results for Problem Setting 4. The bold font indicates the policy (policies) that provides (provide) the best mean performance for each metric in a statistical sense ($\alpha = 0.05$).

Metric	**Patient Class**	**Policy**			
		AOPD	**AOPS**	**FAS**	**M**
DC	–	59376 ± 1495	**57284 ± 1426**	62489 ± 4128	60674 ± 3465
AC	–	534.09 ± 8.59	**509.91 ± 9.19**	787.80 ± 91.60	604.83 ± 27.79
ACU	–	141.57 ± 0.30	141.51 ± 0.30	141.90 ± 0.10	**142.72 ± 0.32**
ATFAS	1	**1.05 ± 0.01**	**1.05 ± 0.01**	2.15 ± 0.07	2.12 ± 0.07
	2	**1.06 ± 0.01**	**1.06 ± 0.01**	2.65 ± 0.11	2.59 ± 0.10
	3	**1.07 ± 0.01**	1.07 ± 0.02	3.34 ± 0.19	3.15 ± 0.15
AWT	1	**1.18 ± 0.03**	1.23 ± 0.04	8.40 ± 1.02	5.76 ± 0.38
	2	**1.93 ± 0.13**	2.02 ± 0.13	8.46 ± 1.01	6.29 ± 0.47
	3	**4.04 ± 0.27**	**4.05 ± 0.27**	8.64 ± 1.02	6.46 ± 0.47
	4	17.09 ± 0.45	16.65 ± 0.49	9.01 ± 1.02	**6.82 ± 0.47**
SL	1	**100.00 ± 0.00**	**99.99 ± 0.01**	14.24 ± 8.71	24.61 ± 8.11
	2	**100.00 ± 0.01**	**100.00 ± 0.00**	49.44 ± 12.79	92.41 ± 2.34
	3	**100.00 ± 0.00**	**100.00 ± 0.00**	88.08 ± 7.07	**100.00 ± 0.00**
	4	**100.00 ± 0.00**	**100.00 ± 0.00**	**100.00 ± 0.00**	**100.00 ± 0.00**

Appendix C. A proof of the form of the optimal affine value function approximation in the case with deterministic service times

We present a proof of the form of the optimal affine value function approximation in the case with deterministic service times. The wait time targets, $T(i)$, for each priority class are assumed to increase with i as a high priority patient is, by definition, a patient who must be served sooner. For completeness, we restate the theorem before giving its proof.

Restating the Theorem

Assuming that $T(i)$ is non-decreasing in i, that the wait time penalties are non-decreasing in n and non-increasing in i, and that the following conditions are satisfied:

$$f^D(i) > \left(\gamma^{T(i)-1} - \gamma^{T(i)}\right)\mu_j h \quad \forall (i,j) \in [I] \times [J] \tag{C.1}$$

$$\sum_{j\in J} \mu_j \left[\sum_{i=1}^{I} \frac{\gamma^{T(i)-1} I(T(i) > 1)\lambda_{ij}}{1-\gamma} + \sum_{m=1}^{N} \gamma^{m-1} E_\alpha[X_{jm}] \right] > \frac{C^R}{1-\gamma} \tag{C.2}$$

$$\sum_{j\in J} \mu_j \left[\sum_{i=1}^{I} \frac{\gamma^{T(i)-n} I(T(i) > n)\lambda_{ij}}{1-\gamma} + \sum_{m=n}^{N} \gamma^{m-n} E_\alpha[X_{jm}] \right] < \frac{C^R + C^{OT}}{1-\gamma} \quad \forall n \in [N] \tag{C.3}$$

Then, the optimal affine value function approximation for the discounted MDP will be given by:

$$V^*_{jn} = \begin{cases} \mu_j h, & \text{n=1;} \\ \gamma V^*_{j(n-1)}, & 2 \le n \le N-1; \\ 0, & n = N. \end{cases} \tag{C.4}$$

$$W^*_{ij} = \begin{cases} V^*_{jT(i)}, & \lambda_{ij} > 0; \\ 0, & \lambda_{ij} = 0. \end{cases} \tag{C.5}$$

$$V^*_0 = \frac{1}{1-\gamma}\left(\sum_{\substack{(i,j)\in\\ [I]\times[J]}} \gamma^{T(i)} \mu_j h E[Y_{ij}] + hC^R \right), \tag{C.6}$$

where μ_j is the number of time units required by a patient of service class j, h is the overtime cost per time unit, γ is the daily discount factor, λ_{ij} is the expected demand from priority class i and service class j patients, and C^R is the system regular-hour capacity in time units.

The Proof

The outline of the proof is as follows:

1. Prove the primal feasibility of the proposed solution.
2. Determine necessary and sufficient conditions under which a dual solution, together with the proposed primal solution, would satisfy complementary slackness.
3. Demonstrate that there exists a dual solution satisfying the necessary and sufficient conditions.

The existence of a dual solution that together with the proposed primal solution satisfies complementary slackness is sufficient to prove the optimality.

Proving Primal feasibility

We begin by proving the feasibility of the hypothesized primal solution. Clearly, it gives non-negative values for $\vec{V}$ and $\vec{W}$. With a little algebraic manipulation, the constraint for the primal LP can be written as:

$$\begin{aligned}(1-\gamma)V_0 \leq f^{AS}(\mathbf{s},\mathbf{a}) &+ \sum_{\substack{(i,j,n)\in\\ [I]\times[J]\times[N]}} \Big(f^{WT}(i,n) + \gamma V_{j,n-1} - \gamma W_{ij} - f^D(i)\Big)a_{ijn} \\ &+ \sum_{\substack{(j,n)\in\\ [J]\times[N]}} \Big(\gamma V_{j,n-1} - V_{jn}\Big)x_{jn} + \sum_{\substack{(i,j)\in\\ [I]\times[J]}} \Big(\gamma W_{ij} + f^D(i) - W_{ij}\Big)y_{ij} + \gamma \sum_{\substack{(i,j)\in\\ [I]\times[J]}} W_{ij}E[Y_{ij}]\end{aligned}$$

For state-action pairs for which $\sum_{j\in[J]} \mu_j(x_{j1} + \sum_{i\in[I]} a_{ij1}) > C^R$, this equates to:

$$\begin{aligned}(1-\gamma)V_0 \leq + &\sum_{\substack{(i,j,n)\in\\ [I]\times[J]\times[N]}} \Big(f^{WT}(i,n) + h\mu_j I(n=1) + \gamma V_{j,n-1} - \gamma W_{ij} - f^D(i)\Big)a_{ijn} \\ &+ \sum_{\substack{(j,n)\in\\ [J]\times[N]}} \Big(h\mu_j I(n=1) + \gamma V_{j,n-1} - V_{jn}\Big)x_{jn} + \sum_{\substack{(i,j)\in\\ [I]\times[J]}} \Big(\gamma W_{ij} + f^D(i) - W_{ij}\Big)y_{ij} \\ &+ \gamma \sum_{\substack{(i,j)\in\\ [I]\times[J]}} W_{ij}E[Y_{ij}] - hC^R\end{aligned}$$

If we substitute into the above equation the hypothesized solution for the approximate value function, we get:

$$(1-\gamma)V_0 \leq \sum_{\substack{(i,j,n)\in\\ [I]\times[J]\times[N]}} \Big(f^{WT}(i,n) + h\mu_j I(n=1) + \gamma^{n-1}\mu_j h - \gamma^{T(i)}\mu_j h - f^D(i)\Big) a_{ijn}$$

$$+ \sum_{\substack{(j,n)\in\\ [J]\times[N]}} \Big(\gamma^{T(i)}\mu_j h + f^D(i) - \gamma^{T(i)-1}\mu_j h\Big) y_{ij} + \gamma \sum_{\substack{(i,j)\in\\ [I]\times[J]}} W_{ij}E[Y_{ij}] - hC^R$$

Since the coefficient of y_{ij} is positive by Equation (C.1), we can substitute $\sum_{n\in[N]} a_{ijn} = y_{ij}$ to get:

$$(1-\gamma)V_0 \leq \sum_{\substack{(i,j,n)\in\\ [I]\times[J]\times[N]}} \Big(f^{WT}(i,n) + h\mu_j I(n=1) + (\gamma^{n-1} - \gamma^{T(i)-1})\mu_j h\Big) a_{ijn}$$

$$+ \sum_{\substack{i,j\in\\ [I]\times[J]}} \gamma^{T(i)}\mu_j h E[Y_{ij}] - hC^R$$

Again, by Equation (C.1), we can see that the coefficient of a_{ijn} is greater than 0 with equality only if $n = T(i)$. Thus, for state-action pairs that satisfy $\sum_{j\in[J]} \mu_j(x_{11} + \sum_{i\in[I]} a_{ij1}) > C^R$ the state action pair that provides the minimum of the right-hand side of the constraint for the primal LP will have $a_{ijn} = 0$ for all $n \neq T(i)$ and yield:

$$V_0 \leq \frac{1}{1-\gamma}\Big(\sum_{\substack{(i,j)\in\\ [J]\times[N]}} \gamma^{T(i)}\mu_j h E[Y_{ij}] - hC^R\Big) \tag{C.7}$$

which is true with equality for the hypothesized value of V_0.

For state-action pairs where $\sum_{j\in[J]} \mu_j(x_{j1} + \sum_{i\in[I]} a_{ij1}) \leq C^R$ the primal constraint simplifies to:

$$(1-\gamma)V_0 \leq \sum_{\substack{(i,j,n)\in\\ [I]\times[J]\times[N]}} \Big(f^{WT}(i,n) + h\mu_j I(n=1) + (\gamma^{n-1} - \gamma^{T(i)})\mu_j h\Big) a_{ijn}$$

$$+ \sum_{j\in[J]} \Big(-(u+h)\mu_j\Big) x_{j1} + \sum_{\substack{(i,j)\in\\ [I]\times[J]}} \gamma^{T(i)}\mu_j h E[Y_{ij}] + uC^R$$

The right-hand side takes its minimum value when $\sum_{j\in[J]} \mu_j x_{j1} = C^R$ and yields:

$$V_0 \leq \frac{1}{1-\gamma}\Bigg(\sum_{\substack{(i,j)\in \\ [J]\times[N]}} \gamma^{T(i)} \mu_j h E[Y_{ij}] - hC^R \Bigg)$$

This is the same lower bound for V_0 as in Equation (C.7) and thus is satisfied by the hypothesized value with equality when $\sum_{j\in[J]} \mu_j x_{j1} = C^R$. This proves the primal feasibility of the hypothesized solution under the given conditions with tight constraints for state-action pairs where:

$$\sum_{j\in[J]} \mu_j x_{j1} \geq C^R, \tag{C.8}$$

$$\sum_{n\in[N]} a_{ijn} = y_{ij} \quad \forall (i,j) \in [I]\times[J], \text{ and} \tag{C.9}$$

$$a_{ijn} = 0 \quad \forall (i,j) \in [I]\times[J] \text{ and } n \neq T(i) \tag{C.10}$$

We can now turn to the second stage of the proof where we demonstrate the existence of a dual solution that together with the hypothesized primal solution satisfies complementary slackness. To do so, we need to demonstrate the existence of a dual solution that is zero for all state-action pairs that do not satisfy the Conditions (C.8) to (C.10) and for which all the dual constraints are tight (since all the primary variables are non-zero). To ease the proof we impose the further condition that a dual variable is positive for a given state-action pair only if $\mu_j x_{jn}$ and $\mu_j a_{ijn}$ equal either zero or $C^R + C^{OT}$ for all i, j and $n > 1$. We first re-state the dual constraints:

$$(1-\gamma) \sum_{\substack{(\mathbf{s},\mathbf{a})\in \\ S\times A(s)}} X(\mathbf{s},\mathbf{a}) = 1 \tag{C.11}$$

$$\sum_{\substack{(\mathbf{s},\mathbf{a})\in \\ S\times A(s)}} X(\mathbf{s},\mathbf{a})\Bigg(x_{jn} - \gamma x_{j,n+1} - \gamma \sum_{i\in[I]} a_{ij,n+1} \Bigg) = E_\alpha[X_{jn}] \tag{C.12}$$

$$\sum_{\substack{(\mathbf{s},\mathbf{a})\in \\ S\times A(s)}} X(\mathbf{s},\mathbf{a})\Bigg(y_{ij} - \gamma E[Y_{ij}] \Bigg) = E_\alpha[Y_{ij}] \tag{C.13}$$

We let $B = \{(\mathbf{s},\mathbf{a}) \in S\times A(\mathbf{s}) | X(\mathbf{s},\mathbf{a}) > 0\}$. For $n = N$ and imposing the conditions above on the dual solution, Equation (C.12) yields:

$$\sum_{(\mathbf{s},\mathbf{a})\in B} X(\mathbf{s},\mathbf{a})x_{jN} = E_\alpha[X_{jN}] \quad \forall j \in [J]$$

$$\Rightarrow \sum_{\substack{(\mathbf{s},\mathbf{a})\in B \\ X_{jN}>0}} = \frac{E_\alpha[X_{jN}]}{C^R + C^{OT}}$$

Proceeding similarly for $n = N-1$, Equation (C.12) yields:

$$\begin{aligned}
&\sum_{(\mathbf{s},\mathbf{a})\in B} X(\mathbf{s},\mathbf{a})(X_{j,N-1} - \alpha X_{jN}) = E_\alpha[X_{j,N-1}] \\
&\qquad\Rightarrow \sum_{\substack{(\mathbf{s},\mathbf{a})\in B \\ X_{j,N-1}>0}} X(\mathbf{s},\mathbf{a})x_{j,N-1} = E_\alpha[X_{j,N-1}] + \alpha \sum_{\substack{(\mathbf{s},\mathbf{a})\in B \\ X_{jN}>0}} X(\mathbf{s},\mathbf{a})x_{jN} \\
&\qquad\Rightarrow \sum_{\substack{(\mathbf{s},\mathbf{a})\in B \\ X_{j,N-1}>0}} X(\mathbf{s},\mathbf{a}) = \frac{\mu_j}{C^R + C^{OT}}\Big(E_\alpha[X_{j,N-1}] + \gamma E_\alpha[X_{jN}]\Big)
\end{aligned}$$

Proceeding similarly, for all $n > T(I) - 1$ we get:

$$\sum_{\substack{(\mathbf{s},\mathbf{a})\in B \\ X_{jn}>0}} X(\mathbf{s},\mathbf{a}) = \frac{\mu_j}{C^R + C^{OT}} \sum_{m=n}^{N} \gamma^{m-n} E_\alpha[X_{jm}]$$

For $n = T(I) - 1$ there is the added complication that $a_{Ij,T(I)}$ may be non-zero. Equation (C.12) now yields:

$$\sum_{(\mathbf{s},\mathbf{a})\in B} X(\mathbf{s},\mathbf{a})\Big(x_{j,T(I)-1} - \gamma(x_{j,T(I)} + a_{Ij,T(I)})\Big) = E_\alpha[X_{j,T(I)-1}]$$

which implies that:

$$\sum_{\substack{(\mathbf{s},\mathbf{a})\in B \\ X_{j,T(I)-1}>0}} X(\mathbf{s},\mathbf{a})x_{j,T(I)-1} = E_\alpha[X_{j,T(I)-1}] + \gamma \sum_{\substack{(\mathbf{s},\mathbf{a})\in B \\ X_{j,T(I)}>0}} X(\mathbf{s},\mathbf{a})x_{j,T(I)} + \gamma \sum_{\substack{(\mathbf{s},\mathbf{a})\in B \\ a_{Ij,T(I)}>0}} X(\mathbf{s},\mathbf{a})a_{Ij,T(I)} \tag{C.14}$$

However, Equations (C.11) and (C.13) yield:

$$\sum_{\substack{(\mathbf{s},\mathbf{a})\in B\\ a_{Ij,T(I)}>0}} X(\mathbf{s},\mathbf{a})a_{Ij,T(I)} = \sum_{\substack{(\mathbf{s},\mathbf{a})\in B\\ a_{Ij,T(I)}>0}} X(\mathbf{s},\mathbf{a})y_{Ij} = E_\alpha[Y_{Ij}] + \frac{\lambda_{Ij}}{1-\gamma}$$

where λ_{Ij} is the arrival rate for patients of priority class I and service class j.

Assuming $\mu_j y_{Ij} = C^R + C^{OT}$ and $E_\alpha[Y_{Ij}]$ is set equal to λ_{Ij} we get:

$$\sum_{\substack{(\mathbf{s},\mathbf{a})\in B\\ a_{Ij,T(I)}>0}} X(\mathbf{s},\mathbf{a}) = \frac{1}{1-\gamma}\left(\frac{\mu_j\lambda_{Ij}}{C^R + C^{OT}}\right)$$

Thus, substituting back into Equation (C.14) we get:

$$\sum_{\substack{(\mathbf{s},\mathbf{a})\in B\\ x_{j,T(I)-1}>0}} X(\mathbf{s},\mathbf{a}) = \frac{\mu_j}{C^R + C^{OT}}\left(\sum_{m=T(I)-1}^{N} \gamma^{m-T(I)+1}E_\alpha[X_{jm}] + \frac{\gamma}{1-\gamma}\lambda_{Ij}\right)$$

Thus, in general, for $n > 1$ we get:

$$\sum_{\substack{(\mathbf{s},\mathbf{a})\in B\\ x_{jn}>0}} X(\mathbf{s},\mathbf{a}) = \frac{\mu_j}{C^R + C^{OT}}\left(\sum_{m=n}^{N} \gamma^{m-n}E_\alpha[X_{jm}] + \sum_{i=1}^{I}\frac{\gamma^{T(i)-n}}{1-\gamma}I(n < T(i))\lambda_{ij}\right) \tag{C.15}$$

For $n = 1$, Equation (C.12) still yields:

$$\begin{aligned}\sum_{\substack{(\mathbf{s},\mathbf{a})\in B\\ x_{j1}>0}} X(\mathbf{s},\mathbf{a})x_{j1} &= E_\alpha[X_{j1}] + \gamma \sum_{\substack{(\mathbf{s},\mathbf{a})\in B\\ x_{j2}>0}} X(\mathbf{s},\mathbf{a})x_{j2}\\ &= \sum_{m=1}^{N}\gamma^{m-1}E_\alpha[X_{jm}] + \sum_{i=1}^{I}\frac{\gamma^{T(i)-1}}{1-\gamma}I(1 < T(i))\lambda_{ij} \end{aligned} \tag{C.16}$$

However, unlike for the cases with $n > 1$ we cannot simply set x_{j1} equal to $C^R + C^{OT}$ or 0 since we need to enforce by Condition (C.8) that $\sum_{j\in J}\mu_j x_{j1} \geq C^R$ for all $(\mathbf{s},\mathbf{a}) \in B$. Thus, along with satisfying Equation (C.16), our dual solution must also satisfy:

$$\begin{aligned}\sum_{(\mathbf{s},\mathbf{a})\in B} X(\mathbf{s},\mathbf{a})\sum_{j\in J}\mu_j x_{j1} &\geq \frac{C^R}{1-\gamma}\\ \Rightarrow \sum_{j\in[J]}\mu_j\left(\sum_{m=1}^{N}\gamma^{m-1}E_\alpha[X_{jm}] + \sum_{i=1}^{I}\frac{\gamma^{T(i)-1}}{1-\gamma}I(1 < T(i))\lambda_{ij})\right) &\geq \frac{C^R}{1-\gamma}\end{aligned} \tag{C.17}$$

which is enforced by Condition (C.2).

The Existence of a Dual Solution Satisfying Complementary Slackness

The above argument suggests a weighting scheme for a dual feasible solution that, together with the proposed primal solution, satisfies complementary slackness. It remains to prove that a dual solution satisfying the above weighting scheme must exist. We can determine the state-action pairs with positive dual weight starting on day 1 and working up to day N. A dual solution exists if:

1. The total dual weight available $\frac{1}{1-\gamma}$ given by Equation (C.11) does not exceed the combined weight assigned to all states where x_{jn} or $\sum_{i=1}^{I} a_{ijn}$ are greater than zero for any n.
2. The total weight assigned to dual variables where $\sum_{n=1}^{N} a_{ijn}$ is positive and is equal to the weight assigned to dual variables where y_{ij} is positive.

This turns out to be straightforward as it is easy to show that, under the above weighting scheme and using Equation (C.3),

$$\sum_{x_{jn}>0} (X(\mathbf{s},\mathbf{a})\mu_j x_{jn} + \sum_{i=1}^{I} \sum_{a_{ijn}>0} X(\mathbf{s},\mathbf{a})\mu_j a_{ijn}) \leq \frac{C_R + C^{OT}}{1-\gamma}$$

$$\Rightarrow \sum_{x_n>0} X(\mathbf{s},\mathbf{a}) + \sum_{i=1}^{I} \sum_{a_{in}>0} X(\mathbf{s},\mathbf{a}) \leq \frac{1}{1-\gamma}$$

for all $n \in [N]$ and

$$\sum_{n=1}^{N} \sum_{(\mathbf{s},\mathbf{a}\in B|a_{ijn}>0} X(\mathbf{s},\mathbf{a}) a_{ijn} = \sum_{(\mathbf{s},\mathbf{a}\in B|y_{ij}>0} X(\mathbf{s},\mathbf{a}) y_{ij}$$

for all $(i,j) \in [I] \times [J]$. Thus, there exist admissible state-action pairs that satisfy the above weighting scheme proving the existence of the required dual solution.