Online Supplements to "Dynamic Type Matching"

A. An alternative formulation of the 2×2 horizontal model

The greedy matching in round 1 allows us to collapse the state space. After round 1, type 1 (resp., type 2) demand and type 1 (resp., type 2) supply cannot be both available. In period t with the state $(\mathbf{x},\mathbf{y})=(x_1,x_2,y_1,y_2)$, we define the transformed state as $\mathbf{z}:=(z_1,z_2)$, where $z_1=x_1-y_1$ and $z_2=y_2-x_2$. The quantity z_1 describes the imbalance between type 1 demand and type 1 supply. A nonnegative z_1 represents the remaining quantity of type 1 demand after greedy matching with type 1 supply in period t (the remaining quantity of type 1 supply will be zero). For a negative value of z_1, $z_1^-=-z_1$ is the remaining quantity of type 1 supply after greedy matching with type 1 demand. Similarly, z_2^+ is the remaining quantity of type 2 supply after greedy matching with type 2 demand, whereas z_2^- is the remaining quantity of type 2 demand after greedy matching with type 2 supply. In the rest of this online appendix, unless otherwise specified, we use the word "state" to refer to the transformed state $\mathbf{z}$. We consider the following cases.

Case 1: $z_1\geq 0$ and $z_2\geq 0$. After round 1 matching, a quantity z_1 of type 1 demand is available to be matched with a quantity z_2 of type 2 supply. Let q be the matching quantity in round 2 between type 1 demand and type 2 supply. We have $0\leq q\leq \min\{z_1,z_2\}$. After round 2 matching, the remaining quantity of type 1 demand is z_1 and that of type 2 supply is z_2-q. The post-matching state is therefore (z_1-q,z_2-q).

Case 2: $z_1<0$ and $z_2<0$. After round 1 matching, a quantity $-z_1$ of type 1 supply is available to be matched with a quantity $-z_2$ of type 2 demand. Let $-q$ be the matching quantity in the round 2 between type 2 demand and type 1 supply. We have $0\leq -q\leq \min\{-z_1,-z_2\}$, or equivalently, $\max\{z_1,z_2\}\leq q\leq 0$. After round 2 matching, the remaining quantity of type 1 supply is $-z_1+q$ and that of type 2 demand is $-z_2+q$. In other words, the post-matching state is (z_1-q,z_2-q).

Case 3: $z_1z_2<0$. After round 1 matching, either there is only demand available or only supply available. The matching quantity in round 2 is $q=0$. The post-matching state is $(z_1-q,z_2-q)=(z_1,z_2)$ (it is identical to the pre-matching state since there is no matching in round 2).

In any of the above cases, the feasible space of matching decisions in round 2 of period t is:

$$M(\mathbf{z})=\{q\mid 0\leq q\leq \min(z_1,z_2)\ \text{ or }\ \max(z_1,z_2)\leq q\leq 0\ \text{or}\ q=0\}. \tag{A.1}$$

To reformulate the problem, we consider the total expected reward received from round 2 matching in period t to the end of period T.

In period t, the matching quantity between type 1 demand and type 2 supply is q^+, and the matching quantity between type 2 demand and type 1 supply is q^-. Thus, a total reward $r_{12}^t q^+ + r_{21}^t q^-$ is received in round 2 of period t.

Since the post-matching state in period t is $(z_1 - q, z_2 - q)$ after round 2, in the beginning of period $t+1$ the available type 1 demand is $\alpha(z_1 - q)^+ + D_1^{t+1}$, available type 2 demand is $\alpha(z_2 - q)^- + D_2^{t+1}$, available type 1 supply is $\beta(z_1 - q)^- + S_1^{t+1}$, and available type 2 supply is $\beta(z_2 - q)^+ + S_2^{t+1}$. In round 1 of period $t+1$, type 1 demand and type 1 supply will be matched greedily, and so will type 2 demand and type 2 supply. This results in the total expected reward $r_{11}^{t+1} E \min\left\{\alpha(z_1 - q)^+ + D_1^{t+1}, \beta(z_1 - q)^- + S_1^{t+1}\right\} + r_{22}^{t+1} E \min\left\{\alpha(z_2 - q)^- + D_2^{t+1}, \beta(z_2 - q)^+ + S_2^{t+1}\right\}$ in round 1 of period $t+1$. The state immediately prior to round 2 of period $t+1$ is $(\alpha(z_1 - q)^+ + D_1^{t+1} - \beta(z_1 - q)^- - S_1^{t+1}, \beta(z_2 - q)^+ + S_2^{t+1} - \alpha(z_2 - q)^- - D_2^{t+1})$.

Let us define $J_t(q, \mathbf{z})$ as the total expected reward received from round 2 of period t until the end of period T if the round 2 matching decision in period t is q. We also define $U_t(\mathbf{z})$ as the optimal total expected reward achievable (by using the optimal q) from round 2 of period t until the end of period T. We are now ready to present the reformulation.

$$
\begin{aligned}
U_t(\mathbf{z}) &= \max_{q \in M(\mathbf{z})} J_t(q, \mathbf{z}) && \text{(A.2)} \\
J_t(q, \mathbf{z}) &= r_{12}^t q^+ + r_{21}^t q^- + \gamma r_{11}^{t+1} E \min\left\{\alpha(z_1 - q)^+ + D_1^{t+1}, \beta(z_1 - q)^- + S_1^{t+1}\right\} \\
&\quad + \gamma r_{22}^{t+1} E \min\left\{\alpha(z_2 - q)^- + D_2^{t+1}, \beta(z_2 - q)^+ + S_2^{t+1}\right\} \\
&\quad + \gamma E U_{t+1}(\alpha(z_1 - q)^+ + D_1^{t+1} - \beta(z_1 - q)^- - S_1^{t+1}, \beta(z_2 - q)^+ + S_2^{t+1} - \alpha(z_2 - q)^- - D_2^{t+1}). && \text{(A.3)}
\end{aligned}
$$

We show the concavity of J_t in the following lemma (for the continuous-valued model).

Lemma A.1. *Consider the problem with continuous-valued state space and matching decisions. $U_t(\mathbf{z})$ is concave in any of the following regions: $\mathbf{z} \in \mathbb{R}_+^2$, $\mathbf{z} \in \mathbb{R}_+ \times \mathbb{R}_-$, $\mathbf{z} \in \mathbb{R}_- \times \mathbb{R}_+$ and $\mathbf{z} \in \mathbb{R}_-^2$ For any given state $\mathbf{z}$, $J_t(q, \mathbf{z})$ is concave in q within its feasible range defined in (A.1).*

Proof of Lemma A.1. We show that $U_t(\mathbf{z})$ is concave for $\mathbf{z} \in \mathbb{R}_+^2$, and its concavity in the other regions are similar. If the original (i.e., untransformed) state in the beginning of period t is given as $x_1 = z_1$, $x_2 = 0$, $y_1 = 0$ and $y_2 = z_2$, the matching quantity in round 1 is zero since there is no type 2 demand or type 1 supply available. By definition, we have $U_t(\mathbf{z}) = V_t(z_1, 0, 0, z_2)$. One can readily show that V_t is concave for the problem with continuous-valued states and decisions. It follows that $U_t(\mathbf{z})$ is concave in $\mathbf{z} \in \mathbb{R}_+^2$.

Next, we show that $J_t(q,\mathbf{z})$ is concave in q for given $\mathbf{z}$. When $z_1z_2<0$, q can only be zero, and thus the result holds trivially. Let us prove that $J_t(q,\mathbf{z})$ is concave in q when $z_1\geq 0$ and $z_2\geq 0$, and the remaining case with $z_1<0$ and $z_2<0$ follows by symmetry.

Following our earlier discussions in this supplementary, after round-2 matching in period t, the available type 1 demand is z_1-q, the available type 2 supply is z_2-q, and there is no available type 2 demand or type 1 supply. Therefore, the (untransformed) post-matching levels in period $t+1$ are $(u_1,u_2,v_1,v_2)=(z_1-q,0,0,z_2-q)$. The sum of the last three terms in (A.3) represents the expected total discounted reward from the beginning of period $t+1$ to the end of period T. Therefore, $J_t(q,\mathbf{z})=r_{12}^tq+\gamma EV_{t+1}(\alpha(z_1-q)+D_1^{t+1},0,0,\beta(z_2-q)+S_2^{t+1})$. Since V_{t+1} is concave, $J_t(q,\mathbf{z})$ is concave in q. $\square$

We now prove Proposition 1 for the continuous-valued model based on the reformulation (A.2)–(A.2), and defer the proof for the discrete-valued model to Online Supplements A.1 (for the case with $\alpha=\beta=1$) and A.2 (for the case with $\alpha=0$ and $\beta=1$).

Proof of Proposition 1 (Continuous-valued model). We focus on the matching in round 2, and only consider the case with $z_1\geq 0$ and $z_1\geq 0$ (the case with $z_1<0$ and $z_2<0$ is symmetric).

Using the reformulation (A.2)–(A.3), the optimal matching quantity solves $\max_{q\in M(\mathbf{z})} J_t(q,\mathbf{z})$. Let us use $p_d:=z_1-q$ and $p_s=z_2-q$ as decision variables in place of q. Then, $p_d=p_s+z_1-z_2=p_s+IB$. Since both p_d and p_s need to be nonnegative, the feasible range of p_s is $IB^-\leq p_s\leq z_2$.

We rewrite $J_t(q,\mathbf{z})$ as a function of p_s, by substituting $q=z_2-p_s$ in (A.3). Given that $0\leq q\leq \min\{z_1,z_2\}$, we have

$$\begin{aligned} J_t(q,\mathbf{z}) =& r_{12}^t(z_2-p_s)+r_{11}^{t+1}E\min\left\{\alpha(p_s+IB)+D_1^{t+1},S_1^{t+1}\right\}+r_{22}^{t+1}E\min\left\{D_2^{t+1},\beta p_s+S_2^{t+1}\right\}\\ &+EU_{t+1}(\alpha(p_s+IB)+D_1^{t+1}-S_1^{t+1},\beta p_s+S_2^{t+1}-D_2^{t+1}), \end{aligned} \tag{A.4}$$

which depends on IB, p_s and also linearly on z_2. The sum of the last three terms in (A.4) represents the expected total discounted reward from period $t+1$ to period T, which is equal to $\gamma EV_{t+1}(\alpha(p_s+IB)+D_1^{t+1},D_2^{t+1},S_1^{t+1},\beta p_s+S_2^{t+1})$ Thus, $J_t(q,\mathbf{z})=r_{12}^t(z_2-p_s)+\gamma EV_{t+1}(\alpha(p_s+IB)+D_1^{t+1},D_2^{t+1},S_1^{t+1},\beta p_s+S_2^{t+1})$. We can write $J_t(q,\mathbf{z})=r_{12}^tz_2+\check{J}_t(p_s,IB)$, where $\check{J}_t(p_s,IB):=-r_{12}^tp_s+\gamma EV_{t+1}(\alpha(p_s+IB)+D_1^{t+1},D_2^{t+1},S_1^{t+1},\beta p_s+S_2^{t+1})$.

It is easy to see that $\check{J}_t$ is concave in p_s (the argument is similar to the proof of concavity of $J_t(q,\mathbf{z})$ with respect to q in Lemma A.1). Let $p_{s_2}(IB)\in\arg\max_{p_s\geq IB^-}\check{J}_t(p_s,IB)$. Since $IB^-\leq p_s\leq z_2$, the optimal decision in terms of p_s is $p_s^*=\min\left\{z_2,p_{s_2}^t(IB)\right\}$. Thus, the optimal matching quantity between type 1 demand and type 2 supply is $q_{12}^{t*}=z_2-p_s^*=z_2-\min\left\{z_2,p_{s_2}^t(IB)\right\}=[z_2-p_{s_2}^t(IB)]^+=[y_2-x_2-p_{s_2}^t(IB)]^+$. $\square$

A.1. The case with equal carry-over rates

We now consider the case $\alpha = \beta$. To begin with, we transform the Bellman equations (A.2)–(A.3). Let us define $\tilde{U}_t(\mathbf{z}) := -r_{11}^t z_1^+ - r_{22}^t z_2^+ + U_t(\mathbf{z})$ and $\tilde{J}_t(q,\mathbf{z}) := -r_{11}^t z_1^+ - r_{22}^t z_2^+ + J_t(q,\mathbf{z})$. In the following lemma, we rewrite the equations (A.2)–(A.3).

LEMMA A.2. *Suppose $\alpha = \beta$. Equations (A.2)–(A.3) are equivalent to the following equations:*

$$\tilde{U}_t(\mathbf{z}) = \max_{q \in M(\mathbf{z})} \tilde{J}_t(q,\mathbf{z}) \tag{A.5}$$

$$\begin{aligned}\tilde{J}_t(q,\mathbf{z}) =& \gamma r_{11}^{t+1} E D_1^{t+1} + \gamma r_{22}^{t+1} E S_2^{t+1} - (r_{11}^t + r_{22}^t - r_{12}^t - r_{21}^t) q^+ - r_{21}^t q \\ & - (r_{11}^t - \gamma\alpha r_{11}^{t+1})(z_1 - q)^+ - (r_{22}^t - \gamma\alpha r_{22}^{t+1})(z_2 - q)^+ \\ & + \gamma E \tilde{U}_{t+1}(\alpha(z_1 - q) + D_1^{t+1} - S_1^{t+1}, \alpha(z_2 - q) + S_2^{t+1} - D_2^{t+1}).\end{aligned} \tag{A.6}$$

Proof of Lemma A.2. Applying the equality $\min\{a,b\} = a - (a-b)^+$, we can rewrite J_t as:

$$\begin{aligned}J_t(q,\mathbf{z}) =& r_{12}^t q^+ + r_{21}^t q^- + \gamma r_{11}^{t+1} E D_1^{t+1} + \gamma r_{22}^{t+1} E S_2^{t+1} \\ & + \gamma\alpha r_{11}^{t+1} E(z_1 - q)^+ - \gamma r_{11}^{t+1} E[\alpha(z_1 - q) + D_1^{t+1} - S_1^{t+1}]^+ \\ & + \gamma\alpha r_{22}^{t+1} E(z_2 - q)^+ - \gamma r_{22}^{t+1} E[\alpha(z_2 - q) + S_2^{t+1} - D_2^{t+1}]^+ \\ & + \gamma E U_{t+1}(\alpha(z_1 - q) + D_1^{t+1} - S_1^{t+1}, \alpha(z_2 - q) + S_2^{t+1} - D_2^{t+1}).\end{aligned}$$

Then, by the definition of $\tilde{J}_t(q,\mathbf{z})$, we have

$$\begin{aligned}\tilde{J}_t(q,\mathbf{z}) =& \gamma r_{11}^{t+1} E D_1^{t+1} + \gamma r_{22}^{t+1} E S_2^{t+1} - r_{11}^t z_1^+ - r_{22}^t z_2^+ + (r_{12}^t + r_{21}^t) q^+ - r_{21}^t q \\ & + \gamma\alpha r_{11}^{t+1} E(z_1 - q)^+ + \gamma\alpha r_{22}^{t+1} E(z_2 - q)^+ \\ & + \gamma E \tilde{U}_{t+1}(\alpha(z_1 - q) + D_1^{t+1} - S_1^{t+1}, \alpha(z_2 - q) + S_2^{t+1} - D_2^{t+1}).\end{aligned}$$

For $q \in M(\mathbf{z})$, we can verify that $z_1^+ = q^+ + (z_1 - q)^+$ and $z_2^+ = q^+ + (z_2 - q)^+$. By substituting z_1^+ and z_2^+ by $q^+ + (z_1 - q)^+$ and $q^+ + (z_2 - q)^+$ respectively, we have

$$\begin{aligned}\tilde{J}_t(q,\mathbf{z}) =& \gamma r_{11}^{t+1} E D_1^{t+1} + \gamma r_{22}^{t+1} E S_2^{t+1} - r_{11}^t [q^+ + (z_1 - q)^+] - r_{22}^t [q^+ + (z_2 - q)^+] + (r_{12}^t + r_{21}^t) q^+ - r_{21}^t q \\ & + \gamma\alpha r_{11}^{t+1} E(z_1 - q)^+ + \gamma\alpha r_{22}^{t+1} E(z_2 - q)^+ \\ & + \gamma E \tilde{U}_{t+1}(\alpha(z_1 - q) + D_1^{t+1} - S_1^{t+1}, \alpha(z_2 - q) + S_2^{t+1} - D_2^{t+1}) \\ =& \gamma r_{11}^{t+1} E D_1^{t+1} + \gamma r_{22}^{t+1} E S_2^{t+1} - (r_{11}^t + r_{22}^t - r_{12}^t - r_{21}^t) q^+ - r_{21}^t q \\ & - (r_{11}^t - \gamma\alpha r_{11}^{t+1})(z_1 - q)^+ - (r_{22}^t - \gamma\alpha r_{22}^{t+1})(z_2 - q)^+\end{aligned}$$

$$+\gamma E\tilde{U}_{t+1}(\alpha(z_1-q)+D_1^{t+1}-S_1^{t+1},\alpha(z_2-q)+S_2^{t+1}-D_2^{t+1}).$$

This completes the proof. □

We will show that both functions $\tilde{J}_t$ and $\tilde{U}_t$ are $L^\natural$-concave. To that end, we first present a lemma, which explores the properties of the transformed value function $\tilde{U}_t$.

LEMMA A.3. *Suppose that Assumption 2 holds. For any transformed state* $\mathbf{z}$ *in period* t *and any* $\varepsilon>0$, *we have* $\tilde{U}_t(\mathbf{z}+\varepsilon\mathbf{1}^2)-\tilde{U}_t(\mathbf{z})\geq-(r_{22}^t-r_{12}^t+r_{11}^t)\varepsilon$.

Proof of Lemma A.3. We prove the lemma by induction. The lemma clearly holds for $t=T+1$ since $\tilde{U}_{T+1}(\mathbf{z})\equiv 0$. Suppose that it holds for period $t+1$. To show that the equality $\tilde{U}_t(\mathbf{z}+\varepsilon\mathbf{1}^2)-\tilde{U}_t(\mathbf{z})\geq-(r_{22}^t-r_{12}^t+r_{11}^t)\varepsilon$ holds in period t, we consider sufficiently small $\varepsilon>0$ such that, if $z_i<0$ $(i=0,1)$ then $\varepsilon<|z_i|$, without loss of generality. (Note that if the inequality holds for any sufficiently small $\varepsilon>0$, then $\tilde{U}_t(\mathbf{z}+K\varepsilon\mathbf{1}^2)-\tilde{U}_t(\mathbf{z})=\sum_{k=1}^K\left[\tilde{U}_t(\mathbf{z}+k\varepsilon\mathbf{1}^2)-\tilde{U}_t(\mathbf{z}+(k-1)\varepsilon\mathbf{1}^2)\right]\geq-\sum_{k=1}^K(r_{22}^t-r_{12}^t+r_{11}^t)\varepsilon=-(r_{22}^t-r_{12}^t+r_{11}^t)K\varepsilon$ for any positive integer K; thus the result will also hold for any $\varepsilon>0$.)

Let us denote by $\hat{q}\in\arg\max_{q\in M(\mathbf{z})}\tilde{J}_t(q,\mathbf{z})$ the optimal matching quantity in round-2 matching of period t, given the transformed state $\mathbf{z}$. We discuss four cases.

Case 1: $z_1\geq 0$ and $z_2\geq 0$.

It is easy to see that $\hat{q}+\varepsilon$ is a feasible matching quantity between type 1 demand and type 2 supply under the state $(z_1+\varepsilon,z_2+\varepsilon)$, for any $\varepsilon>0$.

Thus, for any $\varepsilon>0$,

$$\tilde{U}_t(z_1+\varepsilon,z_2+\varepsilon)-\tilde{U}_t(z_1,z_2)\geq\tilde{J}_t(\hat{q}+\varepsilon,z_1+\varepsilon,z_2+\varepsilon)-\tilde{J}_t(\hat{q},z_1,z_2)=(-r_{11}^t-r_{22}^t+r_{12}^t)\varepsilon,$$

which is equivalent to $\tilde{U}_t(\mathbf{z}+\varepsilon\mathbf{1}^2)-\tilde{U}_t(\mathbf{z})\geq-(r_{22}^t-r_{12}^t+r_{11}^t)\varepsilon$.

Case 2: $z_1<0$ and $z_2<0$.

Let $\varepsilon'=\min\{-\hat{q},\varepsilon\}$. It is easy to see that $\hat{q}+\varepsilon'$ is a feasible decision under the state $\mathbf{z}+\varepsilon\mathbf{1}^2=(z_1+\varepsilon,z_2+\varepsilon)$. Then,

$$\begin{aligned}
&\tilde{U}_t(\mathbf{z}+\varepsilon\mathbf{1}^2)-\tilde{U}_t(\mathbf{z})\\
\geq&\tilde{J}_t(\hat{q}+\varepsilon',\mathbf{z}+\varepsilon\mathbf{1})-\tilde{J}_t(\hat{q},\mathbf{z})\\
=&-r_{21}^t\varepsilon'+\gamma E\tilde{U}_{t+1}(\alpha(z_1-\hat{q})+\alpha(\varepsilon-\varepsilon')+D_1^{t+1}-S_1^{t+1},\alpha(z_2-\hat{q})+\alpha(\varepsilon-\varepsilon')+D_2^{t+1}-S_2^{t+1})\\
&-\gamma E\tilde{U}_{t+1}(\alpha(z_1-\hat{q})+D_1^{t+1}-S_1^{t+1},\alpha(z_2-\hat{q})+D_2^{t+1}-S_2^{t+1})
\end{aligned}$$

$$\geq -r_{21}^t\varepsilon' - \gamma\alpha(r_{22}^{t+1} - r_{12}^{t+1} + r_{11}^{t+1})(\varepsilon - \varepsilon'),$$

where the last inequality follows from the induction hypothesis. It follows that,

$$\begin{aligned}
&(r_{22}^t - r_{12}^t + r_{11}^t)\varepsilon + \tilde{U}_t(\mathbf{z} + \varepsilon\mathbf{1}^2) - \tilde{U}_t(\mathbf{z}) \\
\geq &(r_{22}^t - r_{12}^t + r_{11}^t)\varepsilon - r_{21}^t\varepsilon' - \gamma\alpha(r_{22}^{t+1} - r_{12}^{t+1} + r_{11}^{t+1})(\varepsilon - \varepsilon') \\
= &(r_{22}^t - r_{12}^t + r_{11}^t - r_{21}^t)\varepsilon' + \left[(r_{22}^t - r_{12}^t) - \gamma\alpha(r_{22}^{t+1} - r_{12}^{t+1}) + r_{11}^t - \gamma\alpha r_{11}^{t+1}\right](\varepsilon - \varepsilon') \geq 0,
\end{aligned}$$

which implies the desired result. (The last inequality above follows from Assumption 2.)

Case 3: $z_1 \geq 0$ and $z_2 < 0$.

We have

$$\begin{aligned}
&\tilde{U}_t(\mathbf{z} + \varepsilon\mathbf{1}) - \tilde{U}_t(\mathbf{z}) \\
= &\tilde{J}_t(0, \mathbf{z} + \varepsilon\mathbf{1}) - \tilde{J}_t(0, \mathbf{z}) \\
= &-(r_{11}^t - \gamma\alpha r_{11}^{t+1})\varepsilon \\
&+ \gamma E\tilde{U}_{t+1}(\alpha z_1 + \alpha\varepsilon + D_1^{t+1} - S_1^{t+1}, \alpha z_2 + \alpha\varepsilon + D_2^{t+1} - S_2^{t+1}) - E\tilde{U}_{t+1}(z_1 + D_1^{t+1} - S_1^{t+1}, z_2 + D_2^{t+1} - S_2^{t+1}) \\
\geq &-(r_{11}^t - \gamma\alpha r_{11}^{t+1})\varepsilon - \gamma\alpha(r_{22}^{t+1} - r_{12}^{t+1} + r_{11}^{t+1})\varepsilon,
\end{aligned}$$

where the last inequality follows again from the induction hypothesis. It follows that

$$\begin{aligned}
&(r_{22}^t - r_{12}^t + r_{11}^t)\varepsilon + \tilde{U}_t(\mathbf{z} + \varepsilon\mathbf{1}^2) - \tilde{U}_t(\mathbf{z}) \\
\geq &(r_{22}^t - r_{12}^t + r_{11}^t)\varepsilon - (r_{11}^t - \gamma\alpha r_{11}^{t+1})\varepsilon - \gamma\alpha(r_{22}^{t+1} - r_{12}^{t+1} + r_{11}^{t+1})\varepsilon \\
= &[(r_{22}^t - r_{12}^t) - \gamma\alpha(r_{22}^{t+1} - r_{12}^{t+1})]\varepsilon \geq 0,
\end{aligned}$$

where the last inequality follows from Assumption 2.

Case 4: $z_1 < 0$ and $z_2 \geq 0$.

We have

$$\begin{aligned}
&\tilde{U}_t(\mathbf{z} + \varepsilon\mathbf{1}) - \tilde{U}_t(\mathbf{z}) \\
= &\tilde{J}_t(0, \mathbf{z} + \varepsilon\mathbf{1}) - \tilde{J}_t(0, \mathbf{z}) \\
= &-(r_{22}^t - \gamma\alpha r_{22}^{t+1})\varepsilon + \gamma E\tilde{U}_{t+1}(\alpha z_1 + \alpha\varepsilon + D_1^{t+1} - S_1^{t+1}, \alpha z_2 + \alpha\varepsilon + D_2^{t+1} - S_2^{t+1}) \\
&- \gamma E\tilde{U}_{t+1}(\alpha z_1 + D_1^{t+1} - S_1^{t+1}, \alpha z_2 + D_2^{t+1} - S_2^{t+1})
\end{aligned}$$

$$= -(r_{22}^t - \gamma\alpha r_{22}^{t+1})\varepsilon - \gamma\alpha(r_{22}^{t+1} - r_{12}^{t+1} + r_{11}^{t+1})\varepsilon = -[r_{22}^t + \gamma\alpha(r_{11}^{t+1} - r_{12}^{t+1})]\varepsilon.$$

It follows that

$$\begin{aligned}&(r_{22}^t - r_{12}^t + r_{11}^t)\varepsilon + \tilde{U}_{t+1}(\mathbf{z} + \varepsilon\mathbf{1}^2) - \tilde{U}_{t+1}(\mathbf{z})\\ \geq&(r_{22}^t - r_{12}^t + r_{11}^t)\varepsilon - [r_{22}^t + \gamma\alpha(r_{11}^{t+1} - r_{12}^{t+1})]\varepsilon = (r_{11}^t - r_{12}^t)\varepsilon - \gamma\alpha(r_{11}^{t+1} - r_{12}^{t+1})\varepsilon \geq 0,\end{aligned}$$

where the last inequality holds because of Assumption 2.

Combining the four cases completes the induction. □

LEMMA A.4. *Suppose that Assumption 2 holds. The function $\tilde{J}_t(q, \mathbf{z})$ is $L^\natural$-concave in $(q, \mathbf{z})$ and $\tilde{U}_t(\mathbf{z})$ is $L^\natural$-concave in $\mathbf{z}$.*

Proof of Lemma A.4. We prove the lemma by induction. $\tilde{U}_T(\mathbf{z}) \equiv 0$ is $L^\natural$-concave. Suppose that $\tilde{U}_{t+1}(\mathbf{z})$ is $L^\natural$-concave. To show that $\tilde{U}_t(\mathbf{z})$ is $L^\natural$-concave, it suffices to prove that $\tilde{U}_t(\mathbf{z} - \eta\mathbf{1}^2)$ is supermodular in $(\eta, \mathbf{z})$.

Given the conditions $r_{11}^t \geq \gamma\alpha r_{11}^{t+1}$ and $r_{22}^t \geq \gamma\alpha r_{22}^{t+1}$, the induction hypothesis and the concavity of $-(\cdot)^+$, it is easy to see that $\tilde{J}_t(q - \eta, \mathbf{z} - \eta\mathbf{1}^2)$ is supermodular in $(\eta, q, \mathbf{z})$. This implies that $\tilde{J}_t(q, \mathbf{z})$ is $L^\natural$-concave in $(q, \mathbf{z})$, and thus it is also supermodular in $(q, \mathbf{z})$. Since the set $\{(q, \mathbf{z}) \mid q \in M(\mathbf{z})\}$ is a lattice, $\tilde{U}_t(\mathbf{z}) = \max_{q \in M(\mathbf{z})} \tilde{J}_t(q, \mathbf{z})$ is supermodular. As a result, the function $\tilde{U}_t(\mathbf{z} - \eta\mathbf{1}^2)$ is supermodular in $\mathbf{z}$. To show that $\tilde{U}_t$ is $L^\natural$-concave, it suffices to show that $\tilde{U}(\mathbf{z} - \eta\mathbf{1}^2)$ has increasing difference in (η, z_1) and in (η, z_2). In the followings, we show that this is true within four regions (i.e., $\mathbf{z} - \eta\mathbf{1}^2 \in \mathbb{R}_+^2 := \{\mathbf{z} \mid z_1 \geq 0, z_2 \geq 0\}$, $\mathbf{z} - \eta\mathbf{1}^2 \in \mathbb{R}_-^2 := \{\mathbf{z} \mid z_1 \leq 0, z_2 \leq 0\}$, $\mathbf{z} - \eta\mathbf{1}^2 \in \mathbb{R}_+ \times \mathbb{R}_- := \{\mathbf{z} \mid z_1 \geq 0, z_2 \leq 0\}$ and $\mathbf{z} - \eta\mathbf{1}^2 \in \mathbb{R}_- \times \mathbb{R}_+ := \{\mathbf{z} \mid z_1 \leq 0, z_2 \geq 0\}$), as well as across the four regions.

For $\mathbf{z} - \eta\mathbf{1}^2 \in \mathbb{R}_+^2$, we have $\tilde{U}_t(\mathbf{z} - \eta\mathbf{1}^2) = \max_{q \in M(\mathbf{z} - \eta\mathbf{1}^2)} \tilde{J}_t(q, \mathbf{z} - \eta\mathbf{1}^2) = \max_{q - \eta \in M(\mathbf{z} - \eta\mathbf{1}^2)} \tilde{J}_t(q - \eta, \mathbf{z} - \eta\mathbf{1}^2)$. The feasible set of the maximization problem (on the RHS of the last equality) is $\{(q, \eta, \mathbf{z}) \mid \mathbf{z} - \eta\mathbf{1}^2 \geq 0, q - \eta \in M(\mathbf{z} - \eta\mathbf{1}^2)\} = \{(q, \eta, \mathbf{z}) \mid \eta \leq q \leq \min\{z_1, z_2\}\}$, which is a lattice. Because $\tilde{J}_t(q - \eta, \mathbf{z} - \eta\mathbf{1}^2)$ is supermodular in $(\eta, q, \mathbf{z})$, the function $\tilde{U}_t(\mathbf{z} - \eta\mathbf{1}^2)$ is supermodular in $(\eta, \mathbf{z})$ for $\eta \leq \min\{z_1, z_2\}$. This implies that $\tilde{U}_t(\mathbf{z} - \eta\mathbf{1})$ has increasing differences in (η, z_1) and in (η, z_2) for $\eta \leq \min\{z_1, z_2\}$.

Similarly, for $\mathbf{z} - \eta\mathbf{1}^2 \in \mathbb{R}_-^2$, we have $\tilde{U}_t(\mathbf{z} - \eta\mathbf{1}^2) = \max_{q - \eta \in M(\mathbf{z} - \eta\mathbf{1}^2)} \tilde{J}_t(q - \eta, \mathbf{z} - \eta\mathbf{1}) = \max_{\max\{z_1, z_2\} \leq q \leq \eta} \tilde{J}_t(q - \eta, \mathbf{z} - \eta\mathbf{1})$ is supermodular in $(\eta, \mathbf{z})$ for $\eta \geq \max\{z_1, z_2\}$ because $\{(q, \eta, \mathbf{z}) \mid \max\{z_1, z_2\} \leq q \leq \eta\}$ is a lattice. Thus, $\tilde{U}_t(\mathbf{z} - \eta\mathbf{1}^2)$ has increasing differences in (η, z_1) and in (η, z_2) for $\eta \leq \max\{z_1, z_2\}$.

For $\mathbf{z}-\eta\mathbf{1}^2 \in \mathbb{R}_+ \times \mathbb{R}_-$ or $\mathbf{z}-\eta\mathbf{1}^2 \in \mathbb{R}_- \times \mathbb{R}_+$, we have $M(\mathbf{z}-\eta\mathbf{1}^2) = \{0\}$. Based on the expression in (A.6), it is easy to verify that $\tilde{U}_t(\mathbf{z}-\eta\mathbf{1}) = \tilde{J}_t(0, \mathbf{z}-\eta\mathbf{1})$ is supermodular in $(\eta, \mathbf{z})$ for $(\eta, \mathbf{z}) \in \mathbb{R}_+ \times \mathbb{R}_-$ and for $(\eta, \mathbf{z}) \in \mathbb{R}_- \times \mathbb{R}_+$.

It remains to show that $\tilde{U}_t(\mathbf{z}-\eta\mathbf{1}^2)$ has increasing differences in (η, z_1) and in (η, z_2) across the 4 regions. In the followings, we focus on the difference $\tilde{U}_t(\mathbf{z}+\varepsilon\mathbf{1}^2) - \tilde{U}_t(\mathbf{z})$ across the boundary between $\mathbf{z}-\eta\mathbf{1}^2 \in \mathbb{R}_+^2$ and $\mathbf{z}-\eta\mathbf{1}^2 \in \mathbb{R}_+ \times \mathbb{R}_-$. The same property across the other boundaries can be proved similarly. More specifically, we will prove the following inequality holds for sufficiently small $\varepsilon > 0$.

$$\tilde{U}_t(z_1, 0) - \tilde{U}_t(z_1-\varepsilon, -\varepsilon) \geq \tilde{U}_t(z_1, \varepsilon) - \tilde{U}_t(z_1-\varepsilon, 0),$$

which implies that $\tilde{U}_t(\mathbf{z}-\eta\mathbf{1}^2)$ has increasing differences in (η, z_2) across the boundary between $\mathbf{z}-\eta\mathbf{1}^2 \in \mathbb{R}_+^2$ and $\mathbf{z}-\eta\mathbf{1}^2 \in \mathbb{R}_+ \times \mathbb{R}_-$. (The increasing difference property with respect to (η, z_1) can be proved similarly.)

Let $z_1 > 0$ and $\hat{q} \in \arg\max_{q \in M(z_1,\varepsilon)} \tilde{J}_t(q, z_1, \varepsilon)$. Also, let $\varepsilon' = \varepsilon - \hat{q}$. Then, by using the expression of $\tilde{J}_t$ given in (A.6), we have

$$\begin{aligned}
&\tilde{U}_t(z_1, \varepsilon) - \tilde{U}_t(z_1-\varepsilon, 0)\\
=&\tilde{J}_t(\hat{q}, z_1, \varepsilon) - \tilde{J}_t(0, z_1-\varepsilon, 0)\\
=&-(r_{11}^t + r_{22}^t - r_{12}^t - r_{21}^t)\hat{q} - r_{21}^t\hat{q} - (r_{11}^t - \gamma\alpha r_{11}^{t+1})(z_1-\hat{q}) - (r_{22}^t - \gamma\alpha r_{22}^{t+1})(\varepsilon-\hat{q})\\
&+\gamma E\tilde{U}_{t+1}(\alpha(z_1-\hat{q}) + D_1^{t+1} - S_1^{t+1}, \alpha(\varepsilon-\hat{q}) + S_2^{t+1} - D_2^{t+1})\\
&-[-(r_{11}^t - \gamma\alpha r_{11}^{t+1})(z_1-\varepsilon) + E\tilde{U}_{t+1}(\alpha(z_1-\varepsilon) + D_1^{t+1} - S_1^{t+1}, S_2^{t+1} - D_2^{t+1})]\\
=&-r_{11}^t\varepsilon - r_{22}^t\varepsilon + r_{12}^t\hat{q} + \gamma\alpha r_{11}^{t+1}(\varepsilon-\hat{q}) + \gamma\alpha r_{22}^{t+1}(\varepsilon-\hat{q})\\
&+E\tilde{U}_{t+1}(\alpha(z_1-\hat{q}) + D_1^{t+1} - S_1^{t+1}, \alpha(\varepsilon-\hat{q}) + S_2^{t+1} - D_2^{t+1}) - E\tilde{U}_{t+1}(\alpha(z_1-\varepsilon) + D_1^{t+1} - S_1^{t+1}, S_2^{t+1} - D_2^{t+1})\\
=&-r_{11}^t\varepsilon - r_{22}^t\varepsilon + r_{12}^t(\varepsilon-\varepsilon') + \gamma\alpha r_{11}^{t+1}\varepsilon' + \gamma\alpha r_{22}^{t+1}\varepsilon'\\
&+E\tilde{U}_{t+1}(\alpha(z_1-\varepsilon+\varepsilon') + D_1^{t+1} - S_1^{t+1}, \alpha\varepsilon' + S_2^{t+1} - D_2^{t+1}) - E\tilde{U}_{t+1}(\alpha(z_1-\varepsilon) + D_1^{t+1} - S_1^{t+1}, S_2^{t+1} - D_2^{t+1})
\end{aligned}$$

Also, for $z_1 > 0$,

$$\begin{aligned}
&\tilde{U}_t(z_1, 0) - \tilde{U}_t(z_1-\varepsilon, -\varepsilon)\\
=&\tilde{J}_t(0, z_1, 0) - \tilde{J}_t(0, z_1-\varepsilon, -\varepsilon)\\
=&-(r_{11}^t - \gamma\alpha r_{11}^{t+1})z_1 + \gamma E\tilde{U}_{t+1}(\alpha z_1 + D_1^{t+1} - S_1^{t+1}, S_2^{t+1} - D_2^{t+1})
\end{aligned}$$

$$
\begin{aligned}
&-\left[-(r_{11}^t-\gamma\alpha r_{11}^{t+1})(z_1-\varepsilon)+\gamma E\tilde{U}_{t+1}(\alpha(z_1-\varepsilon)+D_1^{t+1}-S_1^{t+1},-\alpha\varepsilon+S_2^{t+1}-D_2^{t+1})\right]\\
=&-(r_{11}^t-\gamma\alpha r_{11}^{t+1})\varepsilon+\gamma E\tilde{U}_{t+1}(\alpha z_1+D_1^{t+1}-S_1^{t+1},S_2^{t+1}-D_2^{t+1})\\
&-\gamma E\tilde{U}_{t+1}(\alpha(z_1-\varepsilon)+D_1^{t+1}-S_1^{t+1},-\alpha\varepsilon+S_2^{t+1}-D_2^{t+1})\\
=&-(r_{11}^t-\gamma\alpha r_{11}^{t+1})\varepsilon\\
&+\gamma E\tilde{U}_{t+1}(\alpha z_1+D_1^{t+1}-S_1^{t+1},S_2^{t+1}-D_2^{t+1})-\gamma E\tilde{U}_{t+1}(\alpha z_1-\alpha(\varepsilon-\varepsilon')+D_1^{t+1}-S_1^{t+1},-\alpha(\varepsilon-\varepsilon')+S_2^{t+1}-D_2^{t+1})\\
&+\gamma E\tilde{U}_{t+1}(\alpha z_1-\alpha(\varepsilon-\varepsilon')+D_1^{t+1}-S_1^{t+1},-\alpha(\varepsilon-\varepsilon')+S_2^{t+1}-D_2^{t+1})\\
&-\gamma E\tilde{U}_{t+1}(\alpha(z_1-\varepsilon)+D_1^{t+1}-S_1^{t+1},-\alpha\varepsilon+S_2^{t+1}-D_2^{t+1})\\
\geq&-(r_{11}^t-\gamma\alpha r_{11}^{t+1})\varepsilon\\
&+\gamma E\tilde{U}_{t+1}(\alpha z_1+D_1^{t+1}-S_1^{t+1},S_2^{t+1}-D_2^{t+1})-\gamma E\tilde{U}_{t+1}(\alpha z_1-\alpha(\varepsilon-\varepsilon')+D_1^{t+1}-S_1^{t+1},-\alpha(\varepsilon-\varepsilon')+S_2^{t+1}-D_2^{t+1})\\
&+\gamma E\tilde{U}_{t+1}(\alpha z_1-\alpha(\varepsilon-\varepsilon')+D_1^{t+1}-S_1^{t+1},\alpha\varepsilon'+S_2^{t+1}-D_2^{t+1})\\
&-\gamma E\tilde{U}_{t+1}(\alpha(z_1-\varepsilon)+D_1^{t+1}-S_1^{t+1},S_2^{t+1}-D_2^{t+1}),
\end{aligned}
$$

where we obtain the inequality by the increasing difference property of $\tilde{U}_{t+1}$ by induction.

It then follows that

$$
\begin{aligned}
&[\tilde{U}(z_1,0)-\tilde{U}(z_1-\varepsilon,-\varepsilon)]-[\tilde{U}_t(z_1,\varepsilon)-\tilde{U}_t(z_1-\varepsilon,0)]\\
\geq&\gamma\alpha r_{11}^{t+1}(\varepsilon-\varepsilon')+r_{22}^t\varepsilon-r_{12}^t(\varepsilon-\varepsilon')-\gamma\alpha r_{22}^{t+1}\varepsilon'\\
&+\gamma E\tilde{U}_{t+1}(\alpha z_1+D_1^{t+1}-S_1^{t+1},S_2^{t+1}-D_2^{t+1})-\gamma E\tilde{U}_{t+1}(\alpha z_1-\alpha(\varepsilon-\varepsilon')+D_1^{t+1}-S_1^{t+1},-\alpha(\varepsilon-\varepsilon')+S_2^{t+1}-D_2^{t+1})
\end{aligned}
$$

Let us denote $\varepsilon''=\varepsilon-\varepsilon'=\hat{q}$ (which is between 0 and ε by the feasibility of $\hat{q}$), and we have

$$
\begin{aligned}
&[\tilde{U}_t(z_1,0)-\tilde{U}_t(z_1-\varepsilon,-\varepsilon)]-[\tilde{U}_t(z_1,\varepsilon)-\tilde{U}_t(z_1-\varepsilon,0)]\\
\geq&r_{22}^t\varepsilon-r_{12}^t\varepsilon''+\gamma\alpha r_{11}^{t+1}\varepsilon''-\gamma\alpha r_{22}^{t+1}(\varepsilon-\varepsilon'')\\
&+\gamma E\tilde{U}_{t+1}(\alpha z_1+D_1^{t+1}-S_1^{t+1},S_2^{t+1}-D_2^{t+1})\\
&-\gamma E\tilde{U}_{t+1}(\alpha z_1-\alpha\varepsilon''+D_1^{t+1}-S_1^{t+1},-\alpha\varepsilon''+S_2^{t+1}-D_2^{t+1})\\
\geq&r_{22}^t\varepsilon-r_{12}^t\varepsilon''+\gamma\alpha r_{11}^{t+1}\varepsilon''-\gamma\alpha r_{22}^{t+1}(\varepsilon-\varepsilon'')-\gamma\alpha\left(r_{22}^{t+1}-r_{12}^{t+1}+r_{11}^{t+1}\right)\varepsilon''\\
\geq&(r_{22}^t-r_{12}^t)\varepsilon''-\gamma\alpha(r_{22}^{t+1}-r_{12}^{t+1})\varepsilon''+(r_{22}^t-\gamma\alpha r_{22}^{t+1})(\varepsilon-\varepsilon'')\geq 0,
\end{aligned}
$$

where the second inequality follows from Lemma A.3, and the last inequality holds because of Assumption 2.

We note that the above proof applies to both the continuous-valued model and the discrete-valued model (for the latter model, we need to use $\varepsilon=1$ in the proof). □

Proof of Proposition 1 (Discrete-valued model with $\alpha=\beta$). We only consider the case with $z_1>0$ and $z_2>0$, and the proof for the case with $z_1<0$ and $z_2<0$ will be symmetric. By (A.6), for $q\in M(\mathbf{z})$ we have:

$$\begin{aligned}\tilde{J}_t(q,\mathbf{z}) =&\gamma r_{11}^{t+1}ED_1^{t+1}+\gamma r_{22}^{t+1}ES_2^{t+1}-(r_{11}^t+r_{22}^t-r_{12}^t)q-(r_{11}^t-\gamma\alpha r_{11}^{t+1})(z_1-q)-(r_{22}^t-\gamma\alpha r_{22}^{t+1})(z_2-q)\\ &+\gamma E\tilde{U}_{t+1}(\alpha(z_1-q)+D_1^{t+1}-S_1^{t+1},\alpha(z_2-q)+S_2^{t+1}-D_2^{t+1}).\end{aligned}$$

Let $p_s:=z_2-q$ and we can rewrite $\tilde{J}_t(q,\mathbf{z})$ as $\check{J}_t(p_s,IB)-(r_{22}^t-r_{12}^t+\gamma\alpha r_{11}^{t+1})z_2-(r_{11}^t-\gamma\alpha r_{11}^{t+1})(z_1-z_2)$, where

$$\begin{aligned}\check{J}_t(p_s,IB) =&\gamma r_{11}^{t+1}ED_1^{t+1}+\gamma r_{22}^{t+1}ES_2^{t+1}+(r_{11}^t+r_{22}^t-r_{12}^t)p_s-(r_{11}^t-\gamma\alpha r_{11}^{t+1})p_s-(r_{22}^t-\gamma\alpha r_{22}^{t+1})p_s\\ &+\gamma E\tilde{U}_{t+1}(\alpha IB+\alpha p_s+D_1^{t+1}-S_1^{t+1},\alpha p_s+S_2^{t+1}-D_2^{t+1}).\end{aligned}$$

The rest of the proof is identical to the continuous-valued model. Let $p_{s_2}(IB)\in\arg\max_{p_s\geq IB^-}\check{J}_t(p_s,IB)$. Since $\tilde{U}_{t+1}$ is $L^\natural$-concave (Lemma A.4) and any discrete-valued $L^\natural$-concave function can be extended to a continuous-valued concave function, we know that $\check{J}_t(p_s,IB)$ is concave with respect to p_s. Since $IB^-\leq p_s\leq z_2$, the optimal decision in terms of p_s is $p_s^*=\min\left\{z_2,p_{s_2}^t(IB)\right\}$. Thus, the optimal matching quantity between type 1 demand and type 2 supply is $q_{12}^{t*}=z_2-p_s^*=z_2-\min\left\{z_2,p_{s_2}^t(IB)\right\}=[z_2-p_{s_2}^t(IB)]^+=[y_2-x_2-p_{s_2}^t(IB)]^+$. □

A.2. The case with perishable demand

We consider the case with $\alpha=0$ and $\beta>0$. In that case, we can rewrite the expression of $J_t(q,\mathbf{z})$ given in (A.3) as follows:

$$\begin{aligned}J_t(q,\mathbf{z})=&\,r_{12}^tq^++r_{21}^tq^-+\gamma r_{11}^{t+1}E\min\left\{D_1^{t+1},\beta(z_1-q)^-+S_1^{t+1}\right\}\\ &+\gamma r_{22}^{t+1}E\min\left\{D_2^{t+1},\beta(z_2-q)^++S_2^{t+1}\right\}\\ &+\gamma EU_{t+1}(D_1^{t+1}-\beta(z_1-q)^--S_1^{t+1},\beta(z_2-q)^++S_2^{t+1}-D_2^{t+1}).\end{aligned}\tag{A.7}$$

We show that the optimal matching in round 2 of a period is fully determined by state-independent threshold levels.

LEMMA A.5. *In each period t, there exists state-independent threshold levels $\bar{p}_{s_1}^t$ and $\bar{p}_{s_2}^t$ such that,*

(i) If $z_1>0$ and $z_2>0$, the optimal matching quantity between type 1 demand and type 2 supply is $q_{12}^{t*}=\left[z_2-\max\left\{IB^-,\bar{p}_{s_2}^t\right\}\right]^+$;

(ii) If $z_1 < 0$ and $z_1 < 0$, the optimal matching quantity between type 2 demand and type 1 supply is $q_{21}^{t*} = \left[-z_1 - \max\left\{IB^-, \bar{p}_{s_1}^t\right\}\right]^+$.

Proof of Lemma A.5. We focus on proving part (i), and the proof of part (ii) is analogous.

Given that $z_1 \geq 0$ and $z_2 \geq 0$ in part (i), we have $M(\mathbf{z}) = \{q \mid 0 \leq q \leq \min\{z_1, z_2\}\}$. Thus for $q \in M(\mathbf{z})$ we can rewrite $J_t(q, \mathbf{z})$ as:

$$\begin{aligned} J_t(q,\mathbf{z}) =& r_{12}^t q + \gamma r_{11}^{t+1} E \min\left\{D_1^{t+1}, S_1^{t+1}\right\} + \gamma r_{22}^{t+1} E \min\left\{D_2^{t+1}, \beta(z_2 - q) + S_2^{t+1}\right\} \\ & + \gamma E U_{t+1}(D_1^{t+1} - S_1^{t+1}, \beta(z_2 - q) + S_2^{t+1} - D_2^{t+1}). \end{aligned} \tag{A.8}$$

The sum of the last three terms in the above equation represents the expected total discounted reward from period $t+1$ to period T, which is equal to $\gamma EV_{t+1}(D_1^{t+1}, D_2^{t+1}, S_1^{t+1}, \beta(z_2 - q) + S_2^{t+1})$. Let us define $p_s := z_2 - q$. Since $0 \leq q \leq \min\{z_1, z_2\}$, the feasible range of p_s is $(z_2 - z_1)^+ \leq p_s \leq z_2$. Thus, $J_t(q, \mathbf{z}) = r_{12}^t z_2 - r_{12}^t p_s + \gamma EV_{t+1}(D_1^{t+1}, D_2^{t+1}, S_1^{t+1}, \beta p_s + S_2^{t+1})$, and $\max_{0 \leq q \leq \min\{z_1, z_2\}} J_t(q, \mathbf{z})$ is equivalent to

$$\max_{(z_2 - z_1)^+ \leq p_s \leq z_2} \left\{-r_{12}^t p_s + \gamma EV_{t+1}(D_1^{t+1}, D_2^{t+1}, S_1^{t+1}, \beta p_s + S_2^{t+1})\right\}. \tag{A.9}$$

Let us define $\bar{p}_{s_1}^t := \arg\max_{p_s \geq 0} \left\{-r_{12}^t p_s + \gamma EV_{t+1}(D_1^{t+1}, D_2^{t+1}, S_1^{t+1}, \beta p_s + S_2^{t+1})\right\}$. If $EV_{t+1}(D_1^{t+1}, D_2^{t+1}, S_1^{t+1}, \beta p_s + S_2^{t+1})$ is concave with respect to p_s (which we will prove shortly), the optimal solution to (A.9) is $p_s^* = \min\left\{z_2, \max\left\{(z_2 - z_1)^+, \bar{p}_{s_1}^t\right\}\right\}$. Consequently, we can obtain the optimal matching quantity $q_{12}^{t*} = z_2 - p_s^* = z_2 - \min\left\{z_2, \max\left\{(z_2 - z_1)^+, \bar{p}_{s_1}^t\right\}\right\} = \left[z_2 - \max\left\{(z_2 - z_1)^+, \bar{p}_{s_1}^t\right\}\right]^+ = \left[z_2 - \max\{IB^-, \bar{p}_{s_1}^t\}\right]^+$.

Finally, for the continuous-valued model, $EV_{t+1}(D_1^{t+1}, D_2^{t+1}, S_1^{t+1}, \beta p_s + S_2^{t+1})$ is clearly concave with respect to p_s, given the joint concavity of V_{t+1} with respect to all state variables.

For the discrete-valued model, let us define $\tilde{U}_t(\mathbf{z}) := -r_{11}^t z_1^+ - r_{22}^t z_2^+ + U_t(\mathbf{z})$ as in Online Supplement A.1. Following similar analysis as in Online Supplement A.1, we can show that $\tilde{U}_t$ is $L^\natural$-concave for all t. Then, we have

$$\begin{aligned} & EV_{t+1}(D_1^{t+1}, D_2^{t+1}, S_1^{t+1}, \beta p_s + S_2^{t+1}) \\ =& \gamma r_{11}^{t+1} E \min\left\{D_1^{t+1}, S_1^{t+1}\right\} + \gamma r_{22}^{t+1} E \min\left\{D_2^{t+1}, \beta p_s + S_2^{t+1}\right\} + \gamma E U_{t+1}(D_1^{t+1} - S_1^{t+1}, \beta p_s + S_2^{t+1} - D_2^{t+1}) \\ =& \gamma r_{11}^{t+1} E D_1^{t+1} + \gamma \beta r_{22}^{t+1} p_s + \gamma r_{22}^{t+1} E S_2^{t+1} + \gamma E \tilde{U}_{t+1}(D_1^{t+1} - S_1^{t+1}, \beta p_s + S_2^{t+1} - D_2^{t+1}), \end{aligned}$$

which is concave with respect to p_s due to the $L^\natural$-concavity of $\tilde{U}_t$. $\square$

Finally, we note that Lemma A.5 is a stronger result than Proposition 1 for the discrete-valued model with $\alpha=0$. Therefore, the proof of Proposition 1 is also completed.

B. An alternative formulation of the vertical model

We reformulate the vertical model with a *transformed* system state and the total matching quantity Q as the decision variable in each period. For $i=1,\dots,m$ and $j=1,\dots,n$, let $\tilde{x}_i := \sum_{k=1}^{i} x_k$ and $\tilde{y}_j := \sum_{k=1}^{j} y_k$ ($\tilde{x}_0$ and $\tilde{y}_0$ are defined as zero) as the *transformed* system state, $\tilde{u}_i = \sum_{k=1}^{i} u_k$ and $\tilde{v}_j = \sum_{k=1}^{i} v_k$ as the transformed post-matching levels. In addition, let $\tilde{D}_i^t = \sum_{k=1}^{i} D_k^t$ and $\tilde{S}_j^t = \sum_{k=1}^{j} S_k^t$ be the transformed random variables for new arrivals of demand and supply in period t. We write $\tilde{\mathbf{x}} = (\tilde{x}_1,\dots,\tilde{x}_m)$, $\tilde{\mathbf{y}} = (\tilde{y}_1,\dots,\tilde{y}_n)$, $\tilde{\mathbf{u}} = (\tilde{u}_1,\dots,\tilde{u}_m)$, $\tilde{\mathbf{v}} = (\tilde{v}_1,\dots,\tilde{v}_n)$, $\tilde{\mathbf{D}}^t = (\tilde{D}_1^t,\dots,\tilde{D}_m^t)$ and $\tilde{\mathbf{S}}^t = (\tilde{S}_1^t,\dots,\tilde{S}_n^t)$.

Let $\mathbf{U}_k$ be the $k\times k$ upper triangular matrix with all the entries on or above the main diagonal equal to one. Then the state transformation can be written in a matrix form: $\mathbf{x}\mathbf{U}_m = \tilde{\mathbf{x}}$ and $\mathbf{y}\mathbf{U}_n = \tilde{\mathbf{y}}$. Equivalently, we can write $\mathbf{x} = \tilde{\mathbf{x}}\mathbf{U}_m^{-1}$ and $\mathbf{y} = \tilde{\mathbf{y}}\mathbf{U}_n^{-1}$. Here $\mathbf{U}_m^{-1}$ and $\mathbf{U}_n^{-1}$ are the inverse matrices of $\mathbf{U}_m$ and $\mathbf{U}_n$, respectively. One can verify that both $\mathbf{U}_m^{-1}$ and $\mathbf{U}_n^{-1}$ have all their diagonal entries equal to 1 and each off-diagonal entry right above an entry on the main diagonal equal to -1.

Let the total matching quantity Q be the decision variable in period t. Given the transformed state $(\tilde{\mathbf{x}},\tilde{\mathbf{y}})$, the feasible range of Q is $0\le Q\le \min\{\tilde{x}_m,\tilde{y}_n\}$. Under top-down matching, the total matching quantity Q fulfills a total quantity $\min\{Q,\tilde{x}_i\}$ of types $1,\dots,i$ demand combined, and uses a total quantity $\min\{Q,\tilde{y}_j\}$ of types $1,\dots,j$ supply combined. Thus, the quantity of type i demand fulfilled is the total fulfilled quantity of types $1,\dots,i$ demand less the total fulfilled quantity of types $1,\dots,i-1$ demand, i.e.,

$$\min\{\tilde{x}_i,Q\} - \min\{\tilde{x}_{i-1},Q\} = \tilde{x}_i - \tilde{x}_{i-1} - (\tilde{x}_i - Q)^+ + (\tilde{x}_{i-1} - Q)^+.$$

As a result, type i demand contributes the reward $r_{id}^t[\tilde{x}_i - \tilde{x}_{i-1} - (\tilde{x}_i - Q)^+ + (\tilde{x}_{i-1} - Q)^+]$ in period t. Likewise, type j supply contributes the reward $r_{js}^t[\tilde{y}_j - \tilde{y}_{j-1} - (\tilde{y}_j - Q)^+ + (\tilde{y}_{j-1} - Q)^+]$ in period t. Consequently, the total reward received in period t is

$$\begin{aligned}
&\sum_{i=1}^{m} r_{id}^t[\tilde{x}_i - \tilde{x}_{i-1} - (\tilde{x}_i - Q)^+ + (\tilde{x}_{i-1} - Q)^+] + \sum_{j=1}^{n} r_{js}^t[\tilde{y}_j - \tilde{y}_{j-1} - (\tilde{y}_j - Q)^+ + (\tilde{y}_{j-1} - Q)^+]\\
=&\sum_{i=1}^{m}(r_{id}^t - r_{i+1,d}^t)\tilde{x}_i + \sum_{i=1}^{n}(r_{js}^t - r_{j+1,d}^t)\tilde{y}_j - \sum_{i=1}^{m}(r_{id}^t - r_{i+1,d}^t)(\tilde{x}_i - Q)^+ - \sum_{i=1}^{n}(r_{js}^t - r_{j+1,d}^t)(\tilde{y}_j - Q)^+\\
=&\tilde{\mathbf{x}}\mathbf{U}_m^{-1}(\mathbf{r}_d^t)^\intercal + \tilde{\mathbf{y}}\mathbf{U}_n^{-1}(\mathbf{r}_s^t)^\intercal - (\tilde{\mathbf{x}} - Q\mathbf{1}^m)^+\mathbf{U}_m^{-1}(\mathbf{r}_d^t)^\intercal - (\tilde{\mathbf{y}} - Q\mathbf{1}^m)^+\mathbf{U}_n^{-1}(\mathbf{r}_s^t)^\intercal
\end{aligned}$$

where $r^t_{m+1,d} = r^t_{n+1,s} := 0$.

In the end of period t, the remaining quantity of types $1,\dots,i$ demand combined is $(\tilde{x}_i - Q)^+$ and the remaining quantity of types $1,\dots,j$ supply combined is $(\tilde{y}_j - Q)^+$. Thus, the transformed post-matching levels are given by $(\tilde{\mathbf{u}},\tilde{\mathbf{v}}) = ((\tilde{\mathbf{x}} - Q\mathbf{1}^m)^+, (\tilde{\mathbf{y}} - Q\mathbf{1}^n)^+)$. The transformed state in period $t+1$ is $(\tilde{\mathbf{x}}_{t+1},\tilde{\mathbf{y}}_{t+1}) = (\alpha(\tilde{\mathbf{x}} - Q\mathbf{1}^m)^+ + \tilde{\mathbf{D}}^{t+1}, \beta(\tilde{\mathbf{y}} - Q\mathbf{1}^n)^+ + \tilde{\mathbf{S}}^{t+1})$, which can be converted back to the original state as $(\mathbf{x}_{t+1},\mathbf{y}_{t+1}) = (\tilde{\mathbf{x}}_{t+1}\mathbf{U}_m^{-1}, \tilde{\mathbf{y}}_{t+1}\mathbf{U}_n^{-1})$.

With the total matching quantity in period t equal to Q, the maximum total expected reward achievable from period t to period T is

$$\begin{aligned} G_t(Q,\tilde{\mathbf{x}},\tilde{\mathbf{y}}) =& \tilde{\mathbf{x}}\mathbf{U}_m^{-1}(\mathbf{r}^t_d)^\intercal + \tilde{\mathbf{y}}\mathbf{U}_n^{-1}(\mathbf{r}^t_s)^\intercal - (\tilde{\mathbf{x}} - Q\mathbf{1}^m)^+\mathbf{U}_m^{-1}(\mathbf{r}^t_d)^\intercal - (\tilde{\mathbf{y}} - Q\mathbf{1}^m)^+\mathbf{U}_n^{-1}(\mathbf{r}^t_s)^\intercal \\ &+ \gamma E V_{t+1}(\alpha(\tilde{\mathbf{x}} - Q\mathbf{1}^m)^+\mathbf{U}_m^{-1} + \tilde{\mathbf{D}}^{t+1}\mathbf{U}_m^{-1}, \beta(\tilde{\mathbf{y}} - Q\mathbf{1}^n)^+\mathbf{U}_n^{-1} + \tilde{\mathbf{S}}^{t+1}\mathbf{U}_n^{-1}) \end{aligned} \tag{B.10}$$

given the transformed state $(\tilde{\mathbf{x}},\tilde{\mathbf{y}})$ in period t. The optimal total expected reward from period t to period T is thus

$$V_t(\tilde{\mathbf{x}}\mathbf{U}_m^{-1}, \tilde{\mathbf{y}}\mathbf{U}_n^{-1}) = \max_{0\le Q\le \tilde{x}_m\wedge\tilde{y}_n} G_t(Q,\tilde{\mathbf{x}},\tilde{\mathbf{y}}). \tag{B.11}$$

Let us define

$$\tilde{V}_t(\tilde{\mathbf{x}},\tilde{\mathbf{y}}) := V_t(\tilde{\mathbf{x}}\mathbf{U}_m^{-1}, \tilde{\mathbf{y}}\mathbf{U}_n^{-1}) - \tilde{\mathbf{x}}\mathbf{U}_m^{-1}(\mathbf{r}^t_d)^\intercal - \tilde{\mathbf{y}}\mathbf{U}_n^{-1}(\mathbf{r}^t_s)^\intercal. \tag{B.12}$$

$$\tilde{G}_t(Q,\tilde{\mathbf{x}},\tilde{\mathbf{y}}) := -\tilde{\mathbf{x}}\mathbf{U}_m^{-1}(\mathbf{r}^t_d)^\intercal - \tilde{\mathbf{y}}\mathbf{U}_n^{-1}(\mathbf{r}^t_s)^\intercal + G_t(Q,\tilde{\mathbf{x}},\tilde{\mathbf{y}}). \tag{B.13}$$

We see that Equations (B.10) and (B.11) are equivalent to:

$$\tilde{V}_t(\tilde{\mathbf{x}},\tilde{\mathbf{y}}) = \max_{0\le Q\le \min\{\tilde{x}_m,\tilde{y}_n\}} \tilde{G}_t(\tilde{\mathbf{x}},\tilde{\mathbf{y}}). \tag{B.14}$$

$$\begin{aligned} \tilde{G}_t(Q,\tilde{\mathbf{x}},\tilde{\mathbf{y}}) =& \gamma E[\mathbf{D}^{t+1}(\mathbf{r}^{t+1}_d)^\intercal] + \gamma E[\mathbf{S}^{t+1}(\mathbf{r}^{t+1}_s)^\intercal] \\ &- (\tilde{\mathbf{x}} - Q\mathbf{1}^m)^+\mathbf{U}_m^{-1}(\mathbf{r}^t_d - \gamma\alpha\mathbf{r}^{t+1}_d)^\intercal - (\tilde{\mathbf{y}} - Q\mathbf{1}^m)^+\mathbf{U}_n^{-1}(\mathbf{r}^t_s - \gamma\beta\mathbf{r}^{t+1}_s)^\intercal \\ &+ \gamma E\tilde{V}_{t+1}(\alpha(\tilde{\mathbf{x}} - Q\mathbf{1}^m)^+ + \tilde{\mathbf{D}}^{t+1}, \beta(\tilde{\mathbf{y}} - Q\mathbf{1}^n)^+ + \tilde{\mathbf{S}}^{t+1}) \end{aligned} \tag{B.15}$$

Since $V_{T+1}(\mathbf{x},\mathbf{y}) \equiv 0$, we have $\tilde{V}_{T+1}(\tilde{\mathbf{x}},\tilde{\mathbf{y}}) \equiv -\tilde{\mathbf{x}}\mathbf{U}_m^{-1}(\mathbf{r}^t_d)^\intercal - \tilde{\mathbf{y}}\mathbf{U}_n^{-1}(\mathbf{r}^t_s)^\intercal$.

We show the following property for the function $\tilde{V}_t(\tilde{\mathbf{x}},\tilde{\mathbf{y}})$.

Lemma B.6. *Suppose that Assumption 3 holds. Then, for any period $t = 1,\dots,T$, the function $\tilde{V}_t(\tilde{\mathbf{x}},\tilde{\mathbf{y}})$ is decreasing in x_i for $i = 1,\dots,m-1$ and in y_j for all $j = 1,\dots,n-1$.*

Proof of Lemma B.6. By definition of the function $\tilde{V}_t$, for $i=1,\ldots,m-1$, we have

$$\begin{aligned}
&\tilde{V}_t(\tilde{\mathbf{x}}+\varepsilon\mathbf{e}_i^m,\tilde{\mathbf{y}})-\tilde{V}_t(\tilde{\mathbf{x}},\tilde{\mathbf{y}})\\
=&V_t(\mathbf{x}+\varepsilon\mathbf{e}_i^m-\varepsilon\mathbf{e}_{i+1}^m,\mathbf{y})-(\mathbf{x}+\varepsilon\mathbf{e}_i^m-\varepsilon\mathbf{e}_{i+1}^m)(\mathbf{r}_d^t)^\intercal-\mathbf{y}(\mathbf{r}_s^t)^\intercal-V_t(\mathbf{x},\mathbf{y})+\mathbf{x}(\mathbf{r}_d^t)^\intercal-\mathbf{y}(\mathbf{r}_s^t)^\intercal\\
=&V_t(\mathbf{x}+\varepsilon\mathbf{e}_i^m-\varepsilon\mathbf{e}_{i+1}^m,\mathbf{y})-V_t(\mathbf{x},\mathbf{y})-(r_{id}^t-r_{i+1,d}^t)\varepsilon,
\end{aligned}\tag{B.16}$$

where $\mathbf{x}=\tilde{\mathbf{x}}\mathbf{U}_m^{-1}$ and $\mathbf{y}=\tilde{\mathbf{y}}\mathbf{U}_n^{-1}$.

By Lemma E.22, there exists $\lambda_{j'}^\tau\geq 0$ for $j'=1,\ldots,n$ and $\tau=t,\ldots,T$ such that $\sum_{\tau=t}^T\alpha^{-(\tau-t)}\sum_{j'=1}^n\lambda_{j'}^\tau\leq\varepsilon$ and $V_t(\mathbf{x}+\varepsilon\mathbf{e}_i^m-\varepsilon\mathbf{e}_{i+1}^m,\mathbf{y})-V_t(\mathbf{x},\mathbf{y})\leq\sum_{\tau=t}^T\gamma^{\tau-t}\sum_{j'=1}^n\lambda_{j'}^\tau(r_{ij'}^\tau-r_{i+1,j'}^\tau)$. Following (B.16), we have

$$\begin{aligned}
&\tilde{V}_t(\tilde{\mathbf{x}}+\varepsilon\mathbf{e}_i^m,\tilde{\mathbf{y}})-\tilde{V}_t(\tilde{\mathbf{x}},\tilde{\mathbf{y}})\\
\leq&\sum_{\tau=t}^T\gamma^{\tau-t}\sum_{j'=1}^n\lambda_{j'}^\tau(r_{ij'}^\tau-r_{i+1,j'}^\tau)-(r_{id}^t-r_{i+1,d}^t)\varepsilon\\
\leq&\sum_{\tau=t}^T\gamma^{\tau-t}\gamma^{-(\tau-t)}\alpha^{-(\tau-t)}\sum_{j'=1}^n\lambda_{j'}^\tau(r_{ij'}^t-r_{i+1,j'}^t)-(r_{id}^t-r_{i+1,d}^t)\varepsilon\\
=&\sum_{\tau=t}^T\alpha^{-(\tau-t)}\sum_{j'=1}^n\lambda_{j'}^\tau(r_{id}^t-r_{i+1,d}^t)-(r_{id}^t-r_{i+1,d}^t)\varepsilon\\
=&[\sum_{\tau=t}^T\alpha^{-(\tau-t)}\sum_{j'=1}^n\lambda_{j'}^\tau-\varepsilon](r_{id}^t-r_{i+1,d}^t)\leq 0.
\end{aligned}$$

Therefore, $\tilde{V}_t(\tilde{\mathbf{x}},\tilde{\mathbf{y}})$ is decreasing in $\tilde{x}_i$ for $i=1,\ldots,m-1$. Similarly, we can show that $\tilde{V}_t(\tilde{\mathbf{x}},\tilde{\mathbf{y}})$ is decreasing in $\tilde{y}_j$ for $j=1,\ldots,n-1$. □

The following lemma shows the $L^\natural$-concavity of the functions $\tilde{G}_t$ and $\tilde{V}_t$, for the case with equal carry-over rates and the case with perishable demand.

LEMMA B.7. *(i) Suppose that $\alpha=\beta>0$. Then, $\tilde{V}_t(\tilde{\mathbf{x}},\tilde{\mathbf{y}})$ is $L^\natural$-concave in $(\tilde{\mathbf{x}},\tilde{\mathbf{y}})$ for $t=1,\ldots,T+1$, and $\tilde{G}_t(Q,\tilde{\mathbf{x}},\tilde{\mathbf{y}})$ is $L^\natural$-concave in $(Q,\tilde{\mathbf{x}},\tilde{\mathbf{y}})$ for $t=1,\ldots,T$.*

(ii) Suppose that $\alpha=0<\beta$. Then, $\tilde{V}_t(\tilde{\mathbf{x}},\tilde{\mathbf{y}})$ is $L^\natural$-concave in $\tilde{\mathbf{y}}$ for $t=1,\ldots,T+1$, and $\tilde{G}_t(Q,\tilde{\mathbf{x}},\tilde{\mathbf{y}})$ is $L^\natural$-concave in $(Q,\tilde{\mathbf{x}},\tilde{\mathbf{y}})$ for $t=1,\ldots,T$.

Proof of Lemma B.7. (i) We first consider the case with $\alpha=\beta>0$. The proof is by induction on t. Clearly, $\tilde{V}_{T+1}(\tilde{\mathbf{x}},\tilde{\mathbf{y}})=-\tilde{\mathbf{x}}\mathbf{U}_m^{-1}(\mathbf{r}_d^t)^\intercal-\tilde{\mathbf{y}}\mathbf{U}_n^{-1}(\mathbf{r}_s^t)^\intercal$ is $L^\natural$-concave in $(\tilde{\mathbf{x}},\tilde{\mathbf{y}})$. We suppose that $\tilde{V}_{t+1}(\tilde{\mathbf{x}},\tilde{\mathbf{y}})$ is $L^\natural$-concave in $(\tilde{\mathbf{x}},\tilde{\mathbf{y}})$. Then by definition of $L^\natural$-concavity, for any given $\tilde{\mathbf{D}}^{t+1}$ and $\tilde{\mathbf{S}}^{t+1}$, $\tilde{V}_{t+1}(\alpha\tilde{\mathbf{x}}+\tilde{\mathbf{D}}^{t+1},\alpha\tilde{\mathbf{y}}+\tilde{\mathbf{S}}^{t+1})$ is $L^\natural$-concave in $(\tilde{\mathbf{x}},\tilde{\mathbf{y}})$. (Note that since $\alpha=\beta$, we simply replace β by α hereafter.) Since $Q\leq\min\{\tilde{x}_m,\tilde{y}_n\}$, we have

$$\tilde{V}_{t+1}(\alpha(\tilde{\mathbf{x}}-Q\mathbf{1}^n)^+ + \tilde{\mathbf{D}}^{t+1}, \alpha(\tilde{\mathbf{y}}-Q\mathbf{1}^m)^+ + \tilde{\mathbf{S}}^{t+1})$$
$$= \tilde{V}_{t+1}(\alpha(\tilde{\mathbf{x}}_{[1,m-1]}-Q\mathbf{1}^{m-1})^+ + \tilde{\mathbf{D}}^{t+1}_{[1,m-1]}, \alpha(\tilde{x}_m - Q) + \tilde{D}^{t+1}_m, \alpha(\tilde{\mathbf{y}}_{[1,n-1]}-Q\mathbf{1}^{n-1})^+ + \tilde{\mathbf{S}}^{t+1}_{[1,n-1]}, \alpha(\tilde{y}_n - Q) + \tilde{S}^{t+1}_n),$$

which is $L^\natural$-concave in $(Q,\tilde{\mathbf{x}},\tilde{\mathbf{y}})$ by applying Chen et al. (2014, Lemma 4) and noting the monotonicity proved in Lemma B.6. (The notation $\tilde{\mathbf{x}}_{[1,m-1]}$ represents the first $m-1$ entries of the vector $\tilde{\mathbf{x}}$.) By Simchi-Levi et al. (2014, Proposition 2.3.4(c)), $E_{\tilde{\mathbf{D}}^{t+1},\tilde{\mathbf{S}}^{t+1}}[\tilde{V}_{t+1}(\alpha(\tilde{\mathbf{x}}-Q\mathbf{1}^m)^+ + \tilde{\mathbf{D}}^{t+1}, \alpha(\tilde{\mathbf{y}}-Q\mathbf{1}^n)^+ + \tilde{\mathbf{S}}^{t+1})]$ is $L^\natural$-concave in $(Q,\tilde{\mathbf{x}},\tilde{\mathbf{y}})$, thus the last term in (B.15) is $L^\natural$-concave in $(Q,\tilde{\mathbf{x}},\tilde{\mathbf{y}})$. The other terms in (B.15) are $L^\natural$-concave in $(Q,\tilde{\mathbf{x}},\tilde{\mathbf{y}})$, because $-(\tilde{x}_{i'}-Q)^+$ is supermodular in $(Q,\tilde{x}_{i'})$, $-(\tilde{y}_{j'}-Q)^+$ is supermodular in $(Q,\tilde{y}_{j'})$ and $L^\natural$-concavity is preserved under any nonnegative linear combination. Since the other terms are linear, $\tilde{G}_t(Q,\tilde{\mathbf{x}},\tilde{\mathbf{y}})$ is $L^\natural$-concave in $(Q,\tilde{\mathbf{x}},\tilde{\mathbf{y}})$. By Simchi-Levi et al. (2014, Proposition 2.3.4(e)), $\tilde{V}_t(\tilde{\mathbf{x}},\tilde{\mathbf{y}})$ is $L^\natural$-concave in $(\tilde{\mathbf{x}},\tilde{\mathbf{y}})$. This completes the induction.

(ii) The proof is again based on induction. As in the proof of part (i), $\tilde{V}_{T+1}(\tilde{\mathbf{x}},\tilde{\mathbf{y}})$ is $L^\natural$-concave. Suppose that $\tilde{V}_{t+1}(\tilde{\mathbf{x}},\tilde{\mathbf{y}})$ is $L^\natural$-concave in $\tilde{\mathbf{y}}$, for any given $\tilde{\mathbf{x}}$. Since $\alpha=0$, we have $E_{\tilde{\mathbf{D}}^{t+1},\tilde{\mathbf{S}}^{t+1}}[\tilde{V}_{t+1}(\alpha(\tilde{\mathbf{x}}-Q\mathbf{1}^m)^+ + \tilde{\mathbf{D}}^{t+1}, \beta(\tilde{\mathbf{y}}-Q\mathbf{1}^n)^+ + \tilde{\mathbf{S}}^{t+1})] = E_{\tilde{\mathbf{D}}^{t+1},\tilde{\mathbf{S}}^{t+1}}[\tilde{V}_{t+1}(\tilde{\mathbf{D}}^{t+1}, \beta(\tilde{\mathbf{y}}-Q\mathbf{1}^n)^+ + \tilde{\mathbf{S}}^{t+1})]$ is independent of $\tilde{\mathbf{x}}$. Again by applying Chen et al. (2014, Lemma 4) and noting the monotonicity proved in Lemma B.6, it is $L^\natural$-concave in $\tilde{\mathbf{y}}$. It follows that $\tilde{G}_t(Q,\tilde{\mathbf{x}},\tilde{\mathbf{y}})$ is $L^\natural$-concave in $(Q,\tilde{\mathbf{y}})$ for any given $\tilde{\mathbf{x}}$. From (B.15), we see that $\tilde{G}_t(Q,\tilde{\mathbf{x}},\tilde{\mathbf{y}})$ depends on $\tilde{\mathbf{x}}$ only through the term $-(\tilde{\mathbf{x}}-Q\mathbf{1}^m)^+\mathbf{U}_m^{-1}(\mathbf{r}_d^t - \gamma\alpha\mathbf{r}_d^{t+1})^\intercal$, which is $L^\natural$-concave in $(\tilde{\mathbf{x}},Q)$. Thus, $\tilde{G}_t(Q,\tilde{\mathbf{x}},\tilde{\mathbf{y}})$ is $L^\natural$-concave in $(Q,\tilde{\mathbf{x}},\tilde{\mathbf{y}})$. As in the proof of part (i), it follows from Simchi-Levi et al. (2014, Proposition 2.3.4(e)) that $\tilde{V}_t(\tilde{\mathbf{x}},\tilde{\mathbf{y}})$ is $L^\natural$-concave in $\tilde{\mathbf{y}}$, for any given $\tilde{\mathbf{x}}$. □

B.1. The 1-step-lookahead heuristic for the vertical model

The 1-step-lookahead heuristic assumes greedy matching from the next period to the end of the horizon. Recall that we denote by $V_t^g(\mathbf{x},\mathbf{y})$ the total expected reward received under the greedy policy from period t to period T, given that the (original) state in period t is $(\mathbf{x},\mathbf{y})$. The heuristic chooses the matching quantities in period t to maximize the sum of the immediate reward in period t and the future expected reward $V_t^g(\mathbf{x},\mathbf{y})$. In the following, we explore properties of the function $V_t^g(\mathbf{x},\mathbf{y})$ and the 1-step-lookahead heuristic.

As earlier, we consider the transformed state $(\tilde{\mathbf{x}},\tilde{\mathbf{y}})$ defined by $\tilde{x}_i := \sum_{k=1}^i x_i$ and $\tilde{y}_j := \sum_{k=1}^j y_k$, and let $\tilde{V}_t^g(\tilde{\mathbf{x}},\tilde{\mathbf{y}}) := V_t^g(\mathbf{x},\mathbf{y}) - \tilde{\mathbf{x}}\mathbf{U}_m^{-1}(\mathbf{r}_d^t)^\intercal - \tilde{\mathbf{y}}\mathbf{U}_n^{-1}(\mathbf{r}_s^t)^\intercal$. The following lemma presents the recursive equations satisfied by the function $\tilde{V}_t^g(\mathbf{x},\mathbf{y})$.

Given the transformed state $(\tilde{\mathbf{x}},\tilde{\mathbf{y}})$ in period t, we define $G_t^g(Q,\tilde{\mathbf{x}},\tilde{\mathbf{y}})$ as the expected total discounted reward from period t to period T, if we apply top-down matching up to the total matching quantity Q in period t, and apply greedy matching from period $t+1$ to period T. Further, we let $\tilde{G}_t^g(Q,\tilde{\mathbf{x}},\tilde{\mathbf{y}}) := -\tilde{\mathbf{x}}\mathbf{U}_m^{-1}(\mathbf{r}_d^t)^\intercal - \tilde{\mathbf{y}}\mathbf{U}_n^{-1}(\mathbf{r}_s^t)^\intercal + G_t^g(Q,\tilde{\mathbf{x}},\tilde{\mathbf{y}})$. Similar to how we derived $G_t(Q,\tilde{\mathbf{x}},\tilde{\mathbf{y}})$ defined in (B.10), we can express $G_t^g(Q,\tilde{\mathbf{x}},\tilde{\mathbf{y}})$ as:

$$\begin{aligned} G_t^g(Q,\tilde{\mathbf{x}},\tilde{\mathbf{y}}) =& \tilde{\mathbf{x}}\mathbf{U}_m^{-1}(\mathbf{r}_d^t)^\intercal + \tilde{\mathbf{y}}\mathbf{U}_n^{-1}(\mathbf{r}_s^t)^\intercal - (\tilde{\mathbf{x}} - Q\mathbf{1}^m)^+\mathbf{U}_m^{-1}(\mathbf{r}_d^t)^\intercal - (\tilde{\mathbf{y}} - Q\mathbf{1}^m)^+\mathbf{U}_n^{-1}(\mathbf{r}_s^t)^\intercal \\ &+ \gamma E V_{t+1}^g(\alpha(\tilde{\mathbf{x}} - Q\mathbf{1}^m)^+\mathbf{U}_m^{-1} + \tilde{\mathbf{D}}^{t+1}\mathbf{U}_m^{-1}, \beta(\tilde{\mathbf{y}} - Q\mathbf{1}^n)^+\mathbf{U}_n^{-1} + \tilde{\mathbf{S}}^{t+1}\mathbf{U}_n^{-1}). \end{aligned} \tag{B.17}$$

By substituting V_{t+1}^g by $\tilde{V}_{t+1}^g$ according to the relation $V_{t+1}^g(\mathbf{x},\mathbf{y}) = \tilde{V}_{t+1}^g(\tilde{\mathbf{x}},\tilde{\mathbf{y}}) + \tilde{\mathbf{x}}\mathbf{U}_m^{-1}(\mathbf{r}_d^t)^\intercal + \tilde{\mathbf{y}}\mathbf{U}_n^{-1}(\mathbf{r}_s^t)^\intercal$, we have

$$\begin{aligned} \tilde{G}_t^g(Q,\tilde{\mathbf{x}},\tilde{\mathbf{y}}) =& \gamma E\mathbf{D}^{t+1}(\mathbf{r}_d^{t+1})^\intercal + \gamma E\mathbf{S}^{t+1}(\mathbf{r}_s^{t+1})^\intercal \\ & - (\tilde{\mathbf{x}} - Q\mathbf{1}^m)^+\mathbf{U}_m^{-1}(\mathbf{r}_d^t - \gamma\alpha\mathbf{r}_d^{t+1})^\intercal - (\tilde{\mathbf{y}} - Q\mathbf{1}^m)^+\mathbf{U}_n^{-1}(\mathbf{r}_s^t - \gamma\beta\mathbf{r}_s^{t+1})^\intercal \\ & + \gamma E\tilde{V}_{t+1}^g(\alpha(\tilde{\mathbf{x}} - Q\mathbf{1}^m)^+ + \tilde{\mathbf{D}}^{t+1}, \beta(\tilde{\mathbf{y}} - Q\mathbf{1}^n)^+ + \tilde{\mathbf{S}}^{t+1}). \end{aligned} \tag{B.18}$$

LEMMA B.8. *The functions $\tilde{V}_t^g(\mathbf{x},\mathbf{y})$, $t = 1,\ldots,T$, satisfy the following equations:*

$$\begin{aligned} \tilde{V}_t^g(\tilde{\mathbf{x}},\tilde{\mathbf{y}}) =& \gamma E\mathbf{D}^{t+1}(\mathbf{r}_d^{t+1})^\intercal + \gamma E\mathbf{S}^{t+1}(\mathbf{r}_s^{t+1})^\intercal \\ & - (\tilde{\mathbf{x}} - \tilde{y}_n\mathbf{1}^m)^+\mathbf{U}_m^{-1}(\mathbf{r}_d^t - \gamma\alpha\mathbf{r}_d^{t+1})^\intercal - (\tilde{\mathbf{y}} - \tilde{x}_m\mathbf{1}^n)^+\mathbf{U}_n^{-1}(\mathbf{r}_s^t - \gamma\beta\mathbf{r}_s^{t+1})^\intercal \\ & + \gamma E\tilde{V}_{t+1}^g(\alpha(\tilde{\mathbf{x}} - \tilde{y}_n\mathbf{1}^m)^+ + \mathbf{D}^{t+1}\mathbf{U}_m, \beta(\tilde{\mathbf{y}} - \tilde{x}_m\mathbf{1}^n)^+ + \mathbf{S}^{t+1}\mathbf{U}_n). \end{aligned} \tag{B.19}$$

Proof of Lemma B.8. Recall that $G_t^g(Q,\tilde{\mathbf{x}},\tilde{\mathbf{y}})$ as the expected total discounted reward from period t to period T, if we apply top-down matching up to the total matching quantity Q in period t, and apply greedy matching from period $t+1$ to period T. Similar to how we derived $G_t(Q,\tilde{\mathbf{x}},\tilde{\mathbf{y}})$ defined in (B.10), we can express $G_t^g(Q,\tilde{\mathbf{x}},\tilde{\mathbf{y}})$ as:

$$\begin{aligned} G_t^g(Q,\tilde{\mathbf{x}},\tilde{\mathbf{y}}) =& \tilde{\mathbf{x}}\mathbf{U}_m^{-1}(\mathbf{r}_d^t)^\intercal + \tilde{\mathbf{y}}\mathbf{U}_n^{-1}(\mathbf{r}_s^t)^\intercal - (\tilde{\mathbf{x}} - Q\mathbf{1}^m)^+\mathbf{U}_m^{-1}(\mathbf{r}_d^t)^\intercal - (\tilde{\mathbf{y}} - Q\mathbf{1}^m)^+\mathbf{U}_n^{-1}(\mathbf{r}_s^t)^\intercal \\ &+ \gamma E V_{t+1}^g(\alpha(\tilde{\mathbf{x}} - Q\mathbf{1}^m)^+\mathbf{U}_m^{-1} + \tilde{\mathbf{D}}^{t+1}\mathbf{U}_m^{-1}, \beta(\tilde{\mathbf{y}} - Q\mathbf{1}^n)^+\mathbf{U}_n^{-1} + \tilde{\mathbf{S}}^{t+1}\mathbf{U}_n^{-1}). \end{aligned}$$

If greedy matching is used in period t, the total matching quantity Q is $\tilde{x}_m \wedge \tilde{y}_n := \min\{\tilde{x}_m, \tilde{y}_n\}$. Thus,

$$V_t^g(\tilde{\mathbf{x}},\tilde{\mathbf{y}})$$

$$
\begin{aligned}
&= G_t(\tilde{x}_m \wedge \tilde{y}_n, \tilde{\mathbf{x}}, \tilde{\mathbf{y}}) \\
&= \tilde{\mathbf{x}}\mathbf{U}_m^{-1}(\mathbf{r}_d^t)^\intercal + \tilde{\mathbf{y}}\mathbf{U}_n^{-1}(\mathbf{r}_s^t)^\intercal - (\tilde{\mathbf{x}} - \tilde{x}_m \wedge \tilde{y}_n \mathbf{1}^m)^+ \mathbf{U}_m^{-1}(\mathbf{r}_d^t)^\intercal - (\tilde{\mathbf{y}} - \tilde{x}_m \wedge \tilde{y}_n \mathbf{1}^m)^+ \mathbf{U}_n^{-1}(\mathbf{r}_s^t)^\intercal \\
&\quad + \gamma E V_{t+1}^g(\alpha(\tilde{\mathbf{x}} - \tilde{x}_m \wedge \tilde{y}_n \mathbf{1}^m)^+ \mathbf{U}_m^{-1} + \mathbf{D}^{t+1}, \beta(\tilde{\mathbf{y}} - \tilde{x}_m \wedge \tilde{y}_n \mathbf{1}^n)^+ \mathbf{U}_n^{-1} + \mathbf{S}^{t+1}) \\
&= \gamma E \mathbf{D}^{t+1}(\mathbf{r}_d^{t+1})^\intercal + \gamma E \mathbf{S}^{t+1}(\mathbf{r}_s^{t+1})^\intercal + \tilde{\mathbf{x}}\mathbf{U}_m^{-1}(\mathbf{r}_d^t)^\intercal + \tilde{\mathbf{y}}\mathbf{U}_n^{-1}(\mathbf{r}_s^t)^\intercal \\
&\quad - (\tilde{\mathbf{x}} - \tilde{x}_m \wedge \tilde{y}_n \mathbf{1}^m)^+ \mathbf{U}_m^{-1}(\mathbf{r}_d^t - \gamma\alpha \mathbf{r}_d^{t+1})^\intercal - (\tilde{\mathbf{y}} - \tilde{x}_m \wedge \tilde{y}_n \mathbf{1}^m)^+ \mathbf{U}_n^{-1}(\mathbf{r}_s^t - \gamma\beta \mathbf{r}_s^{t+1})^\intercal \\
&\quad + \gamma E \tilde{V}_{t+1}^g(\alpha(\tilde{\mathbf{x}} - \tilde{x}_m \wedge \tilde{y}_n \mathbf{1}^m)^+ + \mathbf{D}^{t+1}\mathbf{U}_m, \beta(\tilde{\mathbf{y}} - \tilde{x}_m \wedge \tilde{y}_n \mathbf{1}^n)^+ + \mathbf{S}^{t+1}\mathbf{U}_n).
\end{aligned} \tag{B.20}
$$

Note that in the last equality above, we have substituted V_{t+1}^g with $\tilde{V}_{t+1}^g$.

Finally, we note that $(\tilde{\mathbf{x}} - \tilde{x}_m \wedge \tilde{y}_n \mathbf{1}^m)^+ = (\tilde{\mathbf{x}} - \tilde{y}_n \mathbf{1}^m)^+$ and $(\tilde{\mathbf{y}} - \tilde{x}_m \wedge \tilde{y}_n \mathbf{1}^n)^+ = (\tilde{\mathbf{y}} - \tilde{x}_m \mathbf{1}^n)^+$.

In fact, if $\tilde{x}_i \le \tilde{y}_n$, then $\tilde{x}_i \le \tilde{x}_m \wedge \tilde{y}_n$ and thus $(\tilde{x}_i - \tilde{x}_m \wedge \tilde{y}_n)^+ = 0$. If $\tilde{x}_i > \tilde{y}_n$, then $\tilde{x}_m \ge \tilde{x}_i > \tilde{y}_n$ and thus $(\tilde{x}_i - \tilde{x}_m \wedge \tilde{y}_n)^+ = (\tilde{x}_i - \tilde{y}_n)^+ = \tilde{x}_i - \tilde{y}_n$. It follows that $(\tilde{x}_i - \tilde{x}_m \wedge \tilde{y}_n)^+ = (\tilde{x}_i - \tilde{y}_n)^+$ and therefore $(\tilde{\mathbf{x}} - \tilde{x}_m \wedge \tilde{y}_n \mathbf{1}^m)^+ = (\tilde{\mathbf{x}} - \tilde{y}_n \mathbf{1}^m)^+$. Likewise, we can show that $(\tilde{\mathbf{y}} - \tilde{x}_m \wedge \tilde{y}_n \mathbf{1}^n)^+ = (\tilde{\mathbf{y}} - \tilde{x}_m \mathbf{1}^n)^+$.

Equation (B.19) then follows from (B.20). □

We further show the monotonicity of the function $\tilde{V}_t^g(\tilde{\mathbf{x}}, \tilde{\mathbf{y}})$.

LEMMA B.9. *Suppose that Assumption 3 holds. For any period $t = 1, \dots, T$, the function $\tilde{V}_t^g(\tilde{\mathbf{x}}, \tilde{\mathbf{y}})$ is decreasing in $\tilde{x}_i$ for all $i = 1, \dots, m-1$ and in y_j for all $j = 1 \dots, n-1$.*

Proof of Lemma B.9. We will show by induction that $\tilde{V}_t^g(\tilde{\mathbf{x}} + \varepsilon \mathbf{e}_i^m, \tilde{\mathbf{y}})$ decreases in $\tilde{x}_i$.

It is trivial to prove for $t = T+1$, given that $\tilde{V}_t^g(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) \equiv -\tilde{\mathbf{x}}\mathbf{U}_m^{-1}(\mathbf{r}_d^t)^\intercal - \tilde{\mathbf{y}}\mathbf{U}_n^{-1}(\mathbf{r}_s^t)^\intercal$. Let us suppose that $\tilde{V}_{t+1}^g(\tilde{\mathbf{x}}, \tilde{\mathbf{y}})$ is decreasing in $\tilde{x}_i$.

To show that $\tilde{V}_t^g(\tilde{\mathbf{x}}, \tilde{\mathbf{y}})$ is decreasing in $\tilde{x}_i$ $(1 \le i \le m-1)$, we note that $-(\tilde{\mathbf{x}} - \tilde{y}_n \mathbf{1}^m)^+ \mathbf{U}_m^{-1}(\mathbf{r}_d^t - \gamma\alpha \mathbf{r}_d^{t+1})^\intercal = -\sum_{i=1}^m (\tilde{x}_i - \tilde{y}_n)^+[(r_{id}^t - r_{i+1,d}^t) - \gamma\alpha(r_{id}^{t+1} - r_{i+1,d}^{t+1})]$ is decreasing in $\tilde{x}_i$. According to the induction hypothesis, the last term in (B.19), $\gamma E \tilde{V}_{t+1}^g(\alpha(\tilde{\mathbf{x}} - \tilde{y}_n \mathbf{1}^m)^+ \mathbf{U}_m^{-1} + \tilde{\mathbf{D}}^{t+1}\mathbf{U}_m^{-1}, \beta(\tilde{\mathbf{y}} - \tilde{x}_m \mathbf{1}^n)^+ \mathbf{U}_n^{-1} + \tilde{\mathbf{S}}^{t+1}\mathbf{U}_n^{-1})$ is decreasing in $\tilde{x}_i$ $(1 \le i \le m-1)$. Thus, all terms in (B.19) are either constant or decreasing in $\tilde{x}_i$. This completes the induction and shows that $\tilde{V}_t^g(\tilde{\mathbf{x}}, \tilde{\mathbf{y}})$ is decreasing in $\tilde{x}_i$. We show that it is also decreasing in $\tilde{y}_j$ $(1 \le j \le n-1)$ similarly. □

The following Lemma B.10 is concerned about the concavity of the functions $\tilde{V}_t(\tilde{\mathbf{x}}, \tilde{\mathbf{y}})$. We note that for the discrete-valued model, concavity is not defined for the multivariate function $\tilde{V}_t(\tilde{\mathbf{x}}, \tilde{\mathbf{y}})$. Nevertheless, it is easy to see from (B.20) that for any integral-valued state $(\tilde{\mathbf{x}}, \tilde{\mathbf{y}})$, the value of $\tilde{V}_t(\tilde{\mathbf{x}}, \tilde{\mathbf{y}})$ coincides with its value in the continuous-valued model. Therefore, the function $\tilde{V}_t(\tilde{\mathbf{x}}, \tilde{\mathbf{y}})$ in the discrete-valued model can be naturally extended to the one in the continuous-valued model. In that sense, Lemma B.10 is also applicable to the discrete-valued model.

Lemma B.10. *Suppose that Assumption 3 holds. For any* $t=1,\ldots,T$, *the function* $\tilde{V}_t^g(\tilde{\mathbf{x}},\tilde{\mathbf{y}})$ *is concave in* $(\tilde{\mathbf{x}}_{[1,m-1]},\tilde{\mathbf{y}}_{[1,n-1]})$.

Proof of Lemma B.10. The proof is based on induction. First, $\tilde{V}_{T+1}^g(\tilde{\mathbf{x}},\tilde{\mathbf{y}})\equiv 0$ is concave. Suppose that $\tilde{V}_{t+1}^g$ is concave in $(\tilde{\mathbf{x}}_{[1,m-1]},\tilde{\mathbf{y}}_{[1,n-1]})$.

The terms $-(\tilde{\mathbf{x}}-\tilde{y}_n\mathbf{1}^m)^+\mathbf{U}_m^{-1}(\mathbf{r}_d^t-\alpha\mathbf{r}_d^{t+1})^\intercal$ and $-(\tilde{\mathbf{y}}-\tilde{x}_m\mathbf{1}^n)^+\mathbf{U}_n^{-1}(\mathbf{r}_s^t-\beta\mathbf{r}_s^{t+1})^\intercal$ are concave due to the concavity of the function $f(x):=-x^+$. It remains to show that $E\tilde{V}_{t+1}^g(\alpha(\tilde{\mathbf{x}}-\tilde{y}_n\mathbf{1}^m)^+ + \tilde{\mathbf{D}}^{t+1},\beta(\tilde{\mathbf{y}}-\tilde{x}_m\mathbf{1}^n)^+ + \tilde{\mathbf{S}}^{t+1})$ is concave in $(\tilde{\mathbf{x}}_{[1,m-1]},\tilde{\mathbf{y}}_{[1,n-1]})$.

Consider two transformed states $(\tilde{\mathbf{x}}_{[1,m-1]},\tilde{x}_m,\tilde{\mathbf{y}}_{[1,n-1]},\tilde{y}_n)$ and $(\tilde{\mathbf{x}}'_{[1,m-1]},\tilde{x}_m,\tilde{\mathbf{y}}'_{[1,n-1]},\tilde{y}_n)$. For $\lambda_1\geq 0$ and $\lambda_2\geq 0$ such that $\lambda_1+\lambda_2=1$, and $i=1,\ldots,m-1$, we have

$$(\lambda_1\tilde{\mathbf{x}}_{[1,m-1]}+\lambda_2\tilde{\mathbf{x}}'_{[1,m-1]}-\tilde{y}_n\mathbf{1}^{m-1})^+\leq\lambda_1(\tilde{\mathbf{x}}_{[1,m-1]}-\tilde{y}_n\mathbf{1}^{m-1})^+ +\lambda_2(\tilde{\mathbf{x}}'_{[1,m-1]}-\tilde{y}_n\mathbf{1}^{m-1})^+, \quad \text{(B.21)}$$

where the inequalities follow from the convexity of the function $f(x):=x^+$.

Similarly, we have

$$(\lambda_1\tilde{\mathbf{y}}_{[1,n-1]}+\lambda_2\tilde{\mathbf{y}}'_{[1,n-1]}-\tilde{x}_m\mathbf{1}^{n-1})^+\leq\lambda_1(\tilde{\mathbf{y}}_{[1,n-1]}-\tilde{x}_m\mathbf{1}^{n-1})^+ +\lambda_2(\tilde{\mathbf{y}}'_{[1,n-1]}-\tilde{x}_n\mathbf{1}^{n-1})^+. \quad \text{(B.22)}$$

It follows from (B.21) and (B.22) and the fact that $\tilde{V}_{t+1}^g$ is decreasing in x_i and y_j for $i=1,\ldots,m-1$, $j=1,\ldots,n-1$ (Lemma B.9)

$$\begin{aligned}
&\tilde{V}_{t+1}^g(\alpha(\lambda_1\tilde{\mathbf{x}}+\lambda_2\tilde{\mathbf{x}}'-\lambda_1\tilde{y}_n\mathbf{1}^m-\lambda_2\tilde{y}'_n\mathbf{1}^m)^+ +\tilde{\mathbf{D}}^{t+1},\beta(\lambda_1\tilde{\mathbf{y}}+\lambda_2\tilde{\mathbf{y}}'-\lambda_1\tilde{x}_m\mathbf{1}^n-\lambda_2\tilde{x}'_m\mathbf{1}^n)^+ +\tilde{\mathbf{S}}^{t+1})\\
=&\tilde{V}_{t+1}^g(\alpha(\lambda_1\tilde{\mathbf{x}}_{[1,m-1]}+\lambda_2\tilde{\mathbf{x}}'_{[1,m-1]}-\tilde{y}_n\mathbf{1}^{m-1})^+ +\tilde{\mathbf{D}}_{[1,m-1]}^{t+1},\alpha(\tilde{x}_m-\tilde{y}_n)^+ +\tilde{D}_m^{t+1},\\
&\quad\beta(\lambda_1\tilde{\mathbf{y}}_{[1,n-1]}+\lambda_2\tilde{\mathbf{y}}'_{[1,n-1]}-\tilde{x}_m\mathbf{1}^{n-1})^+ +\tilde{\mathbf{S}}_{[1,n-1]}^{t+1},\beta(\tilde{y}_n-\tilde{x}_m)^+ +\tilde{S}_n^{t+1})\\
\geq&\tilde{V}_{t+1}^g(\lambda_1[\alpha(\tilde{\mathbf{x}}_{[1,m-1]}-\tilde{y}_n\mathbf{1}^{m-1})^+ +\tilde{\mathbf{D}}_{[1,m-1]}^{t+1}]+\lambda_2[\alpha(\tilde{\mathbf{x}}'_{[1,m-1]}-\tilde{y}_n\mathbf{1}^{m-1})^+ +\tilde{\mathbf{D}}_{[1,m-1]}^{t+1}],\alpha(\tilde{x}_m-\tilde{y}_n)^+ +\tilde{D}_m^{t+1},\\
&\quad\lambda_1[\beta(\tilde{\mathbf{y}}_{[1,n-1]}-\tilde{x}_m\mathbf{1}^{n-1})^+ +\tilde{\mathbf{S}}_{[1,n-1]}^{t+1}]+\lambda_2[\beta(\tilde{\mathbf{y}}'_{[1,n-1]}-\tilde{x}_m\mathbf{1}^{n-1})^+ +\tilde{\mathbf{S}}_{[1,n-1]}^{t+1}],\beta(\tilde{y}_n-\tilde{x}_m)^+ +\tilde{S}_n^{t+1})\\
\geq&\lambda_1\tilde{V}_{t+1}^g(\alpha(\tilde{\mathbf{x}}_{[1,m-1]}-\tilde{y}_n\mathbf{1}^{m-1})^+ +\tilde{\mathbf{D}}_{[1,m-1]}^{t+1},\alpha(\tilde{x}_m-\tilde{y}_n)^+ +\tilde{D}_m^{t+1},\\
&\qquad\beta(\tilde{\mathbf{y}}_{[1,n-1]}-\tilde{x}_m\mathbf{1}^{n-1})^+ +\tilde{\mathbf{S}}_{[1,n-1]}^{t+1},\beta(\tilde{y}_n-\tilde{x}_m)^+ +\tilde{S}_n^{t+1})\\
&+\lambda_2\tilde{V}_{t+1}^g(\alpha(\tilde{\mathbf{x}}'_{[1,m-1]}-\tilde{y}_n\mathbf{1}^{m-1})^+ +\tilde{\mathbf{D}}_{[1,m-1]}^{t+1},\alpha(\tilde{x}_m-\tilde{y}_n)^+ +\tilde{D}_m^{t+1},\\
&\qquad\beta(\tilde{\mathbf{y}}'_{[1,n-1]}-\tilde{x}_m\mathbf{1}^{n-1})^+ +\tilde{\mathbf{S}}_{[1,n-1]}^{t+1},\beta(\tilde{y}_n-\tilde{x}_m)^+ +\tilde{S}_n^{t+1})
\end{aligned}$$

where the last inequality follows from the induction hypothesis of the concavity of $\tilde{V}_{t+1}^g$ with respect to $(\tilde{\mathbf{x}}_{[1,m-1]},\tilde{\mathbf{y}}_{[1,n-1]})$. This proves that $E\tilde{V}_{t+1}^g(\alpha(\tilde{\mathbf{x}}-\tilde{y}_n\mathbf{1}^m)^+ +\tilde{\mathbf{D}}^{t+1},\beta(\tilde{\mathbf{y}}-\tilde{x}_m\mathbf{1}^n)^+ +\tilde{\mathbf{S}}^{t+1})$ is concave in $(\tilde{\mathbf{x}}_{[1,m-1]},\tilde{\mathbf{y}}_{[1,n-1]})$, and therefore completes the induction. □

LEMMA B.11. *Suppose that $\alpha = \beta$. For the given transformed state $(\tilde{\mathbf{x}}, \tilde{\mathbf{y}})$, the function $\tilde{G}^g_t(Q, \tilde{\mathbf{x}}, \tilde{\mathbf{y}})$ defined in (B.18) is concave with respect to the total matching quantity Q within its feasible range $0 \le Q \le \tilde{x}_m \wedge \tilde{y}_n := \min\{\tilde{x}_m, \tilde{y}_n\}$*

Proof of Lemma B.11. Suppose that $Q \le \tilde{x}_m \wedge \tilde{y}_n$. Let $\tilde{\mathbf{x}}^{t+1} := \alpha(\tilde{\mathbf{x}} - Q\mathbf{1}^m)^+ + \tilde{\mathbf{D}}^{t+1}$ and $\tilde{\mathbf{y}}^{t+1} := \alpha(\tilde{\mathbf{y}} - Q\mathbf{1}^n)^+ + \tilde{\mathbf{S}}^{t+1}$. We have

$$
\begin{aligned}
&\tilde{V}^g_{t+1}(\tilde{\mathbf{x}}^{t+1}, \tilde{\mathbf{y}}^{t+1}) \\
=&\gamma E\mathbf{D}^{t+2}(\mathbf{r}_d^{t+2})^\intercal + \gamma E\mathbf{S}^{t+2}(\mathbf{r}_s^{t+2})^\intercal \\
&-(\tilde{\mathbf{x}}^{t+1} - \tilde{y}_n^{t+1}\mathbf{1}^m)^+\mathbf{U}_m^{-1}(\mathbf{r}_d^{t+1} - \gamma\alpha\mathbf{r}_d^{t+2})^\intercal - (\tilde{\mathbf{y}}^{t+1} - \tilde{x}_m^{t+1}\mathbf{1}^n)^+\mathbf{V}_n^{-1}(\mathbf{r}_s^{t+1} - \gamma\beta\mathbf{r}_s^{t+2})^\intercal \\
&+\gamma E\tilde{V}^g_{t+2}(\alpha(\tilde{\mathbf{x}}^{t+1} - \tilde{y}_n^{t+1}\mathbf{1}^m)^+ + \mathbf{D}^{t+2}\mathbf{U}_m, \beta(\tilde{\mathbf{y}}^{t+1} - \tilde{x}_m^{t+1}\mathbf{1}^n)^+ + \mathbf{S}^{t+2}\mathbf{V}_n).
\end{aligned} \tag{B.23}
$$

The first two terms in (B.23) are constants. We examine each of the remaining terms.

For $i = 1, \dots, m-1$, we have

$$
\begin{aligned}
(\tilde{x}_i^{t+1} - \tilde{y}_n^{t+1})^+ &= \left[\alpha(\tilde{x}_i - Q)^+ - \alpha(\tilde{y}_n - Q) + \tilde{D}_i^{t+1} - \tilde{S}_n^{t+1}\right]^+ \\
&= \begin{cases} \left[\alpha(\tilde{x}_i - \tilde{y}_n) + \tilde{D}_i^{t+1} - \tilde{S}_n^{t+1}\right]^+ & \text{if } Q \le \tilde{x}_i, \\ \left[\alpha(Q - \tilde{y}_n) + \tilde{D}_i^{t+1} - \tilde{S}_n^{t+1}\right]^+ & \text{if } Q > \tilde{x}_i. \end{cases}
\end{aligned}
$$

Thus, $(\tilde{x}_i^{t+1} - \tilde{y}_n^{t+1})^+$ is increasing and convex with respect to Q.

For $i = m$, we have

$$
(\tilde{x}_i^{t+1} - \tilde{y}_n^{t+1})^+ = (\tilde{x}_m^{t+1} - \tilde{y}_n^{t+1})^+ = \left[\alpha(\tilde{x}_m - Q) - \alpha(\tilde{y}_n - Q) + \tilde{D}_i^{t+1} - \tilde{S}_n^{t+1}\right]^+ = \left[\alpha(\tilde{x}_m - \tilde{y}_n) + \tilde{D}_m^{t+1} - \tilde{S}_n^{t+1}\right]^+,
$$

which is constant with respect to Q.

Likewise, $(\tilde{y}_j^{t+1} - \tilde{x}_m^{t+1})^+$ is increasing and convex in Q for $j = 1, \dots, n-1$, and is constant (more specifically, equal to $\left[\alpha(\tilde{y}_n - \tilde{x}_m) + \tilde{S}_m^{t+1} - \tilde{D}_n^{t+1}\right]^+$) in Q for $j = n$.

The third term in (B.23) can be written as:

$$
-(\tilde{\mathbf{x}}^{t+1} - \tilde{y}_n^{t+1}\mathbf{1}^m)^+\mathbf{U}_m^{-1}(\mathbf{r}_d^{t+1} - \gamma\alpha\mathbf{r}_d^{t+2})^\intercal = -\sum_{i=1}^{m}(\tilde{x}_i^{t+1} - \tilde{y}_n^{t+1})^+\left[(r_{id}^{t+1} - r_{i+1,d}^{t+1}) - \gamma\alpha(r_{id}^{t+2} - r_{i+1,d}^{t+2})\right],
$$

which is decreasing and concave in Q, if Assumption 3 holds.

Symmetrically, the fourth term in (B.23), $(\tilde{\mathbf{y}}^{t+1} - \tilde{x}_m^{t+1}\mathbf{1}^n)^+\mathbf{V}_n^{-1}(\mathbf{r}_s^{t+1} - \gamma\beta\mathbf{r}_s^{t+2})^\intercal$, is also decreasing and concave in Q, if Assumption 3 holds.

We now examine the last term in (B.23). We have

$$\begin{aligned}
&\tilde{V}^g_{t+2}(\alpha(\tilde{\mathbf{x}}^{t+1}-\tilde{y}^{t+1}_n\mathbf{1}^m)^+ + \tilde{\mathbf{D}}^{t+2}, \beta(\tilde{\mathbf{y}}^{t+1}-\tilde{x}^{t+1}_m\mathbf{1}^n)^+ + \tilde{\mathbf{S}}^{t+2})\\
=&\tilde{V}^g_{t+2}(\alpha(\tilde{\mathbf{x}}^{t+1}_{[1,m-1]}-\tilde{y}^{t+1}_n\mathbf{1}^{m-1})^+ + \tilde{\mathbf{D}}^{t+2}_{[1,m-1]}, \alpha(\tilde{x}^{t+1}_m-\tilde{y}^{t+1}_n)^+ + \tilde{D}^{t+2}_m,\\
&\qquad \alpha(\tilde{\mathbf{y}}^{t+1}_{[1,n-1]}-\tilde{x}^{t+1}_m\mathbf{1}^{n-1})^+ + \tilde{\mathbf{S}}^{t+2}_{[1,n-1]}, \alpha(\tilde{y}^{t+1}_n-\tilde{x}^{t+1}_m) + \tilde{S}^{t+2}_n)\\
=&\tilde{V}^g_{t+2}(\alpha(\tilde{\mathbf{x}}^{t+1}_{[1,m-1]}-\tilde{y}^{t+1}_n\mathbf{1}^{m-1})^+ + \tilde{\mathbf{D}}^{t+2}_{[1,m-1]}, \alpha[\alpha(\tilde{x}_m-\tilde{y}_n)+\tilde{D}^{t+1}_m-\tilde{S}^{t+1}_n]^+ + \tilde{D}^{t+2}_m,\\
&\qquad \alpha(\tilde{\mathbf{y}}^{t+1}_{[1,n-1]}-\tilde{x}^{t+1}_m\mathbf{1}^{n-1})^+ + \tilde{\mathbf{S}}^{t+2}_{[1,n-1]}, \alpha[\alpha(\tilde{y}_n-\tilde{x}_m)+\tilde{S}^{t+1}_n-\tilde{D}^{t+1}_m]^+ + \tilde{S}^{t+2}_n).
\end{aligned}$$

Let $\tilde{\mathbf{u}}_{[1,m-1]} := (\tilde{\mathbf{x}}^{t+1}_{[1,m-1]}-\tilde{y}^{t+1}_n\mathbf{1}^{m-1})^+$ and $\tilde{\mathbf{v}}_{[1,n-1]} := (\tilde{\mathbf{y}}^{t+1}_{[1,n-1]}-\tilde{x}^{t+1}_m\mathbf{1}^{n-1})^+$. Then, $\tilde{V}^g_{t+2}(\alpha(\tilde{\mathbf{x}}^{t+1}-\tilde{y}^{t+1}_n\mathbf{1}^m)^+ + \tilde{\mathbf{D}}^{t+2}, \beta(\tilde{\mathbf{y}}^{t+1}-\tilde{x}^{t+1}_m\mathbf{1}^n)^+ + \tilde{\mathbf{S}}^{t+2})$ can be rewritten as:

$$\begin{aligned}
&\tilde{V}^g_{t+2}(\alpha\tilde{\mathbf{u}}_{[1,m-1]} + \tilde{\mathbf{D}}^{t+2}_{[1,m-1]}, \alpha[\alpha(\tilde{x}_m-\tilde{y}_n)+\tilde{D}^{t+1}_m-\tilde{S}^{t+1}_n]^+ + \tilde{D}^{t+1}_m,\\
&\qquad \alpha\tilde{\mathbf{v}}_{[1,n-1]} + \tilde{\mathbf{S}}^{t+2}_{[1,n-1]}, \alpha[\alpha(\tilde{y}_n-\tilde{x}_m)+\tilde{S}^{t+1}_n-\tilde{D}^{t+1}_m]^+ + \tilde{S}^{t+2}_n).
\end{aligned}$$

By Lemmas B.9 and B.10, the above function is decreasing and concave in $(\tilde{\mathbf{u}}_{[1,m-1]}, \tilde{\mathbf{v}}_{[1,n-1]})$. Earlier, we have shown that both $\tilde{\mathbf{u}}_{[1,m-1]} = (\tilde{\mathbf{x}}^{t+1}_{[1,m-1]}-\tilde{y}^{t+1}_n\mathbf{1}^{m-1})^+$ and $\tilde{\mathbf{v}}_{[1,n-1]} = (\tilde{\mathbf{y}}^{t+1}_{[1,n-1]}-\tilde{x}^{t+1}_m\mathbf{1}^{n-1})^+$ are increasing and convex in Q. We know that the composition of a decreasing concave function and an increasing convex function is still decreasing and concave. Thus, the function $\tilde{V}^g_{t+2}(\alpha(\tilde{\mathbf{x}}^{t+1}-\tilde{y}^{t+1}_n\mathbf{1}^m)^+ + \tilde{\mathbf{D}}^{t+2}, \beta(\tilde{\mathbf{y}}^{t+1}-\tilde{x}^{t+1}_m\mathbf{1}^n)^+ + \tilde{\mathbf{S}}^{t+2})$ is decreasing and concave in Q. Thus, the last term in (B.23) is also concave. It follows that $\tilde{V}^g_{t+1}(\tilde{\mathbf{x}}^{t+1}, \tilde{\mathbf{y}}^{t+1}) = \tilde{V}^g_{t+1}(\alpha(\tilde{\mathbf{x}}-Q\mathbf{1}^m)^+ + \tilde{\mathbf{D}}^{t+1}, \beta(\tilde{\mathbf{y}}-Q\mathbf{1}^n)^+ + \tilde{\mathbf{S}}^{t+1})$ is concave in Q.

Finally, we show that the function

$$\begin{aligned}
\tilde{G}^g_t(Q,\tilde{\mathbf{x}},\tilde{\mathbf{y}}) =& -(\tilde{\mathbf{x}}-Q\mathbf{1}^m)^+\mathbf{U}^{-1}_m(\mathbf{r}^t_d-\alpha\mathbf{r}^{t+1}_d)^\intercal - (\tilde{\mathbf{y}}-Q\mathbf{1}^m)^+\mathbf{U}^{-1}_n(\mathbf{r}^t_s-\beta\mathbf{r}^{t+1}_s)^\intercal\\
&+\gamma E\tilde{V}^g_{t+1}(\alpha(\tilde{\mathbf{x}}-Q\mathbf{1}^m)^+ + \tilde{\mathbf{D}}^{t+1}, \beta(\tilde{\mathbf{y}}-Q\mathbf{1}^n)^+ + \tilde{\mathbf{S}}^{t+1})
\end{aligned}\tag{B.24}$$

is concave in Q. We have just shown that the last term in (B.24) is concave in Q. The term

$$-(\tilde{\mathbf{x}}-Q\mathbf{1}^m)^+\mathbf{U}^{-1}_m(\mathbf{r}^t_d-\alpha\mathbf{r}^{t+1}_d)^\intercal = -\sum_{i=1}^{m}(\tilde{x}^{t+1}_i-Q)^+\left[(r^{t+1}_{id}-r^{t+1}_{i+1,d})-\gamma\alpha(r^{t+2}_{id}-r^{t+2}_{i+1,d}))\right]$$

is concave in Q, and so is the term $-(\tilde{\mathbf{y}}-Q\mathbf{1}^m)^+\mathbf{U}^{-1}_n(\mathbf{r}^t_s-\beta\mathbf{r}^{t+1}_s)^\intercal$ likewise. As a result, $\tilde{G}^g_t(Q,\tilde{\mathbf{x}},\tilde{\mathbf{y}})$ is concave in Q.

We note that even though the proof of the lemma presented above is based on the concavity of $\tilde{V}_t$ and is thus for the continuous-valued model, the function $\tilde{G}_t$ in the discrete-valued model can be naturally extended to its counterpart in the continuous-valued model. Therefore, in the discrete-valued model, $\tilde{G}_t$ is also concave with respect to Q. $\square$

We further present and prove two lemmas.

LEMMA B.12. *Let* $\boldsymbol{\delta} := (\delta_1, \ldots, \delta_{j-1}, 0, \ldots, 0)$ *be* n*-dimension vector such that* $0 \le \delta_1 \le \delta_2 \le \cdots \le \delta_{j-1} \le \varepsilon$*, Among the factors* $(\tilde{\mathbf{x}}, \tilde{\mathbf{y}})$ *and* $\boldsymbol{\delta}$*, the difference* $\tilde{V}_t^g(\tilde{\mathbf{x}}, \tilde{\mathbf{y}} - \boldsymbol{\delta}) - \tilde{V}_t^g(\tilde{\mathbf{x}}, \tilde{\mathbf{y}})$ *depends only on* $\boldsymbol{\delta}$*,* $\tilde{x}_m$ *and* $\tilde{\mathbf{y}}_{[1,j-1]}$*.*

Likewise, for an m*-dimension vector* $\boldsymbol{\theta} := (\theta, \ldots, \theta_{i-1}, 0, \ldots, 0)$ *such that* $0 \le \theta_1 \le \theta_2 \le \cdots \le \theta_{i-1} \le \varepsilon$*, the difference* $\tilde{V}_t^g(\tilde{\mathbf{x}} - \boldsymbol{\theta}, \tilde{\mathbf{y}}) - \tilde{V}_t^g(\tilde{\mathbf{x}}, \tilde{\mathbf{y}})$ *depends only on* $\boldsymbol{\theta}$*,* ε*,* $\tilde{y}_n$ *and* $\tilde{\mathbf{x}}_{[1,i-1]}$*.*

Proof of Lemma B.12. We focus on the first difference $\tilde{V}_t^g(\tilde{\mathbf{x}}, \tilde{\mathbf{y}} - \boldsymbol{\delta}) - \tilde{V}_t^g(\tilde{\mathbf{x}}, \tilde{\mathbf{y}})$, and the other difference will satisfy the desired property by symmetry.

We prove the lemma by induction. Since $\tilde{V}_{T+1}^g(\tilde{\mathbf{x}}, \tilde{\mathbf{y}} - \boldsymbol{\delta}) - \tilde{V}_{T+1}^g(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) \equiv 0$, the difference clearly satisfies the lemma for $t = T+1$. Suppose the desired property is satisfied for $t+1$. We consider two cases for t.

Case 1: $\tilde{x}_m < \tilde{y}_{j-1}$.

In this case, there exists $1 \le j' \le j-1$ such that $\tilde{y}_{j'-1} \le \tilde{x}_m < \tilde{y}_{j'}$. Since $\tilde{x}_m < \tilde{y}_{j-1} \le \tilde{y}_n$, $(\tilde{\mathbf{x}} - \tilde{y}_n \mathbf{1}^m)^+$ reduces to the zero vector. By (B.19) in Lemma B.8, we have

$$
\begin{aligned}
\tilde{V}_t^g(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) =& \gamma E \mathbf{D}^{t+1} (\mathbf{r}_d^{t+1})^\intercal + \gamma E \mathbf{S}^{t+1} (\mathbf{r}_s^{t+1})^\intercal - (\tilde{\mathbf{y}}_{[j',n]} - \tilde{x}_m \mathbf{1}^{n-j'+1})(\mathbf{U}_n^{[j',n]\times[1,n]})^{-1}(\mathbf{r}_s^t - \gamma\beta \mathbf{r}_s^{t+1})^\intercal \\
&+ \gamma E \tilde{V}_{t+1}^g(\tilde{\mathbf{D}}^{t+1}, \beta(\tilde{\mathbf{y}} - \tilde{x}_m \mathbf{1}^n)^+ + \tilde{\mathbf{S}}^{t+1}),
\end{aligned}
$$

where $\mathbf{U}_n^{[j',n]\times[1,n]}$ is a submatrix consist of the j'th to the nth rows and all the columns of U_n.

Also, there exists j'' such that $j' \le j'' \le j-1$ and $\tilde{y}_{j''-1} - \delta_{j''-1} \le \tilde{x}_m < \tilde{y}_{j''} - \delta_{j''}$. It follows that

$$
\begin{aligned}
\tilde{V}_t^g(\tilde{\mathbf{x}}, \tilde{\mathbf{y}} - \boldsymbol{\delta}) =& \gamma E \mathbf{D}^{t+1} (\mathbf{r}_d^{t+1})^\intercal + \gamma E \mathbf{S}^{t+1} (\mathbf{r}_s^{t+1})^\intercal - (\tilde{\mathbf{y}}_{[j'',n]} - \boldsymbol{\delta}_{[j'',n]} - \tilde{x}_m \mathbf{1}^{n-j''+1})(\mathbf{U}_n^{[j'',n]\times[1,n]})^{-1}(\mathbf{r}_s^t - \gamma\beta \mathbf{r}_s^{t+1})^\intercal \\
&+ \gamma E \tilde{V}_{t+1}^g(\tilde{\mathbf{D}}^{t+1}, \beta(\tilde{\mathbf{y}} - \boldsymbol{\delta} - \tilde{x}_m \mathbf{1}^n)^+ + \tilde{\mathbf{S}}^{t+1}).
\end{aligned}
$$

We then have

$$
\begin{aligned}
& \tilde{V}_t^g(\tilde{\mathbf{x}}, \tilde{\mathbf{y}} - \boldsymbol{\delta}) - \tilde{V}_t^g(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) \\
=& - (\tilde{\mathbf{y}}_{[j'',n]} - \boldsymbol{\delta}_{[j'',n]} - \tilde{x}_m \mathbf{1}^{n-j''+1})(\mathbf{U}_n^{[j'',n]\times[1,n]})^{-1}(\mathbf{r}_s^t - \gamma\beta \mathbf{r}_s^{t+1})^\intercal \\
&+ (\tilde{\mathbf{y}}_{[j',n]} - \tilde{x}_m \mathbf{1}^{n-j'+1})(\mathbf{U}_n^{[j',n]\times[1,n]})^{-1}(\mathbf{r}_s^t - \gamma\beta \mathbf{r}_s^{t+1})^\intercal \\
&+ \gamma E \tilde{V}_{t+1}^g(\tilde{\mathbf{D}}^{t+1}, \beta(\tilde{\mathbf{y}} - \boldsymbol{\delta} - \tilde{x}_m \mathbf{1}^n)^+ + \tilde{\mathbf{S}}^{t+1}) - \gamma E \tilde{V}_{t+1}^g(\tilde{\mathbf{D}}^{t+1}, \beta(\tilde{\mathbf{y}} - \tilde{x}_m \mathbf{1}^n)^+ + \tilde{\mathbf{S}}^{t+1})
\end{aligned}
$$

$$
\begin{aligned}
=&-(\tilde{\mathbf{y}}_{[j'',n]}-\boldsymbol{\delta}_{[j'',n]}-\tilde{x}_m\mathbf{1}^{n-j''+1})(\mathbf{U}_n^{[j'',n]\times[1,n]})^{-1}(\mathbf{r}_s^t-\gamma\beta\mathbf{r}_s^{t+1})^{\intercal}\\
&+(\tilde{\mathbf{y}}_{[j'',n]}-\tilde{x}_m\mathbf{1}^{n-j''+1})(\mathbf{U}_n^{[j'',n]\times[1,n]})^{-1}(\mathbf{r}_s^t-\gamma\beta\mathbf{r}_s^{t+1})^{\intercal}\\
&+(\tilde{\mathbf{y}}_{[j',j''-1]}-\tilde{x}_m\mathbf{1}^{j''-j'})(\mathbf{U}_n^{[j',j''-1]\times[1,n]})^{-1}(\mathbf{r}_s^t-\gamma\beta\mathbf{r}_s^{t+1})^{\intercal}\\
&+\gamma E\tilde{V}_{t+1}^g(\tilde{\mathbf{D}}^{t+1},\beta(\tilde{\mathbf{y}}-\boldsymbol{\delta}-\tilde{x}_m\mathbf{1}^n)^++\tilde{\mathbf{S}}^{t+1})-\gamma E\tilde{V}_{t+1}^g(\tilde{\mathbf{D}}^{t+1},\beta(\tilde{\mathbf{y}}-\tilde{x}_m\mathbf{1}^n)^++\tilde{\mathbf{S}}^{t+1})\\
=&\boldsymbol{\delta}_{[j'',n]}(\mathbf{U}_n^{[j'',n]\times[1,n]})^{-1}(\mathbf{r}_s^t-\gamma\beta\mathbf{r}_s^{t+1})^{\intercal}+(\tilde{\mathbf{y}}_{[j',j''-1]}-\tilde{x}_m\mathbf{1}^{j''-j'})(\mathbf{U}_n^{[j',j''-1]\times[1,n]})^{-1}(\mathbf{r}_s^t-\gamma\beta\mathbf{r}_s^{t+1})^{\intercal}\\
&+\gamma E\tilde{V}_{t+1}^g(\tilde{\mathbf{D}}^{t+1},\beta(\tilde{\mathbf{y}}-\boldsymbol{\delta}-\tilde{x}_m\mathbf{1}^n)^++\tilde{\mathbf{S}}^{t+1})-\gamma E\tilde{V}_{t+1}^g(\tilde{\mathbf{D}}^{t+1},\beta(\tilde{\mathbf{y}}-\tilde{x}_m\mathbf{1}^n)^++\tilde{\mathbf{S}}^{t+1}). \qquad \text{(B.25)}
\end{aligned}
$$

Clearly, the first two terms in (B.25) depends only on $\boldsymbol{\delta}$, $\tilde{\mathbf{y}}_{[1,j-1]}$ and $\tilde{x}_m$. It remains to show that the remaining term, $\gamma E\tilde{V}_{t+1}^g(\tilde{\mathbf{D}}^{t+1},\beta(\tilde{\mathbf{y}}-\boldsymbol{\delta}-\tilde{x}_m\mathbf{1}^n)^++\tilde{\mathbf{S}}^{t+1})-\gamma E\tilde{V}_{t+1}^g(\tilde{\mathbf{D}}^{t+1},\beta(\tilde{\mathbf{y}}-\tilde{x}_m\mathbf{1}^n)^++\tilde{\mathbf{S}}^{t+1})$ satisfies the same property.

Let us write $\tilde{\mathbf{Y}}'':=\beta(\tilde{\mathbf{y}}-\boldsymbol{\delta}-\tilde{x}_m\mathbf{1}^n)^++\tilde{\mathbf{S}}^{t+1}$, and $\tilde{\mathbf{Y}}'=\beta(\tilde{\mathbf{y}}-\tilde{x}_m\mathbf{1}^n)^++\tilde{\mathbf{S}}^{t+1}$. We have

$$
\begin{aligned}
\tilde{\mathbf{Y}}''_{[j'',n]}=&\beta(\tilde{\mathbf{y}}_{[j'',n]}-\boldsymbol{\delta}_{[j'',n]}-\tilde{x}_m\mathbf{1}^{n-j''+1})+\tilde{\mathbf{S}}^{t+1}\\
=&\beta(\tilde{\mathbf{y}}_{[j'',n]}-\tilde{x}_m\mathbf{1}^{n-j''+1})+\tilde{\mathbf{S}}^{t+1}-\beta\boldsymbol{\delta}_{[j'',n]}\\
=&\tilde{\mathbf{Y}}'_{[j'',n]}-\beta\boldsymbol{\delta}_{[j'',n]},\\
\tilde{\mathbf{Y}}''_{[j',j''-1]}=&\mathbf{0}^{j''-j'}=\tilde{\mathbf{Y}}'_{[j',j''-1]}-\tilde{\mathbf{Y}}'_{[j',j''-1]}=\tilde{\mathbf{Y}}'_{[j',j''-1]}-\left(\beta(\tilde{\mathbf{y}}_{[j',j''-1]}-\tilde{x}_m\mathbf{1}^{j''-j'})+\tilde{\mathbf{S}}^{t+1}_{[j',j''-1]}\right),
\end{aligned}
$$

and $\tilde{\mathbf{Y}}''_{[1,j'-1]}=\tilde{\mathbf{Y}}'_{[1,j'-1]}=\mathbf{0}^{j'-1}$. (Note that $\mathbf{0}^k$ represents the k-dimension zero vector.)

Let us denote $\boldsymbol{\mu}:=(\mathbf{0}^{j'-1},\beta(\tilde{\mathbf{y}}_{[j',j''-1]}-\tilde{x}_m\mathbf{1}^{j''-j'})+\tilde{\mathbf{S}}^{t+1}_{[j',j''-1]},\beta\boldsymbol{\delta}_{[j'',n]})$. Since $\boldsymbol{\delta}_{[j,n]}=\mathbf{0}^{n-j+1}$, we have $\boldsymbol{\mu}:=(\mathbf{0}^{j'-1},\beta(\tilde{\mathbf{y}}_{[j',j''-1]}-\tilde{x}_m\mathbf{1}^{j''-j'})+\tilde{\mathbf{S}}^{t+1}_{[j',j''-1]},\beta\boldsymbol{\delta}_{[j'',j-1]},\mathbf{0}^{n-j+1})$. We see that $\tilde{\mathbf{Y}}''=\tilde{\mathbf{Y}}'-\boldsymbol{\mu}$. (We note that $\boldsymbol{\mu}$ is a random vector.) By the induction hypothesis, we have the difference

$$
\begin{aligned}
&\gamma E\tilde{V}_{t+1}^g(\tilde{\mathbf{D}}^{t+1},\beta(\tilde{\mathbf{y}}-\boldsymbol{\delta}-\tilde{x}_m\mathbf{1}^n)^++\tilde{\mathbf{S}}^{t+1})-\gamma E\tilde{V}_{t+1}^g(\tilde{\mathbf{D}}^{t+1},\beta(\tilde{\mathbf{y}}-\tilde{x}_m\mathbf{1}^n)^++\tilde{\mathbf{S}}^{t+1})\\
=&\gamma E\tilde{V}_{t+1}^g(\tilde{\mathbf{D}}^{t+1},\tilde{\mathbf{Y}}'-\boldsymbol{\mu})-\gamma E\tilde{V}_{t+1}^g(\tilde{\mathbf{D}}^{t+1},\tilde{\mathbf{Y}}'),
\end{aligned}
$$

depends only on $\tilde{\mathbf{Y}}'_{[1,j-1]}$ and $\boldsymbol{\mu}$.

Further, $\tilde{\mathbf{Y}}'_{[1,j-1]}=\beta(\tilde{\mathbf{y}}_{[1,j-1]}-\tilde{x}_m\mathbf{1}^{j-1})^++\tilde{\mathbf{S}}^{t+1}_{[1,j-1]}$ depends only on $\tilde{\mathbf{y}}_{[1,j-1]}$ and $\tilde{x}_m$. $\boldsymbol{\mu}$ depends only on $\tilde{\mathbf{y}}_{[1,j-1]}$, $\boldsymbol{\delta}$ and $\tilde{x}_m$. Therefore, the difference $\gamma E\tilde{V}_{t+1}^g(\tilde{\mathbf{D}}^{t+1},\beta(\tilde{\mathbf{y}}-\boldsymbol{\delta}-\tilde{x}_m\mathbf{1}^n)^++\tilde{\mathbf{S}}^{t+1})-\gamma E\tilde{V}_{t+1}^g(\tilde{\mathbf{D}}^{t+1},\beta(\tilde{\mathbf{y}}-\tilde{x}_m\mathbf{1}^n)^++\tilde{\mathbf{S}}^{t+1})$ depends only on $\boldsymbol{\delta}$, $\tilde{\mathbf{y}}_{[1,j-1]}$ and $\tilde{x}_m$.

It then follows that the desired result holds for t.

Case 2: $\tilde{x}_m\geq\tilde{y}_{j-1}$.

Since $\boldsymbol{\delta}_{[j,n]} = \mathbf{0}^{n-j+1}$ (i.e., the $(n-j+1)$-dimension zero vector), we have

$$(\tilde{\mathbf{y}} - \boldsymbol{\delta} - \tilde{x}_m \mathbf{1}^n)^+ = \left(\mathbf{0}^{j-1}, (\tilde{\mathbf{y}}_{[j,n]} - \tilde{x}_m \mathbf{1}^{n-j+1})^+\right) = (\tilde{\mathbf{y}} - \tilde{x}_m \mathbf{1}^n)^+.$$

By (B.19) in Lemma B.8, we have $\tilde{V}_t^g(\tilde{\mathbf{x}}, \tilde{\mathbf{y}} - \boldsymbol{\delta}) - \tilde{V}_t^g(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) = 0$. Therefore, the result holds trivially.

Combining both cases, we conclude that $\tilde{V}_t^g(\tilde{\mathbf{x}}, \tilde{\mathbf{y}} - \boldsymbol{\delta}) - \tilde{V}_t^g(\tilde{\mathbf{x}}, \tilde{\mathbf{y}})$ depends only on $\boldsymbol{\delta}$, $\tilde{\mathbf{y}}_{[1,j-1]}$ and $\tilde{x}_m$. $\square$

LEMMA B.13. *For $\epsilon > 0$ and $t = 1, \ldots, T$, the difference $\tilde{V}_t^g(\tilde{\mathbf{x}} + \epsilon \mathbf{1}^m_{[i,m]}, \mathbf{y} + \epsilon \mathbf{1}^n_{[j,n]}) - \tilde{V}_t^g(\tilde{\mathbf{x}}, \tilde{\mathbf{y}})$ depends only on ϵ and $(\tilde{\mathbf{x}}_{[1,i-1]}, \tilde{\mathbf{y}}_{[1,j-1]}, \tilde{x}_m, \tilde{y}_n)$, where $\mathbf{1}^m_{[i,m]}$ represents an m-dimension vector with the ith to mth entries equal to 1 and all other entries equal to 0 (the meaning of $\mathbf{1}^n_{[j,n]}$ is similar).*

Proof of Lemma B.13. We first focus on the case in which $\tilde{x}_m \le \tilde{y}_n$. Under this condition, we have $(\tilde{\mathbf{x}} - \tilde{y}_n \mathbf{1}^m)^+ = 0$ and $[\tilde{\mathbf{x}} + \epsilon \mathbf{1}^m_{[j,m]} - (\tilde{y}_n + \epsilon)\mathbf{1}^m]^+ \le (\tilde{\mathbf{x}} - \tilde{y}_n \mathbf{1}^m)^+ = 0$. We further discuss two possibilities, namely, $\tilde{x}_m \ge \tilde{y}_{j-1}$ and $\tilde{x}_m < \tilde{y}_{j-1}$.

If $\tilde{x}_m \ge \tilde{y}_{j-1}$, then $(\tilde{\mathbf{y}} - \tilde{x}_m \mathbf{1}^n)^+ = \left(\mathbf{0}^{j-1}, (\tilde{\mathbf{y}}_{[j,n]} - \tilde{x}_m \mathbf{1}^{n-j+1})^+\right)$, and

$$\left[\tilde{\mathbf{y}} + \epsilon \mathbf{1}^n_{[j,n]} - (\tilde{x}_m + \epsilon)\mathbf{1}^n\right]^+ = \left(\mathbf{0}^{j-1}, (\tilde{\mathbf{y}}_{[j,n]} - \tilde{x}_m \mathbf{1}^{n-j+1})^+\right) = (\tilde{\mathbf{y}} - \tilde{x}_m \mathbf{1}^n)^+.$$

Then, by (B.19) in Lemma B.8, we can readily verify that $\tilde{V}_t^g(\tilde{\mathbf{x}} + \epsilon \mathbf{1}^m_{[i,m]}, \mathbf{y} + \epsilon \mathbf{1}^n_{[j,n]}) - \tilde{V}_t^g(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) = 0$, which clearly satisfies the lemma.

If $\tilde{x}_m < \tilde{y}_{j-1}$, there exists j' such that $1 \le j' \le j-1$ and $\tilde{y}_{j'-1} \le \tilde{x}_m < \tilde{y}_{j'}$. We have

$$\begin{aligned}
&(\tilde{\mathbf{y}} - \tilde{x}_m \mathbf{1}^n)^+ = \left(\mathbf{0}^{j'-1}, \tilde{\mathbf{y}}_{[j',j-1]} - \tilde{x}_m \mathbf{1}^{j-j'}, (\tilde{\mathbf{y}}_{[j,n]} - \tilde{x}_m \mathbf{1}^{n-j+1})^+\right) \\
&\left[\tilde{\mathbf{y}} + \epsilon \mathbf{1}^n_{[j,n]} - (\tilde{x}_m + \epsilon)\mathbf{1}^n\right]^+ = \left(\mathbf{0}^{j'-1}, [\tilde{\mathbf{y}}_{[j',j-1]} - (\tilde{x}_m + \epsilon)\mathbf{1}^{j-j'}]^+, (\tilde{\mathbf{y}}_{[j,n]} - \tilde{x}_m \mathbf{1}^{n-j+1})^+\right).
\end{aligned}$$

Let $\boldsymbol{\delta} := (\mathbf{0}^{j'-1}, \tilde{\mathbf{y}}_{[j',j-1]} - \tilde{x}_m \mathbf{1}^{j-j'} - [\tilde{\mathbf{y}}_{[j',j-1]} - (\tilde{x}_m + \epsilon)\mathbf{1}^{j-j'}]^+, \mathbf{0}^{n-j+1})$, which depends only on $\tilde{\mathbf{y}}_{[1,j-1]}$, $\tilde{x}_m$ and ϵ. We have $\left[\tilde{\mathbf{y}} + \epsilon \mathbf{1}^n_{[j,n]} - (\tilde{x}_m + \epsilon)\mathbf{1}^n\right]^+ = (\tilde{\mathbf{y}} - \tilde{x}_m \mathbf{1}^n)^+ - \boldsymbol{\delta}$.

By (B.19) in Lemma B.8, we have

$$\begin{aligned}
&\tilde{V}_t^g(\tilde{\mathbf{x}} + \epsilon \mathbf{1}^m_{[i,m]}, \mathbf{y} + \epsilon \mathbf{1}^n_{[j,n]}) - \tilde{V}_t^g(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) \\
=&\boldsymbol{\delta} \mathbf{U}_n^{-1} (\mathbf{r}_s^t - \gamma \beta \mathbf{r}_s^{t+1})^{\intercal} \\
&+ \gamma E \tilde{V}^g_{t+1}(\tilde{\mathbf{D}}^{t+1}, \beta(\tilde{\mathbf{y}} - \tilde{x}_m \mathbf{1}^n)^+ - \beta \boldsymbol{\delta} + \tilde{\mathbf{S}}^{t+1}) - \gamma E \tilde{V}^g_{t+1}(\tilde{\mathbf{D}}^{t+1}, \beta(\tilde{\mathbf{y}} - \tilde{x}_m \mathbf{1}^n)^+ + \tilde{\mathbf{S}}^{t+1}).
\end{aligned}$$

By Lemma B.12, the difference $\gamma E \tilde{V}^g_{t+1}(\tilde{\mathbf{D}}^{t+1}, \beta(\tilde{\mathbf{y}} - \tilde{x}_m \mathbf{1}^n)^+ - \beta \boldsymbol{\delta} + \tilde{\mathbf{S}}^{t+1}) - \gamma E \tilde{V}^g_{t+1}(\tilde{\mathbf{D}}^{t+1}, \beta(\tilde{\mathbf{y}} - \tilde{x}_m \mathbf{1}^n)^+ + \tilde{\mathbf{S}}^{t+1})$ depends only on $\boldsymbol{\delta}$ and the first $j-1$ entries of $(\tilde{\mathbf{y}} - \tilde{x}_m \mathbf{1}^n)^+$. The first $j-1$ entries of $(\tilde{\mathbf{y}}-\tilde{x}_m\mathbf{1}^n)^+$ depends only on $\tilde{\mathbf{y}}_{[1,j-1]}$ and $\tilde{x}_m$. Consequently, the difference $\gamma E\tilde{V}^g_{t+1}(\tilde{\mathbf{D}}^{t+1},\beta(\tilde{\mathbf{y}}-\tilde{x}_m\mathbf{1}^n)^+-\beta\boldsymbol{\delta}+\tilde{\mathbf{S}}^{t+1})-\gamma E\tilde{V}^g_{t+1}(\tilde{\mathbf{D}}^{t+1},\beta(\tilde{\mathbf{y}}-\tilde{x}_m\mathbf{1}^n)^++\tilde{\mathbf{S}}^{t+1})$ depends only on $\tilde{\mathbf{y}}_{[1,j-1]}$, $\tilde{x}_m$ and ϵ. Thus, the difference $\tilde{V}^g_t(\tilde{\mathbf{x}}+\epsilon\mathbf{1}^m_{[i,m]},\mathbf{y}+\epsilon\mathbf{1}^n_{[j,n]})-\tilde{V}^g_t(\tilde{\mathbf{x}},\tilde{\mathbf{y}})$ depends only on $\tilde{\mathbf{y}}_{[1,j-1]}$, $\tilde{x}_m$ and ϵ.

Now we consider the case with $\tilde{x}_m>\tilde{y}_n$. By symmetry, we can infer that difference $\tilde{V}^g_t(\tilde{\mathbf{x}}+\epsilon\mathbf{1}^m_{[i,m]},\mathbf{y}+\epsilon\mathbf{1}^n_{[j,n]})-\tilde{V}^g_t(\tilde{\mathbf{x}},\tilde{\mathbf{y}})$ depends only on $\tilde{\mathbf{x}}_{[1,i-1]}$, $\tilde{y}_n$ and ϵ. This completes the proof. □

C. Extensions and Additional Results

C.1. Modified Monge condition for non-neighboring pairs

As mentioned in the paper, we can extend Definition 2 to allow a pair to weakly precede a non-neighboring pair.

Definition C.1. For two non-neighboring pairs (i,j) and (i',j'), we say that (i,j) weakly precedes (i',j') if there exists a path of demand-supply pairs that connects the two, in one of the following forms,

(i) $(i,j)=(i_1,j_1)\to(i_2,j_1)\to(i_2,j_2)\to\cdots(i_N,j_{N-1})\to(i_N,j_N)=(i,j)$; or

(ii) $(i,j)=(i_1,j_1)\to(i_2,j_1)\to(i_2,j_2)\to\cdots(i_N,j_{N-1})\to(i_N,j_N)\to(i_{N+1},j_N)=(i,j)$; or

(iii) $(i,j)=(i_1,j_1)\to(i_1,j_2)\to(i_3,j_2)\to\cdots(i_N,j_{N-1})\to(i_N,j_N)=(i,j)$; or

(iv) $(i,j)=(i_1,j_1)\to(i_1,j_2)\to(i_3,j_2)\to\cdots(i_N,j_{N-1})\to(i_N,j_N)\to(i_{N+1},j_N)=(i,j)$;

such that along this path, each pair weakly precedes the next pair according to Definition 2.

With Definition C.1, we first prove the following Lemma.

Lemma C.14. *For a given state $(\mathbf{x},\mathbf{y})$ with $x_i>0$ and $y_j>0$, $\epsilon^1_t\in[0,x_i]$ and $\epsilon^2_t\in[0,y_j]$, there exist $\eta^\tau_{j''}\geq 0$ and $\xi^\tau_{i''}\geq 0$ for $j''\in\mathcal{S}$, $i''\in\mathcal{D}$ and $\tau=t,\dots,T+1$ such that $\sum_{\tau=t}^T\alpha^{-(\tau-t)}\sum_{j''=1}^n\eta^\tau_{j''}\leq\epsilon^1_t$, $\sum_{\tau=t}^T\beta^{-(\tau-t)}\sum_{i''\in\mathcal{D}}\xi^\tau_{i''}\leq\epsilon^2_t$ and*

$$V_t(\mathbf{x}-\epsilon^1_t\mathbf{e}^m_i+\epsilon^1_t\mathbf{e}^m_{i'},\mathbf{y}-\epsilon^2_t\mathbf{e}^n_j+\epsilon^2_t\mathbf{e}^n_{j'})-V_t(\mathbf{x},\mathbf{y})\geq\sum_{\tau=t}^T\gamma^{\tau-t}[\sum_{j''=1}^n\eta^\tau_{j''}(r^\tau_{i'j''}-r^\tau_{ij''})+\sum_{i''=1}^m\xi^\tau_{i''}(r^\tau_{i''j'}-r^\tau_{i''j})].$$

Proof of Lemma C.14. We prove the lemma by induction. The result holds trivially for $t=T+1$ by noting that $V_{T+1}(\mathbf{x},\mathbf{y})\equiv 0$. Suppose the result holds for $t+1$.

For period t, we let $\hat{\mathbf{Q}}\in\arg\max_{\mathbf{Q}\in\{\mathbf{Q}\geq\mathbf{0}|\mathbf{u}\geq\mathbf{0},\mathbf{v}\geq\mathbf{0}\}}H_t(\mathbf{Q},\mathbf{x},\mathbf{y})$. For the state $(\mathbf{x}-\epsilon^1_t\mathbf{e}^m_i+\epsilon^1_t\mathbf{e}^m_{i'},\mathbf{y}-\epsilon^2_t\mathbf{e}^n_j+\epsilon^2_t\mathbf{e}^n_{j'})$, $\hat{\mathbf{Q}}$ may not be feasible, since less type i demand and type j supply are available. Based on $\hat{\mathbf{Q}}$, we can construct a feasible decision for $(\mathbf{x}-\epsilon^1_t\mathbf{e}^m_i+\epsilon^1_t\mathbf{e}^m_{i'},\mathbf{y}-\epsilon^2_t\mathbf{e}^n_j+\epsilon^2_t\mathbf{e}^n_{j'})$.

Specifically, we can reduce the matching quantities between type i demand and all supply types by a total amount $[\sum_{j''=1}^n\hat{q}_{ij'}-(x_i-\varepsilon^1_t)]^+$ (which is by how much the decision $\hat{\mathbf{Q}}$ consumes more than the available type i demand $x_i - \varepsilon_t^1$), and reduce the matching quantities between type j supply and demand types by a total amount $[\sum_{i''=1}^m \hat{q}_{i'j} - (y_j - \varepsilon_t^2)]^+$ (which is by how much the decision $\hat{\mathbf{Q}}$ consumes more than the available type j supply $y_j - \varepsilon_t^2$). In other words, there exists $\eta_{j''}^t \geq 0$ and $\xi_{i''}^t \geq 0$ for $i'' = 1, \ldots, m$ and $j'' = 1, \ldots, m$ such that $\sum_{j''=1}^n \eta_{j''}^t \leq \varepsilon_t^1$, $\sum_{i''=1}^m \xi_{i''}^t \leq \varepsilon_t^2$, and the matching quantities $q_{ij''} - \eta_{j''}^t$ $(j'' = 1, \ldots, n)$ and $q_{i''j} - \xi_{i''}^t$ $(i'' = 1, \ldots, m)$ are feasible for the state $(\mathbf{x} - \epsilon_t^1 \mathbf{e}_i^m + \epsilon_t^1 \mathbf{e}_{i''}^m, \mathbf{y} - \epsilon_t^2 \mathbf{e}_j^n + \epsilon_t^2 \mathbf{e}_{j''}^n)$.

In the meantime, we also increase the matching quantity between type i' demand and type j'' supply by $\eta_{j''}^t$ and the matching quantity between type j' supply and type i'' demand by $\xi_{i''}^t$, for all i'' and j''. This leads to the matching decision $\widetilde{\mathbf{Q}} := \hat{\mathbf{Q}} - \sum_{j''=1}^n \eta_{j''}^t \mathbf{e}_{ij''}^{m\times n} - \sum_{i''=1}^m \xi_{i''}^t \mathbf{e}_{i''j}^{m\times n} + \sum_{j''=1}^n \eta_{j''}^t \mathbf{e}_{i'j''}^{m\times n} + \sum_{i''=1}^m \xi_{i''}^t \mathbf{e}_{i''j'}^{m\times n}$, which is feasible for the state $(\mathbf{x} - \epsilon_t^1 \mathbf{e}_i^m + \epsilon_t^1 \mathbf{e}_{i''}^m, \mathbf{y} - \epsilon_t^2 \mathbf{e}_j^n + \epsilon_t^2 \mathbf{e}_{j''}^n)$.

Let us denote by $\mathbf{u}$ and $\mathbf{v}$ the post-matching levels under the state $(\mathbf{x}, \mathbf{y})$ and the decision $\hat{\mathbf{Q}}$ in period t. Define $\epsilon_{t+1}^1 = \alpha(\epsilon_t^1 - \sum_{j''=1}^n \eta_{j''}^t)$ and $\epsilon_{t+1}^2 = \beta(\epsilon_t^2 - \sum_{i''=1}^m \xi_{i''}^t)$. We have:

$$
\begin{aligned}
&V_t(\mathbf{x} - \epsilon_t^1 \mathbf{e}_i^m + \epsilon_t^1 \mathbf{e}_{i'}^m, \mathbf{y} - \epsilon_t^2 \mathbf{e}_j^n + \epsilon_t^2 \mathbf{e}_{j'}^n) - V_t(\mathbf{x}, \mathbf{y}) \\
\geq & H_t(\widetilde{\mathbf{Q}}, \mathbf{x} - \epsilon_t^1 \mathbf{e}_i^m + \epsilon_t^1 \mathbf{e}_{i'}^m, \mathbf{y} - \epsilon_t^2 \mathbf{e}_j^n + \epsilon_t^2 \mathbf{e}_{j'}^n) - H_t(\hat{\mathbf{Q}}, \mathbf{x}, \mathbf{y}) \\
\geq & \sum_{i''=1}^m \xi_{i''}^t (r_{i''j'}^t - r_{i''j}^t) + \sum_{j''=1}^n \eta_{j''}^t (r_{i'j''}^t - r_{ij''}^t) \\
& + \gamma E V_{t+1}(\alpha[\mathbf{u} - (\epsilon_t^1 - \sum_{j''=1}^n \eta_{j''}^t) \mathbf{e}_i^m + (\epsilon_t^1 - \sum_{j''=1}^n \eta_{j''}^t) \mathbf{e}_{i'}^m] + \mathbf{D}^{t+1}, \\
& \qquad \beta[\mathbf{v} - (\epsilon_t^2 - \sum_{i''=1}^m \xi_{i''}^t) \mathbf{e}_j^n + (\epsilon_t^2 - \sum_{i''=1}^m \xi_{i''}^t) \mathbf{e}_{j'}^n] + \mathbf{S}^{t+1}) - \gamma E V_{t+1}(\alpha \mathbf{u} + \mathbf{D}^{t+1}, \beta \mathbf{v} + \mathbf{S}^{t+1}) \\
= & \sum_{i''=1}^m \xi_{i''}^t (r_{i''j'}^t - r_{i''j}^t) + \sum_{j''=1}^n \eta_{j''}^t (r_{i'j''}^t - r_{ij''}^t) \\
& + \gamma E V_{t+1}(\alpha \mathbf{u} - \epsilon_{t+1}^1 \mathbf{e}_i^n + \epsilon_{t+1}^1 \mathbf{e}_{i'}^n + \mathbf{D}^{t+1}, \beta \mathbf{v} - \epsilon_{t+1}^2 \mathbf{e}_j^m + \epsilon_{t+1}^2 \mathbf{e}_{j'}^m + \mathbf{S}^{t+1}) - \gamma E V_{t+1}(\alpha \mathbf{u} + \mathbf{D}^{t+1}, \beta \mathbf{v} + \mathbf{S}^{t+1}).
\end{aligned}
\tag{C.26}
$$

Let $\mathbf{X}_{t+1} = \alpha \mathbf{u} + \mathbf{D}^{t+1}$ and $\mathbf{Y}_{t+1} = \beta \mathbf{v} + \mathbf{S}^{t+1}$. By the induction hypothesis, there exist $K_{j''}^\tau$ and $L_{i''}^\tau$ for $j'' = 1, \ldots, n$, $i'' = 1, \ldots, m$ and $\tau = t+1, \ldots, T$ such that $\sum_{\tau=t+1}^T \alpha^{-(\tau-t-1)} \sum_{j''=1}^n K_{j''}^\tau \leq \epsilon_{t+1}^1$, $\sum_{\tau=t+1}^T \beta^{-(\tau-t-1)} \sum_{i''=1}^m L_{i''}^\tau \leq \epsilon_{t+1}^2$ and

$$
\begin{aligned}
& V_{t+1}(\mathbf{X}_{t+1} - \epsilon_{t+1}^1 \mathbf{e}_i^m + \epsilon_{t+1}^1 \mathbf{e}_{i'}^m, \mathbf{Y}_{t+1} - \epsilon_{t+1}^2 \mathbf{e}_j^n + \epsilon_{t+1}^2 \mathbf{e}_{j'}^n) - V_{t+1}(\mathbf{X}_{t+1}, \mathbf{Y}_{t+1}) \\
\geq & \sum_{\tau=t+1}^T \gamma^{\tau-t-1} [\sum_{j''=1}^n K_{j''}^\tau (r_{i'j''}^\tau - r_{ij''}^\tau) + \sum_{i''=1}^m L_{i''}^\tau (r_{i''j'}^\tau - r_{i''j}^\tau)].
\end{aligned}
\tag{C.27}
$$

Let $\eta_{j''}^\tau = E K_{j''}^\tau$ and $\xi_{i''}^\tau = E L_{i''}^\tau$ for $j'' = 1, \ldots, n$, $i'' = 1, \ldots, m$ and $\tau = t+1, \ldots, T$. We have

$$
\begin{aligned}
&V_t(\mathbf{x}-\epsilon_t^1\mathbf{e}_i^m+\epsilon_t^1\mathbf{e}_{i'}^m,\mathbf{y}-\epsilon_t^2\mathbf{e}_j^n+\epsilon_t^2\mathbf{e}_{j'}^n)-V_t(\mathbf{x},\mathbf{y})\\
\geq&\sum_{i''=1}^m\xi_{i''}^t(r_{i''j'}^t-r_{i''j}^t)+\sum_{j''=1}^n\eta_{j''}^t(r_{i'j''}^t-r_{ij''}^t)\\
&+\gamma EV_{t+1}(\alpha\mathbf{u}-\epsilon_{t+1}^1\mathbf{e}_i^m+\epsilon_{t+1}^1\mathbf{e}_{i'}^m+\mathbf{D}^{t+1},\beta\mathbf{v}-\epsilon_{t+1}^2\mathbf{e}_j^n+\epsilon_{t+1}^2\mathbf{e}_{j'}^n+\mathbf{S}^{t+1})-\gamma EV_{t+1}(\alpha\mathbf{u}+\mathbf{D}^{t+1},\beta\mathbf{v}+\mathbf{S}^{t+1})\\
\geq&\sum_{i''=1}^m\xi_{i''}^t(r_{i''j'}^t-r_{i''j}^t)+\sum_{j''=1}^n\eta_{j''}^t(r_{i'j''}^t-r_{ij''}^t)+\sum_{\tau=t+1}^T\gamma^{\tau-t}[\sum_{j''=1}^n\eta_{j''}^\tau(r_{i'j''}^\tau-r_{ij''}^\tau)+\sum_{i''=1}^m\xi_{i''}^\tau(r_{i''j'}^\tau-r_{i''j}^\tau)]\\
=&\sum_{\tau=t}^T\gamma^{\tau-t}[\sum_{j''=1}^n\eta_{j''}^\tau(r_{i'j''}^\tau-r_{ij''}^\tau)+\sum_{i''=1}^m\xi_{i''}^\tau(r_{i''j'}^\tau-r_{i''j}^\tau)],
\end{aligned}
$$

where the first inequality is (C.26) and the second inequality is due to (C.27).

Moreover, $\sum_{\tau=t}^T\alpha^{-(\tau-t)}\sum_{j''=1}^n\eta_{j''}^\tau=\sum_{j''=1}^n\eta_{j''}^t+E\sum_{\tau=t+1}^T\alpha^{-(\tau-t)}\sum_{j''=1}^n K_{j''}^\tau\leq\sum_{j''=1}^n\eta_{j''}^t+\alpha^{-1}\epsilon_{t+1}^1=\sum_{j''=1}^n\eta_{j''}^t+(\epsilon_t^1-\sum_{j''=1}^n\eta_{j''}^t)=\epsilon_t^1$. Similarly, $\sum_{\tau=t}^T\beta^{-(\tau-t)}\sum_{i''=1}^m\xi_{i''}^\tau=\sum_{i''=1}^m\xi_{i''}^t+E\sum_{\tau=t+1}^T\beta^{-(\tau-t)}\sum_{i''=1}^m L_{i''}^\tau=\sum_{i''=1}^m\xi_{i''}^t+\beta^{-1}\epsilon_{t+1}^2\leq\sum_{i''=1}^m\xi_{i''}^t+(\epsilon_t^2-\sum_{i''=1}^m\xi_{i''}^t)=\epsilon_t^2$. This completes the induction. □

Lemma C.15. *For a state* $(\mathbf{x},\mathbf{y})$*, any* $0<\epsilon_1<x_i$*,* $0<\epsilon_2<y_j$*, we have*

$$
V_{t+1}(\mathbf{x}-\epsilon_1\mathbf{e}_i^m+\epsilon_1\mathbf{e}_{i'}^m,\mathbf{y}-\epsilon_2\mathbf{e}_j^n+\epsilon_2\mathbf{e}_{j'}^n)-V_{t+1}(\mathbf{x},\mathbf{y})\geq\max\left\{(\alpha\gamma)^{-1}\epsilon_1,(\beta\gamma)^{-1}\epsilon_2\right\}(r_{i'j'}^t-r_{ij}^t).
$$

Proof of Lemma C.15. By Lemma C.14, there exist nonnegative numbers $\eta_{j''}^\tau$ and $\xi_{i''}^\tau$ for $j''=1,\ldots,n$, $i''=1,\ldots,m$ and $\tau=t+1,\ldots,T+1$ such that $\sum_{\tau=t+1}^T\alpha^{-(\tau-t-1)}\sum_{j''=1}^n\eta_{j''}^\tau\leq\epsilon_1$, $\sum_{\tau=t}^T\beta^{-(\tau-t-1)}\sum_{i'=1}^m\xi_{i''}^\tau\leq\epsilon_2$ and

$$
\begin{aligned}
&V_{t+1}(\mathbf{x}-\epsilon_1\mathbf{e}_i^m+\epsilon_1\mathbf{e}_{i'}^m,\mathbf{y}-\epsilon_2\mathbf{e}_j^m+\epsilon_2\mathbf{e}_{j'}^m)-V_t(\mathbf{x},\mathbf{y})\\
\geq&\sum_{\tau=t+1}^T\gamma^{\tau-t-1}[\sum_{j''=1}^n\eta_{j''}^\tau(r_{i'j''}^\tau-r_{ij''}^\tau)+\sum_{i''=1}^m\xi_{i''}^\tau(r_{i''j'}^\tau-r_{i''j}^\tau)].
\end{aligned}
\tag{C.28}
$$

Since (i,j) weakly precedes (i',j'), there exists a "zigzag" path connecting the two arcs, along which each pair weakly precedes the next pair. Without loss of generality, we suppose that the path has the form $(i,j)=(i_1,j_1)\to(i_2,j_1)\to(i_2,j_2)\to\cdots\to(i_\ell,j_\ell)=(i',j')$ (proof for the other forms would be analogous). Along the path, (i_k,j_k) weakly precedes (i_{k+1},j_k), which implies that $r_{i_kj_k}^t-r_{i_{k+1}j_k}^t\geq(\gamma\alpha)^{\tau-t}(r_{i_kj''}^\tau-r_{i_{k+1}j''}^\tau)$. Thus

$$
r_{i'j''}^\tau-r_{ij''}^\tau=-(r_{i_1j''}^\tau-r_{i_\ell j''}^\tau)=-\sum_{k=1}^{\ell-1}(r_{i_kj''}^\tau-r_{i_{k+1}j''}^\tau)\geq-(\gamma\alpha)^{-(\tau-t))}\sum_{k=1}^{\ell-1}(r_{i_kj_k}^t-r_{i_{k+1}j_k}^t).
\tag{C.29}
$$

Likewise, since (i_{k+1},j_k) weakly precedes (i_{k+1},j_{k+1}), we have

$$r^{\tau}_{i''j'} - r^{\tau}_{i''j} = -(r^{\tau}_{i''j_1} - r^{\tau}_{i''j_\ell}) = -\sum_{k=1}^{\ell-1}(r^{\tau}_{i''j_k} - r^{\tau}_{i''j_{k+1}}) \geq -(\gamma\beta)^{-(\tau-t)}\sum_{k=1}^{\ell-1}(r^t_{i_{k+1}j_k} - r^t_{i_{k+1}j_{k+1}}). \quad \text{(C.30)}$$

Then,

$$\begin{aligned}
&V_{t+1}(\mathbf{x}-\epsilon_1\mathbf{e}^m_i+\epsilon_1\mathbf{e}^m_{i'},\mathbf{y}-\epsilon_2\mathbf{e}^n_j+\epsilon_2\mathbf{e}^n_{j'})-V_{t+1}(\mathbf{x},\mathbf{y})\\
\geq&\sum_{\tau=t+1}^{T}\gamma^{\tau-t-1}(\gamma\alpha)^{-(\tau-t)}\sum_{j''=1}^{n}\eta^{\tau}_{j''}\sum_{k=1}^{\ell-1}(r^t_{i_{k+1}j_k}-r^t_{i_kj_k})+\sum_{\tau=t+1}^{T}\gamma^{\tau-t-1}(\gamma\beta)^{-(\tau-t)}\sum_{i''=1}^{m}\xi^{\tau}_{i''}\sum_{k=1}^{\ell-1}(r^t_{i_{k+1}j_{k+1}}-r^t_{i_{k+1}j_k})\\
=&\gamma^{-1}\sum_{\tau=t+1}^{T}\alpha^{-(\tau-t)}\sum_{j''=1}^{n}\eta^{\tau}_{j''}\sum_{k=1}^{\ell-1}(r^t_{i_{k+1}j_k}-r^t_{i_kj_k})+\gamma^{-1}\sum_{\tau=t+1}^{T}\beta^{-(\tau-t)}\sum_{i''=1}^{m}\xi^{\tau}_{i''}\sum_{k=1}^{\ell-1}(r^t_{i_{k+1}j_{k+1}}-r^t_{i_{k+1}j_k})\\
=&\gamma^{-1}K\sum_{k=1}^{\ell-1}(r^t_{i_{k+1}j_k}-r^t_{i_kj_k})+\gamma^{-1}L\sum_{k=1}^{\ell-1}(r^t_{i_{k+1}j_{k+1}}-r^t_{i_{k+1}j_k}), \qquad \text{(C.31)}
\end{aligned}$$

where we have denoted $K := \sum_{\tau=t+1}^{T}\alpha^{-(\tau-t)}\sum_{j''=1}^{n}\eta^{\tau}_{j''} \leq \alpha^{-1}\varepsilon_1$ and $L := \sum_{\tau=t+1}^{T}\beta^{-(\tau-t)}\sum_{i''=1}^{m}\xi^{\tau}_{i''} \leq \beta^{-1}\varepsilon_2$, for ease of notation.

If $K \leq L$, we have

$$\begin{aligned}
&V_{t+1}(\mathbf{x}-\epsilon_1\mathbf{e}^m_i+\epsilon_1\mathbf{e}^m_{i'},\mathbf{y}-\epsilon_2\mathbf{e}^n_j+\epsilon_2\mathbf{e}^n_{j'})-V_{t+1}(\mathbf{x},\mathbf{y})\\
\geq&\gamma^{-1}K[\sum_{k=1}^{\ell-1}(r^t_{i_{k+1}j_k}-r^t_{i_kj_k})+\sum_{k=1}^{\ell-1}(r^t_{i_{k+1}j_{k+1}}-r^t_{i_{k+1}j_k})]+\gamma^{-1}(L-K)\sum_{k=1}^{\ell-1}(r^t_{i_{k+1}j_{k+1}}-r^t_{i_{k+1}j_k})\\
=&\gamma^{-1}K(r^t_{i_\ell j_\ell}-r^t_{i_1j_1})+\gamma^{-1}(L-K)\sum_{k=1}^{\ell-1}(r^t_{i_{k+1}j_{k+1}}-r^t_{i_{k+1}j_k})\\
\geq&\gamma^{-1}K(r^t_{i_\ell j_\ell}-r^t_{i_1j_1})+\gamma^{-1}(L-K)\sum_{k=1}^{\ell-1}(r^t_{i_{k+1}j_{k+1}}-r^t_{i_kj_k})\\
=&\gamma^{-1}K(r^t_{i_\ell j_\ell}-r^t_{i_1j_1})+\gamma^{-1}(L-K)(r^t_{i_\ell j_\ell}-r^t_{i_1j_1})\\
=&\gamma^{-1}L(r^t_{i_\ell j_\ell}-r^t_{i_1j_1})=\gamma^{-1}L(r^t_{i'j'}-r^t_{ij})\geq(\gamma\beta)^{-1}(r^t_{i'j'}-r^t_{ij})\varepsilon_2, \qquad \text{(C.32)}
\end{aligned}$$

where the second inequality holds because $r^t_{i_{k+1}j_k} \leq r^t_{i_kj_k}$ (note that (i_k, j_k) weakly precedes (i_{k+1}, j_k)).

If $K > L$, we have

$$\begin{aligned}
&V_{t+1}(\mathbf{x}-\epsilon_1\mathbf{e}^m_i+\epsilon_1\mathbf{e}^m_{i'},\mathbf{y}-\epsilon_2\mathbf{e}^n_j+\epsilon_2\mathbf{e}^n_{j'})-V_{t+1}(\mathbf{x},\mathbf{y})\\
\geq&\gamma^{-1}(K-L)\sum_{k=1}^{\ell-1}(r^t_{i_{k+1}j_k}-r^t_{i_kj_k})+\gamma^{-1}L[\sum_{k=1}^{\ell-1}(r^t_{i_{k+1}j_{k+1}}-r^t_{i_{k+1}j_k})+\sum_{k=1}^{\ell-1}(r^t_{i_{k+1}j_k}-r^t_{i_kj_k})]\\
=&\gamma^{-1}(K-L)\sum_{k=1}^{\ell-1}(r^t_{i_{k+1}j_k}-r^t_{i_kj_k})+\gamma^{-1}L(r^t_{i_\ell j_\ell}-r^t_{i_1j_1})
\end{aligned}$$

$$
\begin{aligned}
&\geq \gamma^{-1}(K-L)\sum_{k=1}^{\ell-1}(r^t_{i_{k+1}j_{k+1}} - r^t_{i_k j_k}) + \gamma^{-1}L(r^t_{i_\ell j_\ell} - r^t_{i_1 j_1}) \\
&= \gamma^{-1}(K-L)(r^t_{i_\ell j_\ell} - r^t_{i_1 j_1}) + \gamma^{-1}L(r^t_{i_\ell j_\ell} - r^t_{i_1 j_1}) \\
&= \gamma^{-1}K(r^t_{i_\ell j_\ell} - r^t_{i_1 j_1}) \\
&\geq (\alpha\gamma)^{-1}(r^t_{i'j'} - r^t_{ij})\epsilon_1.
\end{aligned} \tag{C.33}
$$

Combining the two possibilities, we have $V_{t+1}(\mathbf{x} - \epsilon_1 \mathbf{e}^m_i + \epsilon_1 \mathbf{e}^m_{i'}, \mathbf{y} - \epsilon_2 \mathbf{e}^n_j + \epsilon_2 \mathbf{e}^n_{j'}) - V_{t+1}(\mathbf{x},\mathbf{y}) \geq \max\{(\alpha\gamma)^{-1}\epsilon_1, (\beta\gamma)^{-1}\epsilon_2\}(r^t_{i'j'} - r^t_{ij})$. □

Lemma C.16. *Suppose that (i,j) weakly precedes a non-neighboring pair (i',j'). For two feasible matching decisions $\mathbf{Q}$ and $\mathbf{Q} + \epsilon\mathbf{e}^{m\times n}_{ij} - \epsilon\mathbf{e}^{m\times n}_{i'j'}$ for the state $(\mathbf{x},\mathbf{y})$ in period t, the latter decision leads to a weakly higher expected total discounted reward from period t to period T, i.e., $H_t(\mathbf{Q} + \epsilon\mathbf{e}^{m\times n}_{ij} - \epsilon\mathbf{e}^{m\times n}_{i'j'}, \mathbf{x},\mathbf{y}) \geq H_t(\mathbf{Q},\mathbf{x},\mathbf{y})$.*

Proof of Lemma C.16. Let $\mathbf{u} := \mathbf{x} - \mathbf{1}\mathbf{Q}^\intercal$ and $\mathbf{v} := \mathbf{y} - \mathbf{1}\mathbf{Q}$ be the post-matching levels for using the matching decision $\mathbf{Q}$ in period t under the state $\mathbf{x},\mathbf{y}$). It is easy to see that the post-matching levels by using $\mathbf{Q} + \epsilon\mathbf{e}^{m\times n}_{ij} - \epsilon\mathbf{e}^{m\times n}_{i'j'}$ are given by $\mathbf{u} - \mathbf{e}_i + \mathbf{e}_{i'}$ and $\mathbf{v} - \mathbf{e}_j + \mathbf{e}_{j'}$. We have

$$
\begin{aligned}
&H_t(\mathbf{Q} + \epsilon\mathbf{e}^{m\times n}_{ij} - \epsilon\mathbf{e}^{m\times n}_{i'j'}, \mathbf{x},\mathbf{y}) \\
=&\mathbf{R}^t \circ (\mathbf{Q} + \epsilon\mathbf{e}^{m\times n}_{ij} - \epsilon\mathbf{e}^{m\times n}_{i'j'}) + \gamma E V_{t+1}(\alpha\mathbf{u} - \alpha\epsilon\mathbf{e}_i + \alpha\epsilon\mathbf{e}_{i'} + \mathbf{D}^{t+1}, \beta\mathbf{v} - \beta\epsilon\mathbf{e}_j + \beta\epsilon\mathbf{e}_{j'} + \mathbf{S}^{t+1}) \\
=&(r^t_{ij} - r^t_{i'j'})\epsilon + \mathbf{R}^t \circ \mathbf{Q} + \gamma E V_{t+1}(\alpha\mathbf{u} - \alpha\epsilon\mathbf{e}_i + \alpha\epsilon\mathbf{e}_{i'} + \mathbf{D}^{t+1}, \beta\mathbf{v} - \beta\epsilon\mathbf{e}_j + \beta\epsilon\mathbf{e}_{j'} + \mathbf{S}^{t+1}) \\
\geq&(r^t_{ij} - r^t_{i'j'})\epsilon + \mathbf{R}^t \circ \mathbf{Q} + \gamma E V_{t+1}(\alpha\mathbf{u} + \mathbf{D}^{t+1}, \beta\mathbf{v} + \mathbf{S}^{t+1}) + (r^t_{i'j'} - r^t_{ij})\epsilon \\
=&H_t(\mathbf{Q},\mathbf{x},\mathbf{y}),
\end{aligned}
$$

where the inequality follows from Lemma C.15, with $\epsilon_1 = \alpha\epsilon$ and $\epsilon_2 = \beta\epsilon$. □

We now present the following proposition that adds to the structural property in Theorem 1.

Proposition C.1. *There exists an optimal policy $\pi^* = \{\mathbf{Q}^{t*}\}_{t=1,\ldots,T}$ such that it satisfies the property in Theorem 1, and in addition, for any pair (i,j) weakly preceding a non-neighboring pair (i',j'), in each period t either the matching quantity $q^{t*}_{i'j'} = 0$, the post-matching level $u^{t*}_i = 0$, or $v^{t*}_j = 0$. Further, if Assumption 1 is satisfied, the optimal policy also satisfies the property in Theorem 2.*

Proof of Proposition C.1. The proof is similar to that of Theorems 1 and 2. For any matching policy that does not satisfy the properties in the proposition, we can construct a weakly better policy by successively transferring quantity from a weakly preceded pair to the corresponding preceding pair. In addition to the transferring matching quantities to neighboring pairs, we will also keep transferring quantity from a pair (i',j') to a non-neighboring pair (i,j), if (i,j) weakly precedes (i',j') according to Definition C.1. Following the same analysis as in the proof of Theorem 1, we can eventually obtain a feasible policy that satisfies the stated properties. □

As Theorem 1, Proposition C.1 implies that the optimal policy prioritizes a pair over those pairs it weakly precedes.

C.2. The partial order defined by the modified Monge condition

A partial order $\preceq$ defined over the set of demand-supply pairs $\{\rho=(i,j)\mid i=1,\dots,m, j=1,\dots,n\}$ is a binary relation for comparing (some of) the pairs, such that it satisfies: (i). Reflexivity: $\rho\preceq\rho$ for any pair ρ; (ii). Antisymmetry: If $\rho_1\preceq\rho_2$ and $\rho_2\preceq\rho_1$, then $\rho_1=\rho_2$; (iii). Transitivity: If $\rho_1\preceq\rho_2$ and $\rho_2\preceq\rho_3$, then $\rho_1\preceq\rho_3$.

It is clear from Definition 2 that any pair weakly precedes itself. Therefore, it satisfies Reflexivity. In the following, we will show that the modified Monge condition also satisfies transitivity.

Lemma C.17 below shows that the modified Monge condition satisfies transitivity among pairs that share a common demand/supply type.

Lemma C.17. *If (i,j) weakly precedes (i,j'), and (i,j') weakly precedes (i,j''), then (i,j) weakly precedes (i,j''). Likewise, if (i,j) weakly precedes (i',j) and (i',j) weakly precedes (i'',j), then (i,j) weakly precedes (i'',j).*

Proof of Lemma C.17 We will prove the first claim in the lemma, and the second claim holds by symmetry.

Suppose that (i,j) weakly precedes (i,j') and (i,j') weakly precedes (i,j''). By Definition 2, we have $r^t_{ij}-r^t_{ij'}\geq\beta\gamma(r^{t+1}_{i'''j}-r^{t+1}_{i'''j'})$ and $r^t_{ij'}-r^t_{ij''}\geq\beta\gamma(r^{t+1}_{i'''j'}-r^{t+1}_{i'''j''})$ for any demand type i'''. By adding up those two inequalities, we obtain $r^t_{ij}-r^t_{ij''}\geq\beta\gamma(r^{t+1}_{i'''j}-r^{t+1}_{i'''j''})$ for any demand type i'''.

On the other hand, it is also clear from Definition 2 that $r^t_{ij}\geq r^t_{ij'}\geq r^t_{ij''}$. Thus, $r^t_{ij}-r^t_{ij''}\geq\beta\gamma(r^{t+1}_{i'''j}-r^{t+1}_{i'''j''})^+$ for any demand type i''', which implies that (i,j) weakly precedes (i,j''). □

The following lemma shows that the modified Monge condition satisfies transitivity for any three pairs such that the first and last pairs do not share any demand/supply types.

Lemma C.18. *Suppose that (i,j) weakly precedes (i',j'), and (i',j') weakly (i'',j''). Then, (i,j) weakly precedes (i'',j'') if $i\neq i''$ and $j\neq j''$.*

Proof of Lemma C.18 Since (i,j) weakly precedes (i',j') and (i',j') weakly (i'',j''), there exists a "zig-zag" path of demand-supply pairs connecting (i,j) and (i',j') (if (i,j) and (i',j') are neighboring pairs, the two of them form the path), and another one connecting (i,j) and (i',j'), such that both path have one of the forms given as in Definition C.1 and along both paths each pair weakly precedes the next pair. Combining those two paths, we obtain a path connecting (i,j) and (i'',j''), such that it has one of the forms given as in Definition C.1 and along the path each pair weakly precedes the next pair.

In particular, if the second-to-last pair on the first path (i.e., the one connecting (i,j) and (i',j')) and the second pair on the second path (i.e., the one connecting (i',j') and (i'',j'')) do not share any common demand/supply type, the two paths directly form the combined path (i.e., we travel all the pairs on the two paths to travel from (i,j) and (i'',j'')). If the aforementioned two pairs share a common demand/supply type, say demand type i', then we can combine the two paths but skip the pair (i',j') to form the combined path (note that according to Lemma C.17, the second-to-last pair on the first path weakly precedes the second pair on the second path).

By the modified Monge condition for two non-neighboring pairs (Definition C.1), we can conclude that (i,j) weakly precedes (i'',j''). □

In the next lemma, we show that the modified Monge condition satisfies transitivity for three pairs such that the first and last pair share a common demand/supply type.

LEMMA C.19. *Suppose that (i,j) weakly precedes (i',j'), and (i',j') weakly precedes (i,j''). Then, (i,j) weakly precedes (i,j''). Likewise, if (i,j) weakly precedes (i',j'), and (i',j') weakly precedes (i'',j), then (i,j) weakly precedes (i'',j).*

Proof of Lemma C.19 We prove the first claim of the lemma and the second claim holds by symmetry.

Suppose that (i,j) weakly precedes (i',j'), and (i',j') weakly precedes (i,j''). We can find a "zig-zag" path connecting (i,j) and (i,j''), such that each pair weakly precedes the next pair along the path (analogous to the proof of Lemma C.18, we can combine the path connecting (i,j) and (i',j'), and the path connecting (i',j') and (i,j'') to form the path). Without loss of generality, we assume that the path has the form $(i,j)=(i_1,j_1)\to(i_2,j_1)\to(i_2,j_2)\to\cdots\to(i_\ell,j_\ell)\to(i_1,j_\ell)=(i,j'')$. Since for $k=1,\ldots,l-1$, (i_k,j_k) weakly precedes (i_{k+1},j_k), we have

$$r^t_{i_kj_k}-r^t_{i_{k+1}j_k}\geq\gamma\alpha(r^{t+1}_{i_kj'''}-r^{t+1}_{i_{k+1}j'''}), \tag{C.34}$$

for any supply type j'''. Likewise, since for $k = 1, \ldots, \ell - 1$, (i_{k+1}, j_k) weakly precedes (i_{k+1}, j_{k+1}), we have

$$r^t_{i_{k+1}j_k} - r^t_{i_{k+1}j_{k+1}} \geq \gamma\beta(r^{t+1}_{i'''j_k} - r^{t+1}_{i'''j_{k+1}}), \tag{C.35}$$

for all demand type i'''. Moreover, (i_ℓ, j_ℓ) weakly preceding (i_1, j_ℓ) implies that

$$r^t_{i_\ell j_\ell} - r^t_{i_1 j_\ell} \geq \alpha\gamma(r^{t+1}_{i_\ell j'''} - r^{t+1}_{i_1 j'''}), \tag{C.36}$$

for any supply type j'''.

By summing up (C.34) and (C.35) for $k = 1, \ldots, \ell - 1$ and (C.36), we obtain $r^t_{i_1 j_1} - r^t_{i_1 j_\ell} \geq \gamma\beta(r^{t+1}_{i'''j_1} - r^{t+1}_{i'''j_\ell})$ for any demand type i'''. On the other hand, it is evident that $r^t_{i_1 j_1} \geq r^t_{i_1 j_\ell}$. Since $(i_1, j_1) = (i, j)$ and $(i_1, j_\ell) = (i, j'')$, we have $r^t_{ij} - r^t_{ij''} \geq \gamma\beta(r^{t+1}_{i'''j} - r^{t+1}_{i'''j''})^+$. By Definition 2, this shows that (i, j) weakly precedes (i, j''). $\square$

By Lemmas C.17–C.19, we can conclude that the modified Monge condition satisfies transitivity.

Antisymmetry, however, is not necessarily satisfied by the modified Monge condition. It is possible for two pairs to weakly precede each other. We say that two pairs that precede each other are equivalent. It readily follows from Lemmas C.17–C.19 that this equivalence relation also satisfies transitivity. Clearly, it also satisfies the symmetric property (if ρ_1 is equivalent to ρ_2, then ρ_2 is equivalent to ρ_1) and the reflexive property (any pair ρ is equivalent to itself). It is well known that for an equivalence relation with those properties, the set of all demand-supply pairs can be divided into a number of equivalence classes, which we denote by $\mathcal{A}_1, \ldots, \mathcal{A}_N$. Within each equivalence class, there is at least one pair (since any pair is equivalent to itself), and any two pairs are equivalent.

It follows from Lemma C.16 and Lemma E.23 in Online Supplement E that transferring matching quantity from a pair to an equivalent pair would not change the expected discounted reward. This implies that we can arbitrarily prioritize two equivalent pairs without affecting the expected discounted reward. Formally, let us assign a *unique* integer number, referred to as priority number and denoted by $P_{(i,j)}$, to each demand-supply pair (i, j). The assignment of priority numbers can be arbitrary, but it cannot be altered. If two pairs (i, j) and (i', j') are equivalent, we break the tie by comparing their priority numbers.

Definition C.2. We define the binary relation $\preceq$ over the set of demand-supply pairs as follows. For any two pairs (i, j) and (i', j'), $(i, j) \preceq (i', j')$ if:

(i). (i, j) weakly precedes, but is not equivalent to (i', j'); or (ii). (i,j) is equivalent to (i',j') and $P_{(i,j)} \leq P_{(i',j')}$.

Proposition C.2. *The binary relation $\preceq$ defined in Definition C.2 is a partial order.*

Proof of Proposition C.2 *Reflexivity.* Any pair (i,j) is equivalent to itself. Since $P_{(i,j)} = P_{(i,j)}$, we have $(i,j) \preceq (i,j)$ by condition (ii) of Definition C.2.

Antisymmetry. Suppose that $(i,j) \preceq (i',j')$ and $(i',j') \preceq (i,j)$. We need to show that $(i,j) = (i',j')$. If (i,j) weakly precedes but is not equivalent to (i',j'). Then, according to Definition C.2, it is impossible to have $(i',j') \preceq (i,j)$. Thus, it must be the case that (i,j) and (i',j') are equivalent. By condition (ii) of Definition C.2, $P_{(i,j)} = P_{(i',j')}$. However, since each pair has a unique priority number, $P_{(i,j)} = P_{(i',j')}$ implies that $(i,j) = (i',j')$.

Transitivity. Suppose that $(i,j) \preceq (i',j')$ and $(i',j') \preceq (i'',j'')$. This implies that (i,j) weakly precedes (i',j'), which further weakly precedes (i'',j''). By transitivity of the modified Monge condition, we know that (i,j) weakly precedes (i'',j''). If (i,j) is not equivalent to (i'',j''), we can conclude that $(i,j) \preceq (i'',j'')$ by condition (ii) of Definition C.2. Otherwise, we have (i,j) and (i'',j'') weakly precede each other. Then, (i',j') weakly precedes (i,j), because it weakly precedes (i'',j''), but (i'',j'') weakly precedes (i,j). Thus, (i',j') is equivalent to (i,j). Analogously, (i',j') is also equivalent to (i'',j''). Therefore, the three pairs (i,j), (i',j') and (i'',j'') belong to the same equivalence class. Since $(i,j) \preceq (i',j')$ and $(i',j') \preceq (i'',j'')$, we have $P_{(i,j)} \leq P_{(i',j')} \leq P_{(i'',j'')}$. This further shows that $(i,j) \preceq (i'',j'')$. □

We can readily verify that all our results in the paper remain true if we replace the "weakly preceding" relation with the partial order $\preceq$. In fact, the partial order $\preceq$ is essentially the same as the modified Monge condition. The only difference is that, when two pairs are equivalent according to the modified Monge condition, the partial order $\preceq$ breaks the tie based on the arbitrarily assigned priority numbers. In light of the antisymmetry property of $\preceq$, we can assume without loss of generality that there do not exist two demand-supply pairs that weakly precede each other.

C.3. Endogenous correlation between matching quantities and future supply arrival

We consider an extension that allows the arrival of future supply to depend on the matching decision in the current period. In this extension, we assume that any supply matched in a period t may rejoin the system due to customer cancellation/return.

For simplicity, suppose that demand and supply both arrive in discrete (i.e., integer) quantities, and the matching decision also takes only integer values. Let the supply carry-over rate $\beta = 1$ and the demand carry-over rate α can be either 0 or 1.

For the matching decision $\mathbf{Q} = (q_{ij})_{i=1,\dots,m,j=1,\dots,n}$ used in period t, let $q^t_{\cdot j} := \sum_{i=1}^m q^t_{ij}$ is the total quantity of type j supply matched in that period. We suppose that the arrival of type j supply in the next period $t+1$ is $S_j^{t+1} := S_j^{t+1,0} + \sum_{k=1}^{q^t_{\cdot j}} X_k^{j,t+1}$, where $S_j^{t+1,0}$ is the new supply to join in period $t+1$, and $X_k^{j,t+1}$ ($k = 1,2,\dots,$ $j = 1,\dots,m$, $t = 1,\dots,T$) are i.i.d. binary random variables representing the possible cancellation/return of each unit of type j supply matched in period t. We also assume that $\left\{X_k^{j,t+1}\right\}_{\forall k,i,t}$, $\left\{D_i^{t+1}\right\}_{\forall i,t}$ and $\left\{S_j^{t+1}\right\}_{\forall j,t}$ are independent. Let X be a generic binary random variable that has the same distribution as $X_k^{j,t+1}$ (for all k, j and t), and we define $p_0 := \Pr(X=0)$ and $p_1 := \Pr(X=1)$.

We note that the above settings allow cancellations/returns to happen only in the beginning of the next period after the matching is made.

In the following lemma, we show that transferring matching quantity to a pair (i,j) from a neighboring pair it weakly precedes improves the expected total matching reward.

Lemma C.20. *Suppose that (i,j) weakly precedes (i,j'). For two matching decisions $\mathbf{Q}$ and $\mathbf{Q} + \mathbf{e}_{ij}^{m\times n} - \mathbf{e}_{ij'}^{m\times n}$ that are both feasible for the state $(\mathbf{x},\mathbf{y})$ in period t, we have $H_t(\mathbf{Q} + \mathbf{e}_{ij}^{m\times n} - \mathbf{e}_{ij'}^{m\times n}, \mathbf{x}, \mathbf{y}) \geq H_t(\mathbf{Q},\mathbf{x},\mathbf{y})$.*

To prove Lemma C.20, we require another lemma.

Lemma C.21. *For $t = 1,\dots,T$, $j,j' = 1,\dots,n$, and any $\mathbf{x}$ and $\mathbf{y}$, the inequality $V_t(\mathbf{x},\mathbf{y}+\mathbf{e}_{j'}) - V_t(\mathbf{x},\mathbf{y}+\mathbf{e}_j) \geq -\max_{\tau=t,t+1,\dots,T,i'=1,\dots,m}(r^\tau_{i'j} - r^\tau_{i'j'})^+ \sum_{\tau=t}^T (\gamma p_1)^{\tau-t}$ holds.*

Proof of Lemma C.21. We prove the lemma by induction. Clearly, the inequality holds for $t = T+1$, since $V_{T+1} \equiv 0$. Suppose that it also holds for $t+1$.

For ease of notation, let us define $\Delta^t_{j,j'} = \max_{\tau=t,t+1,\dots,T,i'=1,\dots,m}(r^\tau_{i'j} - r^\tau_{i'j'})^+$.

Let $\hat{\mathbf{Q}}$ be the optimal decision in period t for the state $(\mathbf{x},\mathbf{y}+\mathbf{e}_j)$. With the decision $\hat{\mathbf{Q}}$ in period t, the arrival of type ℓ supply in period $t+1$ is $S_\ell^{t+1,0} + \sum_{k=1}^{\hat{q}_{\cdot\ell}} X_k^\ell$, for any $\ell = 1,\dots,n$. We let $(\mathbf{u},\mathbf{v})$ denote the post-matching levels for using the decision $\hat{\mathbf{Q}}$ in period t.

If there exists some i' such that $\hat{q}_{i'j} > 0$, the decision $\hat{\mathbf{Q}} - \mathbf{e}_{i'j}^{m\times n} + \mathbf{e}_{i'j'}^{m\times n}$ is feasible for the state $(\mathbf{x},\mathbf{y}+\mathbf{e}_{j'})$. With the decision $\hat{\mathbf{Q}} - \mathbf{e}_{i'j}^{m\times n} + \mathbf{e}_{i'j'}^{m\times n}$, the new arrival of type j supply in period $t+1$ is $S_j^{t+1,0} + \sum_{k=1}^{\hat{q}_{\cdot j}-1} X_k^j$, the new arrival of type j' supply in period $t+1$ is $S_{j'}^{t+1,0} + \sum_{k=1}^{\hat{q}_{\cdot j'}+1} X_k^{j'}$, and the arrival of type ℓ supply in period $t+1$ is $S_\ell^{t+1,0} + \sum_{k=1}^{\hat{q}_{\cdot\ell}} X_k^\ell$ for $\ell \neq j,j'$. It is easy to see that the post-matching levels remain $(\mathbf{u},\mathbf{v})$.

Let us denote $\tilde{S}_\ell^{t+1} := S_\ell^{t+1,0} + \sum_{k=1}^{\hat{q}_{\cdot\ell}} X_k^\ell$ for $\ell \neq j$, and $\tilde{S}_j^{t+1} := S_j^{t+1,0} + \sum_{k=1}^{\hat{q}_{\cdot j}-1} X_k^j$. We have

$$V_t(\mathbf{x},\mathbf{y}+\mathbf{e}_{j'}) \geq H_t(\hat{\mathbf{Q}} - \mathbf{e}_{i'j}^{m\times n} + \mathbf{e}_{i'j'}^{m\times n}, \mathbf{x}, \mathbf{y}+\mathbf{e}_{j'})$$

$$= -r^t_{i'j} + r^t_{i'j'} + \mathbf{R}^t \circ \hat{\mathbf{Q}} + \gamma EV_{t+1}(\alpha\mathbf{u} + \mathbf{D}^{t+1}, \mathbf{v} + \tilde{\mathbf{S}}^{t+1} + X^{j'}_{\hat{q}_{\cdot j'}+1}\mathbf{e}_{j'}),$$

and

$$V_t(\mathbf{x}, \mathbf{y} + \mathbf{e}_j) = H_t(\hat{\mathbf{Q}}, \mathbf{x}, \mathbf{y} + \mathbf{e}_j) = \mathbf{R}^t \circ \hat{\mathbf{Q}} + \gamma EV_{t+1}(\alpha\mathbf{u} + \mathbf{D}^{t+1}, \mathbf{v} + \tilde{\mathbf{S}}^{t+1} + X^j_{\hat{q}_{\cdot j}}\mathbf{e}_j),$$

Thus,

$$\begin{aligned}
&V_t(\mathbf{x}, \mathbf{y} + \mathbf{e}_{j'}) - V_t(\mathbf{x}, \mathbf{y} + \mathbf{e}_j) \\
\geq &H_t(\hat{\mathbf{Q}} - \mathbf{e}^{m\times n}_{i'j} + \mathbf{e}^{m\times n}_{i'j'}, \mathbf{x}, \mathbf{y} + \mathbf{e}_{j'}) - H_t(\hat{\mathbf{Q}}, \mathbf{x}, \mathbf{y} + \mathbf{e}_j) \\
= &-r^t_{i'j} + r^t_{i'j'} + \gamma EV_{t+1}(\alpha\mathbf{u} + \mathbf{D}^{t+1}, \mathbf{v} + \tilde{\mathbf{S}}^{t+1} + X^{j'}_{\hat{q}_{\cdot j'}+1}\mathbf{e}_{j'}) - \gamma EV_{t+1}(\alpha\mathbf{u} + \mathbf{D}^{t+1}, \mathbf{v} + \tilde{\mathbf{S}}^{t+1} + X^j_{\hat{q}_{\cdot j}}\mathbf{e}_j) \\
= &-r^t_{i'j} + r^t_{i'j'} + \gamma \left[EV_{t+1}(\alpha\mathbf{u} + \mathbf{D}^{t+1}, \mathbf{v} + \tilde{\mathbf{S}}^{t+1} + \mathbf{e}_{j'}) - EV_{t+1}(\alpha\mathbf{u} + \mathbf{D}^{t+1}, \mathbf{v} + \tilde{\mathbf{S}}^{t+1} + \mathbf{e}_j)\right] p_1 \\
\geq &-\Delta^t_{j,j'} - \gamma p_1 \Delta^{t+1}_{j,j'} \sum_{\tau=t+1}^{T} (\gamma p_1)^{\tau-t-1} \\
= &-\Delta^t_{j,j'} \sum_{\tau=t}^{T} (\gamma p_1)^{\tau-t},
\end{aligned}$$

where the second inequality holds because $-r^t_{i'j} + r^t_{i'j'} \geq -\max_{\tau=1,\ldots,T, i'=1,\ldots,m}(r^\tau_{i'j} - r^\tau_{i'j'})^+ = -\Delta^t_{j,j'}$, and also because of the induction hypothesis.

If $\hat{q}_{i'j} = 0$ for all $i' = 1, \ldots, m$, the decision $\hat{\mathbf{Q}}$ is feasible for the state $(\mathbf{x}, \mathbf{y} + \mathbf{e}_{j'})$. The post matching levels for using $\hat{\mathbf{Q}}$ in period t under the state $(\mathbf{x}, \mathbf{y} + \mathbf{e}_{j'})$ are $(\mathbf{u}, \mathbf{v} - \mathbf{e}_j + \mathbf{e}_{j'})$. The arrivals of supply in period $t+1$ is $\tilde{\mathbf{S}}^{t+1} + X^j_{\hat{q}_{\cdot j}}\mathbf{e}_j$. We have,

$$V_t(\mathbf{x}, \mathbf{y} + \mathbf{e}_{j'}) \geq H_t(\hat{\mathbf{Q}}, \mathbf{x}, \mathbf{y} + \mathbf{e}_{j'}) = \mathbf{R}^t \circ \hat{\mathbf{Q}} + \gamma EV_{t+1}(\alpha\mathbf{u} + \mathbf{D}^{t+1}, \mathbf{v} - \mathbf{e}_j + \mathbf{e}_{j'} + \tilde{\mathbf{S}}^{t+1} + X^j_{\hat{q}_{\cdot j}}\mathbf{e}_j).$$

As shown earlier, we have

$$V_t(\mathbf{x}, \mathbf{y} + \mathbf{e}_j) = H_t(\hat{\mathbf{Q}}, \mathbf{x}, \mathbf{y} + \mathbf{e}_j) = \mathbf{R}^t \circ \hat{\mathbf{Q}} + \gamma EV_{t+1}(\alpha\mathbf{u} + \mathbf{D}^{t+1}, \mathbf{v} + \tilde{\mathbf{S}}^{t+1} + X^j_{\hat{q}_{\cdot j}}\mathbf{e}_j).$$

It follows that

$$\begin{aligned}
&V_t(\mathbf{x}, \mathbf{y} + \mathbf{e}_{j'}) - V_t(\mathbf{x}, \mathbf{y} + \mathbf{e}_j) \\
\geq &\gamma EV_{t+1}(\alpha\mathbf{u} + \mathbf{D}^{t+1}, \mathbf{v} - \mathbf{e}_j + \mathbf{e}_{j'} + \tilde{\mathbf{S}}^{t+1} + X^j_{\hat{q}_{\cdot j}}\mathbf{e}_j) - \gamma EV_{t+1}(\alpha\mathbf{u} + \mathbf{D}^{t+1}, \mathbf{v} + \tilde{\mathbf{S}}^{t+1} + X^j_{\hat{q}_{\cdot j}}\mathbf{e}_j) \\
= &\gamma EV_{t+1}(\alpha\mathbf{u} + \mathbf{D}^{t+1}, \mathbf{v} - \mathbf{e}_j + \mathbf{e}_{j'} + \tilde{\mathbf{S}}^{t+1} + X^j_{\hat{q}_{\cdot j}}\mathbf{e}_j) - \gamma EV_{t+1}(\alpha\mathbf{u} + \mathbf{D}^{t+1}, \mathbf{v} - \mathbf{e}_j + \mathbf{e}_j + \tilde{\mathbf{S}}^{t+1} + X^j_{\hat{q}_{\cdot j}}\mathbf{e}_j)
\end{aligned}$$

$$\geq -\gamma\Delta_{j,j'}^{t+1}\sum_{\tau=t+1}^{T}(\gamma p_1)^{\tau-t-1} \geq -\Delta_{j,j'}^{t}\sum_{\tau=t+1}^{T}(\gamma p_1)^{\tau-t-1} = -\Delta_{j,j'}^{t}\sum_{\tau=t}^{T-1}(\gamma p_1)^{\tau-t} \geq -\Delta_{j,j'}^{t}\sum_{\tau=t}^{T}(\gamma p_1)^{\tau-t},$$

where the second inequality follows from the induction hypothesis. The induction is completed. □

We can now proceed to prove Lemma C.20.

Proof of Lemma C.20. We first consider the decision $\tilde{\mathbf{Q}} := \mathbf{Q} - \mathbf{e}_{ij'}^{m\times n}$. With this decision and the state $(\mathbf{x},\mathbf{y})$, the post-matching levels in period t are given by $\tilde{\mathbf{u}} = \mathbf{x} - \mathbf{1}\mathbf{Q}^{\intercal} + \mathbf{e}_i$ and $\tilde{\mathbf{v}} = \mathbf{y} - \mathbf{1}\mathbf{Q} + \mathbf{e}_{j'}$. The new state in period $t+1$ (after arrival of demand and supply) is $(\tilde{\mathbf{x}}^{t+1}, \tilde{\mathbf{y}}^{t+1}) := (\alpha\tilde{\mathbf{u}} + \mathbf{D}^{t+1}, \beta\tilde{\mathbf{v}} + \tilde{\mathbf{S}}^{t+1})$, where $\tilde{\mathbf{S}}^{t+1} := (\tilde{S}_1^{t+1},\ldots,\tilde{S}_n^{t+1})$, $\tilde{S}_\ell^{t+1} = S_\ell^{t+1,0} + \sum_{k=1}^{q_{\cdot\ell}} X_k^\ell$ for $\ell \neq j'$ and $\tilde{S}_{j'}^{t+1} = S_{j'}^{t+1,0} + \sum_{k=1}^{q_{\cdot j'}-1} X_k^{j'}$.

If we use the matching decision $\mathbf{Q}$ in period t for the state $(\mathbf{x},\mathbf{y})$, the arrival of type j' supply in period $t+1$ is $S_{j'}^{t+1,0} + \sum_{k=1}^{q_{\cdot j'}} X_k^{j'} = \tilde{S}_{j'}^{t+1} + X_{q_{\cdot j'}}^{j'}$, and the arrival of any other type ℓ is equal to $\tilde{S}_\ell^{t+1}$. The post-matching levels in period t are given by $\tilde{\mathbf{u}} - \mathbf{e}_i$ and $\tilde{\mathbf{v}} - \mathbf{e}_{j'}$.

Similarly, if we use the matching decision $\mathbf{Q} - \mathbf{e}_{ij'}^{m\times n} + \mathbf{e}_{ij}^{m\times n}$ in period t for the state $(\mathbf{x},\mathbf{y})$, the arrival of type j supply in period $t+1$ is $S_j^{t+1,0} + \sum_{k=1}^{q_{\cdot j}+1} X_k^j = \tilde{S}_j^{t+1} + X_{q_{\cdot j}+1}^j$, and the new arrival of any other type ℓ is equal to $\tilde{S}_\ell^{t+1}$. The post-matching levels in period t are given by $\tilde{\mathbf{u}} - \mathbf{e}_i$ and $\tilde{\mathbf{v}} - \mathbf{e}_j$.

The expected total discounted reward achievable by the matching decision $\mathbf{Q}$ in period t is:

$$\begin{aligned}
&H_t(\mathbf{Q},\mathbf{x},\mathbf{y})\\
=&\mathbf{R}^t \circ \mathbf{Q} + \gamma EV_{t+1}(\alpha\tilde{\mathbf{u}} - \alpha\mathbf{e}_i + \mathbf{D}^{t+1}, \tilde{\mathbf{v}} - \mathbf{e}_{j'} + \tilde{\mathbf{S}}^{t+1} + X_{q_{\cdot j'}}^{j'}\mathbf{e}_{j'})\\
=&\mathbf{R}^t \circ \mathbf{Q} + \gamma EV_{t+1}(\alpha\tilde{\mathbf{u}} - \alpha\mathbf{e}_i + \mathbf{D}^{t+1}, \tilde{\mathbf{v}} - \mathbf{e}_{j'} + \tilde{\mathbf{S}}^{t+1})p_0 + \gamma EV_{t+1}(\alpha\tilde{\mathbf{u}} - \alpha\mathbf{e}_i + \mathbf{D}^{t+1}, \tilde{\mathbf{v}} + \tilde{\mathbf{S}}^{t+1})p_1
\end{aligned}$$

and the expected total discounted reward achievable by $\mathbf{Q} + \mathbf{e}_{ij}^{m\times n} - \mathbf{e}_{ij'}^{m\times n}$ in period t is:

$$\begin{aligned}
&H_t(\mathbf{Q} + \mathbf{e}_{ij}^{m\times n} - \mathbf{e}_{ij'}^{m\times n},\mathbf{x},\mathbf{y})\\
=&r_{ij}^t - r_{ij'}^t + \mathbf{R}^t \circ \mathbf{Q} + \gamma EV_{t+1}(\alpha\tilde{\mathbf{u}} - \alpha\mathbf{e}_i + \mathbf{D}^{t+1}, \tilde{\mathbf{v}} + \tilde{\mathbf{S}}^{t+1} - \mathbf{e}_j + X_{q_{\cdot j}+1}^{j}\mathbf{e}_j)\\
=&r_{ij}^t - r_{ij'}^t + \mathbf{R}^t \circ \mathbf{Q} + \gamma EV_{t+1}(\alpha\tilde{\mathbf{u}} - \alpha\mathbf{e}_i + \mathbf{D}^{t+1}, \tilde{\mathbf{v}} + \tilde{\mathbf{S}}^{t+1} - \mathbf{e}_j)p_0 + \gamma EV_{t+1}(\alpha\tilde{\mathbf{u}} - \alpha\mathbf{e}_i + \mathbf{D}^{t+1}, \tilde{\mathbf{v}} + \tilde{\mathbf{S}}^{t+1})p_1.
\end{aligned}$$

Then,

$$\begin{aligned}
&H_t(\mathbf{Q} + \mathbf{e}_{ij}^{m\times n} - \mathbf{e}_{ij'}^{m\times n},\mathbf{x},\mathbf{y}) - H_t(\mathbf{Q},\mathbf{x},\mathbf{y})\\
=&r_{ij}^t - r_{ij'}^t
\end{aligned}$$

$$+\gamma\left[EV_{t+1}(\alpha\tilde{\mathbf{u}}-\alpha\mathbf{e}_i+\mathbf{D}^{t+1},\tilde{\mathbf{v}}+\tilde{\mathbf{S}}^{t+1}-\mathbf{e}_j)-EV_{t+1}(\alpha\tilde{\mathbf{u}}-\alpha\mathbf{e}_i+\mathbf{D}^{t+1},\tilde{\mathbf{v}}-\mathbf{e}_{j'}+\tilde{\mathbf{S}}^{t+1})\right]p_0$$
$$=r_{ij}^t-r_{ij'}^t$$
$$+\gamma p_0\left[EV_{t+1}(\alpha\tilde{\mathbf{u}}-\alpha\mathbf{e}_i+\mathbf{D}^{t+1},\tilde{\mathbf{v}}+\tilde{\mathbf{S}}^{t+1}-\mathbf{e}_{j'}-\mathbf{e}_j+\mathbf{e}_{j'})-EV_{t+1}(\alpha\tilde{\mathbf{u}}-\alpha\mathbf{e}_i+\mathbf{D}^{t+1},\tilde{\mathbf{v}}-\mathbf{e}_j-\mathbf{e}_{j'}+\tilde{\mathbf{S}}^{t+1}+\mathbf{e}_j)\right]$$
$$\geq r_{ij}^t-r_{ij'}^t-\max_{\tau=t+1,\ldots,T,i'=1,\ldots,m}(r_{i'j}^\tau-r_{i'j'}^\tau)^+\gamma p_0\sum_{\tau=t+1}^{T}(\gamma p_1)^{\tau-t-1}$$
$$\geq r_{ij}^t-r_{ij'}^t-\max_{\tau=t+1,\ldots,T,i'=1,\ldots,m}(r_{i'j}^\tau-r_{i'j'}^\tau)^+\frac{\gamma p_0}{1-\gamma p_1}$$
$$\geq r_{ij}^t-r_{ij'}^t-\max_{\tau=t+1,\ldots,T,i'=1,\ldots,m}(r_{i'j}^\tau-r_{i'j'}^\tau)^+,$$

where the first inequality follows from Lemma C.21, and the last inequality holds because $1-\gamma p_1\geq\gamma p_0$. By the definition of the modified Monge condition, we have $H_t(\mathbf{Q}+\mathbf{e}_{ij}^{m\times n}-\mathbf{e}_{ij'}^{m\times n},\mathbf{x},\mathbf{y})\geq H_t(\mathbf{Q},\mathbf{x},\mathbf{y})$. □

With Lemma C.20, many of our results in the paper remain true in the current model setting. Specifically, Theorems 1 and 2 remain true (the proofs are identical to those for the baseline model, except that we need to use Lemma C.20 in place of Lemma E.23). Theorem 3 also holds (the proof is similar to the original proof). Consequently, we still have the same structure for the optimal policies in the horizontal model (Proposition 3) and the vertical model (Proposition 4).

C.4. Non-additive reward structure for the vertical model

We consider an extension to the vertical model allow for non-additive reward structure for the vertical model. With the following assumption, all propositions in Section 5 remain true, and the 1-step-lookahead heuristic can be implemented similarly.

Assumption 5. *For $t=1,\ldots,T$,*

(i) The unit matching reward r_{ij}^t is decreasing in i and j;

(ii) For $i=1,\ldots,m-1$ and $j=1,\ldots,n-1$, $r_{ij}^t-r_{i+1,j}^t\geq\gamma\alpha\max_{j''=1,\ldots,m}(r_{ij''}^{t+1}-r_{i+1,j''}^{t+1})$ and $r_{ij}^t-r_{i,j+1}^t\geq\gamma\beta\max_{i''=1,\ldots,m}(r_{i''j}^{t+1}-r_{i'',j+1}^{t+1})$;

(iii) r_{ij}^t is supermodular with respect to i and j, i.e., $r_{ij}^t-r_{i,j+1}^t\geq r_{i+1,j}^t-r_{i+1,j+1}^t$ for $i=1,\ldots,m-1$ and $j=1,\ldots,n-1$.

Part (i) of Assumption 5 innocuously assumes that a demand/supply type with a smaller index has a higher quality level. (ii) requires that the difference in rewards between a high quality demand (supply) type and a low quality demand (supply) type is decreasing over time, regardless of the supply (demand) type they match with. (iii) requires the complementary effect between demand quality and supply quality.

We illustrate the assumption by considering the additive/multiplicative hybrid reward structure $r_{ij}^t = \eta(r_{id}^t + r_{is}^t) + r_{id}^t r_{js}^t$, where $\eta \geq 0$, r_{id}^t decreases in i, and r_{js}^t decreases in j. One can verify that Assumption 5 is satisfied if and only if:

$$\min_{1\leq i,i'\leq m} \frac{\eta + r_{id}^t}{\eta + r_{i'd}^t} \geq \frac{\gamma\beta(r_{js}^{t+1} - r_{j+1,s}^{t+1})}{r_{js}^t - r_{j+1,s}^t} \text{ and } \min_{1\leq j,j'\leq n} \frac{\eta + r_{js}^t}{\eta + r_{j's}^t} \geq \frac{\gamma\alpha(r_{id}^{t+1} - r_{i+1,d}^{t+1})}{r_{id}^t - r_{i+1,d}^t},$$

which holds if Assumption 3 (for the additive-reward model) is satisfied and η is sufficiently large (i.e., the additive component of the reward is sufficiently significant, compared with the multiplicative component).

D. Approximation of the expected total discounted reward under the greedy matching policy by Monte Carlo Simulation

For both the match-down-to heuristic for the horizontal model (see Section 4.2) and the 1-step-lookahead heuristic for the vertical model (see Section 5.2), we need to evaluate the expected discounted reward $EV_{t+1}^g(\alpha\mathbf{u} + \mathbf{D}^{t+1}, \beta\mathbf{v} + \mathbf{S}^{t+1})$ from period t onward by Monte Carlo simulation. In this online supplement, we briefly explain the implementation.

Let us draw N sample paths of demand realizations from period t to the end of the horizon, according to their probability distributions. Let us denote the kth sample path by $\omega_k = (\mathbf{d}_k^{t+1}, \mathbf{s}_k^{t+1}; \mathbf{d}_k^{t+2}, \mathbf{s}_k^{t+2}; \ldots; \mathbf{d}_k^T, \mathbf{s}_k^T)$, for $k = 1, \ldots, N$. For each sample path ω_k, we have the starting state $(\mathbf{x}_k^{t+1}, \mathbf{y}_k^{t+1}) := (\alpha\mathbf{u} + \mathbf{d}_k^{t+1}, \beta\mathbf{v} + \mathbf{s}_k^{t+1})$ in period $t+1$, and apply greedy matching along the sample path ω_k in all subsequent periods to obtain the total discounted matching reward, which we denote by $TR^{t+1}(\mathbf{x}_k, \mathbf{y}_k, \omega_k)$. Then, we approximate $EV_{t+1}^g(\alpha\mathbf{u} + \mathbf{D}^{t+1}, \beta\mathbf{v} + \mathbf{S}^{t+1})$ by $EV_{t+1}^g(\alpha\mathbf{u} + \mathbf{D}^{t+1}, \beta\mathbf{v} + \mathbf{S}^{t+1}) \approx \sum_{k=1}^N TR(\mathbf{x}_k, \mathbf{y}_k, \omega_k)/N$.

For the match-down-to heuristic for the horizontal model, to determine the protection level for matching the pair (i,j) we need to evaluate the expected values $EV_{t+1}^g(\alpha[(\breve{x}_i - \breve{y}_j)^+ + p]\mathbf{e}_i^m + \mathbf{D}^{t+1}, \mathbf{S}^{t+1})$ and $EV_{t+1}^g(\mathbf{D}^{t+1}, \beta[(\breve{x}_i - \breve{y}_j)^- + p]\mathbf{e}_j^n + \mathbf{S}^{t+1})$ in (2), for given $\breve{x}_i$ and $\breve{y}$ (i.e., the available demand and supply immediately before we match the pair (i,j)) and protection level p. We can apply Monte Carlo simulation with $(\mathbf{u}, \mathbf{v}) = ([(\breve{x}_i - \breve{y}_j)^+ + p]\mathbf{e}_i^m, \mathbf{0}^n)$ and $(\mathbf{u}, \mathbf{v}) = (\mathbf{0}^m, [(\breve{x}_i - \breve{y}_j)^- + p]\mathbf{e}_j^n)$ to evaluate the two expected values.

For the 1-step-lookahead heuristic for the vertical model, we need to evaluate the expectation $EV_{t+1}^g(\alpha(IB + p)\mathbf{e}_i^m + \mathbf{D}^{t+1}, \beta p\mathbf{e}_j^n + \mathbf{S}^{t+1})$ in (4) for the given total demand and supply imbalance IB and the protection level p. To that end, we can apply Monte Carlo simulation with $\mathbf{u} = (IB + p)\mathbf{e}_i^m$ and $\mathbf{v} = p\mathbf{e}_j^n$.

E. Proofs

Proof of Theorems 1 and 2. We prove Theorems 1 and 2 simultaneously. To that end, we first present and prove two lemmas (Lemmas E.22 and E.23 as follows).

LEMMA E.22. *The following statements hold for all periods.*

(i) For any $x_i > 0$ and any $\varepsilon \in [0, x_i]$, there exists $(\lambda_1^\tau, \ldots, \lambda_m^\tau) \geq \mathbf{0}$ for $\tau = t, \ldots, T$, such that $\sum_{\tau=t}^{T} \alpha^{-(\tau-t)} \sum_{j'=1}^{m} \lambda_{j'}^{\tau} \leq \varepsilon$ and $V_t(\mathbf{x} - \varepsilon \mathbf{e}_i^m + \varepsilon \mathbf{e}_{i'}^m, \mathbf{y}) - V_t(\mathbf{x}, \mathbf{y}) \geq -\sum_{\tau=t}^{T} \gamma^{\tau-t} \sum_{j'=1}^{m} \lambda_{j'}^{\tau}(r_{ij'}^{\tau} - r_{i'j'}^{\tau})$.

(ii) For any $y_j > 0$ and any $\varepsilon \in [0, y_j]$, there exists $(\xi_1^\tau, \ldots, \xi_n^\tau) \geq \mathbf{0}$ for $\tau = t, \ldots, T$, such that $\sum_{\tau=t}^{T} \beta^{-(\tau-t)} \sum_{i'=1}^{n} \xi_{i'}^{\tau} \leq \varepsilon$ and $V_t(\mathbf{x}, \mathbf{y} - \varepsilon \mathbf{e}_j^n + \varepsilon \mathbf{e}_{j'}^n) - V_t(\mathbf{x}, \mathbf{y}) \geq -\sum_{\tau=t}^{T} \gamma^{\tau-t} \sum_{i'=1}^{n} \xi_{i'}^{\tau}(r_{i'j}^{\tau} - r_{i'j'}^{\tau})$.

Proof of Lemma E.22. We only need to prove part (i), and the proof of part (ii) follows by symmetry. We prove part (i) by induction. The result holds for $t = T+1$. Because $V_{T+1}(\mathbf{x}, \mathbf{y}) \equiv 0$, we can simply set λ_j^T to zero. Suppose that it holds for period $t+1$.

Now consider period t. Let $\hat{\mathbf{Q}} \in \arg\max_{\mathbf{Q}} H_t(\mathbf{Q}, \mathbf{x}, \mathbf{y})$ be an optimal decision in period t under the state $(\mathbf{x}, \mathbf{y})$ in period t. We will construct a decision $\bar{\mathbf{Q}}$ that is feasible under the state $(\mathbf{x} - \varepsilon \mathbf{e}_i^m + \varepsilon \mathbf{e}_{i'}^m, \mathbf{y})$.

Under the latter state $(\mathbf{x} - \varepsilon \mathbf{e}_i^m + \varepsilon \mathbf{e}_{i'}^m, \mathbf{y})$, the capacity of i is reduced by ε compared with the original state $(\mathbf{x}, \mathbf{y})$. We need to adjust the matching decision $\hat{\mathbf{Q}}$ accordingly to make it feasible for that state. In particular, we reduce the matching quantity $\hat{q}_{ij}$ by μ_j for $j = 1, \ldots, n$, where the nonnegative numbers $\mu_1, \ldots, \mu_n$ are defined as follows.

$$\mu_j = \min\{\hat{q}_{ij}, (\varepsilon - \sum_{j'=1}^{j-1} \hat{q}_{ij'})^+\}, \text{ for } j = 1, \ldots, n.$$

If $\sum_{j'=1}^{k-1} \hat{q}_{ij'} < \varepsilon \leq \sum_{j'=1}^{k} \hat{q}_{ij'}$ for some $1 \leq k \leq n$, then one can verify that $\mu_j = \hat{q}_{ij}$ for $j = 1, \ldots, k-1$, $\mu_k = \varepsilon - \sum_{j'=1}^{k} \hat{q}_{ij'}$ and $\mu_j = 0$ for $j = k+1, \ldots, n$. In this case, $\sum_{j=1}^{n} \mu_j = \varepsilon$, and thus $\sum_{j=1}^{n}(\hat{q}_{ij} - \mu_j) = \sum_{j=1}^{n} \hat{q}_{ij} - \sum_{j=1}^{n} \mu_j = \sum_{j=1}^{n} \hat{q}_{ij} - \varepsilon \leq x_i - \varepsilon$.

If $\varepsilon > \sum_{j'=1}^{n} \hat{q}_{ij'}$, then $\mu_j = \hat{q}_{ij}$ for all $j = 1, \ldots, n$. Therefore, in this case we reduce all the matching quantities $\hat{q}_{ij}$, $j = 1, \ldots, n$ to zero. We then have $\sum_{j=1}^{n}(\hat{q}_{ij} - \mu_j) = 0 \leq x_i - \varepsilon$.

On the other hand, under the state $(\mathbf{x} - \varepsilon \mathbf{e}_i^m + \varepsilon \mathbf{e}_{i'}^m, \mathbf{y})$, the capacity of i' is increased by ε compared with the state $(\mathbf{x}, \mathbf{y})$. This allows us to increase the matching quantity $\hat{q}_{i'j}$ by μ_j for all $j = 1, \ldots, n$.

We define

$$\bar{\mathbf{Q}} = \hat{\mathbf{Q}} - \sum_{j=1}^{n} \mu_j \mathbf{e}_{ij}^{m\times n} + \sum_{j=1}^{n} \mu_j \mathbf{e}_{i'j}^{m\times n},$$

which is feasible for the state $(\mathbf{x} - \varepsilon\mathbf{e}_i^m + \varepsilon\mathbf{e}_{i'}^m, \mathbf{y})$. To see this, we have

$$\mathbf{1}^m\bar{\mathbf{Q}} = \mathbf{1}^m\hat{\mathbf{Q}} - \sum_{j=1}^{n} \mu_j \mathbf{1}^m \mathbf{e}_{ij}^{m\times n} + \sum_{j=1}^{n} \mu_j \mathbf{1}^m \mathbf{e}_{i'j}^{m\times n} = \mathbf{1}^m\hat{\mathbf{Q}} - \sum_{j=1}^{n} \mu_j \mathbf{e}_j^n + \sum_{j=1}^{n} \mu_j \mathbf{e}_j^n = \mathbf{1}^m\hat{\mathbf{Q}} \le \mathbf{y},$$

$$\text{and } \bar{\mathbf{Q}}(\mathbf{1}^n)^\intercal = \hat{\mathbf{Q}}(\mathbf{1}^n)^\intercal - \sum_{j=1}^{n} \mu_j \mathbf{e}_{ij}^{m\times n}(\mathbf{1}^n)^\intercal + \sum_{j=1}^{n} \mu_j \mathbf{e}_{i'j}^{m\times n}(\mathbf{1}^n)^\intercal = \hat{\mathbf{Q}}(\mathbf{1}^n)^\intercal - \sum_{j=1}^{n} \mu_j (\mathbf{e}_i^m)^\intercal + \sum_{j=1}^{n} \mu_j (\mathbf{e}_{i'}^m)^\intercal.$$

It follows that $\left(\bar{\mathbf{Q}}(\mathbf{1}^n)^\intercal\right)_i = \sum_{j=1}^n \hat{q}_{ij} - \sum_{j=1}^n \mu_j \le x_i - \varepsilon$, $\left(\bar{\mathbf{Q}}(\mathbf{1}^n)^\intercal\right)_{i'} = \sum_{j=1}^n \hat{q}_{i'j} + \sum_{j=1}^n \mu_j \le \sum_{j=1}^n \hat{q}_{i'j} + \varepsilon \le x_{i'} + \varepsilon$ and $\left(\bar{\mathbf{Q}}(\mathbf{1}^n)^\intercal\right)_{i''} = \sum_{j=1}^n \hat{q}_{i''j} \le x_{i''}$ for all $i'' \ne i, i'$. Thus, $\bar{\mathbf{Q}}(\mathbf{1}^n)^\intercal \le (\mathbf{x} - \varepsilon\mathbf{e}_i^m + \varepsilon\mathbf{e}_{i'}^m)^\intercal$.

Therefore, $\bar{\mathbf{Q}}$ is a feasible decision for the state $(\mathbf{x} - \varepsilon\mathbf{e}_i^m + \varepsilon\mathbf{e}_{i'}^m, \mathbf{y})$. Under the decision $\bar{\mathbf{x}}$, the total reward received in period t is

$$\mathbf{R}^t \circ \bar{\mathbf{Q}} = \mathbf{R}^t \circ (\hat{\mathbf{Q}} - \sum_{j=1}^{n} \mu_j \mathbf{e}_{ij}^{m\times n} + \sum_{j=1}^{n} \mu_j \mathbf{e}_{i'j}^{m\times n}) = \mathbf{R}^t \circ \hat{\mathbf{Q}} - \sum_{j=1}^{n} \mu_j r_{ij}^t + \sum_{j=1}^{n} \mu_j r_{i'j}^t.$$

The post-matching levels in period t are

$$\begin{aligned}
\bar{\mathbf{u}} =& \mathbf{x} - \varepsilon\mathbf{e}_i^m + \varepsilon\mathbf{e}_{i'}^m - \mathbf{1}^n\bar{\mathbf{Q}}^\intercal = \mathbf{x} - \varepsilon\mathbf{e}_i^m + \varepsilon\mathbf{e}_{i'}^m - \mathbf{1}^n\hat{\mathbf{Q}}^\intercal + \sum_{j=1}^{n} \mu_j \mathbf{e}_i^m - \sum_{j=1}^{n} \mu_j \mathbf{e}_{i'}^m \\
&= \hat{\mathbf{u}} - (\varepsilon - \sum_{j=1}^{n} \mu_j)\mathbf{e}_i^m + (\varepsilon - \sum_{j=1}^{n} \mu_j)\mathbf{e}_{i'}^m, \\
\bar{\mathbf{v}} =& \mathbf{y} - \mathbf{1}^m\bar{\mathbf{Q}} = \mathbf{y} - \mathbf{1}^m\hat{\mathbf{Q}} = \hat{\mathbf{v}}.
\end{aligned}$$

Consequently,

$$\begin{aligned}
& V_t(\mathbf{x} - \varepsilon\mathbf{e}_i^m + \varepsilon\mathbf{e}_{i'}^m, \mathbf{y}) - V_t(\mathbf{x}, \mathbf{y}) \\
\ge & H_t(\bar{\mathbf{Q}}, \mathbf{x} - \varepsilon\mathbf{e}_i^m + \varepsilon\mathbf{e}_{i'}^m, \mathbf{y}) - H_t(\hat{\mathbf{Q}}, \mathbf{x}, \mathbf{y}) \\
= & -\sum_{j=1}^{n} \mu_j (r_{ij}^t - r_{i'j}^t) \\
& + \gamma E V_{t+1}(\alpha\hat{\mathbf{u}} - \alpha(\varepsilon - \sum_{j=1}^{n} \mu_j)\mathbf{e}_i^m + \alpha(\varepsilon - \sum_{j=1}^{n} \mu_j)\mathbf{e}_{i'}^m + \mathbf{D}^{t+1}, \beta\hat{\mathbf{v}} + \mathbf{S}^{t+1}) - \gamma E V_{t+1}(\alpha\hat{\mathbf{u}} + \mathbf{D}^{t+1}, \beta\hat{\mathbf{v}} + \mathbf{S}^{t+1}).
\end{aligned}$$

By the induction hypothesis, for each realization of $\mathbf{D}^{t+1}$ and $\mathbf{S}^{t+1}$, there exists $(\Lambda_1^\tau, \ldots, \Lambda_n^\tau)$ for $\tau = t+1, \ldots, T$ such that $\sum_{\tau=t+1}^T \alpha^{-(\tau-t-1)} \sum_{j=1}^n \Lambda_j^\tau \le \alpha(\varepsilon - \sum_{j=1}^n \mu_j)$ and

$$V_{t+1}(\alpha\hat{\mathbf{u}}-\alpha(\varepsilon-\sum_{j=1}^{n}\mu_j)\mathbf{e}_i^m+\alpha(\varepsilon-\sum_{j=1}^{n}\mu_j)\mathbf{e}_{i'}^m+\mathbf{D}^{t+1},\beta\hat{\mathbf{v}}+\mathbf{S}^{t+1})-V_{t+1}(\alpha\hat{\mathbf{u}}+\mathbf{D}^{t+1},\beta\hat{\mathbf{v}}+\mathbf{S}^{t+1})$$
$$\geq-\sum_{\tau=t+1}^{T}\gamma^{\tau-(t+1)}\sum_{j=1}^{n}\Lambda_j^\tau(r_{ij}^\tau-r_{i'j}^\tau).$$

Note that Λ_j^τ is a random variable due to its possible dependency on the random vectors $\mathbf{D}^{t+1}$ and $\mathbf{S}^{t+1}$. It then follows that

$$V_t(\mathbf{x}-\varepsilon\mathbf{e}_i^m+\varepsilon\mathbf{e}_{i'}^m,\mathbf{y})-V_t(\mathbf{x},\mathbf{y})\geq-\sum_{j=1}^{n}\mu_j(r_{ij}^t-r_{i'j}^t)-\gamma\sum_{\tau=t+1}^{T}\gamma^{\tau-(t+1)}\sum_{j=1}^{n}E\Lambda_j^\tau\cdot(r_{ij}^\tau-r_{i'j}^\tau).$$

Since $\sum_{\tau=t+1}^{T}\alpha^{-(\tau-t-1)}\sum_{j=1}^{n}\Lambda_j^\tau\leq\alpha(\varepsilon-\sum_{j=1}^{n}\mu_j)$, we have $\sum_{j=1}^{n}\mu_j+\sum_{\tau=t+1}^{T}\alpha^{-(\tau-t)}\sum_{j=1}^{n}\Lambda_j^\tau\leq\varepsilon$. Let $\lambda_j^t=\mu_j$ for all $j=1,\ldots,n$, and $\lambda_j^\tau=E\Lambda_j^\tau$ for all $j=1,\ldots,n$ and $\tau=t+1,\ldots,T$. The proof is completed. □

Lemma E.23. *(i) Suppose that (i,j) weakly precedes (i',j) by Definition 2. Then, transferring matching quantity from (i',j) to (i,j) weakly improves the total expected reward, i.e., $H_t(\mathbf{Q}+\varepsilon\mathbf{e}_{ij}^{m\times n}-\varepsilon\mathbf{e}_{i'j}^{m\times n},\mathbf{x},\mathbf{y})\geq H_t(\mathbf{Q},\mathbf{x},\mathbf{y})$, if $\mathbf{Q}+\varepsilon\mathbf{e}_{ij}^{m\times n}-\varepsilon\mathbf{e}_{i'j}^{m\times n}$ is a feasible decision under the state $(\mathbf{x},\mathbf{y})$.*

(ii) Similarly, if (i,j) precedes (i,j'), then $H_t(\mathbf{Q}+\varepsilon\mathbf{e}_{ij}^{m\times n}-\varepsilon\mathbf{e}_{ij'}^{m\times n},\mathbf{x},\mathbf{y})\geq H_t(\mathbf{Q},\mathbf{x},\mathbf{y})$.

Proof of Lemma E.23. We prove part (i) only since part (ii) can be proved analogously. The post-matching levels for using $\mathbf{Q}+\varepsilon\mathbf{e}_{ij}^{m\times n}-\varepsilon\mathbf{e}_{ij'}^{m\times n}$ are

$$\begin{aligned}\bar{\mathbf{u}}=&\mathbf{x}-\mathbf{1}^n(\mathbf{Q}+\varepsilon\mathbf{e}_{ij}^{m\times n}-\varepsilon\mathbf{e}_{i'j}^{m\times n})^\intercal=\mathbf{x}-\mathbf{1}^n\mathbf{Q}-\varepsilon\mathbf{e}_i^m+\varepsilon\mathbf{e}_{i'}^m=\mathbf{u}-\varepsilon\mathbf{e}_i^m+\varepsilon\mathbf{e}_{i'}^m,\\ \bar{\mathbf{v}}=&\mathbf{y}-\mathbf{1}^m(\mathbf{Q}+\varepsilon\mathbf{e}_{ij}^{m\times n}-\varepsilon\mathbf{e}_{i'j}^{m\times n})=\mathbf{y}-\mathbf{1}^m\mathbf{Q}=\mathbf{v},\end{aligned}$$

where $(\mathbf{u},\mathbf{v})$ are the post-matching levels by using the decision $\mathbf{Q}$ in period t. Then,

$$\begin{aligned}&H_t(\mathbf{Q}+\varepsilon\mathbf{e}_{ij}^{m\times n}-\varepsilon\mathbf{e}_{ij'}^{m\times n},\mathbf{x},\mathbf{y})-H_t(\mathbf{Q},\mathbf{x},\mathbf{y})\\ =&\varepsilon(r_{ij}^t-r_{ij'}^t)+\gamma EV_{t+1}(\alpha\bar{\mathbf{u}}+\mathbf{D}^{t+1},\beta\bar{\mathbf{v}}+\mathbf{S}^{t+1})-\gamma EV_{t+1}(\alpha\mathbf{u}+\mathbf{D}^{t+1},\beta\mathbf{v}+\mathbf{S}^{t+1}).\end{aligned}\tag{E.1}$$

By Lemma E.22, there exists $(\Lambda_1^\tau,\ldots,\Lambda_n^\tau)$ such that $\sum_{\tau=t+1}^{T}\alpha^{-(\tau-t-1)}\sum_{j'=1}^{n}\Lambda_{j'}^\tau\leq\alpha\varepsilon$ and

$$V_{t+1}(\alpha\bar{\mathbf{u}}+\mathbf{D}^{t+1},\beta\bar{\mathbf{v}}+\mathbf{S}^{t+1})-V_{t+1}(\alpha\mathbf{u}+\mathbf{D}^{t+1},\beta\mathbf{v}+\mathbf{S}^{t+1})\geq-\sum_{\tau=t+1}^{T}\gamma^{\tau-(t+1)}\sum_{j'=1}^{n}\Lambda_{j'}^\tau(r_{ij'}^\tau-r_{i'j'}^\tau).$$

Note that Λ_j^τ is a random variable since it may depend on $\mathbf{D}^{t+1}$ and $\mathbf{S}^{t+1}$.

Since (i,j) weakly precedes (i',j), it is easy to see from Definition 2 that $r^{\tau}_{ij'} - r^{\tau}_{i'j'} \le \gamma^{-(\tau-t)}\alpha^{-(\tau-t)}(r^t_{ij} - r^t_{i'j})$ for all $j' \in \mathcal{S}$ and $\tau = t+1,\dots,T$. Thus,

$$\begin{aligned}
&V_{t+1}(\alpha\bar{\mathbf{u}} + \mathbf{D}^{t+1}, \beta\bar{\mathbf{v}} + \mathbf{S}^{t+1}) - V_{t+1}(\alpha\mathbf{u} + \mathbf{D}^{t+1}, \beta\mathbf{v} + \mathbf{S}^{t+1}) \\
\ge& -\sum_{\tau=t+1}^{T}\gamma^{\tau-(t+1)}\sum_{j'=1}^{n}\Lambda^{\tau}_{j'}(r^{\tau}_{ij'} - r^{\tau}_{i'j'}) \\
\ge& -(r^t_{ij'} - r^t_{i'j'})\sum_{\tau=t+1}^{T}\gamma^{\tau-(t+1)}\gamma^{-(\tau-t)}\alpha^{-(\tau-t)}\sum_{j'=1}^{n}\Lambda^{\tau}_{j'} \\
=& -(\gamma\alpha)^{-1}(r^t_{ij'} - r^t_{i'j'})\sum_{\tau=t+1}^{T}\alpha^{-(\tau-t-1)}\sum_{j'=1}^{n}\Lambda^{\tau}_{j'} \\
\ge& (\gamma\alpha)^{-1}(r^t_{ij} - r^t_{i'j}) \times \alpha\varepsilon \\
=& -\gamma^{-1}(r^t_{ij} - r^t_{i'j})\varepsilon.
\end{aligned} \tag{E.2}$$

Combining (E.1) and (E.2), we have $H_t(\mathbf{Q} + \varepsilon\mathbf{e}^{m\times n}_{ij} - \varepsilon\mathbf{e}^{m\times n}_{ij'}, \mathbf{x}, \mathbf{y}) \ge H_t(\mathbf{Q}, \mathbf{x}, \mathbf{y})$. $\square$

Proof of Theorems 1 and 2. We now prove Theorems 1 and 2.

Let $\mathbf{Q}^{(k)}$ be a feasible decision in period t under the state $(\mathbf{x},\mathbf{y})$, and $(\mathbf{u}^{(k)}, \mathbf{v}^{(k)})$ be the corresponding post-matching levels.

For a pair (i,j), we consider two kinds of transfers of matching quantities from other pairs to (i,j), described as follows.

The first kind of transfers transfer matching quantities from a weakly preceded pair to the corresponding preceding pair. Suppose that $u^{(k)}_i > 0$. For another pair (i',j) such that it is weakly preceded by (i,j) and $q^{(k)}_{i'j} > 0$, we construct the feasible matching decision $\mathbf{Q}^{(k+1)} := \mathbf{Q}^{(k)} + \delta^{(k)}\mathbf{e}_{ij} - \delta^{(k)}\mathbf{e}_{i'j}$, where $\delta^{(k)} := \min\left\{u^{(k)}_i, q^{(k)}_{i'j}\right\}$. By Lemma E.23, $\mathbf{Q}^{(k+1)}$ weakly outperforms $\mathbf{Q}^{(k)}$. Likewise, if $v^{(k)}_j > 0$ and there exists another pair (i,j') weakly preceded by (i,j) such that $q^{(k)}_{ij'} > 0$, we construct the feasible matching decision $\mathbf{Q}^{(k+1)} := \mathbf{Q}^{(k)} + \delta^{(k)}\mathbf{e}_{ij} - \delta^{(k)}\mathbf{e}_{ij'}$, where $\delta^{(k)} := \min\left\{v^{(k)}_j, q^{(k)}_{ij'}\right\}$. Again by Lemma E.23, $\mathbf{Q}^{(k+1)}$ weakly outperforms $\mathbf{Q}^{(k)}$.

The second kind of transfers work as follows. Suppose that there exists two pairs (i',j) and (i,j') such that $r^t_{ij} + r^t_{i'j'} \ge r^t_{i'j} + r^t_{ij'}$, $q^{(k)}_{ij'} > 0$ and $q^{(k)}_{i'j} > 0$. Then, we construct the new feasible matching decision $\mathbf{Q}^{(k+1)} := \mathbf{Q}^{(k)} + \delta^{(k)}\mathbf{e}^{m\times n}_{ij} - \delta^{(k)}\mathbf{e}^{m\times n}_{i'j} - \delta^{(k)}\mathbf{e}^{m\times n}_{ij'}$, where $\delta^{(k)} := \min\left\{q^{(k)}_{i'j}, q^{(k)}_{ij'}\right\}$. It is easy to see that $\mathbf{Q}^{(k+1)}$ weakly outperforms $\mathbf{Q}^{(k)}$, since it leads to a weakly higher matching reward in period t (because $r^t_{ij} + r^t_{i'j'} \ge r^t_{i'j} + r^t_{ij'}$) than but the same post-matching levels as $\mathbf{Q}^{(k)}$.

We consider the procedure that repeatedly performs the first kind of transfers as long as possible for the proof of Theorem 1, and consider the procedure repeatedly apply any of the two kinds of transfers that is still possible for the proof of Theorem 2.

We define a pair (i, j) as a level 1 pair if it is not weakly preceded by any other pair; inductively, we define level ℓ pairs as those pairs weakly preceded only by level $\ell-1$ pairs, for $\ell \geq 2$. Since there are finitely many pairs, the total number of levels is a finite number, which we denote by L.

We can observe that for either kind of transfers, the weakly preceded pair(s) loses matching quantity. Thus, level 1 pairs never loses matching quantity. Since the matching quantity between any pair (i, j) cannot exceed $\min\{x_i, y_j\}$, either the procedure stops transferring matching quantity to level 1 pairs at some time point, or it never stops transfer quantities to level 1 pairs but the quantities transferred converges to zero. In either case, the quantities received by level 1 pairs converge to zero. A level 2 pair may lose matching quantity only when we transfer a quantity to level 1 pairs, and the quantity it loses is equal to the quantity received by the level 1 pair. This implies that the quantities that level 2 pairs lose converge to zero. As a result, the quantities received by level 2 pairs should also be equal to zero, since otherwise the matching quantities for some level 2 pair will grow to infinity.

As the induction hypothesis, let us assume that the matching quantities received by level ℓ pairs converges to zero for $\ell = 1, \ldots, \kappa$. A level $\kappa+1$ pair may lose matching quantity only when the procedure transfers a matching quantity to some level ℓ pair ($\ell \leq \kappa$), and the quantity it loses is equal to that received by the level ℓ pair. Thus, by the induction hypothesis, the quantities lost by level $\kappa+1$ pairs converges to zero. It follows that the quantities received by level $\kappa+1$ pairs should also converge to zero, since otherwise the matching quantity for some level $\kappa+1$ pair will grow to infinity. Therefore, by induction we have shown that the matching quantity transferred converges to zero, i.e., $\lim_{k\to\infty} \delta^{(k)} = 0$.

Let $\mathbf{Q}^{(\infty)}$ be a limiting point of the series $\left\{\mathbf{Q}^{(k)}\right\}_{k=1,2,\ldots}$. Since all $\mathbf{Q}^{(k)}$'s are feasible matching decisions, $\mathbf{Q}^{(\infty)}$ is also feasible. Clearly, $\mathbf{Q}^{(\infty)}$ weakly outperforms any $\mathbf{Q}^{(k)}$. We let $(\mathbf{u}^{(\infty)}, \mathbf{v}^{\infty})$ be the post-matching levels corresponding to $\mathbf{Q}^{(\infty)}$. it is easy to see that $(\mathbf{u}^{(\infty)}, \mathbf{v}^{\infty})$ is a limiting point of the series $(\mathbf{u}^{(k)}, \mathbf{v}^{(k)})$.

For the first kind of transfers, we have $\delta^{(k)} = \min\left\{u_i^{(k)}, q_{i'j}^{(k)}\right\}$ for the type i' such that (i, j) weakly precedes (i', j), or $\delta^{(k)} = \min\left\{v_j^{(k)}, q_{ij'}^{(k)}\right\}$ for the type j' such that (i, j) weakly precedes (i, j'). The fact $\lim_{k\to\infty} \delta^{(k)} = 0$ implies that $\min\left\{u_i^{(\infty)}, q_{i'j}^{(\infty)}\right\} = 0$ and $\min\left\{v_j^{(\infty)}, q_{ij'}^{(\infty)}\right\} = 0$ for the corresponding pairs (i', j) and (i, j').

For the second kind of transfers, we have $\delta^{(k)} = \min\left\{q_{i'j}^{(k)}, q_{ij'}^{(k)}\right\}$ for the types i' and j' such that (i, j) weakly precedes both (i', j) and (i, j') and $r_{ij}^t + r_{i'j'}^t \geq r_{i'j}^t + r_{ij'}^t$. The fact $\lim_{k\to\infty} \delta^{(k)} = 0$ implies that $\min\left\{q_{i'j}^{(\infty)}, q_{ij'}^{(\infty)}\right\} = 0$.

From the above arguments, we see that $\mathbf{Q}^{(\infty)}$ satisfies the properties in Theorems 1 and/or 2. Therefore, we have shown that given any feasible matching decision $\mathbf{Q}^{(1)}$ in period t, we can construct another feasible decision that weakly outperforms $\mathbf{Q}^{(\infty)}$. The above analysis can be applied to any state in any period. This implies that there exists an optimal matching decision that satisfies the desired properties in the two theorems. $\square$

Proof of Special Case 2 We show that the highest-(priority)-level pair always weakly precedes all remaining pairs throughout the procedure of removing demand/supply types. If that is true, we either match the highest-level pair to the maximum (and move on to the new highest-level pair after removing the demand/supply type that is exhausted), or stop the matching after partially matching the highest-level pair.

Suppose that (i,j) is highest-level pair among the remaining pairs, and is of level ℓ. By assumption, (i,j) is the only level ℓ pair among the remaining pairs. Let (i',j') of level $\ell' > \ell$ be of the next-highest level among remaining pairs. (If there are multiple remaining pairs of level ℓ', we can arbitrarily choose a pair).

Suppose to the contrary that (i',j') is not weakly preceded by (i,j). Then, (i',j') has a neighboring pair of level $\ell'-1$, say (i',j''), whose supply type has been previously removed.

If $\ell'-1 > \ell$, then we have removed a lower level pair (i.e., level $\ell'-1$) before a higher level pair (i.e., level ℓ). This is impossible since we always remove the demand or supply type of the highest-level pair.

If $\ell'-1 = \ell$, then we had two level ℓ pairs when we removed j'' (and thus the pair (i',j'')). This contradicts the assumption that there is always just one pair of the highest level among the remaining pairs. $\square$

Proof of Theorem 3. We show that greedy matching between i and j is optimal by induction. It is easy to verify that greedy matching between i and j is optimal in the final period T. Suppose that it is also optimal in period $t+1$.

Let $\mathbf{Q}$ be an optimal decision in period t under the state $(\mathbf{x},\mathbf{y})$, such that it satisfies the properties in Theorems 1 and 2. Suppose to the contrary that $q_{ij}^t < \min\{x_i, y_j\}$. We first show that under this assumption, both the post-matching levels u_i^t and v_j^t corresponding to $\mathbf{Q}$ are positive, if (i,j) weakly precedes all its neighboring pairs. To prove that, let us suppose to the contrary that $u_i^t = 0$. Since $q_{ij} < x_i$, there is a pair (i,j') such that $q_{ij'}^t > 0$. Following Theorem 2, we have $q_{i'j}^t = 0$ for all demand type $i' \neq i$. As a result, $v_j^t = y_j^t - \sum_{i'=1}^m q_{i'j} = y_j^t - q_{ij}^t > 0$. However, both $q_{ij'}^t$ and v_j^t being positive contradicts Theorem 1, given that (i,j) weakly precedes (i,j'). Thus, both u_i^t and v_j^t are positive.

Again by Theorem 1, both u_i^t and v_j^t being positive implies that $q_{ij'}^t = 0$ and $q_{i'j}^t = 0$ for all demand type $i' \neq i$ and all supply type $j' \neq j$. This means that we can increase the matching quantity between i and j until it equals $\min\{x_i, y_j\}$ without the need to change the matching quantity between any other pair. Next, we will show that increasing the matching quantity between i and j by $\varepsilon := \min\{x_i, y_j\}$ does not hurt the optimality of $\mathbf{Q}$.

Increasing the matching quantity between i and j by ε will increase the matching reward in period t by $r_{ij}^t \varepsilon$, but decrease both the post-matching levels of i and j by ε. In other words:

$$\begin{aligned} &H_t(\mathbf{Q} + \varepsilon \mathbf{e}_{ij}^{m\times n}, \mathbf{x}, \mathbf{y}) - H_t(\mathbf{Q} + \varepsilon \mathbf{e}_{ij}^{m\times n}, \mathbf{x}, \mathbf{y}) \\ =& r_{ij}^t \varepsilon + \gamma E V_{t+1}(\alpha \mathbf{u}^t - \alpha\varepsilon \mathbf{e}_i^m + \mathbf{D}^{t+1}, \beta \mathbf{v}^t - \beta \varepsilon \mathbf{e}_j^n + \mathbf{S}^{t+1}) - \gamma E V_{t+1}(\alpha \mathbf{u}^t + \mathbf{D}^{t+1}, \beta \mathbf{v}^t + \mathbf{S}^{t+1}). \end{aligned}$$

Let us consider the case $\beta \geq \alpha$ without loss of generality. We have

$$\begin{aligned} V_{t+1}(\alpha \mathbf{u}^t + \mathbf{D}^{t+1} - \alpha\varepsilon \mathbf{e}_i^m, \beta \mathbf{v}^t + \mathbf{S}^{t+1} - \beta\varepsilon \mathbf{e}_j^n) =& V_{t+1}(\alpha \mathbf{u}^t + \mathbf{D}^{t+1} + (\beta - \alpha)\varepsilon \mathbf{e}_i^m, \beta \mathbf{v}^t + \mathbf{S}^{t+1}) - \beta \varepsilon r_{ij}^{t+1} \\ \geq& V_{t+1}(\alpha \mathbf{u}^t + \mathbf{D}^{t+1}, \beta \mathbf{v}^t + \mathbf{S}^{t+1}) - \beta \varepsilon r_{ij}^{t+1} \end{aligned}$$

where the equality is because of the greedy matching of pair (i, j) for the subsequent periods, and the inequality holds because V_{t+1} is increasing in the state vector (note more demand/supply always leads to weakly higher reward since the firm has the option of never using the extra demand/supply). Therefore,

$$H_t(\mathbf{Q} + \varepsilon \mathbf{e}_{ij}^{m\times n}, \mathbf{x}, \mathbf{y}) - H_t(\mathbf{Q} + \varepsilon \mathbf{e}_{ij}^{m\times n}, \mathbf{x}, \mathbf{y}) \geq (r_{ij}^t - \gamma\beta r_{ij}^{t+1})\varepsilon \geq 0.$$

Therefore, we can always weakly improve $\mathbf{Q}^t$ by increasing the matching quantity q_{ij}^t until it is equal to $\min\{x_i, y_j\}$. It is easy to see that by doing so, properties in Theorems 1 and 2 remain satisfied. Thus, we can do the same for all pairs (i, j) that are not greedily matched until we obtain an optimal solution that satisfies the theorem. □

Proof of Proposition 1. The proof is based on an alternative formulation of the 2×2 horizontal model in Online Supplement A, and the proof itself is also included in Online Supplement A. □

Proof of Proposition 2. (i) We focus on the case with $z_1 \geq 0$ and $z_2 \geq 0$ (or equivalently, $x_1 \geq y_1$ and $x_2 \leq y_2$) to show that $p_{s_2}^t(IB)$ decreases in IB with the rate of decrease less than or equal to 1.

According to Lemma A.4, the function $\tilde{J}_t(q, \mathbf{z})$ (defined by (A.6) in Lemma A.2) is $L^\natural$-concave, the optimal matching quantity q_{12}^{t*} is increasing in both z_1 and z_2, with the rates of increase less than or equal to 1. By Proposition 1, $q_{12}^{t*} = [z_2 - p_{s_2}^t(IB)]^+$. For a given IB, let us choose z_1 and z_2

both sufficiently large such that $z_1 - z_2 = IB$ and $z_2 > p^t_{s_2}(IB)$. Therefore we have $q^{t*}_{12} = z_2 - p^t_{s_2}(IB)$, or equivalently, $p^t_{s_2}(IB) = z_2 - q^{t*}_{12} = z_1 - IB - q^{t*}_{12}$. If we keep z_1 fixed and increase IB, z_2 will increase at the same rate as IB decreases. Since q^{t*}_{12} increases in z_2 at a rate no greater than 1, $-q^{t*}_{12}$ increases in IB at a rate that is no greater than 1. Consequently, $p^t_{s_2}(IB)$ decreases in IB at a rate no greater than 1.

(ii) Part (ii) of the proposition is proved in Lemma A.5 in Online Appendix A. □

Proof of Lemma 1. Let o and d be the two endpoints of the line segment $\mathcal{L}$, and suppose that the direction of $\mathcal{L}$ from o to d. We first show that (i,j) weakly precedes (i',j), if j reaches i before it reaches i', along the direction.

It is clear that $\text{dist}_{i\leftarrow j} \le \text{dist}_{i'\leftarrow j}$. By Assumption 3 i), we have $r^t_{ij} = R^t_i - \text{dist}_{i\leftarrow j} \ge R^t_{i'} - \text{dist}_{i'\leftarrow j} = r^t_{i'j}$ for all t. To show that (i,j) weakly precedes (i',j), it remains to verify that $r^t_{ij} - r^t_{i'j} \ge \gamma\alpha(r^{t+1}_{ij''} - r^{t+1}_{i'j''})$. To that end, we have

$$r^t_{ij} - r^t_{i'j} = (R^t_i - \text{dist}_{i\leftarrow j}) - (R^t_{i'} - \text{dist}_{i'\leftarrow j}) = R^t_i - R^t_{i'} + \text{dist}_{i'\leftarrow j} - \text{dist}_{i\leftarrow j} = R^t_i - R^t_{i'} + \text{dist}_{i'\leftarrow i}.$$

Now consider another supply type j''. We consider the following two possibilities.

If j'' is located between i and endpoint d, we have $r^{t+1}_{i\leftarrow j''} = 0$ since i is not accessible from j''. Then, $\gamma\alpha(r^{t+1}_{ij''} - r^{t+1}_{i'j''}) = -\gamma\alpha r^{t+1}_{i'j''} \le 0 \le r^t_{ij} - r^t_{i'j}$.

If j'' is located between endpoint o and i, then

$$\begin{aligned}\gamma\alpha(r^{t+1}_{ij''} - r^{t+1}_{i'j''}) =&\gamma\alpha[(R^{t+1}_i - \text{dist}_{i\leftarrow j''}) - (R^{t+1}_{i'} - \text{dist}_{i'\leftarrow j''})]\\ =&\gamma\alpha(R^{t+1}_i - R^{t+1}_{i'}) + \gamma\alpha(\text{dist}_{i'\leftarrow j''} - \text{dist}_{i\leftarrow j''}).\end{aligned}$$

Since j reaches i before i' along the direction, it follows for Assumption 3 ii) that $R^t_i - R^t_{i'} \ge \gamma\alpha(R^{t+1}_i - R^{t+1}_{i'})$. We have

$$\begin{aligned}\gamma\alpha(r^{t+1}_{ij''} - r^{t+1}_{i'j''}) =&\gamma\alpha(R^{t+1}_i - R^{t+1}_{i'}) + \gamma\alpha(\text{dist}_{i'\leftarrow j''} - \text{dist}_{i\leftarrow j''})\\ \le& R^t_i - R^t_{i'} + \gamma\alpha(\text{dist}_{i'\leftarrow j''} - \text{dist}_{i\leftarrow j''})\\ \le& R^t_i - R^t_{i'} + \text{dist}_{i'\leftarrow j''} - \text{dist}_{i\leftarrow j''}\\ =&(R^t_i - \text{dist}_{i\leftarrow j''}) - (R^t_{i'} - \text{dist}_{i'\leftarrow j''}) = r^t_{ij''} - r^t_{i'j}.\end{aligned}$$

Therefore (i,j) weakly precedes (i',j).

Conversely, if j reaches i' before i, it is easy to see that $r^t_{ij} < r^t_{i'j}$. In that case, (i,j) does not weakly precede (i',j). Thus, (i,j) weakly precedes (i',j) if and only if j reaches i before i' along the direction, provided that both pairs are matchable.

Analogously, we can show that (i,j) weakly precedes (i,j') if and only if j is closer to i than j' along the direction.

To show that the strong modified Monge condition is satisfied, let us consider a matchable pair (i,j) that weakly precedes two other pairs, (i',j) and (i,j'). We need to show that $r^t_{ij} + r^t_{i'j'} \geq r^t_{i'j} + r^t_{ij'}$. It is easy to see that the inequality holds if either (i',j) or (i,j') is unmatchable. (For instance, if (i',j) is unmatchable, the inequality holds because $r^t_{ij} \geq r^t_{ij'}$ and $r^t_{i'j} = 0$.) Suppose that both (i',j) and (i,j') are matchable. We have $\text{dist}_{i\leftarrow j'} + \text{dist}_{i'\leftarrow j} = (\text{dist}_{i\leftarrow j} + \text{dist}_{j\leftarrow j'}) + (\text{dist}_{i\leftarrow j} + \text{dist}_{i'\leftarrow i}) = \text{dist}_{i\leftarrow j} + (\text{dist}_{j\leftarrow j'} + \text{dist}_{i\leftarrow j} + \text{dist}_{i'\leftarrow i}) = \text{dist}_{i\leftarrow j} + \text{dist}_{i'\leftarrow j'}$. It follows that $r^t_{ij} + r^t_{i'j'} = R^t_i - \text{dist}_{i\leftarrow j} + R_{i'} - \text{dist}_{i'\leftarrow j'} = R^t_i - \text{dist}_{i\leftarrow j'} + R_{i'} - \text{dist}_{i'\leftarrow j} = r^t_{i'j} + r^t_{ij'}$. □

Proof of Proposition 3. The proof follows directly from Lemma 1 and Special Case 1. □

Proof of Corollary 1. The proof follows directly from Theorem 3. □

Proof of Proposition 4. For any demand types i, i' and any supply types j, j', we have $r^t_{ij} + r^t_{i'j'} = r^t_{id} + r^t_{i'd} + r^t_{js} + r^t_{j's} = r^t_{ij'} + r^t_{i'j}$. Thus, the strong modified Monge condition (i.e., Assumption 1) is satisfied. Moreover, if we remove any demand and/or supply types from the bipartite graph, in the remaining graph, the demand type and the supply type of the lowest indices form the only pair of the highest priority level. Thus, the proposition follows from Special Case 2. □

Proof of Lemma 2. A total matching quantity $\bar{Q}$ in period t implies that the quantity of demand fulfilled and that of supply used are both equal to $\bar{Q}$. Under top-down matching, demand and supply types with smaller indices are matched first. Therefore, the firm will fulfill type 1 demand in period t for the quantity $\min\left\{\bar{Q}, x_1\right\}$. Then, the remaining quantity of demand to fulfill is $\bar{Q} - \min\left\{\bar{Q}, x_1\right\} = (\bar{Q} - x_1)^+$. The firm will fulfill type 2 demand for the quantity $\min\left\{(\bar{Q} - x_1)^+, x_2\right\}$ (since demand type 2 is prioritized over all other demand types except demand type 1). Recursively, the quantity of type i demand to fulfill in period t is equal to $\min\left\{(\bar{Q} - \sum_{i'=1}^{i-1} x_{i'})^+, x_i\right\}$. Likewise, the quantity of type j supply to fulfill in period t is $\min\left\{(\bar{Q} - \sum_{j'=1}^{j-1} y_{j'})^+, y_j\right\}$. Since the reward for fulfilling a unit of type i demand is r^t_i and that for fulfilling a unit of type j supply is r^t_j, the matching reward received in period t is $\sum_{i=1}^{m} r^t_i \min\left\{(\bar{Q} - \sum_{i'=1}^{i-1} x_{i'})^+, x_i\right\} + \sum_{j=1}^{n} r^t_j \min\left\{(\bar{Q} - \sum_{j'=1}^{j-1} y_{j'})^+, y_j\right\}$. It is easy to see that the post matching levels in period t are given by $\mathbf{u}=(\mathbf{x}-\bar{Q}\mathbf{1}^m):=\big((x_1-\bar{Q})^+,\dots,(x_m-\bar{Q})^+\big)$ and $\mathbf{v}=(\mathbf{y}-\bar{Q}\mathbf{1}^n)^+:=\big((y_1-\bar{Q})^+,\dots,(y_m-\bar{Q})^+\big)$. Therefore, the maximum expected total discounted reward attainable by the total matching quantity $\bar{Q}$ is given by the expression of $G_t(\bar{Q},\mathbf{x},\mathbf{y})$ given in (3). And the optimal reward $V_t(\mathbf{x},\mathbf{y})$ can be obtained by maximizing $G_t(\bar{Q},\mathbf{x},\mathbf{y})$ with respect to $0\le\bar{Q}\le\min\left\{\sum_{i=1}^m x_i,\sum_{j=1}^n y_j\right\}$ (note that the total matching quantity cannot exceed the total available demand or supply).

Next, we show the concavity of $G_t(\bar{Q},\mathbf{x},\mathbf{y})$ with respect to $\bar{Q}$.

The expected value function $EV_{t+1}(\alpha\mathbf{u}+\mathbf{D}^{t+1},\beta\mathbf{v}+\mathbf{S}^{t+1})$ is concave in $(\mathbf{u},\mathbf{v})$ for both the continuous-valued (state and decision) model and the discrete-valued model with either $\alpha=\beta=1$ or $\alpha=0,\beta=1$ (In Lemma B.7 in Online Supplement B, we show that in the latter case it is $L^\natural$-concave, which implies concavity.) It follows that $G_t(Q,\mathbf{x},\mathbf{y})$ is concave in Q within the interior of the ranges $\tilde{x}_{i-1}\le Q<\tilde{x}_i$ and $\tilde{y}_{j-1}\le Q<\tilde{y}_j$.

Let $\tilde{x}_i:=\sum_{k=1}^i x_k$ and $\tilde{y}_j:=\sum_{k=1}^j y_k$. Without loss of generality, we assume that $\tilde{x}_i\in(\tilde{y}_{j-1},\tilde{y}_j)$. We show that G_t is concave in the neighborhood of a breakpoint $a=\tilde{x}_i$. To this end, it suffices to show that $G_t(a+\epsilon,\mathbf{x},\mathbf{y})-G_t(a,\mathbf{x},\mathbf{y})\le G_t(a,\mathbf{x},\mathbf{y})-G_t(a-\epsilon,\mathbf{x},\mathbf{y})$, where $0<\epsilon<\min\{\tilde{x}_i-\tilde{y}_{j-1},\tilde{y}_j-\tilde{x}_i\}$. On the one hand, we have

$$\begin{aligned}
&G_t(a,\mathbf{x},\mathbf{y})-G_t(a-\epsilon,\mathbf{x},\mathbf{y})\\
=&(r^t_{id}+r^t_{js})\epsilon+\gamma EV_{t+1}(\mathbf{D}_{[1,i]},\alpha\mathbf{x}_{[i+1,n]}+\mathbf{D}_{[i+1,n]},\mathbf{S}_{[1,j-1]},\beta(\tilde{y}_j-\tilde{x}_i)+S_j,\beta\mathbf{y}_{[j+1,m]}+\mathbf{S}_{[j+1,m]})\\
&-\gamma EV_{t+1}(\mathbf{D}_{[1,i-1]},\alpha\epsilon+D_i,\alpha\mathbf{x}_{[i+1,n]}+\mathbf{D}_{[i+1,n]},\mathbf{S}_{[1,j-1]},\beta(\tilde{y}_j-\tilde{x}_i)+\beta\epsilon+S_j,\beta\mathbf{y}_{[j+1,m]}+\mathbf{S}_{[j+1,m]}).
\end{aligned}\tag{E.3}$$

By Lemma E.22, there exists $\lambda^\tau_{j'}$ for $j'=1,\dots,m$ and $\tau=t+1,\dots,T$ such that $\sum_{\tau=t+1}^T\alpha^{\tau-t-1}\sum_{j'=1}^m\lambda^\tau_{j'}\le\alpha\varepsilon$, and

$$\begin{aligned}
&-EV_{t+1}(\mathbf{D}_{[1,i-1]},\alpha\epsilon+D_i,\alpha\mathbf{x}_{[i+1,n]}+\mathbf{D}_{[i+1,n]},\mathbf{S}_{[1,j-1]},\beta(\tilde{y}_j-\tilde{x}_i)+S_j,\beta\mathbf{y}_{[j+1,m]}+\mathbf{S}_{[j+1,m]})\\
&+EV_{t+1}(\mathbf{D}_{[1,i]},\alpha x_{i+1}+D_{i+1}+\alpha\epsilon,\alpha\mathbf{x}_{[i+2,n]}+\mathbf{D}_{[i+2,n]},\mathbf{S}_{[1,j-1]},\beta(\tilde{y}_j-\tilde{x}_i)+\beta\epsilon+S_j,\beta\mathbf{y}_{[j+1,m]}+\mathbf{S}_{[j+1,m]})\\
\ge&-\sum_{\tau=t+1}^T\gamma^{\tau-t-1}\sum_{j'=1}^m\lambda^\tau_{j'}(r^\tau_{ij'}-r^\tau_{i+1,j'}).
\end{aligned}\tag{E.4}$$

Combining (E.3) and (E.4), we have

$$\begin{aligned}
&G_t(a,\mathbf{x},\mathbf{y})-G_t(a-\epsilon,\mathbf{x},\mathbf{y})\\
\ge&(r^t_{id}+r^t_{js})\epsilon-\gamma\sum_{\tau=t+1}^T\gamma^{\tau-t-1}\sum_{j'=1}^m\lambda^\tau_{j'}(r^\tau_{ij'}-r^\tau_{i+1,j'})
\end{aligned}$$

$$
\begin{aligned}
&+\gamma EV_{t+1}(\mathbf{D}_{[1,i]},\alpha\mathbf{x}_{[i+1,n]}+\mathbf{D}_{[i+1,n]},\mathbf{S}_{[1,j-1]},\beta(\tilde{y}_j-\tilde{x}_i)+S_j,\beta\mathbf{y}_{[j+1,m]}+\mathbf{S}_{[j+1,m]})\\
&-\gamma EV_{t+1}(\mathbf{D}_{[1,i]},\alpha x_{i+1}+D_{i+1}+\alpha\epsilon,\alpha\mathbf{x}_{[i+2,n]}+\mathbf{D}_{[i+2,n]},\mathbf{S}_{[1,j-1]},\beta(\tilde{y}_j-\tilde{x}_i)+\beta\epsilon+S_j,\beta\mathbf{y}_{[j+1,m]}+\mathbf{S}_{[j+1,m]})\\
\geq&(r^t_{id}+r^t_{js})\epsilon-\gamma\sum_{\tau=t+1}^{T}\gamma^{\tau-t-1}(\gamma\alpha)^{-(\tau-t)}\sum_{j'=1}^{m}\lambda^\tau_{j'}(r^t_{ij'}-r^t_{i+1,j'})\\
&+\gamma EV_{t+1}(\mathbf{D}_{[1,i]},\alpha\mathbf{x}_{[i+1,n]}+\mathbf{D}_{[i+1,n]},\mathbf{S}_{[1,j-1]},\beta(\tilde{y}_j-\tilde{x}_i)+S_j,\beta\mathbf{y}_{[j+1,m]}+\mathbf{S}_{[j+1,m]})\\
&-\gamma EV_{t+1}(\mathbf{D}_{[1,i]},\alpha x_{i+1}+D_{i+1}+\alpha\epsilon,\alpha\mathbf{x}_{[i+2,n]}+\mathbf{D}_{[i+2,n]},\mathbf{S}_{[1,j-1]},\beta(\tilde{y}_j-\tilde{x}_i)+\beta\epsilon+S_j,\beta\mathbf{y}_{[j+1,m]}+\mathbf{S}_{[j+1,m]})\\
=&(r^t_{id}+r^t_{js})\epsilon-\alpha^{-1}\sum_{\tau=t+1}^{T}\alpha^{-(\tau-t-1)}\sum_{j'=1}^{m}\lambda^\tau_{j'}(r^t_{id}-r^t_{i+1,d})\\
&+\gamma EV_{t+1}(\mathbf{D}_{[1,i]},\alpha\mathbf{x}_{[i+1,n]}+\mathbf{D}_{[i+1,n]},\mathbf{S}_{[1,j-1]},\beta(\tilde{y}_j-\tilde{x}_i)+S_j,\beta\mathbf{y}_{[j+1,m]}+\mathbf{S}_{[j+1,m]})\\
&-\gamma EV_{t+1}(\mathbf{D}_{[1,i]},\alpha x_{i+1}+D_{i+1}+\alpha\epsilon,\alpha\mathbf{x}_{[i+2,n]}+\mathbf{D}_{[i+2,n]},\mathbf{S}_{[1,j-1]},\beta(\tilde{y}_j-\tilde{x}_i)+\beta\epsilon+S_j,\beta\mathbf{y}_{[j+1,m]}+\mathbf{S}_{[j+1,m]})\\
\geq&(r^t_{id}+r^t_{js})\epsilon-(r^t_{id}-r^t_{i+1,d})\varepsilon\\
&+\gamma EV_{t+1}(\mathbf{D}_{[1,i]},\alpha\mathbf{x}_{[i+1,n]}+\mathbf{D}_{[i+1,n]},\mathbf{S}_{[1,j-1]},\beta(\tilde{y}_j-\tilde{x}_i)+S_j,\beta\mathbf{y}_{[j+1,m]}+\mathbf{S}_{[j+1,m]})\\
&-\gamma EV_{t+1}(\mathbf{D}_{[1,i]},\alpha x_{i+1}+D_{i+1}+\alpha\epsilon,\alpha\mathbf{x}_{[i+2,n]}+\mathbf{D}_{[i+2,n]},\mathbf{S}_{[1,j-1]},\beta(\tilde{y}_j-\tilde{x}_i)+\beta\epsilon+S_j,\beta\mathbf{y}_{[j+1,m]}+\mathbf{S}_{[j+1,m]})\\
=&(r^t_{i+1,d}+r^t_{js})\varepsilon+\gamma EV_{t+1}(\mathbf{D}_{[1,i]},\alpha\mathbf{x}_{[i+1,n]}+\mathbf{D}_{[i+1,n]},\mathbf{S}_{[1,j-1]},\beta(\tilde{y}_j-\tilde{x}_i)+S_j,\beta\mathbf{y}_{[j+1,m]}+\mathbf{S}_{[j+1,m]})\\
&-\gamma EV_{t+1}(\mathbf{D}_{[1,i]},\alpha x_{i+1}+D_{i+1}+\alpha\epsilon,\alpha\mathbf{x}_{[i+2,n]}+\mathbf{D}_{[i+2,n]},\mathbf{S}_{[1,j-1]},\beta(\tilde{y}_j-\tilde{x}_i)+\beta\epsilon+S_j,\beta\mathbf{y}_{[j+1,m]}+\mathbf{S}_{[j+1,m]}),
\end{aligned}
$$

where the second inequality holds because $r^t_{ij'}-r^t_{i+1,j'}=r^\tau_{id}-r^t_{i+1,d}\geq(\gamma\alpha)^{\tau-t}(r^\tau_{id}-r^\tau_{i+1,d})=(\gamma\alpha)^{\tau-t}(r^\tau_{ij'}-r^\tau_{i+1,j'})$ for $\tau\geq t+1$ according to Assumption 4. On the other hand, we have

$$
\begin{aligned}
&G_t(a+\epsilon,\mathbf{x},\mathbf{y})-G_t(a,\mathbf{x},\mathbf{y})\\
=&(r^t_{i+1,d}+r^t_{js})\epsilon\\
&+\gamma EV_{t+1}(\mathbf{D}_{[1,i]},\alpha x_{i+1}-\alpha\epsilon+D_{i+1},\alpha\mathbf{x}_{[i+2,n]}+\mathbf{D}_{[i+2,n]},\mathbf{S}_{[1,j-1]},\beta(\tilde{y}_j-\tilde{x}_i)-\beta\epsilon+S_j,\beta\mathbf{y}_{[j+1,m]}+\mathbf{S}_{[j+1,m]})\\
&-\gamma EV_{t+1}(\mathbf{D}_{[1,i]},\alpha x_{i+1}+D_{i+1},\alpha\mathbf{x}_{[i+2,n]}+\mathbf{D}_{[i+2,n]},\mathbf{S}_{[1,j-1]},\beta(\tilde{y}_j-\tilde{x}_i)+S_j,\beta\mathbf{y}_{[j+1,m]}+\mathbf{S}_{[j+1,m]}).
\end{aligned}
$$

By the concavity of V_{t+1}, we have $G_t(a+\epsilon,\mathbf{x},\mathbf{y})-G_t(a,\mathbf{x},\mathbf{y})\leq G_t(a,\mathbf{x},\mathbf{y})-G_t(a-\epsilon,\mathbf{x},\mathbf{y})$. □

Proof of Proposition 5. We use the alternative formulation (B.14)–(B.15) of the model in Online Supplement B. We first focus on the case with $\alpha=\beta>0$, and prove the functional properties in parts i) and ii) for that case.

(i) By Lemma B.7 i) in Online Appendix B, the function $\tilde{G}_t(Q,\tilde{\mathbf{x}},\tilde{\mathbf{y}})$ is $L^\natural$-concave, a fortiori, supermodular in $(Q,\tilde{\mathbf{x}},\tilde{\mathbf{y}})$. (Note that $L^\natural$-concavity implies supermodularity.)

By Simchi-Levi et al. (2014, Theorem 2.2.8), the optimal solution to (B.14), denoted by $\hat{Q}^t(\tilde{\mathbf{x}},\tilde{\mathbf{y}})$, is increasing in x_i and y_j, for $i=1,\ldots,m$ and $j=1,\ldots,n$. Given the relation $\tilde{x}_i=\sum_{k=1}^{i}x_i$ and

$\tilde{y}_j = \sum_{k=1}^{j} y_j$ between the transformed state $(\tilde{\mathbf{x}}, \tilde{\mathbf{y}})$ and the original state $(\mathbf{x}, \mathbf{y})$, we see that both $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{y}}$ increase, as x_i or y_j. Thus, the optimal total matching quantity $\bar{Q}^{t*}(\mathbf{x}, \mathbf{y})$ is increasing in x_i and y_j, for $i = 1, \ldots, m$ and $j = 1, \ldots, n$.

Next, we show that the rates of increase do not exceed 1. Let ϵ be a positive number. By the definition of $L^\natural$-concavity, $\tilde{G}_t(Q - \xi, \tilde{\mathbf{x}} - \xi\mathbf{1}^m, \tilde{\mathbf{y}} - \xi\mathbf{1}^n)$ is supermodular in $(Q, \tilde{\mathbf{x}}, \tilde{\mathbf{y}}, \xi)$. Then, for $Q > \hat{Q}^t(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) + \epsilon$, we have

$$\tilde{G}_t(Q, \tilde{\mathbf{x}} + \epsilon\mathbf{1}^m, \tilde{\mathbf{y}} + \epsilon\mathbf{1}^n) - \tilde{G}_t(\hat{Q}_t(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) + \epsilon, \tilde{\mathbf{x}} + \epsilon\mathbf{1}^m, \tilde{\mathbf{y}} + \epsilon\mathbf{1}^n) \leq \tilde{G}_t(Q - \epsilon, \tilde{\mathbf{x}}, \tilde{\mathbf{y}}) - \tilde{G}_t(\hat{Q}_t(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}), \tilde{\mathbf{x}}, \tilde{\mathbf{y}}) \leq 0,$$

where the first inequality is derived by definition of supermodularity and the second inequality is due to the optimality of $\hat{Q}^t$. This implies that any matching quantity $Q > \hat{Q}_t(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) + \epsilon$ is not better than $\hat{Q}_t(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) + \epsilon$ for the state $(\tilde{\mathbf{x}} + \epsilon\mathbf{1}^m, \tilde{\mathbf{y}} + \epsilon\mathbf{1}^n)$. Therefore, $\hat{Q}_t(\tilde{\mathbf{x}} + \epsilon\mathbf{1}^m, \tilde{\mathbf{y}} + \epsilon\mathbf{1}^n) \leq \hat{Q}_t(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) + \epsilon$. By the monotonicity of $\hat{Q}_t(\tilde{\mathbf{x}}, \tilde{\mathbf{y}})$, $\hat{Q}_t(\tilde{\mathbf{x}} + \epsilon\mathbf{1}^m, \tilde{\mathbf{y}}) \leq \hat{Q}_t(\tilde{\mathbf{x}} + \epsilon\mathbf{1}^m, \tilde{\mathbf{y}} + \epsilon\mathbf{1}^n) \leq \hat{Q}_t(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) + \epsilon$. For $1 \leq i \leq m$, let $\mathbf{1}_{[i,m]}$ be the m-dimension vector with its first $i-1$ entries equal to 0, and the remaining entries equal to 1. We have $\bar{Q}^{t*}(\mathbf{x} + \epsilon\mathbf{e}_i, \mathbf{y}) = \hat{Q}^t(\tilde{\mathbf{x}} + \epsilon\mathbf{1}_{[i,m]}, \mathbf{y}) \leq \hat{Q}^t(\tilde{\mathbf{x}} + \epsilon\mathbf{1}^m, \tilde{\mathbf{y}}) \leq \hat{Q}^t(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) + \epsilon = \bar{Q}^{t*}(\mathbf{x}, \mathbf{y}) + \varepsilon$. Therefore, the rate of increase of $\bar{Q}^{t*}(\mathbf{x}, \mathbf{y})$ with respect to x_i is less than or equal to 1. Analogously, we can show that the rate of increase of $\bar{Q}^{t*}(\mathbf{x}, \mathbf{y})$ with respect to y_j is less than or equal to 1.

(ii) To show that $\bar{Q}^{t*}(\mathbf{x}, \mathbf{y})$ increases faster with respect to x_i than x_{i+1} (for $i = 1, \ldots, m-1$), we consider two original states $(\mathbf{x} + \epsilon\mathbf{e}_i^n, \mathbf{y})$ and $(\mathbf{x} + \epsilon\mathbf{e}_{i+1}^n, \mathbf{y})$. Their transformed states can be ordered as $(\tilde{\mathbf{x}} + \epsilon\mathbf{1}_{[i,m]}, \tilde{\mathbf{y}}) \geq (\tilde{\mathbf{x}} + \epsilon\mathbf{1}_{[i+1,m]}, \tilde{\mathbf{y}})$ (recall that $\mathbf{1}_{[k,m]}$ is the m-dimension vector with the i-th up to m-th entry being 1 and the rest of the entries being 0). By the monotonicity of $\hat{Q}_t(\tilde{\mathbf{x}}, \tilde{\mathbf{y}})$, $\hat{Q}_t(\tilde{\mathbf{x}} + \epsilon\mathbf{1}_{[k,m]}, \tilde{\mathbf{y}}) \geq \hat{Q}_t(\tilde{\mathbf{x}} + \epsilon\mathbf{1}_{[k+1,m]}, \tilde{\mathbf{y}})$. This implies that $\bar{Q}^{t*}(\mathbf{x} + \epsilon\mathbf{e}_i, \mathbf{y}) \geq \bar{Q}^{t*}(\mathbf{x} + \epsilon\mathbf{e}_{i+1}^m, \mathbf{y})$. Thus we have $\bar{Q}^{t*}(\mathbf{x} + \epsilon\mathbf{e}_i, \mathbf{y}) - \bar{Q}^{t*}(\mathbf{x}, \mathbf{y}) \geq \bar{Q}^{t*}(\mathbf{x} + \epsilon\mathbf{e}_{i+1}^m, \mathbf{y}) - \bar{Q}^{t*}(\mathbf{x}, \mathbf{y})$, which implies that $\bar{Q}^{t*}(\mathbf{x}, \mathbf{y})$ increases faster with respect to x_i than x_{i+1}. Analogously we can show that $\bar{Q}^{t*}(\mathbf{x}, \mathbf{y})$ increases faster with respect to y_j than y_{j+1}, for $j = 1, \ldots, n-1$.

It remains to prove i) and ii) for the case with $\alpha = 0 > \beta$. However, the proof is identical, except that we use part (ii) of Lemma B.7 in the proof rather than part (i) of that lemma. □

Proof of Proposition 6. To prove that the 1-step-lookahead heuristic follows the top-down matching structure, It suffices to show that for $i < i'$ and $j < j'$, in any period t it does not match any type i' demand unless type i demand is fully matched, nor any type j' supply unless all type j supply is matched.

For the original state $(\mathbf{x}, \mathbf{y})$ and a matching decision $\mathbf{Q}^t = \{q_{ij}^t\}_{i=1,\ldots,m;\, j=1,\ldots,n}$ in period t, let $\mathbf{u} = (u_1, \ldots, u_m)$ and $\mathbf{v} = (v_1, \ldots, v_n)$ be the corresponding post-matching levels of demand and supply, respectively, $q^t_{i\cdot} = \sum_{j''=1}^m q^t_{ij''}$ be the quantity of type i demand fulfilled in period t, and $q^t_{\cdot j} = \sum_{i''=1}^m q^t_{i''j}$ be the quantity of type j supply fulfilled in period t, for $i = 1, \ldots, m$ and $j = 1, \ldots, n$.

Assume that $\mathbf{Q}^t$ depletes type i demand but fulfills a positive quantity of type i' in period t, i.e., $u_i > 0$ and $q^t_{i'\cdot} := \sum_{j''=1}^m q^t_{i'j''} > 0$. We will modify $\mathbf{Q}^t$ such that it will follow the top-down matching structure after the modification, and the expected discounted reward is (weakly) improved.

More specifically, let $\varepsilon := \min\{u_i, q^t_{i'\cdot}\}$. We modify $\mathbf{Q}^t$ to construct another decision $\hat{\mathbf{Q}}^t = \left\{\hat{q}^t_{ij}\right\}_{i=1,\ldots,m;\, j=1,\ldots,n}$ by reducing the fulfilled quantity of type i' demand by ε and the fulfilled quantity of type i demand by the same amount, i.e., $\hat{q}^t_{i\cdot} = q^t_{i\cdot} + \varepsilon$,$\hat{q}^t_{i'\cdot} = q^t_{i'\cdot} - \varepsilon$, and $\hat{q}^t_{i''j''} = q^t_{i''j''}$ for all $i'' \neq i$ and all j''. By doing so, the reward in period t increases by $(r^t_{id} - r^t_{i'd})\varepsilon$. In the meantime, the post-matching levels under the decision $\hat{\mathbf{Q}}^t$ become $\mathbf{u} - \varepsilon \mathbf{e}_i + \varepsilon \mathbf{e}_{i'}$ and $\mathbf{v}$. Thus, the expected total discounted reward under greedy matching from period $t+1$ to period T changes by

$$\gamma E V^g_{t+1}(\alpha \mathbf{u} + \mathbf{D}^{t+1} - \alpha\varepsilon \mathbf{e}_i + \alpha\varepsilon \mathbf{e}_{i'}, \beta \mathbf{v} + \mathbf{S}^{t+1}) - \gamma E V^g_{t+1}(\alpha \mathbf{u} + \mathbf{D}^{t+1}, \beta \mathbf{v} + \mathbf{S}^{t+1}).$$

In Online Supplement B, we define the transformed state $(\tilde{\mathbf{x}}, \tilde{\mathbf{y}})$ in a period as $\tilde{\mathbf{x}} := (x_1, x_1 + x_2, \ldots, \sum_{k=1}^i x_i, \ldots, \sum_{k=1}^m x_k)$ and $\tilde{\mathbf{y}} := (y_1, y_1 + y_2, \ldots, \sum_{k=1}^j y_k, \ldots, \sum_{k=1}^n y_k)$. We also define $\tilde{V}^g_t(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) := V^g_t(\mathbf{x}, \mathbf{y}) - \mathbf{x}(\mathbf{r}^t_d)^\intercal - \mathbf{y}(\mathbf{r}^t_s)^\intercal$ in Online Supplement B. It follows that

$$\begin{aligned}
&\gamma E V^g_{t+1}(\alpha \hat{\mathbf{u}} + \mathbf{D}^{t+1}, \beta \hat{\mathbf{v}} + \mathbf{S}^{t+1}) - \gamma E V^g_{t+1}(\alpha \mathbf{u} + \mathbf{D}^{t+1}, \beta \mathbf{v} + \mathbf{S}^{t+1}) \\
=&\gamma E V^g_{t+1}(\alpha \mathbf{u} + \mathbf{D}^{t+1} - \alpha\varepsilon \mathbf{e}_i + \alpha\varepsilon \mathbf{e}_{i'}, \beta \mathbf{v} + \mathbf{S}^{t+1}) - \gamma E V^g_{t+1}(\alpha \mathbf{u} + \mathbf{D}^{t+1}, \beta \mathbf{v} + \mathbf{S}^{t+1}) \\
=&-\gamma\alpha(r^{t+1}_{id} - r^{t+1}_{i'd})\varepsilon + E\tilde{V}^g_{t+1}(\alpha \tilde{\mathbf{u}} + \tilde{\mathbf{D}}^{t+1} - \alpha\varepsilon \mathbf{1}^{m,[i,i'-1]}, \beta \tilde{\mathbf{v}} + \tilde{\mathbf{S}}^{t+1}) - \gamma E\tilde{V}^g_{t+1}(\alpha \tilde{\mathbf{u}} + \tilde{\mathbf{D}}^{t+1}, \beta \tilde{\mathbf{v}} + \tilde{\mathbf{S}}^{t+1}) \\
\geq&-\gamma\alpha(r^{t+1}_{id} - r^{t+1}_{i'd})\varepsilon,
\end{aligned}$$

where $\tilde{\mathbf{u}} := (u_1, u_1 + u_2, \ldots, \sum_{k=1}^i u_k, \ldots, \sum_{k=1}^m u_k)$ and $\tilde{\mathbf{v}} := (v_1, v_1 + v_2, \ldots, \sum_{k=1}^j v_k, \ldots, \sum_{k=1}^n v_k)$ are the transformed post-matching levels, and the inequality follows from the monotonicity of $\tilde{V}^g_{t+1}$ (Lemma B.9).

Thus, by modifying $\mathbf{Q}^t$, the change in the expected total discounted reward from period t to period T is at least $[(r^t_{id} - r^t_{i'd}) - \gamma\alpha(r^{t+1}_{id} - r^{t+1}_{i'd})]\varepsilon \geq 0$. Thus, the modification weakly improves the expected total discounted reward (assuming greedy matching from period $t+1$ onwards).

On the other hand, the modification either reduces the post-matching level of demand type i to zero, or reduces the total fulfilled quantity of type i' demand to zero. (This reduces the violation of the top-down matching structure.)

If the matching decision in period t depletes type j supply but fulfills a positive quantity of type j' supply for some $j < j'$, we can implement a similar modification that weakly improves the expected total discounted reward and reduces violation of the top-down structure.

If the resulting matching decision does not follow the top-down structure yet, we will keep implementing similar modifications, and will eventually reach a decision that follows the top-down structure in period t. To show that the 1-step-lookahead heuristic performs weakly better than the greedy matching policy, we let $P^{\text{OSA}[1,t],\text{Greedy}[t+1,T]}$ be the policy that applies the 1-step-lookahead policy up to period t, and uses greedy matching from period $t+1$ to period T.

The two policies, $P^{\text{OSA}[1,t],\text{Greedy}[t+1,T]}$ and $P^{\text{OSA}[1,t-1],\text{Greedy}[t,T]}$ coincide with each other in periods $1,\ldots,t-1$, and therefore have the same expected rewards in those periods.

For any state in the beginning of period t, the policy $P^{\text{OSA}[1,t],\text{Greedy}[t+1,T]}$ uses the 1-step-lookahead policy in that period, which is optimal (for maximizing the total expected reward from period t to period T) given that $P^{\text{OSA}[1,t],\text{Greedy}[t+1,T]}$ will use greedy matching from the next period on. In contrast, the policy $P^{\text{OSA}[1,t-1],\text{Greedy}[t,T]}$ uses greedy matching in period t, which is suboptimal in response to the greedy matching it enforces from period $t+1$ to period T. Consequently, $P^{\text{OSA}[1,t],\text{Greedy}[t+1,T]}$ leads to a higher total expected reward from period t to period T than $P^{\text{OSA}[1,t-1],\text{Greedy}[t,T]}$. The overall total expected matching reward from period 1 to period T is higher under $P^{\text{OSA}[1,t],\text{Greedy}[t+1,T]}$ than under $P^{\text{OSA}[1,t-1],\text{Greedy}[t,T]}$.

The 1-step-lookahead policy coincides with $P^{\text{OSA}[1,T-1],\text{Greedy}[T,T]}$, and the greedy matching policy coincides with $P^{\text{OSA}[1,0],\text{Greedy}[1,T]}$. Thus, the former leads to a higher total expected reward.

For a 2-period problem, it is easy to see that greedy matching (in the descending order of unit rewards) is optimal for period 2. Thus, in period 1, the 1-step-lookahead policy is optimal. □

Proof of Proposition 7. Since the 1-step-lookahead heuristic follows the top-down matching structure, the corresponding matching decision in period t can be determined by the total matching quantity Q. In Online Supplement B.1, given the transformed state $(\tilde{\mathbf{x}}, \tilde{\mathbf{y}})$ in period t, we define $G_t^g(Q, \tilde{\mathbf{x}}, \tilde{\mathbf{y}})$ as the expected total discounted reward from period t to period T, if we follow the top-down matching up to the total quantity Q in period t, and enforce greedy matching thereafter (see (B.17)). We also define $\tilde{G}_t^g(Q, \tilde{\mathbf{x}}, \tilde{\mathbf{y}}) := -\tilde{\mathbf{x}}\mathbf{U}_m^{-1}(\mathbf{r}_d^t)^{\intercal} - \tilde{\mathbf{y}}\mathbf{U}_n^{-1}(\mathbf{r}_s^t)^{\intercal} + G_t^g(Q, \tilde{\mathbf{x}}, \tilde{\mathbf{y}})$ in Online Supplement B.1. To determine the total matching quantity in period t for the 1-step-lookahead heuristic, we solve $\max_{0 \le Q \le \min\{\tilde{x}_m, \tilde{y}_n\}} \tilde{G}_t^g(Q, \tilde{\mathbf{x}}, \tilde{\mathbf{y}})$.

Following the top-down matching structure, type i demand will be matched with type j supply only after all demand of types $1,\ldots,i-1$ and all supply of types $1,\ldots,j-1$ are completely fulfilled in period t. Thus, matching between type i demand and type j supply is possible only if $\tilde{x}_{i-1} < \tilde{y}_j$

and $\tilde{x}_i > \tilde{y}_{j-1}$. (If $\tilde{x}_{i-1} \geq \tilde{y}_j$, type j supply is already depleted before we start to fulfill type i demand; if $\tilde{x}_i \leq \tilde{y}_{j-1}$, type i demand is already depleted before we start to fulfill type j supply.)

Suppose that the 1-step-lookahead heuristic does match (i,j) in period t. Then, the matching of (i,j) starts when the total matching quantity Q reaches $\max\{\tilde{x}_{i-1}, \tilde{y}_{j-1}\}$ (this is when all types $1,\ldots,i-1$ demand and all types $1,\ldots,j-1$ supply are both completely fulfilled), and stops when it reaches $\min\{\tilde{x}_i, \tilde{y}_j\}$ (this is when either types i demand or type j supply is completely exhausted). To determine the matching quantity between i and j for the 1-step-lookahead heuristic in period t, we solve the following problem:

$$\begin{aligned} \max_{Q} \quad & \tilde{G}_t^g(Q, \tilde{\mathbf{x}}, \tilde{\mathbf{y}}), \\ s.t. \quad & \max\{\tilde{x}_{i-1}, \tilde{y}_{j-1}\} \leq Q \leq \min\{\tilde{x}_i, \tilde{y}_j\}. \end{aligned} \tag{E.5}$$

If $\hat{Q}_{(i,j)}$ is an optimal solution to (E.5), the 1-step-lookahead heuristic should match (i,j) for the quantity $\hat{Q}_{(i,j)} - \max\{\tilde{x}_{i-1}, \tilde{y}_{j-1}\}$. Within the feasible range $\max\{\tilde{x}_{i-1}, \tilde{y}_{j-1}\} \leq Q \leq \min\{\tilde{x}_i, \tilde{y}_j\}$, we can rewrite the function $\tilde{G}_t^g(Q, \tilde{\mathbf{x}}, \tilde{\mathbf{y}})$ given in (B.18) as:

$$\begin{aligned} \tilde{G}_t^g(Q, \tilde{\mathbf{x}}, \tilde{\mathbf{y}}) = & \gamma E\mathbf{D}^{t+1}(\mathbf{r}_d^{t+1})^{\intercal} + \gamma E\mathbf{S}^{t+1}(\mathbf{r}_s^{t+1})^{\intercal} \\ & - (\tilde{\mathbf{x}}_{[i,m]} - Q\mathbf{1}^{m-i+1})(\mathbf{U}_m^{-1})^{[i,m]\times[1,m]}(\mathbf{r}_d^t - \gamma\alpha\mathbf{r}_d^{t+1})^{\intercal} \\ & - (\tilde{\mathbf{y}}_{[j,n]} - Q\mathbf{1}^{n-j+1})(\mathbf{V}_n^{-1})^{[j,n]\times[1,n]}(\mathbf{r}_s^t - \gamma\beta\mathbf{r}_s^{t+1})^{\intercal} \\ & + \gamma E\tilde{V}_{t+1}^g(\tilde{\mathbf{D}}_{[1,i-1]}^{t+1}, \alpha(\tilde{\mathbf{x}}_{[i,m]} - Q\mathbf{1}^{m-i+1}) + \tilde{\mathbf{D}}_{[i,m]}^{t+1}, \tilde{\mathbf{S}}_{[1,j-1]}^{t+1}, \beta(\tilde{\mathbf{y}}_{[j,n]} - Q\mathbf{1}^{n-j+1}) + \tilde{\mathbf{S}}_{[j,n]}^{t+1}). \end{aligned} \tag{E.6}$$

To maximize $\tilde{G}_t^g(Q, \tilde{\mathbf{x}}, \tilde{\mathbf{y}})$ within the feasible range, we examine its derivative $\frac{\partial \tilde{G}_t^g(Q, \tilde{\mathbf{x}}, \tilde{\mathbf{y}})}{\partial Q}$ with respect to Q. (In the case where $\tilde{G}_t^g(Q, \tilde{\mathbf{x}}, \tilde{\mathbf{y}})$ is not differentiable with respect to Q, we can consider its maximum subgradient with respect to Q instead; in the case where states, decisions and demand/supply realizations take integer values only, we can consider the difference instead.)

For ease of notation, we write $\tilde{\mathbf{x}}^{t+1}(Q) := \left(\tilde{\mathbf{D}}_{[1,i-1]}^{t+1}, \alpha(\tilde{\mathbf{x}}_{[i,m]} - Q\mathbf{1}^{m-i+1}) + \tilde{\mathbf{D}}_{[i,m]}^{t+1}\right)$ and $\tilde{\mathbf{y}}^{t+1}(Q) := \left(\tilde{\mathbf{S}}_{[1,j-1]}^{t+1}, \beta(\tilde{\mathbf{y}}_{[j,n]} - Q\mathbf{1}^{n-j+1}) + \tilde{\mathbf{S}}_{[j,n]}^{t+1}\right)$. From (E.6), we have

$$\begin{aligned} & \frac{\partial \tilde{G}_t^g(Q, \tilde{\mathbf{x}}, \tilde{\mathbf{y}})}{\partial Q} \\ = & r_{id}^t - \gamma\alpha r_{id}^{t+1} + r_{js}^t - \gamma\alpha r_{js}^{t+1} + \gamma \lim_{\varepsilon\to 0} E\left[\frac{\tilde{V}_{t+1}^g(\tilde{\mathbf{x}}^{t+1}(Q+\varepsilon), \tilde{\mathbf{y}}^{t+1}(Q+\varepsilon)) - \tilde{V}_{t+1}^g(\tilde{\mathbf{x}}^{t+1}(Q), \tilde{\mathbf{y}}^{t+1}(Q))}{\varepsilon}\right] \\ = & r_{id}^t - \gamma\alpha r_{id}^{t+1} + r_{js}^t - \gamma\alpha r_{js}^{t+1} \\ & + \gamma \lim_{\varepsilon\to 0} E\left[\frac{\tilde{V}_{t+1}^g(\tilde{\mathbf{x}}^{t+1}(Q) - \alpha\varepsilon\mathbf{1}_{[i,m]}, \tilde{\mathbf{y}}^{t+1}(Q) - \beta\varepsilon\mathbf{1}_{[j,n]}) - \tilde{V}_{t+1}^g(\tilde{\mathbf{x}}^{t+1}(Q), \tilde{\mathbf{y}}^{t+1}(Q))}{\varepsilon}\right]. \end{aligned}$$

Under the condition $\alpha=\beta$, it follows from Lemma B.13 that the difference $\frac{1}{\varepsilon}\left[\tilde{V}^g_{t+1}(\tilde{\mathbf{x}}^{t+1}(Q)-\alpha\varepsilon\mathbf{1}_{[i,m]},\tilde{\mathbf{y}}^{t+1}(Q)-\beta\varepsilon\mathbf{1}_{[j,n]})-\tilde{V}^g_{t+1}(\tilde{\mathbf{x}}^{t+1}(Q),\tilde{\mathbf{y}}^{t+1}(Q))\right]$ depends only on the first $i-1$ entries of $\tilde{\mathbf{x}}^{t+1}(Q)$ (which is equal to $\tilde{\mathbf{D}}^{t+1}_{[1,i-1]}$), the first $j-1$ entries of $\tilde{\mathbf{y}}^{t+1}(Q)$ (which is equal to $\tilde{\mathbf{S}}^{t+1}_{[1,j-1]}$), the last entry of $\tilde{\mathbf{x}}^{t+1}(Q)$ (which is equal to $\alpha(\tilde{x}_m-Q)+\tilde{D}^{t+1}_m$), and the last entry of $\tilde{\mathbf{y}}^{t+1}(Q)$ (which is equal to $\beta(\tilde{y}_n-Q)+\tilde{S}^{t+1}_n$). Let us define $\tilde{\mathbf{x}}^{t+1,a}(Q):=\left(\tilde{\mathbf{D}}^{t+1}_{[1,i-1]},\alpha(\tilde{x}_m-Q)\mathbf{1}^{m-i+1}+\tilde{\mathbf{D}}^{t+1}_{[i,m]}\right)$ and $\tilde{\mathbf{y}}^{t+1,a}(Q):=\left(\tilde{\mathbf{S}}^{t+1}_{[1,j-1]},\alpha(\tilde{y}_n-Q)\mathbf{1}_{[j,n]}+\tilde{\mathbf{S}}^{t+1}_{[j,n]}\right)$. Since the first $j-1$ entries (resp., first $i-1$ entries) and the last entry are equal for $\tilde{\mathbf{x}}^{t+1}(Q)$ and $\tilde{\mathbf{x}}^{t+1,a}(Q)$ (resp., $\tilde{\mathbf{y}}^{t+1}(Q)$ and $\tilde{\mathbf{y}}^{t+1,a}(Q)$), we have

$$
\begin{aligned}
&\frac{\tilde{V}^g_{t+1}(\tilde{\mathbf{x}}^{t+1}(Q)-\alpha\varepsilon\mathbf{1}_{[i,m]},\tilde{\mathbf{y}}^{t+1}(Q)-\beta\varepsilon\mathbf{1}_{[j,n]})-\tilde{V}^g_{t+1}(\tilde{\mathbf{x}}^{t+1}(Q),\tilde{\mathbf{y}}^{t+1}(Q))}{\varepsilon}\\
=&\frac{\tilde{V}^g_{t+1}(\tilde{\mathbf{x}}^{t+1,a}(Q)-\alpha\varepsilon\mathbf{1}_{[i,m]},\tilde{\mathbf{y}}^{t+1,a}(Q)-\beta\varepsilon\mathbf{1}_{[j,n]})-\tilde{V}^g_{t+1}(\tilde{\mathbf{x}}^{t+1,a}(Q),\tilde{\mathbf{y}}^{t+1,a}(Q))}{\varepsilon}.
\end{aligned}
$$

Let us define

$$
\begin{aligned}
\tilde{G}^{g,a}_{ij,t}(Q,\tilde{x}_m,\tilde{y}_n)=&\gamma E\mathbf{D}^{t+1}(\mathbf{r}^{t+1}_d)^{\intercal}+\gamma E\mathbf{S}^{t+1}(\mathbf{r}^{t+1}_s)^{\intercal}\\
&-(\tilde{x}_m-Q)(r^t_{id}-\gamma\alpha r^{t+1}_{id})-(\tilde{y}_n-Q)(r^t_{js}-\gamma\alpha r^{t+1}_{js})+\gamma E\tilde{V}^g_{t+1}(\tilde{\mathbf{x}}^{t+1,a}(Q),\tilde{\mathbf{y}}^{t+1,a}(Q))\\
=&\gamma E\mathbf{D}^{t+1}(\mathbf{r}^{t+1}_d)^{\intercal}+\gamma E\mathbf{S}^{t+1}(\mathbf{r}^{t+1}_s)^{\intercal}-(\tilde{x}_m-Q)(r^t_{id}-\gamma\alpha r^{t+1}_{id})-(\tilde{y}_n-Q)(r^t_{js}-\gamma\alpha r^{t+1}_{js})\\
&+\gamma E\tilde{V}^g_{t+1}(\tilde{\mathbf{D}}^{t+1}_{[1,i-1]},\alpha(\tilde{x}_m-Q)\mathbf{1}^{m-i+1}+\tilde{\mathbf{D}}^{t+1}_{[i,m]},\tilde{\mathbf{S}}^{t+1}_{[1,j-1]},\beta(\tilde{y}_n-Q)\mathbf{1}^{n-j+1}+\tilde{\mathbf{S}}^{t+1}_{[j,n]}).
\end{aligned}
$$

We can readily verify that $\tilde{G}^{g,a}_{ij,t}(Q,\tilde{x}_m,\tilde{y}_n)=\tilde{G}^g_t(Q,\tilde{\mathbf{x}}',\tilde{\mathbf{y}}')$, where $\tilde{\mathbf{x}}'=(\mathbf{0}^{i-1},\tilde{x}_m\mathbf{1}^{m-i+1})$ and $\tilde{\mathbf{y}}'=(\mathbf{0}^{j-1},\tilde{y}_n\mathbf{1}^{n-j+1})$. Lemma B.11 in Online Supplement B shows that $\tilde{G}^g_t(Q,\tilde{\mathbf{x}},\tilde{\mathbf{y}})$ is concave in Q for any given $(\tilde{\mathbf{x}},\tilde{\mathbf{y}})$, under the condition $\alpha=\beta$. It follows that $\tilde{G}^{g,a}_{ij,t}(Q,\tilde{x}_m,\tilde{y}_n)$ is concave in Q.

Based on the above analysis, we see that for $\max\{\tilde{x}_{i-1},\tilde{y}_{j-1}\}\le Q\le\min\{\tilde{x}_i,\tilde{y}_j\}$, $\frac{\partial\tilde{G}^{g,a}_{ij,t}(Q,\tilde{x}_m,\tilde{y}_n)}{\partial Q}=\frac{\partial\tilde{G}^g_t(Q,\tilde{\mathbf{x}},\tilde{\mathbf{y}})}{\partial Q}$. This implies that to maximize $\tilde{G}^g_t(Q,\tilde{\mathbf{x}},\tilde{\mathbf{y}})$ for $Q\in\max\{\tilde{x}_{i-1},\tilde{y}_{j-1}\}\le Q\le\min\{\tilde{x}_i,\tilde{y}_j\}$, it is equivalent to maximize $\tilde{G}^{g,a}_{ij,t}(Q,\tilde{x}_m,\tilde{y}_n)$ with respect to Q within the same range.

Let us denote $p:=\tilde{y}_n-Q$. Then, $\tilde{x}_m-Q=\tilde{x}_m-\tilde{y}_n+p=IB+p$, where $IB:=\tilde{x}_m-\tilde{y}_n$. We can rewrite $\tilde{G}^{g,a}_{ij,t}(Q,\tilde{x}_m,\tilde{y}_n)$ as a function of p and IB, which we denote by $\tilde{G}^{g,b}_{ij,t}(p,IB)$, as follows.

$$
\begin{aligned}
&\tilde{G}^{g,a}_{ij,t}(Q,\tilde{x}_m,\tilde{y}_n)\\
=&\tilde{G}^{g,b}_{ij,t}(p,IB):=\gamma E\mathbf{D}^{t+1}(\mathbf{r}^{t+1}_d)^{\intercal}+\gamma E\mathbf{S}^{t+1}(\mathbf{r}^{t+1}_s)^{\intercal}-(IB+p)(r^t_{id}-\gamma\alpha r^{t+1}_{id})-p(r^t_{js}-\gamma\alpha r^{t+1}_{js})\\
&+\gamma E\tilde{V}^g_{t+1}(\tilde{\mathbf{D}}^{t+1}_{[1,i-1]},\alpha(IB+p)\mathbf{1}^{m-i+1}+\tilde{\mathbf{D}}^{t+1}_{[i,m]},\tilde{\mathbf{S}}^{t+1}_{[1,j-1]},\beta p\mathbf{1}^{n-j+1}+\tilde{\mathbf{S}}^{t+1}_{[j,n]}). \qquad \text{(E.7)}
\end{aligned}
$$

The variable p represents the post-matching available quantity of supply (of all types) in period

t. The feasible range $\max\{\tilde{x}_{i-1},\tilde{y}_{j-1}\}\le Q\le\min\{\tilde{x}_i,\tilde{y}_j\}$ of Q translates to the feasible range $\tilde{y}_n-\min\{\tilde{x}_i,\tilde{y}_j\}\le p\le\tilde{y}_n-\max\{\tilde{x}_{i-1},\tilde{y}_{j-1}\}$ of p. To simplify the notation, we denote

$$\tilde{v}_{n,L}:=\tilde{y}_n-\min\{\tilde{x}_i,\tilde{y}_j\},\quad \tilde{v}_{n,U}:=\tilde{y}_n-\max\{\tilde{x}_{i-1},\tilde{y}_{j-1}\}.$$

The feasible range of p becomes $\tilde{v}_{n,L}\le p\le\tilde{v}_{n,U}$. We note that $\tilde{v}_{n,U}$ represents the available supply quantity (of all types) when the 1-step-lookahead heuristic starts to match (i,j), and $\tilde{v}_{n,L}$ represents the available supply quantity (of all types) when i and j are matched to the maximum extent (which may or may not happen under the 1-step-lookahead heuristic). We can readily verify that $\tilde{v}_{n,U}$ and $\tilde{v}_{n,L}$ are the same as the quantities $\tilde{v}^{ij}_{n,U}$ and $\tilde{v}^{ij}_{n,L}$ defined in the paper, respectively, if $\tilde{x}_{i-1}<\tilde{y}_j$ and $\tilde{x}_i>\tilde{y}_{j-1}$ (See the discussions following Proposition 6, and recall that matching between type i demand and type j supply is possible only if $\tilde{x}_{i-1}<\tilde{y}_j$ and $\tilde{x}_i>\tilde{y}_{j-1}$).

Thus, to solve $\max_{\max\left\{\tilde{x}_{i-1},\tilde{y}_{j-1}\right\}\le Q\le\min\left\{\tilde{x}_i,\tilde{y}_j\right\}}\tilde{G}^{g,a}_{ij,t}(Q,\tilde{x}_m,\tilde{y}_n)$, it is equivalent to solve:

$$\max_{\tilde{v}_{n,L}\le p\le\tilde{v}_{n,U}}\tilde{G}^{g,b}_{ij,t}(p,IB).\tag{E.8}$$

Given the concavity of $\tilde{G}^{g,a}_{ij,t}(Q,\tilde{x}_m,\tilde{y}_n)$ with respect to Q, $\tilde{G}^{g,b}_{ij,t}(p,IB)$ is concave in p for $p\ge IB^-:=\max\{0,-IB\}$. For a given value of IB, we define

$$p^t_{s_{ij}}(IB):=\inf\arg\max_{p\ge IB^-}\tilde{G}^{g,b}_{ij,t}(p,IB).\tag{E.9}$$

By the definition of $\tilde{V}^g_{t+1}$ (i.e., $\tilde{V}^g_{t+1}(\tilde{\mathbf{x}},\tilde{\mathbf{y}}):=V^g_{t+1}(\mathbf{x},\mathbf{y})-\tilde{\mathbf{x}}\mathbf{U}_m^{-1}(\mathbf{r}_d^{t+1})^\intercal-\tilde{\mathbf{y}}\mathbf{U}_n^{-1}(\mathbf{r}_s^{t+1})^\intercal$; see Online Supplement B.1), we can readily verify that the above equation (E.9) is equivalent to the equation (4) in the paper. By the concavity of $\tilde{G}^{g,b}_{ij,t}(p,IB)$, the optimal solution to (E.8) is given by $\min\left\{\tilde{v}_{n,U},\ p^t_{s_{ij}}(IB)\vee\tilde{v}_{n,L}\right\}$ (where the bivariate operator $\vee$ means to take the maximum of the two inputs). When we start to match (i,j) the available supply quantity (of all types) $\tilde{v}_{n,U}$ is already below $p^t_{s_{ij}}(IB)$, the optimal solution to (E.8) is equal to $\tilde{v}_{n,U}$. In that case, we will not match (i,j) (nor any pair (i',j') such that $i'\le i$ and $j'\le j$, according to the top-down structure), so that the available supply is not further reduced. If the available supply quantity (of all types) $\tilde{v}_{n,U}$ is above $p^t_{s_{ij}}(IB)$, the 1-step-lookahead heuristic matches (i,j) until the available quantity of supply is reduced to either $p^t_{s_{ij}}(IB)$ or $\tilde{v}_{n,L}$, whichever happens first. In the latter situation (i.e., the total supply reduces to $\tilde{v}_{n,L}$ first), we would have matched (i,j) to the maximum extent, whereas the total available supply is still above $p^t_{s_{ij}}(IB)$. Therefore, we either matches i with j to reduce the total supply to the target level $p^t_{s_{ij}}(IB)$ or as close to it as possible. This is equivalent to the matching quantity between i and j given in the proposition. $\square$

References

Simchi-Levi, D, X Chen, J Bramel. 2014. *The Logic of Logistics.* 3rd ed. Springer.

Chen, X, Z Pang, L Pan. 2014. Coordinating inventory control and pricing strategies for perishable products. *Oper. Res.* 62(2) 284300.