Pricing and Perpetual Royalties with Repeated Resale

Matthew Mitchell*

February 22, 2023

Abstract

This paper considers a durable object that is repeatedly resold among a potential buyers that trade bilaterally, so that markets are thin at any point in time. The results highlight differences between possible contracting environments which, in practice, have become especially important as record keeping technologies improve. Traditional ownership, where each owner sets a price unilaterally, leads to reduction in trade through markups; opportunities for future resale increase these inefficiencies relative to one time sales. Markups decline over time as resale opportunities decline. Fixed percentage perpetual royalties paid to the first owner, as mandated in some countries, are counterproductive; they lower the first owner's value. By constrast, a dynamic contract designed to maximize profits of the first owner achieves efficiency in all but the first sale, despite not achieving full surplus extraction at any point. The first sale is distorted exactly as a one time sale, which is a smaller distortion than any transaction under traditional ownership. The dynamic contract can be interpreted as nonlinear perpetual royalties, a form of payment that has increasingly been discussed especially in digital art markets as record keeping technologies improve. Such price discrimination can increase efficiency, especially in resale transactions.

*University of Toronto. Thanks to Heski Bar-Isaac, Rahul Deb, Hugo Hopenhayn, Bálazs Szentes, Gabor Virag and participants at IIOC 2022, the 5th UniBg Industrial Organziation Winter Symposium and seminars at Stony Brook and the University of Toronto for helpful comments.

1 Introduction

This paper studies a perfectly durable object that is repeatedly traded between people with different independent private valuations of the object. Buyer arrivals are infrequent, so the market is thin to the point of only being bilateral and only occasional. Different contractual arrangements are considered in order to explore how the availability of dynamic contracting arrangements impacts the structure and efficiency of allocations.

Examples that fit this description of trade include markets from many forms of intellectual property to complicated financial securities that are traded on OTC markets. Perhaps the most natural example of a market similar to the one modelled here is an art market. In practice, more complicated ownership structures are available due to the ability of information technology to keep records of ownership. The paper shows that advanced recordkeeping can enhance both rewards for creators, which can be important in the intellectual property interpretation, and efficiency of allocations. Perpetual royalties, which apply on every transfer, have been proposed in such markets, and used especially for digital art, based on new recordkeeping technologies including ones that can be decentralized through the blockchain.[1] The nonlinear contract developed here can be interpreted as perpetual royalties. By contrast, a more traditional ownership structure of posted prices by successive owners is less efficient than static trade. This inefficiency is not easily remedied by simple royalties; in fact forced linear royalties on the next sale, as is mandated in some jurisdictions by law, is counterproductive to both efficiency and providing rewards for creators of the good.[2] Increasingly such contracts can be written for a variety of goods, tangible and intangible, and can incorporate more complicated rules for transactions than simple royalties. This paper studies the role of these contracts.

The model delivers several results on the efficiency of trade in these markets, and the impact that contracting opportunities can have on creators (i.e.

[1] See https://news.artnet.com/market/swizz-beatz-sothebys-artist-royalties-1355674. Such rights might be conferred via non-fungible tokens (NFTs), especially (but not only) for digital art. More generally, web3 “smart contracts” encoded in blockchain allow complicated contracts between many parties (see https://ethereum.org/en/smart-contracts/.).

[2] Simple royalties have been used in markets from art to soccer players, commonly called a sell-on clause, and such payments go at least as far back as payments made on transfer of feudal land.

initial owners). In a market with simple posted prices and no royalties, resale opportunities for buyers make sellers more selective, effectively increasing markups as measured by the probability of sale per bilateral meeting. Enhancing contracts with the possibility of fixed or linear royalties on the next transaction can create new inefficiencies including the possibility that objects move from higher valuation to lower valuation consumers in equilibrium. Subsidy is preferred by sellers to royalties, so fixed positive royalties, as has been instituted in some jurisdictions in art markets, are counterproductive for artists. In a fully-nonlinear contracting environment to maximize the value of the first owner, the usual monotone virtual valuation assumption implies that only the first sale is distorted; subsequent transactions occur efficiently. The first sale is distorted exactly as a one time sale, which is a lower distortion than in a simple ownership economy where prices are posted by each owner. These payments can be interpreted as a market with history dependent payments between buyers, as well as perpetual royalties paid to the initial owner which are positive at every history. In other words, nonlinear but positive perpetual royalties are possible, profitable for initial sellers, and increase efficiency, provided sufficient recordkeeping and control over future transactions is possible. The efficiency of subsequent transactions is somewhat surprising given that full extraction is never achieved; the sequential nature of the price discrimination is what makes the model different from the distorted allocations in classic problems of second degree price discrimination like Mussa and Rosen (1978).

These results come from analyzing three different contracting structures. A benchmark environment of *traditional ownership* mimics ownership with posted prices, and no royalties to prior owners. Markov equilibrium can be described by a recursion, which allows a direct comparison to the static problem: resale opportunities discourage trade at any meeting. Then a version of the model is considered where the full nonlinear contract is not available, but owners can encourage or discourage future transactions by collecting a royalty (or paying a subsidy) on a future transaction. The optimal choice is a subsidy. The intuition is that a subsidy, together with a price that induces some marginal type to buy, pays less to all higher types since the higher types hold the object longer, and therefore wait longer to sell. Because the subsidy can extract surplus from inframarginal types, it is beneficial to sellers. This is of note for several reasons. First, some jurisdictions (for instance France) mandate positive royalties paid to the original owner for future transactions in some markets; others (like Canada) are considering similar rules. The results suggest these rules may be counterproductive for sellers. Moreover, many real world contracts for things like ebooks do the opposite: they make resale more difficult.[3]

Although actual subsidies to future transactions might face additional problems (for instance that the buyer could pretend to make a transaction right away, by transacting with themselves or a fake account), it shows that the motive here is for sellers to encourage future transactions. To understand the limits of this force, the paper then considers more sophisticated contracts that map histories of ownership into prices in a possibly nonlinear way. The contract is remarkably simple: static distortion on the first sale, and efficiency thereafter. Since there are always distortions greater than static under simple ownership, the second degree price discrimination unambiguously increases efficiency. The intuition is similar to an optimal auction with constraints: bidders can only be assigned units of time that occur after their arrival.[4]

One can think of the choice of contractual form as being driven by changes in technology, but the model has policy implications concerning what sorts of contracts should be allowed. In economies without complicated contracts where ownership is sold at a fixed price, objects like books and art are subject to the doctrine of copyright exhaustion, which limits the ability of owners to control further use (or sale) of the object after it is transacted. The application of these ideas in digital markets is an active policy question; in Europe, there is discussion over whether the sale of things like ebooks should be subject to the doctrine of first sale, which would limit sellers ability to restrict buyers use of the product.[5] A central question of this paper is how these more complicated contracting environments impact efficiency, which bears on questions like the efficiency implications of the exhaustion doctrines, and the modern world's ability to avoid exhaustion through ownership structures more complicated than were commonly used previously. The model suggests that such policies may have efficiency concerns when sellers can encourage, rather than discourage, future transactions, but that very rich contracts may avoid this concern. On the other hand, in a simpler environment, there is no

[3]This suggests that the motivations may not be enhancing trade on the item in question, but perhaps reducing competition between the used good and a new good for sale by the same seller, as suggested in the used goods literature.

[4]Like an optimal auction, monotone virtual valuations guarantee that higher types are always allocated as much as possible. Consistent with this intuition, without monotonicity of virtual valuations, efficiency disappears.

[5]See Oprysk (2020)

efficiency motive for simply taxing future sales and exhaustion may prevent interference in used markets by sellers that compete with their own used products.

Section 2 introduces the physical environment; Section 3 describes optimal utilitarian allocation as a benchmark. Then Section 4 considers a repeated ownership structure, where a sequence of owners post prices. Then the contracting space is modified to allow the seller to take a share (including a negative share, a subsidy) of future sales, as has been observed in art markets. Finally, Section 5 considers a full nonlinear contract devised by the first owner.

1.1 Literature

1.1.1 Resale Markets

This paper complements the recent work of Condorelli et al. (2021), which is perhaps the paper with the closest setup to this one in terms of studying repeated resale of a durable object to a sequence of owners.The physical environment differs in that, in their model, each owner can only contact a fixed set of potential buyers, and cannot wait for more, whereas this paper highlights the trade-off between waiting and higher prices. Transactions are therefore useful in their model both to reallocate the object to higher valuations, and to find more buyers. Their focus is on efficiency as the frequency of trade grows; that efficiency trivially arises here because future contracting possibilities are indepenent of transactions.

Classic models of second hand markets (for instance Waldman (1993); Hendel and Lizzeri (1999); Gavazza (2011); Hendel and Lizzeri (2015)) focus on thick markets, where trade occurs because of changes in relative values, especially due to depreciation. Although these models do allow market frictions in the form of transaction wedges, they do not explicitly model those frictions the way this model does. Those models were designed especially to think about markets for things like cars and airplanes and shoe machines rather than goods that trade in thinner markets, which leads to new reasons for trade. Another paper in this space is Stolyarov (2002), which has a similar preference structure to the one used here, but will constant opportunities to trade, potentially at cost. One of the key elements of those papers is that the original seller plans to sell many units, and as a result may want to interfere in used markets that could compete with their sales of new goods. Many models of used goods point to a potential downside of this record keeping, as interference in used markets; for instance, whereas books were once sold only as physical items, ebooks are not; along with that has come much more complicated terms and conditions surrounding these items.[6] This paper focuses on the potential benefits of recordkeeping.[7]

Second hand markets have been considered explicitly in price discrimination strategies. Early examples include Swan (1972), who focuses on durability choice as a mechanism. Anderson and Ginsburgh (1994) further this line to consider a thick market for second hand goods, with transaction costs, and how a monopolist can price discriminate in the face of such a market. Beccuti and Moller (2021) study how a firm can price discriminate via time of holding the object, which is similar to what the separating contract does here, but without commitment and when sellers are more patient than buyers. The holding time can be thought of as a resale decision; contracts sort by whether the good is sold or leased.[8]

[6]Ebooks "purchased" from Amazon are in fact not really owned but rather licensed. These licensing agreements replacing ownership extends even to things like software in a car. These sorts of licensing "terms and conditions" apply to many items we buy; even a new car does not entitle the owner to unconstrained ownership of the software that the car's computer uses.

[7]There is a long debate in the legal literature about the benefits to the first sale doctrine, which is designed to reduce interference, and to what extent it is useful to allow contracts that avoid its.Legal scholars including Hovenkamp (2010) and Katz (2014) have discussed the potential merits and drawbacks of licensing contracts that avoid exhaustion in the digital context, but without the ability to analyze what such contracts might look like for long lived assets. This paper follows in the tradition of Waldman (2015) in deriving such tradeoffs from an explicit model. This model abstracts from the usual interference concern, although in several places interference provides a natural contrast to the results here: whereas here dynamic contracting possibilities encourage future transfers, interference generally discourages them, to further the monopolist's future sales of similar products to other buyers. Therefore the model provides no rationale for justification of interference in used markets for the purpose of limiting future transactions. Moreover, at least for simple leasing contracts, our model provides no scope for the seller to improve their position with leasing relative to selling, so we study a different force that may be relevant in modern digital markets but not in cases previously studied. Weyl and Zhang (2022) study a related tradeoff in ownership rights: what is the best way to resolve the tension between delivering surplus to initial owners (who may in turn use that as an incentive to invest) versus markups that result from their continued ability to dictate use.

[8]Mechanisms with resale include models of auctions with resale such as Zheng (2002); Hafalir and Krishna (2008) In those models there is potentially an ex-post allocation question for some mechanisms, but all of the potential owners are present throughout.

1.1.2 Price Discrimination

The comparison of the traditional ownership structure to the dynamic contract belongs to the large literature on the efficiency of price discrimination. Pigou (1920)and Robinson (1933) highlighted that although perfect price discrimination increases efficiency, other forms of price discrimination may or may not. The strand of literature they started was, for the case of static third degree price discrimination that started their investigation, reinvigorated by Schmalensee (1981) and Varian (1985).[9] The contrast between royalties and the fully optimal contract shows that adding a "bit" of price discrimination can have qualitatively different implications from full second degree contract. Here the dynamic contract turns out to have important similarities to static second degree price discrimination, as cast in Mussa and Rosen (1978).

Dynamic price discrimination has a long history. While this model is quite different, there is a relationship between this work and the classic work on dynamic price discrimination with durable goods that dates to Coase (1972). The commitment case, which most closely matches the price discrimination contract constructed here, was formalized by Stokey (1979). In that model there is pooling; Salant (1989) highlights the contrast between those environments, where costs are essentially linear, with Mussa and Rosen (1978), where costs are assumed to be strictly convex, and separation occurs.

In this paper, costs are endogenous and the result of an opportunity cost of foregone transactions in the future, and turn out to be strictly convex due to the nature of the opportunity costs of foregoing future transactions. [10] Strict concavity arises endogenously because increasing allocations both takes away future opportunities, and makes the marginal type that will have the object allocated to them in the future higher. This opportunity cost of future allocation is the difference between this model and standard models of price discrimination.[11] A long literature on dynamic contracting focuses

Here, the fundamental friction is that bidders come in sequence.

[9] This analysis was extended to competitive environments for instance in Holmes (1989) and Corts (1998). More recently contributions include Armstrong and Vickers (2001), Aguirre et al. (2010), and Vickers (2020).

[10] A very important case, but less related to this paper, is the durable goods monopolist without commitment. See for instance Stokey (1981); Bulow (1982).

[11] Conlisk et al. (1984) introduced the arrival of further buyers into durable goods monopoly pricing. With commitment power, because the logic of Salant (1989) applies, there is no change in the Stokey result: prices are constant. Another important feature is that valuations fluctuate as new consumers arrive. In the durable goods case, Biehl (2001)

on the case where buyers are always present but information arrives to those buyers over time. A general structure for those contracts is described in Bergemann and Valimaki (2010); Pavan et al. (2014); further development of these ideas includes Eso and Szentes (2017); Battaglini and Lamba (2019)

Another strand of price discrimination papers that uses waiting to purchase as a discrimination tool when buyers must contract without fully knowing their valuation, for instance as is done with advanced purchase agreements made before consumption for instance in airline markets. Examples include Courty and Hao (2000), who show that in such contracts the nature of the buyer's uncertainty shapes the contract they are offered. Chen (2008) considers the possibility that the same buyer arrives repeatedly, so that price discrimination with time is related to increasing information for sellers about buyers' valuations from repeated purchases.

2 Environment

There is an infinite horizon of continuous time. There is a single, indivisible, perfectly durable private good, and a sequence of people who could eventually posses it. Everyone discounts the future at a rate normalized to one. At Poisson rate $\lambda > 0$, an opportunity for the current holder to trade with a new person arrives.[12] A person's type $\theta \geq 0$ describes their flow utility per unit of time they have the good. Every person draws their valuation from a common, known distribution $F(\theta)$. Valuations are private information but everything else is publicly observable. Trading opportunities are temporary: trade between two people must be taken at the time of arrival, or never.[13] Money can also be transacted between parties; the details of how the outcome

studies a two period model with changing buyer valuation and Deb (2011) studies an infinite horizon model where values change at most once; both find prices that rise over time. Garrett (2016) incorporates both buyers that arrive over time and whose values change over time continuously, and shows that cyclical prices are possible with commitment. A key difference from this paper and those is that in those models there isn't a dynamic allocation of the good to solve; the monopolist can produce more of the good to sell to more buyers, and the question is what time paths do this job most efficiently.

[12] Continuous time here plays no special role relative to discrete time, except to turn comparative statics on the discount factor into more easily interpreted arrival rates of buyers.

[13] This last assumption is consistent with the usual assumption made in search models. In two of the three contracting structures, where decisions are monotone, it is without loss.

depends on what transfers are allowed is the main topic of Sections 4 and 5. Money is valued linearly and separably from ownership benefits.

Throughout it is assumed that F has a continuous density f. Further, it is maintained that the support of F is either compact, and normalized to $[0, 1]$, or is the positive real line with finite mean. Some results apply to the case with increasing virtual valuations, i.e. that $\theta - \frac{1-F(\theta)}{f(\theta)}$ is increasing; when that assumption is made, it will be stated explicitly.

3 Full Information Planning Benchmark

Consider a planner who, with full information, maximizes the present discounted value generated by transactions, and observes valuations directly. Section 5 shows that there exists a contract that can decentralize this allocation even with private information about types. This problem is simple but will introduce some of the notation and concepts used in the various market situations, and provide some relevant intuition. Let the present discounted value to the planner when the current holder is type θ be $W(\theta)$. Since any strategy for transferring the object returns more when the current holder of the object is higher, $W(\theta)$ is strictly increasing. Since any strategy for transferring the object has a higher return when the new potential holder is a higher type, the strategy for transferring the object is clearly a cutoff: transfer if the new type is above y. The value can be described recursively, where the object is transferred to a new owner y above the current owner θ:

$$W(\theta) = \theta + \lambda max_y \int_y (W(x) - W(\theta)) f(x) dx$$

Since $W(\theta)$ is increasing the planner can optimize by setting $y = \theta$. The value can be further described by using the envelope condition:[14]

Proposition 1. *$W(\theta)$ is continuous, convex, and differentiable with $W'(\theta) = \frac{1}{1+\lambda(1-F(\theta))}$*

Relevant to the results for nonlinear pricing is that, even if the planner didn't fully value the object as the people did, at θ per unit of time, but rather the strictly increasing function $w(\theta)$, the value function is increasing and therefore the logic is the same: transfer whenever someone with higher valuation arrives.

[14]Proofs are contained in the appendix.

Corollary 2. *Suppose the planner values ownership at some strictly increasing, differentiable $w(\theta) \geq 0$. Then $y(\theta) = \theta$.*

Also useful is an alternative view of the planning problem. The problem can be re-written as

$$W(\theta) = max_d d\theta + (1-d)W^u(\Theta(d)) \tag{1}$$

where the cutoff y is converted to discounted duration $d \in [\frac{\lambda}{1+\lambda}, 1]$ of ownership described by $d = \frac{1}{1+\lambda(1-F(y))}$. Conditional on the cutoff, when a new arrival is implemented, the planner gets the value conditional on being above y given by W^u:

$$W^u(y) = \frac{\int_y W(x)f(x)dx}{1-F(y)}$$

We can use this formulation to show an important feature of W^u, which again is true even if we replace the planner's payoff with a strictly increasing $w(\theta)$ instead of θ, and will be useful in characterizing a non-linear pricing example below. Let $y = \Theta(d)$ be the cutoff that delivers d:

Lemma 3. $(1-d)W^u(\Theta(d))$ *is strictly concave in d*

The concavity of this object, which corresponds to the negative of the opportunity cost of allocating d to the current user, applies for any strictly increasing $w(\theta)$; no concavity assumption is needed. Intuitively, as the planner allocates the object for longer, there are two effects: fewer future owners are possible (which, for a given marginal owner reduces payoffs linearly) and the marginal user increases (since more future owners must be excluded) as d increases, which generates strict concavity.

4 Sequential Ownership

4.1 Traditional Ownership

Suppose that the person holding the object is an owner; owners post a price p at which they will sell the object. This economy corresponds to what is termed, in copyright law, the doctrine of exhaustion for physical goods: future owners are unencumbered by any conditions, and therefore solve the same problem as the initial owner, but with a possibly different valuation.

It requires no monitoring after the sale, since the buyer has full ownership rights including the price posting.

The analysis focuses on Markov policies where the set of acceptable prices at which to buy, and to set when selling, are a function of the owner's type alone. Denote by $V(\theta)$ the value of owning the good if the owner's type is θ. This value is inclusive of any revenue from selling the good but does not include the price paid for the good. Clearly $V(\theta)$ is strictly increasing since, if a higher type were to post a price identical to a lower type's price, they would make the same revenue from sales, and enjoy more utility in the meantime. A price is accepted, therefore, if $p \leq V(\theta)$ and can be considered as equivalent to a marginal type y that buys the object at price p; $p = V(y)$. The value can then in turn be expressed as

$$V(\theta) = \theta + \lambda max_y (1 - F(y))(V(y) - V(\theta)) \tag{2}$$

Usual contraction arguments guarantee existence and uniqueness of V. Since any equilibria that was Markov as described above would have to satisfy this recursion, existence and uniqueness come directly from the recursion. The following characterizes the solution:

Lemma 4. *V is strictly increasing and strictly convex with $V'(\theta) = \frac{1}{1+\lambda(1-F(y))} = d(\theta)$ defined almost everywhere. Any selection of solutions $y(\theta)$ is strictly increasing with $y(\theta) > \theta$. When the support is compact, $V(1) = 1$. When the support is unbounded, $lim_{\theta \to \infty} V(\theta) = \theta$.*

Convexity of V arises because allocating to higher value types is less useful when the current owner is relatively low value, since even a slightly higher type is likely to derive most of their value from selling the object, and not from the part that depends on the type. Convexity has an immediate implication for the impact of resale opportunities on the time that objects are held. Holding time is a useful measure of the distortions arising from seller market power in this model; holding times will naturally turn out to be higher than the planner's holding time, and so a natural question is by how much. Naturally prices will be higher with resale opportunities since the object can generate more value, so holding time is a more useful measure of monopoly markup.

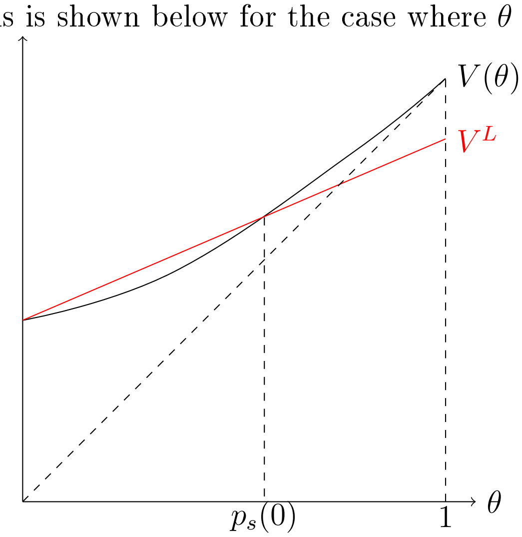
For comparison to a case without resale, consider an economy where an owner with valuation θ has one opportunity to transact with a potential buyer with type drawn from F; no additional trades are possible. This standard monopoly price solves

$$p_s(\theta) = argmax_p(1 - F(p))(p - \theta) \tag{3}$$

This is also the choice of cutoff $y = p_s(\theta)$, and is the cutoff for the dynamic economy with $\lambda = 0$, since in that case $V(\theta) = \theta$. When $\lambda > 0$, sellers are more selective (i.e. have a higher cutoff y) than the static solution:

Proposition 5. *Suppose $\lambda > 0$ and virtual values are increasing. Then for all θ below the maximum of the support of F, $y(\theta) > p^s(\theta)$.*

The intutition for the higher-than-static cutoff comes from convexity of V. The static problem is equivalent to the dynamic one if $V(\theta) = \theta$. In the problem with resale there are two differences: the level of V is higher than θ because of the gains from resale, and it is convex. Let the linear function through both $V(\theta)$ and $V(p^s)$ be V^L. The solution to the recursion in (2) using V^L remains $y = p^s$; formally the linear function simplifies to the case where the value is simply θ. Intuitively it represents a fixed vertical shift, plus a change in "units." The actual function V goes through the same two values at θ and p^s but is *more convex* than the linear function. This changes the return to setting a higher cutoff: since the slope is greater than the linear function at p^s, setting a higher y increases the price that can be charged at a greater rate under the convex V than it does under V^L. The return to holding the object is the same since $V(\theta) = V^L(\theta)$. This therefore leads to a higher cutoff. This is shown below for the case where $\theta = 0$ for simplicity.



When the support is bounded, it is immediate the markups have to be converging to zero (the static market for θ near 1). But this feature of declining markups is more general: According to Proposition 4, when the support of F is unbounded, the markup is converging to the static one as θ gets large, since $V(\theta)$ converges to θ and $V'(\theta)$ converges to one. Long run markups are at their lowest.

Since resale makes sellers more selective, one might wonder whether increasing λ can ever slow transactions, including the positive effect that more meetings per unit of time makes more transactions for a fixed cutoff. It cannot: despite sellers being more selective, more frequent meetings unambiguously speed up transactions, even though transactions per meeting fall.

Proposition 6. *$d(\theta)$ is decreasing in λ.*

Here the argument uses the recursive characterization directly: the slope of the value function, which is $d(\theta)$, is decreasing in λ by a contraction argument.[15]

4.2 Royalties from future sales (and subsidies to future sales)

When considering price discrimination strategies that might make creators better off, one natural starting point is a two-part contract. Here an analogous two-part contract is one which collects (or pays) both at the time of sale, and the time of next sale. Such contracts only rely on monitoring the next transaction, and versions of this structure have been imposed on sellers for instance under France's "droit du suite" policy.

Suppose that an owner post not just p but also a royalty (where a negative royalty corresponds to a subsidy) τ paid at the time of the next owner's sale. Otherwise the economy proceeds as in the the traditional market: each owner faces an amount τ to be paid to the prior owner and can charge a royalty τ' on the next owner of their choosing. For simplicity take the payment to be a

[15] The model of Condorelli et al. (2021) is used to address the question of whether frequent meetings lead to efficiency, which corresponds to increasing λ to infinity. This is an interesting question in their model, where transactions are required to generate further buyers, but is trivially true here: for instance in the compact support case, as λ goes to infinity, a cutoff close enough to 1 generates $V(\theta)$ arbitrarily close to 1 as λ is made big enough, and therefore total surplus, which is at least $V(\theta)$, is converging to social surplus $W(\theta)$, which cannot exceed one.

fixed amount, although the appendix extends the analysis to an ad valorem rate τ on sales revenue, with similar results.

The seller collects $p - \tau$ at the time of the sale, and then collects τ' at a date in the future. Let the payoff, net of the royalty they face but excluding the price they paid, be $V(\theta, \tau)$. They face the recursive problem

$$V(\theta,\tau) = \theta + \lambda max_{y,\tau'}(1-F(y))(V(y,\tau') - \tau + \tau' \int_y s(x,\tau')f(x|x>y)dx - V(\theta,\tau))$$

where $p = V(y, \tau')$ is the net payoff, and therefore the price p that can be charged, to the marginal type y. Here the discounting until next sale for a type θ facing a royalty τ is $s(\theta,\tau) = \frac{\lambda(1-F(y(\theta,\tau)))}{1+\lambda(1-F(y(\theta,\tau)))}$; $s(\theta,\tau) \in [0, \frac{\lambda}{1+\lambda}]$.

Proposition 7. *Suppose that $d(\theta, \tau) < 1$. Then the optimal $\tau'(\theta) < 0$, a subsidy.*

The intuition for subsidy can be seen by considering a fixed y and considering the impact of a subsidy. The marginal consumer is fully extracted regardless; however, a subsidy is less valuable to higher types who intend to hold the object longer. Therefore the subsidy serves to extract from inframarginal types.[16] This channel makes the result different from double marginalization results: even if the decision rules were fixed for the subsequent owner as τ decreases, so that the change in τ doesn't induce more trade, it would still be the case that the subsidy is better because it price discriminates across types by virtue of their duration of ownership, which improves extraction for the seller. The notion that royalties may not be good for rent extraction is counter to the usual intuition that motivates policies that enforce positive royalties.

This result implies that, for an initial owner with θ less than the maximum value and $\tau = 0$, it is optimal to subsidize and have $y < 1$ and $\tau' < 0$. Inductively this implies that trade has subsidy forever almost surely. Subsidies imply a specific deviation from the planners problem that cannot occur under simple ownership. Notice that, if the current owner of some type sets subsidy $-\tau'$, then a new owner who draws θ very close to 1 will set $y < \theta$, since a type that arrives with a type nearly as high as them would be willing

[16]This intuition implies the result might be reversed if buyers had heterogeneous and private values of λ; a buyer with high λ would be hit more by a tax for a given level of θ. Such a model with multi-dimensional heterogeneity is an interesting topic to explore in the future.

to pay at least 1, plus they would receive the subsidy. In other words, they will be willing to sell to someone of a lower type then themselves, due to the subsidy, leading to the good moving from higher to lower valuations. The subsidy encourages trade past the point of efficiency.

It is possible to consider more general subsidies and royalties that apply more than just to the next sale; the space of such possibilities is large. On the one hand, a fixed ad valorem subsidy on all future sales runs up against the same intuition; conditional on a set of marginal types, taxing future sales does a worse job of extracting from inframarginal types.

Corollary 8. *Suppose* $d(\theta, \tau) < 1$*. Then the optimal ad valorem royalty* $\tau'(\theta) < 0$ *is a subsidy.*

On the other hand it is not a surprise that these subsidies are not observed in practice; a buyer who could concoct a sham transaction could immediately collect the subsidy (rather than waiting) and undo all the benefits to the seller. In the next section, a more complicated contract is considered where both subsequent transactions are encouraged, and the ability of buyers to work around the encouragement with sham transactions is eliminated.

5 Nonlinear Contracts

This section considers an initial owner who could prescribe allocations to new arrivals, and payments from those arriving buyers, as a function of their reported type, and the history of previous reports. Although a great deal of generality is allowed, a relatively simple structure emerges, where the next arrival is implemented if it is above a cutoff that depends only on the last arrival, and the cutoff is equal to the current owner's type, except for the first sale. The first sale is distorted exactly like a static, once and for all sale described in (3).

The initial owner will be termed the seller, and will describe a fully history dependent mapping from arrivals and reported types into allocation of the object and prices paid to them. These net payments will later be interpreted below as coming from payments between holders of the object and royalty payments for subsequent sales paid to the seller; therefore it is natural to assume that both the arrival and the reported type is public information, since it needs to be transmitted via the future holders of the object who may be the ones that find buyers. This makes keeping track of histories simpler, and there is nothing payoff relevant for a potential buyer to learn from these details, conditional on the terms they are being offered. In keeping with the smart contracts and NFT motivation, there is full commitment: all terms are encoded in the object at time zero.

To describe this allocation, define a history h^t that lists the times and the reported type of the prior transaction. Therefore h^t is unchanged except at moments when a transaction occurs.[17] The allocation specifies, for any report θ_t at time t given a history leading up to t of h^t_-, whether or not to transact the object and the price the seller receives at this history and report, subject to a mild measurability condition described below. Suppose a person arrives at t and is allocated the object, generating history h^t. Describe their possession of the object (if any) as lasting for any history in the collection $H(h^t)$, which is required to be measurable. Let $\chi()$ be the indicator function. The buyer's payoff from buying is

$$\theta \int E_{h^\tau|h^t}(e^{-\tau}\chi(h^\tau \in H(h^t))d\tau - p(h^t)$$

Let $d(h^t) = \int E_{h^\tau|h^t}(e^{-\tau}\chi(H(h^t), h^\tau)d\tau$; this payoff can then be written as $d(h^t)\theta - p(h^t)$. Since the buyer cares only about $d(h^t)$ for any $p(h^t)$, from the standpoint of incentive compatibility the seller can freely substitute any contract that delivers the same $d(h^t)$ for each history and maintain incentive compatibility at h^t.

It is immediate that if a buyer of type θ finds it optimal to purchase at some history, then so does any buyer with a higher type, since they would get a higher payoff from making the same report. Whether the object is transferred can therefore be described by a measurable function $\theta_{h^t}(\tau)$ which is the cutoff type that is implemented at time $\tau > t$ starting from a purchase at history h^t if no transaction has occurred. Although this can be a complicated object, it is always equivalent, in payoff to the seller and duration of ownership for the owners, to a lottery over fixed cutoffs:

Lemma 9. *For any $\theta_{h^t}(\tau)$ there exists a lottery over constant cutoffs $\Delta\theta$ that delivers the same future payoff to the planner and duration for anyone allocated the object.*

[17] This description excludes from histories dates when an arrival occurs and a report is made that did not lead to a transaction. Allowing contracts to depend on these events amounts to allowing for randomization, which is shown below to not be useful to the seller. It makes notation simpler to not include such arrivals in the history.

From the sellers standpoint, offering different cutoffs at different histories h^τ to a buyer at h^t is equivalent to a lottery over those cutoffs. Therefore, it is sufficient to allow the seller to choose lotteries over cutoffs (which will imply lotteries over duration for the buyer); it will turn out that such lotteries are not optimal, and a deterministic cutoff is optimal. However this consideration of lotteries shows that the problem is allowing for a rich set of history dependent rules. In turn, a lottery over cutoffs is equivalent to a lottery over durations $\Delta d(x)$, which is what will be compute the optimal allocation.

For any cutoff, the payoff to a seller, at the moment a type x arrives above θ, of choosing a future lotteries over allocation $\Delta d(x)$ to those types, can be written recursively as

$$J^u(\theta) = max_{\Delta d(x),p(x)} \int_\theta E_{\Delta d(x)} \left(p(x) + (1-d(x))J^u(\Theta(d(x)))\right) f(x|x>\theta)dx$$

where the expectation for lottery $\Delta d(x)$ is over durations $d(x)$.[18] One can write this as

$$J^u(\theta) = max_{d(x),p(x)} \int_\theta \left(p(x) + conc\left((1-d(x))J^u(\Theta(d(x)))\right)\right) f(x|x>\theta)dx \tag{4}$$

where $conc()$ is the concave envelope, and $d(x)$ is the expected duration across lotteries $\Delta d(x)$. The use of the notation J^u mirrors the function W^u in the planners problem, which is an analogy that is drawn out throughout this section. It will turn out that, under the monotone virtual valuation assumption made below, the function inside the envelope operator is concave, so the solution is solved with a single cutoff.

Incentive compatibility for type x is

$$x \in argmax_{\hat{x}} d(\hat{x})x - p(\hat{x})$$

and IR is that $p(\theta) = d(\theta)\theta$. IC and IR can be replaced, for any increasing $d(x)$, by choosing the appropriate prices so that $p'(x) = d'(x)x$:

$$p(x) = p(\theta) + \int_\theta^x t d'(t)dt$$

[18] With some abuse of notation, which will quickly disappear.

Note that (4) is like the classic formulation of Mussa and Rosen (1978), where $C(d) = -conc(1-d)J^u(\Theta(d))$ and the problem can therefore be written as

$$J^u(\theta) = max_{d(x)} \int_\theta \left(d(x) \left(x - \frac{1-F(x)}{f(x)} \right) - C(d(x)) \right) f(x|x > \theta) dx \quad (5)$$

Where virtual valuations are computed using the conditional distribution but $\frac{1-F(x|x>\theta)}{f(x|x>\theta)} = \frac{1-F(x)}{f(x)}$.

The monotone virtual valuation assumption implies that this maximization can be solved pointwise independent of θ; the implication, combined with the fact that the virtual valuations don't depend on the lower cutoff, is that $d(x)$ does not depend on θ. Since C is concave according to Lemma 3, the solution is monotone in x and IC is satisfied for the pointwise solution. Moreover lotteries are irrelevant; a single cutoff for each duration can be used. Moreover, history impacts allocations only through the cutoff θ and not through allocations of types that report being above the cutoff. This is an important feature of Mussa-Rosen contracts generally: if the seller discovers that the buyer is distributed on $[\theta, 1]$ instead of $[0, 1]$, but follow the conditional distribution of F on that interval, the only change in the optimal contract is that prices shift up by a constant to extract all surplus from the marginal type θ.

The pointwise problem can be written as

$$J(\theta) = max_d d(w(\theta)) + (1-d)J^u(\Theta(d))$$

Because this transformed problem coincides with the modified planning problem with $w(\theta) = \theta - \frac{1-F(\theta)}{f(\theta)}$, it has the same solution:

Proposition 10. *Suppose virual valuations are increasing. Then the solution to (5) is* $d(x) = \frac{1}{1+\lambda(1-F(x))}$.

The price for type x is

$$p(x) = \frac{\theta}{1 + \lambda(1 - F(\theta))} + \int_\theta^x \frac{x\lambda f(x)}{(1 + \lambda(1 - F(x)))^2} ds$$

so that, for incentive compatibility, $p'(x) = d'(x)x = \frac{x\lambda f(x)}{(1+\lambda(1-F(x)))^2}$. .

5.1 Initial d

The initial owner solves, prior to the arrival of the first buyer, a different problem since they can choose a cutoff and know their own type. Their problem, if their type is θ, is

$$J_0(\theta) = max_y \theta + \lambda(1 - F(y))(J^u(y) - J_0(\theta)) \tag{6}$$

Rewrite the initial choice as

$$J_0(\theta) = J(\theta^s)$$

where $w(\theta^s) = \theta$, i.e. $\theta = \theta^s - \frac{1-F(\theta^s)}{f(\theta^s)}$ or $\theta + \frac{1-F(\theta^s)}{f(\theta^s)} = \theta^s$, which is the formula for the static solution $p^s = \theta^s$. Applying the result from Proposition 10, the optimal initial $y_0 = \theta^s = p^s$. The initial owner prices as if solving exactly the static problem, regardless of λ. The dynamic contract generates efficiency on later sales at the expense of earlier ones, but is still more efficient even in the first transaction than simple ownership when repeated ownership occurs.

5.2 The dynamic contract as perpetual royalties

Although the contract in this section is written as a list of payments to the original owner, those net transfers can be described in a variety of ways. One issue with the payments as described in $p(x)$ is that the current holder doesn't get any compensation when they are forced to transact; they would prefer to hide forever and get their type, rather than nothing when the buyer arrives. An alternative is to ask if the payments in the dynamic contract can be redefined with payments to owners when they "sell," such that they are at least as well off transferring the object as not. An additional benefit of such an arrangement is that it has a natural interpretation as prices, together with a (possibly history dependent) perpetual royalty payment or subsidy.

Suppose that the initial owner can describe conditions under which the object is transferred, but cannot force transactions. Therefore it must be the case that any transaction includes a payment to the current holder of at least θ, i.e. their value if they run away with the object and don't trade. We assume, consistent with how some of these record keeping technologies work, that one could not legally transfer the object without following the contract, so subject to the constraint the payments can be made in any way. Modern contracts like NFTs and ethereum "smart contracts" can ensure that contracts cannot be made outside of the rules encoded in the object.

Consider the case where the current owner has type θ. The arriving type reporting type x pays the buyout θ, a royalty (possibly negative) $r(x, \theta)$ to the first owner, and receives, once the next transaction takes place, their own buyout x. Incentive compatibility requires that the net discounted payments equal $p(x)$:

$$\theta + r(x, \theta) - (1 - d(x))x = p(x)$$

Therefore

$$r(x, \theta) = (x - \theta) - (d(x)x - p(x))$$

This expression has a simple and intuitive interpretation: it is the gain from the transfer (i.e. $x - \theta$) less the rents that type x gets from their allocation. Computing

$$\begin{aligned} dr/dx &= 1 - d(x) - (d'(x)x - p'(x)) \\ &= 1 - d(x) > 0 \end{aligned}$$

Since $d'(x)x - p'(x) = 0$ by the IC constraint for type x. Since $r(\theta, \theta) = 0$, the following characterizes the royalties:

Proposition 11. *The royalties $r(x, \theta)$ can be written as*

$$r(x, \theta) = \int_{\theta}^{x} (1 - d(s))ds = \int_{\theta}^{x} \frac{\lambda(1 - F(s))}{1 + \lambda(1 - F(s))} ds \geq 0 \tag{7}$$

This definition of payments has (1) payment such that an owner would (weakly) rather sell than run away and get their type forever, and (2) positive royalties. Notice that since the net payments of every buyer is $p(x)$, it generates the same revenue for the seller. Moreover, unlike a subsidy from the prior section, the structure does not encourage mock transactions. Suppose that instead of reporting their true type x, the buyer could report, in short succession, two arrivals, one of type $m < x$ and then one of type x. This results in the same net payments as reporting x directly: in either case, the buyer pays θ to the prior owner, and receives $(1-d(x))x$ from the true buyer that comes after. They are on both sides of the payment of m. In terms of royalties, instead of paying $r(x, \theta)$, the two reports result in payments of $r(m, \theta)+r(x, m)$. But from the integral description in (7), there two amounts are the same.

Although there are many ways to define payments between buyers and the original monopolist, this one is "minimal" in the sense that it pays as little as possible to have the buyer willing to sell when the time comes (which requires buyouts of at least θ), and buyers are just indifferent to making double reports if they were able to. It results in positive royalties at every history, like perpetual royalties. It looks like payments of regulated prices equal to your reported type, together with royalties. An important difference from the repeated ownership economy, however, is that the monopolist still allows only a fixed menu of possible prices to be sold; in the example of repeated ownership where subsidy was optimal, the future prices of the object could not be directly controlled. This suggests a role for perpetual royalties only if future sales can be regulated in this way.

5.3 Social Value and Gains to Sellers

In order to understand the relationship between social value and the gains sellers can achieve in the optimal contract, this section compares social value and the seller's initial value at time zero for two parametric examples. To focus on the pure seller, consider an object that starts with an agent of type 0, as in a typical seller's problem where the object only has value to the extent that it can generate revenue.

First suppose $F(x)$ is uniform on $[0, 1]$. The planner's problem can be solved via the envelope equation:

$$W'(\theta) = \frac{1}{1 + \lambda(1 - \theta)}$$

and so, since $W(1) = 1$,

$$W(\theta) = 1 - \frac{ln(1 + \lambda(1 - \theta))}{\lambda}$$

For comparison to the problem of the seller, $w(\theta) = 2\theta - 1$ so the envelope equation is

$$J'(\theta) = d(\theta)w'(\theta) = \frac{2}{1 + \lambda(1 - \theta)}$$

and so, since $J(1) = 1$,

$$J(\theta) = 1 - 2\frac{ln(1 + \lambda(1 - \theta))}{\lambda}$$

A useful comparison between the two is between $J_0(0)$, the pure sellers payoff (i.e. they don't value the object) and $W(0)$, the potential social surplus in the same situation, and how it changes with λ. Since increasing λ changes the total value, it is natural to think of the ratio of the two, i.e.

$$J_0(0)/W(0) = J(1/2)/W(0) = \frac{\lambda - 2ln(1+\lambda/2)}{\lambda - ln(1+\lambda)}$$

This is increasing in λ from $1/2$ when λ is near zero to 1 when λ gets large. The latter is immediate from the fact that both must converge to 1 at the top of the support when the support is compact.[19]

When the support is unbounded, there is no longer a necessity of equalization of private and social value as λ grows large. Suppose $F(\theta) = 1 - e^{-\gamma\theta}$. Then

$$W'(\theta) = \frac{1}{1+\lambda e^{-\gamma\theta}} \tag{8}$$

so

$$\begin{aligned} W(\theta) &= \theta + \frac{ln(\frac{1+\lambda e^{-\gamma\theta}}{\gamma})}{\gamma} - \frac{ln(\frac{1}{\gamma})}{\gamma} \\ &= \theta + ln(1+\lambda e^{-\gamma\theta})^{1/\gamma} \end{aligned}$$

Since $w(\theta) = \theta - 1/\gamma$, J shifts this by $1/\gamma$ relative to W, i.e.

$$J(\theta) = \theta + ln(1+\lambda e^{-\gamma\theta})^{1/\gamma} - 1/\gamma$$

But then since type 0 has a static maximizer $1/\gamma$:

$$J_0(0)/W(0) = J(\gamma)/W(0) = \frac{ln(1+\lambda e^{-1})}{ln(1+\lambda)}$$

This is increasing from e^{-1} near $\lambda = 0$ to 1 as λ grows large. This is not due to convergence of the levels of the two: the difference between them is

$$W(0) - J_0(0) = \frac{1}{\gamma} ln(\frac{1+\lambda}{1+\frac{\lambda}{e}})$$

[19] Obviously the two values are zero at $\lambda = 0$, and more generally, for compact support the difference must be zero when λ is zero (since in both cases the object never changes hands) and zero when λ is large, since both the planner and seller get payoff 1. Computing the difference:

$$W(0) - J_0(0) = W(0) - J(1/2) = 2\frac{ln(1+\lambda/2)}{\lambda} - \frac{ln(1+\lambda)}{\lambda}$$

This has an inverted U shape in λ: the gap between social value and the seller's value first increases then decreases.

which is increasing and convering to $1/\gamma$ as λ grows large. In other words, in this case faster transactions always increase social welfare by more than they increase the seller's surplus, but at a slower rate than social surplus grows with λ.

5.4 Discussion: Non-Monotone Virtual Values

A natural question is what happens in the dynamic contract when virtual valuations are not monotone. First, it is immediate that efficiency after first sale cannot be maintained: the efficient allocations are strictly increasing, incentive compatibility is slack, which would imply pointwise maximization and convex costs, but pointwise optimization is not monotone if virtual valuations are not. Second, since virtual valuations are maximized at the top of the distribution, efficiency is eventually reached; this is different from the usual "no distortion at the top" since it applies for any region at the top where virtual valuations are monotone for all higher values, and is consistent with the idea that efficiency is greater later in the contract's life, as was true with monotone virtual valuations.

The standard approach when virtual valuations are not monotone is to iron. Ironing has slightly different implications here because the allocations from the ironed values impact both the payoff directly and indirectly though the endogenous cost function. To see how this manifests itself in this problem, suppose that virtual valuations have a single interior local maximum at θ_1 and a single interior local minimum $\theta_2 > \theta_1$, and then ironed valuations are defined to be equal to the true valuations except on some interval (a, b) where they are constant. The difference from the usual ironing approach is that, not only are the current payoffs payoffs dependent on the choice of ironing, but also indirectly via the cost function.

Consider solving (5) with the ironed values. Since the payoff is weakly increasing, a simple variant of the efficiency result is immediate: the solution is equivalent to treating the ironed values as an atom in the distribution of θ, and the optimal rule is to always transact the object whenever a higher ironed virtual valuation consumer arrives. This, however, doesn't pin down the rule in the ironed region, since values are constant; formally, both the current value and the marginal cost of allocating duration are constant and coincide. In other words, and rule in (a, b) is equally good at maximizing ironed virtual valuation when the current owner is in that region.

As usual, a decreasing rule violates IC, and an increasing rule would imply IC is slack (and therefore in turn not increasing), so the rule must be constant; the constant cutoff (and the ironing point itself) must tradeoff over-rewarding high θ with low virtual valuation, and under-rewarding low θ with high virtual valuation. So the optimum cutoff is above a and below b. This implies departures from efficiency in both directions: near a, the cutoff is above the efficient one (i.e. inefficiently few transactions) and near b it is below (inefficiently many transactions, including transactions to worse types as with the subsidy for the simple royalty case).

Still, outside of the ironed region, there are efficiency benefits from the more complicated contracts. These benefits, including those from perpetual royalties, suggest a potential downside from exhaustion rules that would limit such contracts. Subsidy polices might be possible even with exhaustion, since owner were offered only a free (but not negatively priced) option to accept the subsidy, they would; offering them the right, at the same initial price to own the object without further interaction with the seller is worse for them. However the efficiency benefits come from controlling future transactions in a way that would likely run afoul of exhaustion. There is a trade off between these benefits of dynamic contracts, and the known concern for interference in used markets.

6 Conclusion

This paper introduced a simple model of repeated transactions in a thin market, where buyers come along periodically. Although quite abstract, it highlights the usefulness of dynamic contracts that are now easy to write in encouraging subsequent transactions. A full dynamic contract has an interpretation as perpetual royalties, but bundled with pricing limitations on subsequent owners. Such a contract can achieve efficiency on all but the first sale, while distorting the first sale less than any owner would if they could not include such complicated terms. This shows both the complexity of such arrangements, and the potential limitations to them. The efficiency result contrasts with repeated sellers such as in Hendel and Lizzeri (1999) where sellers discourage resale because it competes with their own sales, and with static problems of price discrimination like Mussa and Rosen (1978) where monotone virtual valuations do not generate efficiency.

The model could be augmented in many ways in the future. One natural concern in used markets is adverse selection a la Akerlof (1970). Because the paper assumes independent private values, this doesn't arise here, but future research could consider such concerns. Another form of heterogeneity might be heterogeneity in owners' ability to find additional buyers (i.e. λ in the model). This would likely change considerably the nature of royalty agreements. Finally, this structure would be a natural one to embed in a search model in order to think about how equilibrium considerations impact these contracts.

References

Aguirre, I., Cowan, S., and Vickers, J. Monopoly price discrimination and demand curvature. *American Economic Review*, 100:1601–1615, 2010.

Akerlof, G. A. The market for 'lemons': Quality uncertainty and the market mechanism. *Quarterly Journal of Economics*, 84:488–500, 1970.

Anderson, S. P. and Ginsburgh, V. A. Price discrimination via second-hand markets. *European Economic Review*, 38:23–44, 1994.

Armstrong, M. and Vickers, J. Competitive price discrimination. *The RAND Journal of Economics*, 32:579, 2001.

Battaglini, M. and Lamba, R. Optimal dynamic contracting: The first-order approach and beyond. *Theoretical Economics*, 14:1435–1482, 2019.

Beccuti, J. and Moller, M. Screening by mode of trade. *Games and Economic Behavior*, 129:400–420, 2021.

Bergemann, D. and Valimaki, J. The dynamic pivot mechanism. *Econometrica*, 78:771–789, 2010.

Biehl, A. R. Durable-goods monopoly with stochastic values. *The RAND Journal of Economics*, 32:565, 2001.

Bulow, J. Durable-goods monopolists. *Journal of Political Economy*, 90: 314–332, 1982.

Chen, Y. Dynamic price discrimination with asymmetric firms*. *The Journal of Industrial Economics*, 56:729–751, 2008.

Coase, R. Durability and monopoly. *Journal of Law and Economics*, 15: 143–149, 1972.

Condorelli, D., Galeotti, A., and Renou, L. On the Efficiency of Large Resale Networks, 2021.

Conlisk, J., Gerstner, E., and Sobel, J. Cyclic pricing by a durable goods monopolist. *The Quarterly Journal of Economics*, 99:489–505, 1984.

Corts, K. Third-degree price discrimination in oligopoly: All-out competition and strategic commitment. *RAND Journal of Economics*, 29:306–323, 1998.

Courty, P. and Hao, L. Sequential screening. *The Review of Economic Studies*, 67:697–717, 2000.

Deb, R. Intertemporal price discrimination with stochastic values. 2011.

Eso, P. and Szentes, B. Dynamic contracting: An irrelevance theorem. *Theoretical Economics*, 12:109–139, 2017.

Garrett, D. F. Intertemporal price discrimination: Dynamic arrivals and changing values. *American Economic Review*, 106:3275–99, 2016.

Gavazza, A. Leasing and secondary markets: Theory and evidence from commercial aircraft. *Journal of Political Economy*, 119:325–377, 2011.

Hafalir, I. and Krishna, V. Asymmetric auctions with resale. *American Economic Review*, 98:87–112, 2008.

Hendel, I. and Lizzeri, A. Interfering with secondary markets. *The RAND Journal of Economics*, 30:1, 1999.

Hendel, I. and Lizzeri, A. The role of leasing under adverse selection. *Journal of Political Economy*, 110:113–143, 2015.

Holmes, T. The effects of third-degree price discrimination in oligopoly. *American Economic Review*, 79:244–250, 1989.

Hovenkamp, H. Post-sale restraints and competitive harm: The first sale doctrine in perspective. *New York University Annual Survey of American Law*, 66, 2010.

Katz, A. The first sale doctrine and the economics of post-sale restraints. *Brigham Young University Law Review*, 2014, 2014.

Mussa, M. and Rosen, S. Monopoly and product quality. *Journal of Economic Theory*, 18:301–317, 1978.

Oprysk, L. 'digital' exhaustion and the eu (digital) single market. *EU Internet Law in the Digital Era: Regulation and Enforcement*, pages 161–180, 2020.

Pavan, A., Segal, I., and Toikka, J. Dynamic mechanism design: A myersonian approach. *Econometrica*, 82:601–653, 2014.

Pigou, A. C. *The economics of welfare*. Macmillan, 1920.

Robinson, J. *Economics of Imperfect Competition*. Macmillan, 1933.

Salant, S. W. When is inducing self-selection suboptimal for a monopolist? *Quarterly Journal of Economics*, 104:391–397, 1989.

Schmalensee, R. Output and welfare implications of monopolistic third-degree price discrimination. *American Economic Review*, 71:242–247, 1981.

Stokey, N. L. Intertemporal price discrimination. *The Quarterly Journal of Economics*, 93:355–371, 1979.

Stokey, N. L. Rational expectations and durable goods pricing. *The Bell Journal of Economics*, 12:112, 1981.

Stolyarov, D. Turnover of used durables in a stationary equilibrium: Are older goods traded more? *Journal of Political Economy*, 110:1390–1413, 2002.

Swan, P. Optimum durability, second-hand markets, and planned obsolescence. *Journal of Political Economy*, 80:575–585, 1972.

Varian, H. Price discrimination and social welfare. *American Economic Review*, 75:870–875, 1985.

Vickers, J. Direct welfare analysis of relative price regulation*. *The Journal of Industrial Economics*, 68:40–51, 2020.

Waldman, M. A new perspective on planned obsolescence. *Quarterly Journal of Economics*, 108:273–283, 1993.

Waldman, M. Eliminating the market for secondhand goods: An alternative explanation for leasing1. *The Journal of Law and Economics*, 40:61–92, 2015.

Weyl, E. G. and Zhang, A. L. Depreciating licenses. *American Economic Journal: Economic Policy*, 2022.

Zheng, C. Z. Optimal auction with resale. *Econometrica*, 70:2197–2224, 2002.

Appendix: Proofs

Proof of Lemma 1

Proof. If the support of f is compact, there exists a unique, bounded $W(\theta)$ by usual contraction arguments. In the case where the support is the real line, the payoff to the planner cannot exceed the payoff from the discounted return to the current type's consumption plue every future arrival receiving (there own copy of) the object forever. Since the expected payoff from each arrival for the planner is the mean of F, the present discounted value of future arrival is bounded because $r > 0$. Therefore denote this bound by $W(\theta) \leq \theta + c$ for some finite c. Then define $G_W(\theta) = W(\theta) - \theta$. Since the solution to the planners problem is bounded by $\theta + c$, it must be the case that the solution to the sequence problem can be written this way for some bounded G_W. But □

$$
\begin{aligned}
G_W(\theta) = W(\theta) - \theta &= \lambda max_y \int_y (W(x) - W(\theta)) f(x) dx \\
&= \lambda max_y \int_y (G_W(x) + x - G_W(\theta) - \theta) f(x) dx
\end{aligned}
$$

Since $lim_{y\to\infty} \int_y x f(x) dx$ is bounded as the mean of f is bounded, this maps bounded, continuous functions from $\mathbb{R}_+$to $\mathbb{R}_+$ into such functions, and therefore usual contraction arguments show that it is the unique such function, and therefore is one that describes the solution.

Guess that $W(\theta)=\int_0^\theta \frac{1}{1+\lambda(1-F(x))}dx+C$. Then it is immediate that $y=\theta$ and, differentiating the Bellman equation,

$$\begin{aligned}W'(\theta) &= 1-\lambda(1-F(\theta))W'(\theta)\\ &= \frac{1}{1+\lambda(1-F(\theta))}\end{aligned}$$

which verifies the guess. Since this is increasing in θ, W is convex as asserted.

Proof of Lemma 3

Proof. Since

$$\begin{aligned}(1-d)W^u &= \frac{\lambda(1-F(\Theta(d)))}{1+\lambda(1-F(\Theta(d)))}\frac{\int_{\Theta(d)} W(x)f(x)dx}{1-F(\Theta(d))}\\ &= \lambda d\int_{\Theta(d)} W(x)f(x)dx\end{aligned}$$

the first derivative is proportional to (with constant of proportionality λ)

$$\int_{\Theta(d)} W(x)f(x)dx - d\cdot W(\Theta(d))f(\Theta(d))\Theta'$$

or

$$\int_{\Theta(d)} W(x)f(x)dx - \frac{(1+\lambda(1-F(y)))}{\lambda}W(y)$$

The second derivative can be signed by taking the derivative with respect to y, since y is increasing in d. It is

$$-W(y)f(y)-W'(y)\frac{(1+\lambda(1-F(y)))}{\lambda}+W(y)f(y)<0$$

since $W'(y)>0$. □

Proof of Proposition 4

Proof. Begin with the case where F is compact. Then the fact that $V(\theta)$ is continuous, strictly increasing, greater than θ, with $V(1)=1$ follows directly from contraction arguments: the operator defined by the right hand side of (9) maps continuous, increasing functions greater than θ into (strictly) increasing functions, so the fixed point must be strictly increasing. The choice of y can also be thought of as determining the present discounted duration of ownership d, where $y = \Theta(d)$ is determined from $d = \frac{1}{1+\lambda(1-F(y))}$; therefore alternatively we can write

$$
\begin{aligned}
V(\theta) &= \theta + \lambda(1 - F(\Theta(d)))(V(\Theta(d)) - V(\theta)) \\
V(\theta) &= max_d d\theta + (1-d)V(\Theta(d)) \qquad (9)
\end{aligned}
$$

We break the rest into a series of claims. Let the set of maximizers of (9) be $D(\theta)$. □

Claim. For all $\theta < 1$ and $d \in D(\theta)$, $\frac{1}{1+\lambda(1-F(\theta))} < d < 1$

Proof. Since $V(\theta) > 0$, it cannot be that If either $d \leq \frac{1}{1+\lambda(1-F(\theta))}$, or $d = 1$, since then then $V(\theta) \leq \theta$; the seller could do better by selling to some higher types to get a value that was a convex combination of θ and a higher value. Therefore $y > \theta$. □

Claim. Suppose $\theta' > \theta$. For all $d(\theta') \in D(\theta')$ and $d(\theta) \in D(\theta')$, $d(\theta') \geq d(\theta)$.

Proof. Optimization implies

$$d(\theta')\theta' + (1 - d(\theta'))V(\Theta(d(\theta'))) \geq d(\theta)\theta' + (1 - d(\theta))V(\Theta(d(\theta)))$$

and

$$d(\theta')\theta + (1 - d(\theta'))V(\Theta(d(\theta'))) \leq d(\theta)\theta + (1 - d(\theta))V(\Theta(d(\theta)))$$

Subtract the second from the first:

$$d(\theta')(\theta' - \theta) \geq d(\theta)(\theta' - \theta)$$

so $d(\theta') \geq d(\theta)$. □

Claim. V is convex, and strictly convex except on intervals where there is a constant solution $d(\theta)$

Proof. Let $\theta = \gamma\theta^h + (1-\gamma)\theta^l$ for $\theta^h > \theta_l$ and $0 < \gamma < 1$. Then

$$V(\theta^h) \geq V(\theta) + d(\theta)(\theta^h - \theta)$$

and

$$V(\theta^l) \geq V(\theta) + d(\theta)(\theta - \theta^l)$$

since at those points the seller could choose the same price, and gain or lose the difference in their value for the duration they held the object. But then

$$\begin{aligned}\gamma V(\theta^h) + (1-\gamma)V(\theta^l) &\geq \gamma\left(V(\theta) + d(\theta)(\theta^h - \theta)\right) + (1-\gamma)\left(V(\theta) + d(\theta)(\theta - \theta^l)\right) \\ &= V(\theta)\end{aligned}$$

□

Claim. Suppose V is convex in the problem $max_y(1-F(y))(V(y)-V(\theta))$. Then, for all θ, the solution occurs at a point where $V(y)$ is differentiable.

Proof. Since V is convex, $V(y)$ always has left (V_l') and right hand (V_r') derivatives. Increasing and convex V implies that $0 \leq V_l'(y) \leq V_r'(y)$ So the question is whether it is possible that $V_l'(y) < V_r'(y)$. But for y to be optimal it must be that

$$(1-F(y))V_r'(y) - f(y)(V(y)-V(\theta)) \leq 0$$

and

$$(1-F(y))V_l'(y) - f(y)(V(y)-V(\theta)) \geq 0$$

which is not possible if $V_l'(y) < V_r'(y)$. Therefore $V_l'(y) = V_r'(y)$ so the function is differentiable at y. □

Claim. V is strictly convex and d is strictly increasing.

Proof. V can only be weakly convex on intervals where the choice of y is constant. But since y occurs at a point of differentiability, a necessary condition for optimality is

$$(1-F(y))V'(y) - f(y)\left(V(y)-V(\theta)\right) = 0 \tag{10}$$

Since $V(\theta)$ is strictly increasing this cannot be satisfied for any y and two values of θ.

When the support is unbounded,, since $V(\theta) \leq W(\theta)$, $V(\theta)$ can be bounded by $\theta + c$. Define

$$\begin{aligned}V_X(\theta) = V(\theta) - \theta &= max_d\,(1-d)\left(V(\Theta(d)) - \theta\right) \\ &= max_d\,(1-d)\left(V_X(\Theta(d)) + \Theta(d) - \theta\right)\end{aligned} \tag{11}$$

(11) maps bounded, continuous, positive functions to the same if $(1-d)\Theta(d)$ can be bounded. But

$$(1-d)\Theta(d) = \frac{\lambda(1-F(y))}{1+\lambda(1-F(y))}y$$

which is bounded since $lim_{y\to\infty}(1-F(y))y = 0$ when F has finite mean.

To see that the solution to (11) is the unique solution to (2), and therefore describes the maximum, define $\psi(\theta) = \theta + max_\theta V_X(\theta)$. Uniqueness follows from using the ψ norm as described in Duran (2000). Once the Bellman equation is established, all of the other facts follow exactly as for the case with compact support. 4 □

Proof of Proposition 5

Proof. Rewrite the optimality condition (10) as

$$\frac{1-F(y)}{f(y)} - \frac{V(y)-V(\theta)}{V'(y)} = 0$$

or

$$\frac{1-F(y)}{f(y)} - \frac{V(y)-V(\theta)}{(y-\theta)V'(y)}(y-\theta) = 0 \tag{12}$$

where

$$\frac{V(y)-V(\theta)}{(y-\theta)V'(y)} < 1$$

since V is strictly convex. Now take any $\theta < y \leq p^s$; then since virtual valuations are increasing, $y - \frac{1-F(y)}{f(y)} < p^s - \frac{1-F(p^s)}{f(p^s)} = \theta$ so

$$\frac{1-F(y)}{f(y)} > y-\theta > \frac{V(y)-V(\theta)}{(y-\theta)V'(y)}(y-\theta)$$

so (12) cannot be satisfied. □

Proof of Proposition 6

Proof. To be explicit about the role of λ, write

$$V(\theta,\lambda) = \theta + \lambda max_y(1-F(y))(V(y,\lambda)-V(\theta,\lambda))$$

or

$$V(\theta, \lambda) = max_d d\theta + (1-d)V(\Theta(d), \lambda) \tag{13}$$

Since, for any y, the right hand side of the first equation is higher for higher λ, it is increasing in λ, $V(\theta, \lambda)$ is increasing in λ as well as θ.

Suppose that the $\frac{\partial V}{\partial \theta}$ is decreasing in λ. (Note that at points of non differentiability, this can be stated in terms of directional derivatives both being decreasing in λ.) Let that be Property P. The following argument shows that the functional equation operator defined by (13) maps functions with Property P on the right hand side into functions with Property P on the left hand side. Since Property P forms a complete metric space, the fixed point of the contraction operator must satisfy Property P.

Since solutions are at a point of differentiability for any λ, they are characterized by the first order condition

$$\theta - V(\Theta(d), \lambda) + (1-d)\frac{dV}{d\theta}\Theta' = 0$$

The solution for d is decreasing in λ if the left hand side is decreasing in λ. Since V is increasing in λ, $-V$ is decreasing. For the second term, Property P implies that $\frac{dV}{d\theta}$ is decreasing in λ. By direct calculation $\Theta' = \frac{1}{\lambda d^2 f(y)}$ is decreasing in λ. Therefore the LHS is decreasing in λ for any d and therefore the solution $d(\theta)$ is decreasing in λ. Since $\frac{\partial V}{\partial \theta} = d(\theta)$, this implies that Property P is indeed satisfied for any value function generated from one where Property P holds, and therefore the fixed point of the contraction operator satisfies property P. $\square$

Proof of Proposition 7

Proof. Differentiating the Bellman equation, $\frac{dV(\theta,\tau)}{d\tau} = -s(\theta, \tau)$. The first order condition for τ' is

$$\frac{dV(y,\tau')}{d\tau'} + \int_y s(x,\tau')f(x|x>y)dx = -\tau' \int_y \frac{ds(x,\tau')}{d\tau'} f(x|x>y)dx$$

By the envelope condition, $\frac{dV(\theta,\tau)}{d\tau} = -s(\theta, \tau) \leq 1$, so marginal return to y is increasing in τ, so y is increasing in τ. This implies that s is decreasing in τ so the integrand on the right hand side of the first order condition is negative. But since $\frac{dV(y,\tau')}{d\tau'} = -s(y,\tau')$ the left hand side is

$$-s(y,\tau') + \int_y s(x,\tau')f(x|x>y)dx$$
$$\leq - s(y,\tau') + \int_y s(y,\tau')f(x|x>y)dx = 0$$

Therefore $\tau' \leq 0$ to make the RHS negative. Since $d(\theta,\tau) < 1$, $\tau' = 0$ implies $s(y,0) > 0$, and since y is less than the maximum value of θ the inequalities must be strict, so $\tau' < 0$. □

To extend the result to an ad valorem royalty, n owner of type θ facing a royalty τ and setting y and future royalty τ' solves

$$V(\theta,\tau) = \theta + \lambda max_{y,\tau'}(1-F(y)\left((1-\tau)(V(y,\tau') + \tau'\int_y s(x,\tau')R(x,\tau')f_y(x)dx) - V(\theta,\tau)\right)$$

Where

$$R(\theta,\tau) = V(y(\theta,\tau),\tau'(\theta,\tau)) + \tau'(\theta,\tau)\int_{y(\theta,\tau)} s(x,\tau'(\theta,\tau))R(x,\tau'(\theta,\tau))$$

The first term is the price collected in state θ,τ from the next buyer, $p = V(y,\tau')$ and the second term is royalties collected from the future owner. Alternatively:

$$\begin{aligned} V(\theta,\tau) &= \theta + \lambda(1-F(y(\theta,\tau))\left((1-\tau)(R(\theta,\tau) - V(\theta,\tau)\right) \\ &= d(\theta,\tau)\theta + (1-\tau)s(\theta,\tau)R(\theta,\tau) \end{aligned}$$

Then the envelope condition is

$$V_2(\theta,\tau) = -s(\theta,\tau)R(\theta,\tau)$$

The first order condition for τ' is

$$V_2(y,\tau') + \int_y s(x,\tau')R(x,\tau')f(x)dx + \tau'\int \frac{d\left(s(x,\tau')R(x,\tau')\right)}{d\tau} = 0$$

In order to follow the same steps as in Proposition 7 we need to show that

Lemma 12. *$s(\theta,\tau)R(\theta,\tau)$ is decreasing in θ and τ*

Proof. For decreasing in θ, take θ and θ^+ with $\theta^+ > \theta$ and suppose that $s(\theta^+,\tau)R(\theta^+,\tau) > s(\theta,\tau)R(\theta,\tau)$. Then by optimality, using

$$d(\theta,\tau)\theta + (1-\tau)s(\theta,\tau)R(\theta,\tau) \geq d(\theta^+,\tau)\theta + (1-\tau)s(\theta^+,\tau)R(\theta^+,\tau) \quad (14)$$

and

$$d(\theta^+,\tau)\theta^+ + (1-\tau)s(\theta^+,\tau)R(\theta^+,\tau) \geq d(\theta,\tau)\theta^+ + (1-\tau)s(\theta,\tau)R(\theta,\tau) \quad (15)$$

But then, taking the LHS of the (15) minus the RHS of (14), which must be greater than the RHS of (15) minus the LHS of the (14):

$$d(\theta^+,\tau)(\theta^+ - \theta) \geq d(\theta,\tau)(\theta^+ - \theta)$$

so $d(\theta^+,\tau) \geq d(\theta,\tau)$, but then clearly (14) is violated since both terms are larger on the RHS.

Similarly, for decreasing in τ, take τ and τ^+ with $\tau^+ > \tau$. Then by optimality

$$d(\theta,\tau)\theta + (1-\tau)s(\theta,\tau)R(\theta,\tau) \geq d(\theta,\tau^+)\theta + (1-\tau)s(\theta,\tau^+)R(\theta,\tau^+)$$

and

$$d(\theta,\tau^+)\theta + (1-\tau^+)s(\theta,\tau^+)R(\theta,\tau^+) \geq d(\theta,\tau)\theta + (1-\tau^+)s(\theta,\tau)R(\theta,\tau)$$

Taking the LHS of the first minus the RHS of the second, which is greater than the RHS of the first minus the LHS of the second:

$$(\tau^+ - \tau)s(\theta,\tau)R(\theta,\tau) \geq (\tau^+ - \tau)s(\theta,\tau^+)R(\theta,\tau^+)$$

so $s(\theta,\tau)R(\theta,\tau) \geq s(\theta,\tau^+)R(\theta,\tau^+)$. □

Proof of Lemma 9

Proof. Conditional on a cutoff, any continuation plan for new owners as a function of $\theta > \theta_{h^t}(\tau)$ is equally feasible and doesn't impact duration for prior innovators given the cutoff; therefore the payoff for the planner cannot vary with h^t for each θ and we can write the expected payoff as $\omega(\theta_\tau)$. The expected payoff from the next arrival, given a sequence of cutoffs $\theta(t)$ is therefore

$$\int e^{-\tau} g_{\theta(t)}(\tau)\omega(\theta(\tau))d\tau$$

where $g_{\theta_t}(\tau)$ is the probability distribution over next transaction given $\theta(t)$. Duration is

$$\int (1-e^{-\tau})g_{\theta_t}(\tau)d\tau$$

Define the measure $\mu(\theta)$ for any measurable subset A of [0,1]:

$$\mu(A) = \int e^{-\tau}\chi(\theta(\tau) \in A)g_{\theta(t)}(\tau)dt$$

so that duration under $\theta(\tau)$ is $d = 1 - \mu([0,1])$. This is the (discounted) measure of instants when cutoffs in A is implemented. We can therefore write the planner's payoff as

$$\int \omega(\theta)d\mu(\theta)$$

Now suppose the seller draws a fixed cutoff from a measure defined by

$$\Delta(A) = \int_A (1/\int g_\theta(t)dt)d\mu(\theta)$$

where g_θ is the distribution over arrival times for a fixed cutoff θ. Then their expected payoff is identical to payoff from $\theta(t)$:

$$\int \omega(\theta)(\int g_\theta(t)dt)d\Delta(\theta) = \int \omega(\theta)d\mu(\theta)$$

and duration provided is the same,

$$\int (1-\int f_\theta(t)dt)d\Delta(\theta) = 1-\mu([0,1]) = d$$

The final step is to show that Δ is a probability measure. Suppose that $\theta(t)$ is a step function taking on two values, θ_1 for $[0,\bar{t}]$ and θ_2 for $[\bar{t},\infty)$. Then, using $G()$ as the cumulative density for g

$$\mu(\theta_1) = (1-G_{\theta_1}(t))\int g_{\theta_1}(t)dt$$

$$\mu(\theta_2) = G_{\theta_1}(t)\int g_{\theta_2}(t)dt$$

and so,

$$\Delta([0,1] = \frac{(1-G_{\theta_{\cdot 1}}(t))\int g_{\theta_1}(t)dt}{\int g_{\theta_1}(t)dt} + \frac{G_{\theta_1}(t)\int g_{\theta_2}(t)dt}{\int g_{\theta_2}(t)dt} = 1$$

The extension to $\theta(t)$ that has N steps, i.e. is a simple function, is immediate. Taking a sequence of simple functions $\theta_n(t) \to \theta(t)$, with the associated measures Δ_n, then since $\int d\Delta_n = 1$ for all n, $\int d\Delta = \Delta([0,1]) = 1$ by monotone convergence for the measure Δ defined by $\theta(t)$. □

Proof of Proposition 10

Proof. Since

$$J^u(\theta) = \int_\theta J(x)f(x)dx/(1-F(\theta))$$

this coincides with (1), except that the planner's payoff θ is replaced with $w(\theta)$. But since that is an increasing function, the choice of d is unchanged. □