Electronic Companion — "Cloud Pricing: The Spot Market Strikes Back" by Ludwig Dierks and Sven Seuken, *Management Science*.

Electronic Companion

EC.1. Additional Statements and Proofs

EC.1.1. Lemma EC.1

A number of our results make use of the following technical Lemma which extends Lemma 7 from Abhishek et al. (2017) to more general g_i and $w()$, which includes our model.

LEMMA EC.1. *Fix a provider strategy ρ with $l_s > 0$. Let $(x_1, \dots, x_k)$ be a weakly decreasing sequence. For $i > k$ let $g_i(x_i), \dots, g_n(x_n)$ be given such that $g_j(x_j)$ is a weakly increasing and semi-differentiable scalar function with left derivative at most $w(\mathcal{S}, x_i, \rho, (x_1, \dots, x_k, x_i \dots, x_i, 0_{i+1}, \dots, 0_n))$. (Here, the notation $(x_1, \dots, x_k, x_i \dots, x_i, 0_{i+1}, \dots, 0_n)$ means that all entries k' with $k < k' \leq i$ are set to x_i.) Further, assume that for all $j \geq i$ it holds that $g_j(\mu v_j) \leq v_j$, as well as $g_j(x) \leq g_i(x)$ for all $x \in \mathbb{R}$. Then there exist unique $x_i, \dots, x_n$ such that for any strategy profile*

$$\sigma' = \overrightarrow{x} = (x_1, \dots, x_k, x_i, \dots, x_i, x_{i+1}, \dots, x_n) \tag{EC.1}$$

where any user of class j joins the spot market if and only if his waiting cost c is below x_j, it holds that

$$\int_0^{x_j} w(\mathcal{S}, c, \rho, \overrightarrow{x}) dc = g_j(x) \quad \text{for all } j \geq i. \tag{EC.2}$$

Proof To see this, assume that the claim holds for $i+1$. Then for any $z \in [0, x_k]$ there exists

$$\sigma(z) = \overrightarrow{x}(z) = (x_1, \dots, x_k, z, \dots, z, x_{i+1}(z), \dots, x_n(z)) \tag{EC.3}$$

satisfying Equation (EC.2) for any $j \geq i+1$. We now show that there exists a unique z^* such that $w(z^*) = \int_0^{z^*} w(\mathcal{S}, c, \rho, \sigma(z^*)) dz$. As a first step, we show that $w(z) = \int_0^z w(\mathcal{S}, c, \rho, \sigma(z)) dz$ is increasing in z. Since for any fixed x, $\int_0^x w(\mathcal{S}, c, \rho, \overrightarrow{x}) dc$ is increasing in each x_j, it follows that $x_{i+1}(z)$ is decreasing in z. Then for any $\hat{z} > z$ it holds by the induction assumption that

$$\int_0^{x_{i+1}(z)} w(\mathcal{S}, c, \rho, \overrightarrow{x}(\hat{z})) dc \geq g_i(x_{i+1}(z)) \tag{EC.4}$$

and therefore

$$\frac{\int_0^{\hat{z}} w(\mathcal{S}, c, \rho, \overrightarrow{x}(\hat{z})) dc - \int_0^z w(\mathcal{S}, c, \rho, \overrightarrow{x}(z)) dc}{\hat{z} - z} \tag{EC.5}$$

$$\geq \frac{g_{i+1}(x_{i+1}(z)) - g_{i+1}(x_{i+1}(z))}{\hat{z} - z} + \frac{\int_{x_{i+1}(z)}^{\hat{z}} w(\mathcal{S}, c, \rho, \overrightarrow{x}(\hat{z})) dc - \int_{x_{i+1}(z)}^{z} w(\mathcal{S}, c, \rho, \overrightarrow{x}(z)) dc}{\hat{z} - z} \tag{EC.6}$$

$$= \frac{\int_z^{\hat{z}} w(\mathcal{S}, c, \rho, \overrightarrow{x}(\hat{z})) dc + \int_{x_{i+1}(z)}^{z} w(\mathcal{S}, c, \rho, \overrightarrow{x}(\hat{z})) - w(\mathcal{S}, c, \rho, \overrightarrow{x}(z)) dc}{\hat{z} - z} \tag{EC.7}$$

$$> \frac{\int_z^{\hat{z}} w(\mathcal{S}, c, \rho, \overrightarrow{x}(\hat{z})) dc}{\hat{z} - z} \tag{EC.8}$$

$$> w(\mathcal{S}, z, \rho, \overrightarrow{x}(\hat{z})). \tag{EC.9}$$

Equation (EC.8) is a direct result of the fact that waiting times of higher priority jobs do not depend on the number of lower priority jobs in the queue. By taking the limit $\hat{z} \to z$ it follows that $w'(z) > w(\mathcal{S}, z, \rho, \overrightarrow{x}(\hat{z})) \geq w(\mathcal{S}, z, \rho, (x_1, \ldots, x_k, z \ldots, z, 0_{i+1}, \ldots, 0_n)$. As $w(0) = 0$ and $w(\mu v_i) \geq v_i \geq g_i(\mu v_i)$, the claim for i follows.

To show the induction base case $i = n$, we introduce a dummy variable x_{n+1} with $g_{n+1} = 0$. This means that for any z it trivially holds that $x_{n+1}(z) = 0$ and the statement for n follows. □

EC.1.2. Proof of Proposition 3

Any equilibrium strategy for users of class i is trivially a threshold strategy because for any fixed strategy profile σ, the payoff of a user of class i, i.e., $\pi_i^c(\mathcal{S}, c, \rho, \sigma)$, is monotone decreasing in the waiting cost c. By setting $g_i(c) = v_i$, the existence of the unique cutoff vector $\overrightarrow{c}^S$ then follows directly from Lemma EC.1. By the incentive compatibility of the payment rule, no user would deviate from $\sigma^* = \overrightarrow{c}^S$. □

EC.1.3. Proof of Proposition 4

Users with waiting close enough to zero always prefer the spot market, no matter how much time they lose compared to the fixed-price market. Since some users join the fixed-price market and both waiting time and payment are continuous in the bid c, for any potential equilibrium strategy profile σ^*, there has to be a lowest point c_1^P such that

$$c_1^P(T + \frac{1}{\mu}) + p_F \frac{1}{\mu} = \int_0^{c_1^P} w(\mathcal{S}, x, \rho, \sigma) dx. \tag{EC.10}$$

Since it holds $\frac{d}{dc}\pi_i^c(\mathcal{S}, c, \rho, \sigma) = w(\mathcal{S}, c, \rho, \sigma) \geq \frac{1}{\mu}\frac{1}{1-\tau\psi_E(l_S)} > T + \frac{1}{\mu} = \frac{d}{dc}\pi_i^c(\mathcal{F}, c, \rho, \sigma)$, the higher a users waiting cost, the worse the spot market compared to the fixed-price market and there cannot exist any $c > c_1^P$ for which users prefer the spot market, i.e. with

$$c(T + \frac{1}{\mu}) + p_F \frac{1}{\mu} > \int_0^{c} w(\mathcal{S}, x, \rho, \sigma) dx. \tag{EC.11}$$

Thus, no user with waiting cost greater than c_1^P joins the spot market. This means that the spot market can be fully defined by the actions of players with $c \leq c_1^P$. Recall that $\overrightarrow{c}^P$ denotes the vector of cutoff points at which a job becomes indifferent between the spot market and either the fixed-price market or balking. It holds that

$$\min\left\{\frac{p_F}{\mu} + c_i^P(\frac{1}{\mu} + T), v_i\right\} = \int_0^{c_i^P} w(\mathcal{S}, x, \rho, \overrightarrow{c}^P) dx \quad \forall i \in \{1, \ldots, n\}, \tag{EC.12}$$

which has a unique solution by Lemma EC.1. Every job of class i with $c < c_i^P$ joins the spot market, and every job with $c_i^P < c < \frac{\mu v_i - p_F}{\mu T + 1}$ joins the fixed-price market and those with $c > \frac{\mu v_i - p_F}{\mu T + 1}$ balk. Setting $c_i^B = max(c_i^P, \frac{\mu v_i - p_F}{\mu T + 1})$, it is clear that every solution of (17) and (18) solves (EC.12) and vice-versa. □

EC.1.4. Lemma EC.2

The following Lemma establishes the broad equilibrium structure when the spot market is faster at the highest bids and shows the existence of two cutoff points in equilibrium between which almost all users join the fixed-price market. It is used in the proof of Proposition 5.

LEMMA EC.2. *For any provider strategy $\rho = (p_F, l_S)$, in any BNE of the user game where some users join the fixed-price market and where the spot market is faster for the highest bids, i.e., $T > \frac{1}{\mu}\left(\frac{1}{1-\tau\psi_E(l_S)} - 1\right)$, there exists an interval $[c^L, c^U]$, such that almost all users (i.e., all besides possibly a set of measure zero that does not influence system dynamics) with waiting costs $c \in [c^L, c^U]$ join the fixed-price market or balk. For bids $c \in [c^L, c^U]$, the total waiting time and the payment in equilibrium are equal in each market, i.e.:*

$$m(\mathcal{S}, c, \rho, \sigma^*) = m(\mathcal{F}, c, \rho, \sigma^*) = p_F \frac{1}{\mu} \tag{EC.13}$$

$$w(\mathcal{S}, c, \rho, \sigma^*) = w(\mathcal{F}, c, \rho, \sigma^*) = T + \frac{1}{\mu} \tag{EC.14}$$

For waiting costs $c \notin [c^L,\ c^U]$ it holds that

$$\pi_i(\mathcal{S}, c, \rho, \sigma^*) > \pi_i(\mathcal{F}, c, \rho, \sigma^*)\ \ \forall i \in \{1, \ldots, n\}\,, \tag{EC.15}$$

and these users join the spot market or balk.

Proof For a job with the highest bid that does not balk, the spot market is faster than the fixed-price market because $T > \frac{1}{\mu}\left(\frac{1}{1-\tau\psi_E(l_S)} - 1\right)$. Any user with such a waiting cost is therefore willing to pay more in the spot market than in the fixed-price market. This means that he strictly prefers the spot market in equilibrium.

Let c^L be the lowest waiting cost for which a job prefers the fixed-price market over the spot market or is indifferent between the two, and let c^U be the highest such waiting cost. c^L and c^U have to exist for any equilibrium where users join both markets. We now show by contradiction that the user's payment in the spot market has to be weakly larger than in the fixed-price market for bids above c^L and that the spot market is weakly slower than the fixed-price market for bids below c^U.

Assume there exists a waiting cost $\bar{c} > c^L$ at which a user would prefer the spot market or be indifferent between spot and fixed-price market, and for which the payment in the spot market is less in expectation than in the fixed-price market, i.e., for which $m(\mathcal{S}, \bar{c}, \rho, \sigma) < m(\mathcal{F}, \bar{c}, \rho, \sigma)$. Then

$$c^L w(\mathcal{S}, c^L, \rho, \sigma) + m(\mathcal{S}, c^L, \rho, \sigma) \tag{EC.16}$$

$$\leq c^L w(\mathcal{S}, \bar{c}, \rho, \sigma) + m(\mathcal{S}, \bar{c}, \rho, \sigma) \tag{EC.17}$$

$$= \frac{c^L}{\bar{c}} \left(\bar{c} w(\mathcal{S}, \bar{c}, \rho, \sigma) + \frac{\bar{c}}{c^L} m(\mathcal{S}, \bar{c}, \rho, \sigma) \right) \tag{EC.18}$$

$$\leq \frac{c^L}{\bar{c}} \left(\bar{c} w(\mathcal{F}, \bar{c}, \rho, \sigma) + m(\mathcal{F}, \bar{c}, \rho, \sigma) + (\frac{\bar{c}}{c^L} - 1) m(\mathcal{S}, \bar{c}, \rho, \sigma) \right) \tag{EC.19}$$

$$< \frac{c^L}{\bar{c}} \left(\bar{c} w(\mathcal{F}, \bar{c}, \rho, \sigma) + m(\mathcal{F}, \bar{c}, \rho, \sigma) + (\frac{\bar{c}}{c^L} - 1) m(\mathcal{F}, \bar{c}, \rho, \sigma) \right) \tag{EC.20}$$

$$= \frac{c^L}{\bar{c}} \left(\bar{c} w(\mathcal{F}, \bar{c}, \rho, \sigma) + \frac{\bar{c}}{c^L} m(\mathcal{F}, \bar{c}, \rho, \sigma) \right) \tag{EC.21}$$

$$= c^L w(\mathcal{F}, \bar{c}, \rho, \sigma) + m(\mathcal{F}, \bar{c}, \rho, \sigma) \tag{EC.22}$$

$$= c^L w(\mathcal{F}, c^L, \rho, \sigma) + m(\mathcal{F}, c^L, \rho, \sigma) \tag{EC.23}$$

(EC.17) holds because the pricing rule is BNIC; (EC.19) holds because at waiting cost $\bar{c}$ the spot market's overall cost has to be lower than the fixed-price market in order for the user to join it. Finally, (EC.20) holds because we assumed the spot market to be cheaper with bid $\bar{c} > c^L$. A job with waiting cost c^L would therefore also strictly prefer the spot market, a contradiction.

Assume there exists a waiting cost $\bar{c} < c^U$ at which a user would prefer the spot market or be indifferent between spot and fixed-price market, and for which the waiting time is lower in the spot market than in the fixed-price market, i.e., for which $w(\mathcal{S}, \bar{c}, \rho, \sigma) < w(\mathcal{F}, \bar{c}, \rho, \sigma)$. Then similarly

$$c^U w(\mathcal{S}, c^U, \rho, \sigma) + m(\mathcal{S}, c^U, \rho, \sigma) \tag{EC.24}$$

$$\leq c^U w(\mathcal{S}, \bar{c}, \rho, \sigma) + m(\mathcal{S}, \bar{c}, \rho, \sigma) \tag{EC.25}$$

$$= \bar{c} w(\mathcal{S}, \bar{c}, \rho, \sigma) + m(\mathcal{S}, \bar{c}, \rho, \sigma) + (c^U - \bar{c}) w(\mathcal{S}, \bar{c}, \rho, \sigma) \tag{EC.26}$$

$$< \bar{c} w(\mathcal{F}, \bar{c}, \rho, \sigma) + m(\mathcal{F}, \bar{c}, \rho, \sigma) + (c^U - \bar{c}) w(\mathcal{F}, \bar{c}, \rho, \sigma) \tag{EC.27}$$

$$= c^U w(\mathcal{F}, \bar{c}, \rho, \sigma) + m(\mathcal{F}, \bar{c}, \rho, \sigma) \tag{EC.28}$$

A user with c^U would therefore also strictly prefer the spot market, a contradiction.

Therefore, for all $c \in [c^L, c^U]$ the spot market can neither be faster nor cheaper than the fixed-price market. If any users with waiting cost $c \in [c^L, c^U]$ join the spot market they have to be indifferent between both markets. Thus, for any σ^* to be a BNE, this means that at most a set of measure zero of such users can join the spot market, and thus the total waiting time and payment stay constant over the whole interval. The statement of the lemma immediately follows. □

EC.1.5. Proof of Proposition 5

It follows from Lemma EC.2 that any equilibrium strategy profile is of the form $\sigma^* = (\overrightarrow{c}^L, \overrightarrow{c}^U, \overrightarrow{c}^H)$: Let the points c^L and c^U be as given by Lemma EC.2 and let $\overrightarrow{c}^H$ denote the waiting costs above which users of each class cannot obtain a positive payoff anymore and balk. Define the cutoff vectors $\overrightarrow{c}^L, \overrightarrow{c}^U$ as $c_i^L = \min\{c^L, c_i^H\}$ and $c_i^U = \min\{c^U, c_i^H\}$. Note that this implies that $c_1^L = c^L$ and $c_1^U = c^U$ (because at least some users from class 1 go to the portion of the spot market that is

faster than the fixed-price market). Then Equations (19), (20) and (21) immediately follow from Lemma EC.2:

1. Equation (19): The payoff at c^L has to be the same for joining the fixed-price or spot market.
2. Equation (20): The payoff at c^U also has to be the same in the fixed-price and spot market.
3. Equation (21): Users do not balk as long as their value for joining one of the markets is greater than 0.

We now show that this system of equations always has a unique solution using a constructive approach. For this, we first introduce some additional notation.

Given provider strategy ρ, we know that in order to satisfy Lemma EC.2, jobs that pay more in the spot market than in the fixed-price marked need to arrive at a rate such that jobs with waiting cost c^L have to queue for exactly T. Denote this arrival rate by $\lambda(T,\rho)$. We now further overload our previous notation for a user strategy profile: for any vector $\hat{c}=(\hat{c}_1,\dots,\hat{c}_n)$ with $\hat{c}_i \geq \hat{c}_j$ for $i<j$, we let $\sigma=(\hat{c},\lambda(T,\rho))$ denote a user strategy profile where every user of class i with waiting cost $c<\hat{c}_i$ joins the spot market, but everyone else balks even if he could obtain a positive payoff in one of the markets. Additionally, we assume that *dummy jobs* of maximal priority arrive with rate $\lambda(T,\rho)$ into the spot market. Thus, $w(\mathcal{S},\hat{c}_1,\rho,(\hat{c},\lambda(T,\rho)))=T$ by definition. Combined with Lemma EC.2, this notational trick allows us to "simulate" the impact users with waiting costs between $\overrightarrow{c}^U$ and $\overrightarrow{c}^H$ have on all other users, without yet knowing $\overrightarrow{c}^U$ and $\overrightarrow{c}^H$. We now need to determine which classes of users join the fixed-price market (in the sense that there exists a c such that a user from that class with waiting cost c joins the fixed-price market) and which do not. Once we know that, we can split the system of equations into two parts that can be solved consecutively.

To check whether the k'th class joins the fixed-price market, i.e., whether $c_k^H > c^L$, we denote by $\hat{c}=(\hat{c}_1,\dots,\hat{c}_n)$ (where each $\hat{c}_i \in [0,\mu v_i]$) the cutoff vector solving the following:

$$0 = v_i - \int_0^{\hat{c}_i} w(\mathcal{S},x,\rho,(\hat{c},\lambda(T,\rho)))dx \quad \forall i \geq k \tag{EC.29}$$

$$\hat{c}_i = \hat{c}_k \quad \forall i < k \tag{EC.30}$$

This has a unique solution according to Lemma EC.1. Note that the cutoff vector $\hat{c}$ here carries an implicit dependence on k, while $\hat{c}_k$ denotes its k'th element.
If it now holds that

$$p_F\frac{1}{\mu} > m(\mathcal{S},\hat{c}_k,\rho,(\hat{c},\lambda(T,\rho))) \tag{EC.31}$$

$$= \int_0^{\hat{c}_k} w(\mathcal{S},x,\rho,(\hat{c},\lambda(T,\rho)))dx - w(\mathcal{S},\hat{c}_k,\rho,(\hat{c},\lambda(T,\rho))), \tag{EC.32}$$

then $c^L \leq \hat{c}_k$ would mean that not enough users join the spot market to reach the price of the fixed-price market at the cutoff point c^L. This means that for any such c^L, the system of equations defining the equilibrium cutoff vectors (i.e., Equations (19), (20) and (21)) cannot be satisfied for any choice of c^U and $\overrightarrow{c}^H$. It follows that in equilibrium, $c^L > \hat{c}_k$. Thus, in equilibrium, no user of class k joins the fixed-price market, i.e., $c_k^H < c^L$.

Conversely, if Equation (EC.31) does not hold, then setting $c^L > \hat{c}_k$ would mean that too many users join the spot market, and the payment in the spot market at the cutoff point c^L is larger than the payment in the fixed-price market. Thus, in any equilibrium, some users of class k join the fixed-price market and $c^L \leq \hat{c}_k$. As $m(\mathcal{S}, \hat{c}_k, \rho, (\hat{c}, \lambda(T, \rho)))$ is monotone decreasing in k, it follows that either there exists a lowest class k^* such that (EC.32) is satisfied and for which no user joins the fixed-price market, or all classes join the fixed-price market in which case we set $k^* = n+1$. Splitting the system of equations that defines the equilibrium strategy profile at this k^*, Lemma EC.1 yields that

$$0 = \hat{c}_i(T + \frac{1}{\mu}) + p_F \frac{1}{\mu} - \int_0^{\hat{c}_i} w(\mathcal{S}, x, \rho, (\hat{c}, \lambda(T, \rho)))dx \quad \forall i < k^* \tag{EC.33}$$

$$0 = v_i - \int_0^{\hat{c}_i} w(\mathcal{S}, x, \rho, (\hat{c}, \lambda(T, \rho)))dx \quad \forall i \geq k^* \tag{EC.34}$$

has a unique solution with $\hat{c}_{k^*} < \hat{c}_{k^*-1}$ and for any equilibrium $(\overrightarrow{c}^L, \overrightarrow{c}^U, \overrightarrow{c}^H)$ it holds that $c^L = \hat{c}_{k^*-1}$ and $c_i^H = \hat{c}_i$ for all $i \geq k^*$.

Given the solution to (EC.33) and (EC.34), we can now equivalently find the highest class k^{**} that joins the upper portion of the spot market (i.e. for which $c_{k^{**}}^U < c_{k^{**}}^H$). To this end, fix any $k < k^*$. Again carrying an implicit dependence on k, we define temporary cutoff vectors $\hat{c}^U$ and $\hat{c}^H$. Set $\hat{c}_i^H = c^L$ for all $k < i < k^*$ and $\hat{c}_i^H = \overrightarrow{c}_i^H$ for all $i \geq k^*$. Further let $\hat{c}^U$ and $\hat{c}_i^H$ for $i \leq k$ be given as the solution to

$$0 = v_{k+1} - \hat{c}_{k+1}^U(T + \frac{1}{\mu}) + p_F \frac{1}{\mu} \tag{EC.35}$$

$$0 = v_i - \int_0^{\hat{c}_i^H} w(\mathcal{S}, x, \rho, (\overrightarrow{c}^L, \hat{c}^U, \hat{c}^H))dx \quad \forall i \leq k \tag{EC.36}$$

$$\hat{c}_i^U = min(\hat{c}_k^U, \hat{c}_i^H) \quad \forall i \neq k+1. \tag{EC.37}$$

This system of equations has a unique solution for every $k < k^*$ according to Lemma EC.1. Intuitively, $(\overrightarrow{c}^L, \hat{c}^U, \hat{c}^H)$ can be seen as the strategy profile where users with waiting costs below c^L play the equilibrium strategy, no user joins the fixed-price market, and users with waiting costs above the point where class $k+1$ would obtain zero payoff in the fixed-price market join the spot market (if their payoff for doing so is positive). This means that under this strategy profile *more* users

join the spot market than would under any potential equilibrium strategy profile where $c^U > \hat{c}^U_{k+1}$. Analogous to k^*, there now exists a lowest class $k^{**} < k^*$, such that if only jobs of classes $i \leq k^{**}$ join the upper part of the spot market, there are still enough users that potentially (i.e., as long as the fixed-price market isn't better) join the spot market, such that the waiting time in the spot market at $\hat{c}^U_k$ is at least as high as the waiting time in the fixed-price market, i.e., k^{**} is the smallest k for which it holds that

$$T + \frac{1}{\mu} \leq w(\mathcal{S}, \hat{c}^U_k, \rho, (\overrightarrow{c}^L, \hat{c}^U, \hat{c}^H)). \tag{EC.38}$$

Conversely, if users of classes higher than k^{**} would join the upper portion of the spot market (i.e. $c^U \leq \hat{c}^U_{k^{**}+1}$) then the waiting time in the spot market at c^U is always above $T + \frac{1}{\mu}$. Consequently, we can calculate $\overrightarrow{c}^H_i$ for $k^{**} < i < k^*$ as the solution to

$$0 = v_i - \overrightarrow{c}^H_i (T + \frac{1}{\mu}) + p_F \frac{1}{\mu}. \tag{EC.39}$$

Then finally, we can calculate c^U and $\overrightarrow{c}^H_i$ for $i \leq k^{**}$ as the solution to

$$0 = c^U (T + \frac{1}{\mu}) + p_F \frac{1}{\mu} - \int_0^{c^U} w(\mathcal{S}, x, \rho, (\overrightarrow{c}^L, \overrightarrow{c}^U, \overrightarrow{c}^H)) dx \tag{EC.40}$$

$$0 = v_i - \int_0^{c^H_i} w(\mathcal{S}, x, \rho, (\overrightarrow{c}^L, \overrightarrow{c}^U, \overrightarrow{c}^H)) dx \quad \forall i \leq k^{**}, \tag{EC.41}$$

which, given c^L and c^H_i for all $i > k^{**}$ now also has a unique solution according to Lemma EC.1.

As each of the successively solved subsystems of equations was, at the time it was solved, independent of the then unsolved parts, $(\overrightarrow{c}^L, \overrightarrow{c}^U, \overrightarrow{c}^H)$ solves the whole system of equations.

EC.1.6. Proof of Lemma 1

For $l_S > 0$, all users with waiting costs in some neighborhood around zero prefer the spot market. Let ρ be a provider strategy with $l_S > 0$ that is proper. Assume we have an equilibrium where no one joins the fixed-price market, i.e., where the hybrid market degenerates to the spot market. A user of class 1 (i.e., the class with maximal value for completion) with waiting cost c^S_1 (if $\frac{1}{\mu}\left(\frac{1}{1-\tau\psi_E(l_S)} - 1\right) \geq T$) or c' (if $\frac{1}{\mu}\left(\frac{1}{1-\tau\psi_E(l_S)} - 1\right) < T$) could then obtain a better payoff by switching to the fixed-price market, leading to a contradiction. Any BNE therefore has some users joining the fixed-price market.

Now assume ρ is not proper. We first show the statement for $\frac{1}{\mu}\left(\frac{1}{1-\tau\psi_E(l_S)} - 1\right) \geq T$ (i.e., when the spot market is always slower than the fixed-price market). Proposition 4 gives us that any equilibrium where some users join the fixed-price market is of the form $\sigma = (\overrightarrow{c}^P, \overrightarrow{c}^B)$. At any waiting cost c for which users of some class i balk under $\sigma = (\overrightarrow{c}^S)$ but join the spot market under $\sigma = (\overrightarrow{c}^P, \overrightarrow{c}^B)$, their payoff in the spot market needs to be higher under $\sigma = (\overrightarrow{c}^P, \overrightarrow{c}^B)$, i.e.,

$$\pi^c_i(\mathcal{S}, c, \rho, (\overrightarrow{c}^P, \overrightarrow{c}^B)) \geq 0 \geq \pi^c_i(\mathcal{S}, c, \rho, (\overrightarrow{c}^S)). \tag{EC.42}$$

The payment and waiting time in the spot market are only *changing* in the bid c at bids for which any users go into the spot market. It directly follows that a user's payoff in the spot market (if he were to choose it) is (weakly) higher for *any* user (and thus also for users with waiting cost c_1^S) under $\sigma = (\overrightarrow{c}^S)$ than under $\sigma = (\overrightarrow{c}^P, \overrightarrow{c}^B)$. If ρ is not proper, it follows

$$\pi_1^{c_1^S}(\mathcal{S}, c_1^S, \rho, (\overrightarrow{c}^P, \overrightarrow{c}^B)) \geq \pi_1^{c_1^S}(\mathcal{S}, c_1^S, \rho, (\overrightarrow{c}^S)) > \pi_1^{c_1^S}(\mathcal{F}, c_1^S, \rho, (\overrightarrow{c}^S)) = \pi_1^{c_1^S}(\mathcal{F}, c_1^S, \rho, (\overrightarrow{c}^P, \overrightarrow{c}^B)), \tag{EC.43}$$

i.e., users of class 1 with waiting cost c_1^S would deviate from $\sigma = (\overrightarrow{c}^P, \overrightarrow{c}^B)$, contradicting that $\sigma^* = (\overrightarrow{c}^P, \overrightarrow{c}^B)$ is a BNE.

Now we show the statement for when ρ is not proper and when $\frac{1}{\mu}\left(\frac{1}{1-\tau\psi_E(l_S)} - 1\right) < T$ (i.e., when the spot market is faster for the highest bids). If no c' exists for which the expected waiting time in both queues is equal (i.e., for which condition (a) from Definition 1 holds), the spot market is trivially faster for every user (and consequently also cheaper) and the statement follows by the same argument as for $\frac{1}{\mu}\left(\frac{1}{1-\tau\psi_E(l_S)} - 1\right) \geq T$.

Now assume there exists a c' satisfying condition (a), but it does not satisfy condition (b), i.e.

$$c'(T + \frac{1}{\mu}) + p\frac{1}{\mu} \geq \int_0^{c'} w(\mathcal{S}, x, \rho, \sigma)dx \tag{EC.44}$$

for c' such that $T + \frac{1}{\mu} = w(\mathcal{S}, c', \rho, \sigma)$. It follows that

$$p_F\frac{1}{\mu} \geq \int_0^{c'} w(\mathcal{S}, x, \rho, \sigma)dx - c'w(\mathcal{S}, c', \rho, \sigma). \tag{EC.45}$$

This means that even with bid c', for which the fixed-price market has the same total waiting time as the spot market, joining the spot market is still cheaper. We now show that, in this case, there only exist BNEs where no user joins the fixed-price market. Since the payoff in the spot market for every user is monotone decreasing in the number of users that join, it is enough to show that when playing $\sigma = \overrightarrow{c}^S$, no user has an incentive to switch to the fixed-price market.

A user with waiting cost c' clearly has no reason to switch. Assume that a user of class i with waiting cost $c \neq c'$ would prefer to switch to the fixed-price market. Misreporting his class as c' and joining the spot market would then lead to a payoff of

$$\begin{aligned}
\pi_i^c(\mathcal{S}, c', \rho, \overrightarrow{c}^S) &= v_i - \int_0^{c'} w(\mathcal{S}, x, \rho, \sigma)dx + c'w(\mathcal{S}, c', \rho, \sigma) - cw(\mathcal{S}, c', \rho, \sigma) && \text{(EC.46)}\\
&= v_i - \int_0^{c'} w(\mathcal{S}, x, \rho, \sigma)dx + c'w(\mathcal{S}, c', \rho, \sigma) - c(T + \frac{1}{\mu}) && \text{(EC.47)}\\
&\geq v_i - p_F\frac{1}{\mu} - c(T + \frac{1}{\mu}) && \text{(EC.48)}\\
&= \pi_i^c(\mathcal{F}, c, \rho, \overrightarrow{c}^S) && \text{(EC.49)}\\
&> \pi_i^c(\mathcal{S}, c, \rho, \overrightarrow{c}^S) && \text{(EC.50)}
\end{aligned}$$

Misreporting in the spot market would therefore be beneficial over reporting truthfully, contradicting the pricing rule being Bayes-Nash incentive compatible. Consequently, no user prefers the fixed-price market and by Theorem 3 it holds that $\sigma^* = \overrightarrow{c}^S$. □

EC.1.7. Lemma EC.3

The proof of Theorem 2 requires the introduction of an additional technical Lemma. The following Lemma establishes that the average payments in the spot market approach the payments in the fixed-price market for small enough l_S.

LEMMA EC.3. *For any strategy $\sigma^* = (\overrightarrow{c}^P, \overrightarrow{c}^B)$ or $\sigma^* = (\overrightarrow{c}^L, \overrightarrow{c}^U, \overrightarrow{c}^H)$, denote by $\bar{m}(\mathcal{S}, \rho, \sigma^*)$ the average payment in the spot market, i.e., respectively*

$$\bar{m}(\mathcal{S}, \rho, (\overrightarrow{c}^P, \overrightarrow{c}^B)) := \frac{\sum_i \lambda_i \int_0^{c_i^P} m(\mathcal{S}, x, \rho, \sigma^*) f_i(x) dx}{\sum_i \lambda_i \int_0^{c_i^P} f_i(x) dx} \tag{EC.51}$$

and

$$\bar{m}(\mathcal{S}, \rho, (\overrightarrow{c}^L, \overrightarrow{c}^U, \overrightarrow{c}^H)) := \frac{\sum_i \lambda_i \left[\int_0^{c_i^L} m(\mathcal{S}, x, \rho, \sigma^*) f_i(x) dx + \int_{c_i^U}^{c_i^H} m(\mathcal{S}, x, \rho, \sigma^*) f_i(x) dx\right]}{\sum_i \lambda_i \left[\int_0^{c_i^L} f_i(x) dx + \int_{c_i^U}^{c_i^H} f_i(x) dx\right]}. \tag{EC.52}$$

For every setting, $p_F < \mu v_1$ and $\varepsilon > 0$, there exists a (possibly fractional) number of spot instances $l_S \leq l$ such that for $\rho = (p_F, l_S)$ it holds that $\bar{m}(\mathcal{S}, \rho, \sigma^*)$ is greater than the expected payment in the fixed-price market minus ε, i.e.

$$\bar{m}(\mathcal{S}, \rho, \sigma^*) \geq \frac{p_F}{\mu} - \varepsilon. \tag{EC.53}$$

Proof For any fixed p_F there exists some number of spot instances l' such that all provider strategies $\rho = (p_F, l_S)$ with $0 < l_S \leq l'$ result in an equilibrium that is either of the form $\sigma^* = (\overrightarrow{c}^P, \overrightarrow{c}^B)$ or of the form $\sigma^* = (\overrightarrow{c}^L, \overrightarrow{c}^U, \overrightarrow{c}^H)$. We first present the case where $T \leq \frac{1}{\mu}\left(\frac{1}{1-\tau\psi_E(l_S)} - 1\right)$ for all $0 < l_S \leq l'$ and therefore $\sigma^* = (\overrightarrow{c}^L, \overrightarrow{c}^U, \overrightarrow{c}^H)$. To keep the proof readable, we introduce new notation to draw all waiting costs in the spot market from a single distribution instead of first drawing a job's class and then its waiting cost. Note that this does not change the number of jobs or their bids nor their waiting costs in the market. For provider strategy ρ and equilibrium strategy profile $\sigma^* = (\overrightarrow{c}^L, \overrightarrow{c}^U, \overrightarrow{c}^H)$, we define the distribution

$$F_S(c) := \frac{\sum_i \lambda_i \left[\int_0^{\min\{c, c_i^L\}} f_i(x) dx + \int_{\min\{c, c_i^U\}}^{c} f_i(x) dx\right]}{\sum_i \lambda_i \left[\int_0^{c_i^L} f_i(x) dx + \int_{c_i^U}^{c_i^H} f_i(x) dx\right]} \tag{EC.54}$$

and the arrival rate

$$\lambda_S := \sum_i \lambda_i \left[\int_0^{c_i^L} f_i(x) dx + \int_{c_i^U}^{c_i^H} f_i(x) dx\right] \tag{EC.55}$$

with similarly constructed PDF $f_S(c)$. Now consider an artificial spot market with arrival rate λ_S, where every arriving job's waiting cost is drawn from F_S and everyone joins. From the provider's point of view, this market is the same as the normal spot market that would result from her playing ρ, including users on average having the same expected payments. To analyze the provider's profit from the spot market when playing ρ, we can thus instead analyze this artificial market.

The per-user-average profit $\bar{m}(\mathcal{S}, \rho, \sigma^*)$ of the artificial spot market is given by taking the expectation of the payment $m(\mathcal{S}, c, \rho, \sigma^*)$, where the expectation is taken over c drawn from the PDF $f_S(c)$:

$$\bar{m}(\mathcal{S}, \rho, \sigma^*) = \frac{\lambda_S \left[\int_0^{c_1^H} m(\mathcal{S}, x, \rho, \sigma^*) f_S(x) dx\right]}{\lambda_S} \tag{EC.56}$$

$$= \int_0^{c_1^H} m(\mathcal{S}, x, \rho, \sigma^*) f_S(x) dx \tag{EC.57}$$

$$= \int_{-\infty}^{\infty} m(\mathcal{S}, x, \rho, \sigma^*) f_S(x) dx \tag{EC.58}$$

$$= E_{c\sim f_S}[m(\mathcal{S}, c, \rho, \sigma^*)] \tag{EC.59}$$

Now, for any l_S and any $0 < \xi < 1$ define $c_\xi^{l_S}$ as the waiting cost with $F_S(c_\xi^{l_S}) = \xi$. It then follows by Markov's inequality that

$$\bar{m}(\mathcal{S}, \rho, \sigma^*) \geq (1 - F_S(c_\xi^{l_S})) m(\mathcal{S}, c_\xi^{l_S}, \rho, \sigma^*) \tag{EC.60}$$

$$= (1 - \xi) m(\mathcal{S}, c_\xi^{l_S}, \rho, \sigma^*). \tag{EC.61}$$

Further, because the total waiting time is monotone increasing, by the integral upper bound, the following holds:

$$\frac{p_F}{\mu} = \int_0^{c^L} w(\mathcal{S}, x, \rho, \sigma^*) dx + c^L \left(\frac{1}{\mu} - T\right) \tag{EC.62}$$

$$\leq c^L \, max_{c\in[0,c^L]} \, w(\mathcal{S}, c, \rho, \sigma^*) + c^L \left(\frac{1}{\mu} - T\right) \tag{EC.63}$$

$$= c^L w(\mathcal{S}, 0, \rho, \sigma^*) + c^L \left(\frac{1}{\mu} - T\right) \tag{EC.64}$$

Now observe that the cutoff point c^L goes to zero as the spot market becomes sufficiently small (i.e., $c^L \xrightarrow[l_S\to 0]{} 0$). Combined with Equation (EC.64), it follows that the waiting time goes to infinity for users with bid 0, i.e.:

$$w(\mathcal{S}, 0, \rho, \sigma^*) \xrightarrow[l_S\to 0]{} \infty. \tag{EC.65}$$

As a job with bid 0 is served exactly when there is an idle instance in the spot queue (i.e., there are fewer than l_S jobs of higher priority in the spot queue), the instance utilization of the spot queue has to go to full as the size of the spot market becomes sufficiently small, i.e.

$$\frac{\lambda_S}{l_S \mu} \xrightarrow[l_S\to 0]{} 1. \tag{EC.66}$$

Now fix some $\xi > 0$. It holds that

$$\frac{(1 - F_S(c_\xi^{l_S}))\lambda_S}{l_S \mu} \xrightarrow[l_S \to 0]{} (1-\xi)1 \tag{EC.67}$$

i.e., as the size of the spot market goes towards zero, the (average) utilization of the spot instances by jobs with priority over $c_\xi^{l_S}$ will always at most be $(1-\xi)$. For a given ξ (but independent of l_S), this limits the total waiting time at $c_\xi^{l_S}$ to some possibly very high but finite value $\bar{w}_\xi$. For any $c_\xi^{l_S}$ it further either holds $c_\xi^{l_S} > c_1^L$ (and $m(\mathcal{S}, c_1^L, \rho, \sigma^*) < m(\mathcal{S}, c_\xi^{l_S}, \rho, \sigma^*)$ trivially) or the following holds:

$$\begin{aligned}
& m(\mathcal{S}, c_1^L, \rho, \sigma^*) && \text{(EC.68)}\\
&= \int_0^{c_\xi^{l_S}} w(\mathcal{S}, x, \rho, \sigma^*)dx + \int_{c_\xi^{l_S}}^{c_1^L} w(\mathcal{S}, x, \rho, \sigma^*)dx - c_1^L w(\mathcal{S}, c_1^L, \rho, \sigma^*) && \text{(EC.69)}\\
&\leq \int_0^{c_\xi^{l_S}} w(\mathcal{S}, x, \rho, \sigma^*)dx + (c_1^L - c_\xi^{l_S}) w(\mathcal{S}, c_\xi^{l_S}, \rho, \sigma^*) - c_1^L w(\mathcal{S}, c_1^L, \rho, \sigma^*) && \text{(EC.70)}\\
&= \int_0^{c_\xi^{l_S}} w(\mathcal{S}, x, \rho, \sigma^*)dx - c_\xi^{l_S} w(\mathcal{S}, c_\xi^{l_S}, \rho, \sigma^*) + c_1^L (w(\mathcal{S}, c_\xi^{l_S}, \rho, \sigma^*) - w(\mathcal{S}, c_1^L, \rho, \sigma^*)) && \text{(EC.71)}\\
&\leq \int_0^{c_\xi^{l_S}} w(\mathcal{S}, x, \rho, \sigma^*)dx - c_\xi^{l_S} w(\mathcal{S}, c_\xi^{l_S}, \rho, \sigma^*) + c_1^L (\bar{w}_\xi - w(\mathcal{S}, c_1^L, \rho, \sigma^*)) && \text{(EC.72)}\\
&= m(\mathcal{S}, c_\xi^{l_S}, \rho, \sigma^*) + c_1^L (\bar{w}_\xi - w(\mathcal{S}, c_1^L, \rho, \sigma^*)) && \text{(EC.73)}
\end{aligned}$$

As $c_1^L \xrightarrow[l_S \to 0]{} 0$, it follows that, for all $0 < \xi < 1$ and all $\delta > 0$ there exists an l_S with $m(\mathcal{S}, c_\xi^{l_S}, \rho, \sigma^*) \geq m(\mathcal{S}, c_1^L, \rho, \sigma^*) - \delta$ and therefore

$$\begin{aligned}
\bar{m}(\mathcal{S}, \rho, \sigma^*) &\geq (1-\xi) m(\mathcal{S}, c_\xi^{l_S}, \rho, \sigma^*) && \text{(EC.74)}\\
&\geq (1-\xi)(m(\mathcal{S}, c_1^L, \rho, \sigma^*) - \delta) && \text{(EC.75)}
\end{aligned}$$

Choosing ξ and δ such that $\frac{1}{2}\varepsilon \geq \xi m(\mathcal{S}, c_1^L, \rho, \sigma^*) + (1-\xi)\delta$ and noting that by Lemma EC.2 it holds $m(\mathcal{S}, c_1^L, \rho, \sigma^*) = \frac{p_F}{\mu}$ then yields

$$\bar{m}(\mathcal{S}, \rho, \sigma^*) \geq m(\mathcal{S}, c_1^L, \rho, \sigma^*) - \frac{1}{2}\varepsilon = \frac{p_F}{\mu} - \frac{1}{2}\varepsilon \tag{EC.76}$$

and the statement of the lemma follows.

When $T > \frac{1}{\mu}\left(\frac{1}{1-\tau\psi_E(l_S)} - 1\right)$ and $\sigma^* = (\overrightarrow{c}^P, \overrightarrow{c}^B)$, we analogously (only replacing the relevant notation) obtain

$$\bar{m}(\mathcal{S}, \rho, \sigma^*) \geq m(\mathcal{S}, c_1^P, \rho, \sigma^*) - \frac{1}{2}\varepsilon. \tag{EC.77}$$

Because users with waiting cost $\overrightarrow{c}_1^P$ are indifferent between both markets, it has to hold that $\frac{p_F}{\mu} + c_1^P(\frac{1}{\mu}+T) = c_1^P \frac{1}{\mu}\frac{1}{1-\tau\psi_E(l_S)} + m(\mathcal{S}, c_1^P, \rho, \sigma^*)$. Solving this for $m(\mathcal{S}, c_1^P, \rho, \sigma^*)$ and substituting it into Equation (EC.77) yields

$$\bar{m}(\mathcal{S}, \rho, \sigma^*) \geq \frac{p_F}{\mu} - c_1^P \left(\frac{1}{\mu} \left(\frac{1}{1-\tau\psi_E(l_S)} - 1 \right) - T \right) - \frac{1}{2}\varepsilon. \tag{EC.78}$$

Lastly note that $c_1^P \left(\frac{1}{\mu} \left(\frac{1}{1-\tau\psi_E(l_S)} - 1 \right) - T \right) \xrightarrow[l_S \to 0]{} 0$ because $\frac{1}{\mu} \left(\frac{1}{1-\tau\psi_E(l_S)} - 1 \right) - T$ increases in l_S and it holds that $c_1^P \xrightarrow[l_S \to 0]{} 0$. For any $\varepsilon > 0$, l_S small enough therefore yields the statement of the lemma. □

EC.2. Basis of the Numerical Example

In this section, we give a precise description of how we calculate waiting times and preemption probabilities for the numerical examples. In the following we let $\varphi(l, \frac{\lambda}{\mu})$ denote the probability that fewer than l jobs are currently in a queue with l instances, arrival rate λ and with expected job service time $\frac{1}{\mu}$. For memoryless queues, $\varphi(l, \frac{\lambda}{\mu})$ is given by the well-known Erlang C formula:[20]

$$\varphi(l, \frac{\lambda}{\mu}) = 1 - \left(1 + (1 - \frac{\lambda(c)}{l\mu}) \frac{l!}{\frac{\lambda(c)}{\mu}^l} \sum_{k=0}^{l-1} \frac{\frac{\lambda(c)}{\mu}^k}{k!} \right)^{-1} \tag{EC.79}$$

Given this, the required calculations of the numerical example for the fixed-price queue are straightforward, while we need to make additional simplifying assumptions for the spot queue.

EC.2.1. Fixed-price Queue

For a fixed-price queue, the total waiting time $w(\mathcal{F}, c, \rho, \sigma)$ and payment $m(\mathcal{F}, c, \rho, \sigma)$ are directly determined by the parameters of the setting. The only thing left to calculate is the minimal number of instances $l_F(\rho, \sigma)$ required to serve all users in the fixed-price market while observing the upper bound on the queuing time T. This can easily be done using the Erlang C formula, as it is well known that the expected queuing time $q(l, \frac{\lambda}{\mu})$ of a user joining a FIFO queue with l instances, arrival rate λ and service time $\frac{1}{\mu}$ is given by

$$q(l, \frac{\lambda}{\mu}) = \frac{1 - \varphi(l, \frac{\lambda}{\mu})}{l\mu - \lambda}. \tag{EC.80}$$

See Cooper (1981) for a proof. Plugging in the arrival rate into the fixed-price market for any pair of strategies (ρ, σ) and solving $q(l_F(\rho, \sigma)) = T$ then yields $l_F(\rho, \sigma)$.

[20] See for example Cooper (1981); a proof of the Erlang C formula can be found in Takagi (2008).

EC.2.2. Spot Queue

For the spot queue, the total waiting time $w(\mathcal{S}, c, \rho, \sigma)$ and payment $m(\mathcal{S}, c, \rho, \sigma)$ are not directly determined by the setting because they depend on the dynamics of the queue. Unfortunately, we cannot directly use the Erlang C to derive those terms because this would require that the running times and queuing times for all users are equal (Buzen and Bondi 1983). Since with priorities, and especially when costly preemptions are present, they are *not* equal, we make the following two simplifying assumptions (for the numerical examples only). This then allows us to calculate waiting times and preemption probabilities.

ASSUMPTION EC.1. *For calculating $w(\mathcal{S}, c, \rho, \sigma)$, we assume that jobs, while running, see the steady state distribution over states of the spot queue (when looking at the queue not including itself).*

Note that Assumption EC.1 would be exactly satisfied if a given job ran for an infinite amount of time. Since jobs start in a random state but end in a state in which the queue has free capacity for a job of the same priority, jobs in practice see more "busy" states than in steady state and consequently take slightly longer to run. Importantly, we only make this assumption when calculating any single jobs's runtime, but we still calculate the steady state of the queue exactly to avoid the accumulation of approximation errors.

ASSUMPTION EC.2. *Any additional running time above and beyond a job's service time $\frac{1}{\mu}$ is run on "abstract" additional instances and does not influence the spot queue's steady state. However, while run on these abstract instances, a job still causes load-dependent costs for the provider and is still (internally and externally) preempted as if it was in the queue, as denoted by $\psi_I(c, \rho, \sigma)$ and $\psi_E(l_S)$.*

Effectively, Assumption EC.2 dynamically gives the spot market more instances than it actually has, to accommodate the additional running time needed due to preemptions.

Taken together, these two assumptions give us the following very useful Lemma.

LEMMA EC.4. *During its time in the spot queue, a job with bid c sees the steady state distribution of a FIFO queue with arrival rate $\lambda(c)$ and service time $\frac{1}{\mu}$. Here, $\lambda(c)$ denotes the arrival rate of jobs with a higher priority; i.e., during any time unit, on average, $\lambda(c)$ jobs with a higher priority than c arrive into the queue.*

Proof Note that jobs with a lower bid do not influence the total waiting time of a user and can thus be ignored. As the probability that any other user in the system also has a waiting cost of exactly c is zero, we can assume that every other job has a strictly higher bid and thus a strictly higher priority. Combining this with Assumption EC.1, we can assume that the job, while running,

sees the steady state probabilities of the queue consisting of only those users with higher priorities than itself. Furthermore, by Assumption EC.2, these steady state probabilities are the same as the steady state probabilities with zero preemption costs, which in turn are the same as the steady state probabilities of a FIFO queue consisting of all users with higher priority (see Buzen and Bondi (1983)). Taken together, the statement follows. □

Given Lemma EC.4, we can now derive expressions for the waiting time and the expected number of internal preemptions per time unit.

PROPOSITION EC.1. *Given Assumptions EC.1 and EC.2, provider strategy* $\rho = (p_F, l_S)$*, and user strategy profile* σ*, the total waiting time of a user with bid* c *is given by*

$$w(\mathcal{S}, c, \rho, \sigma) = \frac{r(\mathcal{S}, c, \rho, \sigma)}{\varphi(l_S, \frac{\lambda(c)}{\mu})}, \tag{EC.81}$$

where $\lambda(c)$ *denotes the arrival rate of jobs with a bid higher than* c *into the spot queue (given* σ*).*

Proof Recall that the waiting time is defined as

$$w(\mathcal{S}, c, \rho, \sigma) = q(\mathcal{S}, c, \rho, \sigma) + r(\mathcal{S}, c, \rho, \sigma). \tag{EC.82}$$

Observe that when a job with bid c is in the spot queue, it is run whenever there are fewer than l_S jobs with a higher bid in the system. By Lemma EC.4, during its runtime, a jobs sees the steady state distribution of a FIFO queue with arrival rate $\lambda(c)$ and service time $\frac{1}{\mu}$. Thus, to be running for one full time unit, the job with bid c will, on average, have a waiting time of $\frac{1}{\varphi(l_S, \frac{\lambda(c)}{\mu})}$ time units. The statement of the Proposition now follows by noting that the running time of a job is given by $r(\mathcal{S}, c, \rho, \sigma)$. □

PROPOSITION EC.2. *Given Assumptions EC.1 and EC.2, provider strategy* $\rho = (p_F, l_S)$*, and user strategy profile* σ*, the expected number of internal preemptions per time unit of a user with bid* c *is given by*

$$\psi_I(c, \rho, \sigma) = \frac{(\mu l_S - \lambda(c))(1 - \varphi(l_S, \frac{\lambda(c)}{\mu}))}{\varphi(l_S, \frac{\lambda(c)}{\mu})}, \tag{EC.83}$$

where $\lambda(c)$ *denotes the arrival rate of jobs with a bid higher than* c *into the spot queue (given* σ*).*

Proof While a job with bid c is running, it will be internally preempted whenever the system contains exactly $l_S - 1$ jobs of higher priority and another job of higher priority arrives. By Lemma EC.4, during its runtime, the job sees the steady state distribution of a FIFO queue with arrival rate $\lambda(c)$ and service time $\frac{1}{\mu}$. Thus, given a newly-arriving job (with priority higher than c), the

probability that this job preempts the job with bid c is equal to the probability that the FIFO queue contains exactly $l_S - 1$ jobs, which is given by

$$\frac{(1-\frac{\lambda(c)}{\mu l_S})}{\frac{\lambda(c)}{\mu l_S}}(1-\varphi(l_S,\frac{\lambda(c)}{\mu})) \tag{EC.84}$$

(see Cooper (1981)). Since we are interested in the preemption rate taken over the running time of the job (as opposed to the total time the job is in the system), we normalize this term by the probability that less than l_S jobs of higher priority are in the system. Because $\lambda(c)$ jobs with higher priority arrive per time unit, we also multiply with $\lambda(c)$, which yields

$$\psi_I(c,\rho,\sigma) = \lambda(c)\frac{\frac{(1-\frac{\lambda(c)}{\mu l_S})}{\frac{\lambda(c)}{\mu l_S}}(1-\varphi(l_S,\frac{\lambda(c)}{\mu}))}{\varphi(l_S,\frac{\lambda(c)}{\mu})} \tag{EC.85}$$

$$= \frac{(\mu l_S - \lambda(c))(1-\varphi(l_S,\frac{\lambda(c)}{\mu}))}{\varphi(l_S,\frac{\lambda(c)}{\mu})}. \tag{EC.86}$$

□

Taking the expression for the waiting time from Proposition EC.1, plugging in the expression for the running time from Proposition 1, and lastly plugging in the expression for the internal preemptions from Proposition EC.2, we can now write the expected total waiting time $w(\mathcal{S},c,\rho,\sigma)$ of a user joining the spot queue as

$$w(\mathcal{S},c,\rho,\sigma) = \frac{1}{\mu\varphi(l_S,\frac{\lambda(c)}{\mu})}\frac{1}{1-\tau(\psi_I(c,\rho,\sigma)+\psi_E(l_S))} \tag{EC.87}$$

$$= \frac{1}{\mu\varphi(l_S,\frac{\lambda(c)}{\mu})}\frac{1}{1-\tau(\frac{(\mu l_S-\lambda(c))(1-\varphi(l_S,\frac{\lambda(c)}{\mu}))}{\varphi(l_S,\frac{\lambda(c)}{\mu})}+\psi_E(l_S))}. \tag{EC.88}$$

Using the iterative approach described in the proof of Proposition 5, we can calculate the cutoff vectors of the user equilibrium strategies by solving a number of non-linear root searches. This allows us to calculate payments and profits and search for the optimal provider strategy $\rho = (p_F, l_S)$.

EC.3. User Welfare in Example 1

To help us better understand how the hybrid strategy with $p_F = p_F^{*0}$ affects the users, Figure EC.1 shows the user welfare for this strategy and compares it against the fixed-price strategy (we omit the other two strategies because plotting all four strategies makes the figure very hard to read). As we can see, for the fixed-price strategy (dashed blue line), the user welfare monotonically decreases in the instance costs κ. While the instance costs do not directly affect the users, they influence the optimal provider strategies. This leads to the observed discontinuities in the welfare, since instances can only be bought in discrete units and the provider consequently changes her strategy in discrete steps.

It is clearly visible that the hybrid strategy with $p_F = p_F^{*0}$ (the dotted black line) shares many large discontinuity points with the fixed-price strategy (the dashed blue line). This happens because both strategies share the same price p_F^{*0}, which at these points increases, causing some users to balk and making everyone that remains in the fixed-price market worse off. However, the *size* of the discontinuities may differ, because under the hybrid strategy, some users might additionally move from the fixed-price to the spot market. Perhaps surprisingly, *between* these shared discontinuity points, the welfare corresponding to the hybrid strategy often increases. This happens because the provider is restricted to $p_F = p_F^{*0}$ and cannot optimally increase her price, so has to change l_S instead. To understand this in more detail, consider a situation where a cost increase does not lead to a change of p_F, but where the provider still wants users to move out of the relatively less profitable fixed-price market and into the spot market. In this situation, she then has to increase the attractiveness of the spot market, which increases the user welfare. Finally, as expected, the hybrid strategy with $p_F = p_F^{*0}$ always leads to a higher user welfare than the fixed-price strategy, which can be seen by the dotted black line always lying above the dashed blue line.

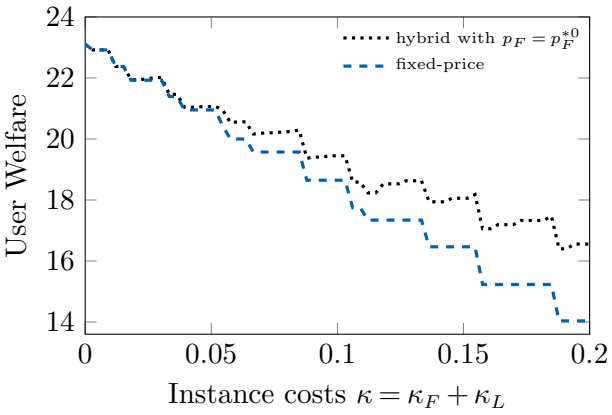

Figure EC.1 User welfare under different strategies while varying the instance costs κ

EC.4. Using Idle Fixed-price Instances for the Spot Market

In this section, we consider an alternative model where the provider takes the spot instances from the pool of currently idle *fixed-price instances* instead of using other idle capacity (long-term reserved instances, maintenance capacity, etc.). In Section EC.4.1., we first discuss the importance of using instances that are *reliably idle* for the spot market and why the fixed-price market is typically *not* the best source for such instances. Nevertheless, as some providers may want to use idle instances from the fixed-price market, we then show how the new cross-channel interactions change our results compared to our main model. In Section EC.4.2, we first present the required changes to the model. In Section EC.4.3, we then show how the equilibria of the user game change. Finally, in Section EC.4.4, we then derive a similar condition for how a provider can simultaneously achieve a profit increase and a Pareto improvement for the users (as we did for our main model).

EC.4.1. The Availability of Reliably Idle Instances from the Fixed-Price Market

As preemptions are costly for the users, the "usefulness" of an idle instance critically relies on how *reliably idle* it is (i.e., for how long the instance will remain idle). Therefore, the provider should use the most reliably idle instances for the spot market. While fixed-price markets of large cloud providers do contain a reasonable number of idle instances *on average*, only few of those instances are reliably idle and providers should not simply put any idle fixed-price instance on the spot market. This is due to two effects: larger markets require relatively smaller buffers and these buffers will be used more frequently the larger the market (but for increasingly shorter durations). To see this, we now provide a simple but striking numerical example.

EXAMPLE EC.1. Consider two M/M/l queues, one with arrival rate 100 and one with arrival rate 1000. Assume for both an expected service time of $\frac{1}{\mu} = 1$ and an SLA of $T = 0.0001$. We can use the Erlang C formula to derive that we need a buffer of 22 instances to satisfy the SLA for the first queue with arrival rate of 100 (i.e., the provider needs $l_F = 122$ fixed-price instances). For the second queue with arrival rate 1000 the required buffer only grows from 22 to 55 (i.e., the provider now needs $l_F = 1055$ fixed-price instances).

Now assume that the provider uses up to 1 idle instance for a spot queue, i.e., $l_s = 1$. Assume that this idle instance comes from the first queue (with arrival rate 100) and whenever the provider has at least 1 idle fixed-price instance and no spot job is running, he starts a new spot job. Then a job in the spot market would in expectation get externally preempted at a rate of $\psi_E = 0.47$. For the second queue (with arrival rate 1000), the rate of external preemptions for spot jobs rises to $\psi_E = 3.07$. For both queues, the high preemption rate occurs because a large queue can frequently reach zero idle capacity while still satisfying the SLA, as it is highly likely that another instance becomes free shortly thereafter (and because newly-arriving users are willing to wait a little bit). However, any time this happens, the spot instance is immediately preempted. This effect becomes more pronounced the larger the fixed-price queue, which explains the large rise of the preemption rate for the second queue with arrival rate 1000.

In practice, the provider wants to keep the preemption rate reasonably low. To this end, she can decide to only start spot jobs whenever the expected time until any running spot job will be externally preempted is above some threshold t_E, resulting in $\psi_E < \frac{1}{t_E}$. While doing this reduces the number of external preemptions, this of course also further reduces the supply of instances for the spot market. For example, if the provider wanted to ensure that jobs (in expectation) run for at least $t_E = 20$ before they get preempted (leading to $\psi_E < \frac{1}{t_E} = 0.05$), then she would have to only start spot jobs when the fixed-price queue contains less than 105 jobs (for the queue with arrival

rate 100 and $l_F = 122$ fixed-price instances) and less than 957 jobs (for the queue with arrival rate 1000 and $l_F = 1055$ fixed-price instances).[21]

Note that the size of the effects observed in the example are particularly large because the example considers memoryless service processes. While real-world service processes are usually heavy-tailed (which leads to larger and more reliably idle buffers), the observation that larger fixed-price markets lead to less reliably idle instances remains true in practice. Thus, a provider who wants to limit the number of external preemptions can only offer relatively few reliably idle fixed-price instances on the spot market. Additionally, the provider has to consider the cross-channel effects that occur when users move from the fixed-price market to the spot market, which decreases the number of fixed-price instances but not the number of users. While a provider can nevertheless choose to offer idle fixed-price instances on the spot market, most providers typically have access to many alternative instances that are more reliably idle than most idle instances from the fixed-price market. This includes instances from other business areas (e.g., long-term reserved instances), maintenance instances (which make up 5-10% of the capacity of a cloud computing center), etc. At least some of these alternatives are available for all current major cloud providers.

EC.4.2. Required Model Changes

Even though most idle instances from the fixed-price market are typically not reliably idle, some cloud providers may still want to use them for a secondary spot market. Therefore, we now show how to adapt our model to using idle fixed-price instances. The most immediate change is that an *external preemption* now happens whenever a spot user is preempted in favor of a fixed-price user. Thus, while previously the number of external preemptions $\psi_E(l_S)$ was a function given by the setting that only depended on the provider's strategy ρ, the number of external preemptions $\psi_E(c, \rho, \sigma)$ now arises from the queueing system. Specifically, it now also depends on the strategies of all other users σ and, because lower bids get preempted first, also on a user's bid c. While in our main model, l_S denotes the (average) number of offered spot instances, it now denotes the maximum number of idle fixed-price instances the provider offers on the spot market, i.e., l_S is now an *upper bound* on the number of offered spot instances.

To control the number of external preemptions and only offer sufficiently reliably idle instances on the spot market, we introduce an additional strategy variable for the provider t_E, which denotes that the provider only starts a new spot job whenever, after starting this job, the expected time until the next external preemption for any running spot job is above the threshold t_E. Thus, given provider strategy $\rho = (p_F, l_S, t_E)$, for a job to be started in the spot market, four conditions have

[21] The preemption rates and the expected time until the next preemption can be calculated by solving the difference equations of the corresponding Markov chains.

to be satisfied: (1) no job with a higher priority is waiting; (2) if the job started, there would be at most l_S spot jobs running, (3) there has to be an idle fixed-price instance or there is currently a spot job with a lower priority running, and (4) if the job started now, the expected time until the next external preemption for any running spot job would be at least t_E. Note that this implies that a spot job with low priority is *not* immediately preempted when a spot job with higher priority is waiting if the expected time until the next external preemption is currently too low. Due to the new cross-channel interactions that arise because the spot instances are now taken from the fixed-price market, both the running time $r(\mathcal{S}, c, \rho, \sigma)$ and the queuing time $q(\mathcal{S}, c, \rho, \sigma)$ in the spot market are now highly dependent on the number of users that join the fixed-price market. Additionally, if the threshold t_E is set too high (for a given user strategy profile σ), then the provider may never start a spot job (effectively not offering a spot market). A setting in our alternative model is now fully defined by $(n, v, \lambda, \mu, F, \tau, \kappa_F, \kappa_L, T)$ because the alternative model does not contain a maximum number of available spot instances l nor an exogenous function for the number of external preemptions.

Because larger fixed-price markets can have less reliably idle capacity than smaller markets (see Example EC.1), we may observe the counterintuitive effect that the waiting time of the users with the highest priority in the spot market can decrease in the number of people that join the spot market. However, the overall costs of any user joining the spot market, i.e, $\int_0^c w(\mathcal{S}, x, \rho, \sigma)dx$ typically increase for any fixed c if more users move to the spot market. To see this, note that when users move from the fixed-price market to the spot market, the total number of instances decreases, but the number of users does not. Yet we cannot say with certainty that $\int_0^c w(\mathcal{S}, x, \rho, \sigma)dx$ *always* increases because the service discipline of the whole market is not work-conserving (i.e., there can be an idle instance even though jobs are waiting when the time until the next external preemption is too low) and these dynamics change whenever users move from the fixed-price to the spot market. While this effect is typically negligible compared to the reduction in the number of instances in the system, we cannot fully exclude the possibility that there could be some parameterizations for which there is a σ where $\int_0^c w(\mathcal{S}, x, \rho, \sigma)dx$ is *locally* decreasing in the number of users that choose the spot market. To avoid having to handle those cases (which do not change the form of the potential equilibria, but could in rare cases potentially lead to the existence of multiple equilibria) we therefore make the following assumption for the rest of the paper:

ASSUMPTION EC.3. *The overall cost of a user with any fixed bid c joining the spot market, i.e. $\int_0^c w(\mathcal{S}, x, \rho, \sigma)dx$, increases if additional users (compared to σ) move to the spot market.*

EC.4.3. Equilibria

Whenever some instances are actually offered on the spot market, we obtain an equilibrium structure similar to the one derived in Subsection 4.3.1:

PROPOSITION EC.3. *For any provider strategy $\rho = (p_F, l_S, t_E)$, in any BNE of the user game where any user joins the spot market, any equilibrium strategy profile is of the form $\sigma^* = (\overrightarrow{c}^P, \overrightarrow{c}^B)$. Here, $\sigma = (\overrightarrow{c}^P, \overrightarrow{c}^B)$ denotes that a user of class i with waiting cost c joins the spot market when $c < c_i^P \leq c_i^B$ and the fixed-price market when $c_i^P < c < c_i^B$; when $c > c_i^B$, he balks and does not join any market. The cutoff point c_1^P and the cutoff vector $\overrightarrow{c}^B$ are the unique solution to the following system of equations:*

$$0 = c_1^P(T + \frac{1}{\mu}) + \frac{p_F}{\mu} - \int_0^{c_1^P} w(\mathcal{S}, x, \rho, (\overrightarrow{c}^P, \overrightarrow{c}^B))dx \tag{EC.89}$$

$$0 = v_i - \min\left\{ c_i^B(\frac{1}{\mu} + T) + \frac{p_F}{\mu}, \int_0^{c_i^B} w(\mathcal{S}, x, \rho, (\overrightarrow{c}^P, \overrightarrow{c}^B))dx \right\} \quad \forall i \in \{1, \ldots, n\} \tag{EC.90}$$

The rest of the cutoff vector $\overrightarrow{c}^P$ is given as $c_i^P = min(c_1^P, c_i^B)$.

Proof Even users with the highest bid in the spot market have to queue longer than users in the fixed-price market, because (by definition) a user only gets served in the spot market when the fixed-price market has idle capacity. Thus, all users in the spot market are willing to pay strictly less than what they would have to pay in the fixed-price market. The remainder of the proof is equivalent to the proof of Proposition 4. □

While this gives us the structure of the equilibrium when some spot instances are offered and utilized, the following proposition tells us when that is the case.

PROPOSITION EC.4. *For any provider strategy $\rho = (p_F, l_S, t_E)$, the equilibrium strategy profile of the users is*

1. *$\sigma^* = \overrightarrow{c}^F$ (i.e., no user joins the spot market, as described in Proposition 2) if and only if $l_S = 0$ or t_E is "too high," i.e., the fixed-price queue arising from $\sigma = \overrightarrow{c}^F$ has no state for which the expected time until the next external preemption of a hypothetically starting spot job would be more than t_E.*
2. *$\sigma^* = (\overrightarrow{c}^P, \overrightarrow{c}^B)$ otherwise.*

Proof Recall from Proposition 2 that $\sigma = \overrightarrow{c}^F$ is the equilibrium user strategy profile when no spot market is offered. We denote by $(\overrightarrow{x}, \overrightarrow{c}^F)$ a different strategy profile where any user of class i with waiting cost c joins the spot market if $c < x_i$. We now look at different provider strategies and classify the corresponding user equilibrium strategy profiles. If $l_S = 0$, then $\sigma^* = \overrightarrow{c}^F$ trivially. Now assume that the fixed-price queue arising from $\sigma = \overrightarrow{c}^F$ has no state for which the

expected time until the next external preemption of a hypothetically starting spot job would be more than t_E. Then, as long as almost all users (i.e., all besides at most a null set) play $\sigma = \overrightarrow{c}^F$, the provider would never start a spot job, even if a single user deviated to the spot market and $l_S > 0$. Consequently, it holds that $\int_0^x w(\mathcal{S}, c, \rho, \sigma)dc = \infty$. By Assumption EC.3, it immediately follows that $\int_0^x w(\mathcal{S}, c, \rho, (\overrightarrow{x}, \overrightarrow{c}^F))dc = \infty$ for any $\overrightarrow{x}$ and thus, in equilibrium, no user joins the spot market. On the other hand, if $l_s > 0$ and if the fixed-price queue arising from user strategy profile $\sigma = \overrightarrow{c}^F$ has a state for which the expected time until the next external preemption of a hypothetically starting spot job would be more than t_E, then users with waiting cost very close to 0 prefer the spot market and by Proposition EC.3 it holds that $\sigma^* = (\overrightarrow{c}^P, \overrightarrow{c}^B)$. □

EC.4.4. Well-behaved Settings: Increasing Provider Profit and User Welfare

We now show how the profit and welfare result from our main model translates to the alternative model. First note that the bound from Lemma 2 on the number of saved fixed-price instances per fixed-price user who moves to the spot market still holds, as the mechanics of the fixed-price market did not change. Next, we translate Lemma 3 to the alternative model.

LEMMA EC.5. *The average running time in the spot market (i.e., the left-hand side of the following inequality) is bounded above as follows:*

$$\frac{\sum_i \lambda_i \int_{x:\sigma_{i,1}(x)=\mathcal{S}} r(\mathcal{S}, x, \rho, \sigma) f_i(x)dx}{\sum_i \lambda_i \int_{x:\sigma_{i,1}(x)=\mathcal{S}} f_i(x)dx} < (\frac{1}{\mu} + \tau)\frac{1}{1 - \tau\frac{1}{t_E}} \tag{EC.91}$$

Proof Recall that $\psi_I(y, \rho, \sigma)r(\mathcal{S}, y, \rho, \sigma)$ denotes the number of internal preemptions a job suffers in expectation. By the same arguments as in Lemma 3, it holds that

$$\frac{\sum_i \lambda_i \int_{x:\sigma_{i,1}(x)=\mathcal{S}} \psi_I(x, \rho, \sigma) r(\mathcal{S}, x, \rho, \sigma) f_i(x)dx}{\sum_i \lambda_i \int_{x:\sigma_{i,1}(x)=\mathcal{S}} f_i(x)dx} < 1 \tag{EC.92}$$

and

$$r_I(\mathcal{S}, c, \rho, \sigma) = \psi_I(c, \rho, \sigma) r(\mathcal{S}, c, \rho, \sigma)\tau \frac{1}{1 - \tau\psi_E(c, \rho, \sigma)} \tag{EC.93}$$

$$\leq \psi_I(c, \rho, \sigma) r(\mathcal{S}, c, \rho, \sigma)\tau \frac{1}{1 - \tau\frac{1}{t_E}}, \tag{EC.94}$$

where (EC.94) follows from (EC.93) because, whenever a job starts to run, the expected time until the next preemption is bounded by t_E. Combining these two inequalities we obtain

$$\frac{\sum_i \lambda_i \int_{x:\sigma_{i,1}(x)=\mathcal{S}} r_I(\mathcal{S}, x, \rho, \sigma) f_i(x)dx}{\sum_i \lambda_i \int_{x:\sigma_{i,1}(x)=\mathcal{S}} f_i(x)dx} < \frac{\tau}{1 - \tau\frac{1}{t_E}}. \tag{EC.95}$$

Similar as in the proof of Lemma 3, we finally obtain

$$\frac{\sum_i \lambda_i \int_{x:\sigma_{i,1}(x)=\mathcal{S}} r(\mathcal{S},x,\rho,\sigma) f_i(x)dx}{\sum_i \lambda_i \int_{x:\sigma_{i,1}(xx)=\mathcal{S}} f_i(x)dx} < (\frac{1}{\mu}+\tau)\frac{1}{1-\tau\frac{1}{t_E}}. \tag{EC.96}$$

□

Given these bounds, we can now state a well-behavedness condition analogous to Definition 2 for our main model, where the new parameter t^w corresponds to a lower bound on the strategy variable t_E (capturing the reliablity of the fixed-price instances):

DEFINITION EC.1. We say that a setting $(n,v,\lambda,\mu,F,\tau,\kappa_F,\kappa_L,T)$ is *t^w-well-behaved* if t^w is the infimum over the t_E for which the following holds:

$$\frac{1+\tau\mu}{1-\tau\frac{1}{t_E}} - 1 < \frac{\kappa_F}{\kappa_L} \tag{EC.97}$$

With this definition in hand, we can now show a profit and welfare result analogous to Theorem 2 for our main model.

THEOREM EC.1. *Given a t^w-well-behaved setting $(n,v,\lambda,\mu,F,\tau,\kappa_F,\kappa_L,T)$ and any fixed-price strategy $\rho_0=(p_F^0,0,\infty)$ that results in a positive profit and for which the queue arising from the corresponding equilibrium user strategy profile σ_0^* has any state for which the expected time until the next external preemption of a hypothetically starting spot job would be more than t^w, then there exists a strategy $\rho=(p_F^0,l_S,t_E)$ with the same price p_F^0, with $0<l_S$ and with $t_E \geq t^w$ that yields a higher profit for the provider, i.e.,*

$$\Pi((p_F^0,l_S,t_E),\sigma^*) > \Pi((p_F^0,0,\infty),\sigma_0^*), \tag{EC.98}$$

and the same strategy also yields a Pareto improvement for the users, i.e.,

$$\forall i \in \{1,\dots,n\}\, \forall c \in [0,\mu v_i]: \quad \pi_i^c(\alpha,\beta,\rho,\sigma^*) \geq \pi_i^c(\alpha,\beta,\rho_0,\sigma_0^*), \text{ and} \tag{EC.99}$$

$$\exists i \in \{1,\dots,n\}\, \exists c \in [0,\mu v_i]: \quad \pi_i^c(\alpha,\beta,\rho,\sigma^*) > \pi_i^c(\alpha,\beta,\rho_0,\sigma_0^*). \tag{EC.100}$$

Proof By Proposition EC.4, any such strategy $\rho=(p_F^0,l_S,t_E)$ leads to some users joining the spot market in equilibrium. The proof of the theorem is then equivalent to the proof of Theorem 2 after replacing the general well-behavedness bound on the running time with $b(l_S)=\frac{1+\tau\mu}{1-\tau\frac{1}{t_E}}$. □

Informally, Theorem EC.1 says that if the provider's current fixed-price market has some instances that are sufficiently reliably idle, then she can obtain a profit increase and achieve a Pareto improvement for the users by offering a spot market alongside her existing fixed-price market (as in our main model). Note that, in contrast to our main model, executing the provider's

strategy in practice is now more difficult, because it will typically be intractable to exactly calculate, for every possible state, whether t_E would be satisfied when starting a new job. However, in this case, the provider could still approximate t_E (e.g., by using historical or simulated data).

While our analysis shows that offering idle fixed-price instances on the spot market can (in principle) be advantageous for the provider, recall from Section EC.4.1 that a provider typically only has relatively few fixed-price instances that are sufficiently reliably idle. In contrast, instances from other areas of the cloud computing center (e.g., long-term reserved instances, maintenance instances, or capacity buffers intended for hardware failure) usually offer a better stock of idle capacity. We therefore recommend using idle instances from the fixed-price market only to bolster the supply of instances for the spot market when the utilization of the fixed-price market is particularly low and to instead primarily use other sources of idle capacity for the spot market.