Supplementary Appendix

B Additional Proofs

Proof of Proposition 2. Constructing an instance for $S = 1$ is trivial, so we focus on the case where $S \geq 2$. Consider the following instance. Let the set of prices be $\mathcal{D} = \{1, 3, 200\}$ and assume that $\gamma_0 = 1, \gamma_{S-1} = 100, \gamma_S = 300$ and $\gamma_w = 0$ for all $w \in \{1, ..., S-2\}$. Assume also that impatient customers ($w = 0$) buy at all three price points, semi-patient customers ($w = S-1$) buy at the prices $p = 1$ and $p = 3$ and that patient customers ($w = S$) can be divided into two subgroups: 1/9 of them buy at both prices $p = 1$ and $p = 3$, while the other 8/9 of them only buy at the lowest price, $p = 1$.

We now show that the optimal policy that the shortest optimal policy for this family of instances is $2S$ periods long for any value of $S \geq 2$. We first consider policies that involve a single price, then move on to policies that involve two prices and, finally, we look at policies that involve all three prices. Among the three available static pricing policies, the optimal one is to offer a price of $p = 3$, which produces average revenue of 403. From Lemma 2, we know that we only need to use the lowest price in a given cyclic policy once. Therefore, among policies with two prices, it's sufficient to consider policies where the higher price is used for $T - 1$ periods and the lower price is only used in period T. We now consider all three possible pairs of prices. The price pair 3 and 200 yields revenue $600 - \frac{197}{T}$ for $T \leq S$ and $600 - \frac{497}{S+1}$ for $T = S + 1$. The price pair 1 and 200 yields revenue $600 - \frac{199}{T}$ for $T \leq S$ and $600 - \frac{299}{S+1}$ for $T = S + 1$ and the price pair 1 and 3 yields lower revenues. The optimal two-price policy thus yields an average revenue of $600 - \frac{197}{S}$.

We now consider policies that involve 3 different prices. By Theorem 1, it's sufficient to search among policies that are at most 2S periods long. For any $T \in \{3, ..., 2S\}$, we can assign the lowest price, $p = 1$, exclusively to period T by Lemma 2. For a given T, the decision that we need to make is to what subset of $\{1, 2, ..., T-1\}$ to assign the price $p = 3$, with the rest of the periods having price $p = 200$. Note that for $w = 0$, the effective price $e_{t,0}$ is higher with $p = 200$ than with $p = 3$ and for $w = S-1$ and $w = S$, the effective price $e_{t,w}$ is higher with $p = 3$ than with $p = 200$. This implies that any 3 price policy with length $T \leq S$ and $p_T = 1$ can be replaced by a better 2 price policy since any period t with price $p_t = 3$ adds a period with effective price $e_{t,0} = 3$ but not does add any effective price $e_{t,w} = 3$ for $w = S-1$ or $w = S$. For any $T \in \{S+1, ..., 2S\}$, the best case scenario is to have a single period t with $e_{t,0} = 3$ (there must be one period where the price $p = 3$ is offered) and all t' with $e_{t',w} \in \{1, 3\}$ for $w = S-1$ and $w = S$. As Figure 9 shows, this can be accomplished by having the price $p = 3$ exactly at period $t = T - S$, giving rise to a pricing policy of the form $(200, ..., 200, 3, 200, ..., 200, 1)$. For such a policy, the average revenue of the firm is $600 - \frac{198}{T}$. By optimizing over $T \in \{S+1, ..., 2S\}$, we find that $T = 2S$ is optimal, yielding a revenue of $600 - \frac{96}{S}$, which is superior than the best revenue that can be obtained by a static or a two-price policy.

	t=1	t=2				t=T-S						t=T
w=0	200	200	200	200	200	3	200	200	200	200	200	1
					3	3					1	1
				3	3	3				1	1	1
			3	3	3	3			1	1	1	1
		3	3	3	3	3		1	1	1	1	1
w=S-1	3	3	3	3	3	3	1	1	1	1	1	1
w=S	3	3	3	3	3	1	1	1	1	1	1	1

Figure 9: The optimization of a 3 price policy involves choosing in which subset of periods $1, ..., T-1$ the firm will offer price $p = 3$ (price $p = 1$ must be offered exclusively at time T). The firm's revenue is decreasing in the number of times $p = 3$ shows up in the first row of the effective prices table and increasing in the number of times $p = 3$ shows up in the bottom two rows of the table. The optimal solution is clearly to let $p = 3$ only at $t = T - S$.

□

Proof of Theorem 3. For any $\varepsilon > 0$, we use an iterative procedure to construct two sets that we denote by $\mathcal{D}_\varepsilon$ and $\mathcal{D}_\varepsilon^X$. Initialize the sets so that $\mathcal{D}_\varepsilon = \emptyset$ and $\mathcal{D}_\varepsilon^X = \mathcal{D}$. Let x be the smallest element in $\mathcal{D}_\varepsilon^X$ (such an operation is valid since $\mathcal{D}$ is closed). Include x in the set $\mathcal{D}_\varepsilon$ and remove $[x, x + \frac{\varepsilon}{\Gamma L})$ from $\mathcal{D}_\varepsilon^X$. The remaining elements in $\mathcal{D}_\varepsilon^X$ still form a closed set so we can repeat the procedure above. Repeat this process until $\mathcal{D}_\varepsilon^X = \emptyset$. Once this process is over, $\mathcal{D}_\varepsilon$ is a subset of $\mathcal{D}$ with at most $\frac{\overline{V}\Gamma L}{\varepsilon} + 1$ elements such that any element in $\mathcal{D}$ is at most $\frac{\varepsilon}{\Gamma L}$ away from an element in $\mathcal{D}_\varepsilon$.

By Theorem 2, one can compute a policy $\mathbf{p}_\varepsilon$ that is optimal for the set of prices $\mathcal{D}_\varepsilon$ in time $\mathcal{O}(\Gamma\overline{V}LS^2/\varepsilon)$. Let $\mathbf{p}$ be the optimal policy when the set of available prices is $\mathcal{D}$, and let $\hat{\mathbf{p}}_\varepsilon$ be the rounding of the prices in $\mathbf{p}$ down to the closest element in the set $\mathcal{D}_\varepsilon$. By Lipschitz continuity, $R(\hat{\mathbf{p}}_\varepsilon) \geq R(\mathbf{p}) - \varepsilon$. The proof is complete since $\mathbf{p}_\varepsilon$ yields higher performance than $\hat{\mathbf{p}}_\varepsilon$, i.e., $R(\mathbf{p}_\varepsilon) \geq R(\hat{\mathbf{p}}_\varepsilon)$. □

Proof of Proposition 3. Let $\mathbf{p}$ be an optimal policy that is cyclic and whose cycle is of length $S < L \leq 2S$. Assume without loss of generality that the lowest price is offered last in the cycle. Let k be an index such that p_k is the second lowest price in the policy. If $k < T - S$, then k is a reset period and, by the Policy Decomposition Lemma, there exists a shorter optimal policy. If $k > S$, then consider the reflected policy $\mathbf{p}^r$. By the Reflection Lemma, this policy yields as much revenue as $\mathbf{p}$. The period $T+1-k$ is a reset period in the reflected policy, so by the Policy Decomposition Lemma, there exists a shorter optimal policy. □

Proof of Proposition 4. A corollary of the Reflection Lemma is that for any cyclic pricing policy with monotonically increasing prices, there exists a cyclic policy yielding the same revenues with monotonically decreasing prices. Without loss of generality, let us prove the result for non-decreasing policies. Consider an arbitrary cyclic non-decreasing policy with length $T \leq 2S$. If $T \leq S+1$, there is nothing to prove. Suppose $T > S+1$. Then the system necessarily resets at periods 1 and $T-S-1$ since $p_1 = \min\{p_1, ..., p_{S+1}\}$, and $p_{T-S-1} = \min\{p_{T-S-1}, ..., p_T\}$. By the Policy Decomposition Lemma, there exists a cyclic policy with length less or equal than $\max\{T-S-1, S+1\} \leq S+1$ that dominates the original policy. □

Proof of Proposition 5. Consider an arbitrary cyclic monotonic policy with length $T > j$. Without loss of generality, by the Reflection Lemma and Proposition 4, one may assume that the policy is non-increasing over the cycle and that $T \leq S+1$.

We next establish that one may construct a policy with weakly higher revenues such that $p_1 = v_0$, ..., $p_j = v_{j-1}$.

Suppose that $p_1 > v_0$. If $p_T \geq v_0$, then one may increase revenues by setting $p_1 = p_2 = ... = p_T = v_0$. Otherwise, let $k = \min\{i \in \{1, ..., S+1\} : p_i < v_0\}$. In such a case, one may again increase revenues by setting $p_1 = p_2 = ... = p_k = v_0$. Suppose now $p_1 \leq v_0$. Note that since the policy is non-increasing, one may assume that only impatient customers (with $w = 0$) purchase in period 1 when computing revenues. Hence, one may increase revenues by setting $p_1 = v_0$. We conclude that one may always weakly increase revenues by setting $p_1 = v_0$ while maintaining the non-increasing structure.

Assuming that $p_1 = v_0$ and the policy is non-increasing, and using the fact that $\overline{R}(v_0) \leq \overline{R}(v_1) \leq ... \leq \overline{R}(v_{j-1})$, one may show in a recursive fashion that one may weakly increase revenues by setting $p_2 = v_1$,, ..., $p_j = v_{j-1}$.

We now assume without loss of generality that $p_i = v_{i-1}$, $i = 1, ..., j$.

Case 1: $p_{j+1} \leq v_j$. Then increasing p_{j+1} to v_{j-1} does not alter the non-increasing structure of the policy and yields an additional $\overline{R}(v_{j-1}) - \overline{R}(v_j) > 0$ per cycle and hence strictly increases revenues and the initial policy was suboptimal.

Case 2: $p_{j+1} > v_j$. Since the policy is non-increasing, one has that $p_{j+1} \leq p_j = v_{j-1}$. If $p_{j+1} < v_{j-1}$, then one may strictly increase revenues by increasing it to v_{j-1}. Suppose that $p_{j+1} = v_{j-1}$. In such a case, consider the policy $q_1, ..., q_T$ that coincides with p with the exception of period $j+1$. In particular, we set $q_{j+1} = v_0$.

$$R(\mathbf{q}) - R(\mathbf{p}) = \frac{1}{T}\gamma_0(v_0 - v_{j-1}) > 0.$$

We deduce that p was necessarily suboptimal.

We conclude that any optimal cyclic policy with length $T > j$ is non-monotone. $\square$

Proof of Proposition 6. We establish the result by constructing a sequence of problem instances indexed by n such that cyclic monotone policies perform arbitrarily poorly as n increases to ∞.

Preliminaries: Class of instances under consideration. Let k be a positive integer, and let

$$w_{i,n} \quad = \quad 2^{ni} - 1, \qquad i = 1, ..., k$$

Fixing k, we consider the following family of instances indexed by n. The price set is $[0, \overline{V}]$. The maximum willingness to wait is given by $S_n = w_{k,n}$. Valuations are deterministic denoted by v_w. Furthermore, for all $w \in \{0, 1, ..., w_{k,n}\} \setminus \{w_{i,n} : i = 1, ..., k\}$, $v_w = 0$ and $\gamma_w = 0$, and for $i = 1, ..., k$,

$$\begin{aligned} v_{w_{i,n}} &= 2^{-in} \\ \gamma_{w_{i,n}} &= 2^{in}. \end{aligned}$$

Next, we first lower bound the optimal performance, and then we upper bound the performance of any cyclic monotone policy.

Step 1: Lower bound on optimal performance. We construct a particular cyclic policy $\mathbf{p}$, with cycle length $S_n + 1$, and lower bound its performance. Let

$$\begin{aligned} \mathcal{T}_{i,n} &= \{j 2^{in} : j = 1, ..., 2^{(k-i)n}\}, \qquad i = 1, ..., k, \\ \mathcal{T}_{k+1,n} &= \emptyset. \end{aligned}$$

Consider the policy that applies the following prices

$$\begin{aligned} p_t &= v_{w_{i,n}} \text{ for } t \in \mathcal{T}_{i,n} \setminus \cup_{j=i+1}^{k+1} \mathcal{T}_{j,n}, \qquad i = 1, ..., k \\ p_t &= v_{w_{1,n}} + 1, \text{ otherwise.} \end{aligned}$$

In particular, no customer would ever purchase at time periods that do not belong to $\cup_{i=1}^{k} \mathcal{T}_{i,n}$.

Consider the revenues generated by segment k, the most patient customers. All customers with patience level $w_{k,n}$ purchase at time $w_{k,n} + 1$ and the revenues generated are given by:

$$(w_{k,n} + 1)\gamma_{w_{k,n}} v_{w_{k,n}} = (w_{k,n} + 1) = 2^{kn}.$$

Consider now the revenues generated by segment $k - 1$. All customers purchase at times in $\mathcal{T}_{k-1,n}$ except for those customers who could purchase at times in $\mathcal{T}_{k,n} = \{w_{k,n} + 1\}$. Hence, the revenues generated from segment $k-1$ are lower bounded by

$$[w_{k,n}+1-(w_{k-1,n}+1)]\gamma_{w_{k-1,n}}v_{w_{k-1,n}} \;=\; (w_{k,n}+1)\left[1-\frac{w_{k-1,n}+1}{w_{k,n}+1}\right] \;=\; 2^{kn}(1-2^{-n}).$$

Repeating a similar argument for an arbitrary segment i, all segment i customers purchase at times in $\mathcal{T}_{i,n}$ except for those customers who could purchase at times in $\cup_{j=i+1}^{k+1}\mathcal{T}_{j,n}$. Hence, the revenues generated from segment i are lower bounded by

$$[w_{k,n}+1-(w_{i,n}+1)|\cup_{j=i+1}^{k+1}\mathcal{T}_{j,n}|]\gamma_{w_{i,n}}v_{w_{i,n}} \;=\; w_{k,n}+1-(w_{i,n}+1)2^{(k-i-1)n} \;=\; 2^{kn}(1-2^{-n})$$

Adding up the revenues generated by the k segments, one has that the total revenues generated by the proposed cyclic policy are lower bounded by:

$$R(\mathbf{p}) \geq k(1-2^{-n}). \tag{16}$$

Step 2: Upper bound on the performance of the best policy in $\mathcal{M}$. Without loss of generality, one may restrict attention to policies with cycle length at most $S+1$; furthermore, by the reflection lemma (Lemma 3), one may further restrict attention to non-increasing policies. Consider an arbitrary cyclic monotone non-increasing policy $\mathbf{p}$ in $\mathcal{M}$ with cycle length $T \leq S_n+1$. For $j = 0,...,k'$, let $m_{t,n} = \max\{\ell : v_{w_{\ell,n}} > e_{t,w_{i,n}}(\mathbf{p})\}$. Given that the policy is non-increasing over a cycle, one may assume for revenue computations that all customers arriving at t purchase at time $\min\{t+w,T\}$. The average revenues over a cycle may be written as

$$R(\mathbf{p}) \;=\; \frac{1}{T}\left[\sum_{t=1}^{T-1}\sum_{w=0}^{t-1}\gamma_w p_t \mathbf{1}\{v_w > p_t\} + \sum_{w=0}^{w_{k,n}}\min\{w+1,T\}p_T\mathbf{1}\{v_w > p_T\}\right]$$

For $t \leq T-1$,

$$\sum_{w=0}^{t-1}\gamma_w p_t\mathbf{1}\{v_w > p_t\} \;=\; \sum_{i=1}^{k-1}\gamma_{w_{i,n}}p_t\mathbf{1}\{v_{w_{i,n}} > p_t\} \;\leq\; \sum_{i=1}^{k-1}\gamma_{w_{i,n}}v_{w_{m_{t,n},n}}\mathbf{1}\{v_{w_{i,n}} > p_t\} \;\leq\; 1+(k-2)^+2^{-n}.$$

In addition,

$$\begin{aligned}\sum_{w=0}^{w_{k,n}}\min\{w+1,T\}\gamma_w p_T\mathbf{1}\{v_w > p_T\} \;&\leq\; \sum_{i=1}^{k}\min\{w_{i,n}+1,T\}\gamma_{w_{i,n}}v_{w_{m_{T,n},n}}\mathbf{1}\{v_{w_{i,n}} > p_T\} \\ &\leq\; T(1+(k-1)2^{-n}).\end{aligned}$$

We deduce that

$$\sup_{\mathbf{p}\in\mathcal{M}} R(\mathbf{p}) \;\leq\; 2(1+(k-1)2^{-n}). \tag{17}$$

Combining (16) and (17), one obtains

$$\frac{\sup_{\mathbf{p}\in\mathcal{M}} R(\mathbf{p})}{\sup_{\mathbf{p}\in\mathcal{P}} R(\mathbf{p})} \;\leq\; \frac{2+(k-1)2^{-n}}{k(1-2^{-n})}.$$

Since this bound is valid for arbitrary values of n, and for arbitrary values of k, (9) follows; the proof is complete. □

Proof of Proposition 7. To simplify notation, we assume without loss of generality that $p_{-C+1} = ... = p_0 = \bar{V}+1$. Without loss of generality, we may assume that units are consumed in a FIFO (First in First Out) fashion, as units are indistinguishable. Assuming such an order in consumption will enable one to track units in the system and in particular track the purchasing price of a unit consumed in a given period t.

For any policy $\mathbf{y}$ and $T>0$, we let $U_T^{\mathbf{y}} = \sum_{t=1}^{T} vx_t - p_t y_t$ denote the utility generated over the T first time periods.

Step 1: An upper bound on performance. Let A_t represent the set of periods such that, if the consumer would stop purchasing at that period, they would still have enough inventory left over to consume at period t, i.e., $A_t = \{\ell \leq t : I_\ell \geq t-\ell\}$. Then, under the FIFO rule for consumption and given the fact that consumption is never delayed, the price paid for a unit consumed in period t (if consumption takes place) is given by p_{j_t} where $j_t = \min\{\ell : \ell \in A_t\}$.

Note that for all $\ell \in A_t$, by the feasibility of the policy, $t-\ell \leq I_\ell \leq c$, and hence, $\ell \geq t-c$. This implies that if a unit is consumed in period t, i.e., $x_t>0$, then the price paid for that unit, p_{j_t}, is such that $p_{j_t} \geq \widetilde{e}_{t,c}(\mathbf{p})$.

For any $T>0$,

$$\frac{U_T^{\mathbf{y}}}{T} \;\leq\; \frac{1}{T}\sum_{t=1}^{T} x_t(v-p_{j_t}) \;\leq\; \frac{1}{T}\sum_{t=1}^{T} x_t(v-\widetilde{e}_{t,c}(\mathbf{p})) \;\leq\; \frac{1}{T}\sum_{t=1}^{T}(v-\widetilde{e}_{t,c}(\mathbf{p}))^+.$$

We deduce that for any feasible policy $\mathbf{y}$,

$$\liminf_{T\to\infty} \frac{U_T^{\mathbf{y}}}{T} \leq \liminf_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T}(v-\widetilde{e}_{t,c}(\mathbf{p}))^+. \tag{18}$$

Step 2: An optimal consumer policy. Consider a policy $\mathbf{y}$ where the consumer purchases a unit to consume at period l at the first period starting from $l-c$ when she observes a price that is equal to $\widetilde{e}_{l,c}(\mathbf{p})$, assuming this price is at most v.

We represent whether the consumer buys a unit at time t to consume at time l by $z_{l,t}$ and wether the consumer owns a unit at time t to consume at time l by $q_{l,t}$. Using this representation, the total purchases at time t equal

$$y_t = \sum_{l=t}^{t+c} z_{l,t}, \tag{19}$$

and whether consumption occurs at time t is $x_t = q_{t,t}$. The policy we are considering is given by the following algorithm:

Set $I_1 = 0$, $q_{\ell,s} = 0$, $z_{\ell,s} = 0$, for $\ell, s \in \mathbb{Z}$.

For all $t \geq 1$,

– **For all** $\ell \in \{t, t+1, ..., t+c\}$,

— **If** $p_t = \widetilde{e}_{\ell,c}(\mathbf{p})$, $v \geq p_t$ and $q_{\ell,t-1} = 0$, **Then** set $z_{\ell,t} = 1$.

— Set $q_{\ell,t} = q_{\ell,t-1} + z_{\ell,t}$.

We first show that the policy is feasible and then derive its performance. The first step is to establish that $I_t = \sum_{\ell=t+1}^{t+c} q_{\ell,t}$. We proceed by induction. For $t = 0$, the result is trivial. Suppose the result is true for $t-1$. One has

$$I_t = I_{t-1} + y_t - x_t = \sum_{\ell=t}^{t+c-1} q_{\ell,t-1} + \sum_{\ell=t}^{t+c} z_{\ell,t} - x_t = \sum_{\ell=t+1}^{t+c} q_{\ell,t} + (q_{t,t-1} + z_{t,t}) - x_t$$

where the second equality follows from the induction hypothesis and Eq. (19); and the third equality follows from the definition of $q_{\cdot,\cdot}$ and the fact that $q_{t+c,t-1} = 0$. Note that $(q_{t,t-1} + z_{t,t}) - x_t = q_{t,t} - x_t = 0$. One deduces that $I_t = \sum_{\ell=t+1}^{t+c} q_{\ell,t}$ and the induction step is complete.

The result above in particular implies that $I_t \geq 0$ for all $t \geq 1$ and that $I_t = \sum_{\ell=t+1}^{t+c} q_{\ell,t} \leq c$, and hence the policy is feasible.

We next now derive the performance of the policy. We establish by induction that

$$U_T^{\mathbf{y}} = \sum_{t=1}^{T} (v - \widetilde{e}_{t,c}(\mathbf{p}))^+ - \sum_{t=T+1}^{T+c} \sum_{s=t-c}^{T} z_{t,s} p_s.$$

Consider the base case. For $T = 1$, one has that

$$U_1^{\mathbf{y}} = x_1 v - p_1 y_1 = z_{1,1} v - p_1 z_{1,1} - p_1 \sum_{t=2}^{1+c} z_{t,1} p_1 = (v - p_1)^+ - \sum_{t=2}^{1+c} \sum_{s=t-c}^{1} z_{t,s} p_s,$$

where the last term has all elements equal to 0 for all $s \neq 1$. The base case is verified. Suppose the result is true for $T-1 \geq 1$. One has that

$$\begin{aligned}
U_T^{\mathbf{y}} &= U_{T-1}^{\mathbf{y}} + x_T v - p_T y_T \\
&\overset{(a)}{=} \sum_{t=1}^{T-1} (v - \widetilde{e}_{t,c}(\mathbf{p}))^+ - \sum_{t=T}^{T-1+c} \sum_{s=t-c}^{T-1} z_{t,s} p_s + x_T v - p_T \sum_{\ell=T+1}^{T+c} z_{\ell,T} - p_T z_{T,T} \\
&= \sum_{t=1}^{T-1} (v - \widetilde{e}_{t,c}(\mathbf{p}))^+ + x_T v - p_T z_{T,T} - \sum_{s=T-c}^{T-1} z_{t,s} p_s - \sum_{t=T+1}^{T-1+c} \sum_{s=t-c}^{T-1} z_{t,s} p_s - p_T \sum_{\ell=T+1}^{T+c} z_{\ell,T} \\
&= \sum_{t=1}^{T-1} (v - \widetilde{e}_{t,c}(\mathbf{p}))^+ + x_T v - \sum_{s=T-c}^{T} z_{t,s} p_s - \sum_{t=T+1}^{T+c} \sum_{s=t-c}^{T} z_{t,s} p_s \\
&\overset{(b)}{=} \sum_{t=1}^{T} (v - \widetilde{e}_{t,c}(\mathbf{p}))^+ - \sum_{t=T+1}^{T+c} \sum_{s=t-c}^{T} z_{t,s} p_s,
\end{aligned}$$

where (a) follows from the induction hypothesis and Eq. (19); and (b) follows from the fact that $x_T v - \sum_{s=T-c}^{T} z_{t,s} p_s = (v - \widetilde{e}_{t,c}(\mathbf{p}))^+$.

The average utility generated by the policy by time T is given by

$$\begin{aligned}
\frac{1}{T} \sum_{t=1}^{T} (v - \widetilde{e}_{t,c}(\mathbf{p}))^+ &= \frac{1}{T} \sum_{t=1}^{T} (v - \widetilde{e}_{t,c}(\mathbf{p}))^+ - \frac{1}{T} \sum_{t=T+1}^{T+c} \sum_{s=t-c}^{T} z_{t,s} p_s \\
&\geq \frac{1}{T} \sum_{t=1}^{T} (v - \widetilde{e}_{t,c}(\mathbf{p}))^+ - \frac{c \max\{p : p \in \mathcal{D}\}}{T}.
\end{aligned}$$

Hence the long-run average performance is given by

$$\liminf_{T \to \infty} \frac{U_T^{\mathbf{y}}}{T} = \liminf_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} (v - \widetilde{e}_{t,c}(\mathbf{p}))^+.$$

The latter, in conjunction with the upper bound on the performance on the performance of any policy (see Eq. (18)), establishes the optimality of the policy $\mathbf{y}$. □