Online Supplement for "Dealership or Marketplace"

EC.1. Remarks for Assumption 1 and Proposition 1

Assumption 1. The first assumption is common in the literature of dynamic pricing, e.g., Gallego and van Ryzin (1994). The second assumption guarantees the existence of a unique equilibrium price $p_e > 0$ such that $\lambda_b d_b(p_e) = \lambda_s d_s(p_e)$; in a competitive equilibrium without dealers, it is given by the intersection of demand and supply curves. The last technical assumption is that the suprema are attained in the definitions of $\mathcal{R}_b$ and $\mathcal{R}_s$. A sufficient condition is that $d_b(p)$ and $d_s(p)$ are continuous in $p \geq 0$ and $d_b(p) = o(1/p)$. Most classic demand (supply) functions satisfy this condition, e.g., linear demand and exponential demand. For any $z > 0$, define $P_b(z) \triangleq \arg\max_{p\geq 0} \mathcal{R}_b(p,z)$ and $P_s(z) \triangleq \arg\max_{p\geq 0} \mathcal{R}_s(p,z)$. Elementary comparative statics reveal that $P_b(z)$ and $P_s(z)$ are nonnegative and increasing in $z \geq 0$. Note that this standard assumption also guarantees the sufficiency and necessity of the HJB equation, similar to Gallego and van Ryzin (1994), Ghuloum et al. (2022).

Proposition 1. In our model, the bid-ask spread arises from the concavity of $V_{\mathcal{D}}(x)$ as it implies that $p_b^*(x) \geq \Delta V_{\mathcal{D}}(x) \geq \Delta V_{\mathcal{D}}(x+1) \geq p_s^*(x)$, i.e., the marginal value of inventory must be straddled by the spread. Note that the bid-ask spread does not necessarily straddle the equilibrium price p_e, at which demand and supply cross, i.e., $\lambda_b d_b(p_e) = \lambda_s d_s(p_e)$. The monotonicity of $p_s^*(x)$ in the dealer's inventory level is intuitive: As x increases, the marginal value of inventory decreases and the dealer is unwilling to buy more, thus the bid price decreases. Combining with $\lim_{x\to\infty} p_s^*(x) = 0$, it implies that if the lowest valuation of sellers is above 0, there are no sellers willing to sell to the dealer when the dealer's inventory is higher than a threshold. The monotonicity of $p_b^*(x)$ in the dealer's inventory level follows from Gallego and van Ryzin (1994): The ask price increases in the "scarcity" of inventory, as $\Delta V_{\mathcal{D}}(x)$ decreases in x. However, the bid-ask spread is *not* monotone in x in general. With unlimited inventory, the dealer does not need to replenish the inventory from the sellers and simply optimizes the revenue rate from the buyers. When the interest rate increases (or the market sizes decrease), the dealer lowers the bid and ask prices to avoid the high inventory risk, or maintain the turnover rate. The mean reversion of the inventory is consistent with the findings in Amihud and Mendelson (1980), Li et al. (2019). This phenomenon occurs due to the inventory risk: The dealer is unwilling to keep a very low level of inventory, as potential buyers might be turned away if $x = 0$ due to stochastic fluctuation; on the other hand, he is unwilling to stock too much, as the inventory risk is high because of the discounting. As the market size increases (λ_b and λ_s increase), the dealer needs more inventory as a "buffer" to counter the increasing stochastic fluctuation. This explains the monotonicity of x_p. It implies that for a thicker market, the dealer is willing to keep more inventory to hedge the inventory risk.

EC.2. Proofs for Section 2.1

Proof of Proposition 1: Part one is just the same as part one of Proposition EC.3 in Section EC.7. Recall that $p_b^*(x) = \arg\max_{p\geq 0}\{d_b(p)(p - \Delta V_{\mathcal{D}}(x))\}$, and $p_s^*(x) = \arg\max_{p\geq 0}\{d_s(p)(\Delta V_{\mathcal{D}}(x+1) - p)\}$. For part two, it is straightforward as $p_s^*(x) \leq \Delta V_{\mathcal{D}}(x+1) \leq \Delta V_{\mathcal{D}}(x) \leq p_b^*(x)$.

For part three, as $\lim_{x\to\infty} \Delta V_{\mathcal{D}}(x) = 0$, $\lim_{x\to\infty} p_s^*(x) = 0$ and $\lim_{x\to\infty} p_b^*(x) = \arg\max_p\{d_b(p)p\}$. From Proposition EC.3, $\Delta V_{\mathcal{D}}(x)$ is decreasing in x and $\frac{d\Delta V_{\mathcal{D}}(x)}{dr} < 0$. Then by Lemma EC.3 in Section EC.7, we have the monotonicity of $p_b^*(x)$ and $p_s^*(x)$. The dependence on the market size when λ_b and λ_s change in proportion, is just the opposite to that of r.

The properties about the preferred inventory level follow directly from the monotonicity of $p_s^*(x)$ and $p_b^*(x)$. $\square$

Then, we show the properties of the value function of the fluid approximation.

PROPOSITION EC.1. *The value function $\bar{V}_{\mathcal{D}}(x)$ satisfies the following properties:*

1. *$\bar{V}_{\mathcal{D}}(x)$ is increasing and concave in x.*
2. *$\bar{V}_{\mathcal{D}}(x)$ is an upper bound of $V_{\mathcal{D}}(x)$ and $\lim_{x\to\infty} \bar{V}_{\mathcal{D}}(x) = \lim_{x\to\infty} V_{\mathcal{D}}(x) = \lambda_b \mathcal{R}_b(0)/r$.*

Proof of Proposition EC.1: By the coupling argument, it is easy to prove that $\bar{V}_{\mathcal{D}}(x)$ is strictly increasing in x. For concavity, note that we must have $\lambda_b d_b(\bar{p}_b^*(x)) \geq \lambda_s d_s(\bar{p}_s^*(x))$ for all $x \geq 0$, i.e., the inventory is always being depleted. This is because in the fluid approximation, there is no randomness and it is never optimal to stock more inventory. Because $\bar{V}_{\mathcal{D}}(x)$ is strictly increasing in x, we must have that for any $x > 0$,

$$0 < \frac{dr\bar{V}_{\mathcal{D}}(x)}{dx} = \frac{d[\lambda_b \mathcal{R}_b(\bar{V}_{\mathcal{D}}'(x)) + \lambda_s \mathcal{R}_s(\bar{V}_{\mathcal{D}}'(x))]}{dx} = [\lambda_s d_s(\bar{p}_s^*(x)) - \lambda_b d_b(\bar{p}_b^*(x))] V_{\mathcal{D}}''(x),$$

where the second equality follows from the envelope theorem. Therefore, $\bar{V}_{\mathcal{D}}''(x) < 0$ and $\bar{V}_{\mathcal{D}}(x)$ is increasing concave.

For part two, it is easy to show that $\lim_{x\to\infty} \bar{V}_{\mathcal{D}}(x) = \lambda_b \mathcal{R}_b(0)/r$.

Next, we show that $\bar{V}(x)$ is an upper bound for $V(x)$. Because $\bar{V}(x)$ is concave, $\Delta \bar{V}_{\mathcal{D}}(x) > \bar{V}_{\mathcal{D}}'(x) > \Delta \bar{V}_{\mathcal{D}}(x+1)$. By the property of $\mathcal{R}_b$ and $\mathcal{R}_s$ in Lemma EC.3 in Section EC.7, we have

$$r\bar{V}_{\mathcal{D}}(x) = \lambda_b \mathcal{R}_b(\bar{V}_{\mathcal{D}}'(x)) + \lambda_s \mathcal{R}_s(\bar{V}_{\mathcal{D}}'(x)) \geq \lambda_b \mathcal{R}_b(\Delta \bar{V}_{\mathcal{D}}(x)) + \lambda_s \mathcal{R}_s(\Delta \bar{V}_{\mathcal{D}}(x+1)),$$

for $x \in \mathbb{Z}^+$. Moreover, by the continuous differentiability of $\bar{V}_{\mathcal{D}}$, we have

$$r\bar{V}_{\mathcal{D}}(0) = \lambda_b \mathcal{R}_b(\bar{V}_{\mathcal{D}}'(0)) + \lambda_s \mathcal{R}_s(\bar{V}_{\mathcal{D}}'(0)) \geq \lambda_s \mathcal{R}_s(\bar{V}_{\mathcal{D}}'(0)) \geq \lambda_s \mathcal{R}_s(\Delta \bar{V}_{\mathcal{D}}(1)).$$

The HJB equation in (1) can be formulated as linear programming,

$$\begin{aligned} V_{\mathcal{D}}(y) = \min_{F(\cdot)} \quad & F(y) \\ \text{subject to} \quad & rF(0) \geq \lambda_s \mathcal{R}_s(\Delta F(1)), \\ & rF(x) \geq \lambda_b \mathcal{R}_b(\Delta F(x)) + \lambda_s \mathcal{R}_s(\Delta F(x+1)), \text{ for any } x \in \mathbb{Z}^+. \end{aligned}$$

Clearly, $\bar{V}_{\mathcal{D}}(x)$ is a feasible solution to the problem above, then $V_{\mathcal{D}}(x) \le \bar{V}_{\mathcal{D}}(x)$. $\square$

Proof of Proposition 2: When the dealer adopts the fixed pricing policies in (6), the dynamics of the inventory follow that of a state-independent M/M/1 queue. Denote $\tilde{V}_{\mathcal{D}}(x)$ as the expected discounted profit for this policy with x units of initial inventory. By the Markovian structure, we have

$$\begin{aligned}(r+\lambda_b d_b(\hat{p}_b)+\lambda_s d_s(\hat{p}_s))\tilde{V}_{\mathcal{D}}(x) &= \lambda_b d_b(\hat{p}_b)\tilde{V}_{\mathcal{D}}(x-1)+\lambda_s d_s(\hat{p}_s)\tilde{V}_{\mathcal{D}}(x+1)\\ &\quad +\hat{p}_b\lambda_b d_b(\hat{p}_b)-\hat{p}_s\lambda_s d_s(\hat{p}_s) \quad x>0,\\ (r+\lambda_s d_s(\hat{p}_s))\tilde{V}_{\mathcal{D}}(0) &= \lambda_s d_s(\hat{p}_s)\tilde{V}_{\mathcal{D}}(1)-\hat{p}_s\lambda_s d_s(\hat{p}_s). \end{aligned} \tag{EC.1}$$

From the first equation above, $\tilde{V}_{\mathcal{D}}(x)$ must be of the form $\tilde{V}_{\mathcal{D}}(x) = C_0 s_0^x + C_1 s_1^x + \frac{1}{r}(\hat{p}_b\lambda_b d_b(\hat{p}_b) - \hat{p}_s\lambda_s d_s(\hat{p}_s))$, where $0 < s_0 < 1 < s_1$ and s_0 and s_1 are the two solutions to $\lambda_s d_s(\hat{p}_s)x^2 - (r+\lambda_b d_b(\hat{p}_b)+\lambda_s d_s(\hat{p}_s))x + \lambda_b d_b(\hat{p}_b) = 0$, with

$$s_0 = \frac{r+\lambda_b d_b(\hat{p}_b)+\lambda_s d_s(\hat{p}_s)-\sqrt{(r+\lambda_b d_b(\hat{p}_b)+\lambda_s d_s(\hat{p}_s))^2-4\lambda_s\lambda_b d_s(\hat{p}_s)d_b(\hat{p}_b)}}{2\lambda_s d_s(\hat{p}_s)} \in (0,1).$$

Since $\tilde{V}_{\mathcal{D}}(x)$ is bounded above by the same argument as in Proposition EC.3 in Section EC.7, we must have $C_1 = 0$. By the boundary condition, the last equation in (EC.1),

$$C_0 = -\frac{\hat{p}_b\lambda_b d_b(\hat{p}_b)}{r+\lambda_s d_s(\hat{p}_s)(1-s_0)} < 0.$$

Now scale demand and supply simultaneously by n, i.e., $(\lambda_b^n, \lambda_s^n) = n(\lambda_b, \lambda_s)$ and denote $V_D^{(n)}(x)$,$\bar{V}_D^{(n)}(x)$,$\tilde{V}_D^{(n)}(x)$ as the corresponding value function in the original problem, fluid approximation and the fixed pricing policy, respectively. Denote $d_n = n\lambda_b d_b(\hat{p}_b) = n\lambda_s d_s(\hat{p}_s)$. Then

$$|\tilde{V}_{\mathcal{D}}^{(n)}(x) - \bar{V}_{\mathcal{D}}^{(n)}(0)| = |C_0|s_0^x \le |C_0| = \frac{\hat{p}_b d_n}{r+\left(\sqrt{r^2+4rd_n}-r\right)/2} = O(\sqrt{n}).$$

As $\bar{V}_{\mathcal{D}}^{(n)}(x) = \Omega(n)$, $\lim_{n\to\infty}\frac{\tilde{V}_{\mathcal{D}}^{(n)}(x)}{\bar{V}_{\mathcal{D}}^{(n)}(0)} = 1$ and $\tilde{V}_{\mathcal{D}}^{(n)}(x) \le V_{\mathcal{D}}^{(n)}(x) \le \bar{V}_{\mathcal{D}}^{(n)}(x)$, we have the result. $\square$

EC.3. Proofs for Section 2.2

Proof of Lemma 1: Given Assumptions 1 and 2, we first provide a lower bound for the expected waiting time when there are m sellers, $\mathbb{E}[\tau_m]$. Consider a case when no new sellers will come, buyers arrive with the largest effective rate λ_b, and all sellers have the same chance to sell their products. In this case, each seller's expected waiting time is

$$\sum_{i=1}^{m}\mathbb{P}(\text{sell to } i\text{-th buyer})\mathbb{E}[\text{Arrival Time of the } i\text{-th buyer}] \ge \sum_{i=1}^{m}\frac{1}{m}\frac{i}{\lambda_b} = \frac{m+1}{2\lambda_b}, \tag{EC.2}$$

which is apparently a lower bound for $\mathbb{E}[\tau_m]$.

Then, the capacity M can be upper bounded by a constant $2\bar{p}\lambda_b/\kappa$ because when $m \geq 2\bar{p}\lambda_b/\kappa - 1$, we have

$$f_m = \mathbb{E}[\tilde{p}_m - c_s - \kappa\tau_m] \leq \bar{p} - \kappa\mathbb{E}[\tau_m] \leq \bar{p} - \kappa\frac{m+1}{2\lambda_b} \leq 0,$$

where the first inequality holds because $\tilde{p} \leq \bar{p}$ and $c_s \geq 0$. The second inequality holds due to (EC.2). □

EC.3.1. Proof of Theorem 1

We first introduce some notations, and a few auxiliary results used in the proof. Define f^* as the unique solution to $\lambda_b\tilde{\mathcal{R}}_b(f^*) = \lambda_b \max_{p\geq 0}\{d_b(p+w)(p-f^*)\} = \kappa$.[12] Define f^{**} as the smallest positive solution to $\lambda_b d_b(f^{**}+w)f^{**} = \kappa$ if there exists such a solution or $+\infty$ otherwise. Note that when $f^* > 0$, $f^{**} < +\infty$ exists and $\frac{d\lambda_b d_b(p+w)p}{dp}\,|_{p=f^{**}} > 0$.[13] We also define the following function

$$R(x; y, z) \triangleq \lambda_b d_b(x+w)(x-y) - \kappa - \lambda_s d_s(z)(y-z). \tag{EC.3}$$

Using this function, the recursive system (8) can be rewritten as

$$\begin{aligned} \sup_x R(x; f_1, f_2) = 0, \\ R(f_{m-1}; f_m, f_{m+1}) = 0, m \geq 2. \end{aligned} \tag{EC.4}$$

Now we are ready to prove Theorem 1. Note that from the definitions of f^* and f^{**}, we only need to establish the proof for three separate cases: $f^* \leq 0$, $0 < f^* \leq f^{**}$ and $f^* > f^{**} > 0$. The proof is organized as follows: In Section EC.3.1.1, we prove the existence of solutions to (8) when $f^* \leq 0$ or $0 < f^* \leq f^{**}$. In Section EC.3.1.2, we prove the existence of solutions to (8) when $f^* > f^{**} > 0$. In Section EC.3.1.3, we demonstrate the properties that any solution to (8) will satisfy. Using these properties, we then prove that the solutions that have been found in Section EC.3.1.1 and Section EC.3.1.2 are the unique solutions to (8). Thus there exists a unique Markov perfect equilibrium of the dynamic game.

EC.3.1.1. Existence: $f^* \leq 0$ or $0 < f^* \leq f^{**}$. When $f^* \leq 0$ or $0 < f^* \leq f^{**}$, then we can show that

$$f_1 = f^*, \quad f_m = f_{m-1} - \frac{\kappa}{\lambda_b d_b(f_{m-1}+w)} \quad m \geq 2 \tag{EC.5}$$

[12] To show that f^* always exists, note that $\lim_{z\to+\infty}\tilde{\mathcal{R}}_b(z) = 0$ and $\lim_{z\to-\infty}\tilde{\mathcal{R}}_b(z) = +\infty$, and that $\tilde{\mathcal{R}}_b(z)$ is continuous. To show the uniqueness, note that by the similar proof in Lemma EC.3 in Section EC.7, $\tilde{\mathcal{R}}_b(\cdot)$ is monotone. Moreover, $\tilde{\mathcal{R}}_b'(z)|_{z=f^*} = -\lambda_b d_b(P_b(f^*)+w) < 0$, so $\tilde{\mathcal{R}}_b(z)$ is strictly decreasing at $z = f^*$. This implies that f^* is unique.

[13] If $f^* > 0$, then $\max_p\{\lambda_b d_b(p+w)p\} > \max_p\{\lambda_b d_b(p+w)(p-f^*)\} = \kappa$. Therefore, f^{**} exists. Moreover, by Assumption 1, $\lambda_b d_b(p+w)p$ is unimodal in $p \in [0,\infty)$. Because f^{**} is the smaller solution, we have $\frac{d\lambda_b d_b(p+w)p}{dp}\,|_{p=f^{**}} > 0$.

is a solution to (8). In fact, if $f^* \leq 0$, then $f_2 \leq f_1 = f^* \leq 0$. Therefore, $M = 0$ and (8) degenerates to a single equation $\kappa = \max_{f_0}\{\lambda_b d_b(f_0 + w)(f_0 - f_1)\}$, which holds when $f_1 = f^*$ by the definition of f^*.

If $0 < f^* \leq f^{**}$, then from the definition of f^{**} (see Footnote 13), $f^* \leq f^{**} < p^* \triangleq \arg\max_{p\geq 0}\{\lambda_b d_b(p+w)p\}$. By Assumption 1, $\lambda_b d_b(p+w)p$ is increasing in $[0, p^*)$ and thus $\lambda_b d_b(f^*+w)f^* \leq \lambda_b d_b(f^{**}+w)f^{**} = \kappa$ (by definition of f^{**}). Therefore, $f_2 = f^* - \frac{\kappa}{\lambda_b d_b(f^*+w)} \leq 0$, so $M = 1$ and (8) again degenerates to a single equation $\kappa = \max_{f_0}\{\lambda_b d_b(f_0 + w)(f_0 - f_1)\}$, which holds for $f_1 = f^*$.

EC.3.1.2. Existence: $f^* > f^{**} > 0$. To show the existence of solutions to (8) when $f^* > f^{**} > 0$, we design Algorithm 1 (see Section EC.6 in the additional material). The algorithm is based on the following intuition: Assign a value to f_M, then recursively compute f_{M-1}, f_{M-2}, ..., f_1 using the Mth, $(M-1)$th, ..., second equation of (8). Then check whether the first equation of (8) holds. Since M is not known in advance, in the algorithm, we use X_n to denote f_{M+1-n}. When backwards calculating X_{n+1}, we use $R(X_{n+1}; X_n, X_{n-1}) = 0$ due to (EC.4). The additional constraint in the computation $\frac{dR(x;X_n,X_{n-1})}{dx}|_{x=X_{n+1}} > 0$ will be specified in Lemma EC.2 as a necessary condition for any solution to (8).

The algorithm takes an input a, which is a candidate solution for f_M. If the output is $R_N = 0$, $N < \infty$, $X_0 \leq 0$ and $X_i \geq 0$ for $i > 0$, then $[X_N, \cdots, X_1, X_0]$ must satisfy

$$\sup_x R(x; X_N, X_{N-1}) = 0,$$
$$R(f_{n+1}; f_n, f_{n-1}) = 0, \quad 1 \leq n < N.$$

It solves (EC.4), or equivalently (8) , by substituting $M = N$, $f_1 = X_N$, $\cdots$, $f_M = X_1$, $f_{M+1} = X_0$.

Denote $N(a)$ as the output of N when the input is a. Similarly, we introduce $X_n(a)$ and $R_n(a)$ for fixed n. First, we point out some facts for the proposed algorithm, based on Assumption 1 and Lemma 1 and the following conditions: $f^* \geq f^{**} > 0$ and $a \in [0, f^{**}]$. These facts are then used to prove the existence of solutions to (8) when $f^* > f^{**} > 0$ based on Algorithm 1.

Fact 1: $X_0(a) \leq 0$, $X_0(a) < X_1(a)$ and $\frac{d\lambda_b d_b(x+w)(x-X_0(a))}{dx}|_{x=X_1(a)} > 0$.

This fact states that if $f_M = a \in [0, f^{**}]$, then $f_{M+1} = X_0(a) \leq 0$. So the choice of a is proper.

Proof of Fact 1: By their definitions in the algorithm, $X_1(a) = a \geq 0$ and $X_0(a) < X_1(a)$. Let $p^* \triangleq \arg\max_{p\geq 0}\{\lambda_b d_b(p + w)p\}$. By the definition (Footnote 13), $f^{**} < p^*$. By Assumption 1, $\lambda_b d_b(p + w)p$ is increasing in $[0, p^*)$. Then as $a \leq f^{**} < p^*$, $\lambda_b d_b(a + w)a \leq \lambda_b d_b(f^{**} + w)f^{**} = \kappa$ and thus $X_0(a) = a - \frac{\kappa}{\lambda_b d_b(a+w)} \leq 0$, where the equality holds only when $a = f^{**}$. When $a = f^{**}$, $\frac{d\lambda_b d_b(x+w)(x-X_0(a))}{dx}|_{x=X_1(a)} = \frac{d\lambda_b d_b(x+w)x}{dx}|_{x=f^{**}} > 0$, where the inequality is due to the definition of f^{**}. When $a < f^{**}$, $X_0(a) < 0$ and thus

$$\lambda_b d_b(X_1(a) + w)(X_1(a) - X_0(a)) = \kappa = \lambda_b d_b(f^{**} + w)f^{**} < \lambda_b d_b(f^{**} + w)(f^{**} - X_0(a)),$$

where the first equality is due to the definition of $X_0(a)$ in the algorithm and the second equality is due to the definition of f^{**}. Due to Assumption 1, $\lambda_b d_b(x+w)(x-X_0(a))$ is unimodal in x and as $X_1(a)=a<f^{**}$, the inequality above leads to $\frac{d\lambda_b d_b(x+w)(x-X_0(a))}{dx}|_{x=X_1(a)}>0$. $\square$

Fact 2: If $R_n(a)>0$, $X_{n+1}(a)$ exists and is unique. Moreover, $X_{n+1}(a)\geq X_n(a)+\frac{\kappa}{\lambda_b d_b(w)}>X_n(a)\geq 0$.

This fact states that for any input $a\in[0,f^{**}]$, there is a unique output by the algorithm.

Proof of Fact 2: As $X_1(a)=a\geq 0\geq X_0(a)$, we have $R(0;X_1(a),X_0(a))<0$. By Assumption 1, $R(x;X_1(a),X_0(a))$ is unimodal in $x\in[0,\infty)$. Then if $R_1(a)=\sup_{x\geq 0}R(x;X_1(a),X_0(a))>0$, there exists a unique $X_2(a)>0$ satisfying $R(X_2(a);X_1(a),X_0(a))=0$ and $\frac{dR(x;X_1(a),X_0(a))}{dx}|_{x=X_2(a)}>0$. As $R(X_2(a);X_1(a),X_0(a))=\lambda_b d_b(X_2(a)+w)(X_2(a)-X_1(a))-\kappa-\lambda_s d_s(X_0(a))(X_1(a)-X_0(a))=0$ and $X_1(a)>X_0(a)$, we have $\lambda_b d_b(X_2(a)+w)(X_2(a)-X_1(a))-\kappa\geq 0$. That is, $X_2(a)-X_1(a)\geq \frac{\kappa}{\lambda_b d_b(X_2(a)+w)}>\frac{\kappa}{\lambda_b d_b(w)}>0$. By induction, if $R_n(a)>0$, $X_{n+1}(a)$ exists and is unique. Besides, $X_{n+1}(a)\geq X_n(a)+\frac{\kappa}{\lambda_b d_b(w)}>X_n(a)\geq 0$. $\square$

Fact 3: $N(f^{**})=N(0)-1>0$, $R_{N(0)}(0)=R_{N(0)-1}(f^{**})\leq 0$ and $R_{N(0)-1}(0)>0$.

This fact checks the property of two extremes of the input $a\in[0,f^{**}]$.

Proof of Fact 3: By the step in the algorithm, $R_1(0)=\sup_x R(x;0,-\frac{\kappa}{\lambda_b d_b(w)})=\max_p\{\lambda_b d_b(p+w)p\}-\kappa>0$, where the inequality is due to $f^*>0$. From the algorithm, it will lead to $N(0)>1$. Clearly, $X_0(f^{**})=0$, $X_1(f^{**})=f^{**}$, $X_1(0)=0$ and $X_2(0)=f^{**}$. Therefore, there is always a lag of one when using the input $a=0$ and $a=f^{**}$. One can check that $X_n(0)=X_{n-1}(f^{**})$ and thus $N(f^{**})=N(0)-1>0$. From the condition that the algorithm stops, we have $R_{N(0)}(0)=R_{N(f^{**})}(f^{**})=R_{N(0)-1}(f^{**})\leq 0$ and $R_{N(0)-1}(0)>0$. $\square$

Fact 4: $N(a)<+\infty$.

Proof of Fact 4: From *Fact 2*, $X_n(a)\geq X_1(a)+\frac{(n-1)\kappa}{\lambda_b d_b(w)}$. If $N(a)$ is infinite, then when $n\to\infty$, $X_n(a)\geq X_{n-1}(a)=+\infty$, which will lead to $R_n<0$, in contradiction with the condition that the algorithm stops. $\square$

Lemma EC.1. *Under Assumption 1 and 2, when $f^*\geq f^{**}>0$, there exists a unique $a^*\in(0,f^{**}]$ such that when the input is a^*, Algorithm 1 outputs $R_N=0$, $N<\infty$, $X_0\leq 0$ and $X_i\geq 0$ for $i>0$. That is, one solution to (8) can be found by setting $M=N(a^*)$, $f_1=X_N(a^*)$, $\cdots$, $f_M=X_1(a^*)$, $f_{M+1}=X_0(a^*)$.*

Proof of Lemma EC.1: We divide the proof into several steps.

Increase the input $a\to a+\Delta a$ by an infinitesimal positive amount Δa. Denote $\tilde{N}=\min\{N(a),N(a+\Delta a)\}$ and $X_n(a+\Delta a)=X_n(a)+\Delta X_n$ for $n\leq\tilde{N}$. With a slight abuse of notation, we use X_n to refer to $X_n(a)$ below.

Step (1): $\Delta X_{\tilde{N}}>\cdots>\Delta X_n>\cdots>\Delta X_1>\Delta X_0>0$.

Note that $X_0 = X_1 - \frac{\kappa}{\lambda_b d_b(X_1+w)}$ and $X_0 + \Delta X_0 = X_1 + \Delta X_1 - \frac{\kappa}{\lambda_b d_b(X_1+\Delta X_1+w)}$ from Step 3 in the algorithm when the input is a and $a+\Delta a$ respectively. These equalities can be rewritten as $\lambda_b d_b(X_1+w)(X_1-X_0)=\kappa$ and $\lambda_b d_b(X_1+\Delta X_1+w)((X_1+\Delta X_1)-(X_0+\Delta X_0))=\kappa$. Therefore

$$\begin{aligned} 0 =& \lambda_b d_b(X_1+\Delta X_1+w)((X_1+\Delta X_1)-(X_0+\Delta X_0)) - \lambda_b d_b(X_1+w)(X_1-X_0) \\ =& \frac{d\lambda_b d_b(x+w)(x-X_0)}{dx}\,|_{x=X_1}\,\Delta X_1 - \lambda_b d_b(X_1+w)\Delta X_0 \end{aligned}$$

where the second equality follows from a Taylor expansion. From *Fact 1*, $\frac{d\lambda_b d_b(x+w)(x-X_0)}{dx}\,|_{x=X_1} > 0$. Then due to $\Delta X_1 = \Delta a > 0$, the equation above will lead to $\Delta X_0 > 0$. Besides, the equation above can be rewritten as

$$\lambda_b d_b'(X_1+w)(X_1-X_0)\Delta X_1 + \lambda_b d_b(X_1+w)(\Delta X_1 - \Delta X_0) = 0.$$

From *Fact 1*, $X_1 > X_0$. As $\lambda_b d_b'(X_1+w) < 0$ and $\Delta X_1 > 0$, the first term in the left hand above is negative, and thus the second term in the left hand above is positive, which leads to $\Delta X_1 > \Delta X_0 > 0$.

Now from $\Delta X_1 > \Delta X_0$ and $X_1 > X_0$, one can show that $(X_1+\Delta X_1)-(X_0+\Delta X_0) > X_1 - X_0 > 0$, $(X_1+\Delta X_1) - X_1 > (X_0+\Delta X_0) - X_0 > 0$. Then from Lemma EC.4 in Section EC.8, by setting $y_2 = X_1+\Delta X_1$, $z_2 = X_0+\Delta X_0$, $y_1 = X_1$ and $z_1 = X_0$, we have $\Delta X_2 - \Delta X_1 > 0$. By induction, for all $1 \le n \le \tilde{N}$, $\Delta X_n > \Delta X_{n-1} > \cdots > \Delta X_1 > 0$.

Step (2): $N(a+\Delta a) \le N(a)$ and $R_n(a+\Delta a) < R_n(a)$ for all $1 \le n \le N(a+\Delta a)$.

By the definition in the algorithm $R_n(a) = \mathcal{R}_b(X_n(a)) - \kappa - \lambda_s d_s(X_{n-1}(a))(X_n(a) - X_{n-1}(a))$. Then for all $1 \le n \le \tilde{N}$, by Taylor's expansion, $R_n(a+\Delta a) = R_n(a) + \mathcal{R}_b'(X_n+w)\Delta X_n - [\lambda_s d_s(X_{n-1})(\Delta X_n - \Delta X_{n-1}) + \lambda_s d_s'(X_{n-1})(X_n - X_{n-1})\Delta X_{n-1}]$. As $X_n \ge X_{n-1}$ by *Fact 2* and $\Delta X_n \ge \Delta X_{n-1} > 0$ by Step (1), we have $R_n(a+\Delta a) < R_n(a)$.

Next, we prove $N(a+\Delta a) \le N(a)$ by contradiction. Suppose $N(a+\Delta a) > N(a)$. Then by the condition that the algorithm stops, $R_{N(a)}(a+\Delta a) > 0$ and $R_{N(a)}(a) \le 0$, contradicting $R_{N(a)}(a+\Delta a) < R_{N(a)}(a)$ we have just proved. As a result, $N(a+\Delta a) \le N(a)$.

Step (3): There exists a unique value, $a^* \in (0, f^{**}]$, such that $R_{N(a^*)}(a^*) = 0$, $N(a^*) < \infty$, $X_0(a^*) \le 0$ and $X_i(a^*) \ge 0$ for $i > 0$.

From Step (2), $N(a)$ is non-increasing in a. From *Fact 3*, $N(f^{**}) = N(0) - 1$. As a result, $N(0) \ge N(a) \ge N(f^{**}) = N(0) - 1$ when $a \in (0, f^{**}]$. That is, when $a \in (0, f^{**}]$, the algorithm terminates with $N(a) \in \{N(0)-1, N(0)\}$. Next we show that $N(a^*) = N(0) - 1$. Suppose that the algorithm terminates with $N = N(0)$. From Step (2), the termination of the algorithm must satisfy $R_N < R_{N(0)}(0) \le 0$. As a result, if the output of algorithm satisfies $R_{N(a^*)}(a^*) = 0$, we must have $N(a^*) = N(0) - 1$. From *Fact 3*, $R_{N(0)-1}(f^{**}) \le 0$ and $R_{N(0)-1}(0) > 0$. Then because of Step

(2), $R_{N(0)-1}(a)$ is decreasing in a and thus there exists a unique value $a^* \in (0, f^{**}]$ such that $R_{N(a^*)}(a^*) = 0$. By *Fact 1*, *Fact 2* and *Fact 4*, $N(a^*) < \infty$, $X_0(a^*) \le 0$ and $X_i(a^*) \ge 0$ for $i > 0$.

By the discussion of the intuition to design the algorithm, one solution to (8) can be derived by setting $M = N(a^*)$, $f_1 = X_N(a^*)$, $\cdots$, $f_M = X_1(a^*)$, $f_{M+1} = X_0(a^*)$. $\square$

EC.3.1.3. Structural Properties In this subsection, we show the structural properties of any solution to (8), using the definition of the last section.

Lemma EC.2. *Under Assumption 1 and 2, any solution to (8) satisfies the following properties, (1) f_m is decreasing and convex in m. (2) For $1 < m \le M+1$, we have $\frac{d}{dx}[\lambda_b d_b(x+w)(x - f_m)]|_{x=f_{m-1}} > 0$, and $\frac{\lambda_b d_b(x+w)x + \lambda_s d_s(f_{m+1}) f_{m+1} - \kappa}{\lambda_b d_b(x+w) + \lambda_s d_s(f_{m+1})} < f_m$ when $x < f_{m-1}$. (3) $f_1 \le f^*$ and $f_M \le f^{**}$. (4) $M = 0$ if and only if $f^* \le 0$; $M = 1$ if and only if $0 < f^* \le f^{**}$; $M > 1$ if and only if $f^* > f^{**}$.*

Proof of Lemma EC.2: To show part (1), we first show f_m is decreasing in m. For $m \ge M$, according to the m-th equation in (8), $f_{m-1} - f_m = \frac{\kappa}{\lambda_b d_b(f_{m-1}+w)} > 0$. Suppose there exists $m < M$ such that $f_{m-1} - f_m \le 0$. Then according to the mth equation in (8), $\lambda_s d_s(f_{m+1})(f_m - f_{m+1}) + \kappa = \lambda_b d_b(f_{m-1} + w)(f_{m-1} - f_m) \le 0$. Therefore, $f_m \le f_{m+1}$. By induction, from the mth to the $(M-1)$-th equations in (8), we have $f_{M-1} \le f_M$. However, from the M-th equation in (8), $f_{M-1} - f_M = \frac{\kappa}{\lambda_b d_b(f_{M-1}+w)} > 0$. This leads to a contradiction and proves that f_m must be decreasing.

Next we show that $f_m - f_{m+1}$ is decreasing in m in part (1). For $m \ge M$, $f_m - f_{m+1} = \frac{\kappa}{\lambda_b d_b(f_m+w)}$, which is clearly decreasing in m as f_m is decreasing in m and $d_b(\cdot)$ is a decreasing function on its support. For $m < M$, starting from the $(M-1)$-th and Mth equations in (8),

$$\begin{aligned} \lambda_s d_s(f_M)(f_{M-1} - f_M) + \kappa &= \lambda_b d_b(f_{M-2} + w)(f_{M-2} - f_{M-1}), \\ \kappa &= \lambda_b d_b(f_{M-1} + w)(f_{M-1} - f_M). \end{aligned}$$

We have $\lambda_b d_b(f_{M-2} + w)(f_{M-1} - f_M) < \lambda_b d_b(f_{M-1} + w)(f_{M-1} - f_M) = \kappa < \lambda_b d_b(f_{M-2} + w)(f_{M-2} - f_{M-1})$, where the first inequality follows from $f_{M-2} > f_{M-1}$, the equality follows from the second equation above, and the second inequality follows from the first equation above and $f_{M-1} \ge f_M$. As a result, $f_{M-1} - f_M < f_{M-2} - f_{M-1}$. Suppose $f_{m-1} - f_m < f_{m-2} - f_{m-1}$ for some $m < M$. Then

$$\begin{aligned} \lambda_b d_b(f_{m-2} + w)(f_{m-2} - f_{m-1}) &= \lambda_s d_s(f_m)(f_{m-1} - f_m) + \kappa \\ &< \lambda_s d_s(f_m)(f_{m-2} - f_{m-1}) + \kappa < \lambda_s d_s(f_{m-1})(f_{m-2} - f_{m-1}) + \kappa \\ &= \lambda_b d_b(f_{m-3} + w)(f_{m-3} - f_{m-2}) < \lambda_b d_b(f_{m-2} + w)(f_{m-3} - f_{m-2}), \end{aligned}$$

where the equalities follow from (8), the first inequality follows from $f_{m-1} - f_m < f_{m-2} - f_{m-1}$, the second inequality follows from $f_{m-1} > f_m$ and the last inequality follows from $f_{m-2} < f_{m-3}$. As a result, $f_{m-2} - f_{m-1} < f_{m-3} - f_{m-2}$. By induction, we have shown the claim.

Next we show part (2): $\frac{d}{dx}[\lambda_b d_b(x+w)(x-f_m)]|_{x=f_{m-1}} > 0$ for $1 < m \le M$. Note that

$$\begin{aligned}\lambda_b d_b(f_{m-1}+w)(f_{m-1}-f_m) =& \kappa + \lambda_s d_s(f_{m+1})(f_m - f_{m+1}) < \kappa + \lambda_s d_s(f_m)(f_{m-1}-f_m)\\ =& \lambda_b d_b(f_{m-2}+w)(f_{m-2}-f_{m-1}) < \lambda_b d_b(f_{m-2}+w)(f_{m-2}-f_m),\end{aligned}$$

where the equalities follow from (8), the first inequality follows from $f_{m+1} < f_m$ and $f_m - f_{m+1} < f_{m-1} - f_m$, and the last inequality is due to $f_m < f_{m-1}$. Let $x^* = \arg\max_x d_b(x+w)(x-f_m)$. By Assumption 1, $d_b(x+w)(x-f_m)$ is unimodal and increasing in $x \in [0, x^*)$. Since $f_{m-1} < f_{m-2}$, the inequality above indicates that $f_{m-1} < x^*$. Thus, $\frac{d}{dx}[\lambda_b d_b(x+w)(x-f_m)]|_{x=f_{m-1}} > 0$. Moreover, when $x < f_{m-1}$, we have

$$\begin{aligned}&\frac{\lambda_b d_b(x+w)x + \lambda_s d_s(f_{m+1})f_{m+1} - \kappa}{\lambda_b d_b(x+w) + \lambda_s d_s(f_{m+1})} - f_m\\ =&\frac{\lambda_b d_b(x+w)(x-f_m) + \lambda_s d_s(f_{m+1})(f_{m+1}-f_m) - \kappa}{\lambda_b d_b(x+w) + \lambda_s d_s(f_{m+1})}\\ <&\frac{\lambda_b d_b(f_{m-1}+w)(f_{m-1}-f_m) + \lambda_s d_s(f_{m+1})(f_{m+1}-f_m) - \kappa}{\lambda_b d_b(x+w) + \lambda_s d_s(f_{m+1})}\\ =&0,\end{aligned}$$

where the inequality is because $\frac{d}{dx}[\lambda_b d_b(x+w)(x-f_m)] > 0$ when $x < f_{m-1}$, and the last equality is due to (8).

Next, we show part (3): $f_1 \le f^*$ and $f_M \le f^{**}$. Note that $\max_p\{\lambda_b d_b(p+w)(p-f_1)\} = \lambda_s d_s(f_2)(f_1-f_2)+\kappa \ge \kappa = \max_p\{\lambda_b d_b(p+w)(p-f^*)\}$, where the first equality is due to (8) and the second equality is due to the definition of f^*, and the inequality is due to the fact that f_m is decreasing in m. As $\max_p\{\lambda_b d_b(p+w)(p-z)\}$ is decreasing in z, $f_1 \le f^*$. To show $f_M \le f^{**}$, we only need to focus on the case where f^{**} is finite. According to the definition of f^{**}, $\frac{d}{dx}[\lambda_b d_b(x+w)x]|_{x=f^{**}} \ge 0$. Besides,

$$\lambda_b d_b(f^{**}+w)f^{**} = \kappa = \lambda_b d_b(f_M+w)(f_M - f_{M+1}) \ge \lambda_b d_b(f_M+w)f_M,$$

where the first equality is due to the definition of f^{**}, the second equality is due to the $(M+1)$th equation in (8), and the inequality is due to $f_{M+1} \le 0$. Also, since $\frac{d}{dx}[\lambda_b d_b(x+w)(x-f_{M+1})]|_{x=f_M} \ge 0$ by part (2), we have $\frac{d}{dx}[\lambda_b d_b(x+w)x]|_{x=f_M} \ge 0$. As $\lambda_b d_b(x+w)x$ is unimodal by Assumption 1, the inequality above indicates $f_M \le f^{**}$.

Finally, we show part (4): $M=0$ if and only if $f^* \le 0$; $M=1$ if and only if $0 < f^* \le f^{**}$; $M > 1$ if and only if $f^* > f^{**}$.

We first show that they are *necessary conditions*. a) If $M=0$, $f_1 = f^* \le 0$. b) If $M=1$, due to the monotonicity of f_m, it requires that $f_0 > f_1 = f^* > 0 > f_2$. Then $\lambda_b d_b(f_1+w)f_1 \le \kappa = \lambda_b d_b(f_0+w)(f_0-f_1) < \lambda_b d_b(f_0+w)f_0$, where the first inequality is due to $f_2 = f_1 - \frac{\kappa}{\lambda_b d_b(f_1+w)} \le 0$, the equality is due to the first equation in (8), and the last inequality is due to $f_1 > 0$. By Assumption 1, $\lambda_b d_b(p+w)p$ is unimodal. As $f_1 < f_0$, the inequality above leads to $\frac{d}{dp}[\lambda_b d_b(p+w)p]|_{p=f_1} \geq 0$. Besides, $\lambda_b d_b(f^{**}+w)f^{**} = \kappa \geq \lambda_b d_b(f_1+w)f_1$, where the equality is due to the definition of f^{**} and the inequality is due to $f_2 = f_1 - \frac{\kappa}{\lambda_b d_b(f_1+w)} \leq 0$. Then we have $f^{**} \geq f_1 = f^*$.

c) If $M > 1$, we know $\lambda_b d_b(f_1+w)(f_1-f_2) = \kappa + \lambda_s d_s(f_3)(f_2-f_3) \geq \kappa$, where the equality is due to (8) and the inequality is due to $f_2 > f_3$. As $f_2 > 0$ for $M > 1$, it leads to $\lambda_b d_b(f_1+w)f_1 > \kappa$. Then by the definition of f^{**}, $\lambda_b d_b(f_1+w)f_1 > \lambda_b d_b(f^{**}+w)f^{**}$. As $\frac{d}{dx}[\lambda_b d_b(x+w)x]|_{x=f^{**}} \geq 0$ and $\lambda_b d_b(x+w)x$ is unimodal in x, we have $f_1 > f^{**}$. Thus by part (3) $f^* > f^{**}$.

Next we show they are *sufficient conditions*. a) For $f^* < 0$, since $f_1 \leq f^* \leq 0$ by part (3), we have $M = 0$. b) For $0 < f^* \leq f^{**}$, suppose $M > 1$, we must have $f^* \geq f_1 > f^{**}$ from the necessary condition, leading to contradiction. Suppose $M = 0$, then $f_1 = f^* \leq 0$, leading to contradiction. Therefore, $M = 1$. c) For $f^* > f^{**}$, suppose $M \leq 1$, we must have $f^* \leq 0$ or $0 < f^* \leq f^{**}$ from the necessary condition, leading to contradiction. Therefore, $M > 1$. □

Proof of Theorem 1: We establish the existence in Section EC.3.1.1 and Section EC.3.1.2. We then need the following steps to show the uniqueness.

Step (1): The solution to (8) is unique when $f^* \leq 0$ or $0 < f^* \leq f^{**}$. By Lemma EC.2, $M \leq 1$. Then the unique solution to (8) can be directly solved, having the form in (EC.5).

Step (2): The solution to (8) is unique when $f^* > f^{**}$. We will show that any solution can be derived from the proposed algorithm with the output $R_N = 0$. Let $[f_1, \cdots, f_M, \cdots]$ be a solution to (8). By part (2) in Lemma EC.2, for $1 < m \leq M$, $\frac{dR(x; f_m, f_{m+1})}{dx}|_{x=f_{m-1}} = \frac{d}{dx}[\lambda_b d_b(x+w)(x - f_m)]|_{x=f_{m-1}} > 0$. Due to (8) , for $1 < m \leq M$, $R(f_{m-1}; f_m, f_{m+1}) = 0$, then $\sup_x R(x; f_m, f_{m+1}) > 0$. Besides, the first equation in (8) can be rewritten as $\sup_x R(x; f_1, f_2) = 0$. Therefore if we set the input a to be f_M in the solution, then the output of the algorithm satisfies $R_N = 0$, $N = M$, $X_0 = f_{M+1}$, $X_1 = f_M$, $\cdots$, $X_N = f_1$. Due to $f_M \leq f^{**}$ by part (3) in Lemma EC.2 and Lemma EC.1, the solution to (8) is uniquely output by Algorithm 1. This proves Theorem 1. □

Proof of Proposition 3 It is straightforward to prove part one and part two by part one in Lemma EC.2.

For part three, suppose M is unchanged when we change κ. We will examine how f_m changes with respect to κ. Taking derivative with respect to κ on both sides of the recursive system (8) yields,

$$\begin{aligned} \lambda_s d_s(f_2)(f_1' - f_2') + a_1 f_2' + 1 &= b_1 f_1', \\ \lambda_s d_s(f_{m+1})(f_m' - f_{m+1}') + a_m f_{m+1}' + 1 &= b_m f_{m-1}' + \lambda_b d_b(f_{m-1}+w)(f_{m-1}' - f_m')\ M-1 \geq m \geq 2, \\ 1 &= b_M f_{M-1}' + \lambda_b d_b(f_{M-1}+w)(f_{M-1}' - f_M'), \end{aligned} \tag{EC.6}$$

where $f_i' = \frac{df_i}{d\kappa}$, $a_m = \lambda_s d_s'(f_{m+1})(f_m - f_{m+1})$, and

$$b_m = \begin{cases} \frac{d}{df_1}[\max_{f_0}\{\lambda_b d_b(f_0+w)(f_0-f_1)\}], & m=1, \\ \lambda_b d_b'(f_{m-1}+w)(f_{m-1}-f_m), & m>1. \end{cases}$$

From the properties of $d_b(\cdot)$, $d_s(\cdot)$ and the monotonicity of f_m, we have $a_m > 0$, $b_m < 0$ for any m. As $\frac{d}{dx}[\lambda_b d_b(x+w)(x-f_m)]|_{x=f_{m-1}} > 0$ by part (2) in Lemma EC.2, rewriting this inequality, we have $b_m + \lambda_b d_b(f_{m-1}+w) > 0$.

Next, we prove $f_M' < 0$ by contradiction. Suppose $f_M' \geq 0$. From the last equation in (EC.6), we have $1 = (b_M + \lambda_b d_b(f_{M-1}+w))f_{M-1}' - \lambda_b d_b(f_{M-1}+w)f_M'$. As $b_M + \lambda_b d_b(f_{M-1}+w) > 0$, we must have $f_{M-1}' > 0$. Then as $b_M < 0$ and $f_{M-1}' > 0$, this equation also shows $1 = b_M f_{M-1}' + \lambda_b d_b(f_{M-1}+w)(f_{M-1}' - f_M') \leq \lambda_b d_b(f_{M-1}+w)(f_{M-1}' - f_M')$. Therefore $f_{M-1}' > f_M' \geq 0$. By recursively applying the same argument to the other equations in (EC.6) (replacing the left-hand side 1 by $\lambda_s d_s(f_{m+1})(f_m' - f_{m+1}') + a_m f_{m+1}' + 1$, which is still positive), we have $f_1' > f_2' > 0$. This cannot happen, as in the first equation of (EC.6), if $f_1' > f_2' > 0$, as $a_1 > 0$, the left hand is positive; however, as $b_1 < 0$, the right hand is negative, leading to a contradiction. As a result, $f_M' < 0$.

The above proof shows that when κ increases, f_M decreases. When f_M becomes zero, then the endogenous capacity M decreases by one. Therefore, M is decreasing in κ. Similarly we can prove how f_M changes in λ_b and λ_s. The endogenous capacity is decreasing (increasing, decreasing, decreasing) in κ (λ_b, λ_s, w).

Due to part four in Lemma EC.2, $M \geq 1$ if and only if $\max_p\{d_b(p+w)p\} \geq \kappa/\lambda_b$.

For part four, we divide the proof into several steps.

(1) $V_{\mathcal{M}}(m)$ is increasing in m. From the first equation in (9), $V_{\mathcal{M}}(0) < V_{\mathcal{M}}(1)$. We prove the claim for $m > 0$ by contradiction. Suppose $V_{\mathcal{M}}(l-1) < V_{\mathcal{M}}(l)$ holds for $l \leq m$, but $V_{\mathcal{M}}(m) \geq V_{\mathcal{M}}(m+1)$ for some $m \geq 1$. Then we have

$$\begin{aligned} &[r + \lambda_s d_s(f_{m+1}) + \lambda_b d_b(f_{m-1}+w)]V_{\mathcal{M}}(m) \\ =&\lambda_s d_s(f_{m+1})V_{\mathcal{M}}(m+1) + \lambda_b d_b(f_{m-1}+w)[V_{\mathcal{M}}(m-1)+w] \\ <&\lambda_s d_s(f_{m+1})V_{\mathcal{M}}(m) + \lambda_b d_b(f_{m-1}+w)[V_{\mathcal{M}}(m)+w], \end{aligned}$$

where the equality is due to (9), and the inequality is due to our induction hypothesis that $V_{\mathcal{M}}(m) \geq V_{\mathcal{M}}(m+1)$ and $V_{\mathcal{M}}(m-1) \leq V_{\mathcal{M}}(m)$. After arranging the terms, we have $rV_{\mathcal{M}}(m) < \lambda_b d_b(f_{m-1}+w)w$. Due to the monotonicity of f_m, $rV_{\mathcal{M}}(m) < \lambda_b d_b(f_m+w)w$. Since $V_{\mathcal{M}}(m) \geq V_{\mathcal{M}}(m+1)$, we have $rV_{\mathcal{M}}(m+1) < \lambda_b d_b(f_m+w)w$. As a result,

$$\begin{aligned} &(\lambda_s d_s(f_{m+2}) + \lambda_b d_b(f_m+w))V_{\mathcal{M}}(m+1) + \lambda_b d_b(f_m+w)w \\ >&(\lambda_s d_s(f_{m+2}) + \lambda_b d_b(f_m+w) + r)V_{\mathcal{M}}(m+1) \\ =&\lambda_s d_s(f_{m+2})V_{\mathcal{M}}(m+2) + \lambda_b d_b(f_m+w)[V_{\mathcal{M}}(m)+w] \\ \geq&\lambda_s d_s(f_{m+2})V_{\mathcal{M}}(m+2) + \lambda_b d_b(f_m+w)(V_{\mathcal{M}}(m+1)+w), \end{aligned}$$

where the equality is due to (9), the first inequality is due to $rV_{\mathcal{M}}(m+1) < \lambda_b d_b(f_m + w)w$, the second inequality is due to $V_{\mathcal{M}}(m+1) \le V_{\mathcal{M}}(m)$. Arranging the terms, we have $V_{\mathcal{M}}(m+1) \ge V_{\mathcal{M}}(m+2)$. Moreover, as $rV_{\mathcal{M}}(m+1) < \lambda_b d_b(f_m + w)w$ and $f_m > f_{m+1}$, we have $rV_{\mathcal{M}}(m+2) < \lambda_b d_b(f_m + w)w < \lambda_b d_b(f_{m+1} + w)w$. By induction, we can show $V_{\mathcal{M}}(M-1) \ge V_{\mathcal{M}}(M)$ and $rV_{\mathcal{M}}(M) < \lambda_b d_b(f_{M-1} + w)w$. However, this leads to contradiction as the last equation in (9) can not hold. Thus, $V_{\mathcal{M}}(m)$ is increasing in m.

(2) $V_{\mathcal{M}}(m) - V_{\mathcal{M}}(m-1)$ is decreasing in m. By the last two equations of 9, we have

$$\begin{aligned}
&\lambda_b d_b(f_{M-2}+w)[V_{\mathcal{M}}(M-1) - V_{\mathcal{M}}(M-2) - w] \\
=& -rV_{\mathcal{M}}(M-1) + \lambda_s d_s(f_M)[V_{\mathcal{M}}(M) - V_{\mathcal{M}}(M-1)] \\
>& -rV_{\mathcal{M}}(M) = \lambda_b d_b(f_{M-1}+w)[V_{\mathcal{M}}(M) - V_{\mathcal{M}}(M-1) - w]
\end{aligned}$$

where the inequality is due to $V_{\mathcal{M}}(M) > V_{\mathcal{M}}(M-1)$. Since f_m is decreasing, $\lambda_b d_b(f_{M-1}+w) > \lambda_b d_b(f_{M-2}+w)$, and the inequality above leads to $V_{\mathcal{M}}(M-1) - V_{\mathcal{M}}(M-2) > V_{\mathcal{M}}(M) - V_{\mathcal{M}}(M-1)$. Now we use induction. Suppose $V_{\mathcal{M}}(m-1) - V_{\mathcal{M}}(m-2) > V_{\mathcal{M}}(m) - V_{\mathcal{M}}(m-1)$ for some $m \le M$. Then

$$\begin{aligned}
&\lambda_b d_b(f_{m-3}+w)[V_{\mathcal{M}}(m-2) - V_{\mathcal{M}}(m-3) - w] \\
=& -rV_{\mathcal{M}}(m-2) + \lambda_s d_s(f_{m-1})[V_{\mathcal{M}}(m-1) - V_{\mathcal{M}}(m-2)] \\
>& -rV_{\mathcal{M}}(m-1) + \lambda_s d_s(f_m)[V_{\mathcal{M}}(m) - V_{\mathcal{M}}(m-1)] \\
=& \lambda_b d_b(f_{m-2}+w)[V_{\mathcal{M}}(m-1) - V_{\mathcal{M}}(m-2) - w]
\end{aligned},$$

where we have used $\lambda_s d_s(f_{m-1}) \ge \lambda_s d_s(f_m)$ and $V_{\mathcal{M}}(m-1) - V_{\mathcal{M}}(m-2) > V_{\mathcal{M}}(m) - V_{\mathcal{M}}(m-1)$ in the inequality and the equalities are due to (9). Similarly, as $d_b(f_{m-2}+w) > d_b(f_{m-3}+w)$, we have $V_{\mathcal{M}}(m-2) - V_{\mathcal{M}}(m-3) > V_{\mathcal{M}}(m-1) - V_{\mathcal{M}}(m-2)$. This completes the proof.

(3) $V_{\mathcal{M}}(m) - V_{\mathcal{M}}(m-1) \le w$. As f_m is decreasing in m, the arrival rate of the sellers is bounded by $\lambda_s d_s(f_1)$, then we must have that for any m, $V_{\mathcal{M}}(m) \le \frac{\lambda_s d_s(f_1)w}{r}$. We then prove the claim by contradiction.

Suppose $V_{\mathcal{M}}(m) - V_{\mathcal{M}}(m-1) > w$ for some m. Then due to the concavity of $V_{\mathcal{M}}(m)$, we have $V_{\mathcal{M}}(1) - V_{\mathcal{M}}(0) > w$. As $(r + \lambda_s d_s(f_1))V_{\mathcal{M}}(0) = \lambda_s d_s(f_1)V_{\mathcal{M}}(1)$ from (9), then $(r + \lambda_s d_s(f_1))V_{\mathcal{M}}(0) > \lambda_s d_s(f_1)[V_{\mathcal{M}}(0) + w]$. That is $V_{\mathcal{M}}(0) > \frac{\lambda_s d_s(f_1)w}{r}$. This leads to a contradiction. □

EC.3.2. Proofs for Section 2.2.2

Proof of Proposition 4: For part (1), from (8) , we have

$$d_s(f^{(n)}_{M^*+2})(f^{(n)}_{M^*+1} - f^{(n)}_{M^*+2}) + \frac{\kappa}{n\lambda_s} = \frac{\lambda_b}{\lambda_s} d_b(f^{(n)}_{M^*} + w)(f^{(n)}_{M^*} - f^{(n)}_{M^*+1}) > \frac{\lambda_b}{\lambda_s} d_b(f^{(n)}_{M^*} + w)(f^{(n)}_{M^*+1} - f^{(n)}_{M^*+2}).$$

Arranging the terms of the two sides of the inequality, we have $\lim_{n\to\infty}(\frac{\lambda_b}{\lambda_s} d_b(f^{(n)}_{M^*} + w) - d_s(f^{(n)}_{M^*+2}))((f^{(n)}_{M^*+1} - f^{(n)}_{M^*+2})) \le 0$. We then prove $\lim_{n\to\infty}(f^{(n)}_{M^*+1} - f^{(n)}_{M^*+2}) = 0$ by contradiction. If for a subsequence $n_k \to \infty$, $\lim_{k\to\infty}(f^{(n_k)}_{M^*+1} - f^{(n_k)}_{M^*+2}) > 0$, to make the equation above hold, we have $\lim_{k\to\infty}\left(\frac{\lambda_b}{\lambda_s}d_b(f_{M^*}^{(n_k)}+w)-d_s(f_{M^*+2}^{(n_k)})\right)\le 0$. However, from the definition of M^*, we have $\frac{\lambda_b}{\lambda_s}d_b(f_{M^*}^{(n_k)}+w)\ge d_s(f_{M^*+1}^{(n_k)})\ge d_s(f_{M^*+2}^{(n_k)})$. Then we must have $\lim_{k\to\infty}(\frac{\lambda_b}{\lambda_s}d_b(f_{M^*}^{(n_k)}+w)-d_s(f_{M^*+2}^{(n_k)}))=0$. As a result, $\lim_{k\to\infty}(d_s(f_{M^*+1}^{(n_k)})-d_s(f_{M^*+2}^{(n_k)}))=0$ and thus $\lim_{k\to\infty}(f_{M^*+1}^{(n_k)}-f_{M^*+2}^{(n_k)})=0$, leading to contradiction. Therefore, $\lim_{n\to\infty}(f_{M^*+1}^{(n)}-f_{M^*+2}^{(n)})=0$. Then using (8), by taking the limit of $n\to\infty$, the waiting cost is negligible and one can easily prove that for any finite k, $\lim_{n\to\infty}[f_{M^*\pm k}^{(n)}-f_{M^*}^{(n)}]=0$. As $f_{M^*+1}^{(n)}\le\hat{f}$ and $f_{M^*-1}^{(n)}\ge\hat{f}$, $\lim_{n\to\infty}f_{M^*\pm k}^{(n)}=\hat{f}$.

For part (2), we need several steps to complete the proof.

1) $\lim_{n\to\infty}M^{*(n)}=+\infty$. From the first equation in (8), we have $\lambda_s d_s(f_2^{(n)})(f_1^{(n)}-f_2^{(n)})+\frac{\kappa}{n}=\max_{f_0}\{\lambda_b d_b(f_0+w)(f_0-f_1^{(n)})\}$. If the claim does not hold, then from part (1), we have $\lim_{n\to\infty}f_1^{(n)}=\lim_{n\to\infty}f_2^{(n)}=\hat{f}$ (for a subsequence of n with a slight abuse of notation), then the left-hand side of the equation above is 0, whereas the right-hand side is positive, which leads to a contradiction.

2) For any positive ϵ, $\lim_{n\to\infty}\pi_{n_1}=\lim_{n\to\infty}\pi_{n_2}=0$, where $n_1=\max\{m:f_m^{(n)}>\hat{f}+\epsilon\}$ and $n_2=\min\{m:f_m^{(n)}<\hat{f}-\epsilon\}$. Recall that π_m is the steady-state distribution of state m, which depends on n as well. Clearly, $n_1<M^*$. Since $\lim_{n\to\infty}f_{M^*\pm k}^{(n)}=\hat{f}$ for finite k, $\lim_{n\to\infty}(M^*-n_1)=\infty$. Besides, for any $m\in[n_1,M^*]$, we know $\pi_m\ge\pi_{n_1}$, because of the definition of π_m in (10), the monotonicity of f_m and the definition of M^*. Then, we have $1\ge\lim_{n\to\infty}\sum_{m=n_1}^{M^*}\pi_m\ge\lim_{n\to\infty}\sum_{m=n_1}^{M^*}\pi_{n_1}=\lim_{n\to\infty}(M^*-n_1)\pi_{n_1}$. As $\lim_{n\to\infty}(M^*-n_1)=\infty$, $\lim_{n\to\infty}\pi_{n_1}=0$. Similarly, $\lim_{n\to\infty}\pi_{n_2}=0$.

3) $\lim_{n\to\infty}\sum_{m:|f_m^{(n)}-\hat{f}|>\epsilon}\pi_m=0$. For $m\in[0,n_1]$, because of (10) and $f_m\le f_{m-1}$, we have $\pi_{m-1}<a\pi_m$, where $a_n\le\frac{\lambda_b d_b(f_{n_1-1}^{(n)}+w)}{\lambda_s d_s(f_{n_1+1}^{(n)})}$. As a result, we have $\lim_{n\to\infty}\sum_{m:f_m^{(n)}>\hat{f}+\epsilon}\pi_m\le\lim_{n\to\infty}\sum_{n=0}^{\infty}a_n^n\pi_{n_1}=\lim_{n\to\infty}\frac{1}{1-a_n}\pi_{n_1}$. As $\lim_{n\to\infty}\frac{\lambda_b d_b(f_{n_1-1}^{(n)}+w)}{\lambda_s d_s(f_{n_1}^{(n)})}<\lim_{n\to\infty}\frac{\lambda_b d_b(f_{M^*}^{(n)}+w)}{\lambda_s d_s(\hat{f}+\epsilon)}=\frac{\lambda_b d_b(\hat{f}+w)}{\lambda_s d_s(\hat{f}+\epsilon)}<\frac{\lambda_b d_b(\hat{f}+w)}{\lambda_s d_s(\hat{f})}=1$, we have $a_n<1-\delta$ being bounded away from 1 and thus $\lim_{n\to\infty}\sum_{m:f_m^{(n)}>\hat{f}+\epsilon}\pi_m=0$ because of step 2) above. Similarly, $\lim_{n\to\infty}\sum_{m:f_m^{(n)}<\hat{f}-\epsilon}\pi_m=0$. Thus we complete the proof. □

Proposition EC.2. *Given $w=c_s+c_b$, in the thick market, the firm's long-run average revenue converges:* $\lim_{n\to\infty}\frac{V_{\mathcal{M}}^{(n)}}{n}=\frac{\lambda_s d_s(\hat{f})w}{r}$.

Proof of Proposition EC.2: For any positive ϵ, define $n_1=\max\{m:f_m^{(n)}>\hat{f}+\epsilon\}$ and $n_2=\min\{m:f_m^{(n)}<\hat{f}-\epsilon\}$. Then

$$\begin{aligned}\frac{rV_{\mathcal{M}}^{(n)}}{nw}&=\sum_{m=1}^{M^{(n)}}\pi_m\lambda_b d_b(f_{m-1}^{(n)}+w)\\&=\sum_{m=1}^{n_1}\pi_m\lambda_b d_b(f_{m-1}^{(n)}+w)+\sum_{m=n_1+1}^{n_2-1}\pi_m\lambda_b d_b(f_{m-1}^{(n)}+w)+\sum_{m=n_2}^{M^{(n)}}\pi_m\lambda_b d_b(f_{m-1}^{(n)}+w).\end{aligned}$$

For the first term,

$$\lim_{n\to\infty}\sum_{m=1}^{n_1}\pi_m\lambda_b d_b(f_{m-1}^{(n)}+w)\le\lim_{n\to\infty}\sum_{m=1}^{n_1}\pi_m\lambda_b d_b(f_{n_1-1}^{(n)}+w)\le\lambda_b d_b(\hat{f}+\epsilon+w)\lim_{n\to\infty}\sum_{m=1}^{n_1}\pi_m=0,$$

where we have used $\lim_{n\to\infty}\sum_{m=1}^{n_1}\pi_m = 0$ from Proposition 4. Similarly, the last term is also zero when $n\to\infty$. For the second term, as for any $m\in[n_1+1,n_2-1]$, $d_b(\hat f+\epsilon+w)\le d_b(f_{m-1}^{(n)}+w)\le d_b(\hat f-\epsilon+w)$, we have

$$\lambda_b d_b(\hat f+\epsilon+w)\lim_{n\to\infty}\sum_{m=n_1+1}^{n_2-1}\pi_m \le \lim_{n\to\infty}\sum_{m=n_1+1}^{n_2-1}\pi_m\lambda_b d_b(f_{m-1}^{(n)}+w)\le \lambda_b d_b(\hat f-\epsilon+w)\lim_{n\to\infty}\sum_{m=n_1+1}^{n_2-1}\pi_m.$$

From Proposition 4, $\lim_{n\to\infty}\sum_{m=n_1+1}^{n_2-1}\pi_m = 1$. Therefore,

$$\lambda_b d_b(\hat f+\epsilon+w)\le \lim_{n\to\infty}\frac{rV_{\mathcal M}^{(n)}}{nw}\le \lambda_b d_b(\hat f-\epsilon+w).$$

Since it holds for an arbitrary $\epsilon>0$, we have completed the proof. □

EC.4. Proofs for Section 3

Proof of Theorem 2: By Lemma 1, we have that the maximal number of sellers can be bounded, i.e., $M\le\frac{2\bar p\lambda_b}{\kappa}$. When the market size of buyers is relatively small ($\lambda_b<\frac{\kappa}{2\bar p}$) or the waiting cost κ is relatively high ($\kappa>2\bar p\lambda_b$), the cost of selling a product on the marketplace is too large such that no seller will join the marketplace, i.e., $M=0$ and $V_{\mathcal M}=0$. In contrast, when $d_s(p)>0$ for any $p>0$, the dealer can always achieve a positive expected profit, $V_{\mathcal D}(0)>0$, by setting

$$p_b(0)=\infty,\ p_b(x)=\eta\quad\forall\, x\ge 1,\ p_s(0)=e^{-\frac{r}{\lambda_b d_b(\eta)}}\eta-\epsilon,\ p_s(x)=0\quad\forall\, x\ge 1,$$

where η is a positive value such that $d_b(\eta)>0$ and $d_s(e^{-\frac{r}{\lambda_b d_b(\eta)}}\eta)>0$, and ϵ is a small positive constant close to 0. In this case, the dealer will hold at most one inventory. Then, we analyze the sign of the expected profit of each inventory as follows: $\mathbb E[e^{-r\tau}p_b(1)-p_s(0)]\ge e^{-\frac{r}{\lambda_b d_b(\eta)}}\eta-\left(e^{-\frac{r}{\lambda_b d_b(\eta)}}\eta-\epsilon\right)=\epsilon>0$, where τ is the random time the dealer holds the inventory, and the inequality is due to Jensen's inequality. Therefore, the dealer earns more profits than the platform in these cases, i.e., $V_{\mathcal D}(0)>0=V_{\mathcal M}$.

By (8), as κ decreases, the resulting $\boldsymbol f$ is the same as that after scaling up λ_b and λ_s proportionally. When $\kappa\to 0$, by a similar proof to Proposition 4 and Proposition EC.2 in the online supplement, we have $\lim_{\kappa\to 0}\sum_{m:|f_m-\hat f|>\epsilon}\pi_m=0$ and hence $\lim_{\kappa\to 0}V_{\mathcal M}=\frac{\lambda_s d_s(\hat f)w}{r}$, which is the optimal value of problem (12) after optimizing w. From Proposition EC.1 in the online supplement, $V_{\mathcal D}(0)$ is upper bounded by $\bar V_{\mathcal D}(0)$, the optimal value of problem (6). In Theorem 3, we will show that the optimal values in problems (6) and (12) are equivalent. Therefore, $V_{\mathcal D}(0)<\lim_{\kappa\to 0}V_{\mathcal M}^*$. That is, the platform earns more profits than the dealer. □

Proof of Theorem 3: The convergences to the static optimization problems (6) and (12) are established in Section 2.1 and 2.2. Therefore, we only need to compare (6) and (12) in the theorem. It is easy to see that the decision variables of (6), p_b and p_s, can be transformed to those of (12),

$c_s + c_b$ by the following equations: $\hat{f} = p_s, \quad \hat{f} + c_s + c_b = p_b$. There is a one-to-one correspondence between (6) and (12). Part (1) and part (2) immediately follow from this observation. Part (3) and part (4) follow from the additional observation that the average utility of sellers is $\hat{f}$ and the average utility of buyers is $\hat{f} + w$, from Proposition 4. □

EC.5. Case Studies and Economic Insights

We shall present some case studies to illustrate economic insights.[14]

Second-hand Luxury Goods Market. According to a market survey,[15] the market share of second-hand luxury goods is well above 20% in developed countries such as Japan, U.S., and France. In China, the number is only 2% but growing rapidly in the last few years. The different stages of the market also lead to different business models. In developed countries, consumers have seen the coexistence of dealership and marketplace. Designer Exchange, Resellfridges, Cocoon, BrandOff and Komehyo operate as dealers: They buy high-end luxury goods from consumers and resell it, controlling the price and inventory on their own. ThredUp, Lampoo, and Luxury Closet operate as marketplaces: They allow customers to list their used luxury products on the websites and to be directly matched with buyers. In contrast, in the Chinese market, the dealership model is dominant, including MilanStation and PonhuLuxury.

This is consistent with our theory: Theorem 2 predicts that when the market size is small, especially that of buyers, then the dealership model is more profitable, as is the case in China. Theorem 3 predicts that when the market is thick, then both models are equivalent and may coexist, as is the case in developed countries. In fact, the CEO of PonhuLuxury pointed out explicitly in an interview[16] that in the current Chinese market, a marketplace is hard to survive because of the difficulty to attract customers and make the market thick.

Used-Car Market. The used-car market is a major component of the automobile industry. It is estimated that more than 42 million used cars were sold in 2014 (Manheim 2015), and 2/3 of them were through dealers. The marketplace of used cars is growing, but eclipsed by the market share of dealership. Besides the concern for quality and information asymmetry, which is not captured by our model, Theorem 2 explains the dominance of dealership in the used-car market from the angle

[14] Our theoretical results are also consistent with empirical findings in the finance literature, although the boundary between dealership and marketplace in financial markets is blurred and the limit order books in financial markets are in general two-sided. Mayhew (2002) analyzes the data of equity option prices from 1986 to 1997, and find that dealers such as designated primary market maker works in thin markets, while traditional open outcry platform works well for liquid markets Mayhew (2002) show that for high-volume options the difference between the two market structures becomes slighter. This is consistent with our finding (Theorem 3), as for high-volume options, the market is thick and the two models are fundamentally the same.

[15] http://baogao.chinabaogao.com/baihuo/392626392626.html

[16] https://36kr.com/p/5112915

of waiting cost. The unit prices of cars are usually substantially higher than other second-hand products. This can be translated to a higher waiting cost for sellers. From Theorem 2, the dealership model is more profitable when consumers are impatient. For instance, haoche51, a Chinese online platform for used cars, started in 2014 as a marketplace to match sellers and buyers directly, and later transformed into a model that is more similar to dealership in 2016. According to an interview[17] of the CEO, the marketplace model fails because "the C2C model is not efficient ... it takes a long time for the buyers and sellers to find a match". Their new dealership is believed to better accommodate the needs of consumers.

P2P Lending Markets. In money markets, we can view lenders as sellers and borrowers as buyers. Traditional banking attracts deposits and makes loans to individuals or small and medium enterprises (SMEs). This can be viewed as a dealership model, as a bank manages its own "inventory" and makes profits from the spread. Over the past decade, technology has revolutionized the financial-service sectors. Thanks to the emergence of efficient online communication technology and data and tools available for credit modeling, P2P lending platforms, in which individual borrowers (or owners of SMEs) are matched with individual investors, have become popular since 2005, when the UK-based Zopa opened its doors. The industry has grown to $3.5bn, and continues to gain attraction.[18] These platforms are designed as a marketplace model. For example, LendingClub,[19] a US-based P2P lending company, allows borrowers to post their loan needs on the website with a target rate they would pay, and then individual lenders can browse the loan listings and select one to invest in; thus, the roles of buyers/sellers are flipped, as the buyers/borrowers are placing orders, as opposed to sellers on traditional platforms. The recent emergence of this business model is consistent with the numerical study in Table 3, in which we show that the marketplace model becomes relatively profitable for a larger market size. Indeed, the rapid increase of the consumer base is crucial for P2P lending platforms, which is only made possible by recent technology. Interestingly, when the consumer base becomes saturated, we find some companies turning to a hybrid model, consistent with our thick market analysis. Indeed, after eight years since its foundation, LendingClub began partnering with banks to enable them to purchase loans directly through the LendingClub platform or offer LendingClub products to their customers.[20]

[17] https://m.qctt.cn/news/120581

[18] https://igniteoutsourcing.com/fintech/peer-to-peer-lending-industry-trends/

[19] LendingClub is the largest P2P lending platform with a 46% market share in 2014 according to https://www.statista.com/statistics/468469/market-share-of-lending-companies-by-loans/. The second largest has 12%. LendingClub is largely a market monopoly.

[20] https://www.lendingclub.com/investing/institutional/banks