

# Dynamic Knapsack Problem

AS, WM, KV

May 25, 2021

## Contents

<b>1</b>	<b>Questions/Comments</b>	<b>2</b>
1.1	Questions . . . . .	2
1.2	Comments . . . . .	2
<b>2</b>	<b>Assumptions</b>	<b>2</b>
<b>3</b>	<b>MDP Model</b>	<b>3</b>
3.1	Decision Epochs . . . . .	3
3.2	State Space . . . . .	3
3.2.1	Description . . . . .	3
3.3	Action Sets . . . . .	3
3.3.1	Description . . . . .	3
3.3.2	Auxiliary Variables . . . . .	3
3.3.3	Auxiliary Variable Definition . . . . .	3
3.3.4	State-Action Constraints . . . . .	4
3.4	Transition Probabilities . . . . .	4
3.4.1	Uncertainty Sources . . . . .	4
3.4.2	Transition Constraints . . . . .	4
3.5	Costs . . . . .	5
<b>4</b>	<b>LP Model</b>	<b>5</b>
4.1	Full LP . . . . .	5
<b>5</b>	<b>Aproximate Dynamic Programming Model</b>	<b>6</b>
5.1	Full ADP Model . . . . .	6
5.2	Expectation ADP Model . . . . .	6
5.3	ADP Master Problem . . . . .	7
5.4	ADP Pricing Problem . . . . .	8
<b>6</b>	<b>Solution Explanation</b>	<b>8</b>
6.1	Algorithm for solving . . . . .	8
6.2	Generating an Action . . . . .	8

# 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 than  $TL$ , 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 accomodate changes due to variability.
  - If we have violation of PPE units - the extra capacity comes from external source

### 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

$$\vec{S} = (u\vec{e}, u\vec{u}, p\vec{w}, p\vec{s})$$

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

$$\vec{A} = (\vec{s}\vec{c}, r\vec{s}\vec{c}, u\vec{v}, \text{auxiliary variables})$$

- $\vec{s}\vec{c} = sc_{tmdc}$  - Patients of complexity  $d$ , CPU  $c$ , waiting for  $m$  periods, to schedule in period  $t$
- $r\vec{s}\vec{c} = rsc_{tt'mdc}$  - Patients of complexity  $d$ , CPU  $c$ , waiting for  $m$  periods, to reschedule from period  $t$  to period  $t'$
- $u\vec{v} = uv_{tp}$  - 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
- $\hat{p}w_{mdc}$  - post-decision patients waiting
- $\hat{p}s_{tmdc}$  - post-decision patients scheduled

##### 3.3.3 Auxiliary Variable Definition

$$\hat{u}u_{tp} = \sum_{mdc} U_{pdc} \hat{p}s_{tmdc} \quad \forall tp \quad (1)$$

$$\hat{p}w_{mdc} = pw_{mdc} - \sum_t sc_{tmdc} \quad \forall mdc \quad (2)$$

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

### 3.3.4 State-Action Constraints

- Consistency Constraint

$$uu_{tp} = \sum_{mdc} U_{pdc} ps_{tmdc} \quad \forall tp \quad (4)$$

- Resource Usage Constraint

$$\hat{u}u_{tp} \leq ue_{tp} + uv_{tp} \quad \forall tp \quad (5)$$

- Custom bounds on when reschedules are allowed

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

$$rsc_{tt'mdc} = 0 \quad \forall tt' mdc, \text{ where } t = t' \quad (7)$$

- Bounds on Schedules/Reschedules

$$\sum_{t'} rsc_{tt'mdc} \leq ps_{tmdc} \quad \forall tmdc \quad (8)$$

$$\sum_t sc_{tmdc} \leq pw_{mdc} \quad \forall mdc \quad (9)$$

- Bounds on states

$$ue_{tp} \leq uen_p + \max(ued_p) \quad \forall tp \quad (10)$$

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

$$pw_{mdc} \leq pea_{dc} * 4 \quad \forall mdc \quad (12)$$

$$ps_{tmdc} \leq pea_{dc} * 4 \quad \forall tmdc \quad (13)$$

## 3.4 Transition Probabilities

### 3.4.1 Uncertainty Sources

1. Number of patients arriving this period -  $pw_{0dc}$ 
  - $pea_{dc}$  - represents the random number of patients arriving. Follows a poisson distribution.
2. Transition between patient complexities within the wait list -  $pw_{m>=1,dc}$ 
  - $pwt_{tmdc}$  represents the random number of patients of complexity  $d$ , CPU  $c$ , waiting for  $m$  period, who became more complex. Follows binomial distribution.
  - $ptp_{tmdc}$  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. Amount 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

$$ue'_{1p} = \overbrace{ue_{2p} + ued_p}^{\text{deviation}} + \overbrace{ue_{1p} - \hat{u}u_{1p}}^{\text{unused}} + uv_{1p} \quad \forall p \in \{\text{Carry Over Resources}\} \quad (14)$$

$$ue'_{1p} = ue_{2p} + ued_p \quad \forall p \in \{\text{Non Carry Over Resources}\} \quad (15)$$

$$ue'_{tp} = ue_{t+1,p} \quad \forall t \in \{2...T-1\}p \quad (16)$$

$$ue'_{Tp} = uen_p \quad \forall p \quad (17)$$

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

$$uu'_{tp} = \hat{u}_{t+1,p} + \overbrace{\sum_{mdc} pst_{tmdc}(U_{pd+1c} - U_{pdc})}^{\text{change in complexities}} \quad \forall t \in \{1 \dots T-1\} \quad (18)$$

$$uu'_{Tp} = 0 \quad \forall p \quad (19)$$

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

$$pw'_{0dc} = pea_{dc} \quad \forall dc \quad (20)$$

$$pw'_{mdc} = \hat{p}w_{m-1,dc} + \overbrace{pwt_{m-1,d-1,c} - pwt_{m-1,dc}}^{\text{change in complexities}} \quad \forall m \in \{1 \dots M-1\}dc \quad (21)$$

$$pw'_{Mdc} = \sum_{M-1}^M (\hat{p}w_{mdc} + \overbrace{pwt_{md-1,c} - pwt_{mdc}}^{\text{change in complexities}}) \quad \forall dc \quad (22)$$

## 4. Transition from $\vec{p}$ to $\vec{p}'$ - Flow of patients scheduled

$$ps'_{t0dc} = 0 \quad \forall tdc \quad (23)$$

$$ps'_{Tmdc} = 0 \quad \forall mdc \quad (24)$$

$$ps'_{tmdc} = \hat{p}s_{t+1,m-1,dc} + \overbrace{pst_{t+1,m-1,d-1,c} - pst_{t+1,m-1,dc}}^{\text{change in complexities}} \quad \forall t \in \{1 \dots T-1\}m \in \{1 \dots M-1\}dc \quad (25)$$

$$ps'_{tMdc} = \sum_{M-1}^M (\hat{p}s_{t+1mdc} + \overbrace{pst_{t+1,md-1,c} - pst_{t+1,mdc}}^{\text{change in complexities}}) \quad \forall t \in \{1 \dots T-1\}dc \quad (26)$$

## 3.5 Costs

Cost will come from 4 sources:

- waiting ( $cw$ )
- rescheduling ( $cc$ )
- scheduling ( $cs$ )
- resource violation ( $M$ )

$$\begin{aligned} C = c(\vec{p}, \vec{p}', r\vec{s}, \vec{u}) &= \overbrace{\sum_{mdc} cw^m(\hat{p}w_{mdc})}^{\text{cost of waiting}} + \overbrace{\sum_{tmdc} cw^M(\hat{p}s_{tMdc})}^{\text{prevent infinite reschedules}} + \\ &+ \overbrace{\sum_{tmdc} cs^t(sc_{tmdc})}^{\text{Prefer earlier appointments}} \\ &+ \overbrace{(1.5cc) \sum_{\substack{tt'mdc \\ t' > t}} rsc_{tt'mdc}}^{\text{Bad Reschedule}} - \overbrace{(0.5cc) \sum_{\substack{tt'mdc \\ t' < t}} rsc_{tt'mdc}}^{\text{Good Reschedule}} + M \sum_{tp} uv_{tp} \end{aligned} \quad (27)$$

## 4 LP Model

### 4.1 Full LP

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

$$\max_{\vec{v}} \sum \alpha(\vec{S})v(\vec{S}) \quad (28)$$

subject to

$$c(\vec{p}, \vec{p}', r\vec{s}, \vec{u}) + \gamma \sum_{\vec{p}} p(\vec{p}ea, \vec{p}wt, \vec{p}st, \vec{u}ed)v(\vec{S}'|\vec{S}, \vec{A}, \vec{p}ea, \vec{p}wt, \vec{p}st, \vec{u}ed) \geq v(\vec{S}) \quad \forall \vec{S}, \vec{A} \quad (29)$$

## 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:

$$\begin{aligned} v(\vec{u}\vec{e}, \vec{u}\vec{u}, \vec{p}\vec{w}, \vec{p}\vec{s}) = & \beta^0 + \sum_{tp} \beta_{tp}^{ue} ue_{tp} + \sum_{tp} \beta_{tp}^{uu} uu_{tp} + \\ & + \sum_{mdc} \beta_{mdc}^{pw} pw_{mdc} + \sum_{tmdc} \beta_{tmdc}^{ps} pst_{mdc} \end{aligned} \quad (30)$$

This gives the following LP

$$\max_{\vec{\beta}} \sum \alpha(\vec{u}\vec{e}, \vec{u}\vec{u}, \vec{p}\vec{w}, \vec{p}\vec{e}, \vec{p}\vec{s}) v(\vec{u}\vec{e}, \vec{u}\vec{u}, \vec{p}\vec{w}, \vec{p}\vec{e}, \vec{p}\vec{s}) \quad (31)$$

Subject to:

$$\begin{aligned} & c(\vec{p}\vec{w}, \vec{p}\vec{s}, r\vec{s}c, \vec{u}\vec{v}) + \gamma \sum_{\vec{p}} p \left( \beta^0 + \right. \\ & \left( \sum_p \beta_{1p}^{ue} (ue_{2p} + ued_p + ue_{1p} - \hat{u}u_{1p} + uv_{1p}) \right) + \left( \sum_{t=2,p}^{T-1} \beta_{tp}^{ue} ue_{t+1,p} \right) + \left( \sum_p \beta_{Tp}^{ue} uen_p \right) + \\ & \left( \sum_{tp}^{T-1} \beta_{tp}^{uu} (\hat{u}u_{t+1,p} + \sum_{mdc} pst_{tmdc} (U_{pd+1c} - U_{pdc})) \right) + \left( \sum_p \beta_{Tp}^{uu} * 0 \right) + \\ & \left( \sum_{dc} \beta_{0dc}^{pw} pea_{dc} \right) + \left( \sum_{m=1,dc}^{M-1} \beta_{mdc}^{pw} (\hat{p}w_{m-1dc} + pwt_{m-1d-1c} - pwt_{m-1dc}) \right) + \\ & \left( \sum_{dc} \beta_{Mdc}^{pw} \left( \sum_{M-1}^M \hat{p}w_{mdc} + pwt_{md-1c} - pwt_{mdc} \right) \right) + \\ & \left( \sum_{tdc} \beta_{t0dc}^{ps} 0 \right) + \left( \sum_{mdc} \beta_{Tmdc}^{ps} 0 \right) + \left( \sum_{t=1m=1dc}^{T-1M-1} \beta_{tmdc}^{ps} (\hat{p}s_{t+1m-1dc} + pst_{t+1m-1d-1c} - pst_{t+1m-1dc}) \right) + \\ & \left. \left( \sum_{t=1Mdc}^{T-1} \beta_{tmdc}^{ps} \left( \sum_{M-1}^M (\hat{p}s_{t+1mdc} + pst_{t+1md-1c} - pst_{t+1mdc}) \right) \right) \right) \geq v(\vec{u}\vec{e}, \vec{p}\vec{w}, \vec{p}\vec{s}) \quad \forall \vec{S} \vec{A} \\ & \vec{\beta} \geq 0 \quad \forall \beta \end{aligned} \quad (32)$$

$$\vec{\beta} \geq 0 \quad \forall \beta \quad (33)$$

And all auxiliary constraints in "Auxiliary Variable Definition"

### 5.2 Expectation ADP Model

Steps for rearranging and converting the constraint (32) into expectation:

- $E[ued_p] = 0$
- $E[pwt_{mdc}] = ptp_{mdc} * \hat{p}w_{mdc}$
- $E[pst_{tmdc}] = ptp_{mdc} * \hat{p}s_{tmdc}$

$$\beta^0 : \quad (1 - \gamma) \beta^0 \quad (34)$$

$$\beta_{tp}^{ue} : \quad \sum_p \beta_{1p}^{ue} (ue_{1p} - \gamma(uen_p + ue_{1p} - \hat{u}u_{1p} + uv_{1p})) + \sum_{t=2p}^T \beta_{tp}^{ue} (ue_{tp} - \gamma(uen_p)) \quad (35)$$

$$\beta_{tp}^{uu} : \quad \sum_{tp}^{T-1} \beta_{tp}^{uu} (uu_{tp} - \gamma(\hat{u}u_{t+1p} + \sum_{mdc} E[pst_{t+1mdc}] (U_{pd+1c} - U_{pdc}))) + \sum_p \beta_{Tp}^{uu} (uu_{Tp}) \quad (36)$$

$$\beta_{mdc}^{pw} : \sum_{dc} \beta_{0dc}^{pw} (pw_{0dc} - \gamma(pea_{dc})) + \sum_{m=1dc}^{M-1} \beta_{mdc}^{pw} (pw_{mdc} - \gamma(\hat{p}w_{m-1dc} + E[pwt_{m-1d-1c}] - E[pwt_{m-1dc}])) + \sum_{dc} \beta_{Mdc}^{pw} (pw_{Mdc} - \gamma \sum_{M-1}^M (\hat{p}w_{mdc} + E[pwt_{md-1c}] - E[pwt_{mdc}])) \quad (37)$$

$$\beta_{tmdc}^{ps} : \sum_{tdc} \beta_{t0dc}^{ps} (ps_{t0dc}) + \sum_{mdc} \beta_{Tmdc}^{ps} (ps_{Tmdc}) + \sum_{t=1m=1dc}^{T-1M-1} \beta_{tmdc}^{ps} (ps_{tmdc} - \gamma(\hat{p}s_{t+1m-1dc} + E[pst_{t+1m-1d-1c}] - E[pst_{t+1m-1dc}])) + \sum_{t=1dc}^{T-1} \beta_{tMdc}^{ps} (ps_{tMdc} - \gamma \sum_{M-1}^M (\hat{p}s_{t+1mdc} + E[pst_{t+1md-1c}] - E[pst_{t+1mdc}])) \quad (38)$$

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

$$\max_{\beta} \left( \sum \beta^0 + \sum \beta^{ue} E[ue] + \sum \beta^{uu} E[uu] + \sum \beta^{pw} E[pw] + \sum \beta^{ps} E[ps] \right) \quad (39)$$

Subject to:

$$E[V] \leq c(p\vec{w}, \vec{s}\vec{c}, r\vec{s}\vec{c}, u\vec{v}) \quad \forall \vec{S}\vec{A} \quad (40)$$

### 5.3 ADP Master Problem

Converting to Dual

$$\min_w \sum_{\vec{S}\vec{A}} w(\vec{S}\vec{A})c(\vec{S}\vec{A}) \quad (41)$$

Subject To:

$$\beta^0 : \sum_{\vec{S}\vec{A}} \vec{w}(1 - \gamma) = 1 \quad (42)$$

$$\beta^{ue} : \sum_{\vec{S}\vec{A}} \vec{w} (ue_{1p} - \gamma(uen_p + ue_{1p} - \hat{u}u_{1p} + uv_{1p})) \geq E[ue_{tp}] \quad \forall p \quad (43)$$

$$\sum_{\vec{S}\vec{A}} \vec{w} (ue_{tp} - \gamma(uen_p)) \geq E[ue_{tp}] \quad \forall t \in \{2 \dots T\}p \quad (44)$$

$$\beta^{uu} : \sum_{\vec{S}\vec{A}} \vec{w} (uu_{tp} - \gamma(\hat{u}u_{t+1p} + \sum_{mdc} E[pst_{t+1mdc}](U_{pd+1c} - U_{pdc}))) \geq E[uu_{tp}] \quad \forall t \in \{1 \dots T-1\}p \quad (45)$$

$$\sum_{\vec{S}\vec{A}} \vec{w} (uu_{Tp}) \geq E[uu_{Tp}] \quad \forall p \quad (46)$$

$$\beta^{pw} : \sum_{\vec{S}\vec{A}} \vec{w} (pw_{0dc} - \gamma(pea_{dc})) \geq E[ps_{0dc}] \quad \forall dc \quad (47)$$

$$\sum_{\vec{S}\vec{A}} \vec{w} (pw_{mdc} - \gamma(\hat{p}w_{m-1dc} + E[pwt_{m-1d-1c}] - E[pwt_{m-1dc}])) \geq E[pw_{mdc}] \quad \forall m \in \{1 \dots M-1\}dc \quad (48)$$

$$\sum_{\vec{S}\vec{A}} \vec{w} (pw_{Mdc} - \gamma \sum_{M-1}^M (\hat{p}w_{mdc} + E[pwt_{md-1c}] - E[pwt_{mdc}])) \geq E[pw_{Mdc}] \quad \forall dc \quad (49)$$

$$\beta^{ps} : \sum_{\vec{S}\vec{A}} \vec{w} (ps_{t0dc}) \geq E[ps_{t0dc}] \quad \forall tdc \quad (50)$$

$$\sum_{\vec{S}\vec{A}} \vec{w} (ps_{Tmdc}) \geq E[ps_{Tmdc}] \quad \forall mdc \quad (51)$$

$$\sum_{\vec{S}\vec{A}} \vec{w} (ps_{tmdc} - \gamma(\hat{p}s_{t+1m-1dc} + E[pst_{t+1m-1d-1c}] - E[pst_{tdc}])) \geq E[ps_{tmdc}] \quad \forall \{1 \dots T-1\}\{1 \dots M-1\}dc \quad (52)$$

$$\sum_{\vec{S}\vec{A}} \vec{w} \left( ps_{tMdc} - \gamma \sum_{M-1}^M (\hat{p}s_{t+1mdc} + E[ps_{t+1md-1c}] - E[ps_{td}]) \right) \geq E[ps_{tMdc}] \forall \{1 \dots T-1\}dc \quad (53)$$

$$\text{State Action Bounds:} \quad w \geq 0 \quad \forall w \quad (54)$$

## 5.4 ADP Pricing Problem

$$\min_{(\vec{u}\vec{e}, \vec{u}\vec{u}, \vec{p}\vec{w}, \vec{p}\vec{e}, \vec{p}\vec{s}) \in S, (\vec{s}\vec{c}, r\vec{s}\vec{c}) \in A} c(\vec{p}\vec{w}, \vec{s}\vec{c}, r\vec{s}\vec{c}, \vec{u}\vec{v}) - E[V] \quad (55)$$

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.

$$\begin{aligned} & \min_{\vec{A}} c(\vec{S}, \vec{A}) - \gamma v(S') \\ v(S') = & \left( \beta^0 + \left( \sum_p \beta_{1p}^{ue} (uen_d + ue_{1p} - \hat{u}u_{1p} + uv_{1p}) \right) + \left( \sum_{t=2,p}^T \beta_{tp}^{ue} uen_d \right) + \right. \\ & \left( \sum_{tp}^{T-1} \beta_{tp}^{uu} (\hat{u}u_{t+1,p} + \sum_{mdc} E[ps_{tmdc}](U_{pd+1c} - U_{pdc})) \right) + \\ & \left( \sum_{dc} \beta_{0dc}^{pw} pea_{dc} \right) + \left( \sum_{m=1,dc}^{M-1} \beta_{mdc}^{pw} (\hat{p}w_{m-1dc} + E[pwt_{m-1d-1c}] - E[pwt_{m-1dc}]) \right) + \\ & \left( \sum_{dc} \beta_{Mdc}^{pw} \left( \sum_{M-1}^M \hat{p}w_{mdc} + E[pwt_{md-1c}] - E[pwt_{mdc}] \right) \right) + \\ & \left( \sum_{t=1m=1dc}^{T-1M-1} \beta_{tmdc}^{ps} (\hat{p}s_{t+1m-1dc} + E[ps_{t+1m-1d-1c}] - E[ps_{t+1m-1dc}]) \right) + \\ & \left. \left( \sum_{t=1Mdc}^{T-1} \beta_{tmdc}^{ps} \left( \sum_{M-1}^M (\hat{p}s_{t+1mdc} + E[ps_{t+1md-1c}] - E[ps_{t+1mdc}]) \right) \right) \right) \end{aligned} \quad (56)$$

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