Dynamic Knapsack Problem

AS, WM, KV

May 15, 2021

Contents

1	Questions/Comments	
	.1 Questions	
	.2 Comments	
2	Assumptions	
3	MDP Model	
	.1 Decision Epochs	
	.2 State Space	
	3.2.1 Description	
	.3 Action Sets	
	3.3.1 Description	
	3.3.2 Auxiliary Variables	
	3.3.3 Auxiliary Variable Definition	
	3.3.4 State-Action Constraints	
	.4 Transition Probabilities	
	3.4.1 Uncertainty Sources	
	3.4.2 Transition Constraints	
	.5 Costs	
4	P Model	
	.1 Full LP	
5	Aproximate Dynamic Programming Model	
	.1 Full ADP Model	
	.2 Expectation ADP Model	
	.3 ADP Master Problem	
	.4 ADP Pricing Problem	
6	Solution Explanation	
	.1 Algorithm for solving	
	.2 Generating an Action	

1 Questions/Comments

1.1 Questions

• Priority need to be added

1.2 Comments

•

2 Assumptions

- Assume all new patients are arriving at the beginning of the period
 - Assume arrivals follow poisson distribution
- It is assumed that there is no difference between how long a patient waits for their appointment within some limit (TL).
 - Let's assume our limit is 10 periods, there is no difference in terms of cost between patients who has been waiting for 1 period and patients waiting 10 period for their appointment (assuming they are the same category).
 - However, if a patient waits for 11 period for their appointment there should be some kind of penalty for the
 wait
 - The penalty comes from increasing patient complexity. Higher complexities require more resources and are thus more costly overall in the system
- We assume patient complexity transitions follow binary distribution each period a patient can become more complex with a certain probability
 - While patient is waiting less that TL, the transition probability is 0.
 - After TL, transition probability is some arbitrary value
- It is assumed if a patient is scheduled in period one, it means they are served "immediately", regardless if in practice the appointment is at the beginning or end of the period.
- We distinguish between two types of reschedules: good and bad reschedules
 - Good reschedules are reschedules where a patient is rescheduled to an earlier period
 - Bad reschedules are reschedules where a patient is rescheduled to a later period
- We assume that only specific reschedules are allowed (to simplify the model, and remove redundancies)
 - Good reschedules are allowed from any period after 2 into period 1
 - Bad reschedules are only allowed from period 1 to period 2
- It is assumed that there is a certain default expected number of PPE units available for all periods
 - However, in the period 1, there is some random deviation from the expected number of units
 - This random deviation follows some uniform distribution
- We allow some violation of PPE units, in rare cases (with high cost), in order to accommodate changes due to variability.
 - If we have violation of PPE units it reduces capacity in the future (Meaning violations are borrowed from future capacity)

3 MDP Model

3.1 Decision Epochs

Decisions are made at the beginning of each time period. There are 3 relevant time intervals to look at:

- Pre-decision state (S)
 - Pre-decision state defines the initial state on which a decision is required
 - This information drives decision making
- Post-decision state (\hat{S})
 - Post-decision state is the time when the decision has been executed, but no new info came in
 - This state defines the immediate cost of an action
- Post-transition state (S')
 - This is the state when new information has come in (transition randomness). Defines next pre decision state
 - It is primarily used to generate expectation for the ADP

3.2 State Space

3.2.1 Description

State is defined by current available and used resources, patient waitlist/demand, and patients already scheduled

$$\vec{S} = (\vec{ue}, \vec{uu}, \vec{pw}, \vec{ps})$$

- $\vec{ue} = ue_{tp}$ Expected unist for period t, and resource p
- $u\bar{u} = uu_{tp}$ Used units for period t, and resource p
- $\vec{pw} = pw_{mdc}$ Patients of complexity d, CPU c, on a wait list for m periods (m of 0 just arrived)
- $\vec{ps} = ps_{tmdc}$ Patients of complexity d, CPU c, scheduled to period t, waiting for m periods (m of 0 just arrived)

3.3 Action Sets

3.3.1 Description

Decision consists of rescheduling currently scheduled patients, and scheduling patients on waitlist. There are also some goal and auxiliary variables

$$\vec{A} = (\vec{sc}, \vec{rsc}, \vec{uv}, \text{auxiliary variables})$$

- $\vec{sc} = sc_{tmdc}$ Patients of complexity d, CPU c, waiting for m periods, to schedule in period t
- $r\vec{s}c = rsc_{tt'mdc}$ -Patients of complexity d, CPU c, waiting for m periods, to reschedule from period t to period t
- $\vec{uv} = uv_{tp}$ goal variable, violation on number of resources used for period t, of resource p

3.3.2 Auxiliary Variables

- $\hat{u}u_{tp}$ post-decision units used
- \hat{ps}_{tmdc} post-decision patients scheduled

3.3.3 Auxiliary Variable Definition

$$\hat{u}u_{tp} = \sum_{mdc} U_{pdc} \hat{ps}_{tmdc} \qquad \forall tp \qquad (1)$$

$$\hat{pw}_{mdc} = pw_{mdc} - \sum_{t} sc_{tmdc} \qquad \forall mdc \qquad (2)$$

$$\hat{ps}_{tmdc} = ps_{tmdc} + sc_{tmdc} + \sum_{t} rsc_{tt'mdc} - \sum_{t'} rsc_{tt'mdc} \quad \forall tmdc$$
 (3)

3.3.4 State-Action Constraints

• Consistency Constraint

$$uu_{tp} = \sum_{mdc} U_{pdc} p s_{tmdc} \quad \forall tp \tag{4}$$

• Resource Usage Constraint

$$\hat{uu}_{tp} \le ue_{tp} + uv_{tp} \quad \forall tp \tag{5}$$

• Custom bounds on when reschedules are allowed

$$rsc_{tt'mdc} = 0 \quad \forall t \in \{2...T\}, t' \in \{2...T\} mdc$$
 (6)

$$rsc_{tt'mdc} = 0 \quad \forall t \in \{1\}, t' \in \{3...T\} mdc \tag{7}$$

$$rsc_{tt'mdc} = 0 \quad \forall tt'mdc, \text{ where } t = t'$$
 (8)

• Bounds on Schedules/Reschedules

$$\sum_{t'} rsc_{tt'mdc} \le ps_{tmdc} \quad \forall tmdc \tag{9}$$

$$\sum_{t'} rsc_{tt'mdc} \le ps_{tmdc} \quad \forall tmdc$$

$$\sum_{t} sc_{tmdc} \le pw_{mdc} \quad \forall mdc$$

$$(9)$$

• Bounds on states

$$ue_{tp} \leq uen_p + max(ued_p) \quad \forall tp$$
 (11)

$$uu_{tp} \leq uen_p + max(ued_p) \quad \forall tp$$
 (12)

$$pw_{mdc} \le pea_{dc} * 4 \qquad \forall mdc \tag{13}$$

$$ps_{tmdc} \le pea_{dc} * 4 \qquad \forall tmdc$$
 (14)

3.4 Transition Probabilities

Uncertainty Sources 3.4.1

- 1. Number of patients arriving this period pw_{0dc}
 - peadc represents the random number of patients arriving. Follows a poisson distribution.
- 2. Transition between patient complexities within the wait list $pw_{m>=1,dc}$
 - pwt_{mdc} represents the random number of patients of complexity d, CPU c, waiting for m period, who became more complex. Follows binomial distribution.
 - ptp_{mdc} represents transition probability to a higher complexity
- 3. Transition between patient complexities within the scheduled list ps_{tmdc}
 - pst_{tmdc} represents the random number of patients of complexity d, CPU c, waiting for m period, scheduled into t period, who became more complex. Follows binomial distribution.
- 4. Amout of expected units of PPE resource for the next time period ue_{1p}
 - ued_p represents the random deviation of resource p from the expectation. Follows uniform distribution.
 - uen_p expected units of resource p per period

3.4.2 **Transition Constraints**

1. Transition from \vec{ue} to $\vec{ue'}$ - Expected Resources

$$ue'_{1p} = \underbrace{ue_{2p} + ued_p}_{\text{deviation}} + \underbrace{ue_{1p} - \hat{uu}_{1p}}_{\text{unused}} \quad \forall p \in \{\text{Carry Over Resources}\}$$
 (15)

$$ue'_{1p} = ue_{2p} + ued_p$$
 $\forall p \in \{\text{Non Carry Over Resources}\}$ (16)

$$ue'_{tp} = ue_{t+1,p}$$
 $\forall t \in \{2...T-1\}p$ (17)

$$ue'_{T_p} = uen_p$$
 $\forall p$ (18)

2. Transition from \vec{uu} to $\vec{uu'}$ - Used Resources

$$uu'_{tp} = \hat{u}u_{t+1,p} + \sum_{mdc} pst_{tmdc}(U_{pd+1c} - U_{pdc}) \quad \forall t \in \{1...T - 1\}$$

$$(19)$$

$$uu'_{Tp} = 0 \forall p (20)$$

3. Transition from $p\vec{w}$ to $p\vec{w}'$ - Flow of patients on waitlist

$$pw'_{0dc} = pea_{dc} \forall dc (21)$$

$$pw'_{mdc} = p\hat{w}_{m-1,dc} + pwt_{m-1,d-1,c} - pwt_{m-1,dc} \quad \forall m \in \{1...M-1\}dc$$
 (22)

$$pw'_{mdc} = p\hat{w}_{m-1,dc} + pwt_{m-1,d-1,c} - pwt_{m-1,dc} \quad \forall m \in \{1...M-1\}dc$$

$$pw'_{Mdc} = \sum_{M=1}^{M} \left(p\hat{w}_{mdc} + pwt_{md-1,c} - pwt_{mdc}\right) \quad \forall dc$$

$$(22)$$

4. Transition from \vec{ps} to $\vec{ps'}$ - Flow of patiensts scheduled

$$ps'_{t0dc} = 0 \forall tdc (24)$$

$$ps'_{Tmdc} = 0$$
 $\forall mdc$ (25)

$$ps'_{tmdc} = \hat{ps}_{t+1,m-1,dc} + \underbrace{pst_{t+1,m-1,d-1,c} - pst_{t+1,m-1,dc}}_{\text{change in complexities}} \quad \forall t \in \{1...T-1\}m \in \{1...M-1\}dc$$
 (26)

change in complexities
$$ps'_{tmdc} = \hat{ps}_{t+1,m-1,dc} + \underbrace{pst_{t+1,m-1,d-1,c} - pst_{t+1,m-1,dc}}_{\text{change in complexities}} \quad \forall t \in \{1...T-1\}m \in \{1...M-1\}dc \qquad (26)$$

$$ps'_{tMdc} = \sum_{M=1}^{M} \left(\hat{ps}_{t+1mdc} + \underbrace{pst_{t+1,md-1,c} - pst_{t+1,mdc}}_{\text{change in complexities}}\right) \quad \forall t \in \{1...T-1\}dc \qquad (27)$$

3.5 Costs

Cost will come from 3 sources:

- waiting (cw)
- rescheduling $(cc = \lceil cw^M \rceil)$
- \bullet resource violation (M)

$$C = c(p\hat{\vec{w}}, p\hat{\vec{s}}, r\vec{s}c, u\vec{v}) = \sum_{mdc}^{\text{cost of waiting}} cw^{m}(p\hat{w}_{mdc}) + \sum_{tdc}^{\text{cov}} cw^{M}(p\hat{s}_{tMdc}) + \sum_{tdc}^{\text{Good Reschedule}} cw^{M}(p\hat$$

LP Model 4

Full LP 4.1

Given a full MDP model, the equivalent LP would look as follows:

$$\max_{\vec{s}} \sum \alpha(\vec{S}) v(\vec{S}) \tag{29}$$

subject to

$$c(p\hat{\vec{w}}, p\hat{\vec{s}}, r\vec{s}c, \vec{uv}) + \gamma \sum_{\vec{n}} p(p\vec{e}a, p\vec{w}t, p\vec{s}t, \vec{ued}) v(\vec{S'}|\vec{S}, \vec{A}, p\vec{e}a, p\vec{w}t, p\vec{s}t, \vec{ued}) \ge v(\vec{S}) \quad \forall \vec{S}\vec{A}$$
 (30)

5 Aproximate Dynamic Programming Model

5.1 Full ADP Model

Let's convert it into ADP. We do that by changing $v(\vec{S})$ to an approximation as follows:

$$v(\vec{ue}, \vec{uu}, \vec{pw}, \vec{ps}) = \beta^0 + \sum_{tp} \beta_{tp}^{ue} ue_{tp} + \sum_{tp} \beta_{tp}^{uu} uu_{tp} + \sum_{mdc} \beta_{mdc}^{pw} pw_{mdc} + \sum_{tmdc} \beta_{tmdc}^{ps} ps_{tmdc}$$

$$(31)$$

This gives the following LP

$$\max_{\vec{\beta}} \sum \alpha(\vec{ue}, \vec{uu}, \vec{pw}, \vec{pe}, \vec{ps}) v(\vec{ue}, \vec{uu}, \vec{pw}, \vec{pe}, \vec{ps})$$
(32)

Subject to:

$$c(p\vec{w}, p\vec{s}, r\vec{s}c, u\vec{v}) + \gamma \sum_{\vec{p}} p \left(\beta^{0} + \left(\sum_{p} \beta_{1p}^{ue}(ue_{2p} + ued_{p} + ue_{1p} - \hat{u}u_{1p})\right) + \left(\sum_{t=2,p}^{T-1} \beta_{tp}^{ue}ue_{t+1,p}\right) + \left(\sum_{p} \beta_{Tp}^{ue}uen_{d}\right) + \left(\sum_{tp} \beta_{tp}^{ue}(\hat{u}u_{t+1,p} + \sum_{mdc} pst_{tmdc}(U_{pd+1c} - U_{pdc}))\right) + \left(\sum_{p} \beta_{Tp}^{uu} * 0\right) + \left(\sum_{dc} \beta_{0dc}^{pw}pea_{dc}\right) + \left(\sum_{m=1,dc} \beta_{mdc}^{pw}(\hat{p}\hat{w}_{m-1dc} + pwt_{m-1d-1c} - pwt_{m-1dc})\right) + \left(\sum_{dc} \beta_{Mdc}^{pw}(\sum_{M-1}^{M} \hat{p}\hat{w}_{mdc} + pwt_{md-1c} - pwt_{mdc})\right) + \left(\sum_{tdc} \beta_{t0dc}^{ps}(0) + \left(\sum_{mdc} \beta_{Tmdc}^{ps}(0) + \left(\sum_{t=1m=1dc} \beta_{tmdc}^{ps}(\hat{p}\hat{s}_{t+1m-1dc} + pst_{t+1m-1d-1c} - pst_{t+1m-1dc})\right) + \left(\sum_{t=1Mdc} \beta_{tmdc}^{ps}(\sum_{M-1}^{M} (\hat{p}\hat{s}_{t+1mdc} + pst_{t+1md-1c} - pst_{t+1mdc})\right)\right) + \left(\sum_{t=1Mdc} \beta_{tmdc}^{ps}(\sum_{M-1}^{M} (\hat{p}\hat{s}_{t+1mdc} + pst_{t+1md-1c} - pst_{t+1mdc})\right)\right) \geq v(u\vec{e}, p\vec{w}, p\vec{s}) \quad \forall \vec{S}\vec{A}$$

$$\vec{\beta} \geq 0 \quad \forall \beta \qquad (34)$$

And all auxiliary constraints in "Auxiliary Variable Definition"

5.2 Expectation ADP Model

Steps for rearranging and converting the constraint (33) into expectation:

- $E[ued_p] = 0$
- $E[pwt_{mdc}] = ptp_{mdc} * p\hat{w}_{mdc}$
- $E[pst_{tmdc}] = ptp_{mdc} * \hat{ps}_{tmdc}$

$$\beta^0: \qquad (1-\gamma)\beta^0 \tag{35}$$

$$\beta_{tp}^{ue}: \qquad \sum_{p} \beta_{1p}^{ue} \left(ue_{1p} - \gamma (uen_p + ue_{1p} - \hat{u}u_{1p}) \right) + \sum_{t=2p}^{T} \beta_{tp}^{ue} \left(ue_{tp} - \gamma (uen_p) \right)$$
(36)

$$\beta_{tp}^{uu}: \sum_{tp}^{T-1} \beta_{tp}^{uu} \left(uu_{tp} - \gamma (\hat{u}u_{t+1p} + \sum_{mdc} E[pst_{t+1mdc}](U_{pd+1c} - U_{pdc})) \right) + \sum_{p} \beta_{Tp}^{uu} \left(uu_{Tp} \right)$$
(37)

$$\beta_{mdc}^{pw}: \sum_{dc} \beta_{0dc}^{pw} \Big(pw_{0dc} - \gamma(pea_{dc}) \Big) + \sum_{m=1dc}^{M-1} \beta_{mdc}^{pw} \Big(pw_{mdc} - \gamma(\hat{pw}_{m-1dc} + E[pwt_{m-1d-1c}] - E[pwt_{m-1dc}]) \Big) + \sum_{dc} \beta_{Mdc}^{pw} \Big(pw_{Mdc} - \gamma \sum_{M-1}^{M} (\hat{pw}_{mdc} + E[pwt_{md-1c}] - E[pwt_{mdc}]) \Big)$$
(38)

$$\beta_{tmdc}^{ps} : \sum_{tdc} \beta_{t0dc}^{ps} \left(ps_{t0dc} \right) + \sum_{mdc} \beta_{Tmdc}^{ps} \left(ps_{Tmdc} \right) + \sum_{tdc} \beta_{tmdc}^{ps} \left(ps_{tmdc} - \gamma(\hat{p}s_{t+1m-1dc} + E[pst_{t+1m-1d-1c}] - E[pst_{t+1m-1dc}]) \right) + \sum_{t=1}^{T-1} \beta_{tmdc}^{ps} \left(ps_{tmdc} - \gamma(\hat{p}s_{t+1m-1dc} + E[pst_{t+1m-1d-1c}] - E[pst_{t+1m-1dc}]) \right) + \sum_{t=1}^{T-1} \beta_{tmdc}^{ps} \left(ps_{tMdc} - \gamma \sum_{M-1}^{M} (\hat{p}s_{t+1mdc} + E[pst_{t+1md-1c}] - E[pst_{t+1mdc}]) \right)$$
(39)

Let's say E[V] is the addition of all parts above. Then the ADP model converted to expectation would look as follows:

$$\max_{\vec{\beta}} \left(\sum \beta^0 + \sum \beta^{ue} E[ue] + \sum \beta^{uu} E[uu] + \sum \beta^{uv} E[uv] + \sum \beta^{pw} E[pw] + \sum \beta^{ps} E[ps] \right) \tag{40}$$

Subject to:

$$E[V] \le c(p\vec{w}, \vec{sc}, r\vec{sc}, \vec{uv}) \quad \forall \vec{S}\vec{A}$$

$$\tag{41}$$

5.3 ADP Master Problem

Converting to Dual

$$\min_{w} \sum_{\vec{S}\vec{A}} w(\vec{S}\vec{A})c(\vec{S}\vec{A}) \tag{42}$$

Subject To:

$$\beta^{0}: \sum_{\vec{S}\vec{A}} \vec{w}(1-\gamma) = 1 \tag{43}$$

$$\beta^{ue}: \sum_{\vec{S}\vec{A}} \vec{w} \Big(ue_{1p} - \gamma (uen_p + ue_{1p} - \hat{u}u_{1p}) \Big) \ge E[ue_{tp}]$$
 $\forall p$ (44)

$$\sum_{\vec{S}\vec{A}} \vec{w} \Big(ue_{tp} - \gamma (uen_p) \Big) \ge E[ue_{tp}] \qquad \forall t \in (2...T)p$$
(45)

$$\beta^{uu} : \sum_{\vec{S}\vec{A}} \vec{w} \Big(uu_{tp} - \gamma (\hat{u}u_{t+1p} + \sum_{mdc} E[pst_{t+1mdc}](U_{pd+1c} - U_{pdc})) \Big) \ge E[uu_{tp}] \qquad \forall t \in \{1...T - 1\}p$$
(46)

$$\sum_{\vec{S}\vec{A}} \vec{w} \Big(u u_{Tp} \Big) \ge E[u u_{Tp}] \tag{47}$$

$$\beta^{pw} : \sum_{\vec{c} \in \vec{s}} \vec{w} \Big(pw_{0dc} - \gamma(pea_{dc}) \Big) \ge E[ps_{0dc}]$$
 $\forall dc$ (48)

$$\sum_{\vec{c} \cdot \vec{k}} \vec{w} \Big(p w_{mdc} - \gamma (\hat{p} w_{m-1dc} + E[p w t_{m-1d-1c}] - E[p w t_{m-1dc}]) \Big) \ge E[p w_{mdc}] \ \forall m \in \{1...M-1\} dc$$
 (49)

$$\sum_{\vec{c},\vec{\Lambda}} \vec{w} \Big(pw_{Mdc} - \gamma \sum_{M=1}^{M} (\hat{p}w_{mdc} + E[pwt_{md-1c}] - E[pwt_{mdc}]) \Big) \ge E[pw_{Mdc}] \quad \forall dc$$
 (50)

$$\beta^{ps}: \sum_{\vec{S}\vec{A}} \vec{w} \Big(ps_{t0dc} \Big) \ge E[ps_{t0dc}]$$
 $\forall tdc$ (51)

$$\sum_{\vec{S}\vec{A}} \vec{w} \Big(ps_{Tmdc} \Big) \ge E[ps_{Tmdc}] \qquad \forall mdc \qquad (52)$$

$$\sum_{\vec{S}\vec{A}} \vec{w} \Big(ps_{tmdc} - \gamma (\hat{ps}_{t+1m-1dc} + E[pst_{t+1m-1d-1c}] - E[pst_d]) \Big) \ge E[ps_{tmdc}] \ \forall \{1...T-1\} \{1...M-1\} dc \qquad (53)$$

$$\sum_{\vec{S}\vec{A}} \vec{w} \Big(p s_{tMdc} - \gamma \sum_{M=1}^{M} (\hat{p} s_{t+1mdc} + E[p s t_{t+1md-1c}] - E[p s t_d]) \Big) \ge E[p s_{tMdc}] \, \forall \{1...T - 1\} dc$$
 (54)

State Action Bounds:
$$w \ge 0 \quad \forall w$$
 (55)

5.4 ADP Pricing Problem

$$\min_{\substack{(\vec{u}e, \vec{u}u, \vec{p}\vec{w}, \vec{p}e, \vec{p}\vec{s}) \in S, (\vec{sc}, r\vec{s}c) \in A}} c(\vec{p}\vec{w}, \vec{s}c, r\vec{s}c, \vec{u}v) - E[V]$$
(56)

Subject to: constraints in sections "State-Action Constraints" and "Auxiliary Variable Definition"

6 Solution Explanation

6.1 Algorithm for solving

To get $\vec{\beta}$ values, which will be used to generate an action follow steps below:

- 1. Perform a monte-carlo simulation (following some arbitrary policy) to get E[ue], E[uu], E[pw]E[pw]
- 2. Create an initial feasible set of state-action pairs \vec{w}
- 3. Solve model in section "ADP Master Problem" where each state-action pairs in \vec{w} corresponds to a variable and parameters for all the constraints for a specific action.
- 4. Solve model in section "ADP Pricing Problem", where duals from problem in step 3 correspond to $\vec{\beta}$ values.
 - If objective function is less than 0, add solution as a single state-action pair to \vec{w} and go to step 3
 - If objective function is greater than 0, continue to next step
- 5. Duals from problem in step 3 correspond to final $\vec{\beta}$ values

6.2 Generating an Action

Once $\vec{\beta}$ values have been approximated, you may use the model below to generate a recommended action for a specific state.

$$\min_{\vec{A}} c(\vec{S}, \vec{A}) - \gamma v(S')$$

$$\begin{split} v(S') &= \left(\beta^0 + \sum_{t=2p}^T \beta_{tp}^{ue} uen_p + \\ &\sum_{p} \beta_{1p}^{ue} \left(uen_p + ue_{1p} - uu_{1p} - \sum_{dc} U_{pdc} (\sum_{m} sc_{1mdc} + \sum_{t'm} rsc_{1t'mdc} - \sum_{tm} rsc_{t1mdc}) \right) + \\ &\sum_{p} \beta_{1p}^{ue} \left(uu_{t+1p} + \sum_{mdc} ps_{tmdc} pwp(U_{pd+1c} - U_{pdc}) + U_{pdc} (\sum_{mdc} sc_{t+1} - \sum_{t'mdc} rsc_{t+1t'} + \sum_{tmdc} rsc_{tt'+1}) \right) + \\ &\sum_{tp} \beta_{0dc}^{pw} E[pea_{dc}]) + \\ &\sum_{m=1,dc} \beta_{mdc}^{pw} \left(pw_{m-1dc} (1-pwp) + pw_{m-1d-1c} pwp - \sum_{t} sc_{tm-1dc} \right) + \\ &\sum_{dc} \beta_{mdc}^{pw} \sum_{M-1} \left(pw_{mdc} (1-pwp) + pw_{md-1c} pwp - \sum_{t} sc_{tmdc} \right) + \\ &\sum_{t=1m=1dc} \beta_{tmdc}^{pw} \left(ps_{t+1m-1dc} (1-pwp) + ps_{t+1m-1d-1c} pwp - \sum_{t} sc_{t+1m-1dc} - \\ &- \sum_{t'} rsc_{t+1t'm-1dc} + \sum_{t} rsc_{tt'+1m-1dc} \right) + \end{split}$$

$$\sum_{t=1dc}^{T-1} \beta_{tMdc}^{ps} \sum_{M-1}^{M} \left(ps_{t+1mdc} (1 - pwp) + ps_{t+1md-1c} pwp - \sum_{t} sc_{t+1mdc} - \sum_{t'} rsc_{t+1t'mdc} + \sum_{t} rsc_{tt'+1mdc} \right)$$

Subject to: constraints in section "State-Action Constraints"