Exploding versus Recallable Offers:
Price Competition and Welfare under Time-Pressured Search

S. Nuray Akin* and Brennan C. Platt†‡

Abstract

We analyze equilibrium search pricing in continuous time, where buyers incur a utility penalty if they fail to meet their deadline. Sellers, unaware of buyer types, provide quotes that are either exploding (expire immediately) or recallable (last for T periods). Under mild restrictions on parameters, a novel equilibrium emerges: some sellers concentrate on a single, aggressive exploding price, while others disperse across higher, recallable prices. When this occurs, exploding offers induce greater inter-temporal competition than direct comparison of recallable offers. Our findings reveal a regulatory tension: banning recallable offers strictly benefits consumers, whereas banning exploding offers boosts total welfare.

Keywords: Equilibrium search, price formation, deadline, recall, discounts, exploding offers

JEL Classification: D40, D83

---

*Department of Economics, Ozyegin University, Cekmekoy, Istanbul, Türkiye (216) 594-0000, nuray.akin@ozyegin.edu.tr

†Department of Economics, Brigham Young University, 2111 WVB, Provo, UT 84602, (801) 422-8904, brennan_platt@byu.edu

‡We thank Alli Wilkinson for her excellent support as a research assistant. We are grateful for comments from Cole Williams, Jose Luis Moraga-Gonzales, Maarten Janssen, Luis Cabral, Mark Armstrong, Aleksei Parakhonyak, as well as many participants and discussants at the 2022 Consumer Search and Switching Costs Workshop, 2022 International Industrial Organization Conference, 2022 International Conference on Public Economic Theory, 2022 Econometric Society European Summer Meeting, 2023 Econometric Society Asian Meetings, and 2023 Society for the Advancement of Economic Theory Conference. S. Akin thanks the University of Iowa Department of Economics for their hospitality during her sabbatical visit and acknowledges financial support from the Technological and Scientific Research Council of Türkiye under grant 2219. All remaining errors are our own.

1 Introduction

Buyers often feel time pressure, stemming from the need to complete their purchase by a specific time.[1] This pressure is exacerbated by the uncertainty inherent in the process of search: buyers are unsure of how soon they will get another quote, what price it will list, and how long the quote will be valid. Indeed, although sellers may not know how motivated a particular buyer is, they can use the price and recallability of their offer to engage in probabilistic price discrimination.

For instance, a take-it-or-leave-it ("exploding") price could deliver profits right away and stop the buyer from discovering competitors' prices, if it is low enough that the buyer is willing to abandon further search. Armstrong and Zhou (2016) provides many examples of tactics used to get buyers to purchase immediately: sellers may offer a discount that only applies today, or require the purchase now or never. In the labor market, academic positions or legal clerkships are sometimes extended contingent on an almost immediate response.

On the other hand, a recallable offer can list a higher price while allowing the buyer to shop around. This delays profits and exposes the seller to future competition via price comparisons, but can ultimately capture buyers who do not find better alternatives before their deadline. Many markets feature some firms offering recall. Mortgage lenders, auto insurers, and rental car agencies often specify a window during which their quote remains valid (sometimes referred to as a rate lock). If a car buyer leaves a dealership, the dealer must decide whether to keep an offer on the table for a period. In all of these settings, sellers cannot perfectly observe buyers' deadlines or past quotes.

In this paper, we investigate equilibrium pricing strategies and the resulting welfare in a frictional competitive market where sellers quote a price and recall policy. Our interest lies in questions that are not straightforward to answer without careful modeling. When do firms benefit from an exploding price strategy? Are consumers always better off with recallable offers? When will exploding offers and recallable offers coexist in the same market? Does industry-wide profit maximization align with that of a single firm? In each case, the answers depend heavily on the type of pricing that emerges in equilibrium.

---

[1] Lemieux and Peterson (2011) find that shoppers for rental truck reservations become less sensitive to price and more likely to purchase as their deadline approaches. Akin, *et al* (2013) show that real estate investment trusts, which face regulatory deadlines to deploy new capital within a year, consequently end up paying more than individual buyers of identical properties. Coey, *et al* (2020) document buyer deadlines across a wide variety of consumer products, both in self-reported survey data from online shoppers and by following each bidder across multiple auctions of the same item, finding that their bids increased after each failure to win.

In our model, a continuous flow of buyers seeks to buy one unit of a homogenous good. Each buyer has a *grace period* T, after which she incurs a per-period cost d until purchase.[2] A continuum of sellers does not observe a buyer's remaining grace period or past quotes. Upon meeting, the buyer learns the details of the seller's quote and decides whether to buy or continue searching. We study three settings: (i) *recallable*, where quotes can be exercised anytime; (ii) *exploding*, where quotes must be accepted immediately; and (iii) *endogenous*, where sellers choose whether to permit recall.

We characterize the equilibrium outcomes in each setting based on who will accept the offered prices. In an *early equilibrium*, some sellers offer the *early price* that any buyer will immediately accept, even those who have just started their search spell. Other sellers may offer *late prices* which are initially rejected by buyers early in their search, hoping for better quotes later, but are eventually accepted by buyers late in their search, whether by receiving a new exploding offer or by exercising an old recallable offer. In a *late equilibrium*, sellers only offer late prices.

In the recallable setting, only a late equilibrium can occur. Sellers offer a continuous distribution of prices, and buyers wait until their deadline to compare all offers. Higher quotes have more risk of being undercut by a competitor, but earn a higher markup if eventually exercised; these forces balance in equilibrium. As offers become more plentiful, the dispersion shrinks but never collapses to a single price.

In the exploding setting, sellers face a similar tradeoff,[3] but here, buyers' price comparisons are all prospective, considering the current offer versus potential future quotes. As offers become more plentiful, the equilibrium shifts from late to early. That is, greater prospective competition pushes sellers to lower their prices, effectively targeting buyers earlier in their search. Unlike the recallable setting, the equilibrium distribution of prices always includes at least one atom of sellers offering identical prices, and can even produce a degenerate distribution in which all firms offer the early price.

In the endogenous setting, sellers consider both options. When buyers receive offers at

---

[2] If a buyer needs housing when moving to a new city as part of employment, the grace period is the time up until the job starts, while d reflects the higher cost of short-term housing or long-distance commuting after the job starts. For a buyer in search for a gift for a special occasion, the grace period is the advance notice of the need for a gift, while d represents lower utility when expectations of a gift are not met. In labor search, T might represent the duration of unemployment benefits while d would be the cost of self-financing after exhausting benefits.

[3] This balanced tradeoff between markup and volume of sales is common to many models of price posting that generate dispersed prices, such as Burdett and Judd (1983), Rob (1985), Diamond (1987), etc. The unique feature of our model is that the impending deadline interacts with the offer's recallability to endogenously determine the volume of buyers willing to accept a given price.

a moderate rate, a mass of firms offer the early price with an exploding policy, while the remainder post a dispersed set of late prices with a recallable policy. Thus, the two policies can coexist for a generic set of parameters. Low arrival rates mimic the recallable setting; high arrival rates replicate the early equilibrium in the exploding setting. Notably, this environment never yields a mass at the monopoly price.

These equilibrium properties are instrumental for understanding welfare implications of recall in search. One might intuitively expect that exploding offers would always work against the buyers' interests, particularly when they already face time pressure. However, we show otherwise. This happens when offers are frequent enough, motivating firms to offer the early price. This competition to capture early buyers is more effective at reducing average prices than the competition among recallable offers at expiration.

In a similar vein, recallable offers may seem to expose sellers to self-inflicted competition; yet sellers realize that recallable offers can be priced higher than exploding offers, which more than compensates for the chance of being undercut, generating greater producer surplus. This happens when offers are infrequent. Indeed, an industry cartel could increase profits by forbidding exploding offers. Significantly, with a high arrival rate, exploding offers dominate recallable offers in terms of profit, despite the fact that the exploding offer is the early price.

For societal welfare, since the price is just a transfer between a buyer and a seller, the goal is to minimize the frequency of consumers incurring deadline penalties, which are deadweight losses. Thus, we find that a market with recall is always more efficient than one without. Any exploding late prices will be rejected by some early buyers, who risk not getting another offer before their deadline. In contrast, an early buyer who receives either the early price or a recallable offer will completely avoid expiration. Under moderate arrival rates, this creates an interesting tension for a regulator who could help consumers by forbidding recallable offers or increase market welfare by allowing them.

We proceed with a review of related work. Section 2 presents the model in which all offers are exogenously recallable, while Section 3 considers the case with exploding offers. Section 4 merges these models to study endogenous recall. Section 5 computes and compares welfare across the models. Section 6 concludes.

1.1 Related Literature

Stigler (1961) initiated the formal modeling of consumer search behavior, taking as given the distribution of prices offered by sellers. The goal of equilibrium search theory has been to complete the model, so that sellers also behave optimally, given the search strategy of buyers. Diamond (1971) highlights the difficulty in sustaining price dispersion in equilibrium, so most models introduce heterogeneity or uncertainty — such as in search costs (e.g. Salop and Stiglitz, 1976; Butters, 1977; Wilde and Schwartz, 1979; Rob, 1985; Janssen and Moraga-Gonzalez, 2004), production costs (e.g Reinganum, 1979; Janssen, *et al*, 2011), demand (e.g Reinganum, 1979), or valuations (e.g Diamond, 1987; Choi, *et al*, 2018). For a comprehensive overview, see Baye, *et al* (2006).

Our model assumes buyers are initially identical in valuation, demand, and search cost. Heterogeneity arises ex-post as buyers experience different search spells or encounter different types of offers, driven by both chance and strategic choices. This generates a continuum of prices under a wide range of parameters, which contrasts with Curtis and Wright (2004) where (in a monetary search setting) price posting generically results in no more than two prices that maximize profits, despite many types of buyers. We overcome this limited price dispersion because un-targeted buyer types build up in steady state, eventually making a range of them equally profitable.

In our exploding setting, a positive mass of sellers may offer the early price or the lowest late price, despite a continuum of buyer types.[4] This phenomenon, absent in prior equilibrium search literature, may explain jumps in offer acceptance rates during search. In our bimodal equilibrium, only the early price and highest late price are offered, without targeting any of the buyers in between. This is analogous to Salop and Stiglitz (1976), but here this pattern arises endogenously with an ex-post continuum of buyer types, rather than only two types ex-ante differing in search cost.

Coey, *et al* (2020) study buyer deadlines in auctions, where sellers offer either a fixed buy-it-now price or an auction. Note that auction prices are necessarily non-recallable. Moreover, there, the posted price is exogenously given, not determined through competition, and buyers can exercise it at any time. In our setting, all prices must be found through search and are subject to competition, with continued search as the fall-back option.[5]

Our recallable setting incorporates a flavor of simultaneous search (e.g. Burdett and Judd, 1983; Janssen and Moraga-Gonzalez, 2004) and perfect recall in Stahl (1989), since these models generate a continuum of prices based on the tradeoff between higher markups

---

[4] A third atom comes from buyers who pass their deadline.

[5] In models of revenue management, deadlines usually apply to the seller, and buyers may wait for rare flash sales (see Dilmé and Li, 2019). Liu and Lu (2025) allows buyers to vary search intensity, prompting sellers to offer menus. Our model differs: deadlines are buyer-specific, capacity is unlimited, and sellers face competition.

and chance of being undercut.[6] Here, richer pricing strategies emerge with buyer deadlines: for example, the early price can end search, a phenomenon absent in Stahl (1989) because shoppers have no discounting or search costs, so might as well receive all possible quotes.

Akin and Platt (2014) study recall in a three-period model, where each offer may be available in the next period with an exogenous probability. There, more-likely recall will benefit late buyers but can hurt early ones. Here, we use a continuous time setting with discounting to allow sellers more richness in who they target in their strategies, including to allow them to endogenously choose their recall policy. Notably, when arrival rates are very high, equilibrium yields only early prices (which cannot occur Akin and Platt, 2014), and consumers can be better off by forbidding exploding offers.

Buyers can recall past prices for an extra cost in Janssen and Parakhonyak (2014), producing a continuous price distribution as in Stahl (1989). Our buyers avoid search costs if they purchase before their deadline. Our endogenous recall effectively bridges the spectrum from costless recall to infinitely-costly recall (in the form of exploding offers). Choi, *et al* (2022) models recall in sequential search for exogenous price and match-quality distributions. Since each match is idiosyncratic, no match is accepted by all buyers as with our early price. Menzio and Trachter (2015) consider asymmetric recall: buyers can always revisit large sellers' quotes but must accept or reject small sellers' offers immediately. Both seller types offer continuous price distributions, but buyers benefit most when both firm types compete, especially when there are frequent interactions with small sellers.

Our early prices mimic the search deterrence in Armstrong and Zhou (2016), where a monopolist induces immediate purchases by raising recall costs. Our exploding setting is an extreme example of this. However, when offers are infrequent (high search frictions), search-deterring discounts do not arise; rather, sellers prefer high prices targeting buyers near their deadline. This can also occur with endogenous recall: the seller chooses recallable offers rather than the early price, when the buyer is unlikely to get another quote. Full recall is welfare improving in Armstrong and Zhou (2016) compared to exploding offers or restricted recall, which inefficiently lowers search and match quality.

Welfare is strongly influenced by the frequency of offers. Janssen and Moraga-Gonzalez

---

[6] This tradeoff also occurs in Burdett and Mortensen (1998). There, workers sequentially receive exploding offers, but can accept a current job offer and continue searching while working, upgrading each time a better offer is made. This process continues indefinitely, since the highest wage in the support is almost never offered. This resembles our buyers in the recall setting, who upgrade whenever a lower price is found, but this process is constrained by the buyer's deadline. Moreover, we find that there can be a positive mass of offers that convince buyers to abandon search.

(2004) show that, in an equilibrium in which some searchers do not participate, having more firms can harm consumers by discouraging participation. In contrast, our model shows that higher arrival rates always benefit consumers and total welfare by increasing competition and reducing the likelihood of incurring deadline penalties, though they also affect targeting and the choice of recall policies across firms.

2 Search with Recallable Offers

Consider a continuous time environment, with infinitesimal buyers and sellers each entering the market at rate δ. All agents discount future utility at rate ρ, and each buyer seeks one unit of the homogenous good being sold. This good provides value x to any buyer, but because sellers may ask different prices for the good, buyers may find it worthwhile to search. Buyers encounter a seller at Poisson rate μ. Upon encounter, the buyer draws an asking price p from the distribution of offered prices, $F(p)$. The buyer can either make the purchase, obtaining $x - p$ surplus and exiting the market, or continue searching while able to costlessly recall past offers at any point. We refer to this as the *recallable setting*, treating recall as exogenously given; later, we make recallability an endogenous choice of sellers in Section 4. If a buyer receives two identical quotes, she randomly selects one to retain.

2.1 Buyer's Problem with Recallable Offers

The buyer has T units of time to search without penalty, which we refer to as the *grace period*. If search extends beyond that period, the buyer incurs a flow cost $d > 0$ until a purchase is made. In other words, search doesn't have any direct costs until after T. Even so, buyers will anticipate those costs as their search continues, particularly because quotes from sellers only arrive sporadically.

We characterize the expected utility of a buyer, $V(z, q)$, as a function of the buyer's remaining grace period z and lowest quote received q. If no quotes have been offered yet, we set $q = +\infty$. We depict this search recursively in Bellman equations, beginning with those who have exhausted their grace period.

$$\rho V(0,q) = \max\left\{\rho(x-q),\ -d + \mu \int_{-\infty}^{q} (V(0,p) - V(0,q))dF(p)dp\right\}. \tag{1}$$

Once the grace period is exhausted ($z = 0$), their search becomes stationary; meaning it no longer matters how much time has elapsed, since there is no further changes in search parameters. At any moment, the buyer can exercise their best option q, which if they did, would yield utility $V(0, q) = x - q$. If they instead choose additional search, they incur cost d each instant, and receive new quotes at rate μ. The new quote is ignored if $p > q$; otherwise, p replaces q, increasing the expected utility by $V(0, p) - V(0, q)$. Note that $V(0, p)$ allows the buyer to exercise this new quote immediately if they so choose.

For a buyer who still has z remaining of the grace period, the Bellman equation alters slightly:

$$\rho V(z, q) = \max\left\{\rho(x - q),\ -V_z(z, q) + \mu \int_{-\infty}^{q} (V(z, p) - V(z, q))dF(p)dp\right\}. \tag{2}$$

As before, the buyer can transact with their best quote q at any time, and will replace it with any lower quote p as it is received. The key addition is that the state variable z steadily falls over time, reflected in the derivative of the first component, $-V_z(z, q)$.

We can immediately derive conditions for when a buyer will make a purchase during their grace period.

Lemma 1. *If the lowest price in the support of F is ℓ, a buyer who is offered quote ℓ will immediately accept it if $\ell < x$, will never accept any price if $\ell > x + \frac{d}{\rho}$, and will exercise the price at time $z = 0$ if $\ell \in \left(x, x + \frac{d}{\rho}\right)$.*

This lemma sets bounds on when a buyer acts on the best possible offer. By virtue of ℓ being the lowest price, the integrals in Eqs. 1 and 2 disappear, so the buyer only contemplates when (if ever) to exercise this price. If there were a price at or below x, it would constitute an *early price* that any buyer would accept immediately; in the absence of those, the buyer waits until after T, then chooses the lowest among offered quotes (provided it is less than $x + \frac{d}{\rho}$). Notably, the buyer makes the same decision at any point in the grace period. We will show that no early price will be profitable in equilibrium in the current setting.

2.2 Steady State Population with Recallable Offers

We consider a steady state equilibrium, where the measure of buyers in each state stays constant over time. One can interpret steady state as the long run outcome of a large market. Sellers correctly anticipate the steady state distribution to select a pricing strategy.

First, we track the measure of buyers who have no quotes with z or less time remaining, which we label $H_u(z)$. Thus, $H_u'(z)$ indicates the relative density of those quoteless buyers. Buyers enter without a quote, so $H_u'(T) = \delta$. They receive quotes at rate μ, so the population of quote-less buyers falls at rate $H_u''(z) = \mu H_u'(z)$. Finally, in steady state, the flow of quote-less buyers newly reaching their deadline must equal the flow of quote-less buyers at their deadline who receive a quote: $H_u'(0) = \mu H_u(0)$. This system of differential equations yields a cumulative quote-less population of:

$$H_u(z) = \frac{\delta}{\mu} e^{-\mu(T-z)}. \tag{3}$$

We must also track how many buyers are in the market at all (with or without quotes), with their cumulative density denoted $H(z)$. Again, $H'(T) = \delta$. Early exit (with $z > 0$ remaining of the grace period) only occurs when the buyer is offered the early price ℓ. Assuming that fraction $\alpha \equiv F(\ell)$ of all offers are the early price, the rate of exit would be $H''(z) = \alpha\mu H'(z)$. At the deadline, however, all buyers with quotes immediately exit (because otherwise, the price quote is never acted upon and thus should not be offered in equilibrium). Therefore, only those without quotes remain: $H(0) = H_u(0) = \frac{\delta}{\mu} e^{-\mu T}$. This solves as:

$$H(z) = \frac{\delta}{\mu} \left(\frac{e^{\alpha\mu(z-T)} - e^{-\alpha\mu T}}{\alpha} + e^{-\mu T} \right). \tag{4}$$

The fractional term becomes μz as $\alpha \to 0$. Note that all buyers still in the market are relevant to sellers because they are still considering price offers; even those with a quote can be lured with a better offer.

2.3 Seller's Problem with Recallable Offers

Sellers produce the good at cost c per unit at the time of the transaction. Throughout the paper, we set $c = x$ for tractability, though similar behavior holds when $c < x$. The sellers are unable to observe the state of the buyer with whom they have been paired (neither their time remaining nor their prior quotes).

Let $\underline{p}$ and $\bar{p}$ denote the infimum and supremum of the support of $F(p \mid p > x)$. That is, these indicate the range of *late prices* offered: prices that are too high to be accepted during the grace period, but will be compared at the deadline. If the seller offers price $p \in [\underline{p}, \bar{p}]$, he must be concerned with the possibility of being undercut by another offer. Competing offers arrive at a Poisson rate μ, and are lower than p with probability $F(p)$. Thus, if a buyer has z periods until expiration, the probability that she gets no better offers *after* this one is $e^{-\mu F(p) z}$. On the other hand, the buyer may already have a better offer in hand at time z, though we know that offer is not the early price, because she would have already accepted it and exited the market. The probability that a buyer received no better price in the $T - z$ time *before* this offer p is $e^{-\mu(F(p)-\alpha)(T-z)}$.

Even if there were no better offers before or after, the sale will not occur for another z periods, and thus profit must be discounted accordingly. Thus, upon making an offer to a buyer of type z, the profit from sale is discounted by $e^{-\rho z-\mu F(p)z-\mu(F(p)-\alpha)(T-z)}$. Finally, there are those buyers already past their deadline who accept any offer immediately, so no discounting is necessary.

Combining the components of the preceding paragraphs, the expected profit upon offering a late price p, averaged over all types z (including those at the deadline) is:

$$\Pi(p) = \frac{H(0) + \int_0^T e^{-\rho z-\mu F(p)z-\mu(F(p)-\alpha)(T-z)} H'(z)dz}{H(T)}(p - c). \tag{5}$$

In contrast, any time the early price ℓ is offered, it will be accepted for sure, generating profit $\ell - c$.

2.4 Equilibrium Definition with Recallable Offers

In this recallable setting, a steady state equilibrium consists of seller profit π, the measure of buyers $H(z)$, and the distribution of sellers' offered prices $F(p)$ such that:

1. All prices in the support of F produce the same maximal profit π, while all other prices produce no more than π.
2. $H(z)$ satisfies the steady state conditions captured in Eq. 4.

The first requirement ensures that sellers only quote prices that are profit maximizing. Among such prices, sellers are indifferent and can randomize over which is chosen.[7] In particular, the early price ℓ can only be offered if $\pi = \ell - c$; otherwise $\alpha^* = 0$. Conversely, if every late price generates strictly less profit than the early price (i.e. $\pi = \ell - c > \Pi(p)$ for all $p \in [\underline{p}, \bar{p}]$), then no such p is offered, so $\alpha^* = 1$.

The second requirement simply states that sellers understand the population law of motion and correctly anticipate its steady-state outcome. Buyers' decisions are defined by Lemma 1 as noted earlier, and are embedded in the sellers' profit calculation.

---

[7] One can interpret $F(p)$ as the mixed strategy that all sellers employ. Alternatively, each seller could use a pure strategy, while $F(p)$ indicates the distribution of such pure strategies across all sellers.

2.5 Equilibrium Characterization with Recallable Offers

We now present the unique solution for equilibrium in the recallable setting. First, by substituting the population (Eq. 4) into profits (Eq. 5), we can evaluate the integral to obtain:

$$\Pi(p) = \frac{\left(\mu\left(1 - e^{-\rho T}\right) e^{\mu T(1-F(p))} + \rho\right) \alpha e^{\alpha\mu T}}{\left(\alpha e^{\alpha\mu T} + e^{(\alpha+1)\mu T} - e^{\mu T}\right)\rho}(p - c). \tag{6}$$

Since all late prices need to yield equal profits, the distribution of prices can easily be solved for by taking the derivative of Eq. 6 with respect to p, and equating it to zero. The resulting differential equation (reported in the Technical Appendix[8]) is uniquely solved with the boundary that $F(\bar{p}) = 1$, yielding:

$$F(p) = 1 - \frac{\ln\left(\frac{\left(\mu\left(1-e^{-\rho T}\right)+\rho\right)\frac{\bar{p}-c}{p-c}-\rho}{\mu\left(1-e^{-\rho T}\right)}\right)}{\mu T}. \tag{7}$$

Note that there are no atoms in this portion of the distribution, because all such offers are evaluated after the grace period, potentially against other offers, so ties become relevant. If a positive fraction of sellers offer the same price, there will be a coin flip to see which actually makes the sale at the deadline (if it is the best offer). But then any seller could lower their offer by ϵ to avoid the coin flip, strictly increasing their chance of sale.

This also allows us to use the solution $F(p)$ in order to calculate $\underline{p}$, since $F(\underline{p}) = \alpha$:

$$\underline{p} = c + (\bar{p} - c)\frac{e^{-\mu(1-\alpha)T}\left(\rho + \mu\left(1 - e^{-\rho T}\right)\right)}{\rho e^{-\mu(1-\alpha)T} + \mu\left(1 - e^{-\rho T}\right)}. \tag{8}$$

Next, we note that to be profit maximizing, the highest late price must make a buyer whose grace period has expired indifferent between accepting $\bar{p}$ and continuing search (that is, $\bar{p} = p^*$ where $V(0, p^*) = x - p^*$). If a seller tried offering a price higher than p^*, even the most desperate of buyers would reject it indefinitely, earning the seller zero profit. If the maximum price $\bar{p}$ were strictly lower than p^*, a seller could deviate by offering $\bar{p} + \epsilon$, which would be accepted whenever $\bar{p}$ is (namely, by those with no other offers at deadline), but will earn additional profit ϵ.

Moreover, since buyers at expiration are indifferent about accepting $\bar{p}$ or continuing search, they strictly prefer to accept if their best offer is lower than $\bar{p}$. This allows us to

---

[8] Available at `https://economics.byu.edu/faculty-and-staff/platt-research`

simplify the post-expiration Bellman Eq. 1 to

$$\rho V(0,q) = \begin{cases} \rho(x-q) & \text{if } q \leq \bar{p} \\ -d + \mu\left(\alpha\ell + \int_{\underline{p}}^{\bar{p}} (x - p - V(0,q))dF(p)dp\right) & \text{if } q > \bar{p}. \end{cases} \tag{9}$$

That is, any price below $\bar{p}$ is immediately executed. Higher prices are ignored and search continues, but the next offer (among prices offered in equilibrium) will be accepted immediately, whether the early price or one of the late prices. The utility from both cases must equate if $q = \bar{p}$ as explained above. We can then insert the solution for F (Eq. 7) and $\underline{p}$ (Eq. 8) into Eq. 9 and solve for the $\bar{p}$ that generates indifference. Intermediate steps are reported in the Technical Appendix, but this results in:

$$\bar{p} = c + \frac{T\rho(\alpha\mu(\ell - c) + \rho(x-c) + d)}{\rho(\mu+\rho)T - (\rho + \mu\left(1 - e^{-\rho T}\right)) \ln\left(1 - \frac{\rho\left(1 - e^{-(1-\alpha)\mu T}\right)}{\rho + \mu\left(1 - e^{-\rho T}\right)}\right)}. \tag{10}$$

We can also substitute for $F(p)$ in the profit equation to get the expected discounted profit from offering a late price:

$$\Pi(\alpha) = \frac{\alpha\left(\rho + \mu\left(1 - e^{-\rho T}\right)\right)}{\rho\left(\alpha + e^{\mu T} - e^{(1-\alpha)\mu T}\right)}(\bar{p} - c). \tag{11}$$

Profit across all late prices are now constant by construction, so with slight abuse of notation, we define Π as a function of the atom at ℓ.

If the seller offers the early price, the buyer's timeline and best price are irrelevant; any buyer will immediately accept, generating profit $\ell - c$ each time a quote is generated. In light of Lemma 1, any early price will result in zero or negative profit and hence will not be offered ($\alpha^* = 0$).[9] Thus all offers received during the grace period will be deferred, with the buyer acting on the best at the time of the deadline. We refer to this as a *late equilibrium* since only late prices are offered. The following proposition reports this unique equilibrium, while Figure 1 provides a numerical example.

---

[9] In our preceding analysis, we allow for the possibility of the early price in preparation for Section 4, where one can occur in equilibrium.

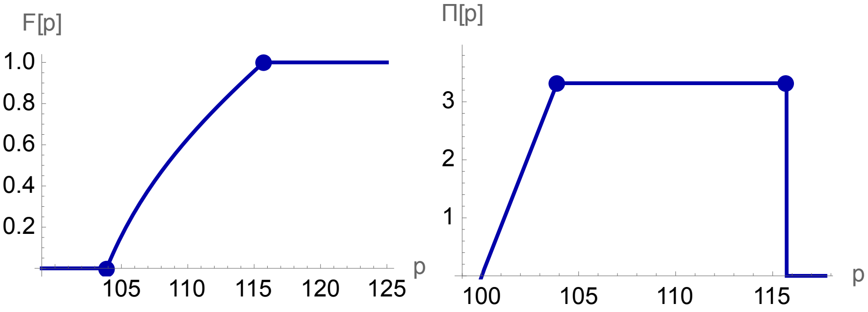


Figure 1: **Equilibrium Solution in the Recallable Setting:** price offer distribution (left), and profits (right), under parameters $x = c = 100$, $\rho = 0.03$, $\mu = 0.15$, $d = 1.5$, and $T = 12$. Dots indicate $\underline{p}$ and $\bar{p}$, respectively.

Proposition 1. *In the recallable setting with $x = c$, the unique equilibrium has a distribution of buyers $H(z)$ as reported in Eq. 4, a distribution of prices $F(p)$ as reported in Eq. 7, and*

$$\bar{p} = x + \frac{d}{\mu + \rho + \left(\frac{\rho+\mu(1-e^{-\rho T})}{T\rho}\right) \ln\left(1 - \frac{\rho(1-e^{-\mu T})}{\rho+\mu(1-e^{-\rho T})}\right)} \tag{12}$$

$$\underline{p} = x + (\bar{p} - x)\frac{e^{-\mu T}\left(\rho + \mu\left(1 - e^{-\rho T}\right)\right)}{\rho e^{-\mu T} + \mu\left(1 - e^{-\rho T}\right)} \tag{13}$$

$$\pi = \frac{\rho + \mu\left(1 - e^{-\rho T}\right)}{\rho\left(1 + \mu T e^{\mu T}\right)}(\bar{p} - x). \tag{14}$$

Note that $F(\underline{p}) = 0$, so $\alpha^* = 0$.

The recallable setting always generates dispersed prices (offering the same good for distinct prices, ranging from $\underline{p}$ to $\bar{p}$) because of the uncertain amount of competition that each seller faces. The number of quotes a buyer obtains over the grace period as well as the quotes drawn are random. Thus, sellers face a tradeoff of higher markup but lower probability of being the best offer. Indeed, this mirrors the uncertainty sellers face in traditional models of simultaneous search,[10] except in our model, receiving those competing offers takes time, which sellers account for in discounting expected profits. In this sense, recall in our model

---

[10] In Burdett and Judd (1983), buyers select a fixed number of quotes which is not known to the seller. Price dispersion only occurs there because some buyers seek only one quote; otherwise, Bertrand-like competition pushes prices to marginal cost. In our setting, the search friction naturally ensures that some buyers will only get one quote over the grace period.

imports the flavors of simultaneous search into a sequential search framework.[11] Moreover, in our setting, the highest willingness to pay ($\bar{p}$) is derived from the option to continue search beyond T, rather than exogenously fixing the reservation price.

We underscore that seller behavior is the driving force of the equilibrium features in the recallable setting. Sellers are making complicated pricing choices, while buyers simply collect all offers and evaluate them at the end of the grace period. Hence, the order in which they are received carries no information; the recalled price is uniformly likely to have been received at any point.

3 Search with Exploding Offers

Now consider another benchmark, where all offers must be used immediately or they are lost, which we refer to as the *exploding setting.* The lack of recall could arise because a seller cannot identify recipients of past quotes, so each buyer draws a new quote even if she does return.[12] As before, sellers are unable to observe a buyer's remaining grace period, but anticipate what fraction of buyers are willing to accept a given offer.

3.1 Buyer's Problem with Exploding Offers

A buyer must choose whether to use a quote as it is received. With exploding offers, the only relevant buyer state is the remaining time until the grace period expires, z; the number or details of rejected past offers are irrelevant. The search problem of those whose grace period has expired ($z = 0$) can be recursively formulated as follows:

$$\rho V(0) = -d + \mu \int_{-\infty}^{\infty} \max\{x - p - V(0),\ 0\}\, dF(p). \tag{15}$$

Here, $V(0)$ represents the expected net present utility of a buyer searching after the grace period. Each instant, she incurs cost d. She is quoted a price p at rate μ, and can either exercise it immediately (for net surplus $x - p$) and abandon further search ($-V(0)$), or can ignore the offer (for 0 net change in utility).

---

[11] Stahl (1989) has a similar feel in that some buyers never search more than once while others seek quotes from all N firms. In our model, buyers are ex-ante identical, differing only ex-post in their luck as to how many quotes they received.

[12] In a less literal sense, exploding offers could arise if sellers have capacity constraints, forcing them to serve their immediate clients only.

During the grace period, the recursive problem takes the following form:

$$\rho V(z) = -V'(z) + \mu \int_{-\infty}^{\infty} \max\{x - p - V(z),\ 0\}\ dF(p). \tag{16}$$

Note two changes in this Bellman equation, compared to decisions after the grace period. First, there is no search cost d. Second, the state variable z deterministically falls as the grace period ticks down, which is reflected in the term $-V'(z)$.

By defining the Bellman equation in this way, we are assuming that both $V(z)$ and $V'(z)$ are continuous and differentiable; thus we do not examine possible equilibria with discontinuous value functions. Even though the instantaneous utility abruptly falls once $z = 0$, the present expected cost of these penalties grows smoothly as expiration approaches.

At each state z, a buyer will simply compare $x-p-V(z)$ to 0. If we define her *reservation price* $R(z)$ such that for all $z \in [0, T]$:

$$R(z) = x - V(z), \tag{17}$$

then the buyer will optimally accept any price $p \leq R(z)$, while rejecting any $p > R(z)$.

3.2 Steady State Conditions with Exploding Offers

We again let $H(z)$ denote the measure of buyers with z or less time remaining in their grace period. Buyers enter the market at rate δ, which is thus the relative density of buyers with their full grace period ahead of them:

$$H'(T) = \delta. \tag{18}$$

At $z > 0$, buyers exit the market only when they find an acceptable price, which happens at rate $\mu F(R(z))$. Thus, the density of buyers at z must fall at that rate:

$$H''(z) = \mu F(R(z)) H'(z). \tag{19}$$

Finally, for those with $z = 0$, all prices offered in equilibrium are acceptable. Thus, they exit upon encountering a seller, *i.e.* at rate μ. At the same time, this population of expired buyers is replenished by the flow of buyers whose grace period has just expired, $H'(0)$:

$$H'(0) = \mu H(0). \tag{20}$$

3.3 Seller's Problem with Exploding Offers

A seller who offers price $p = R(z)$ can be described as *targeting* buyers in state z, though all buyers with less time remaining will also accept. Thus, the expected profit from targeting them is represented as follows:

$$\pi(z) = \frac{H(z)}{H(T)}(R(z) - c). \tag{21}$$

Since the measure of buyers and Bellman equation (and hence reservation values) are continuously differentiable, the expected profit function is also continuously differentiable. If multiple prices produce the same maximal expected profit, sellers can randomize over these prices, represented in the cumulative price distribution $F(p)$.

3.4 Equilibrium Definition with Exploding Offers

An equilibrium in the exploding setting consists of seller profit π, reservation price function $R(z)$, the measure of buyers $H(z)$, and the distribution of sellers' offered prices $F(p)$, such that:

1. All prices in the support of F produce the same maximal profit π, while all other prices produce no more than π.
2. $H(z)$ satisfies the steady state conditions in Eqs. 18 through 20.
3. $R(z)$ maximizes the utility of a buyer of type z, given $F(p)$.

3.5 Equilibrium Characterization with Exploding Offers

Equilibria in the exploding setting can be categorized by two features. First, the price distribution can be *degenerate*, where all sellers offer the same price, or *dispersed*, where a variety of prices are offered. If exactly two prices are offered, we refer to this dispersion as *bimodal*.

Second, in the exploding setting, there is a critical state $Z^* \in [0, T]$ such that $R(Z^*)$ is the lowest late price offered, while $R(0)$ is the highest late price offered. In a *late equilibrium*, buyers with $z > Z^*$ have too low a reservation price, making it unprofitable for sellers to target them; these buyers reject all offers until at least time Z^*. Even buyers with $z < Z^*$ will reject some offers, preferring to continue searching.

An *early equilibrium* adds an atom α^* at the early price $R(T)$, indicating that a fraction of sellers are targeting early buyers. This early price is almost always strictly less than $R(Z^*)$, leaving a gap in the support of the price offer distribution. For $z \in (Z^*, T)$, the relative density of buyers $H(z)$ is initially falling faster than the reservation price rises, making them unprofitable to target.

Below, we first present the possible equilibria, then demonstrate that these are the only ones that can occur (in Proposition 2, which also provides the process of constructing this equilibrium), and show that only one will occur for a given set of parameters (in Proposition 3). First, the reservation price at the deadline is always the highest any buyer is willing to pay, and its solution depends on whether there is a continuous portion of the distribution ($Z^* > 0$) or not:

$$R(0) = \begin{cases} x + \frac{d}{\rho} \cdot \frac{\rho + \alpha^* \mu e^{-(\alpha^* \mu + \rho)T}}{\rho + \alpha^* \mu} & \text{if } Z^* = 0 \\ x + \frac{d}{\mu} & \text{if } Z^* > 0. \end{cases} \tag{22}$$

During the grace period, the reservation price for an exploding offer depends on whether they fall in the late price range $z < Z^*$ or not:

$$R(z) = \begin{cases} x + \frac{d}{\mu} e^{-\frac{2\mu}{\rho}\left(1 - e^{-\frac{\rho z}{2}}\right)} & \text{if } z \in (0, Z^*] \\ x + (R(Z^*) - x)\, \frac{\alpha^* \mu + \rho e^{(\alpha^* \mu + \rho)(T - z)}}{\alpha^* \mu + \rho e^{(\alpha^* \mu + \rho)(T - Z^*)}} & \text{if } z \in (Z^*, T]. \end{cases} \tag{23}$$

In equilibrium, buyers will only encounter late prices between $R(Z^*)$ and $R(0)$ (and possibly the early price $R(T)$). Yet we can still compute what buyers would be *willing* to pay at any point in their search process. Sellers can then consider (but reject, in equilibrium) the option of making offers that would target these untargeted buyers.

Two properties of reservation prices in the late range are worth emphasizing. First, these reservation prices always cover the cost of production x. Any additional willingness to pay comes from the impending costs of searching after expiration, d, which is moderated by the arrival rate μ and exacerbated by the impatience of buyers ρ. Effectively, the search friction allows sellers to partially extract surplus, based on the idiosyncratic time until the deadline.

Second, we find that $R'(z) < 0$ and $R''(z) > 0$. That is, buyers accept higher prices (and the increase becomes more pronounced) as their deadline approaches.[13] This acceleration in

---

[13] Discounting is one essential ingredient for this result. The present value of the penalty, $de^{-\rho z}$ is bigger and grows faster as the deadline approaches. At the same time, this is more than just a mechanical effect of discounting, because the reservation price solution also anticipates the buyer's own willingness to accept more offers as time runs out.

reservation prices is directly proportional to the search costs incurred after expiration.

Next, we report the solution for the equilibrium distribution of seller asking prices:

$$F(p) = \begin{cases} 0 & \text{if } p < R(T) \\ \alpha^* & \text{if } R(T) < p < R(Z^*) \\ 1 - \frac{\rho}{2\mu}\left(1 - \ln \frac{\mu(p-x)}{d}\right) & \text{if } R(Z^*) < p < R(0) \\ 1 & \text{if } p \geq R(0). \end{cases} \tag{24}$$

This solution can also be reframed as the distribution of offers targeting a specific type of buyer, $F(R(z))$. This distribution generates positive masses targeting up to three different types: a mass of α^* at the early price $R(T)$, a mass of $e^{-\frac{\rho Z^*}{2}} - \frac{\rho}{2\mu} - \alpha^*$ at the lowest late price $R(Z^*)$, and a mass of $\frac{\rho}{2\mu}$ at the highest late price $R(0)$. In the continuous, late range of the distribution (the third case of Eq. 24), we see that $F''(p) < 0$, meaning that higher prices are relatively less frequent.

The population of buyers in equilibrium is:

$$H(z) = \begin{cases} H(Z^*) e^{-\frac{2\mu}{\rho}\left(e^{-\frac{\rho z}{2}} - e^{-\frac{\rho Z^*}{2}}\right)} & \text{if } 0 \leq z < Z^* \\ H(Z^*) + \delta(z - Z^*) & \text{if } Z^* \leq z \leq T \text{ and } \alpha^* = 0 \\ H(Z^*) + \frac{\delta}{\alpha^* \mu}\left(e^{\alpha^* \mu (z-T)} - e^{\alpha^* \mu (Z^* - T)}\right) & \text{if } Z^* \leq z \leq T \text{ and } \alpha^* > 0, \end{cases} \tag{25}$$

where

$$H(Z^*) = \frac{\delta}{\mu} e^{-\alpha^* \mu (T - Z^*) + \frac{\rho Z^*}{2}}. \tag{26}$$

This solution accounts for the rate at which buyers make purchases, $\mu F(R(z))$, across all possible types.

The only remaining equilibrium variables are the critical time Z^* and the atom α^*. The solution for Z^* ensures that the reservation price solution in Eq. 23 (and hence, the buyer's Bellman equations) is continuous at $z = Z^*$. Effectively, this requires that early buyers correctly anticipate the expected benefit of search, though this is an equilibrium condition, not a choice by buyers. For this, we define:

$$\zeta(Z, \alpha) \equiv e^{-\frac{2\mu}{\rho}\left(1 - e^{-\frac{\rho Z}{2}}\right)} \left(\frac{\alpha\mu \, e^{(\alpha\mu + \rho)(Z-T)} + \rho}{\alpha\mu + \rho} e^{-\frac{\rho Z}{2}} - \frac{\rho}{\mu} \right). \tag{27}$$

The size of the atom α^* is determined by the relative profitability of offering $R(T)$ compared to $R(Z)$, which is:

$$\phi(Z,\alpha) \equiv (\alpha\mu+\rho)\left(e^{\alpha\mu(T-Z)}-1\right)-\alpha\rho\left(e^{(T-Z)(\alpha\mu+\rho)}-1\right)e^{\frac{\rho Z}{2}}. \tag{28}$$

Thus, when $\phi(Z^*,\alpha^*)=0$, these prices are equally profitable, but when $\phi(0,\alpha)>0$ for all $\alpha\in[0,1]$, only $R(T)$ will be offered.

Proposition 2. *In the exploding setting, a solution $R(z)$, $F(p)$, $H(z)$, and $G(z)$ is an equilibrium if and only if it satisfies Eqs. 22 through 28 with one of the following cases:*

- *(late degenerate) $Z^*=\alpha^*=0$ with $\zeta(0,0)\leq 0$*
- *(late dispersed) $Z^*\in(0,T]$ and $\alpha^*=0$ with $\zeta(Z^*,0)=0$*
- *(early dispersed) $Z^*\in(0,T]$ and $\alpha^*\in(0,1)$ with $\zeta(Z^*,\alpha^*)=0$ and $\phi(Z^*,\alpha^*)=0$*
- *(early bimodal) $Z^*=0$ and $\alpha^*\in(0,1)$ with $\zeta(0,\alpha^*)\geq 0$ and $\phi(0,\alpha^*)=0$*
- *(early degenerate) $Z^*=0$ and $\alpha^*=1$ with $\phi(0,1)\leq 0$.*

Which equilibrium type occurs depends on parameter values ρ, μ, and T. Increases in the utility parameters x and d will raise price levels or slopes, but do not affect the equilibrium type because they proportionally affect all prices. It is worth noting each equilibrium type coincides with the one in the neighboring row, in the limit.

When a late degenerate equilibrium occurs, sellers offer $R(0)=x+\frac{d}{\rho}$. Indeed, $\zeta(0,0)\leq 0$ indicates that earlier buyers are not willing to pay enough to warrant targeting them. When an early degenerate equilibrium occurs, sellers offer $R(T)=x+\frac{d}{\mu}e^{-(\rho+\mu)T}$. Even though later buyers are willing to pay more, $\phi(0,1)\leq 0$ indicates that there are too few of them to make targeting them profitable.

This equilibrium is unique under any set of parameters: only one of the five cases can occur, and only one equilibrium in that case can occur.

Proposition 3. *The equilibrium (Z^*,α^*) pair is unique.*

The equilibria can be partitioned by μ in relation to other parameters. For instance, the late degenerate equilibrium occurs only if $\mu\leq\rho$; that is, quotes are received infrequently. The late dispersed equilibrium occurs only if $\rho<\mu\leq\rho\ e^{\frac{\rho T}{2}}$, while an early equilibrium occurs if $\mu>\rho\ e^{\frac{\rho T}{2}}$. We can also partition among the early equilibria, but the conditions are

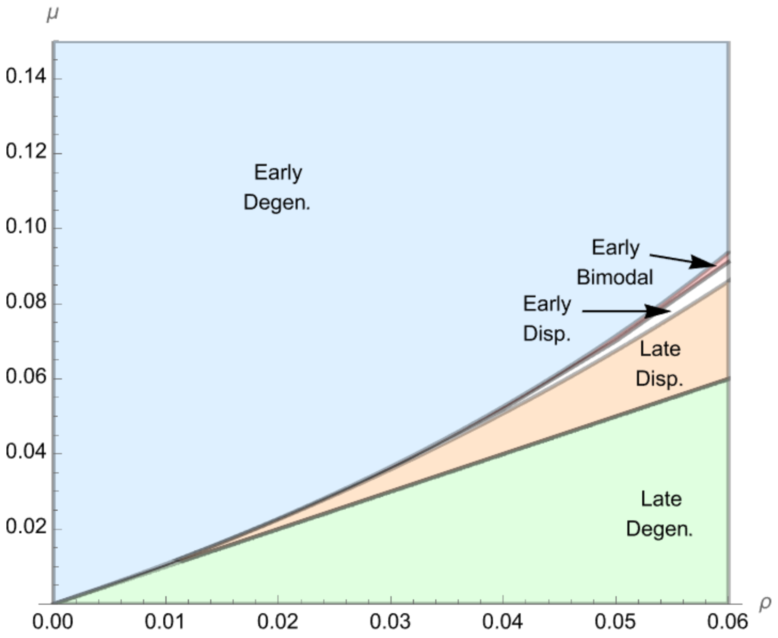


Figure 2: **Equilibrium Regions in the Exploding Setting:** Each region indicates the type of equilibrium that occurs for combinations of ρ and μ, holding $T = 12$.

more cumbersome. For example, the early degenerate occurs if and only if $(\mu+\rho)\left(e^{\mu T}-1\right) \geq \rho\left(e^{T(\mu+\rho)}-1\right)$, which requires a larger μ than any of the other equilibria.

In Figure 2, we illustrate how various parameters determine the type of equilibrium. There is no overlap; each equilibrium is unique. Also, each type of equilibrium occurs across a generic parameter space, not simply for knife-edge conditions.

It is surprising that a degenerate equilibrium does not always exist, perhaps as one of multiple equilibria. In many sequential search models, if sellers only offer a monopoly price (such as $R(0)$), buyers have no reason to wait for anything else. Here, even if all sellers offer $R(0)$, buyers are unwilling to accept it until their deadline. This leaves buyers in their grace period as potential targets for other sellers, and when $\zeta(0,0) > 0$ (i.e. $\mu > \rho$), they are too profitable to pass up. Thus, the late degenerate equilibrium does not always exist, but only when quotes arrive slower than the discount rate, $\mu \leq \rho$. Similar reasoning drives all of the uniqueness result. While buyers are ex-ante identical when they enter the market, they differ ex-post as some remain in the market longer than others. Those differences pin down the unique pricing strategy.[14]

---

[14] It is also noteworthy that, by Proposition 2, no equilibrium can exist where sellers target buyers with types $z \in [Z, T]$ while ignoring those $z \in [0, z)$. This is because buyers' reservation prices accelerate near the

In Figure 3, we illustrate $R(z)$, $H(z)$, $F(p)$, and $\pi(z)$ for each possible equilibrium. Across the rows, we increase the frequency of quotes (holding all other parameters fixed), thereby reducing the search friction.

The left panels of Figure 3 show that buyers' willingness to pay rises as they get closer to the deadline and the price increase becomes more pronounced as the grace period expires. Also, these reservation prices shift downward as μ increases, reducing sellers monopoly power.

As we compare across equilibria, the arrival rate affects who the sellers target (seen in the center right panels). When buyers get very few quotes, all sellers charge the highest price $R(0)$, and buyers wait until after expiration to accept it (Panel A). As more quotes become available, price dispersion emerges, with sellers willing to target a continuous range of earlier (but still the most desperate) buyers, some of whom will successfully purchase before their deadline (Panel B).

With even more plentiful quotes, a mass of sellers will offer the early price $R(T)$. At first, this includes a continuous distribution from $R(Z^*)$ to $R(0)$, meaning that the remaining sellers are only targeting those closest to their deadline (Panel C). As the arrival rate increases, this late price range shrinks, so that eventually sellers are either pricing for those who have hit their deadline, or for those who have just entered (Panel D).

Once the arrival rate is large enough, only the early price $R(T)$ is offered; no buyer passes up an offer (Panel E). Indeed, more frequent quotes push sellers to more heavily target early buyers. It is thus unsurprising that profits are falling as μ increases.

It is noteworthy that price dispersion always occurs (*i.e.* for any μ) in the recallable setting, but only for intermediate values of μ in the exploding setting. In either setting, though, more frequent quotes will compress the price distribution towards lower prices. We provide a thorough comparison of the resulting prices and welfare in Section 5 below.

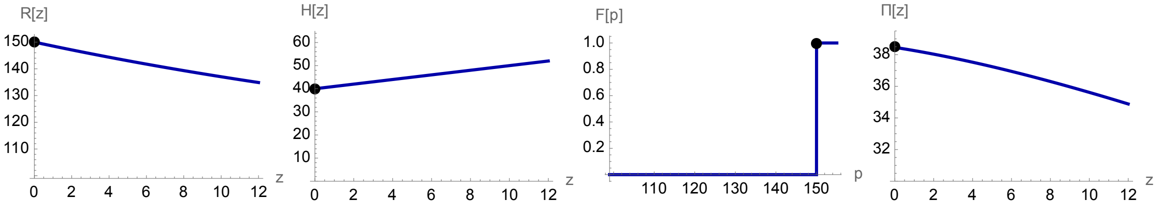
4 Search with Endogenous Recall

The preceding sections commit all sellers to follow the same recall policy. This might be reasonable in certain situations. For instance, if the item being sold cannot be easily reproduced, the seller cannot guarantee its future availability and thus would not be able to honor past quotes, consistent with Section 3. On the other hand, a government might treat

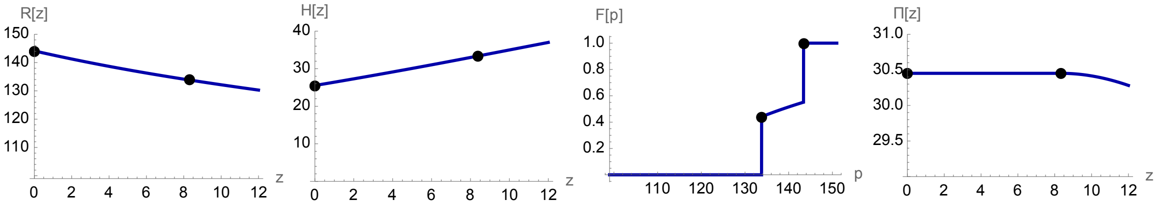
---

deadline — which would go untargeted in such an equilibrium. Thus, the willingness to pay at $Z - \epsilon$ rises faster than the decline in buyer population, so sellers would deviate to offer higher prices.

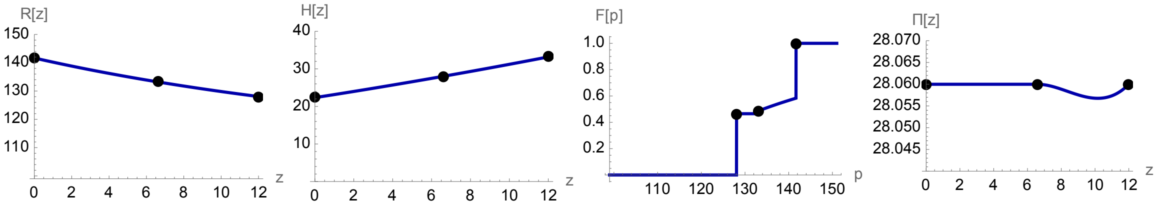
A. Late Degenerate ($\mu = 0.025$)



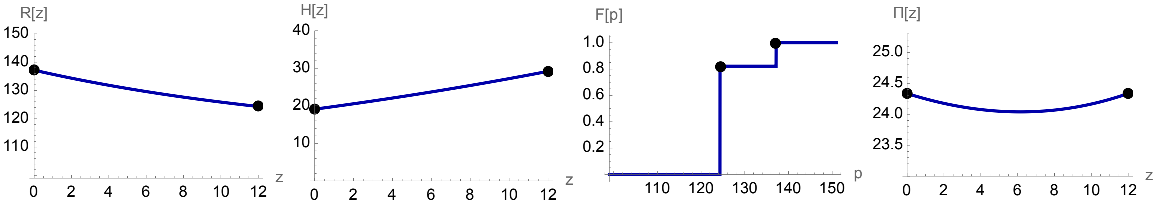
B. Late Dispersed ($\mu = 0.034$)



C. Early Dispersed ($\mu = 0.036$)



D. Early Bimodal ($\mu = 0.0365$)



E. Early Degenerate ($\mu = 0.10$)

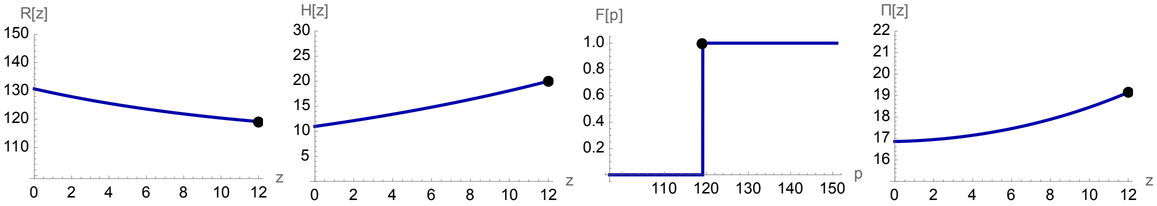


Figure 3: **Equilibrium Solution in the Exploding Setting:** Reservation prices (left), population of buyers (left center), price offer distribution (right center), and profits (right), varying the arrival rate μ, while holding other parameters at $x = c = 100$, $d = 1.5$, $\rho = 0.03$, and $T = 12$.

any written quote as a contract,[15] requiring sellers to honor them consistent with Section 2. But what if sellers had full freedom to choose their recall policy as they issue a price quote, informing the customer whether the offer must be used immediately, or if the price remains valid for a fixed time period?

In this *endogenous setting*, we allow sellers to quote both a price and a recall policy upon meeting a buyer. The resulting model weaves together the previous two sections. Fraction α of sellers offer an exploding offer drawn from c.d.f. $\bar{F}(p)$, which must be used immediately or lost. The remaining $1 - \alpha$ of sellers offer a recallable offer drawn from $\hat{F}(p)$, under which the seller promises to honor the quote for T periods.

The solution for recallable prices proceeds as in Section 2. For exploding prices, the combination of time remaining z and best recallable quote q create a complicated state space; expected profit would require determining a reservation price for exploding offers $R(z, q)$ and computing the measure $h(z, q)$ of the buyers willing to accept a given price. Fortunately, we are able to use a simpler approach to construct an equilibrium.[16] We compute the reservation price $R(z)$ for someone who has no recallable offers in hand at time z, and act as if all buyers with z or less time would accept such an offer, even those with recallable quotes. Since this is an overestimate of the true reservation price for an exploding offer, the resulting expected profit, $\frac{H(z)}{H(T)} \cdot (R(z) - c)$, is an upper bound on the more-complicated true profit, but we show that this upper bound is strictly maximized at $z = T$. Moreover, the upper bound coincides with the true profit at $z = T$ because no one has any quotes at time T. We further show that all buyers prefer $R(T)$ over any recallable quote. Thus, $R(T)$ is the only exploding price potentially offered in our equilibria.

We proceed by constructing the upper bound, then define and characterize the resulting equilibria. We show that when quotes arrive infrequently, the market only provides recallable quotes just as in Section 2. When quotes arrive frequently, the market only offers a single exploding offer, identical to the early degenerate equilibrium in Section 3. For intermediate rates of quote arrival, we obtain a new type of equilibrium with dispersed recallable prices and a single exploding offer. We provide precise bounds on those arrival rates in what follows.

---

[15] Price transparency laws or truth-in-lending disclosures are potential examples.

[16] Proposition 4 establishes that our candidate solution is an equilibrium, but does not claim uniqueness, though we still conjecture that it holds. We have found no counter examples, but the equilibrium conditions are too dense to allow proof.

4.1 Buyers with Endogenous Recall

A buyer in this setting must decide how to proceed after receiving any offer. This is greatly simplified since, as in Lemma 1, a buyer who receives a recallable price $q \geq x$ will always wait before acting on it, choosing the best recallable price at expiration. Even so, the buyer may receive an exploding offer in the interim and could act on it instead. Thus, we must derive what a buyer is willing to pay for an exploding offer.

Consider a buyer who has reached expiration with no quotes, whose expected utility from continued search is denoted as U. Since these are the most desperate, any price offered in equilibrium is immediately accepted, whether recallable or exploding. This results in Bellman equation:

$$\rho\, U = -d + \mu \left(x - \alpha \int_{-\infty}^{\bar{p}} p \cdot d\bar{F}(p) - (1-\alpha) \int_{-\infty}^{\bar{p}} p \cdot d\hat{F}(p) - U \right). \tag{29}$$

The buyer incurs the flow cost d, but if any offer is obtained (at rate μ) it is accepted, obtaining utility x minus the average price offered across the two types, and abandons the value of further search U.

Next, consider a buyer's expected value derived only from exploding offers during the grace period, denoted $V(z)$.

$$\rho V(z) = -V'(z) + \alpha\mu \int_{-\infty}^{\bar{p}} \max\left\{(x - p - V(z)), 0\right\} \cdot d\bar{F}(p). \tag{30}$$

This is almost identical to Equation 16 for the exploding setting, except that only fraction α of offers are exploding.

We now construct the expected utility for a buyer once she reaches expiration. Upon reaching expiration, with probability $e^{-\mu T}$ the buyer will not have received any quotes over the T periods. With probability $1 - e^{-\alpha\mu T}$, the buyer received an acceptable exploding offer at some point and thus left the market before expiration. Otherwise (with probability $e^{-\alpha\mu T} - e^{-\mu T}$), the buyer received at least one late price (and no exploding offers) during the grace period, with the best late price being distributed with density $\mu T e^{-\mu T \hat{F}(p)} \hat{F}'(p)$. This means the expected utility at expiration is:

$$V(0) = \frac{e^{-\mu T} U + \left(e^{-\alpha\mu T} - e^{-\mu T}\right) \int_{-\infty}^{\bar{p}} (x-p) \mu T e^{-\mu T \hat{F}(p)} \hat{F}'(p) dp}{e^{-\alpha\mu T}}. \tag{31}$$

This is divided by $e^{-\alpha\mu T}$ because it is conditional on not having received an acceptable exploding offer, which would have prevented reaching expiration.

This construction of $V(z)$ presents a pessimistic view of the value of search because it neglects whether a buyer at time z holds any recallable quotes or which ones. Rather, the buyer pretends as if she has no quotes currently, though anticipating the possibility of having recallable quotes at expiration. Thus, $R(z) = x - V(z)$ is an upper bound on what a buyer with z time remaining is willing to pay for an exploding offer, since some of those buyers may hold better quotes and would insist on a lower price to cut their search short. Even so, newly-entered buyers necessarily have no quotes, so $R(T)$ accurately reflects their maximum willingness to pay.

4.2 Sellers with Endogenous Recall

Upon encountering a buyer, a seller chooses what price to quote and whether to make it recallable for T periods. If it is a recallable quote, the seller knows that the buyer will not act on it until expiration, and in the meantime could be displaced by lower recallable offers or an accepted exploding offer. Expected profits Π_r must thus discount for the time delay and adjust for the probability of acceptance, which is the same as depicted in Eq. 5 for the recallable setting.

If making an exploding offer $R(z)$, sellers must anticipate what fraction of buyers would accept such a price. In the recallable setting, the measure of buyers in state z was computed using Eq. 4. This is still accurate under endogenous recall; it merely neglects to track the best quote the buyers hold. Since $R(z)$ is an upper bound on the willingness to pay, profit from an exploding offer meant for type z is bounded above by:

$$\Pi_e(z) = \frac{H(z)}{H(T)}(R(z) - c). \tag{32}$$

Even though $\Pi_e(z)$ is an over-estimate of expected profit from an exploding offer, we later show that T always maximizes $\Pi_e(z)$. Moreover, $\Pi_e(T)$ is the exact profit from offering $R(T)$, so sellers only need to compare $\Pi_e(T)$ versus Π_r to decide on the type of offer. If equal, sellers can randomize over the two types.

This profit comparison is particularly simple for an exploding offer $R(0)$ versus a recallable offer $\bar{p}$; indeed, these prices are necessarily the same, targeting those who reach expiration with no other quotes.

Lemma 2. *The recallable quote $\bar{p}$ always generates strictly more profit than the exploding offer $R(0)$.*

Intuitively, the exploding offer $R(0)$ would only be accepted after expiration. The recallable offer $\bar{p}$ would be accepted by those same buyers, plus the ones who received $\bar{p}$ during their grace period but received no better offers later. This larger volume of sales at the same price ensures strictly higher profit.

4.3 Equilibrium Definition with Endogenous Recall

An equilibrium in the endogenous setting consists of seller profit π, a reservation price function $R(z)$, the measure of buyers $H(z)$, the fraction of exploding offers α, the distribution of exploding offers $\bar{F}(p)$, and the distribution of recallable quotes $\hat{F}(p)$, such that:

1. If $\alpha > 0$, all exploding prices in the support of $\bar{F}$ produce the same maximal profit π.
2. If $\alpha < 1$, all recallable prices in the support of $\hat{F}$ produce the same maximal profit π.
3. Any recallable or exploding offers not offered produce no more profit than π.
4. $H(z)$ satisfies the steady state conditions in Eqs. 3 and 4.
5. $R(z)$ defines the utility maximizing reservation price on exploding offers for a buyer in state z and no recallable offers, per Eqs. 29 through 31, given $\bar{F}$ and $\hat{F}$.
6. If offered, exploding offer $R(T)$ will be accepted by any buyer, including those with the lowest recallable offer $\underline{p}$ in hand.

This definition merges requirements for the recallable and exploding settings. Any prices offered must be equally profitable, and any not offered must be weakly less profitable. The steady state population is adopted from the exploding setting as it is more general. The final requirement verifies that all buyers accept the early price when offered.

4.4 Equilibrium Characterization with Endogenous Recall

This setting of endogenous recall generates stark results. When arrival rates are low, only late prices are offered, generating the same dispersed prices as the recallable setting. When arrival rates are high, all sellers offer a single exploding early price, which coincides with the early degenerate equilibrium in the exploding setting. For moderate arrival rates, however, we obtain a variation of the early dispersed equilibrium: sellers offer a range of recallable offers along with a single exploding offer, $R(T)$. We now formally report this equilibrium solution.

Since there is at most one price offered as an exploding offer $\ell < \underline{p}$, we note that $\bar{F}(p)$ is degenerate at ℓ and we define:

$$F(p) = \begin{cases} (1-\alpha)\hat{F}(p) & \text{if } p \geq \underline{p} \\ \alpha & \text{if } p \geq \ell \\ 0 & \text{if } p < \ell \end{cases}. \tag{33}$$

This returns us to using the same notation as in Section 2. Moreover, the (exploding) reservation price for a person in state z and no quotes becomes:

$$R(z) = x + \frac{d\left(\alpha\mu + \rho e^{(\alpha\mu+\rho)(T-z)}\right)}{\mu\left(\alpha\mu+\rho\right)} \Bigg/ \left(\alpha + \frac{\left(1-e^{-\rho T}\right)e^{\mu(1-\alpha)T}\left(\alpha\mu + \rho e^{T(\alpha\mu+\rho)}\right)}{\rho T(\alpha\mu+\rho)\left(\frac{\mu(e^{-\rho T}-1)(\mu(1-\alpha)T+1)+\rho(\alpha\mu+\rho)T}{Q} - 1\right)}\right) \tag{34}$$

where

$$Q \equiv \rho T(\mu+\rho) + \left(\rho + \mu\left(1-e^{-\rho T}\right)\right)\ln\left(1 - \frac{\rho\left(1-e^{\mu(\alpha-1)T}\right)}{\rho+\mu\left(1-e^{-\rho T}\right)}\right). \tag{35}$$

This allows us to characterize the resulting equilibrium as follows:

Proposition 4. *In the endogenous setting, assuming* $\rho < \frac{\mu}{\mu T + e^{-\mu T}}$*, a solution* $R(z)$*,* $F(p)$*, and* Π *is an equilibrium if it satisfies Eqs. 7 through 11 and Eq. 34 with one of the following cases:*

- *(late dispersed)* $\alpha^* = 0$ *if* $\Pi \geq R(T) - x$
- *(early dispersed)* $\alpha^* \in (0,1)$ *with* $\ell = R(T)$ *if* $\Pi = R(T) - x$
- *(early degenerate)* $\alpha^* = 1$ *with* $\ell = R(T)$ *if* $\Pi \leq R(T) - x$

Figure 4 illustrates the circumstances under which each of these equilibria occur. For a given ρ, one can numerically solve for a $\underline{\mu}$ below which the first case occurs, and a $\bar{\mu}$ above which the third case occurs, with the second case occurring for $\underline{\mu} < \mu < \bar{\mu}$. The condition imposed on ρ in Proposition 4 is illustrated by the dashed line in Figure 4, satisfying the condition with any (ρ, μ) pair above the dashed line. When it holds, the exploding offer $R(T)$

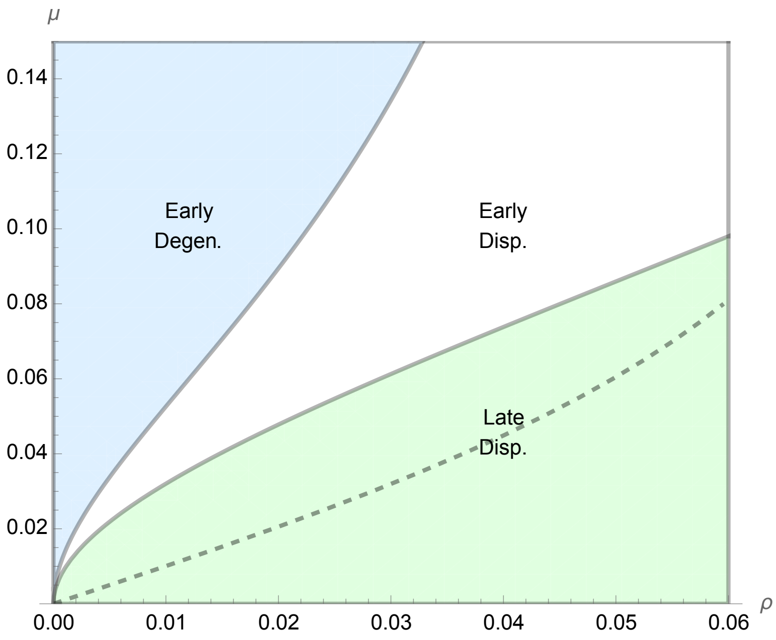


Figure 4: **Equilibrium Regions with Endogenous Recall:** Each region indicates the type of equilibrium that occurs for combinations of ρ and μ, holding $T = 12$. The dashed line indicates where $\rho = \mu / \left(\mu T + e^{-\mu T}\right)$.

yields higher profit than any other exploding offer, simplifying the comparison of exploding offers versus recallable offers.[17]

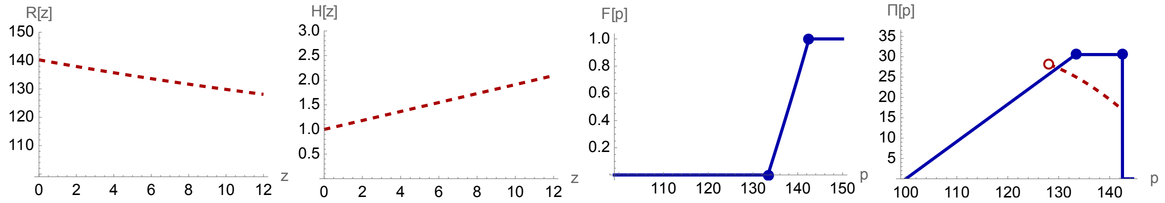
Figure 5 illustrates the equilibrium solution for several sets of parameters, which are best understood in comparison to the benchmark recallable and exploding settings.

First, the late dispersed equilibrium is identical to the recallable setting presented in Section 2. Expected profits from recallable prices are reported in the solid line of the right-most graph of Figure 5.A. If there are no exploding offers, sellers offer the same set of late prices as if all prices were recallable, because in equilibrium, they are. In the endogenous setting, sellers contemplate deviation by offering an exploding price (with expected profits reported by the dashed line); but in a late dispersed equilibrium, they find such deviation less profitable. Even though buyers strictly prefer the early price (\$129 in the example) over all recallable prices, a seller earns an average of \$2 more by offering a higher recallable price

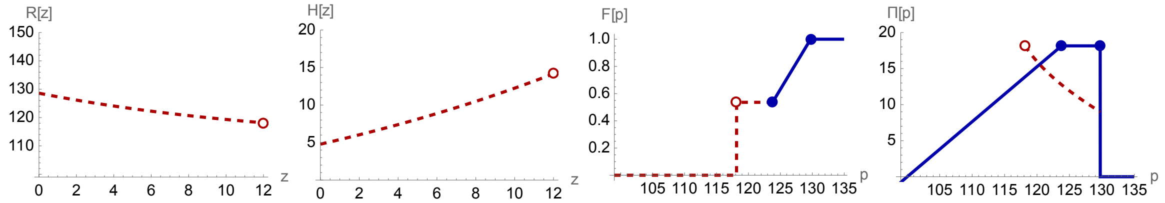
---

[17] This assumption is stronger than needed, however. If it does not hold (because participants are very impatient), one can compute the exploding offer $\ell = R(z^*)$ that uniquely maximizes profit. Even so, we have verified that, for the parameter values under the dashed line, $R(z^*)$ will still be less profitable than any recallable offer, maintaining the same late dispersed equilibrium.

A. Late Dispersed ($\mu = 0.05$)



B. Early Dispersed ($\mu = 0.08$)



C. Early Degenerate ($\mu = 0.15$)

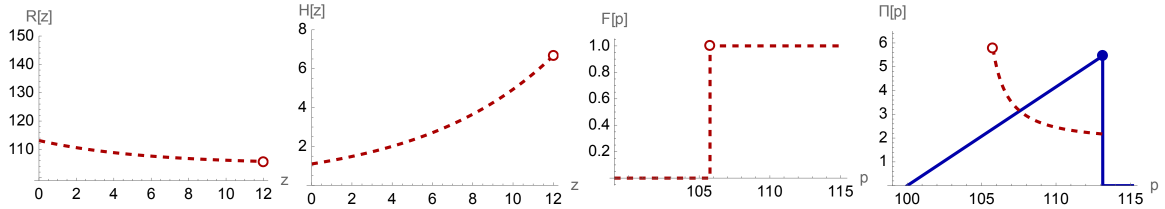


Figure 5: **Equilibrium Solution in the Endogenous Setting:** Reservation prices (left), population of buyers (left center), price offer distribution (right center), and profits (right), varying the arrival rate μ, while holding other parameters at $x = c = 100$, $d = 1.5$, $\rho = 0.03$, and $T = 12$. Solid lines indicate recallable offer, while dashed lines indicate exploding offers. Solid dots indicate the (recallable) lowest and highest late prices, while open dots indicate the (exploding) early price.

while hoping to be the best price at the end of the grace period. As seen in Figure 4, such an equilibrium can be sustained as long as price offers are not too frequent. This is because recipients of a recallable offer will rarely have more than one offer to compare, resulting in limited competition and more profit than possible from the exploding early price.

Second, the early degenerate case (shown in Figure 5.C) exactly coincides with the exploding setting presented in Section 3, offering only the exploding early price, $\ell = R(T)$. Other exploding offers are not as profitable for the same reason as in Section 3, but it turns out that deviation to a recallable offer is more binding (*i.e.* the early degenerate region is smaller in Figure 4 than in Figure 2). In both settings, sellers contemplate deviating to other exploding offers, including $R(0)$; but $R(0)$ is more profitable when offered with recall because it captures not only those who have expired but also those who receive it during their grace period but never find the early price.[18] Even so, in the early degenerate equilibrium, early prices are offered with sufficiently frequency that any recallable offer is likely to be undercut before it is exercised, discouraging sellers from deviating to recallable offers.

These extreme equilibria highlight the advantages of each type of offer. Compared to exploding offers, recallable offers can capture additional buyers that, in the absence of better alternatives, will eventually make a purchase. This is most important for the highest price offered in the market. Yet recallable offers also lead to delayed purchase (discounting that future profit) and introduce the potential to be undercut at the end, which creates the best opportunities for exploding offers at a low price that everyone is willing to accept.

These forces are exactly balanced in the early dispersed equilibrium (shown in Figure 5.B). Some sellers offer recallable prices, distributed as they would have been under the late dispersed equilibrium, except shifted lower because of the potential for early prices increasing the value of search. Moreover, $\underline{p}$ is pinned down where $F(\underline{p}) = \alpha$ (in the figure, $F(123.8) = 0.536$), rather than $F(\underline{p}) = 0$ as in the late dispersed solution.

The possibility of the early price creates added nuance between the sequential and simultaneous search behavior. Offers in the late region are evaluated by buyers simultaneously at the deadline, disciplined by competition to beat other offers at that time. Exploding offers are evaluated by buyers sequentially, being acted on immediately. These offers are never compared against another offer, but they are disciplined by competition against *prospective* offers through further search. Indeed, the willingness to continue search is the competitive force in the exploding setting, and is also the force that limits the highest price $\bar{p}$ in the recallable setting. In the early dispersed equilibrium in the endogenous setting, we see both forms of competition in operation.

It is worth noting the absence of a late degenerate or early bimodal equilibrium. Both would require a positive mass of sellers offering $\bar{p}$ (with recall, given Lemma 2). But this creates incentive to undercut one another, leading to dispersion in recallable prices. While we can rule out those behaviors, Proposition 4 does not claim uniqueness of equilibrium though we still conjecture that it holds. We have found no counter examples, but the equilibrium conditions are too dense to allow proof.[19]

Finally, note also that the early dispersed region is substantially larger than in Section 3. Indeed, if we set $\rho = 0.06$ in our example, any value of $\mu > 0.081$ leads to an early dispersed

---

[18] Formally, this is more profitable when $(\mu + \rho)\left(e^{\mu T} - 1\right) \geq \rho\left(e^{T(\mu+\rho)} - 1\right) + \frac{\mu^2\left(e^{\mu T} - e^{-\rho T}\right)}{\mu+\rho}$. Notably, this is a tighter constraint than faced in the recallable setting because of the addition of the fractional term.

[19] The Technical Appendix provides an interactive version of the graphs, allowing the reader to explore equilibrium under various parameters values.

equilibrium; the early degenerate never occurs. Thus, the early dispersed equilibrium could easily be the most relevant of the potential outcomes.

5 Welfare

In this section, we consider the welfare implications of search frictions, with a particular focus on who benefits from the ability to recall prices and when. We first define welfare metrics in this setting, and then explore how these are affected by parameter values.

5.1 Welfare metrics

This search market generates value for buyers and sellers, which we compute as the expected gain from trade per newly-entering buyer. When a buyer enters the market, her expected consumer surplus is precisely captured by $V(T) = x - R(T)$, averaging over the various quotes she may receive, accounting for her response to them, and discounting for the time elapsed and potential penalties after expiration. This is computed by applying the equilibrium solution for the reservation price: Eq. 23 for exploding offers and Eq. 34 for recallable offers. In the latter, α is set to 0 in the recallable setting, or is set to the equilibrium α for endogenous recall.

Total welfare is defined as the expected total gains from trade (to both buyer and seller, making transfers between them irrelevant), measured for each newly-entered buyer (to remain consistent with consumer surplus). When the buyer eventually obtains the good, the net welfare is the utility from consumption minus the cost of production, $x - c$, which is 0 by assumption.[20] The only other impact on total welfare are the expected costs in the penalty phase.

If the buyer had no market in which to search, she would receive no offers and would thus incur the cost d forever after the grace period, generating expected welfare $-\frac{d}{\rho}e^{-\rho T}$. In the first-best scenario, buyers are always able to find a seller and transact before their grace period expires, thus avoiding the penalty phase for a total welfare of 0. Therefore, the maximum potential gains[21] in this market would be: $\frac{d}{\rho}e^{-\rho T}$.

---

[20] If we allow $x > c$, consuming early in search would provide additional welfare gains since $x - c$ would be less discounted than when consumption occurs late. Even so, the results remain qualitatively similar

[21] One could rightly say that $x + \frac{d}{\rho}e^{-\rho T}$ is the entering consumer's true value from the good, since obtaining it avoids those penalties..

In a setting with recallable quotes (whether exogenous or endogenous), a buyer who receives any offer prior to expiration is guaranteed to avoid all penalties. If it is the early price, it is accepted immediately. If it is a late price, the buyer will accept it or a better offer at the deadline. Either way, anyone with a pre-expiration quote generates a total welfare of 0. This directs our attention to prospective buyers with no quotes and z periods remaining. We define the expected welfare from such a buyer as $W_r(z)$, as depicted in the following differential equations:

$$\begin{aligned} \rho W_r(0) &= -d + \mu(0 - W_r(0)) \qquad (36) \\ \rho W_r(z) &= -W_r'(z) + \mu(0 - W_r(z)). \qquad (37) \end{aligned}$$

If any offer is received after expiration, it is immediately accepted, generating a net change in welfare of $0 - W_r(0)$. While awaiting an offer, however, the buyer incurs a penalty flow d. During the grace period, no cost is incurred, but the remaining time is declining (as indicated by $-W_r'(z)$). Once a first offer arrives, the buyer escapes the penalty phase, so the net change in welfare is $0 - W_r(z)$. When solved and evaluated for the new entrant, we find an expected welfare of $W_r(T) = -\frac{d}{\rho+\mu}e^{-(\rho+\mu)T}$. When expressed relative to the maximum welfare gains possible in this market, we obtain:

$$W_r\% = \frac{\frac{d}{\rho}e^{-\rho T} - \frac{d}{\rho+\mu}e^{-(\rho+\mu)T}}{\frac{d}{\rho}e^{-\rho T}} = 1 - \frac{\rho}{\rho+\mu}e^{-\mu T}. \qquad (38)$$

Note that this computation is the same for the recallable or endogenous setting. Any offer received has the same welfare effect on avoiding further search costs, and the exact price paid is just a transfer between buyer and seller.

For the exploding setting, expected welfare $W_e(z)$ is derived from the following differential equations:

$$\begin{aligned} \rho W_e(0) &= -d + \mu(0 - W_e(0)) \qquad (39) \\ \rho W_e(z) &= -W_e'(z) + \mu F(R(z))(0 - W_e(z)). \qquad (40) \end{aligned}$$

The post-deadline welfare proceeds the same as with recall, but during the grace period, an offer above $R(z)$ is unacceptable, so the buyer only accepts at rate $\mu F(R(z))$. For the new entrant, this results in the solution:

$$W_e(T) = -\frac{d}{\rho+\mu} e^{\frac{\rho Z^*}{2} - \rho T - \alpha^* \mu (T - Z^*) - \frac{2\mu}{\rho}\left(1 - e^{-\frac{\rho Z^*}{2}}\right)}, \tag{41}$$

or expressed relative to the maximum welfare gains possible in this market,

$$W_e\% = 1 - \frac{\rho}{\rho+\mu} e^{\frac{\rho Z^*}{2} - \alpha^* \mu (T - Z^*) - \frac{2\mu}{\rho}\left(1 - e^{-\frac{\rho Z^*}{2}}\right)}. \tag{42}$$

Returning to consumer surplus, we can likewise evaluate the buyer's expected utility relative to the maximum welfare gains, obtaining:

$$CS\ \% = \frac{x - R(T) + \frac{d}{\rho} e^{-\rho T}}{\frac{d}{\rho} e^{-\rho T}} = 1 - \frac{R(T) - x}{d} \rho\, e^{-\mu T}. \tag{43}$$

Finally, producer surplus (per entrant) is necessarily the difference between expected welfare and expected utility: $W_i(T) - V(T)$, where $i = r$ or e. Of course, $\frac{d}{\rho} e^{-\rho T}$ is the most profit per entrant that the seller could hope to capture, so the fraction of potential welfare gains captured as producer surplus is:

$$PS\ \% = \frac{W_i(T) - V(T)}{\frac{d}{\rho} e^{-\rho T}} = \frac{W_i(T) - V(T)}{d} \rho\, e^{-\mu T}. \tag{44}$$

In all three versions of our model, the potential delay between quotes (on average, $1/\mu$) is the market friction. It forces some buyers to exhaust their grace period and incur costs, and also enables sellers to exercise probabilistic price discrimination. Thus, it is particularly important to consider how reducing this friction (with a higher μ) would impact participants, including a comparison across our models.

5.2 Total Welfare

First, we consider the response of total welfare to an increase in μ. We graph a parameterized example in the right panel of Figure 6. First consider settings with recallable offers, depicted with the dotted line. In such settings (including all equilibria of the endogenous setting), anyone who receives any offer before time T will avoid the penalty phase. Thus, it is not surprising that welfare is strictly increasing in μ, because fewer people reach expiration without an offer. In fact, by the time $\mu = 0.21$, this market achieves 99% of potential

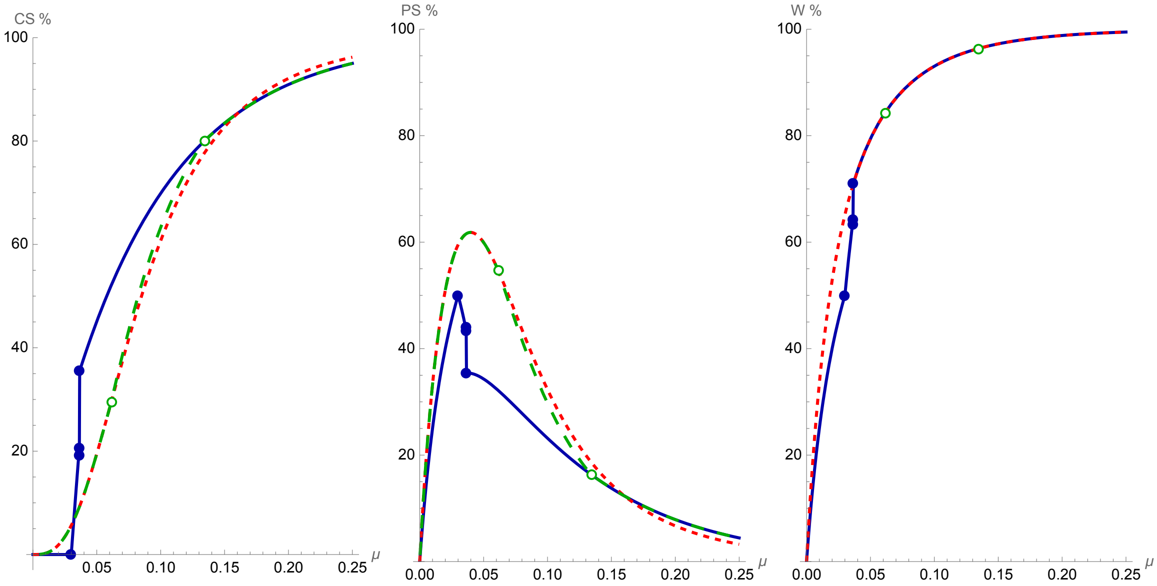


Figure 6: **Equilibrium welfare** in the Exploding (blue solid), Recallable (red dotted), and Endogenous (green dashed) settings, as a function of the arrival rate μ. Consumer surplus (left), producer surplus (center), and total welfare (right) are shown as a percentage of the first-best welfare gains if this were a frictionless market. All other parameters as in Figure 3. On each graph, solid dots indicate the transitions from (left to right) late degenerate, late dispersed, early dispersed, early bimodal, and early degenerate equilibria. Open dots indicate transitions between (left to right) late dispersed, early dispersed, and early degenerate equilibria under endogenous recall.

welfare. Inefficiency is only very pronounced when μ is fairly low (*e.g.* below 3ρ); that is, quotes are rather scarce compared to the impatience of buyers.

In the exploding setting, we obtain the welfare denoted by the solid line. Welfare is still strictly increasing in μ, but is lower than settings with recallable quotes — strictly lower when the exploding setting does not produce an early degenerate equilibrium. This is because the other forms of equilibria involve some buyers receiving a quote but not acting on it. This effectively wastes a scarce opportunity to help a buyer avoid the penalty phase. While socially inefficient, this is individually rational because the rejecting buyer sees the price as too high at that time in her search. The early degenerate equilibrium (occurring when $\mu > 0.05$ at the right-most solid dot) achieves the same welfare as recallable offers because all sellers offer the early price accepted by all buyers, achieving the same number of buyers who avoid penalties as with recallable offers. We formalize this result in the following claim.

Proposition 5. *All else equal, $W_r\% \geq W_e\%$, with strict inequality iff $\alpha^* < 1$ in the exploding setting.*

5.3 Consumer Welfare

Next, we turn to the consumer surplus, which is also increasing in μ, as depicted in the left panel of Figure 6. The most striking result here is that consumers can be better off in the exploding setting than the recallable or endogenous settings, which happens for intermediate values of μ.[22] That is, consumers can be strictly better off if sellers are not allowed to honor past quotes!

This is because sequential search without recall cultivates inter-temporal competition. As the search friction is reduced, sellers are driven to target earlier buyers, moving from the late degenerate equilibrium through the dispersed equilibria and eventually the early degenerate equilibrium (with transitions marked by closed dots). For moderate values of μ, this inter-temporal competition is more effective than the competition via comparing prices at the deadline with recallable quotes.

We formalize this result in the following proposition, proved in the Technical Appendix.

Proposition 6. *If $\rho T \left(\frac{\rho}{\mu\left(1-e^{\rho(-T)}\right)} + 1 \right) > 1$ and $\alpha^* = 1$ in an exploding setting but $\alpha^* < 1$ in an endogenous setting, then $CS_e\% > CS_r\%$.*

That is, if the exploding setting result in all firms offering the early price while the endogenous setting does not, then consumers are strictly better off without recall. Note that the first assumption on parameters is a sufficient condition, but not necessary. Indeed, we have yet to find parameters for which the rest of the claim does not hold.

As for the recallable versus endogenous settings, the former provides more consumer surplus only when μ is high. This ranking depends on whether the competition among late prices is a greater check on monopoly power than the early price in the endogenous setting.[23] Eventually, a higher rate of offers will make recallable offers dominate, though the two asymptotically approach each other in the limit, with either setting delivering all potential market gains to the consumers.

---

[22] Note that consumers have the same surplus under the recallable (dotted) or endogenous (dashed) settings when $\mu \leq 0.061$ because it is literally the same late dispersed equilibrium. Likewise, their surplus is the same in the exploding (solid) and endogenous (dashed) for $\mu \geq 0.135$ because both are in an early degenerate equilibrium.

[23] The early price in the early degenerate equilibrium is only competitive in the sense that it must appeal to buyers with the maximum search time remaining.

5.4 Producer Welfare

Finally, we examine the producer surplus depicted in the center panel of Figure 6. Some results carry over immediately: for instance, the recallable and endogenous settings are equally profitable in the late dispersed range (lower than the first open dot), but for larger μ, the two diverge in the opposite pattern of consumer surplus. But this is necessarily so because the total welfare is the same in these two settings at any μ. Also, the exploding and endogenous settings are equally profitable in the early degenerate range (above the second open dot).

Perhaps the most surprising result is that seller profit is maximized at a very low level of μ. For the exploding setting, profits are highest when $\mu = \rho = 0.03$, because this is the transition between late degenerate and late dispersed equilibria — anything higher begins to erode sellers' market power by offering lower prices, but anything lower makes a longer wait to obtain those profits. In either setting with recall, profits are maximized at $\mu = 0.04$, well inside the late dispersed equilibrium. The same tradeoff occurs here but more smoothly — a higher μ means more competition at expiration, but ensures that more buyers make a purchase then instead of being quote-less at expiration.

The formal results here are limited by the complicated solutions under recall, but we can prove the most surprising fact: that offering recall can be more profitable than a setting with only exploding offers. This comparison is most easily made when $\mu = \rho$, which is also a local maximum of producer surplus in the exploding setting. Of course, given the continuity of the equilibrium solution, if recall strictly outperforms exploding offers for sellers, the same comparison will hold for an open set of parameters around $\mu = \rho$.

Proposition 7. *In the exploding setting, $PS_e\%$ has a local maximum when $\mu = \rho$. Moreover, if $\rho T < 1$ and $\mu = \rho$, then $PS_r\% > PS_e\%$.*

This is proven in the Technical Appendix. The condition $\rho T < 1$ ensures a sufficiently short timeline that buyers will not have numerous recallable offers to compare at the deadline. When T is substantially larger than this, we have found counterexamples where the exploding setting generates more producer surplus than settings with recall.

The examination of welfare underscores the importance of a complete characterization of possible equilibria. While the ability for recall may seem like an unambiguous advantage to buyers, this is not necessarily true. Indeed, what flips the conventional wisdom is that the type of equilibrium will determine the targeting of buyers. The exploding setting can induce early buyers to pass up some high-priced offers intended for later buyers, which creates inefficiency (since some of those buyers will hit expiration before finding another offer), but yet can still generate higher utility than the recallable setting. In a similar vein, when offers are abundant (high μ), the endogenous setting is worse for buyers than the recallable setting, because the early price in the former is higher than most of the late prices in the latter.

6 Conclusion

We analyze the strategic pricing and recall decisions of sellers in settings where participants face time pressure. If quotes are required to be recallable, all prices are subject to comparison at the deadline, so sellers balance the risk of being undercut against higher markups as in simultaneous search. If quotes are required to explode, the offer distribution could include the early price that anyone would accept, though sellers can also charge higher prices that only the most desperate buyers accept.

With endogenous recall, all sellers opt for recallable prices when quotes are infrequent because buyers have few prices to compare at their deadline. When quotes are frequent, all sellers offer the same exploding early price, since any recallable offer would likely be undercut before the deadline. With moderate arrival rates, high recallable prices and a low exploding price are both offered.

The types of equilibria are pivotal in understanding welfare consequences. Indeed, it is surprising that recall can harm consumers. This always holds when quotes are somewhat infrequent, so recallable offers do not induce much competition, but are frequent enough that exploding offers include the early price. Even so, exploding prices are always inefficient when not all sellers offer the early price, because some of those who reject higher prices will reach expiration and incur penalties.

One virtue of our model is that search provides the consistent disciplining force throughout. Buyers' willingness to pay is endogenously determined by the value of continued search. Rather than fixing a number of quotes, buyers always have the option to gather more, even after crossing the deadline. Indeed, even the highest price is disciplined by buyers being able to seek another quote, rather than setting an exogenous outside option (as in Burdett and Judd, 1983; Coey, *et al*, 2020). This unified framework makes for a cleaner comparison across all three settings, where recallable quotes generate a price distribution similar to those in simultaneous search, but derived from a sequential setting and allows the possibility of the early price.

Our model also assumes homogeneity in buyer valuations, grace period length, penalties, and seller costs, which is an intentional modeling choice to isolate the pricing incentives created by the proximity to the impending deadline or by the best offer in-hand (when recall is possible), which differ ex-post across the buyers. These incentives would still be present after adding exogenous heterogeneity, though it would affect who sellers target. If the heterogeneous buyers have reservation price ranges that overlap, we expect that the equilibrium price distribution will be qualitatively similar to those depicted here. With wider heterogeneity, it could be that the market will ignore some less-populous types.

Our model can easily be adjusted to incorporate other plausible features of search on a deadline. Service or subscription contracts sometimes give new customers an introductory rate in exchange for committing to the service for a fixed period of time, after which the monthly price could be much higher. As the contract nears its conclusion, the customer becomes interested in shopping for a new contract. Of course, while the customer wants to avoid the rate increase after the contract expires, she is already committed to service with her current provider through the end of the contract, and would like to avoid early termination fees. These features mostly just change the interpretation of post-expiration penalties; the one modeling difference is that the flow of new buyers now endogenously arises from the rate at which buyers sign new contracts.

Another interesting application could be the job market for doctoral candidates. Universities, research institutions, or firms in the private sector might need to fill a vacancy. Candidates might look similar on paper, but may differ in their time looking for a job or might already have some offers. In such a market where firms hurry to fill a vacancy and candidates hurry to find a job, our model lays out optimal strategies for each side of the market and provides welfare consequences of those equilibrium strategies.

A Proofs

Proof of Lemma 1. Suppose a buyer with z periods remaining is offered price ℓ which is the lowest price in the support of $F(p)$. If she immediately accepts this price, she receives utility $x - \ell$.

If instead she continues her search, note that she already has a quote for the best price available, so she will not find a better price. Rather, she is merely delaying the purchase at that price. This "search" will give expected present utility of:

$$(x - \ell)e^{-\rho z}.$$

Note that immediate acceptance is strictly preferred if $\ell < x$. If $\ell > x$, the buyer will postpone the purchase until the deadline. The buyer will purchase at the deadline so long as $x - \ell > -\frac{d}{\rho}$; that is, the consumer surplus is greater than incurring the penalty forever. □

Proof of Proposition 1. The steady state conditions for a system of differential equations with the unique solution reported in $H(z)$. Likewise, the offered prices are equally profitable iff the derivative of Eq. 6 is zero; the resulting differential equation with boundary $F(\bar{p}) = 1$ also has the unique solution expressed in $F(p)$. Since any ℓ that would be accepted immediately is smaller than x, profits from offering it are weakly negative, so $\alpha^* = 0$. The solutions for $\bar{p}$, $\underline{p}$, and π simply follow from substituting $\alpha^* = 0$ into Eq. 10, Eq. 8, and (taking the limit as $\alpha \to 0$) Eq. 11, respectively.

Note that the profit π from offering a late price is strictly positive, unlike offering the early price. Any price greater than $\bar{p}$ will never be accepted, generating 0 profit, and any price below $\underline{p}$ will reduce revenue while still being accepted with the same frequency as $\underline{p}$, thus sacrificing expected profit.

Finally, we verify that any price offered in equilibrium is preferred over abstaining indefinitely. For a seller who has reached expiration, this would yield a present discounted utility of $-\frac{d}{\rho}$, whereas purchasing at price $\bar{p}$ would yield:

$$x - \bar{p} = -\frac{d}{\mu + \rho - \left(\frac{\rho+\mu(1-e^{-\rho T})}{T\rho}\right)\ln\left(1 - \frac{\rho(1-e^{-\mu T})}{\rho+\mu(1-e^{-\rho T})}\right)} > -\frac{d}{\rho}, \tag{45}$$

where the inequality holds because the term in the logarithm is less than one and hence the last term in the denominator is positive. □

Proof of Proposition 2. We start by translating the equilibrium conditions into functions of the reservation prices $R(z)$, using Eq. 17 to substitute for the value function and its derivatives.[24] The equilibrium conditions are equivalent to the following system of equations. When $Z > 0$, for $z \in [0, Z)$, we have:

$$\begin{aligned} R''(z) &= -(\rho + \mu F(R(z)))R'(z) && (46)\\ R'(0) &= -d && (47)\\ F(R(z)) &= \frac{R''(z)}{\mu R'(z)} - \frac{2R'(z)}{R(z) - c} && (48)\\ R'(0) &= -\mu(R(0) - c). && (49) \end{aligned}$$

Eqs. 46 and 48 are derived from buyer optimization (Eq. 16) and equal profits (Eq. 21), respectively, for interior values of z. Eqs. 47 and 49 ensure that buyer optimization and profits at $z = 0$ are continuous (*i.e.* $V(\epsilon)$ approaches $V(0)$ and $\Pi(\epsilon) = \Pi(0)$).

When $Z = 0$, there is no interior of the late price range; however, the post-expiration Bellman equation (15) can be translated and solved directly as:

$$R(0) = x + \frac{d + \alpha\mu(R(T) - x)}{\rho + \alpha\mu}. \tag{50}$$

Whether $Z = 0$ or $Z > 0$, in the range of $z \in [Z, T]$, buyer optimization translates to:

$$R'(z) = \rho(x - R(z)) + \alpha\mu(R(T) - R(z)). \tag{51}$$

This system of equations solves as follows. In the late region $z \in [0, Z^*]$, we substitute for $F(R(z))$ from Eq. 48 into Eq. 46 and get a second-order differential equation of the reservation prices:

$$\rho R'(z) + 2R''(z) + \frac{2R'(z)^2}{c - R(z)} = 0. \tag{52}$$

This differential equation has a unique solution, up to two constants which are pinned down using Eqs. 47 and 49 as boundary conditions at $z = 0$. This provides the equilibrium reservation prices.

In the early region $z \in (Z^*, T]$, the first-order differential Eq. 51 yields a unique solution

---

[24] For brevity, here we only report the translated equilibrium conditions used to solve for reservation prices and the price distribution; the steady state conditions on population and profit follow from reservation prices, and are relegated to the Technical Appendix, along with the algebraic manipulations used to obtain the following solutions.

as well, up to one constant determined by the boundary condition that $R(z)$ is continuous at Z^*. Indeed, the same Eq. 51 pins down the critical state Z^* by requiring that $R'(z)$ is also continuous at Z^* (relative to the Eq. 52 solution).

We then use these reservation price solutions to compute the distribution of buyers and expected profits. Finally, the atom α^* is pinned down by comparing the profits from targeting buyers with $R(Z^*)$ versus $R(T)$; if these can be equated for an $\alpha \in (0,1)$, it generates an early dispersed or bimodal equilibrium.

This translation process ensures that the $R(z)$ function in Eqs. 22 and 23 are equivalent to conditions 1 and 3 in the equilibrium definition, and demonstrates that all prices in the support of $F(p)$ in Eq. 24 are equally profitable. Indeed, in an early dispersed or bimodal equilibrium, $R(0)$ and $R(T)$ are equally profitable by construction.

We only need to verify that any price outside the support generates weakly lower profits. Of course, any price above $R(0)$ will be rejected by all buyers and hence generates zero profit. Any price below $R(T)$ will reduce revenue per sale without increasing the number of potential buyers, and hence is strictly not preferred. Last we inspect the impact of offering a price below $R(Z^*)$, taking as given $H(z)$ and $R(z)$. We proceed assuming that this is a dispersed equilibrium; the same logic applies in a degenerate equilibrium, after appropriately substituting the equations used.

We proceed by computing $\Pi(z,\alpha^*) = (R(z)-c)H(z)$ for $z \in (Z^*,T)$ from the proposed solution in Eqs. 23 and 25. The Technical Appendix provides algebraic manipulations.

In the case of a late dispersed equilibrium (with $\alpha^* = 0$), offering a price $R(z)$ where $z \in (Z^*,T)$ would generate profit:

$$\Pi(z) \equiv \frac{e^{\rho(Z^*-z)}\left(\mu(z-Z^*)+e^{\frac{\rho Z^*}{2}}\right)}{\mu\left(\mu(T-Z^*)+e^{\frac{\rho Z^*}{2}}\right)} \cdot d\, e^{-\frac{2\mu}{\rho}\left(1-e^{-\frac{\rho Z^*}{2}}\right)}.$$

We take the derivative w.r.t z, and since $\zeta(Z^*,0)=0$ is equivalent to $\mu = \rho\, e^{\frac{\rho Z^*}{2}}$, we substitute for μ, obtaining:

$$\Pi'(z) \equiv \frac{\rho(Z^*-z)}{\rho(T-Z^*)+1} \cdot d\, e^{\rho(Z^*-z)-\frac{\rho Z^*}{2}-\frac{2\mu}{\rho}\left(1-e^{-\frac{\rho Z^*}{2}}\right)}.$$

Because $z > Z^*$, profit is always decreasing when shifting the target earlier to some $z > Z^*$. We use the same strategy for the late degenerate equilibria (reported in the Technical Appendix), showing that profit strictly decreases as the targeted z increases.

In an early dispersed equilibrium, targeting $z \in (Z^*, T)$ would generate profit:

$$\Pi(z) \equiv \frac{\left(\alpha\mu + \rho e^{(T-z)(\alpha\mu+\rho)}\right)\left(e^{\alpha\mu(z-Z^*)} - 1 + \frac{\alpha\mu\left(\rho+\alpha\mu e^{(Z^*-T)(\alpha\mu+\rho)}\right)}{\rho(\alpha\mu+\rho)}\right)}{\mu\left(\alpha\mu + \rho e^{(T-Z^*)(\alpha\mu+\rho)}\right)\left(e^{\alpha\mu(T-Z^*)} + \alpha e^{\frac{\rho Z^*}{2}} - 1\right)} \cdot d\, e^{-\frac{2\mu}{\rho}\left(1-e^{-\frac{\rho Z^*}{2}}\right)}.$$

Note that the denominator and final terms are both positive and constant w.r.t. z and are therefore omitted in the following derivatives.

In an early dispersed equilibrium, we find $\Pi'(Z^*) = 0$ and

$$\Pi''(Z^*) = 2\alpha^2\mu^3\left(\alpha^* + \frac{\rho}{2\mu} + \frac{\rho e^{-\frac{\rho Z^*}{2}}}{2\mu\left(e^{-\frac{\rho Z^*}{2}} - \frac{\rho}{\mu} - \alpha^*\right)}\right) < 2\alpha^2\mu^3\left(\alpha^* + \frac{\rho}{2\mu} - e^{-\frac{\rho Z^*}{2}}\right) < 0.$$

Both inequalities must hold because $F(R(Z^*)) = e^{-\frac{\rho Z^*}{2}} - \frac{\rho}{2\mu}$ and must be larger than $F(R(T)) = \alpha^*$, which rearranges to yield $e^{-\frac{\rho Z^*}{2}} - \frac{\rho}{2\mu} > \alpha^*$. Thus, any price just below $R(Z^*)$ will be strictly less profitable.

For the remaining potential targets $z \in (Z^*, T)$, we show that $\Pi''(z)$ will change sign only once (from negative to positive). Suppose that $\Pi''(\hat{z}) = 0$ at some $\hat{z} \in (Z^*, T)$, yielding:

$$\alpha^{*2}\mu^2\left(\alpha^*\mu e^{\hat{z}(2\alpha^*\mu+\rho)} + (\alpha^*\mu+\rho)e^{Z^*(2\alpha^*\mu+\rho)}\right) = \rho^2 e^{T(\alpha^*\mu+\rho)}\left((\alpha^*\mu+\rho)e^{\alpha^*\mu Z^*} - \rho e^{\alpha^*\mu\hat{z}}\right).$$

We then take the third derivative evaluated at $\hat{z}$:

$$\Pi'''(\hat{z}) \equiv \alpha^*\mu\left(\rho^3 e^{\rho T+\alpha^*\mu(T+z)} + \alpha^{*2}\mu^2(2\alpha^*\mu+\rho)e^{\hat{z}(2\alpha^*\mu+\rho)}\right) > 0.$$

Thus, $\Pi'''(\hat{z}) > 0$ whenever $\Pi''(\hat{z}) = 0$. Thus, there is only one such $\hat{z}$.

Thus, as z increases, profit initially falls near Z^*, but eventually will increase (as the second derivative turns positive, then eventually the first derivative). However, it can never rise above the profit $\Pi(T)$, because if it did, then it would have to later decrease before reaching $\Pi(T)$. This would generate a point where $\Pi''(\hat{z}) = 0$ but $\Pi'''(\hat{z}) < 0$. In the Technical Appendix, we use this same strategy for a bimodal or an early degenerate equilibrium, showing that the third derivative is positive when the second is zero.

Thus, the proposed solution satisfies all necessary conditions for equilibrium. $\square$

Proof of Proposition 3. We proceed by demonstrating that if the conditions for one of the five equilibria holds, it precludes any of the others. First, note that the derivative of ζ w.r.t. Z is:

$$\zeta_Z\left(Z,\alpha\right) = -\left(e^{-\frac{\rho Z}{2}} - \frac{\rho}{2\mu} - \alpha\right)\frac{\mu\left(\alpha\mu\, e^{(\alpha\mu+\rho)(Z-T)} + \rho\right)e^{-\frac{2\mu}{\rho}\left(1-e^{-\frac{\rho Z}{2}}\right)-\frac{\rho Z}{2}}}{\alpha\mu+\rho} < 0.$$

The sign is always negative because $F(R(Z^*)) = e^{-\frac{\rho Z^*}{2}} - \frac{\rho}{2\mu}$ and must be larger than $F(R(T)) = \alpha$ (in a late or early dispersed equilibrium), making the first parenthetical term positive. Note that ζ is not binding in an early bimodal or degenerate equilibrium.

We also find that the derivative w.r.t. α is:

$$\zeta_\alpha\left(Z,\alpha\right) = -\left(\rho\left(1-e^{(\rho+\alpha\mu)(Z-T)}\right) + \alpha\mu(\rho+\alpha\mu)(T-Z)e^{(\rho+\alpha\mu)(Z-T)}\right)\frac{\mu\, e^{-\frac{2\mu}{\rho}\left(1-e^{-\frac{\rho Z}{2}}\right)}}{(\rho+\alpha\mu)^2} < 0,$$

where the sign holds because $T \geq Z$.

First, suppose a late degenerate equilibrium occurs ($\zeta(0,0) \leq 0$). Since $\zeta_Z(Z,0) < 0$, there is no $Z > 0$ for which $\zeta(Z,0) = 0$. Moreover, $\zeta_\alpha(Z,\alpha) < 0$, so likewise, any $\alpha > 0$ would result in $\zeta(Z,\alpha) < 0$, precluding an early equilibrium.

Next, suppose a late dispersed equilibrium occurs ($\zeta(Z^*,0) = 0$). Since $\zeta_Z(Z,0) < 0$ for all $Z \in [0,T]$, so no other late dispersed equilibrium can exist, nor can a degenerate equilibrium since $\zeta(0,0) > 0$. Likewise, $\zeta_\alpha(Z,\alpha) < 0$, so any increase in α will ensure that $\zeta(Z,\alpha) < 0$, thereby precluding an early equilibrium.

The same approach applies to the early dispersed equilibria. However, since $\zeta_Z(Z,\alpha) < 0$ and $\zeta_\alpha(Z,\alpha) < 0$, it is possible that an increase in Z can be offset with a decrease in α so as to maintain $\zeta(Z,\alpha) = 0$. However, this will also disrupt the $\phi(Z,\alpha) = 0$ condition because $\phi_Z(Z,\alpha) < 0$ and $\phi_\alpha(Z,\alpha) > 0$, as shown below (with details in the Technical Appendix):

$$\phi_Z\left(Z,\alpha\right) \;=\; \frac{1}{2}\alpha\left(\rho(2\alpha\mu+\rho)e^{(T-Z)(\alpha\mu+\rho)+\frac{\rho Z}{2}} - 2\mu(\alpha\mu+\rho)e^{\alpha\mu(T-Z)} + \rho^2 e^{\frac{\rho Z}{2}}\right) < 0.$$

This is negative because $\mu > \rho e^{\frac{\rho T}{2}}$ for the early dispersed equilibrium to occur. Thus $\phi_Z\left(T,\alpha\right) = \alpha(\alpha\mu+\rho)\left(\rho e^{\frac{\rho T}{2}} - \mu\right) < 0$ and, for all $Z < T$,

$$\phi_{ZZ}\left(Z,\alpha\right) = \frac{1}{2}\alpha\left(2\alpha^2\mu^2 + 3\alpha\mu\rho + \rho^2\right)\left(\mu e^{\frac{\rho Z}{2}} - \rho e^{\rho T}\right)e^{\alpha\mu(T-Z)-Z\rho} > 0.$$

Similarly, when we take the derivative w.r.t. α and substitute for $\phi = 0$ and $\zeta = 0$, we get:

$$\phi_\alpha\left(Z,\alpha\right) \;=\; (T-Z)\mu\left(\alpha\mu+\rho-\alpha\rho e^{\frac{\rho Z}{2}}\right)-\rho\left(\mu-\rho e^{\frac{\rho Z}{2}}\right)\Big/\left(\alpha\mu+\rho-\mu e^{-\frac{\rho Z}{2}}\right)>0.$$

This is positive because $\phi_\alpha\left(Z,0\right)=\rho\left(\mu(T-Z)+e^{\frac{\rho Z}{2}}\right)>0$, and

$$\phi_{\alpha\alpha}\left(Z,\alpha\right)=\mu\left(\mu-\rho e^{\frac{\rho Z}{2}}\right)\left(T-Z+\rho\Big/\left(\alpha\mu+\rho-\mu e^{-\frac{\rho Z}{2}}\right)^2\right)>0.$$

Thus, if (Z^*,α^*) constitutes an early dispersed equilibrium, any alternative $(\hat{Z},\hat{\alpha})$ will either violate $\zeta(\hat{Z},\hat{\alpha})=0$ or $\phi(\hat{Z},\hat{\alpha})=0$. Indeed, if followed until $\alpha=0$, we rule out a late dispersed equilibrium, or until $Z=0$, we rule out an early bimodal equilibrium.

The early bimodal equilibrium follows the same procedure as the early dispersed (using the requirement that $\phi\geq 0$) to establish $\phi_\alpha\left(0,\alpha\right)>0$. Thus, if an early bimodal equilibrium exists ($\phi(0,\alpha^*)=0$), then an early degenerate cannot (since $\phi(0,1)>0$) and vice versa. $\square$

Proof of Lemma 2. By construction, $\bar{p}=x-V(0)$ is the highest price that anyone is willing to pay, whether with or without recall. It would only exercised by people with $z=0$. Thus offering $\bar{p}$ without recall would only capture measure $H(0)$ of buyers who receive the quote after their deadline, whereas offering $\bar{p}$ with recall will capture those buyers, as well as those who receive the quote at some time $z>0$ but receive no other quotes in the meantime. Thus, the expected profit is strictly higher under the recallable price. This applies even if $\bar{p}$ is not offered in equilibrium (because other prices are more profitable). $\square$

Proof of Proposition 4. The proposed early degenerate equilibrium coincides with the early degenerate equilibrium in the exploding setting, and $R(T)$ is more profitable than any other exploding price under the same condition as in the exploding setting: $(\mu+\rho)\left(e^{\mu T}-1\right)\;\geq\;\rho\left(e^{T(\mu+\rho)}-1\right)$. We also must verify that a recallable offer is not more profitable, with $R(0)$ being the best option among recallable prices, since higher prices will be universally rejected while lower prices will be accepted at exactly the same rate (by those who received no early price before expiration). This would generate the same profit as in Eq. 11 when setting $\alpha=1$ and $\bar{p}=R(0)$, which simplifies to:

$$\Pi_r=\frac{de^{-2T(\mu+\rho)}\left(\mu+\rho e^{T(\mu+\rho)}\right)\left((\mu+\rho)e^{\rho T}-\mu\right)}{\rho^2(\mu+\rho)}. \tag{53}$$

Comparing this against $\Pi_e = \frac{d}{\rho} e^{-T(\mu+\rho)}$, we find that this equilibrium exists iff $(\mu+\rho)\left(e^{\mu T} - 1\right) \geq \rho\left(e^{T(\mu+\rho)} - 1\right) + \frac{\mu^2\left(e^{\mu T} - e^{-\rho T}\right)}{\mu+\rho}$. If this is satisfied, deviation to an exploding price is also not optimal.

For the late or early dispersed equilibria, the late prices coincide with the pure recall equilibrium in Section 2, which establishes that no other recallable price is profitable. We must compute the possibilities for the exploding prices that could be offered. Under the proposed equilibria, we can simplify the buyer's Bellman equations to contemplate accepting the early price if it occurs at rate α, or taking the best late price at expiration if not. Under that perspective, someone with no quotes at expiration has expected utility U defined as:

$$\rho U = -d + \mu \left(x - \alpha \ell + \int_{\underline{p}}^{\bar{p}} p dF(p) - U \right). \tag{54}$$

This indicates that the individual will accept the first quote she receives, whether ℓ with probability α, or one of the late prices p which collectively occur with probability $1 - \alpha$. We recall that $\bar{p}$ is pinned down by indifference so that $U = x - \bar{p}$.

Someone who has just begun their search at time $z = T$ would anticipate a probability $e^{-\mu T}$ of having no quotes by expiration, a probability $1 - e^{-\alpha \mu T}$ of having received the early price quote and thus leaving before expiration, and otherwise getting at least one late price, with the best of these distributed with density $\mu T e^{-\mu T F(p)} F'(p)$. This means the expected utility at expiration is

$$V(0) = \frac{e^{-\mu T} U + \int_{\underline{p}}^{\bar{p}} (x - p) \mu T e^{-\mu T F(p)} F'(p) dp}{e^{-\alpha \mu T}}. \tag{55}$$

This is divided by $e^{-\alpha \mu T}$ because it is conditional on not having received the early price, which would have prevented reaching expiration. During the grace period, no late price is accepted, so the expected utility would be:

$$\rho V(z) = -V'(z) + \alpha \mu (x - \ell - V(z)). \tag{56}$$

We then solve this differential equation, paired with $\ell = R(T)$ and $V(z) = x - R(z)$ (with intermediate steps in the Technical Appendix), to obtain Eqs. 34 and 35. Note that the buyer population in the market with z time remaining follows Eq. 25 with $Z^* = 0$. This then generates the following expected profit from an exploding offer:

$$\Pi_e(z) = \frac{e^{-\alpha^*\mu(T-z)} - e^{-\alpha^*\mu T} + \alpha^* e^{-\mu T}}{1 - e^{-\alpha^*\mu T} + \alpha^* e^{-\mu T}}(R(z) - x). \tag{57}$$

In a late dispersed equilibrium, we set $\alpha = 0$ and verify that $\Pi_e'(z) > 0$ for all z, which holds so long as $\rho < \frac{\mu}{\mu T + e^{-\mu T}}$. In an early dispersed equilibrium, we first verify that $\Pi_e'(T) > 0$, making $\ell = R(T)$ the most profitable exploding price (locally). This again holds if $\rho < \frac{\mu}{\mu T + e^{-\mu T}}$. We then demonstrate that $\Pi_e(z) \leq \Pi(z)/B$, where $\Pi(z)$ is the profit function from the early bimodal equilibrium in the exploding setting. Indeed, $\Pi_e(T) = \Pi(T)/B$ because both settings have $R(T)$ always accepted, and $\Pi_r = \Pi(0)/B$ because both settings have $R(0) = \bar{p}$ accepted at the deadline. Even so, $\Pi_e(0) < \Pi(0)/B$ because it offers the same price $R(0)$, but it is rejected by buyers who receive it prior to expiration, rather than being retained with recall as in Π_r. We then follow the proof of Proposition 2 for the bimodal equilibrium, showing that $\Pi(z)$ is U-shaped. Moreover, since $\Pi_e(T) = \Pi(T)/B$ and $\Pi_e(z) < \Pi(z)/B$ for $z < T$, this shows that $\ell = R(T)$ is the unique profitable exploding offer. Algebraic details are provided in the Technical Appendix.

Finally, we verify that every buyer would accept the exploding offer, as assumed. This holds so long as $\ell < \underline{p}$, but equal profits requires $Pi_e(T) = \ell - x = \Pi_r$, so we only need to verify that $\Pi_r < \underline{p} - x$. But this must hold, because a seller who offers a recallable $\underline{p}$ will only get $\underline{p} - x$ if the buyer does not encounter the early price before the deadline, and with discounting for the elapsed time. Thus, the expected profit must be lower than $\underline{p} - x$, which would be the profit if it was immediately accepted. $\square$

Proof of Proposition 5. The inequality $W_r\% \geq W_e\%$ is equivalent to $e^{\alpha\mu T} - e^{\mu T + Z\alpha\mu + \frac{\rho Z}{2} + \frac{2\mu\left(e^{-\frac{\rho Z}{2}} - 1\right)}{\rho}} < 0$. This is increasing in α, and is thus hardest to satisfy at $\alpha = 1$. Evaluating this and taking logs, the inequality becomes:

$$\frac{2\mu\left(e^{-\frac{\rho Z}{2}} - 1\right)}{\rho} + \mu Z + \frac{\rho Z}{2} > 0.$$

The left hand side equals 0 when Z = 0 and is strictly increasing in Z, with a derivative $\mu\left(1 - e^{-\frac{\rho Z}{2}}\right) + \frac{\rho}{2} > 0$. Thus, at $Z = 0$ and $\alpha = 1$, $W_r\% = W_e\%$, and if $Z > 0$ or $\alpha < 1$, $W_r\% > W_e\%$. $\square$

References

Akin, S. Nuray, V. Lambson, G. McQueen, B. Platt, B. Slade, and J. Wood. (2013) "Rushing to Overpay: Modeling and Measuring the REIT Premium," *Journal of Real Estate Finance and Economics*, **47**, 506-537.

Akin, S. Nuray, and B. Platt. (2014) "A Theory of Search with Deadlines and Uncertain Recall," *Economic Theory*, **55**, 101-133.

Armstrong, M, and J. Zhou. (2016). "Search Deterrence," *Review of Economics Studies* **83**, 26-57.

Baye, Michael, J. Morgan, and P. Scholten. (2006). "Information, Search, and Price Dispersion." *Handbook on Economics and Information Systems* **1**, 323-375.

Burdett, Kenneth, and K. Judd. (1983) "Equilibrium Price Dispersion," *Econometrica* **51**, 955-970.

Burdett, Kenneth, and D. Mortensen. (1998) "Wage Differentials, Employer Size, and Unemployment," *International Economic Review* **39**, 257-273.

Butters, Gerard. (1977) "Equilibrium Distribution of Sales and Advertising Prices," *Review of Economics Studies* **44**, 465-491.

Choi, Michael, A. Y. Dai, K. Kim. (2018) "Consumer Search and Price Competition," *Econometrica* **86**, 1257-1281.

Choi, Michael, and L. Smith. (2022) "Optimal Sequential Search with Recall," working paper, University of Wisconsin.

Coey, Dominic, B. Larsen, and B. Platt. (2020) "Discounts and Deadlines in Consumer Search," *American Economic Review* **110**, 3748-85.

Curtis, E., and R. Wright. (2004) "Price setting, price dispersion, and the value of money: or, the law of two prices," *Journal of Monetary Economics*, **51**, 1599-1621.

Diamond, Peter. (1971) "A Model of Price Adjustment," *Journal of Economic Theory* **3**, 156-168.

Diamond, Peter. (1987) “Consumer Differences and Prices in a Search Model,” *Quarterly Journal of Economics* **102**, 429-436.

Dilmé, Francesc, and F. Li. (2019). “Revenue Management without Commitment: Dynamic Pricing and Periodic Flash Sales,” *Review of Economics Studies* **86**, 1999-2034.

Janssen, Maarten, and J. L. Moraga-Gonzalez. (2004). “Strategic Pricing, Consumer Search, and the Number of Firms,” *Review of Economics Studies* **71**, 1089-1118.

Janssen, Maarten, and A. Parakhonyak. (2014) “Consumer Search Markets with Costly Revisits,” *Economic Theory* **55**, 481-514.

Janssen, Maarten, P. Pichler, and S. Weidenholz. (2011) “Oligopolistic markets with sequential search and production cost uncertainty,” *RAND Journal of Economics* **42**, 444-470.

Lemieux, James, and R. Peterson. (2011) “Purchase Deadline as a Moderator of the Effects of Price Uncertainty on Search Duration,” *Journal of Economic Psychology* **32**, 33-44.

Liu, Dongri, and J. Lu. (2025) “Buyers Hidden Search with Continuous Intensity: The Seller-Optimal Mechanism and A Perverse Effect,” *RAND Journal of Economics*, forthcoming.

Menzio, Guido, and N. Trachter. (2015) “Equilibrium Price Dispersion with Sequential Search,” *Journal of Economic Theory*, **160**, 188-215.

Reinganum, Jennifer. (1979) “A Simple Model of Equilibrium Price Dispersion,” *Journal of Political Economy* **87**, 851-858.

Rob, Rafael. (1985) “Equilibrium Price Distributions,” *Review of Economic Studies* **52**, 487-504.

Salop, Steven, and J. Stiglitz. (1976) “Bargains and Ripoffs: A Model of Monopolistically Competitive Prices,” *Review of Economic Studies* **44**, 493-510.

Stahl, Dale O. (1989) “Monopolistic Pricing with Sequential Consumer Search,” *American Economic Review* **79**, 700-712.

Stigler, George. (1961) “The Economics of Information,” *Journal of Political Economy* **69**, 213-225.

Wilde, Louis, and A. Schwartz. (1979) “Equilibrium Comparison Shopping,” *Review of Economic Studies* **46**, 543-554.