Intertemporal Allocation with Unknown Discounting

Justin Burkett* and Kyle Woodward†

December 2020

Click here for latest version

Abstract

We consider the problem faced by a durable good monopolist who can allocate a single good at any time, but is uncertain of buyers' values, future arrival times, and temporal preferences.We derive conditions under which it is optimal for the monopolist to ignore buyer discount rates and to not temporally discriminate. Our analytical approach allows for nonexponential discount rates, and our results also apply when sellers have ambiguity regarding buyers' temporal preferences. Our results provide a novel justification for temporal nondiscrimination when the seller is incompletely informed about buyers' patience.

*Georgia Institute of Technology; justin.burkett@gatech.edu

†University of North Carolina at Chapel Hill; kyle.woodward@unc.edu

We thank Simon Board, Nima Haghpanah, Peter Norman, and Rakesh Vohra, as well as seminar audiences at INFORMS and SEA for valuable feedback and comments.

1 Introduction

A seller in a dynamic environment must evaluate the trade-off between selling today and potentially selling tomorrow. This trade-off raises a surprising number of potential complications relating not only to the seller's direct benefits from selling or delaying, but also to the need to provide incentives for potential buyers to purchase. In response, the literature on dynamic mechanism design has successfully identified optimal selling mechanisms in rich and complex environments.[1], and in this literature it is standard to assume that the buyers all share the same commonly known discount factor. Intuition and empirical work suggest, in contrast, that there may be significant heterogeneity in buyers' time preferences.[2] Our goal in this paper is to generalize the scope of these optimal selling mechanisms by showing how to relax strong assumptions about what the seller knows about the buyers' temporal preferences.

We study a mechanism design problem in which buyers' temporal preferences are heterogeneous and privately known. In our model, buyers are privately informed of their values for receiving a good immediately or at some future date. Buyers' types are multidimensional, and most of our results are flexible enough to incorporate a wide variety of possible time preferences, including exponential and hyperbolic discounting as well as nonstandard temporal preferences.[3] The only substantive restrictions on buyer types are that the set of possible types is finite, and that all buyers weakly prefer to consume today versus any point in the future. Our arguments do not require any assumptions about the ordering of the types, such as that the type space is a lattice, nor do they depend on assumptions guaranteeing the sufficiency of local incentive constraints.[4]

We focus on two natural cases of dynamic sales. In the main case (the *independent arrivals* case), buyers arrive stochastically over time, and the mechanism continues until a sale is made. As it turns out, all intuition in the independent arrivals case may be derived from the *no arrivals* case, in which all potential buyers are present when the mechanism opens and the seller is deciding only when a particular types should receive an allocation.

[1]For example, Board and Skrzypacz [2016] show how to construct an optimal mechanism for the sale of multiple goods to forward-looking buyers who arrive stochastically with private information about their valuations.

[2]Empirical work has shown that exponential discount rates differ across agents (Mischel et al. [1989], Kirby and Maraković [1995], Green and Myerson [2004], Hakimi [2013]), and that even the form of temporal preferences may be unknown (Benzion et al. [1989]).

[3]Our main analysis is carried out under standard exponential discounting. When we analyze non-exponential discounting, buyers' preferences may be time-inconsistent. In this case, we assume that both the seller and buyers have commitment power, and a quantity and transfer agreed to at time $t = 0$ cannot be reneged upon at time $t' > 0$. When immediate sale is optimal, commitment is irrelevant in equilibrium.

[4]Carroll [2012] shows that convexity of the type space, which does not hold in our model, is tightly connected to the sufficiency of local incentive constraints in quasilinear models.

Our main case is related to Board and Skrzypacz [2016], in which a random number of forward-looking buyers arrive in each period, but our model is distinguished by the presence of private information about heterogeneous discount rates.

Our main result (Theorem 1) gives conditions under which the seller optimally does not screen on buyers' temporal preferences, i.e., conditions under which temporal nondiscrimination is optimal.[5] In the settings we study, the optimal mechanism that does not depend on temporal preferences is a Myerson auction, either an immediate optimal auction (in the no arrivals case) or a sequence of identical auctions with high reserve prices (in the independent arrivals case).[6] In other words, we give conditions under which a standard Myerson auction remains optimal in the presence of heterogeneous time preferences, potentially adjusting the optimal reserve price.

The condition we obtain for optimality of temporal nondiscrimination is given in terms of virtual values. Virtual values are a standard tool in the mechanism design literature, but the condition is nonetheless potentially opaque. We therefore derive an intuitive sufficient condition for this optimality, given only in terms of properties of the joint distribution of values and discount rates: temporal nondiscrimination is optimal when buyers with a higher value immediate consumption are more likely to be patient, in the sense that their relative value of future consumption is also likely to be high (Corollary 1). Our sufficient condition is intuitive and corresponds to the notion of first-order stochastic dominance in a multidimensional setting.Taken together, our results show that in a complex environment comparatively a simple condition can suffice to make straightforward and well-understood mechanisms optimal.

Our main result in the independent arrivals case is closely related to an identical condition for the no arrivals case. Mathematical approaches to the no arrivals case are uncomplicated by considerations of which buyer arrived when, and who might arrive in the future. For this reason, in the main text we mostly give intuition for the no arrivals case, and show how to adapt the technical arguments to the independent arrivals case. In the no arrivals case, optimal temporal nondiscrimination is equivalent to a standard Myerson optimal auction. When the seller does not separate discount types, the relevant distribution over value types is the marginal distribution over values, equivalent to the average conditional distribution of values given discount types. Then allocating according to the Myerson rule in our model

[5]The equivalence of these two statements is easily observed. When agents discount future consumption, an mechanism which sells at only a single point in time must sell immediately. When the good is sold immediately or not at all, buyers cannot be screened on discount rate.

[6]When, as is typically assumed, buyers strictly discount the future, the optimal mechanism is a sequence of identical Myerson auctions in the no arrivals case as well. Because buyers discount the future and delay will only lead to identical terms, they participate in the immediate auction or not at all.

amounts to allocating to the highest ex post *average* virtual valuation, where the average is taken over the various time preferences held by a buyer with a particular (first-period) value. The average virtual valuation contrasts with the *conditional* virtual valuation, which is the virtual value conditional on a particular discount type.

If not for incentive constraints, the seller would like to allocate to the buyer with the highest conditional virtual valuation ex post.[7] Whether the conditional virtual valuation exceeds the average virtual valuation for a type depends on the probabilistic dependence between temporal preferences and valuations. If the dependence is positive, meaning higher valuations tend to be more patient, *impatient* types will have relatively high conditional virtual values: it is unlikely that a buyer with a high value is impatient, and therefore impatient high-value buyers receive relatively low information rents. But it is these buyers for whom temporal incentive constraints are binding, because increasing the allocation of an impatient type requires increasing the allocation of at least all of the more patient types with the same valuation. Then a positive statistical dependency between patience and value makes temporal screening relatively costly, and tends to drive allocation to the first period.[8]

A natural reading of the multidimensional mechanism design literature suggests that complete solutions are elusive, and that optimal mechanisms can be unwieldy and complicated. Instead of characterizing solutions for all instances of the problem, we focus on a canonical, clearly feasible rule and determine when that rule is optimal.[9] To do this we employ standard techniques from the theory of linear programming and the theory of feasible flows on a network. Our proof approach employs complementary slackness to identify a system of inequalities in the dual variables on the incentive compatibility constraints, and provide conditions under which this system has a valid solution. One important conceptual advantage of this approach is that we never need to identify the particular incentive constraints that bind, which effectively sidesteps a key difficulty inherent in multidimensional mechanism design.[10]

[7]Because agents typically discount future consumption, deferring consumption reduces an agent's willingness to pay, and is costly to the seller. Without incentive constraints on intertemporal preferences, this drives the mechanism to immediate allocation, and it is optimal to award the item to the agent with the highest virtual value immediately.

[8]As a counterpoint, suppose that patience and value are negatively related, so that high-value buyers strongly prefer immediate consumption while low-value buyers are indifferent across consumption in different periods. The seller can discriminate against high-value buyers today, and against low-value buyers tomorrow.

[9]In this regard, our approach is similar to Manelli and Vincent [2006] and Haghpanah and Hartline [2019].

[10]Our approach even allows for non-local incentive constraints to bind. A standard (but not necessary) requirement for local incentive compatibility to imply global incentive compatibility is that the type space be convex (see, e.g., Carroll [2012] and Archer and Kleinberg [2014]). With discrete type spaces arguments from convexity do not apply. We focus our analysis on a subset of incentive constraints, irrespective of whether or not these are sufficient for global incentive compatibility; global incentive compatibility is immediate when the Myerson auction is optimal.

Our main result is derived under the assumption that the binding constraints are the downward constraints in the discount type space. These constraints prevent buyers from misreporting as less patient than they truly are. Ceteris paribus, less-patient buyers receive earlier allocations, and since more-patient buyers still value immediate consumption more than deferred consumption they have a natural incentive to report a less-patient type. Thus it is intuitive that these constraints might bind. Once we establish optimality of temporal nondiscrimination given this set of potentially binding incentive constraints, it is immediate that temporal nondiscrimination is optimal given the full set of potentially binding incentive constraints: more constraints make discrimination more difficult and less profitable, hence the addition of constraints weakly improves the favorability of temporal nondiscrimination.

After deriving our main result we relax the assumption that only downward constraints bind, and we allow for all possible misreports of discount type. This relaxation comes with some technical trade-offs. By allowing for more constraints to bind and inhibit temporal discrimination, the results under this relaxation are naturally tighter than when we exogenously assume that only downward discount constraints may bind. Furthermore, the new form of the problem allows us to fully consider that the seller must use time—and hence reduce revenue through future allocation—to screen on discount type, in comparison to our main result which is independent of the levels of discount types. However, the relaxation breaks the network flow interpretation of the seller's problem, and requires an analysis based on direct application of Farkas' lemma. Ultimately, our main result is neither weaker nor stronger than our result allowing for all possible misreports of discount type; accordingly, we report both.

In our baseline model discount types are multidimensional, but are derived from a single-dimensional set of exponential discount rates. Our main arguments depend on identifying the set of types more patient than a given type, and are not sensitive to the definition of a discount type; thus only the partial ordering on the discount type space matters for our results, and not the total ordering on the discount rate space. We show that our main results remain valid for any model of discounting, so long as buyers uniformly value immediate consumption no less than future consumption. The analysis under nonstandard discounting is distinguished from the analysis of exponential discounting by the observation that, when discounting is nonstandard, whether one buyer is more patient than another may depend on the number of periods available. For standard models of discounting, either one buyer is more patient than another, or there is some number of periods such that one buyer is initially more patient than another, but later the ranking is reversed. This observation, plus the fact that our main analysis assumes that only downward discount constraints may bind, implies that there is a number of periods sufficient to apply our results and show that temporal discrimination is never optimal.

Finally, an immediate consequence of our analysis is that a seller who is uncertain the statistical relationship between value and patience should optimally rely on a temporally nondiscriminatory mechanism, so long as they believe that our main conditions are potentially satisfied.[11] Given the ambiguity regarding the heterogeneity of individual temporal preferences (see our discussion above), our results are therefore consistent with a relative lack of temporal screening by mechanism designers.[12]

Our results contribute to the ongoing study of why and when auctions are valuable, and how simple mechanisms can persist in relatively complicated settings. In our model the space of available mechanisms, which may discriminate on both value and temporal preference, is large and complex. Nonetheless, the introduction of temporal incentive constraints drives allocation away from utilization of this dimension. Even though a rich set of mechanisms is available, consideration of agents' incentives encourages the use of a relatively simple auction, which does not make use of all (or even most) of the information potentially available to the designer. We believe the interaction between incentive constraints and simplicity merits further study.

Our paper proceeds in Section 2 by defining our model. Section 3 establishes our main result, Section 4 provides applications to standard models of temporal discounting, and Section 5 considers the seller's ambiguity regarding the distribution of temporal preferences. Related literature is deferred to Section 6, where we use our model to compare our work to that of others. Section 6.1 concludes.

2 Model

A seller offers one unit of an indivisible good for sale to buyers potentially arriving over time. Time is discrete, $t \in \{0, 1, \dots, T\}$, where T is possibly infinite; allocation may take place in any period. To simplify notation, where there is no chance of confusion we write $t \in T$ for $t \in \{0, 1, \dots, T\}$. The seller commits to a mechanism in the first period, $t = 0$, and discounts the future at (exponential) rate δ_s.

Buyers arrive (potentially) stochastically over time. There are n *kinds* of bidders, $i \in \{1, \dots, n\}$, and the probability a bidder of type i arrives in period τ is $g_i(\tau)$. We consider two cases of bidder arrival. In the *no arrivals* case, one buyers of each kind is present at time $t = 0$, and $g_i(\tau) = 1[\tau = 0]$. In the *independent arrivals* case time is infinite, $T = \infty$,

[11] Carroll [2017] establishes a version of this claim.

[12] Additionally, a lack of temporal screening by mechanism designers is consistent with little learning about temporal preferences, which in turn supports ambiguous beliefs regarding temporal preferences.

and a buyer of kind i arrives in period τ with probability $g_i(\tau) = g_i$, independent of τ.[13] Where there is no confusion, we refer to a buyer of kind i simply as buyer i. Regardless of the arrival structure, the seller observes the number of buyers who arrive in period τ but the buyers do not, and a buyer who arrives in period t remains until the game ends, either via allocation or reaching period T.[14] Buyer i has *value type* $v^i \in \mathcal{V}^i \subset [0,1]$ and discount factor $\hat{\delta}_i \in [0,1]$; the buyer's *discount type* is $\delta^i = (1, \hat{\delta}_i, \ldots, \hat{\delta}_i^T) \in \mathcal{D}^i$.[15] The buyer discounts consumption at time t by discount factor δ^i_t. Each buyer's utility is quasilinear in transfers, and if her allocation is $q \in [0,1]^{T+1}$ her interim utility is

$$u\left(q, p \middle| v, \delta\right) = q \cdot \delta v - p.$$

We assume that the support of value types $\mathcal{V}$ and the support of discount types $\mathcal{D}$ are both finite, and for simplicity we assume further that $\mathcal{V}^i = \mathcal{V} \equiv \{\varepsilon, 2\varepsilon, \ldots, 1-\varepsilon, 1\}$ for some $\varepsilon > 0$. The buyer's (value-relevant) type space is $\Theta^i \equiv \mathcal{V} \times D^i$. The realized type $(v^i, \delta^i) \in \Theta^i$ is known only by agent i, and buyers' types are independently distributed.

We assume that each buyer's private type (v, δ) is statistically indepedent of her arrival time τ, and we define $f^i(v, \delta)$ to be the (commonly known) probability that buyer i has type (v, δ).[16] Let $f^i(v) \equiv \sum_{\delta \in \mathcal{D}^i} f^i(v, \delta)$ so that $F^i(v) \equiv \sum_{v' \leq v} f^i(v')$ is the cumulative (marginal) distribution of valuation types for buyer i. Similarly, let $f^i(\delta) \equiv \sum_{v \in \mathcal{V}} f^i(v, \delta)$ and $f^i(v|\delta) \equiv f^i(v, \delta)/f^i(\delta)$ so that $F^i(v|\delta) \equiv \sum_{v' \leq v} f^i(v'|\delta)$ is the cumulative (marginal) distribution of valuation types for buyer i, conditional on her having discount type δ. Define $f^i(\delta)$, $f^i(\delta|v)$, $F^i(\delta)$, and $F^i(\delta|v)$ analogously. To distinguish random variables, we add a tilde. For example, $\tilde{\theta} = (\tilde{v}^i, \tilde{\delta}^i)$ is the random variable corresponding to buyer i's type, and $\tilde{\tau}$ is the random variable corresponding to her arrival time. We use $\mathbb{E}_i$ for the expectation taken with respect to buyer i's random attributes.

In the *symmetric case*, there are Θ and f such that $\Theta^i = \Theta$ for all $i \in \mathcal{I}$, and $f^i(v, \delta) = f(v, \delta)$ for all $(v, \delta) \in \Theta$. When discussing the symmetric case, we drop the i superscripts.

[13] The independent arrivals case incorporates the case where buyers have symmetric type distributions, and the number of buyers arriving in period τ is i.i.d. over time.

[14] The partial-observability assumption simplifies our analysis but is otherwise inessential. See Board and Skrzypacz [2016] for a related argument.

[15] In our main results we consider only standard exponential discounting, and $\mathcal{D}^i$ can be identified with the set of feasible discount factors $\hat{\delta}_i$. In our extensions, we allow general discount types δ^i, in which case $\mathcal{D}^i$ may be a fundamental. See our discussion of parameterized discounting in Section 4.2.

[16] We distinguish the buyer's exogenously-determined arrival time τ from the passage of clock time t. The former is a fundamental characteristic of an agent, the latter is a tool to be possibly employed by the mechanism designer.

2.1 Mechanisms

Because our seller can commit to a mechanism ex ante, the revelation principle applies. It is without loss of generality to consider direct mechanisms in which the buyers' type reports determine a probability of receiving the good in each period and an expected payment to be made to the seller. We let $q_t^i(v, \delta, \tau)$ denote the (interim) probability that buyer i receives the good in period t having reported the type (v, δ) and arrived in period τ. We use $q_t^i(v, \delta, \tau) \in [0, 1]^{T+1}$ to indicate the vector of probabilities across time periods. We assume without loss of generality that the seller collects a payment of $p^i(v, \delta, \tau)$ from the buyer who reports the type (v, δ) immediately after arriving in period τ and regardless of whether the good is allocated in period τ or later.[17] If buyer i has type (v, δ) and arrives in period τ, and reports the type (v', δ'), her expected payoff from the mechanism is therefore

$$u^i(v', \delta' | v, \delta, \tau) \equiv q^i(v', \delta', \tau) \cdot \delta v - p^i(v', \delta', \tau),$$

where $\cdot$ is a dot product. We use $u^i(v, \delta, \tau) \equiv u^i(v, \delta | v, \delta, \tau)$ for the equilibrium payoff of the type (v, δ) bidder who arrives in period τ.

Remark 1. *The seller may screen on arrival time by offering different mechanisms to buyers who arrive at different times. Because the seller observes buyer arrivals, there is no incentive compatibility constraint associated with such screening, and there is a fundamental distinction between the seller-unobserved type $(v, \delta) \in \Theta^i$ and the seller-observed arrival time τ. Under the conditions identified in our main results, the optimal mechanism in this setting does not screen on the time a buyer arrives. Since buyers (weakly) discount the future, the optimal mechanism satisfies incentive compatibility even when arrival times are private information.*

2.2 Seller's problem

Our analysis considers when the seller (optimally) sells the good irrespective of buyers' discount types. If a mechanism's allocation depends only on buyer values and not on discount types, we say that the mechanism *does not temporally discriminate.* In the no arrivals case, the optimal mechanism without temporal discrimination is an optimal static auction: at time $t = 0$, bidders have highest willingness to pay for consumption at time $t = 0$, and deferring a buyer's consumption is costly for the seller. In the independent arrivals case, the optimal mechanism without temporal discrimination is identified by a period-by-period application

[17] Restricting transfers between the seller and a buyer to the period in which the buyer arrives rules out the possibility of unbounded utility through intertemporal transfers.

of Myerson's rule, with an adjusted cutoff type:[18] a buyer has a positive probability of receiving a good only in the period in which she enters the mechanism.

The general revenue maximization problem is

$$\begin{aligned}
\max_{\{q^i,p^i\}_{i\in I}} \quad & \sum_{i\in I}\sum_{\tau\in T}\mathbb{E}_{\Theta^i,\tau}\left[\delta_s^{\tau}p_{\tau}^{i}\left(\tilde{v},\tilde{\delta},\tilde{\tau}\right)\right] \\
\text{s.t.} \quad & u^i\left(v,\delta,\tau\right)\geq u^i\left(v',\delta'|v,\delta,\tau\right) \quad \forall i\in I,\tau\in T,(v,\delta)\in\Theta^i,(v',\delta')\in\Theta^i \qquad \text{(IC)}\\
& u^i\left(v,\delta,\tau\right)\geq 0 \quad \forall i\in I,\tau\in T,(v,\delta)\in\Theta^i, \qquad \text{(IR)}\\
& q^i \text{ is feasible.} \qquad \text{(F)}
\end{aligned}$$

Since we formulate the problem using interim allocation probabilities and payments, we require a set of feasibility constraints on the $q_t^i(v,\delta)$. By definition it must be that $q^i(v,\delta)\in[0,1]^{T+1}$ and $0\leq\sum_t q_t^i(v,\delta,\tau)\leq 1$ for all types (v,δ) and arrival times τ, but we also require that the probabilities be consistent with the type distribution. Here we use the characterization of these feasibility conditions (the Border constraints) developed in Che et al. [2013], building on the previous work of Border [1991, 2007].

Definition 1. *Given type $(v,\delta)\in\Theta^i$, buyer i's* conditional virtual value *is*

$$m^i\left(v|\delta\right)=v-\left[\frac{1-F^i\left(v|\delta\right)}{f^i\left(v|\delta\right)}\right]\varepsilon.$$

Buyer i's average virtual value *is*

$$m^i\left(v\right)=\mathbb{E}_i\left[m^i\left(\tilde{v}\middle|\tilde{\delta}\right)\middle|v\right]=v-\left[\frac{1-F^i\left(v\right)}{f^i\left(v\right)}\right]\varepsilon.$$

We assume that the unconditional virtual value $m^i(\cdot)$ is weakly increasing. For the no arrivals case, denote by $v_i^\star$ the minimum value type with a strictly positive average virtual value, $v_i^\star=\min\{v\in\mathcal{V}\colon m^i(v)>0\}$. When the seller is constrained to allocate in the first period, discount types are irrelevant (since $\delta_0^i=1$ for all δ^i), and the optimal allocation rule without temporal discrimination is equivalent to allocating to the agent with the highest average virtual valuation.[19] Because the value-relevant type (v,δ) is independent of arrival time τ, virtual values do not depend on arrival time.

In the no arrivals case, the threshold virtual value $v_i^\star$ (essentially) defines the implemen-

[18] See, e.g., Board and Skrzypacz [2016].

[19] When the virtual value curve is nonmonotone, ironing may be necessary [Mussa and Rosen, 1978, Myerson, 1981], and it may not be possible to allocate to the agent with the highest (unironed) virtual value. With slight adjustment of our feasibility constraints and the cutoff type $v_i^\star$ below, our results extend to the case with nonmonotone virtual values.

tation of an optimal nondiscriminatory mechanism: buyers with values $v \geq v_i^\star$ participate in a standard auction for the good, and other buyers cannot receive the good.[20] In the independent arrivals case, where buyers arrive independently and stochastically over time, the optimal mechanism is definable in terms of a slightly higher threshold $v_i^\star$ satisfying

$$v_i^\star = \min \left\{ v' \colon m^i \left(v_i^\star \right) \geq \delta \mathbb{E} \left[\max \left\{ m^i \left(v \right), m^i \left(v_i^\star \right) \right\} \right] \right\}.$$

For details, see equation (6) of Board and Skrzypacz [2016]. Because $v_i^\star$ serves the same role in both the arrivals and no arrivals cases, we do not provide separate notation for the two thresholds. In slight abuse of notation, we write $v_i^\star(v) = \max\{v, v_i^\star\}$ to be the larger of v and $v_i^\star$.

To provide conditions under it is optimal to not temporally discriminate even though allocation of the good is feasible in any period, we begin with a few simple observations. It is obvious that temporal non-discrimination is feasible. To prove that it is optimal, it is sufficient to find a subset of the (IC), (IR) and (F) constraints under which it is optimal, since any ignored constraints are satisfied implicitly. That is, temporal non-discrimination is incentive compatible, individually rational, and feasible, so if it is uniquely optimal given a subset of the constraint set, it is uniquely optimal given the full constraint set. Our results in the following section are primarily distinguished by the particular subset of constraints that we choose to impose on the seller's problem.

3 Analysis

We now derive results establishing when it is optimal for the seller to disregard information on buyers' discount rates. We begin by stating our main result (Theorem 1) and a natural corollary (Corollary 1), and we then provide intuition for these results in the no arrivals case, where all potential buyers are present at time $t = 0$. Analyzing the no arrivals case allows us to shut down complications such as deferring allocation in hopes of receiving better buyers in later periods. As we show, straightforward alterations of the argument in the no arrivals case establish the general result.

We make use of the following definition in Theorem 1. For $v \in \mathcal{V}$ define

$$\mu^i \left(v, \delta \right) \equiv \left(m^i \left(v | \delta \right) - m_+^i \left(v \right) \right) f^i \left(v, \delta \right),$$

[20]If there are value types $v < v_i^\star$ such that $m^i(v) = 0$, there will be multiple optimal mechanisms, some of which allocate to types with zero virtual value and some of which do not. This nonuniqueness is appropriately handled in our proofs and has no effect on our results. For expositional purposes we therefore assume that the mechanism allocates only to types with strictly positive virtual value; this mechanism selection is unique.

where $m^i_+(v) = \max\{m^i(v), m^i(v^\star_i)\}$. We define $\mu^i(v', \delta) = 0$ for $v' \notin \mathcal{V}$. The quantity $\mu^i(v, \delta)$ measures the difference between the virtual value of type (v, δ) and the (truncated) average virtual value of types with the same value, weighted by the probability that (v, δ) occurs. As we discuss below, the term $m^i_+(v)$ captures the implicit cost of deviating from temporal nondiscrimination.

Theorem 1. *In the independent arrivals case, the optimal mechanism does not temporally discriminate if for all buyers i and all types $(v, \delta) \in \Theta^i$,*

$$\sum_{\delta' \geq \delta} \left\{ \mu^i \left(v^\star_i \left(v \right), \delta \right) - \mu^i \left(v - \varepsilon, \delta \right) \right\} \geq 0. \tag{1}$$

For one interpretation of (1), note that a sufficient condition for (1) is

$$\sum_{\delta' \geq \delta} \left\{ \mu^i \left(v, \delta' \right) - \mu^i \left(v - \varepsilon, \delta' \right) \right\} \geq 0, \ \forall v \in \mathcal{V}. \tag{2}$$

Summing up (2) over all $v' \in \{v, \ldots, v^\star_i(v)\}$ gives inequality (1). Observe that summing (2) across $v' \geq v + \varepsilon$ for some v yields $-\sum_{\delta' \geq \delta} \mu^i(v, \delta) \geq 0$, which is equivalent to

$$\mathbb{E}_{\Theta^i} \left[m^i_+ \left(v \right) - m^i \left(v \middle| \tilde{\delta} \right) \middle| v, \tilde{\delta} \geq \delta \right] \geq 0. \tag{3}$$

That is, temporal nondiscrimination is optimal if average censored virtual values weakly exceed conditional virtual values, where the average is taken over all types more patient than a given type. In a posted price mechanism, expected revenue is the sum of (average) virtual values of types who purchase, and the optimal price sells only to types with positive virtual value.[21] An upper bound on the revenue lost, in comparison to temporal nondiscrimination, by screening on type δ is given on the left-hand side of (3). Arrival times τ do not feature in any of these conditions, since the distribution of type (v, δ) is independent of τ, and arrivals are identically distributed across time periods. More directly, when inequality (3) is satisfied the seller can only improve revenue in a nondiscriminatory mechanism by discriminating against relatively patient buyers. This discrimination requires deferring the allocations of patient buyers, who must buy on worse terms than impatient buyers. As patient buyers are at least as willing to consume in every period as impatient buyers, profitable discrimination violates incentive compatibility, and is therefore not feasible.

The inequalities in (1) and (3) are given in terms of virtual values. While intuition from virtual values is standard in the mechanism design literature, we show now that a simpler,

[21] In this intuition, we elide the possibility of ironing.

purely statistical condition is sufficient. For example, inequality (1) is satisfied if buyer types with higher values are more likely to have higher discount types in the following sense.

Corollary 1. *In the independent arrivals case, the optimal mechanism does not temporally discriminate if, for all buyers i and all $(v, \delta) \in \Theta^i$, $F^i(\delta|v)$ is nonincreasing in v.*

The condition given in Corollary 1 can be understood as a first-order stochastic dominance condition on the distribution of discount types in response to an increase in the valuation. Corollary 1 also implies that temporal nondiscrimination is optimal when discount types are common knowledge.[22] When Corollary 1 is satisfied, value types are positively correlated with discount types, but in general correlation is not sufficient for the optimality of temporal nondiscrimination.

3.1 The no arrivals case: intuition and an example

Intuition for our results is most easily understood in the no arrivals case, where all potential buyers are present in the initial period. Consider the symmetric case where there are two discount types, a patient type δ and an impatient type δ' with $\delta \geq \delta'$. Suppose $m^i(v|\delta') > m^i(v|\delta)$. This relation is determined by the probabilistic dependence of v on δ. Intuitively, $m^i(v|\delta)$ is lower than $m^i(v|\delta')$ if for $v' > v$ there are relatively more (v', δ) types than there are (v', δ') types, because the virtual valuation incorporates the necessary reduction in payment by all higher value types when one increases the allocation of that type.[23] There are relatively more (v, δ) types when value and patience are positively dependent (recall that we assumed that $\delta \geq \delta'$). Notice that $m(v|\delta') > m(v|\delta)$ if and only if

$$\begin{aligned} (1 - f(\delta|v)) \sum_{w>v} f(\delta|w) f(w) &> f(\delta|v) \sum_{w>v} (1 - f(\delta|w)) f(w) \\ \iff \qquad \sum_{w>v} f(\delta|w) f(w) &> f(\delta|v) \sum_{w>v} f(w) . \end{aligned}$$

The first line uses the definition of $m(v|\delta)$ and the identity $f(\delta'|v) + f(\delta|v) = 1$. The inequality holds for all v if $f(\delta|\cdot)$ is nondecreasing.

[22] This does not imply the "no haggling" result of Riley and Zeckhauser [1983]. We assume that $\delta_1^i = 1$ for all $\delta^i \in \mathcal{D}^i$, so each buyer is willing to consume immediately, which rules out correlated random arrivals. Our results are similarly distinct from the dynamic pricing literature; see our discussion of related literature in Section 6.

[23] This tradeoff is familiar from other mechanism design contexts, and is equivalent to a standard monopoly pricing argument: to increase the probability of a sale, the monopolist must reduce the posted price. Reducing the posted price uniformly increases the utility of every buyer with value above the posted price. We do not assume a posted price mechanism.

When patience and value are positively dependent, the seller cannot increase the allocation of the patient buyer without also increasing the allocation of the patient buyer. When discount type δ is more patient than discount type δ', $\delta \geq \delta'$, the (v, δ) type has a weakly higher value for (v, δ')'s allocation across all time periods. Any increase in (v, δ')'s allocation without a corresponding increase in (v, δ)'s will violate the (IC) constraint on misreport of discount type. The ultimate conclusion is that the seller can do no better than to pool the two types together. When he does this, he receives the average virtual value from each type, $m(v)$.

To extend the logic of the previous example to more discount types, consider how incentive compatible reallocations from temporal nondiscrimination influence the seller's payoff. Specifically, consider increasing the allocation of all discount types with some value v by some amount ξ. Assuming that the (IC) constraints bind from higher to lower discount types, the allocation of some type (v, δ) can be increased as long as it is increased for all types (v, δ') with $\delta' \geq \delta$ as well. This suggests that a reallocation will be profitable if there is a (v, δ) such that

$$\mathbb{E}\left[m^i\left(v\middle|\tilde{\delta}\right)\middle|v, \tilde{\delta} \geq \delta\right] > m^i\left(v\right). \tag{4}$$

The left-hand side is the marginal increase in expected revenue from increasing the allocation of (v, δ) and all higher (i.e., more patient) discount types, and the right-hand side is the marginal increase in expected revenue from increasing the allocation to all types with valuation v, irrespective of discount type. Notice that the opposite of inequality (4)—which will hold when increasing the allocation to (v, δ) is not advantageous—is nearly equivalent to (3).

We now give a version of Theorem 1 for the no arrivals case.

$$\mu^i\left(v, \delta\right) \equiv \left(m^i\left(v|\delta\right) - m^i_+\left(v\right)\right) f^i\left(v, \delta\right),$$

Theorem 2. *In the no arrivals case, the optimal mechanism does not temporally discriminate if for all buyers $i \in I$ and all types $(v, \delta) \in \Theta^i$,*

$$\sum_{\delta' \geq \delta} \left\{\mu^i\left(v^\star_i\left(v\right), \delta'\right) - \mu^i\left(v - \varepsilon, \delta'\right)\right\} \geq 0. \tag{5}$$

Because Corollary 1 follows algebraically from Theorem 1, it remains valid in the no arrivals context.

Corollary 2. *In the no arrivals case, the optimal mechanism does not temporally discriminate if, for all $i \in I$ and all $(v, \delta) \in \Theta^i$, $F^i(\delta|v)$ is nonincreasing in v.*

The primary complication in establishing Theorem 2 (and, hence, Theorem 1) is that we do not know *a priori* which IC constraints will bind at the optimal selling mechanism. The idea behind the proof is to first introduce dual variables (Lagrange multipliers) associated with each of the IC and feasibility constraints. When temporal nondiscrimination is optimal, we can assign feasible values to the dual variables so that a standard first-order condition is satisfied. We then apply duality to construct a system of inequalities in these dual variables. The system of inequalities represents the requirements imposed by complementary slackness and feasibility on the dual linear program. Each inequality corresponds to a variable $q_0^i(v, \delta^i)$. When there is a solution to the resulting linear system, temporal nondiscrimination is optimal.

The ensuing argument relies on an assignment of a value to the feasibility constraint (F). In line with the intuition given above, we say that the shadow cost of the feasibility constraint is $m_+^i(v)$. That is, a slight slackening of the feasibility constraint for value type v would increase the seller's revenue by $m_+^i(v)$, which takes into account the increased allocation to higher types associated with maintaining incentive compatibility. This point is the only meaningful distinction between Theorem 1 and Theorem 2: in the no arrivals case, feasibility does not need to take into account when a buyer arrives, while in the independent arrivals case it does.

Having translated the conditions under which temporal nondiscrimination is optimal to a system of inequalities in the dual variables, the next step is to determine conditions under which this system has a feasible solution.[24] We show in the proof of Theorem 2 that when we consider a particular subset of the (IC) constraints the problem of determining whether this system of inequalities in the dual variables has a feasible solution is isomorphic to the problem of determining whether there exists a feasible flow in a canonical network flow problem. The subset of constraints we analyze is the set of downward (IC) constraints on the misreport of the discount type, which prevent buyers from misrepresenting as less patient than they are. The network consists of nodes, which are identified with types on our model, and arcs (or links) that carry "flow," where the flow between two types with different discount types determines the value of the dual variable on the corresponding (IC) constraints, and the flow between adjacent value types with identical discount types determines the value of the dual variable on the corresponding positivity constraint. The former dual variable can be interpreted as the shadow price of loosening an (IC) constraint between two types.

The network flow argument in the proof involves results well-known within the network flow literature, and our central insight is an application of Gale [1957]. The results have the

[24]In the primal problem, feasibility is a constraint on allocations. In the dual problem, feasibility is a constraint on the sign of the dual variables, which must be weakly positive.

	$(1,0)$	$\left(1,\hat{\delta}\right)$
1	$\frac{1}{3}$	$\frac{1}{2}\pi$
$\frac{1}{2}$	$\frac{1}{6}$	$\frac{1}{2}(1-\pi)$

$F(v,\delta)$

Figure 1: The type distribution for Example 1.

advantage of being straightforward to describe and understand, and the crux of our argument is in making appropriate definitions so that the seller's problem is represented as a network flow problem. In all, we show that the potentially complex problem of determining which (IC) constraints bind in the optimal solution can be broken down into a series of steps, each of which is comparatively simple.

An example clarifies the relative contributions of Theorem 2 and Corollary 2.

Example 1. *Consider the no arrivals case with a single buyer, $n=1$. The type space is $\Theta=\{1/2,1\}\times\{0,\hat{\delta}\}$, where $\delta\in(0,1]$. There are two time periods, $t\in\{0,1\}$. The joint distribution F over valuation and discount types is parameterized by probability $\pi\in[0,1]$ and is shown in Figure 1. If discount types are common knowledge, the seller will offer a price of $p^\star(0)=1$ to a buyer with discount type $\delta=0$ and a price of $p^\star(\hat{\delta})=1$ to a buyer with discount type $\delta=\hat{\delta}$ when $\pi\geq 1/2$ and a price of $p^\star(\hat{\delta})=1/2$ otherwise.*

Applying Theorem 2, we find that temporal nondiscrimination is optimal if

$$-\mu^i\left(\frac{1}{2}\middle|\hat{\delta}\right)\geq 0,\ \text{and}\ -\mu^i\left(\frac{1}{2}\middle|0\right)-\mu^i\left(\frac{1}{2}\middle|\hat{\delta}\right)\geq 0.$$

Straightforward calculation gives that temporal nondiscrimination is optimal when $\pi\geq 1/2$; see Appendix B for additional details. Intuitively, because the optimal sales mechanism is independent of discount type when $\pi\geq 1/2$, there is no incentive to screen buyers on their discount types.

Applying Corollary 2, we find that $F(\delta|v)$ is nonincreasing in v when

$$\frac{2}{2+3\pi}\leq\frac{1}{4-3\pi}.$$

Then Corollary 2 shows that temporal nondiscrimination is optimal when $\pi\geq 2/3$. Thus Theorem 2 (and, hence, Theorem 1) is strictly more general than Corollary 2 (respectively, Corollary 1).

3.2 General discount misreporting

The proofs of Theorems 1 and 2 consider only the ability of the seller to satisfy a subset of the agents' (IC) constraints, and not whether the seller wants to defer a particular type's consumption. Since deferring consumption exogenously reduces willingness to pay (when $\delta < 1$), the sufficient conditions in Theorems 1 and 2 are overly strong. In this section we take a distinct analytical approach and simultaneously allow for additional binding (IC) constraints while adjusting for the seller's incentives, leading to a characterization distinct from our ealier results.

By allowing for all possible (IC) constraints to bind, including IC constraints from lower to higher types and those between unordered types, the system of linear inequalities we obtain in the proof of Theorem 3 is comparable to the one found in the proof of Theorem 2, but does not correspond to a standard network flow problem. Theorem 3 thus necessitates a different approach to feasibility. In this case, we employ Farkas' Lemma to derive a distinct sufficient condition for the optimality of temporal nondiscrimination.

Theorem 3. *In both the independent and no arrivals cases, the optimal mechanism does not temporally discriminate if for all buyers i and all types $(v, \delta) \in \Theta^i$,*

$$\sum_{\delta' > \delta} \frac{1}{\varepsilon} \left(\mu^i\left(v, \delta'\right) - \mu^i\left(v - \varepsilon, \delta'\right)\right)\left(\delta' - \delta\right) v + \left(1 - \delta'\right) m^i_+\left(v\right) f^i\left(v, \delta'\right) \geq 0. \tag{6}$$

Note that for $v \in \mathcal{V}$ with $m^i(v) > 0$, condition (6) is strictly more general than conditions (1) and (5). However, for $v \in \mathcal{V}$ with $m^i(v) < 0$, condition (6) is neither more nor less general than its equivalents in Theorems 1 and 2.

Although self-evidently more technical than our earlier results, intuition for inequality (6) may be readily drawn from the discussion surrounding Theorem 2. In particular, the difference $\mu^i(v, \delta) - \mu^i(v - \varepsilon, \delta)$ still represents the relative cost of increasing the allocation of type $(v - \varepsilon, \delta)$. Inequality (6) adjusts this cost to account for the fact that increasing the allocation to a given type in an incentive-compatible manner requires increasing the allocation at some time $t \geq 1$, and not at time $t = 0$; thus buyers' discount rates scale the terms μ. Finally, there is an additional term associated with the fact that discrimination must occur in some future period, resulting in a revenue loss proportional to the minimum difference in temporal taste from period $t = 0$, which is $1 - \delta$.

Continuing Example 1 illustrates additional the power of Theorem 3.

Example 1 (continued)**.** *Recall that Theorem 2 implies that temporal nondiscrimination is optimal when $\pi \geq 1/2$, while Corollary 2 implies that temporal nondiscrimination is optimal when $\pi \geq 2/3$.*

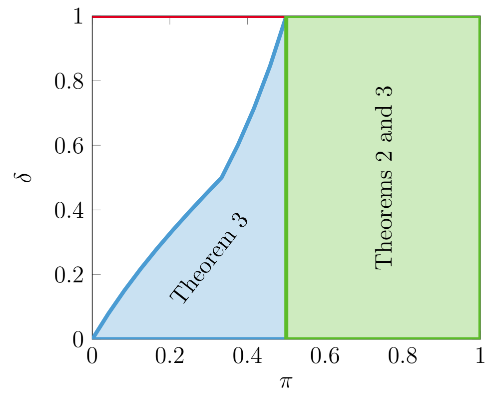

Figure 2: Optimal mechanisms in Example 1. Theorem 3 implies that temporal nondiscrimination is optimal in all shaded regions. Theorem 2 is weaker in this example, and only implies that temporal nondiscrimination is optimal in the rightmost shaded region. The red line ($\delta = 1$ and $\pi < 1/2$) is the unique region where perfect temporal separation is strictly optimal.

On the other hand, Theorem 3 implies that temporal nondiscrimination is optimal when

$$\pi \geq \frac{1}{2}, \text{ or } \frac{1}{2} > \pi \geq \frac{\hat{\delta}}{1+\hat{\delta}}, \text{ or } \frac{1}{3} > \pi \geq \frac{4 - \hat{\delta} - \sqrt{16 - 32\hat{\delta} + 25\hat{\delta}^2}}{6\left(1-\hat{\delta}\right)}.$$

Intuitively, temporal nondiscrimination is optimal when patient buyers are sufficiently likely to have high values, or when even patient buyers significantly discount the future ($\hat{\delta} \approx 0$). The kink at $\pi = 1/3$ arises from the fact that, at $\pi = 1/3$, the optimal temporally nondiscriminating mechanism goes from a posted price of $p_0^\star = 1$ (for $\pi > 1/3$) to a posted price of $p_0^\star = 1/2$ (for $\pi < 1/3$). This relationship is depicted in Figure 2.

The seller will only engage in perfect temporal separation if $\delta = 1$. That is, even if temporal discrimination is optimal, the seller will offer patient buyers some probability of allocation in period $t = 0$ and some probability of allocation in period $t = 1$, unless patient buyers are infinitely patient. This occurs because when $\delta < 1$, deferring consumption reduces the seller's revenue; by shifting some (but not all) of the patient buyer's consumption to date $t = 0$, the seller can improve revenue while ensuring that the impatient buyer does not want to misrepresent his discount type. Thus in the unshaded region of Figure 2 the optimal mechanism sells to all buyers in period $t = 0$, and to patient buyers in period $t = 1$.

In the constrained problem where the seller can only allocate to a buyer in a particular period (and cannot offer lotteries across periods), perfect temporal separation is optimal when patient types are sufficiently patient and are significantly likely to have a low value. That is, *when*

$$0 \leq \pi \leq \frac{1}{3} \text{ and } \delta \geq \frac{2}{3}, \text{ or } \frac{1}{3} \leq \pi \leq \frac{1}{2}\delta.$$

4 Non-exponential discounting

The proof of our main result, Theorem 1, does not depend on the assumption that buyers have exponential discount rates. All our technical results go through provided buyers identically value immediate consumption, $\delta_0 = 1$, no buyer values future consumption more than immediate consumption, $\delta_t \leq 1$, and the seller is constrained to accept transfers only immediately upon arrival.[25] Then, with minimal alterations, our Theorems 1 and 2 remain valid when buyers are non-exponential discounters.[26]

Theorem 4. *Suppose buyers are potentially non-exponential discounters. In both the independent and no arrivals cases, the optimal mechanism does not temporally discriminate if for all buyers i and all types $(v, \delta) \in \Theta^i$,*

$$\sum_{\delta' \geq \delta} \left\{ \mu^i \left(v_i^\star \left(v\right), \delta\right) - \mu^i \left(v - \varepsilon, \delta\right) \right\} \geq 0. \tag{7}$$

Note that inequality (7) is identical to inequality (1) in Theorem 1. Thus Corollary 1 remains valid when buyers are non-exponential discounters.

4.1 Unordered discount types

When discount types are non-exponential, the discount type space may be unordered, and the condition $\delta' \geq \delta$ may be degenerate. Theorem 4 remains nonetheless nontrivial. First, inequality (7) still must consider discount types which are ordered. And, second, for a value type $v \in \mathcal{V}$, all conditional virtual values are compared to the same unconditional virtual value $m^i(v)$, implicitly tying together inequality (7) across multiple discount types.

Define a *discount type chain* to be a set of discount types $\mathcal{D}_C^i$ such that any two $\delta, \delta' \in \mathcal{D}_C^i$ are comparable (either $\delta \geq \delta'$ or $\delta' \geq \delta$) and any $\delta' \in \mathcal{D}^i \setminus \mathcal{D}_C^i$ is incomparable with any $\delta \in \mathcal{D}_C$ (neither $\delta \geq \delta'$ nor $\delta' \geq \delta$).

[25] Economically, buyers who are non-exponential discounters will have temporally inconsistent preferences (Strotz [1955]), which the seller might exploit to improve revenue (Spiegler [2011]). In our model we assume that the seller is constrained to make/accept transfers only when a buyer arrives, and dynamic inconsistency remains irrelevant.

[26] Theorem 3, on the other hand, depends on monotonicity of the expression $w(t; \delta, \delta') \equiv (1 - \delta_t)/(\delta_t - \delta_t')$ in t. This expression is monotone as well when buyers are hyperbolic discounters, and thus Theorem 3 can be extended to certain environments, but in general monotonicity of w is not guaranteed and hence Theorem 1 cannot be immediately extended.

Proposition 1. *Suppose that the discount type space can be partitioned into distinct discount type chains, $\mathcal{D}^i = \cup_{k=1}^{K} \mathcal{D}_k^i$. The optimal mechanism does not temporally discriminate if*

1. *For each $\mathcal{D}_k^i$, $F^i(\delta|v, \mathcal{D}_k^i)$ is nonincreasing in v; and*

2. *For each $\mathcal{D}_k^i$, $F^i(\max \mathcal{D}_k^i|v)$ is constant in v.*

Proof. For any $\delta \in \mathcal{D}_k^i$ and any valuation type $v \in \mathcal{V}$, the set $U_\delta = \{\delta' \in \mathcal{D}^i \colon \delta' \geq \delta\}$ is a subset of D_k^i. Normalizing $F^i(\delta|v)$ by the probability that the realized discount type is in $\mathcal{D}_k^i$ gives the first condition. The second condition ensures that $F^i(\max \mathcal{D}_k^i|v)$ is constant in v, and thus Theorem 1 obtains. □

The following example illustrates Proposition 1.

Example 2. *The seller is allocating in one of three periods. There are two value types, $v \in \{1, 2\}$, and two discount types, $\delta \in \{(1,1,0), (1,0,1)\}$: all buyers are willing to consume at time $t = 0$, but a buyer with discount type $\delta = (1,1,0)$ is willing to consume the object at time $t = 1$ but not at time $t = 2$, and a buyer with discount type $\delta = (1,0,1)$ is willing to consume the object at time $t = 2$ but not at time $t = 1$.*

Because the discount types are unordered, the only upper sets to check in condition (1) *are $-\mu(1,(1,1,0)) \geq 0$ and $-\mu(1,(1,0,1)) \geq 0$. These inequalities will hold if*

$$\max\left\{m^i(1), 0\right\} \geq m^i(v|(1,1,0)), \text{ and } \max\left\{m^i(1), 0\right\} \geq m^i(v|(1,0,1)).$$

These inequalities are trivially satisfied if the seller never wants to allocate to a buyer with value $v = 1$, even if he knows the buyer's discount type. This buyer has a negative conditional virtual value, and temporal screening has no value. In this case, the seller will allocate the good immediately at a price of 2.

Otherwise, these inequalities are satisfied only if $m^i(1|(1,1,0)) = m^i(1|(1,0,1))$. Since $m^i(1) = \mathbb{E}_\delta[m^i(1|\delta)]$, the only way that both conditional virtual values can be weakly below the expected virtual value is if they are equal. Then the seller cannot improve revenues by temporal screening if the optimal sales mechanism is unaffected by knowledge of the discount type.

On the other hand, if $m^i(1|(1,1,0)) < 0 < m^i(1|(1,0,1))$, the seller strictly prefers to screen on time, and will offer a posted price of 2 *at time $t \in \{0, 1\}$ and a posted price of* 1 *at time $t = 2$.*

As with our earlier corollaries, Proposition 1 is weaker than is necessary to ensure optimality of temporal non-discrimination. This is readily observable in application to Example 2: when $F(v|(1,1,0))$ and $F(v|(1,0,1))$ are constant in v, the seller wants to allocate either to both values $v \in \{1, 2\}$, or just to value $v = 2$, even if he has full knowledge of the buyer's discount type.

Corollary 3. *Suppose that the discount type space can be partitioned into distinct discount type chains, $\mathcal{D}^i = \cup_{k=1}^{K} \mathcal{D}_k^i$, and each discount type chain is a singleton, $\mathcal{D}_k^i = \{\delta^{ik}\}$. It is optimal for the seller not to temporally screen if for each discount type δ^{ik}, $F^i(\delta^{ik}|v) = f^i(\delta^{ik}|v)$ is constant in v.*

4.2 Parameterized discounting

When buyer discount types are exponential, our sufficient condition for temporal non-discrimination in the no arrivals case is independent of the number of available periods. Our conditions depend on the set of types more patient than a given type, henceforth *more-patient sets*. A buyer with a higher (exponential) discount rate at time $t = 1$, $\delta_1' > \delta_1$, also has a higher discount rate in all time periods, $\delta_t' > \delta_t$ for all $t \geq 1$, so more-patient sets are independent of the number of periods in the model. In the case of general temporal discounting, it may no longer be the case that a buyer with a higher preference for consumption at time $t = 1$ also has a higher preference for consumption at time $t = 2$. In other words, with general temporal discounting the question of who is more patient may be nontrivially multi-dimensional. We now address this multidimensionality, as well as the question of extending nonexponential discount types to additional periods.

Let B be a (finite) set of *discount parameters*, and let $h : \mathbb{N}_0 \times B \to [0, 1]$ give the discount rate at time t associated with a given discount parameter. That is, a buyer with discount parameter $\beta \in B$ discounts consumption at time t by $h(t; \beta)$. We denote the discount type of a buyer with discount parameter β by $\vec{\delta}\,(\beta; T) = (h(0; \beta), h(1; \beta), \ldots, h(T; \beta))$. We refer to (B, h) as a parameterized model of discounting, and the T-period discount type space generated by (B, h) is $\mathcal{D}_T = \{\vec{\delta}\,(\beta; T) \colon \beta \in B\}$.

For example, exponential and hyperbolic discounting may be generated by a single-dimensional parameter space $B \subseteq \mathbb{R}_+$. Under exponential discounting, for any $\delta \in \mathcal{D}$ there is $\beta \geq 0$ so that $\delta_t = \exp(-\beta t)$, and under hyperbolic discounting, for any $\delta \in \mathcal{D}$ there is $\beta > 0$ so that $\delta_t = 1/(1 + \beta t)$. Quasi-hyperbolic discounting and generalized hyperbolic discounting are typically generated by a two-dimensional parameter space $B \subseteq \mathbb{R}_+^2$: under quasi-hyperbolic discounting, for any $\delta \in \mathcal{D}$ there are $\beta_1 > 0$ and $\beta_2 \in [0, 1]$ such that $\delta = (1, \beta_1\beta_2, \ldots, \beta_1\beta_2^T)$, and under generalized hyperbolic discounting, for any $\delta \in \mathcal{D}$ there are $\beta_1, \beta_2 > 0$ such that $\delta_t = (1 + \beta_2 t)^{-\beta_1}$.

For any discount type δ, let $U_{\mathcal{D}}(\delta)$ be a set consisting of higher discount types in $\mathcal{D}$,

$$U_{\mathcal{D}}(\delta) = \{\delta' \in \mathcal{D} \colon \delta' \geq \delta\}.$$

Given a parameterized model of discounting, the more-patient set $U_{\mathcal{D}_T}(\delta)$ is eventually constant in T (see Lemma 2 in Appendix C). Our main results give sufficient conditions for optimality of temporal nondiscrimination, expressed only in terms of more-patient sets and the distribution of types; these conditions do not depend directly on T. Because the distribution of discount parameters is unchanging as $\mathcal{D}_T$ is extended to $\mathcal{D}_{T'}$,[27] the sufficient condition is unchanging if the more-patient set is unchanging. Then the following is an immediate corollary of Lemma 2.

Corollary 4. *Let (B, h) be a parameterized model of discounting. There is $\bar{T}$ such that if (7) is satisfied for $\mathcal{D}^i_{\bar{T}}$, it is satisfied for $\mathcal{D}^i_{T'}$ for all $T' > \bar{T}$.*

Note that Corollary 4 does not imply that if temporal nondiscrimination is optimal for $\mathcal{D}^i_{\bar{T}}$, it is optimal for $\mathcal{D}^i_{T'}$ for all $T' > \bar{T}$. In particular, we do not address the possibility that temporal nondiscrimination is optimal while the antecedent of Theorem 4 is unsatisfied.

Now let $\mathcal{D}^i_\infty$ be the infinite-period extension of $\mathcal{D}^i$, so that for any $\delta \in \mathcal{D}^i_\infty$ and any $T' \in \mathbb{N}$, $\delta_{(1:T')} \in \mathcal{D}^i_{T'}$. Corollary 4 holds as well when the number of allocation periods is taken to infinity.

Proposition 2. *Let (B, h) be a parameterized model of discounting. In the no arrivals case, there is $\bar{T}$ such that if (7) is satisfied for $\mathcal{D}_{\bar{T}}$, then temporal nondiscrimination is optimal for $\mathcal{D}_T$, for all $T \in \mathbb{N}$.*

One direction of Proposition 2 follows immediately from Corollary 4. Because more-patient sets are eventually constant in T, the set of discount types accounted for in condition (7) is eventually constant in T, and condition (7) is itself therefore either satisfied for all $T' \geq T$, or for no $T' \geq T$. For the other direction, observe that more-patient sets are shrinking in T. Then if condition (7) is satisfied for large T, and hence for relatively small more-patient sets, the relevant over larger more-patient sets may be obtained by summing over smaller more-patient sets. Since each of the latter is weakly positive, so is the former, and temporal nondiscrimination is optimal for all T.

Our model of discounting does not permit for allocation at time $t = \infty$, however under the regularity condition $\vec{\delta}_\infty(\beta; \infty) = \lim_{t \nearrow \infty} \vec{\delta}_t(\beta; t)$, Proposition 2 extends immediately to allocation across times $t \in \mathbb{N} \cup \{\infty\}$.

[27] As $\mathcal{D}_T$ is extended to $\mathcal{D}_{T'}$, the support of the discount type space is changing. Nonetheless, types in $\mathcal{D}_T$ are in one-to-one correspondence with discount parameters $\beta \in B$, which are in one-to-one correspondence with discount types in $\mathcal{D}_{T'}$ and, as a result of Lemma 2, more-patient sets shift in a regular way.

Remark 2. *For hyperbolic discounting, Proposition 2 gives a simple sufficient condition for optimality of temporal nondiscrimination even when the time horizon is infinite. In particular, hyperbolic discounting is generated by a one-dimensional parameter space $B \subseteq \mathbb{R}_+$, and*

$$h^{hyp}(\beta;t) \geq h^{hyp}(\beta';t) \iff \beta \leq \beta'.$$

Then under hyperbolic discounting, more-patient sets are constant for all $T \geq 1$. Proposition 2 implies that if the condition of Theorem 4 is satisfied when $T = 1$, temporal nondiscrimination is optimal no matter the number of periods.

Remark 3. *Under quasi- and generalized hyperbolic discounting, an argument similar to the cases of exponential and hyperbolic discounting applies. However, it may not longer be the case that satisfaction of Theorem 4 when $T = 1$ implies optimality of temporal nondiscrimination when $T \nearrow \infty$. Quasi- and generalized hyperbolic discounting are typically generated by a two-dimensional parameter space $B \subseteq \mathbb{R}^2_+$. In particular,*

$$h^{qhyp}(\beta;t) \geq h^{qhyp}(\beta';t) \iff \frac{\beta_1}{\beta_1'} \geq \left(\frac{\beta_2}{\beta_2'}\right)^{-t}, \quad \textit{and}$$
$$h^{ghyp}(\beta;t) \geq h^{ghyp}(\beta';t) \iff (1+\beta_2 t)^{-\beta_1} \geq (1+\beta_2' t)^{-\beta_1'}.$$

Both of these inequalities depend on t, and more-patient sets in $\mathcal{D}_{T'}$ may not be constant until T' is large, depending on the underlying set of discount parameters. If the discount parameter space is ordered, so that for any feasible parameters $\beta, \beta' \in B$, $\beta_1 \leq \beta_1'$ if and only if $\beta_2 \leq \beta_2'$, it is again the case that satisfaction of Theorem 4 when $T = 1$ implies temporal nondiscrimination when $T \nearrow \infty$.

5 Ambiguous temporal preferences

In practice, it may be difficult for the seller to evaluate the marginal distribution of discount types. We now consider the possibility that the seller knows only the marginal distribution of valuation types. We abstract from buyer ambiguity aversion and assume there is a single buyer, $N = 1$. The seller is ambiguity averse, and optimizes maxmin expected utility [Gilboa and Schmeidler, 1989]. Given a type distribution F, let F_v and F_δ be the marginal distributions of valuation and discount types, respectively, and for the moment assume that the the seller knows the marginal distribution of valuation types F_v and the support of discount types $\mathcal{D}$, but knows neither the joint distribution F nor the marginal distribution F_δ.[28] The seller

[28]Carroll [2017] analyzes the case in which the seller knows F_δ, and finds that (applied to our setting) temporal nondiscrimination is optimal.

believes that the feasible set of joint distributions is $\mathcal{F} \subseteq \{\hat{F} \colon \hat{F}_v = F_v \text{ and } \operatorname{Supp} \hat{F}_\delta = \mathcal{D}\}$. In the no arrivals case the seller's problem is[29]

$$\max_{\{q,p\}} \inf_{F \in \mathcal{F}} \mathbb{E}_F \left[p \left(\tilde{v}, \tilde{\delta}, \tilde{\tau} \right) \right],$$

$$\text{s.t.} \quad u^i (v, \delta, \tau) \geq u^i (v', \delta' | v, \delta, \tau) \quad \forall i \in I, \tau \in T, (v, \delta) \in \Theta^i, (v', \delta') \in \Theta^i \qquad \text{(AIC)}$$

$$u^i (v, \delta, \tau) \geq 0 \quad \forall i \in I, \tau \in T, (v, \delta) \in \Theta^i, \qquad \text{(IR)}$$

$$q^i \text{ is feasible.} \qquad \text{(AF)}$$

Proposition 3. *Suppose that the seller strictly discounts the future, $\delta_s < 1$, and that there is $F \in \mathcal{F}$ that satisfies the condition of Theorem 4. Then temporal nondiscrimination is optimal in the seller's problem with no arrivals and ambiguous temporal preferences.*

When the statistical relationship between valuation and discount types is ambiguous, the seller's (minimum) expected revenue is weakly bounded above by the revenue arising under any given type distribution, including those which satisfy Theorem 4. In this case, revenue is strictly optimized with a temporally nondiscriminatory mechanism. Since temporal nondiscrimination generates the same revenue regardless of the joint distribution of valuation and discount types, the optimal mechanism does not temporally discriminate.

The assumption that the seller strictly discounts the future is mild but essential. In the setting of Example 2 above, when the seller does not discount the future temporal discrimination is only weakly optimal; the seller could obtain identical revenue by selling in periods $t \in \{1, 2\}$ rather than immediately in period $t = 0$. When the seller (or, all buyers) strictly discounts the future deferring allocation exogenously decreases revenue, and immediate allocation is strictly optimal whenever the antecedents of Theorem 4 are satisfied.

6 Related literature and conclusion

Our environment bears a similarity to the literature on dynamic pricing. A common assumption in the dynamic pricing literature is that agents arrive at random times. When buyers with symmetric and known discount rates can choose when to purchase (but not when to arrive), Board and Skrzypacz [2016] show that a gradually declining reserve price is optimal.[30] Pai and Vohra [2013] and Mierendorff [2016] consider the possibility that agents have

[29] Proposition 1 of di Tillio et al. [2016] holds in this setting, and the revelation principle applies.

[30] Stokey [1979] finds a declining price curve when there is a continuum of buyers. Riley and Zeckhauser [1983] show that, against a stream of buyers, the seller's optimal mechanism is a fixed price in each period. A new buyer arrives in every period before the good is sold, which is possible in our model only if discount types are not independent across agents.

privately-known deadlines. A key distinguishing feature of our work is that the literature on dynamic pricing asks how to optimally sell a good over time, while we ask when it is not optimal to sell a good over time.

Our work ties most directly to previous work on bundling. Traditional bundling models consider when it is optimal to package multiple goods together, and when it is optimal to sell them individually. McAfee and McMillan [1988] consider the monopolist's problem when selling multiple goods to agents with multidimensional types. Rochet and Choné [1998] show that in optimal multidimensional mechanisms, there are typically collections of types receiving identical allocations; Manelli and Vincent [2006] provide conditions under which bundling (i.e., identical allocations for all types) is optimal, and Manelli and Vincent [2007] characterize the full set of optimal mechanisms when types are multidimensional. In our model, the set of goods corresponds to the ability to allocate a fixed unit at different points in time; the implied bundle therefore has an allocation constraint that is slightly nonstandard: a little more of tomorrow's good means a little less of today's good.[31]

Our work is closely related to Haghpanah and Hartline [2019], which gives conditions under which a monopolist sells only a "grand bundle" of all products. An agent's fundamental ($t = 0$) value v in our model corresponds to the value for the grand bundle in Haghpanah and Hartline [2019], and their Theorem 1 corresponds to our Corollary 1. Our Theorems 2 and 3 are stronger than our Corollary 1 (see Example 1), therefore our results are stronger in our context.[32] Our approach to Theorem 3, via Farkas' Lemma, is methodologically distinct. Additionally, our no arrival results consider optimality of the Myerson rule when there are n buyers, while Haghpanah and Hartline [2019] consider the optimality of a single posted price when there is one buyer. More importantly, our set of goods is interrelated and subject to an adding-up constraint, $0 \leq \sum_t q_t^i \leq 1$; Haghpanah and Hartline [2019] consider the more traditional question of bundling existing goods, and there is no inherent relationship between them, $0 \leq q_t^i \leq 1$.

Our proof strategy follows from the observation that temporal nondiscrimination is feasible, independent of the relationship between discount types and valuation types. This allows us to avoid the complication of evaluating which (IC) constraints bind. Border [1991] and Border [2007] give conditions necessary for the implementation of a particular outcome rule; we utilize the Border constraints to analyze when it is suboptimal to defer allocation to later periods. Previous work has examined which incentive constraints will bind in optimal mechanisms (Carroll [2012], Archer and Kleinberg [2014], Mishra et al. [2016]). Our

[31] This approach is distinct from, e.g., Basov [2001], since our seller has a number of "goods" equal to the number of periods, which is equal to the dimension of the agents' discount type space.

[32] The model of Haghpanah and Hartline [2019] is more general than ours (for example, it does not assume additive valuations), and our results neither imply nor are implied by theirs.

approach is distinct, in that we assume a particular set of constraints binds, show that this is consistent with (unique) optimality of temporal nondiscrimination, which is consistent with these constraints binding.

6.1 Conclusion

In dynamic environments, sellers may be imperfectly aware of buyers' temporal preferences. We model a mechanism design problem in which buyers have private information about values and temporal preferences, and the seller can potentially improve revenue by screening on temporal preferences. We show that when values and discount rates are positively related, the optimal mechanism is a Myerson auction which allocates immediately upon arrival. In this case, temporal preferences are ignored. We further show that when buyers are hyperbolic discounters, a seller who does not temporally discriminate in a two-period mechanism never temporally discriminates, regardless of the number of periods available; and, when the seller has ambiguous beliefs regarding buyer temporal preferences, a nondiscriminatory mechanism is optimal so long as the seller believes it might be optimal. Our results suggest that the incentive constraints associated with complicated design settings may imply that comparatively simple mechanisms are optimal. We believe this intuition merits further study.

References

Ravindra K. Ahuja, Thomas L. Magnanti, and James B. Orlin. *Network Flows: Theory, Algorithms, and Applications*. Prentice-Hall, Inc., 1993.

Aaron Archer and Robert Kleinberg. Truthful germs are contagious: a local-to-global characterization of truthfulness. *Games and Economic Behavior*, 86:340–366, 2014.

Suren Basov. Hamiltonian approach to multi-dimensional screening. *Journal of Mathematical Economics*, 36(1):77–94, 2001.

Uri Benzion, Amnon Rapoport, and Joseph Yagil. Discount rates inferred from decisions: An experimental study. *Management science*, 35(3):270–284, 1989.

Simon Board and Andrzej Skrzypacz. Revenue management with forward-looking buyers. *Journal of Political Economy*, 124(4):1046–1087, 2016.

Kim C. Border. Implentation of reduced form auctions: A geometric approach. *Econometrica*, 59:1175–1187, 1991.

Kim C. Border. Reduced form auctions revisited. *Economic Theory*, 31:167–181, 2007.

Gabriel Carroll. When are local incentive constraints sufficient? *Econometrica*, 80(2):661–686, 2012.

Gabriel Carroll. Robustness and separation in multidimensional screening. *Econometrica*, 85(2):453–488, 2017.

Yeon-Koo Che, Jinwoo Kim, and Konrad Mierendorff. Generalized reduced-form auctions: A network-flow approach. *Econometrica*, 81(6):2487–2520, 2013.

Alfredo di Tillio, Nenad Kos, and Matthias Messner. The design of ambiguous mechanisms. *The Review of Economic Studies*, 84(1):237–276, 2016.

David Gale. A theorem on flows in networks. *Pacific Journal of Mathematics*, 7:1073–1082, 1957.

I Gilboa and D Schmeidler. Maxmin expected utilty theory with a non-unique prior. *Journal of Mathematical Economics*, pages 18–141, 1989.

Leonard Green and Joel Myerson. A discounting framework for choice with delayed and probabilistic rewards. *Psychological bulletin*, 130(5):769, 2004.

Nima Haghpanah and Jason Hartline. When is pure bundling optimal? Technical report, Working paper, 2019.

Shabnam Hakimi. *Characterization of the Neural Mechanisms Supporting the Implementation of Cognitive Control in Human Decision Making.* PhD thesis, California Institute of Technology, 9 2013.

Kris N Kirby and Nino N Maraković. Modeling myopic decisions: Evidence for hyperbolic delay-discounting within subjects and amounts. *Organizational Behavior and Human decision processes*, 64(1):22–30, 1995.

Alejandro M Manelli and Daniel R Vincent. Bundling as an optimal selling mechanism for a multiple-good monopolist. *Journal of Economic Theory*, 127(1):1–35, 2006.

Alejandro M Manelli and Daniel R Vincent. Multidimensional mechanism design: Revenue maximization and the multiple-good monopoly. *Journal of Economic Theory*, 137(1): 153–185, 2007.

R Preston McAfee and John McMillan. Multidimensional incentive compatibility and mechanism design. *Journal of Economic theory*, 46(2):335–354, 1988.

Konrad Mierendorff. Asymmetric reduced form auctions. *Economics Letters*, 110:41–44, 2011.

Konrad Mierendorff. Optimal dynamic mechanism design with deadlines. *Journal of Economic Theory*, 161:190–222, 2016.

Walter Mischel, Yuichi Shoda, and Monica I Rodriguez. Delay of gratification in children. *Science*, 244(4907):933–938, 1989.

Debasis Mishra, Anup Pramanik, and Souvik Roy. Local incentive compatibility with transfers. *Games and Economic Behavior*, 100:149–165, 2016.

Michael Mussa and Sherwin Rosen. Monopoly and product quality. *Journal of Economic theory*, 18(2):301–317, 1978.

Roger B. Myerson. Optimal auction design. *Mathematics of Operations Research*, 6(1): 58–73, February 1981.

Mallesh M. Pai and Rakesh Vohra. Optimal dynamic auctions and simple index rules. *Mathematics of Operations Research*, 38(4):682–697, 2013.

John Riley and Richard Zeckhauser. Optimal selling strategies: When to haggle, when to hold firm. *The Quarterly Journal of Economics*, 98(2):267–289, 1983.

Jean-Charles Rochet and Philippe Choné. Ironing, sweeping, and multidimensional screening. *Econometrica*, pages 783–826, 1998.

Ran Spiegler. *Bounded rationality and industrial organization.* Oxford University Press, 2011.

Nancy L Stokey. Intertemporal price discrimination. *The Quarterly Journal of Economics*, pages 355–371, 1979.

Robert Henry Strotz. Myopia and inconsistency in dynamic utility maximization. *The review of economic studies*, 23(3):165–180, 1955.

A Proofs of main results

A.1 Technical background: network flows

A network consists of a set of nodes, $\mathcal{N}$, and a set of directed arcs, $\mathcal{A}$, which may carry "flow" between two nodes. A nonnegative flow across arcs is feasible if it satisfies nodespecific requirements and any arc-specific capacity constraints. The variation of feasible flow theorem that we use in Theorem 2 is due to Gale [1957].[33] Let $g(x, x')$ represent the flow between two nodes $x, x' \in \mathcal{N}$ (or the flow across the (x, x') arc). Each arc has capacity $k(x, x') \geq 0$, which limits the corresponding flow, and each node has a net demand of $b(x)$[34]. The feasible flow problem is to determine when there exists a flow in a network satisfying the capacity constraints and the net demand requirements. Stated formally, we want to determine when there exists a solution in $g(x, x')$ to the following problem.

$$\sum_{\{x' | (x', x) \in \mathcal{A}\}} g(x', x) - \sum_{\{x' | (x, x') \in \mathcal{A}\}} g(x, x') = b(x) \qquad \forall x \in \mathcal{N} \tag{8}$$

$$0 \leq g(x, x') \leq k(x, x') \qquad \forall (x, x') \in \mathcal{A}, \tag{9}$$

where $\sum_{x \in \mathcal{N}} b(x) = 0$. Gale [1957] provides the answer in the following result.

Theorem 5. *There exists a solution, g, to the system in* (8) *and* (9) *if and only if*

$$\sum_{x \in X, x' \in \bar{X}} k(x, x') \geq \sum_{x' \in \bar{X}} b(x') \qquad \forall X \subseteq \mathcal{N}, \tag{10}$$

where $\bar{X} = \mathcal{N} \setminus X$.

Intuitively, there is a feasible flow if and only if the capacity for sending flow from any set of nodes, X, to its complement, $\bar{X}$, exceeds the net demand of the receiving nodes.

A.2 Proof of Proposition 2 (No arrivals case)

Proof. Assume for the moment that time is finite, $T \in \mathbb{N}$. This assumption ensures that our dual variables are all of finite dimension. We relax this assumption at the end of the proof.

For some buyer i with type (v, δ) consider the IC constraints preventing the misreport of a lower discount type $\delta' < \delta$,

$$\varepsilon \sum_{v' < v} q^i (v', \delta) \cdot \delta \geq \varepsilon \sum_{v' < v} q (v', \delta') \cdot \delta' + q (v, \delta') \cdot (\delta - \delta') v,$$

The left- and right-hand sides of this inequality are interim utility to a buyer with type (v, δ) who reports type (v, δ'); these expressions arise from the IC constraints requiring truthful reporting of values. Attach to each such constraint the dual variable $\lambda^i(v, \delta, \delta')$.

[33] We report the version of this theorem stated as Theorem 6.12 of Ahuja et al. [1993]. We have adjusted the notation and the statement of the theorem.

[34] If $b(x) < 0$, x is a supply node, but we use the term net demand for both cases.

Intuitively, the feasibility (Border) constraints in our problem will bind for sets of buyer types that include all buyers with an equal or higher average virtual valuation. For buyers i and j (possibly equal to i) and value $v \in \mathcal{V}$, define $M^{ij}(v) \equiv \{v' \colon m^j(v') \geq m^i(v)\}$ to be the set of valuations such that buyer j's average virtual value is higher than the average virtual value of buyer i with value v. For a buyer i and value v, the family of feasibility constraints relevant for our problem are

$$\sum_j \sum_{v' \in M^{ij}(v)} \sum_\delta \sum_{t \in T} q_t^j\left(v', \delta\right) f^j\left(v', \delta\right) \leq 1 - \prod_j \Pr\left(\tilde{v}^j \notin M^{ij}\left(v\right)\right). \tag{11}$$

The left-hand side of (11) is the ex ante probability of assigning the good to a buyer type with an average virtual valuation of at least $m^i(v)$ (given i and v), while the right-hand side is the probability that at least one buyer's type realization has such a virtual valuation. See Theorems 2 and 3 in Mierendorff [2011] as well as Border [2007] for a justification of this constraint. Let $\rho^i(v)$ be the dual variable associated with the constraint in (11) written with respect to bidder i and valuation v, and define $R^i(v) \equiv \sum_{v'<v} \rho^i(v')$. Letting $\gamma_t^i(v,\delta)$ be the multiplier on the positivity constraint $q_t^i(v,\delta) \geq 0$, the coefficient on $q_0^i(v,\delta)$, denoted $c_0^i(v,\delta)$, in the resulting linear programming problem is given by

$$c_0^i\left(v,\delta\right) = \left(m^i\left(v|\delta\right) - R^i\left(v\right)\right) f^i\left(v,\delta\right) + \gamma_0^i\left(v,\delta\right) + \sum_{\substack{\delta > \delta' \\ v' > v}} \lambda^i\left(v', \delta, \delta'\right)\varepsilon - \sum_{\substack{\delta' > \delta \\ v' > v}} \lambda^i\left(v', \delta', \delta\right)\varepsilon.$$

To prove the optimality of temporal nondiscrimination, it is sufficient to find (feasible) values for the dual variables such that for all types, (v,δ), such that the following condition is satisfied:

$$c_t^i\left(v,\delta\right) = 0 \;\forall i,t,v,\delta, \text{ and } q_0^i\left(v,\delta\right) > 0 \text{ implies } \gamma_0^i\left(v,\delta\right) = 0. \tag{LP}$$

We first show that feasible multipliers at time $t = 0$ imply feasible multipliers for all subsequent time periods $t > 0$. The coefficients c_t can be written as

$$c_t^i\left(v,\delta\right) = \delta^t c_0^i(v,\delta) - \left(1-\delta_t\right) R^i\left(v\right) f^i\left(v,\delta\right) - \delta_t \gamma_0^i\left(v,\delta\right) - \sum_{\delta' > \delta} \lambda^i\left(v,\delta',\delta\right)\left(\delta'_t - \delta_t\right) v + \gamma_t^i\left(v,\delta\right).$$

The dual variables R^i, λ^i, and γ_0^i are all weakly positive. Therefore, $c_0^i(v,\delta) = 0$ implies there is always a nonnegative value for $\gamma_t^i(v,\delta)$ making $c_t^i(v,\delta) = 0$. Consequently, we only need to consider the $t = 0$ terms in (LP).

Given valid choices for γ_0^i, condition (LP) is implied by $c_0^i(1,\delta) = 0$, $c_0^i(v-\varepsilon,\delta) - c_0^i(v,\delta) = 0$. Fixing a value for the summation of the feasibility dual variables, $R^i(v) = m_+^i(v)$, we analyze the remaining problem in the λ^i and γ_0^i variables. Define $\mu^i(v,\delta) \equiv (m^i(v|\delta) -$

$R^i(v))f^i(v,\delta) = \left(m^i(v|\delta) - m^i_+(v)\right) f^i(v,\delta)$, and for convenience let $\mu^i(-\varepsilon,\delta) = 0$.[35] Note first that $\{v' \in \mathcal{V} \colon v' > 1\} = \emptyset$; then since $m^i(1) = m^i(1|\delta) = 0$ for all $\delta \in \mathcal{D}^i$, $c^i_0(1,\delta) = 0$ for all agents i and discount types δ. It follows that condition (LP) is satisfied if the following equation has a nonnegative solution in the dual variables λ^i:

$$\sum_{\delta'>\delta} \lambda^i\left(v,\delta',\delta\right)\varepsilon - \sum_{\delta>\delta'} \lambda^i\left(v,\delta,\delta'\right)\varepsilon - \gamma^i_0\left(v-\varepsilon,\delta\right) + \gamma^i_0\left(v,\delta\right) = \mu^i\left(v-\varepsilon,\delta\right) - \mu^i\left(v,\delta\right). \quad (12)$$

Here, we require that $\gamma^i_0(v,\delta) = 0$ when $m^i(v) > 0$, consistent with $q^i_0(v,\delta) > 0$ for $m^i(v) > 0$.[36]

To apply Theorem 5, we represent this system as a network in which each type (v,δ) is associated with a node in $\mathcal{N}$, each $\lambda^i(v,\delta,\delta')$ is associated with a (nonnegative) flow from (v,δ) to (v,δ'), and each $\gamma^i(v-\varepsilon,\delta)$ is associated with a (nonnegative) flow from (v,δ) to $(v-\varepsilon,\delta)$. We define a separate network for each buyer i. First, write

$$g^i\left(\hat{v},\hat{\delta},\hat{v}',\hat{\delta}'\right) = \begin{cases} \lambda^i\left(\hat{v},\hat{\delta},\hat{\delta}'\right)\varepsilon & \text{if } \hat{v}' = \hat{v} \text{ and } \hat{\delta} \neq \hat{\delta}', \\ \gamma^i\left(\hat{v}',\hat{\delta}\right) & \text{if } \hat{v}' \in \{\hat{v},\hat{v}-\varepsilon\} \text{ and } \hat{\delta} = \hat{\delta}', \\ 0 & \text{otherwise.} \end{cases}$$

Now, define capacities $k(\hat{v},\hat{\delta},\hat{v}',\hat{\delta}')$ for the network arcs,

$$k\left(\hat{v},\hat{\delta},\hat{v}',\hat{\delta}'\right) = \begin{cases} +\infty & \text{if } \hat{v}' = \hat{v} \text{ and } \hat{\delta} > \hat{\delta}', \\ +\infty & \text{if } \hat{v}' \in \{\hat{v},\hat{v}-\varepsilon\}, \hat{\delta} = \hat{\delta}', \text{ and } m^i\left(\hat{v}'\right) \leq 0 \\ 0 & \text{otherwise.} \end{cases}$$

Finally, define b so that $b(v,\delta) = \mu^i(v-\varepsilon,\delta) - \mu^i(v,\delta)$. We apply Theorem 5 to this network.

Let $X \subseteq \Theta^i$ be a set of types (v,δ) and let $\bar{X} = \Theta^i \setminus X$ be its complement. Given the functions b and k defined above, observe that inequality (10) is slack, because the left-hand side is infinite, in two cases:

- There are two types $(v,\delta) \geq (v,\delta')$ such that $(v,\delta) \in X$ and $(v,\delta') \in \bar{X}$;
- There are types $(v,\delta) \in X$ and $(v',\delta) \in \bar{X}$ such that $v' < v < v^\star_i$ (i.e., $m^i(v) \leq 0$).

Therefore, to consider satisfaction of (10) we need only consider sets X such that if $(v,\delta) \in \bar{X}$, then $(v',\delta') \in \bar{X}$ for all $v' \in \{v,\ldots,v^\star_i(v)\}$ and $\delta' \geq \delta$. We refer to such $\bar{X}$ as *limited upper*

[35] This is analogous to adding a dummy type reflecting the individual rationality constraint.
[36] Recall that in the no arrivals case analyzed here, $m^i(v) > 0$ if and only if $m^i(v) \geq m^i(v^\star_i)$.

sets. Theorem 5 implies that there is a solution for the multipliers λ^i and γ^i if and only if, for any limited upper set $\bar{X}$,

$$0 \geq \sum_{(v,\delta)\in\bar{X}} b(v,\delta) = \sum_{(v,\delta)\in\bar{X}} \mu^i(v-\varepsilon,\delta) - \mu^i(v,\delta). \tag{13}$$

Since inequality (13) must hold for all limited upper sets $\bar{X}$, it must hold for limited upper sets such that there is $(v,\delta) \in \bar{X}$ with $(v',\delta') \in \bar{X}$ if and only if $v' \in \{v,\ldots,v_i^\star(v)\}$ and $\delta' \geq \delta$. Thus a necessary condition for inequality (13) is

$$\sum_{\delta'\geq\delta} \left\{\mu^i\left(v^\star(v),\delta'\right) - \mu^i(v-\varepsilon,\delta)\right\} \geq 0, \tag{14}$$

for all $v \in \mathcal{V}$ and $\delta \in \mathcal{D}^i$.

Now, consider an arbitrary upper set $\bar{X} \subseteq \Theta^i$, and let $\bar{X} = \bar{X}_- \cup \bar{X}_+$, where $(v,\delta) \in \bar{X}_-$ if $m^i(v) \leq 0$ and $(v,\delta) \in \bar{X}_+$ if $m^i(v) > 0$. Let $D(\tilde{X}) = \{d \in \mathcal{D}^i \colon \exists v \in \mathcal{V} \text{ s.t. } (v,\delta) \in \tilde{X}\}$ be the set of discount types appearing in $\tilde{X}$, and for $\delta \in D(\tilde{X})$ let $V(\delta;\tilde{X}) = \{v \in \mathcal{V} \colon (\delta,v) \in \tilde{X}\}$ be the set of valuation types associated with discount type δ in $\tilde{X}$. Define $V(\tilde{X})$ and $D(v;\tilde{X})$ similarly. Observing that for all $\delta \in D(\bar{X}_-)$, $v_i^\star(\min V(\delta,\bar{X}_-)) = v_i^\star$, we write

$$\begin{aligned}
&\sum_{(v,\delta)\in\bar{X}} \left\{\mu^i(v,\delta) - \mu^i(v-\varepsilon,\delta)\right\} \\
&\quad= \sum_{(v,\delta)\in\bar{X}_+} \left\{\mu^i(v,\delta) - \mu^i(v-\varepsilon,\delta)\right\} + \sum_{(v,\delta)\in\bar{X}_-} \left\{\mu^i(v,\delta) - \mu^i(v-\varepsilon,\delta)\right\} \\
&\quad= \sum_{v\in V\left(\bar{X}_+\right)} \sum_{\delta\in D\left(v;\bar{X}_+\right)} \left\{\mu^i(v,\delta) - \mu^i(v-\varepsilon,\delta)\right\} + \sum_{\delta\in D\left(\bar{X}_-\right)} \sum_{v\in V\left(\delta;\bar{X}_-\right)} \left\{\mu^i(v,\delta) - \mu^i(v-\varepsilon,\delta)\right\} \\
&\quad= \sum_{v\in V\left(\bar{X}_+\right)} \sum_{\delta\in D\left(v;\bar{X}_+\right)} \left\{\mu^i(v,\delta) - \mu^i(v-\varepsilon,\delta)\right\} + \sum_{\delta\in D\left(\bar{X}_-\right)} \left\{\mu^i(v_i^\star,\delta) - \mu^i\left(\min V\left(\delta;\bar{X}_-\right)-\varepsilon,\delta\right)\right\}.
\end{aligned}$$

Since $\bar{X}_+$ and $\bar{X}_-$ are upper sets, $D(\bar{X}_-) = \{\delta' \in \mathcal{D}^i \colon \delta' \geq \delta\}$ for some $\delta \in \mathcal{D}^i$, and the same is true of $D(v;\bar{X}_+)$. Then satisfaction of inequality (14) implies that the above expression is weakly positive, and hence inequality (14) is necessary and sufficient for condition (13).

Finally, condition (14) is independent of the number of available periods, and if it is satisfied for any finite T, then temporal nondiscrimination is optimal for any length $T \in \mathbb{N}$. Since buyers are exponential discounters, the revenue obtainable by allocating to some type at time $t = \infty$ is approximated by the revenue obtainable by allocating to this same type at time $t = T$, for T large.[37] It follows that when the latter is never optimal, neither is the

[37]When all discount types strictly discount the future, $\delta < 1$ for all $\delta \in \mathcal{D}^i$, this result is immediate:

former, and thus Theorem 2 extends to the case of infinite T. $\square$

A.3 Proof of Theorem 1

Proof of Theorem 1. We follow the proof of Proposition 2, and initially (for the proof only) assume that T is finite. To begin, we respecify the (IC) and (IR) constraints in the seller's problem. For buyer i with type (v, δ, τ) consider all downward IC constraints preventing the misreport of $\delta' < \delta$; for $\delta' < \delta$, these constraints take the form

$$\varepsilon \sum_{v'<v} \sum_{t \in T} \delta^t q_t^i (v', \delta, \tau) \geq \varepsilon \sum_{v'<v} \sum_{t \in T} \delta'^t q_t^i (v', \delta', \tau) + \sum_{t \in T} q_t^i (v', \delta', \tau) \left(\delta^t - \delta'^t\right) v,$$

Attach to each such constraint the dual variable $\lambda^i(v, \delta, \delta', \tau)$.[38]

We must also adapt the feasibility constraints to dynamic arrivals. As in our proof of Theorem 2, we adapt the (MRM) feasibility constraint of Border [2007] to state that the probability of allocating to buyers higher virtual values who arrive before time τ cannot exceed the probability that such a buyer exists. We define an analogue of M^{ij}, accounting for the possibility of past buyers,

$$M_\tau^{ij} (v) \equiv \left\{(v', \tau') : \tau' < \tau \text{ and } v' \geq v_j^\star, \text{ or } \tau' = \tau \text{ and } m^j (v') \geq m^i (v)\right\}.$$

For a buyer j, the value type v' and arrival time $\tau' \leq \tau$ are (jointly) in $M_\tau^{ij}(v)$ if either the buyer arrived strictly before time τ and had virtual value above the cutoff virtual value $v_j^\star$, or if the buyer arrived at time τ and had virtual value above $m^i(v)$. That is, this buyer is in $M_\tau^{ij}(v)$ if they are more profitable for the seller. The feasibility constraint that limits allocation to buyers with higher virtual values is, for all $v \in \mathcal{V}$ and all $\tau \in T$,

$$\sum_j \sum_{(v',\tau') \in M_\tau^{ij}(v)} \sum_{\delta \in \mathcal{D}^j} \sum_{t \in T} q_t^j (v', \delta, \tau') g_j f^j (v', \delta) \leq \Pr\left(\exists (j, v', \tau') \text{ s.t. } (v', \tau') \in M_\tau^{ij} (v)\right). \tag{15}$$

The left-hand side of inequality (15) is the probability an agent with a higher virtual value v' arrives at time $\tau' \leq \tau$ and receives the item; the right-hand side of the inequality is the probability that such an agent exists, given a value v and an arrival time τ. For a given value

deferring allocation to $t = \infty$ reduces willingness to pay down to zero. Otherwise, when some discount type is perfectly patient, $\delta = 1$, their willingness to pay is identical across all time periods; if the seller can improve revenue by infinitely deferring consumption of one type to a point where all other exponential discount types have zero value of consumption, a strict revenue improvement is available when T is large enough that all other exponential discount types have approximately zero value of consumption.

[38] As discussed in the main text, arrivals are common knowledge and thus there is no need to consider misreports of arrival time τ.

v and arrival time τ, we attach the multiplier $\rho(v,\tau)$ to the feasibility constraint (15); the sum of these multipliers is $R(v,\tau) = \sum_{\tau' \geq \tau} \sum_{v'<v} \rho(v',\tau')$, which represents the aggregation of these constraints for all later arrivals with lower values. We let $\gamma_t^i(v,\delta,\tau)$ be the multiplier on the constraint $q_t^i(v,\delta,\tau) \geq 0$.

We now consider the effect of temporal discrimination on revenue, considering the given constraints. The coefficient on $q_\tau^i(v,\delta,\tau)$, denoted $c_\tau^i(v,\delta,\tau)$, in the linear programming problem representing the seller's revenue maximization is given by

$$\begin{aligned} c_\tau^i(v,\delta,\tau) = {} & \left(\delta^\tau m^i(v|\delta) - R(v,\tau)\right) g(\tau) f(v,\delta) + \gamma_\tau^i(v,\delta,\tau) \\ & + \sum_{\substack{(w,\delta)\geq(w',\delta') \\ w\geq v}} \delta^\tau \lambda^i(w,\delta,\delta',\tau)\,\varepsilon - \sum_{\substack{(w',\delta')\geq(w,\delta) \\ w\geq v}} \delta^\tau \lambda^i(w',\delta',\delta,\tau)\,\varepsilon. \end{aligned}$$

This expression is an adaption of $c_0^i(v,\delta)$ from the proof of Theorem 2, adjusting for the observable arrival time τ.

Recall that under the dynamic Myerson rule, a buyer who arrives in period τ receives the good in period τ if she has the highest virtual value above the period's reserve price $v_\tau^\star$, implying that $q_\tau^i(v,\delta,\tau) > 0$ for $v > v_\tau^\star$. We therefore require that $\gamma_\tau^i(v,\delta,\tau) = 0$ for $v > v_\tau^\star$. Following the logic in the proof of Theorem 2, we set $R(v,\tau) = \delta^\tau m_+(v)$. Note that since $R(v,\tau) = \sum_{\tau'\geq\tau}\sum_{v'<v}\rho(v',\tau')$, this is equivalent to setting $\sum_{v'<v}\rho(v',\tau) = (\delta^\tau - \delta^{\tau+1})m_+(v)$.

As in the proof of Theorem 2, requiring $c_\tau^i(v,\delta,\tau) = 0$ is equivalent to requiring $c_\tau^i(1,\delta,\tau) = 0$ and $c_\tau^i(v-\varepsilon,\delta,\tau) - c_\tau^i(v,\delta,\tau) = 0$ for all $v > 0$. This implies a system in the λ variables analogous to (12), where we now define $\mu^i(v,\delta,\tau) \equiv \left(m^i(v|\delta) - m_+^i(v,\tau)\right)g(\tau)f(v,\delta)$ with $\mu^i(0,\delta,\tau) = 0$,

$$\sum_{\delta'\geq\delta}\lambda^i(v,\delta',\delta,\tau) - \sum_{\delta\geq\delta'}\lambda^i(v,\delta,\delta',\tau) - \gamma_\tau^i(v-\varepsilon,\delta,\tau) + \gamma_\tau^i(v,\delta,\tau) = \mu^i(v-\varepsilon,\delta,\tau) - \mu^i(v,\delta,\tau). \tag{16}$$

Since arrival times are publicly observable, they have no effect on the central argument of Theorem 2, which analyzes the problem buyer-by-buyer and arrival time-by-arrival time. Thus even though there are an infinite number of dual variables λ^i, the network associated with arrival time τ is finite, as in the proof of Theorem 2. Thus from this point on, the proof is identical to the proof of Theorem 2 (including the extension to $T = \infty$), and is omitted here. □

Proof of Corollary 1. By definition, $m^i(v^\star(v)) > 0$. Then since $m_+^i(v) = \max\{0, m^i(v)\} \geq$

$m^i(v)$, for any $v \in \mathcal{V}$ we have Let $v \in \mathcal{V}$ be such that $m^i(v) \geq 0$. We calculate

$$\begin{aligned}
&\mu\left(v^{\star}\left(v\right),\delta\right)-\mu\left(v-\varepsilon,\delta\right)\\
&\quad\geq\left(m^{i}\left(v^{\star}\left(v\right)|\delta\right)-m^{i}\left(v^{\star}\left(v\right)\right)\right)f^{i}\left(v^{\star}\left(v\right),\delta\right)-\left(m^{i}\left(v-\varepsilon|\delta\right)-m^{i}\left(v-\varepsilon\right)\right)f^{i}\left(v-\varepsilon,\delta\right)\\
&\quad=\left(\frac{1-F^{i}\left(v^{\star}\left(v\right)\right)}{f^{i}\left(v^{\star}\left(v\right)\right)}-\frac{1-F^{i}\left(v^{\star}\left(v\right)|\delta\right)}{f^{i}\left(v^{\star}\left(v\right)|\delta\right)}\right)f^{i}\left(v^{\star}\left(v\right),\delta\right)\\
&\qquad\qquad-\left(\frac{1-F^{i}\left(v-\varepsilon\right)}{f^{i}\left(v-\varepsilon\right)}-\frac{1-F^{i}\left(v-\varepsilon|\delta\right)}{f^{i}\left(v-\varepsilon|\delta\right)}\right)f^{i}\left(v-\varepsilon,\delta\right)\\
&\quad=\left(1-F^{i}\left(v^{\star}\left(v\right)\right)\right)\left(f^{i}\left(\delta|v^{\star}\left(v\right)\right)-f^{i}\left(\delta|v-\varepsilon\right)\right)\\
&\qquad+\sum_{v^{\star}(v)>v'\geq v-\varepsilon}\left(f^{i}\left(\delta|v'\right)-f^{i}\left(\delta|v-\varepsilon\right)\right)f^{i}\left(v'\right). \qquad (17)
\end{aligned}$$

Summing (17) over all $\delta' > \delta$ gives

$$\begin{aligned}
&\left(1-F^{i}\left(v^{\star}\left(v\right)\right)\right)\left(F^{i}\left(\delta|v-\varepsilon\right)-F^{i}\left(\delta|v^{\star}\left(v\right)\right)\right)\\
&\qquad+\sum_{v^{\star}(v)>v'\geq v-\varepsilon}\left(F^{i}\left(\delta|v-\varepsilon\right)-F^{i}\left(\delta|v'\right)\right)f^{i}\left(v'\right)\geq 0,
\end{aligned}$$

where the inequality follows from the corollary's assumption that $F^i(\delta|\cdot)$ is nonincreasing. □

A.4 Proof of Theorem 3

Lemma 1. *Let $\delta > \tilde{\delta}$, and define $w : \mathbb{R}_{++} \to \mathbb{R}$ by*

$$w\left(t\right)=\frac{1-\delta^{t}}{\delta^{t}-\tilde{\delta}^{t}}.$$

Then w is increasing in t.

Proof. The first derivative of w takes the same sign as

$$\begin{aligned}
w'\left(t\right)&\overset{\text{sign}}{=}-\left(\delta^{t}-\tilde{\delta}^{t}\right)\delta^{t}\ln\delta-\left(1-\delta^{t}\right)\left(\delta^{t}\ln\delta-\tilde{\delta}^{t}\ln\tilde{\delta}\right)\\
&=-\left(1-\tilde{\delta}^{t}\right)\delta^{t}\ln\delta+\left(1-\delta^{t}\right)\tilde{\delta}^{t}\ln\tilde{\delta}\overset{\text{sign}}{=}-\underbrace{\frac{\delta^{t}\ln\delta}{1-\delta^{t}}}_{\hat{w}(\delta)}+\frac{\tilde{\delta}^{t}\ln\tilde{\delta}}{1-\tilde{\delta}^{t}}.
\end{aligned}$$

The derivative of $\hat{w}$ with respect to δ is

$$\hat{w}'\left(\delta\right)\overset{\text{sign}}{=}\left(t\delta^{t-1}\ln\delta+\delta^{t-1}\right)\left(1-\delta^{t}\right)+t\delta^{t-1}\delta^{t}\ln\delta\overset{\text{sign}}{=}t\ln\delta+\left(1-\delta^{t}\right)=\ln\delta^{t}-\left(\delta^{t}-1\right).$$

Since ln is concave, a standard Taylor approximation of $\hat{w}'(\delta)$ implies that $\hat{w}'(\delta) \leq 0$. The assumption that $\delta > \tilde{\delta}$ then implies that $w'(t) \geq 0$, as desired. □

Proof of Theorem 3. We build on the preparatory work done in the proof of Theorem 2. Allowing for all possible deviations to alternate discount rates, the coefficient c_t^i on the allocation q_t^i in the linear programming problem is

$$c_t^i(v,\delta) = \left(\delta_t m^i(v|\delta) - R^i(v)\right) f^i(v,\delta) + \sum_{\substack{\delta'\neq\delta \\ v'>v}} \lambda^i(v',\delta,\delta')\,\delta_t\varepsilon - \sum_{\substack{\delta'\neq\delta \\ v'>v}} \lambda^i(v',\delta',\delta)\,\delta'_t\varepsilon.$$

As in the proof of Theorem 2, we look for multipliers so that $c_t^i(v,\delta) = 0$ implies that $t = 0$ and $m^i(v) > 0$. It is sufficient to solve the following system:

$$\sum_{\delta'\neq\delta} \left[\lambda^i(v,\delta,\delta') - \lambda^i(v,\delta',\delta)\right]\varepsilon + \gamma_0^i(v,\delta) = \mu^i(v,\delta) - \mu^i(v-\varepsilon,\delta) \tag{18}$$

$$\sum_{\delta'\neq\delta} \lambda^i(v,\delta',\delta)\left(\delta_t - \delta'_t\right)v + \gamma_t^i(v,\delta) = (1-\delta_t)\,m_+^i(v)\,f(v,\delta) \quad \forall t > 0. \tag{19}$$

Applying Farkas' Lemma, there exists a nonnegative solution (λ, γ) to the system (18)–(19) if and only if there are no $y^i(v,\delta)$ and $z_t^i(v,\delta)$ satisfying

$$\left[y^i(v,\delta) - y^i(v,\delta')\right]\varepsilon \geq \sum_{t>0} z_t^i(v,\delta')\left(\delta_t - \delta'_t\right)v \qquad \forall \delta \neq \delta' \tag{20}$$

$$y^i(v,\delta) \geq 0 \qquad \forall (v,\delta) \in \Theta^i \text{ s.t. } m(v) \leq 0 \tag{21}$$

$$z_t^i(v,\delta) \geq 0 \qquad \forall (v,\delta) \in \Theta^i, \forall t > 0, \tag{22}$$

and

$$\sum_{(v,\delta)\in\Theta^i} \left[y^i(v,\delta)\left(\mu(v,\delta) - \mu(v-\varepsilon,\delta)\right) + m_+^i(v)\,f^i(v,\delta)\sum_{t>0} z_t^i(v,\delta)(1-\delta_t)\right] < 0. \tag{23}$$

Under exponential discounting we may write any discount type $\delta \in \mathcal{D}^i$ as $\delta = (1, \hat{\delta}^1, \ldots, \hat{\delta}^T)$; for simplicity we work directly with the (single-dimensional) discount rates $\hat{\delta}$, and order them so that $\hat{\delta}_1 > \hat{\delta}_2 > \cdots > \hat{\delta}_d$, where $d = \#\mathcal{D}^i$. To simplify notation, we make the following substitutions:

$$\Delta\mu_k \equiv \mu^i\left(v,\hat{\delta}_k\right) - \mu^i\left(v-\varepsilon,\hat{\delta}_k\right), \quad mf_k \equiv m_+^i(v)\,f^i\left(v,\hat{\delta}_k\right), \quad y_k \equiv y^i\left(v,\hat{\delta}_k\right), \text{ and } z_{tk} \equiv z_t^i\left(v,\hat{\delta}_k\right).$$

Furthermore, we let $y_k = \sum_{k'\leq k} x_{k'}$, so that $y_k - y_{\tilde{k}} = \sum_{\tilde{k}<k'\leq k} x_{k'}$ for all $\tilde{k} < k$.

Consider choosing values z_{tk} to solve the following problem:

$$\min \sum_{t>0} \left(1-\hat{\delta}_k^t\right) z_{tk} \text{ s.t. } \sum_{t>0} \left(\hat{\delta}_k^t - \hat{\delta}_{k+1}^t\right) v z_{tk} \geq x_k \varepsilon.$$

Lemma 1 shows that $(1-\hat{\delta}_k^t)/(\hat{\delta}_k^t - \hat{\delta}_{k+1}^t)$ is increasing in t, thus the solution is

$$z_{tk} = \begin{cases} \frac{x_k \varepsilon}{\left(\hat{\delta}_k - \hat{\delta}_{k+1}\right) v} & \text{if } t = 1, \\ 0 & \text{if } t > 1. \end{cases}$$

Substituting these values into inequality (23) gives

$$\begin{aligned} &\sum_{(v,\delta)\in\Theta^i} \left[\Delta\mu_k y_k + m f_k \sum_{t>0} \left(1-\hat{\delta}_k^t\right) z_{tk}\right] \\ &= \sum_{(v,\delta)\in\Theta^i} \left[\Delta\mu_k \sum_{k'\leq k} x_{k'} + \left(\frac{x_k \varepsilon}{\left(\hat{\delta}_k - \hat{\delta}_{k+1}\right) v}\right) \left(1-\hat{\delta}_k\right) m f_k\right] \\ &= \sum_{(v,\delta)\in\Theta^i} \left[\sum_{k'\geq k} \Delta\mu_{k'} + \left(\frac{1-\hat{\delta}_k}{\hat{\delta}_k - \hat{\delta}_{k+1}}\right) \frac{\varepsilon}{v} m f_k\right] x_k. \end{aligned} \tag{24}$$

Now fix $\bar{v}$ and $\bar{k} \in \{1,\ldots,d\}$, and let x_k be

$$x_k = x^i\left(v, \hat{\delta}_k\right) = \begin{cases} \hat{\delta}_k - \hat{\delta}_{k+1} & \text{if } k \leq \bar{k} \text{ and } v = \bar{v}, \\ 0 & \text{otherwise.} \end{cases}$$

Substituting into (24) gives

$$\begin{aligned} &\sum_{k\leq\bar{k}} \left[\left(\hat{\delta}_k - \hat{\delta}_{k+1}\right) \sum_{k'\geq k} \Delta\mu_{k'} + \left(1-\hat{\delta}_k\right) m f_k \frac{\varepsilon}{v}\right] \\ &\overset{\text{sign}}{=} \sum_{k\leq\bar{k}} \left[\frac{\Delta\mu_k}{\varepsilon} \left(\hat{\delta}_k - \hat{\delta}_{\bar{k}+1}\right) v + \left(1-\hat{\delta}_k\right) m f_k\right] \equiv M_{\bar{k}}(v). \end{aligned}$$

Then when $M_{\bar{k}}(v) \geq 0$ for all v and $\bar{k}$, there are values for y_k and z_{tk} so that inequalities (20)–(22) are satisfied but inequality (23) is not. It follows that the when $M_{\bar{k}}(v) \geq 0$ for all v and $\bar{k}$ there is no solution to the given system, and hence (by Farkas' Lemma) there is a solution to the system (18)–(19). □

B Calculations for Example 1

Let (exponential) discount types $\delta \in \{0, \hat{\delta}\}$ and values be $v \in \{1/2, 1\}$. For $\pi \in [0,1]$, let the distribution over types be

	0	$\hat{\delta}$
1	$\frac{1}{3}$	$\frac{1}{2}\pi$
$\frac{1}{2}$	$\frac{1}{6}$	$\frac{1}{2}(1-\pi)$

Note that the marginal probability of either discount rate is $\Pr(\delta = 0) = \Pr(\delta = \hat{\delta}) = 1/2$.

Straightforward calculation gives

$$m(v) = \begin{cases} 1 & \text{if } v = 1, \\ \frac{1-3\pi}{4-3\pi} & \text{if } v = \frac{1}{2}; \end{cases} \quad m(v|0) = \begin{cases} 1 & \text{if } v = 1, \\ -\frac{1}{2} & \text{if } v = \frac{1}{2}; \end{cases} \quad m\left(v \middle| \hat{\delta}\right) = \begin{cases} 1 & \text{if } v = 1, \\ \frac{1}{2}\left(\frac{1-2\pi}{1-\pi}\right) & \text{if } v = \frac{1}{2}. \end{cases}$$

In turn,

$$m(1|\delta) - m_+(1) = \begin{cases} 0 & \text{if } \delta = 0, \\ 0 & \text{if } \delta = \hat{\delta}; \end{cases} \quad \begin{array}{l} m\left(\frac{1}{2}\middle|\delta\right) - m_+\left(\frac{1}{2}\right) = \begin{cases} -\frac{1}{2} & \text{if } \delta = 0, \\ \frac{1}{2}\left(\frac{1-2\pi}{1-\pi}\right) & \text{if } \delta = \hat{\delta}, \end{cases} \quad \text{if } \pi \geq \frac{1}{3}; \\ m\left(\frac{1}{2}\middle|\delta\right) - m_+\left(\frac{1}{2}\right) = \begin{cases} -\frac{6-9\pi}{8-6\pi} & \text{if } \delta = 0, \\ \frac{2-3\pi}{2(1-\pi)(4-3\pi)} & \text{if } \delta = \hat{\delta}, \end{cases} \quad \text{if } \pi < \frac{1}{3}. \end{array}$$

Finally,

$$\mu(1|\delta) = \begin{cases} 0 & \text{if } \delta = 0, \\ 0 & \text{if } \delta = \hat{\delta}; \end{cases} \quad \begin{array}{l} \mu\left(\frac{1}{2}\middle|\delta\right) = \begin{cases} -\frac{1}{12} & \text{if } \delta = 0, \\ \frac{1}{4}(1-2\pi) & \text{if } \delta = \hat{\delta}, \end{cases} \quad \text{if } \pi \geq \frac{1}{3}; \\ \mu\left(\frac{1}{2}\middle|\delta\right) = \begin{cases} -\frac{2-3\pi}{16-12\pi} & \text{if } \delta = 0, \\ \frac{1}{4}\left(\frac{2-3\pi}{4-3\pi}\right) & \text{if } \delta = \hat{\delta}, \end{cases} \quad \text{if } \pi < \frac{1}{3}. \end{array}$$

To apply Theorem 1, we check

$$\mu(1|\delta) - \mu\left(\frac{1}{2}\middle|\delta\right) = \begin{cases} \frac{1}{12} & \text{if } \delta = 0, \\ -\frac{1}{4}(1-2\pi) & \text{if } \delta = \hat{\delta}, \end{cases} \quad \text{if } \pi \geq \frac{1}{3};$$

$$\mu(1|\delta) - \mu\left(\frac{1}{2}\middle|\delta\right) = \begin{cases} -\frac{1}{4}\left(\frac{2-3\pi}{4-3\pi}\right) & \text{if } \delta = 0, \\ -\frac{1}{4}\left(\frac{2-3\pi}{4-3\pi}\right) & \text{if } \delta = \hat{\delta}, \end{cases} \quad \text{if } \pi < \frac{1}{3}.$$

Note that when $\pi < 1/3$, $\mu(1|\hat{\delta}) - \mu(1/2|\hat{\delta}) < 0$ and Theorem 2 does not apply. On the other hand, when $\pi \geq 1/3$, $\mu(1|0) - \mu(1/2|0) = 1/12 > 0$, and

$$\mu\left(1\middle|\hat{\delta}\right) - \mu\left(\frac{1}{2}\middle|\hat{\delta}\right) = -\frac{1}{4}(1-2\pi) \geq 0 \quad \Longleftrightarrow \quad \pi \geq \frac{1}{2}.$$

Then Theorem 2 applies when $\pi \geq 1/2$.

Remark 4. *Note that the conditional cdf of discount type given value, $F(\delta|v)$, is*

	0	$\hat{\delta}$
1	$\frac{2}{2+3\pi}$	1
$\frac{1}{2}$	$\frac{1}{4-3\pi}$	1

Then $F(0|1) > F(0|1/2)$ whenever $\pi < 2/3$. Hence for $\pi \in [1/2, 2/3)$ our Theorem 1 implies that temporal nondiscrimination is optimal, but neither our own Corollary 1 nor Haghpanah and Hartline [2019]'s Theorem 1 apply.

We now show that the condition in Theorem 3 relaxes the above result. Because $\delta \in \{0, \hat{\delta}\}$ and $m_+(v) \geq 0$, the only relevant inequality is when $\delta_j = \hat{\delta}$. When $\pi \geq 1/3$ we check,

$$\hat{\delta}\left(\mu\left(1\middle|\hat{\delta}\right) - \mu\left(\frac{1}{2}\middle|\hat{\delta}\right)\right)\frac{1}{\frac{1}{2}} + \left(1-\hat{\delta}\right)(1)\left(\frac{1}{2}\pi\right) \geq 0$$

$$\Longleftrightarrow \quad -\frac{1}{2}\hat{\delta}(1-2\pi) + \frac{1}{2}\pi\left(1-\hat{\delta}\right) \geq 0 \quad \Longleftrightarrow \quad \pi \geq \frac{\hat{\delta}}{1+\hat{\delta}}.$$

Then when $\hat{\delta} < \frac{1}{2}$, Theorem 3 applies for all $\pi \geq 1/3$, and immediate sale is optimal.

When $\pi < 1/3$ we check

$$\begin{aligned} & \hat{\delta}\left(\mu\left(1\middle|\hat{\delta}\right) - \mu\left(\frac{1}{2}\middle|\hat{\delta}\right)\right)\frac{1}{\frac{1}{2}} + \left(1-\hat{\delta}\right)(1)\left(\frac{1}{2}\pi\right) \geq 0 \\ \Longleftrightarrow \quad & -(2-3\pi)\hat{\delta} + \left(1-\hat{\delta}\right)(4-3\pi)\pi \geq 0 \\ \Longleftrightarrow \quad & -3\left(1-\hat{\delta}\right)\pi^2 + \left(4-\hat{\delta}\right)\pi - 2\hat{\delta} \geq 0. \end{aligned}$$

Then immediate sale will be optimal when

$$\frac{1}{3} > \pi \geq \frac{4-\hat{\delta} \pm \sqrt{16-32\hat{\delta}+25\hat{\delta}^2}}{6\left(1-\hat{\delta}\right)}.$$

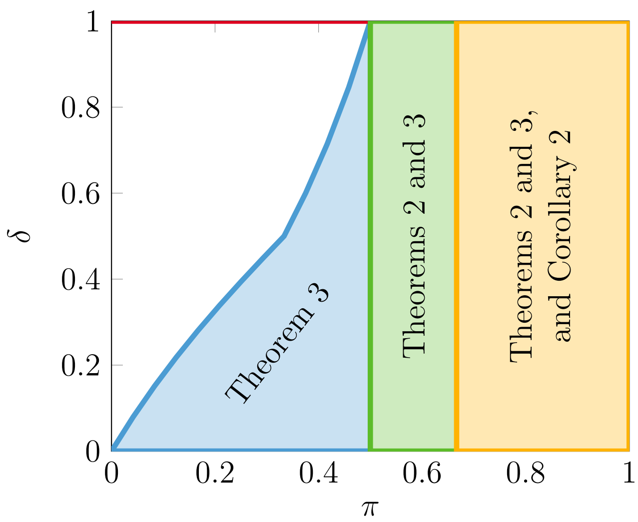

Figure 3: Temporal nondiscrimination is optimal in all shaded regions. In the rightmost (orange) region, value and patience are positively correlated, and Corollary 2 applies. In the middle (green and orange) region, the optimal monopoly price is independent of patience. In the full region, optimality of nondiscrimination follows from Theorem 3. Perfect separation of buyers into distinct time periods is weakly optimal when $\delta = 1$ (red line). In the remaining unshaded region, patient bidders receive some allocation at time $t = 0$ and some at time $t = 1$.

Note that this has a solution if and only if

$$2\left(1-\hat{\delta}\right) \geq 4-\hat{\delta}-\sqrt{16-32\hat{\delta}+25\hat{\delta}^2} \iff 2-6\hat{\delta}+4\hat{\delta}^2 \geq 0 \iff \hat{\delta} < \frac{1}{2}.$$

By contrast, assume that $\hat{\delta} = 1$ and $\pi = 0$, so that the relatively patient type is infinitely patient and all buyers are either high-value and impatient, or low-value and patient. In this case, it is straightforward to see that the optimal mechanism is to sell immediately to any high-value buyer, or to a low-value buyer in the final period if no high-value buyer arrives.[39]

B.1 Perfect separation

Our main results consider the optimality of temporal nondiscrimination. We now show, in this example, that perfect temporal separation—where discount type is allocated in one period, and ther other is allocated in another—is generically nonoptimal. Because we cannot

[39]This mechanism is optimal because it extracts all surplus from the buyers.

apply our main results, we respecify the seller's problem:

$$\max_{q,t} \sum_{v} t(v,\delta) f(v,\delta),$$

$$\text{s.t. } v\delta \cdot q(v,\delta) - t(v,\delta) \geq 0 \qquad \mu \qquad \text{(IR)}$$

$$\sum_{\delta} \sum_{v' \geq v} \sum_{t} q_t(v,\delta) \leq B(v) \qquad \beta \qquad \text{(Feas.)}$$

$$q_t(v,\delta) \geq 0 \qquad \gamma \qquad \text{(Feas.)}$$

$$v\delta \cdot q(v',\delta') - t(v',\delta') \leq v\delta \cdot q(v,\delta) - t(v,\delta) \qquad \lambda. \qquad \text{(IC)}$$

The right-hand variables are the multipliers on the constraints in the respective Lagrangian. The first-order conditions of this problem are:

$$\begin{aligned} 0 &= f(v,\delta) - \mu(v,\delta) + \sum_{(v',\delta') \neq (v,\delta)} \lambda(v,\delta|v',\delta') - \sum_{(v',\delta') \neq (v,\delta)} \lambda(v',\delta'|v,\delta), \\ 0 &= v\delta_t \mu(v,\delta) - \sum_{v' \leq v} \beta(v') f(v',\delta) + \gamma_t(v,\delta) \\ &\qquad - \sum_{(v',\delta') \neq (v,\delta)} v'\delta'_t \lambda(v,\delta|v',\delta') + \sum_{(v',\delta') \neq (v,\delta)} v\delta_t \lambda(v',\delta'|v,\delta). \end{aligned}$$

We say that the seller engages in perfect temporal separation if whenever a buyer with one discount type receives an allocation in period t, no buyer with another discount type ever receives an allocation in period t. Because in our example the impatient type $\hat{\delta} = (1,0)$ does not value consumption in period $t = 1$, if the seller engages in perfect temporal discrimination they will sell to the impatient type only in period $t = 0$, and to the patient type only in period $t = 1$.

Observe that, when $\pi \geq 1/2$ and $\delta < 1$, the optimal mechanism is to sell at a posted price of $p_0^\star = 1$ at time $t = 0$, and to not sell in period $t = 1$. This is because the optimal mechanism sells only to a buyer with value $v = 1$, even when discount types are common knowledge. Delaying sale to a patient buyer sacrifices some available surplus, which is achievable through immediate sale. Thus if there is perfect temporal separation, it must be that $\pi < 1/2$. In this case, the optimal mechanism is to sell in period $t = 0$ at price $p_0^\star = 1$ and in period $t = 1$ at price $p_1^\star = 1/2$. In this optimal mechanism, it follows that the following constraints are slack:

$$\gamma_0\left(1,(1,0)\right)=0,\ \gamma_1\left(\cdot,(1,\delta)\right)=0,\ \mu\left(1,(1,\delta)\right)=0,$$
$$\lambda\left(1,\cdot\middle|\frac{1}{2},(1,0)\right)=0,\ \lambda\left(\cdot,(1,\delta)|\cdot,(1,0)\right)=0,\ \text{and } \lambda\left(1,(1,0)|\cdot,(1,\delta)\right)=0.$$

Substitute these multipliers into the first-order conditions with respect to $q_t(1,(1,\delta))$,

$$\frac{1}{2}\pi\beta\left(1\right)+\frac{1}{2}\left(1-\pi\right)\beta\left(\frac{1}{2}\right)=\gamma_0\left(1,(1,\delta)\right)-\frac{1}{2}\lambda\left(1,(1,\delta)\middle|\frac{1}{2},(1,\delta)\right)+\lambda\left(\frac{1}{2},(1,\delta)\middle|1,(1,\delta)\right),$$
$$\frac{1}{2}\pi\beta\left(1\right)+\frac{1}{2}\left(1-\pi\right)\beta\left(\frac{1}{2}\right)=\gamma_1\left(1,(1,\delta)\right)-\frac{1}{2}\delta\lambda\left(1,(1,\delta)\middle|\frac{1}{2},(1,\delta)\right)+\delta\lambda\left(\frac{1}{2},(1,\delta)\middle|1,(1,\delta)\right).$$

Note that these equations are jointly satisfied only if $\delta=1$ or all multipliers are 0. Since we are looking to show that these equations are satisfied only if $\delta=1$, we assume for now that all multipliers are 0, and in particular that $\beta(1)=\beta(1/2)=0$.

Substituting $\beta(1)=\beta(1/2)=0$ into the first-order condition with respect to $q_0(1,(1,0))$ gives

$$0=\mu\left(1,(1,0)\right)+\lambda\left(\frac{1}{2},(1,0)\middle|1,(1,0)\right).$$

Thus $\mu(1,(1,0))=\lambda(1/2,(1,0)|1,(1,0))=0$. And substituting these values in turn into the first-order condition with respect to $t(1,(1,0))$ gives $0=1/3$, a contradiction. It follows that perfect temporal separation is optimal only when $\delta=1$.

B.2 Constrained optimality

Suppose the seller is constrained to offer a mechanism which is temporally simple: if a buyer receives a probabilistic allocation in period t, she cannot receive a probabilistic allocation in any other period. And, all other buyers with the same discount type receive allocations only in period t. That is, for all $(v,\delta)\in\Theta$,

$$q_t\left(v,\delta\right)>0\quad\Longrightarrow\quad q_{t'}\left(v',\delta\right)\ \forall t'\neq t, v'\in\mathcal{V}.$$

In this example, there are two relevant temporally simple mechanisms: either selling to all buyers at a single posted price at time $t=0$, or selling to impatient buyers $\delta=0$ at a posted price $p_0^\star=1$ at time $t=0$ and selling to patient buyers $\delta=\hat{\delta}$ at a posted price $p_1^\star=1/2$ at time $t=1$. Because impatient buyers receive no value from consuming in period $t=1$, we do not need to consider mechanisms in which their consumption is deferred. And when the optimal posted price does not depend on the buyer's discount type, there is no advantage to deferring any type's consumption.

Observe that there are two cases for optimal posted prices in a mechanism which allocates only in period $t = 0$. When $p > 1/3$ the optimal posted price is $p_0^\star = 1$, and when $p < 1/3$ the optimal posted price is $p_0^\star = 1/2$. We consider the cases in turn.

- When $\pi > 1/3$, the maximum revenue from a mechanism which allocates only in period $t = 0$ is $\Pi_0 = 1/3 + \pi/2$. If the seller implemented a mechanism which sold to patient types in period $t = 1$, the optimal revenue would be $\Pi_1 = 1/3 + \delta/4$. Then selling in only period $t = 0$ is (constrained) optimal when

$$\frac{1}{3} + \frac{1}{2}\pi > \frac{1}{3} + \frac{1}{4}\delta \quad \Longleftrightarrow \quad \pi > \frac{1}{2}\delta.$$

- When $\pi < 1/3$, the maximum revenue from a mechanism which allocates only in period $t = 0$ is $\Pi_0 = 1/2$. If the seller implemented a mechanism which sold to patient types in period $t = 1$, they would post a higher price in period $t = 0$, and the optimal revenue would be $\Pi_1 = 1/3 + \frac{1}{4}\delta$. Then selling in only period $t = 0$ is (constrained) optimal when

$$\frac{1}{2} > \frac{1}{3} + \frac{1}{4}\delta \quad \Longleftrightarrow \quad \delta < \frac{2}{3}.$$

C Proofs for extensions

C.1 Parameterized models of discounting

Lemma 2. *Let (B, h) be a parameterized model of discounting. There is $\bar{T}$ such that for all $T', T'' > \bar{T}$, and all $\beta \in B$,*

$$U_{\mathcal{D}_{T'}}\left(\vec{\delta}\ (\beta; T')\right) = U_{\mathcal{D}_{T''}}\left(\vec{\delta}\ (\beta; T'')\right).$$

Proof. This follows immediately from the fact that the discount parameter space B is finite, and either $\vec{\delta}\ (\beta; t) \geq \vec{\delta}\ (\beta'; t)$ for all t, or there is some finite t such that $\vec{\delta}_t\ (\beta; t) < \vec{\delta}_t\ (\beta'; t)$. □

C.2 Ambiguous temporal preferences

Proof of Proposition 3. Note that optimal temporally-nondiscriminatory mechanisms (Myerson's rule, or Board and Skzrypacz's rule) are feasible in this context, since the marginal distribution of valuation types is known. Then it is sufficient to show that any other mechanism will yield strictly lower maxmin revenue. For any fixed $F \in \mathcal{F}$,

$$\inf_{F' \in \mathcal{F}} \mathbb{E}_{F'} \left[p\left(\tilde{v}, \tilde{\delta}, \tilde{\tau}\right) \right] \leq \mathbb{E}_{F} \left[p\left(\tilde{v}, \tilde{\delta}, \tilde{\tau}\right) \right].$$

Then the seller's revenue under any mechanism (p, t) is bounded above by what would be obtained if the true distribution of values was F. When F satisfies Corollary 1 temporal nondiscrimination is strictly optimal, in the sense that any mechanism which alters the allocation strictly reduces the seller's revenue. □

C.3 Future consumption contracts

Assume now that $\mathcal{D}' \subseteq [0,1]^T$ is a set of discount types, with the restriction that for any $\delta \in \mathcal{D}$ there is t such that $\delta_t = 1$. This restriction is equivalent to requiring that an agent's value v may be realized in some period.[40] Let $\mathcal{F}'$ be the set of type distributions consistent with $\operatorname{Supp} F_\delta \subseteq \mathcal{D}'$.

Define a *future consumption contract* so that an allocation at time t may be consumed either at time t, or once at any time thereafter. Given allocation q, consumption under a future consumption contract is

$$x(\delta) \in \operatorname{argmax}_{x' \in [0,1]^T} x' \cdot \delta, \text{ s.t. } \sum_{t'=0}^{t} x'_{t'} \leq \sum_{t'=0}^{t} q_{t'} \ \ \forall t.$$

Lemma 3. *Given an allocation q and discount type δ, utility under a future consumption contract is identical to the utility obtained by an agent with discount type δ', $\delta'_t = \max_{t' \geq t} \delta_{t'}$ receiving allocation q.*

A future consumption contract yields buyer utility identical to the utility of a buyer with a standard decreasing discount type. Then properties of future consumption contracts follow from properties of ambiguous joint type distributions (Proposition 3).

Proposition 4. *Suppose that for any $\mathcal{D}$ such that there exists $F \in \mathcal{F}'$ with $\operatorname{Supp} F_\delta = \mathcal{D}$, there is $F' \in \mathcal{F}'$ with $\operatorname{Supp} F'_\delta = \mathcal{D}$ and such that $F'(\cdot|v) = F'_\delta$ for all $v \in \mathcal{V}$. Then temporal nondiscrimination over future consumption contracts is optimal in the ambiguous seller's problem.*

Proof. Fix a distribution $F' \in \mathcal{F}$ such that $F'(\cdot|v) = F'_\delta$; optimal revenue given distribution F' weakly exceeds optimal revenue under ambiguous F. The proof of Theorem 2 does not

[40] In our main analysis, we assume that $\delta_0 = 1$ for all $\delta \in \mathcal{D}$, and that time-0 consumption is weakly optimal. In this section, we relax this assumption.

depend on the period t in which $\delta_t = 1$, only on the existence of such a period. Then when $\delta_t = 1$, a period-t temporally-nondiscriminatory mechanism is optimal. Moreover, any other mechanism which induces a different allocation reduces revenue. Because $F'(\cdot|v) = F'_\delta$, the period-t temporally-nondiscriminatory rule depends on neither δ nor t. An identical temporally-nondiscriminatory rule is being applied in each period (in which an auction is run), and the resulting allocation may be implemented by a temporally nondiscriminatory mechanism for future consumption contracts at time $t = 0$. $\square$

Proposition 4 implies that when temporal preferences are ambiguous and independence is plausible, it is optimal for the seller to sell the good today and wait for the buyer to receive the good at any later date. When there is a single buyer, there is no competition for the good and the optimal mechanism without temporal discrimination is equivalent to a posted price. In this case a temporally-nondiscriminatory mechanism for future consumption contracts is equivalent to committing to a posted price, then waiting for the buyer to arrive and purchase. This is familiar from Riley and Zeckhauser [1983], among others.