Intertemporal Allocation with Unknown Discounting

Justin Burkett* and Kyle Woodward†

March 2021

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 heterogeneity in buyers' discount factors; for example, discriminating on discount factor is not optimal when buyers with higher values are also more patient. These conditions also apply when sellers face ambiguity regarding buyers' discount factors. Our results provide a novel justification for temporal nondiscrimination when the seller is incompletely informed about buyers' temporal preferences.

*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, Marek Pycia, and Rakesh Vohra, as well as seminar audiences at INFORMS and SEA for valuable feedback and comments.

1 Introduction

Determining how to optimally sell a durable good is surprisingly difficult, due to in large part to the temporal scope of the problem: the seller must decide not only to whom to sell, but also when to sell. Considerable progress has been made towards characterizing solutions to important cases of this problem,[1] but many analyses remain restrictive in potentially important ways. For example, typical models assume that buyers share a known discount factor,[2] in contrast with mounting evidence of heterogeneity of temporal preferences (Mischel et al., 1989; Kirby and Maraković, 1995; Green and Myerson, 2004; Hakimi, 2013).[3] In principle, a seller might profitably exploit heterogeneity in buyers' discount factors by engaging in *temporal discrimination*. On the one hand, any information the seller has about variation in discount factors will help the seller determine whom to sell to when; but on the other hand, such considerations may make already complex problems intractable for the seller to solve. Furthermore, existing solutions that effectively ignore heterogeneity in time preference have intuitively appealing features and seem to explain observed behavior (Board and Skrzypacz, 2016).

In this paper, we explore the possibility that a seller who is aware of buyer heterogeneity in time preference might (optimally) implement a mechanism that does not discriminate along this dimension. We start from the simple observation that even when potential buyers have heterogeneous discount factors, it remains feasible for the seller to use a mechanism which does not discriminate on discount factor. In other words, the seller is free to ignore this heterogeneity. We then take a candidate nondiscriminatory mechanism — optimal with respect to a common discount factor — and apply tools from the theory of linear programming to obtain conditions under which this candidate mechanism is optimal when buyers have heterogeneous discount factors. That is, we describe conditions under which the seller optimally ignores heterogeneity in temporal preferences.

We show that the conditions under which the seller does not discriminate on discount factor can be stated simply, and allow for a broad set of statistical relationships between buyers' discount factors and initial valuations. In our model, buyers arrive stochastically and have private information about both their initial valuation for the good and the discount factor by which they devalue future consumption. We show that if these two pieces of information are positively related, in the sense of first-order stochastic dominance, it is

[1] The literature on dynamic selling mechanisms is extensive. See, for example, Stokey (1979), Riley and Zeckhauser (1983), Gallego and Van Ryzin (1994), Gallien (2006), and Dilme and Li (2019).

[2] Exceptions include Pai and Vohra (2013) and Mierendorff (2016), in which buyers have identical discount factors but heterogeneous deadlines for consumption.

[3] Within individuals, discount factors are also observed to vary across commodities (Ubfal, 2016).

optimal for the seller to ignore the heterogeneity in discount factors by discriminating only on initial valuation and arrival time. The seller does this by using a mechanism that would be optimal if all buyers had the same discount factor as herself. In plain language, the seller does not use information about buyers' discount rates if buyers with high values tend also to be more patient.[4]

More specifically, we consider the problem faced by a seller with one unit of a durable good facing a stream of buyers who arrive stochastically over an infinite horizon. In each period, the seller observes each new arrival, but the buyers are privately informed of both their (initial) valuation and their discount factor.[5] We assume that the seller can commit to a sales mechanism ex ante, and that the buyers' types are i.i.d. draws from a finite set, but our arguments require few restrictions beyond these. Importantly, our model allows for the set of types and the statistical relation between value and discount factor to be arbitrary.

The intuition behind our results can be understood in terms of weighing the benefits to discriminating on temporal preferences against the costs. As is typical in the study of optimal sales mechanisms, the "virtual value" of a buyer — the buyer's value, adjusted downward to account for information rents — can be understood as a measure of the marginal revenue available from that buyer (Bulow and Roberts, 1989). In our model, the difference between a buyer's "conditional virtual value," which we define as the virtual value of a buyer with a known discount factor, and their "average virtual value," which is the virtual value of a buyer after integrating out the discount factor, determines the improvement in marginal revenue to the seller from increasing the allocation for this particular buyer. The seller weighs this improvement in revenue against the implicit cost of increasing this buyer's allocation, which enters through other buyers' incentive constraints. Intuitively, when the seller increases the allocation of one buyer type, she must also increase the allocation for all buyer types who want to mimic that type. Unless conditional virtual values are equal to average virtual values for all types,[6] the seller can always find types that would increase revenue relative to the nondiscriminatory mechanism. She will not separate these types if she loses sufficient

[4]By "more patient," we mean that the buyer has a higher discount factor. Patience could also be defined in an absolute sense by calling a buyer more patient if they are more willing to defer consumption to the next period, but this notion of patience confounds values and discount factors. In the latter sense, increasing a buyer's value for immediate consumption will make them less patient (by increasing the utility they lose by waiting to consume), while in our analysis increasing a buyer's value, holding fixed their discount factor, has no effect on their patience. Pai and Vohra (2013) and Mierendorff (2016) define buyers as less patient the more immediate is their exogenous purchase deadline, largely in line with our definition.

[5]We constrain attention to buyers who discount the future at an exponential rate (Samuelson, 1937). Empirical work has shown that even the form of temporal discounting may be idiosyncratic (Benzion et al., 1989). A working version of this paper provides results for buyers who are potentially non-exponential discounters.

[6]Conditional virtual values equal average virtual values if and only if discount factors are independent of valuations.

revenue from increasing the utility of the mimicking types.

Determining which types will want to mimic (i.e., which incentive constraints bind), and hence which allocations need to be adjusted in response to adjusting one type's allocation, is a famously difficult question in multidimensional mechanism design. However, due to the nature of the question we ask and the methods we use to answer it, we are able to prove our results without ever identifying the precise incentive constraints that bind.[7] We start from the observation that the candidate mechanism, which ignores heterogeneity in discount rates, is feasible. To prove that it is optimal we then need only identify a subset of incentive constraints under which it is optimal. Furthermore, using duality results from linear programming, it is sufficient for some subset of incentive constraints to identify choices for the dual variables (or multipliers) under which the appropriate complementary slackness conditions are satisfied.In other words, optimality of the candidate mechanism is equivalent to resolving the question of whether a set of linear inequalities has a solution, which is a well-understood problem.

Our main result shows that the candidate, nondiscriminatory mechanism is optimal under a condition relating the conditional and average virtual values of each type (Theorem 1). We derive this result by showing that if we only include what we refer to as "downward" incentive constraints, the complementary slackness conditions yield a system that is equivalent to a standard problem on the existence of a "flow" in a network that consists of nodes with demands for flow and arcs with capacities for carrying flow between nodes.[8] Once the analogy is established, the argument we use is an immediate consequence of Gale's feasible flow theorem (Gale, 1957). Relating conditional and average virtual values is natural given the structure of the problem, but the condition in Theorem 1 can be difficult to interpret. We therefore show that this condition is implied by the simple statistical condition given in Corollary 1, which states that the distribution of discount factors is increasing in the buyer's valuation in the sense of first-order stochastic dominance. Intuitively, when discount factor and valuation are positively related, buyers with higher discount factors are more likely to have higher values and consequently obtain higher information rents, because their information rents derive from the ability of types with even higher values to mimic them. In other words, under the conditions presented in Theorem 1 or Corollary 1, when the seller

[7]It is well-known that incentive constraints in multidimensional mechanism design problems might bind in multiple "directions," meaning multiple incentive constraints involving the same type might bind. Because the type space in our problem is finite it is non-convex, we cannot rule out a priori that non-local incentive constraints bind at the optimal solution either (Carroll, 2012). Our approach does not rely on eliminating either of these possibilities.

[8]The downward constraints in our model are the ones that prevent types with higher willingness to pay in every period from mimicking those with lower willingnesses to pay. In particular, more-patient buyers cannot misrepresent as less-patient buyers. In Section 4 we allow for arbitrary incentive constraints to bind.

wants to increase the allocation of some type, she must also — through binding downward incentive constraints — increase the allocation of higher types, who are likely to have larger information rents and hence reduce the seller's average revenue.

An immediate consequence of our analysis is that a seller who is uncertain of 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.[9] 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 sellers of durable goods. Practically, our results show that sellers should not use the passage of time to screen buyers when value and patience are positively related, or when an uncertain, ambiguity-averse seller believes that value and patience might be positively related. We know of no studies separately examining buyer values and discount factors, an absence which is consistent with ambiguous beliefs regarding their interrelation as well as a general absence of discrimination on discount factor in applied settings.

Our results contribute to the ongoing study of why simple mechanisms can persist in relatively complicated settings. A natural reading of the multidimensional mechanism design literature suggests that complete solutions are elusive, and that optimal mechanisms can be unwieldy and complicated. Indeed, 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. That is, in spite of the rich set of available mechanisms, full consideration of agents' incentives encourages the use of a relatively simple sales mechanism, 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.

This paper proceeds in Section 2 by defining our model. Section 3 establishes our main result, and Section 4 extends the analysis to allow for arbitrary incentive constraints on discount rate. Section 5 considers the optimal mechanism when the seller faces ambiguity regarding the distribution of temporal preferences. Related literature is deferred to Section 6.

[9] Carroll (2017) establishes a version of this claim for the case where the seller knows the marginal distributions of buyers' types but is of uncertain the joint distribution. In our anlaysis, the seller does not even need to know the marginal distribution of buyers' discount rates.

2 Model

A seller offers one unit of an indivisible good for sale to buyers potentially arriving over time. Time is discrete, $t \in \{0, 1, \ldots, \}$, and allocation may take place in any period. The seller commits to a mechanism in the first period, $t = 0$, and discounts the future at (exponential) rate δ_s.

Buyers arrive stochastically over time. There are n *kinds* of bidders, $i \in I = \{1, \ldots, n\}$, and the probability a bidder of kind i arrives in period τ is g_i, independent of time t. Where there is no risk of confusion, we refer to a buyer of kind i simply as buyer i. The seller observes the kinds of buyers who arrive in period τ but the buyers do not, and a buyer who arrives in period τ remains until the good is allocated (at which point the game ends).[10] A buyer of kind i who arrives at time τ has *value type* $v^i \in \mathcal{V}^i \subset [0, 1]$ and *discount type* $\delta^i = (1, \hat{\delta}_i, \hat{\delta}_i^2, \ldots) \in \mathcal{D}^i$, where $\hat{\delta}_i \in [0, 1]$ is the buyer's exponential discount factor. The buyer discounts consumption t periods in the future by discount factor δ_t^i. Each buyer's utility is quasilinear in transfers, and if her allocation is $q = (q_t)_{t=\tau}^{\infty}$ her interim utility is

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

We assume that the support of value types $\mathcal{V}^i$ and the support of discount types $\mathcal{D}^i$ are both finite, and for simplicity we assume further that there is $\varepsilon > 0$ so that $\mathcal{V}^i = \mathcal{V} \equiv \{\varepsilon, 2\varepsilon, \ldots, 1 - \varepsilon, 1\}$. 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. To distinguish random variables, we add a tilde. For example, $\tilde{\theta}^i = (\tilde{v}^i, \tilde{\delta}^i)$ is the random variable corresponding to buyer i's type, and $\tilde{\tau}$ is the random variable corresponding to her arrival time.

We 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).[11] 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. We use $\mathbb{E}_{\Theta^i}$ 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 $f^i(v, \delta) =$

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

[11] 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 seller.

$f(v,\delta)$ for all $(v,\delta)\in\Theta$. When discussing the symmetric case, we omit the i superscripts.

2.1 Mechanisms

Because the seller commits 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 arrived in period τ and reported the type (v,δ). We use $q^i(v,\delta,\tau)$ to indicate the vector of probabilities across time periods, from arrival time τ onward. 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.[12] 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

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

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 and our optimal mechanism is a time-independent threshold rule, the optimal mechanism satisfies incentive compatibility even when arrival times are private information.*

2.2 The seller's problem

Our analysis considers when it is optimal for the seller to ignore buyers' discount types when allocating the good. If a mechanism depends only on buyer values and not on discount types, we say that it *does not temporally discriminate.* In our model, the optimal mechanism without temporal discrimination is a series of auctions with time-independent reserve prices,

[12]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.

and a buyer has a positive probability of receiving a good only in the period in which she enters the mechanism.[13]

The general revenue maximization problem is

$$\max_{\{q^i,p^i\}_{i\in I}} \quad \sum_{i\in I}\sum_{\tau=0}^{\infty}\mathbb{E}_{\Theta^i}\left[g_i\delta_s^{\tau}p^i\left(\tilde{v},\tilde{\delta},\tilde{\tau}\right)\right]$$

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

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

$$q^i \text{ is feasible.} \tag{F}$$

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

The seller's problem considers revenue maximization net of buyers' incentives to reveal their private information. As in other mechanism design contexts, the revenue a seller can extract from a given type is captured by the type's virtual value.

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 average virtual value $m^i(\cdot)$ is weakly increasing. When the seller does not temporally discriminate, the optimal mechanism is definable in terms of a threshold $v_i^{\star}\geq v_i^{+}$, where

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

The cutoff value $v_i^{\star}$ is the value for which the seller is (nearly) indifferent between allocating

[13] See, e.g., Board and Skrzypacz (2016).

today and waiting for new arrivals tomorrow; for details, see equation (6) of Board and Skrzypacz (2016). Because the value-relevant type (v, δ) is independent of arrival time τ, virtual values (and hence thresholds $v_i^\star$) do not depend on arrival time. 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 nondiscrimination 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 nondiscrimination 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 the optimality of temporal nondiscrimination. We first provide a detailed intuitive argument in a simplified, semi-static model in which all potential buyers are present at time $t = 0$. This simplificiation removes the need to consider the possibility that the seller would rather not sell today in hopes of drawing a better set of buyers tomorrow. The crux of our analysis is the allocational feasibility constraint (Border, 1991). When we moving from the semi-static to the fully-dynamic model the feasibility constraint changes but the intuitive arguments remain unchanged.

3.1 Intuition from the semi-static model

To develop intuition, we consider the *semi-static* model in which all potential buyers are present at time $t = 0$. Aside from the presumption that there are no future arrivals, the semi-static model is identical to our main model. For simplicity we restrict attention to the symmetric case, where $f^i = f$ for all buyer kinds i.

In the semi-static model, the optimal temporally nondiscriminatory mechanism is a canonical optimal auction (Myerson, 1981) without any future allocation, $q_t(v, \delta) = 0$ for all $t > 0$ and all types (v, δ).[14] To evaluate whether nondiscrimination is optimal, it is

[14]Under temporal nondiscrimination, the buyer's allocation can depend only on her value type and not on her discount type. Because buyers discount the future (we allow $\delta = 1$, but since there are at least two discount types there must be a discount type $\delta < 1$) deferring allocation into the future strictly reduces the expected value of consumption, and in turn reduces revenue. Then the optimal temporally nondiscriminatory

therefore sufficient to consider whether a small deviation from the optimal static auction can improve the seller's profits. The optimal static auction allocates to all buyers whose average virtual values are strictly positive and never allocates to buyers whose average virtual values are strictly negative: $q_0(v,\delta) > 0$ if $m(v) > 0$, and $q_0(v,\delta) = 0$ if $m(v) < 0$.

Consider a small deviation from the optimal static mechanism, which increases the allocation allocation probability of type (v,δ) in time $t > 0$ by amount $\xi > 0$. In the static mechanism, the IC constraints bind from higher values to lower values, and increasing the allocation of type (v,δ) requires increasing the allocation of all types (v',δ) with value $v' > v$.[15] However, it also requires increasing the allocation of all more-patient types (v,δ') with $\delta' > \delta$: because the less-patient type (v,δ) is willing to accept the adjusted allocation, and all discount types receive the same allocation in the temporally nondiscriminatory mechanism, the more-patient type (v,δ') must strictly prefer the adjusted allocation. In turn, all higher-value, more-patient types must have their allocations improved, or incentive compatibility will be violated.

The additional revenue obtained by increasing the allocation to value-type v, net of incentive compatibility for all value-types $v' > v$, is the virtual value $m(v)$; similarly, the additional revenue obtained by increasing the allocation to type (v,δ), net of incentive compatibility for all types (v',δ) with $v' > v$, is the conditional virtual value $m(v|\delta)$. As argued above, incentive compatibility with respect to reported discount factors means that increasing the allocation to type (v,δ) is also associated with increasing the allocation to all types (v,δ') with $\delta' > \delta$, and the additional revenue from this increase is $m(v,\delta')$. Temporal discrimination is only (potentially) profitable if the expected benefit from increasing the allocation to all more-patient types, net of incentive compatibility, is positive and yields higher revenue benefit than increasing the allocation of any buyer with value-type v, irrespective of discount rate. That is, temporal discrimination of type (v,δ) is potentially strictly beneficial only if

$$\mathbb{E}_\Theta\left[m\left(v\middle|\tilde{\delta}\right)\middle|v,\tilde{\delta}\geq\delta\right] > \max\left\{0, m\left(v\right)\right\}. \tag{1}$$

Then temporal nondiscrimination is optimal if inequality (1) does not hold for any type (v,δ). Our main result, Theorem 1, is an algebraic rearrangement of this condition.

Note that this derivation presumes that the incentive constraints are binding only from higher types to lower types, and increasing the allocation of a type (v,δ) requires increasing the allocations of all types $(v',\delta') \geq (v,\delta)$. It does not consider the possibility that the

mechanism allocates immediately and depends only on the buyers' value types.

[15] We say that the IC constraint binds from one type to another if the constraint which prevents the former from misreporting as the latter is binding. If the IC constraints from higher values to lower values are not binding a small increase in transfers can improve the seller's revenue without violating incentive compatibility.

incentive constraint is binding from a lower-value, more-patient type to a higher-value, less-patient type.[16] Accounting for these "diagonal" constraints strenthens, rather than weakens, our results: expanding the set of incentive constraints considered raises the implicit costs of increasing the allocation of type (v, δ), and if increasing this allocation does not improve revenue given a smaller set of constraints it will not improve revenue given a larger set of constraints.

Having built intuition from the semi-static case, we next formalize our results in the full model with dynamic arrivals. The preceding intution suggests that, given an arrival time τ, the seller will only temporally discriminate against buyers who arrive at time τ by potentially "calling back" buyers at a later time after subsequent arrivals have proved unprofitable. However, as shown in Board and Skrzypacz (2016), it is never optimal for the seller to recall buyers who are tentatively unallocated, and the threshold type for allocation remains constant over time. Thus inequality (1) remains valid in the case with dynamic arrivals.

3.2 Formal results for the fully-dynamic model

We now return to the fully-dynamic model set out in Section 2. We make use of the following definition our subsequent results. For a buyer i with type $(v, \delta) \in \Theta^i$, define the seller's *incremental anticipated virtual value* μ^i as

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

where $m^i_+(v) = \max\{m^i(v), m^i(v^\star_i)\}$.[17] 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. Note that

$$\mathbb{E}\left[m^i\left(\tilde{v}\middle|\tilde{\delta}\right)\middle|\tilde{v} = v, \tilde{\delta} \geq \delta\right] - m^i_+(v) = \frac{1}{\Pr\left(\tilde{v} = v, \tilde{\delta} \geq \delta\right)} \sum_{\delta' \geq \delta} \mu^i(v, \delta'),$$

and the partial sums of incremental anticipated virtual value correspond to inequality (1) above. That is, the partial sums of $\mu^i(v, \delta)$ are the benefit associated with increasing the allocation to type (v, δ), net of binding incentive constraints for higher-value, more-patient

[16] Deferring the consumption of an impatient buyer while holding their utility constant will improve the utility a patient buyer obtains from the impatient buyer's allocation. When values are similar and discount rates are significantly different, the patient type may strictly prefer the impatient type's deferred allocation, even with a lower initial valuation v.

[17] Equivalently, we may write $m^i_+(v) = m^i(v^\star_i(v))$.

types. Negativity of the left-hand side is equivalent to negativity of the right-hand sum, and the intuitive derivation and condition from the preceding subsection is formalized in Theorem 1.

Theorem 1 (Optimality of nondiscrimination)**.** *Temporal nondiscrimination is optimal if for all buyers i and all types $(v, \delta) \in \Theta^i$,*

$$\sum_{\delta' \geq \delta} \left\{ \mu^i \left(v_i^\star \left(v \right), \delta \right) - \mu^i \left(v - \varepsilon, \delta \right) \right\} \geq 0. \tag{2}$$

The primary complication in establishing 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, given in Appendix A, 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 $q_\tau^i(v, \delta^i, \tau)$. 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. In line with the intuition given above, we say that the shadow cost of the feasibility constraint for type v is the average marginal value $m_+^i(v)$. 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.

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.[18] We show in the proof of Theorem 1 that when we consider the downward IC constraints — corresponding to more-patient buyers not misreporting as less-patient buyers, as considered in the intuitive derivation — 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 network consists of nodes, which are identified with types on our model, and arcs (directed links) that carry "flow," where the flow between two types with different discount types determines the value of the dual variable on the corresponding

[18]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.

IC constraints, and the flow between adjacent value types with identical discount types determines the value of the dual variable on the corresponding constraint ensuring weakly positive allocations. The former dual variable can be interpreted as the shadow price of loosening an IC constraint between two types.[19] 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 comparatively simple steps.

For one interpretation Theorem 1, note that a sufficient condition for (2) is

$$\sum_{\delta' \geq \delta} \left\{ \mu^i \left(v, \delta'\right) - \mu^i \left(v - \varepsilon, \delta'\right) \right\} \geq 0, \ \forall v \in \mathcal{V}. \tag{3}$$

This follows because summing (3) over all $v' \in \{v, \ldots, v_i^\star(v)\}$ gives inequality (2). Fixing a value type v, summing (3) over $v' \geq v + \varepsilon$ yields $-\sum_{\delta' \geq \delta} \mu^i(v, \delta) \geq 0$, which is equivalent to

$$\mathbb{E}\left[m_+^i \left(v\right) - m^i \left(v \middle| \tilde{\delta}\right) \middle| v, \tilde{\delta} \geq \delta \right] \geq 0. \tag{4}$$

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. An optimal temporally nondiscriminatory mechanism will sell only to value types with positive average virtual value. 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 (4).[20] More directly, when inequality (4) 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 (2) and (4) 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, purely statistical condition is sufficient. For example, inequality (2) is satisfied if buyer types with higher values are more likely to have higher discount types in the following sense.

Corollary 1 (Statistical condition for nondiscrimination)**.** *Temporal nondiscrimination is*

[19] 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 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.

[20] Arrival times τ do not feature in any of these conditions, since the distribution of the value-relevant type (v, δ) is independent of τ, and arrivals are identically distributed across time periods.

	0	$\hat{\delta}$
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.

optimal 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.[21] 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.

An example clarifies the relative strengths of Theorems 1 and Corollary 1.

Example 1. *Consider case with a single kind of buyer, $n = 1$, which is guaranteed to arrive in each period, $g_i = 1$, and suppose that the seller significantly discounts the future, $\delta_s \approx 0$.[22]. The type space is $\Theta = \{1/2, 1\} \times \{0, \hat{\delta}\}$, where $\delta \in (0, 1]$. 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.*

Application of Theorem 1 gives 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.

[21]This does not imply the "no haggling" result of Riley and Zeckhauser (1983). We assume that $\delta_0^i = 1$ 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.

[22]Because buyers make transfers upon arrival the seller's discount rate does not affect the optimality of temporal nondiscrimination, and the only effect of the seller's impatience is to encourage sale to the first buyer who arrives (i.e., the seller's impatience lowers $v^\star$).

Applying Corollary 1, we find that $F(\delta|v)$ is nonincreasing in v when

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

Then Corollary 1 shows that temporal nondiscrimination is optimal when $\pi \geq 2/3$, which is exactly when valuation and discount factor are weakly positively correlated. Since $\pi \geq 2/3$ is more restrictive than $\pi \geq 1/2$, Theorem 1 is strictly more general than Corollary 1.

4 Generic misreports of discount rate

The proof of Theorem 1 considers 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 condition in Theorem 1 is overly strong. In this section we take a distinct analytical approach, allowing for additional binding IC constraints which simultaneously 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 2 no longer corresponds to a standard network flow problem. Theorem 2 thus necessitates an approach to feasibility distinct from that applied to Theorem 1. In this case, we employ Farkas' Lemma to derive a distinct sufficient condition for the optimality of temporal nondiscrimination.

Theorem 2. *The optimal mechanism does not temporally discriminate if for all buyers i and all types $(v, \delta) \in \Theta^i$,*

$$\sum_{\delta' > \delta} \frac{1}{\varepsilon} \left(\mu^i\left(v, \delta'\right) - \mu^i\left(v - \varepsilon, \delta'\right)\right)\left(\delta' - \delta\right) v + \left(1 - \delta'\right) m^i_+\left(v\right) f^i\left(v, \delta'\right) \geq 0. \tag{5}$$

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

Although self-evidently more analytically complicated than our earlier results, intuition for inequality (5) may be readily drawn from the discussion surrounding Theorem 1. In particular, the difference $\mu^i(v, \delta) - \mu^i(v - \varepsilon, \delta)$ represents the relative cost of increasing the allocation of type $(v - \varepsilon, \delta)$. Inequality (5) adjusts this cost to account for the fact that increasing the allocation to a given type in an incentive-compatible manner requires

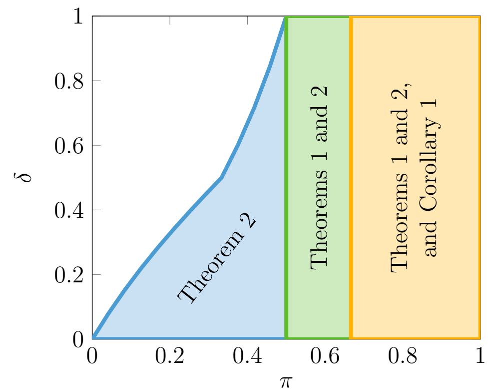

Figure 2: Optimal mechanisms in Example 1. Theorem 2 implies that temporal nondiscrimination is optimal in all shaded regions. Theorem 1 is weaker in this example, and only implies that temporal nondiscrimination is optimal in the right-hand regions, $\pi \geq 1/2$. Corollary 1 is weaker still.

increasing the allocation at some time $t \geq 1$, and not at time $t = 0$; thus buyers' discount factors scale the incremental anticipated virtual values 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 $1 - \delta$.

Returning to Example 1 illustrates additional the power of Theorem 2.

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

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

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.

Remark 2. *In this example, the seller will only engage in perfect temporal separation — that is, will only allocate to distinct discount types in distinct periods — when $\hat{\delta} = 1$. Otherwise, even when temporal discrimination is optimal, the seller will offer patient buyers some probability of allocation in period $t = \tau$ and some probability of allocation in period $t' > \tau$, unless* *patient buyers are infinitely patient. This occurs because when $\hat{\delta} < 1$, 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 interior of the unshaded region of Figure 2 the optimal mechanism sells to all buyers in period $t = \tau$, and to patient buyers in period $t' > \tau$.*[23]

5 Ambiguous temporal preferences

In practice it may be difficult for the seller to evaluate the marginal distribution of discount types, so 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 kind of buyer. 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.[24] The seller believes that the feasible set of joint distributions is $\mathcal{F} \subseteq \{\hat{F} \colon \hat{F}_v = F_v \text{ and } \operatorname{Supp} \hat{F}_\delta = \mathcal{D}\}$. The seller's problem is[25]

$$\begin{aligned}
&\max_{\{q,p\}} \inf_{F \in \mathcal{F}} \sum_{\tau=0}^{\infty} \mathbb{E}_F \left[g \delta_s^\tau p \left(\tilde{v}, \tilde{\delta}, \tau \right) \right], \\
&\text{s.t.} \quad u^i(v, \delta, \tau) \geq u^i(v', \delta' | v, \delta, \tau) \quad \forall i \in I, \tau, (v, \delta) \in \Theta^i, (v', \delta') \in \Theta^i \qquad \text{(AIC)} \\
&\qquad u^i(v, \delta, \tau) \geq 0 \quad \forall i \in I, \tau, (v, \delta) \in \Theta^i, \qquad \text{(IR)} \\
&\qquad q^i \text{ is feasible.} \qquad \text{(AF)}
\end{aligned}$$

Proposition 1. *Suppose that there is $F \in \mathcal{F}$ that satisfies the condition of Theorem 1. Then temporal nondiscrimination is optimal in the seller's problem ambiguous temporal preferences.*

[23] Note that the seller's discount factor δ_s does not factor in to the choice to temporally discriminate, since all transfers are made immediately upon arrival and a buyer is guaranteed to arrive, $g = 1$.

[24] Carroll (2017) analyzes the case in which the seller knows F_δ, and finds that (applied to our setting) temporal nondiscrimination is optimal. Madarász and Prat (2017) show that a seller with a misspecified model can obtain better outcomes with a contingent profit-sharing scheme; by contrast, our seller suffers only from an incomplete understanding of the distribution of patience, and does not need to hedge against unforeseen types. Assuming that the seller knows the set of feasible discount types $\mathcal{D}$ simplifies analysis but is otherwise inessential to our results.

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

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 1. 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 condition of Theorem 1 is satisfied when there is a weakly positive statistical relationship between value and patience. Thus when the seller believes that discount types may be independent of value types, temporal nondiscrimination is optimal.

Corollary 2. *Suppose that there is $F \in \mathcal{F}$ such that value type and discount type are independent, $F(v, \delta) = F_v(v)F_\delta(\delta)$. Then temporal nondiscrimination is optimal in the seller's problem with ambiguous temporal preferences.*

6 Related literature and conclusion

Our model builds on work on dynamic pricing. Our buyers arrive independently over time, a common assumption in the dynamic pricing literature. 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.[26] Pai and Vohra (2013) and Mierendorff (2016) consider the possibility that agents have 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 formal analysis ties most directly to previous work on bundling. Traditional bundling models consider when it is optimal to package multiple goods (or attributes) together, and when it is optimal to sell them individually. McAfee and McMillan (1988) consider the problem faced by a monopolist 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; Fang and Norman (2006) compare the seller's preference for full bundling versus separate sales, and Pycia (2006) shows that "simple" mechanisms are generically

[26] Stokey (1979) finds a declining price curve only when the seller faces positive marginal costs which decline over time. Riley and Zeckhauser (1983) show that, against a stream of buyers, the seller's optimal mechanism is a fixed price in each period.

nonoptimal.[27] In our model, the set of goods corresponds to the ability to allocate a fixed unit at different points in time, and a little more of tomorrow's good comes at the cost of a little less of today's good. In mathematical shorthand, a dynamic allocation of a single good is feasible if $0 \le \sum_t q_t \le 1$, while the feasibility constraint in most bundling analyses is $0 \le q_k \le 1$ for all goods k.[28] 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.

Our main result 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 initial value v in our model, obtainable by consuming at time $t = \tau$, corresponds to the value for the grand bundle in Haghpanah and Hartline (2019), and their Theorem 1 corresponds to our Corollary 1. Our Theorem 1 is stronger than our Corollary 1 (see Example 1), therefore our results are stronger in our context; otherwise, our results neither imply nor are implied by theirs. Our approach to Theorem 2, via Farkas' Lemma, is methodologically distinct.

The proof of our main result 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 address the suboptimality of deferring 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).[29] Our approach is distinct, in that we initially allow only the set of downward discount constraints to bind and derive a condition for the optimality of temporal nondiscrimination given only these constraints; adding unconsidered constraints back to the problem does not affect the feasibility of immediate allocation and therefore does not affect the optimality of temporal nondiscrimination.[30] Our Theorem 2 expands the set of potentially-binding constraints and obtains a sufficient condition which is neither weaker nor stronger than our main result.

[27]With a single buyer, an allocation is feasible in our model only if $0 \le \sum_t q_t \le 1$. This contrasts the feasibility constraint in standard bundling problems, $0 \le q_t \le 1$, and the simple mechanisms of Pycia (2006) are infeasible in our context. See our discussion of Haghpanah and Hartline (2019) below.

[28]Pycia (2006) considers simple mechanisms, where the constraint is $q_t \in \{0, 1\}$.

[29]In the related problem of dynamic contracting, Battaglini and Lamba (2019) show that local incentive constraints are frequently insufficient for global incentive compatibility.

[30]Pavan et al. (2014) observe that incentive compatibility is easier to satisfy in dynamic models than in static models. This follows from the slow revelation of private information in their model, and is not at odds with our finding that incentive constraints cause the seller to not screen on discount factor.

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 buyers' discount factors. We show that when values and discount factors are positively related, the optimal mechanism ignores temporal preferences and allocates to a given buyer either upon arrival or never. We further show that 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.

Marco Battaglini and Rohit Lamba. Optimal dynamic contracting: The first-order approach and beyond. *Theoretical Economics*, 14(4):1435–1482, 2019.

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.

Jeremy Bulow and John Roberts. The simple economics of optimal auctions. *Journal of Political Economy*, 97:1060–1090, 1989.

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.

Francesc Dilme and Fei Li. Revenue management without commitment: Dynamic pricing and periodic flash sales. *The Review of Economic Studies*, 86(5):1999–2034, 2019.

Hanming Fang and Peter Norman. To bundle or not to bundle. *The RAND Journal of Economics*, 37(4):946–963, 2006.

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

Guillermo Gallego and Garrett Van Ryzin. Optimal dynamic pricing of inventories with stochastic demand over finite horizons. *Management science*, 40(8):999–1020, 1994.

Jérémie Gallien. Dynamic mechanism design for online commerce. *Operations Research*, 54 (2):291–310, 2006.

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? *Review of Economic Studies*, 1, 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.

Kristóf Madarász and Andrea Prat. Sellers with misspecified models. *The Review of Economic Studies*, 84(2):790–815, 2017.

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. 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.

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.

Alessandro Pavan, Ilya Segal, and Juuso Toikka. Dynamic mechanism design: A Myersonian approach. *Econometrica*, 82(2):601–653, 2014.

Marek Pycia. Stochastic vs deterministic mechanisms in multidimensional screening. *Report, Univer*, 2006.

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.

Paul A Samuelson. A note on measurement of utility. *The review of economic studies*, 4(2): 155–161, 1937.

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

Diego Ubfal. How general are time preferences? Eliciting good-specific discount rates. *Journal of Development Economics*, 118:150–170, 2016.

A Proofs for Section 3 (Analysis)

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 node-specific requirements and any arc-specific capacity constraints. The variation of feasible flow theorem that we use in Theorem 1 is due to Gale (1957).[31] 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 $b(x)$[32]. 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.

$$\sum_{\{x'|(x',x)\in\mathcal{A}\}} g(x', x) - \sum_{\{x'|(x,x')\in\mathcal{A}\}} g(x, x') = b(x) \qquad \forall x \in \mathcal{N} \tag{6}$$

$$0 \leq g(x, x') \leq k(x, x') \qquad \forall (x, x') \in \mathcal{A}, \tag{7}$$

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

Theorem 3. *There exists a solution, g, to the system in* (6) *and* (7) *if and only if*

$$\sum_{x\in X, x'\in\bar{X}} k(x, x') \geq \sum_{x'\in\bar{X}} b(x') \qquad \forall X \subseteq \mathcal{N}, \tag{8}$$

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.

[31] 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.

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

A.2 Proof of Theorem 1

Proof of Theorem 1. We initially assume that the seller may only allocate to buyers in the first T periods after they arrive, and later extend the analysis to the case where the seller may allocate to a buyer in any period after she arrives. 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 $\delta' < \delta$; these constraints take the form

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

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 in the optimal temporally nondiscriminatory mechanism. Attach to each such constraint the dual variable $\lambda^i(v, \delta, \delta', \tau)$.[33]

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.[34] For buyers i and j (possibly equal to i) value $v \in \mathcal{V}$, and arrival time τ, define $M_\tau^{ij}(v)$ by

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

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, ignoring heterogeneity in discount factors. The feasibility constraint that limits allocation to buyers with higher virtual values is, for all $v \in \mathcal{V}$ and all arrival times τ,

$$\sum_j \sum_{(v', \tau') \in M_\tau^{ij}(v)} \sum_{\delta \in \mathcal{D}^j} \sum_{t \in T} q_t^j \left(v', \delta, \tau'\right) g_j f^j\left(v', \delta\right) \leq \Pr\left(\exists \left(j, v', \tau'\right) \text{ s.t. } \left(v', \tau'\right) \in M_\tau^{ij}(v)\right). \quad (9)$$

The left-hand side of inequality (9) 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 v and arrival time τ, we attach the multiplier $\rho^i(v, \tau)$ to the feasibility constraint (9); the sum of these multipliers is $R^i(v, \tau) = \sum_{\tau' \geq \tau} \sum_{v'<v} \rho^i(v', \tau')$, which represents the aggregation of

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

[34] See Section 3.1 in the main text for a discussion of this intuition.

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

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

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,τ) the following condition is satisfied:

$$c_{\tau+t}^i\left(v,\delta,\tau\right) = 0 \;\forall i,t,v,\delta,\tau,\text{ and } q_\tau^i\left(v,\delta,\tau\right) > 0 \text{ implies } \gamma_\tau^i\left(v,\delta,\tau\right) = 0. \tag{LP}$$

We first show that feasible multipliers at time τ imply feasible multipliers for all subsequent time periods $\tau+t>\tau$. The coefficients $c_{\tau+t}^i$ can be written as

$$\begin{aligned} c_{\tau+t}^i\left(v,\delta,\tau\right) = &\,\delta^t c_\tau^i\left(v,\delta,\tau\right) + \gamma_{\tau+t}^i\left(v,\delta,\tau\right) \\ &- \left[\left(1-\delta^t\right) R^i\left(v,\tau\right) f^i\left(v,\delta\right) + \delta^t \gamma_\tau^i\left(v,\delta,\tau\right) + \sum_{\delta'>\delta} \lambda^i\left(v,\delta',\delta\right)\left(\delta'^t - \delta^t\right) v\right]. \end{aligned}$$

The dual variables R^i, λ^i, and γ_τ^i are all weakly positive. Therefore, $c_\tau^i(v,\delta,\tau)=0$ implies there is always a nonnegative value for $\gamma_{\tau+t}^i(v,\delta,\tau)$ so that $c_{\tau+t}^i(v,\delta,\tau)=0$. Consequently, we only need to consider the time-τ terms in (LP).

Under the optimal temporally-nondiscrimatory mechanism, a buyer who arrives in period τ receives the good in period τ if she has the highest virtual value above the threshold $v_i^\star$, implying that $q_\tau^i(v,\delta,\tau)>0$ when $v>v_i^\star$ (Board and Skrzypacz, 2016). We therefore require that $\gamma_\tau^i(v,\delta,\tau)=0$ for $v>v_i^\star$. Since in the optimal temporally-nondiscriminatory mechanism the feasibility constraint will be binding for value-types with strictly positive virtual value, we set $R^i(v,\tau)=\delta^\tau m_+^i(v)$. Note that since $R^i(v,\tau)=\sum_{\tau'\geq\tau}\sum_{v'<v}\rho^i(v',\tau')$, this is equivalent to setting $\sum_{v'<v}\rho^i(v',\tau)=(\delta^\tau-\delta^{\tau+1})m_+^i(v)$.

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 relationship implies a system of equations in the λ

variables; defining $\mu^i(v,\delta) \equiv \big(m^i(v|\delta) - m^i_+(v)\big)g_i f(v,\delta)$ with $\mu^i(0,\delta)=0$, the system is

$$\sum_{\delta'\geq\delta}\lambda^i\left(v,\delta',\delta,\tau\right)\varepsilon - \sum_{\delta\geq\delta'}\lambda^i\left(v,\delta,\delta',\tau\right)\varepsilon - \gamma^i_\tau\left(v-\varepsilon,\delta,\tau\right) + \gamma^i_\tau\left(v,\delta,\tau\right) = \mu^i\left(v-\varepsilon,\delta\right) - \mu^i\left(v,\delta\right). \tag{10}$$

To apply Theorem 3, we represent this system as a network in which each type (v,δ) is associated with a node in $\mathcal{N}$, each $\lambda^i(v,\delta,\delta',\tau)$ is associated with a (nonnegative) flow from (v,δ,τ) to (v,δ',τ), and each $\gamma^i_\tau(v-\varepsilon,\delta,\tau)$ is associated with a (nonnegative) flow from (v,δ,τ) to $(v-\varepsilon,\delta,\tau)$. Because we have assigned a value to the feasbility constraint, we are concerned only with constraints which are relevant across a buyer's types and not between buyers, and we define a separate network for each buyer i and arrival time τ.[35] First, write

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

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^i so that $b^i(v,\delta) = \mu^i(v-\varepsilon,\delta) - \mu^i(v,\delta)$. We apply Theorem 3 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^i and k defined above, there are two cases in which inequality (8) is slack, because the left-hand side is infinite:

- 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., $m^i(v) \leq m^i(v^\star_i)$).

Therefore, to consider satisfaction of (8) 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 sets*. Theorem 3 implies that there is a solution for the multipliers λ^i and γ^i if and only if,

[35] Since arrival times are publicly observable, they do not explicitly enter into the buyer's incentive constraints. Thus even though there are an infinite number of dual variables λ^i, the network associated with arrival time τ is finite.

for any limited upper set $\bar{X}$,

$$0 \geq \sum_{(v,\delta)\in\bar{X}} b^i(v,\delta) = \sum_{(v,\delta)\in\bar{X}} \mu^i(v-\varepsilon,\delta) - \mu^i(v,\delta). \tag{11}$$

Since inequality (11) 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 $v' \in \{v,\dots,v_i^\star(v)\}$ and $\delta' \geq \delta$. Thus a necessary condition for inequality (11) is

$$\sum_{\delta'\geq\delta} \left\{\mu^i(v^\star(v),\delta') - \mu^i(v-\varepsilon,\delta)\right\} \geq 0, \tag{12}$$

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 $m^i(v) \leq m^i(v_i^\star)$ and $(v,\delta) \in \bar{X}_+$ if $m^i(v) > m^i(v_i^\star)$. Let $D(\tilde{X}) = \{\delta \in \mathcal{D}^i \colon \exists v \in \mathcal{V} \text{ s.t. } (v,\delta) \in \tilde{X}\}$ 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

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

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 (12) implies that the above expression is weakly positive, and hence inequality (12) is necessary and sufficient for condition (11).

Finally, condition (12) 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.[36] It follows that when the latter is never optimal, neither is the

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

former.

□

Proof of Corollary 1. By definition, $m^i(v_i^\star(v)) > 0$. Then since $m_+^i(v) = \max\{m^i(v), m^i(v_i^\star)\} \geq m^i(v)$, for any $v \in \mathcal{V}$ we have Let $v \in \mathcal{V}$ be such that $m^i(v) \geq 0$. We calculate

$$\begin{aligned}
&\mu^i\left(v_i^\star(v), \delta\right) - \mu^i\left(v-\varepsilon, \delta\right) \\
&\quad \geq \left(m^i\left(v_i^\star(v)|\delta\right) - m^i\left(v_i^\star(v)\right)\right) f^i\left(v_i^\star(v), \delta\right) - \left(m^i\left(v-\varepsilon|\delta\right) - m^i\left(v-\varepsilon\right)\right) f^i\left(v-\varepsilon, \delta\right) \\
&\quad = \left(\frac{1-F^i\left(v_i^\star(v)\right)}{f^i\left(v_i^\star(v)\right)} - \frac{1-F^i\left(v_i^\star(v)|\delta\right)}{f^i\left(v_i^\star(v)|\delta\right)}\right) f^i\left(v_i^\star(v), \delta\right) \\
&\qquad\qquad - \left(\frac{1-F^i\left(v-\varepsilon\right)}{f^i\left(v-\varepsilon\right)} - \frac{1-F^i\left(v-\varepsilon|\delta\right)}{f^i\left(v-\varepsilon|\delta\right)}\right) f^i\left(v-\varepsilon, \delta\right) \\
&\quad = \left(1-F^i\left(v_i^\star(v)\right)\right)\left(f^i\left(\delta|v_i^\star(v)\right) - f^i\left(\delta|v-\varepsilon\right)\right) \\
&\qquad + \sum_{v-\varepsilon \leq v' < v_i^\star(v)} \left(f^i\left(\delta|v'\right) - f^i\left(\delta|v-\varepsilon\right)\right) f^i\left(v'\right). \qquad (13)
\end{aligned}$$

Summing (13) over all $\delta' > \delta$ gives

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

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

B Calculations for Example 1

Let discount factors be $\delta \in \{0, \hat{\delta}\}$ 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 $\Pr(\delta = 0) = \Pr(\delta = \hat{\delta}) = 1/2$.

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.

Straightforward calculation gives

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

$$m(v) = \begin{cases} 1 & \text{if } v = 1, \\ \frac{1-3\pi}{4-3\pi} & \text{if } v = \frac{1}{2}. \end{cases}$$

In turn,

$$m(1|\delta) - m_+(1) = \begin{cases} 0 & \text{if } \delta = 0, \\ 0 & \text{if } \delta = \hat{\delta}; \end{cases}$$

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

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

Finally,

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

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

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

Remark 3. *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 2 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,

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

Then when $\hat{\delta} < \frac{1}{2}$, Theorem 2 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

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

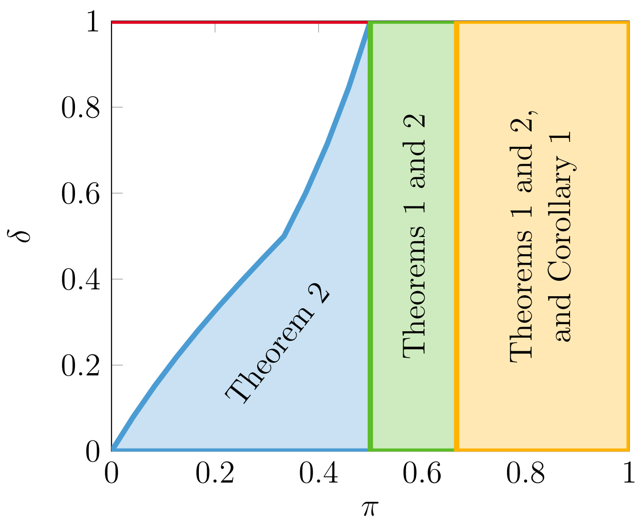

Figure 3: Temporal nondiscrimination is optimal in all shaded regions. In the rightmost (orange) region, value and patience are positively correlated, and Corollary 1 applies. In the middle (green and orange) region, the optimal monopoly price is independent of patience. In the full region, optimality of nondiscrimination follows from Theorem 2. Perfect separation of buyers into distinct time periods is weakly optimal when $\hat{\delta} = 1$ (red line). In the remaining unshaded region, patient bidders receive some allocation at time $t = \tau$ and some at time $t' > \tau$.

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 potentially to a low-value buyer in the next period if no high-value buyer arrives.

B.1 Perfect separation

Our main results consider the optimality of temporal nondiscrimination. We now show, in this example, that perfect temporal separation — where one discount type receives an allocation in one period, and the other receives an allocation in another — is generically nonoptimal.[37] Because we cannot apply our main results, we respecify the seller's problem:

$$
\begin{aligned}
&\max_{q,t} \sum_{v} t\left(v,\delta\right) f\left(v,\delta\right), && \\
&\text{s.t. } v\delta \cdot q\left(v,\delta\right) - t\left(v,\delta\right) \geq 0 && \mu \qquad \text{(IR)} \\
&\sum_{\delta}\sum_{v' \geq v}\sum_{t} q_t\left(v,\delta\right) \leq B\left(v\right) && \beta \qquad \text{(Feas.)} \\
&q_t\left(v,\delta\right) \geq 0 && \gamma \qquad \text{(Feas.)} \\
&v\delta \cdot q\left(v',\delta'\right) - t\left(v',\delta'\right) \leq v\delta \cdot q\left(v,\delta\right) - t\left(v,\delta\right). && \lambda \qquad \text{(IC)}
\end{aligned}
$$

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

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

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 consumption in period $t = 0$ at price $p_0^\star = 1$ and consumption in period $t = 1$ at price $p_1^\star = 1/2$. In this optimal mechanism, it follows that

[37]For similar claims regarding the suboptimality of degenerate mechanisms, see Pycia (2006) and Fang and Norman (2006), among others.

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 $\delta=1$.

C Proofs for Section 4 (Generic misreports)

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[38]

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

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

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

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 2. We build on the preparatory work done in the proof of Theorem 1. Again, we temporarily assume that the seller must allocate within the first T periods after a given buyer arrives, and subsequently remove this assumption. 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

$$
c^i_{\tau+t}(v,\delta,\tau) = \left(\delta^t m^i(v|\delta) - R^i(v,\tau)\right) f^i(v,\delta) + \sum_{\substack{\delta'\neq\delta \\ v'>v}} \lambda^i(v',\delta,\delta',\tau)\,\delta^t\varepsilon - \sum_{\substack{\delta'\neq\delta \\ v'>v}} \lambda^i(v',\delta',\delta,\tau)\,\delta'^t\varepsilon.
$$

As in the proof of Theorem 1, we look for multipliers so that $c^i_{\tau+t}(v,\delta,\tau) = 0$ implies that $t = 0$ and $m^i(v) > m^i(v_i^\star)$. It is sufficient to solve the following system:

$$
\sum_{\delta'\neq\delta}\left[\lambda^i(v,\delta,\delta',\tau) - \lambda^i(v,\delta',\delta,\tau)\right]\varepsilon + \gamma^i(v,\delta,\tau) = \mu^i(v,\delta) - \mu^i(v-\varepsilon,\delta) \tag{14}
$$

$$
\sum_{\delta'\neq\delta}\lambda^i(v,\delta',\delta,\tau)\left(\delta^t - \delta'^t\right)v + \gamma^i(v,\delta) = \left(1-\delta^t\right)m^i_+(v)\,f(v,\delta) \quad \forall t>0. \tag{15}
$$

Applying Farkas' Lemma, there exists a nonnegative solution (λ,γ) to the system (14)–(15)

[38] We say $a \overset{\text{sign}}{=} b$ if $a, b \neq 0$ implies $ab > 0$.

if and only if there are no $y^i(v,\delta)$ and $z_t^i(v,\delta)$ satisfying[39]

$$\left[y^i\left(v,\delta\right)-y^i\left(v,\delta'\right)\right]\varepsilon\geq\sum_{t>0}z_t^i\left(v,\delta'\right)\left(\delta_t-\delta'_t\right)v \qquad \forall\delta\neq\delta' \tag{16}$$

$$y^i\left(v,\delta\right)\geq 0 \qquad \forall\left(v,\delta\right)\in\Theta^i \text{ s.t. } m\left(v\right)\leq 0 \tag{17}$$

$$z_t^i\left(v,\delta\right)\geq 0 \qquad \forall\left(v,\delta\right)\in\Theta^i,\forall t>0, \tag{18}$$

and

$$\sum_{(v,\delta)\in\Theta^i}\left[y^i\left(v,\delta\right)\left(\mu\left(v,\delta\right)-\mu\left(v-\varepsilon,\delta\right)\right)+m_+^i\left(v\right)f^i\left(v,\delta\right)\sum_{t>0}z_t^i\left(v,\delta\right)\left(1-\delta_t\right)\right]<0. \tag{19}$$

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:

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

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:

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

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

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

[39] Since we are concerned with feasibility holding fixed the observable arrival time τ, we henceforth simplify notation by dropping τ from all function arguments.

Substituting these values into inequality (19) gives

$$\begin{aligned}
&\sum_{(v,\delta)\in\Theta^i}\left[\Delta\mu_k y_k + m f_k \sum_{t>0}\left(1-\hat{\delta}_k^t\right) z_{tk}\right] \\
&= \sum_{(v,\delta)\in\Theta^i}\left[\Delta\mu_k \sum_{k'\le k} x_{k'} + \left(\frac{x_k \varepsilon}{\left(\hat{\delta}_k-\hat{\delta}_{k+1}\right) v}\right)\left(1-\hat{\delta}_k\right) m f_k\right] \\
&= \sum_{(v,\delta)\in\Theta^i}\left[\sum_{k'\ge k}\Delta\mu_{k'} + \left(\frac{1-\hat{\delta}_k}{\hat{\delta}_k-\hat{\delta}_{k+1}}\right)\frac{\varepsilon}{v} m f_k\right] x_k. \qquad (20)
\end{aligned}$$

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

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

Substituting into (20) gives

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

Then when $M_{\bar{k}}(v) \ge 0$ for all v and $\bar{k}$, there are values for y_k and z_{tk} so that inequalities (16)–(18) are satisfied but inequality (19) is not. It follows that the when $M_{\bar{k}}(v) \ge 0$ 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 (14)–(15). As in the proof of Theorem 1, when these conditions are satisfied the seller will not discriminate on temporal preference even when she may allocate to a buyer at any time after arrival. □

D Proofs for Section 5 (Ambiguous temporal preferences)

Proof of Proposition 1. Note that the optimal temporally-nondiscriminatory mechansismis 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 $F \in \mathcal{F}$,

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

Then the seller's revenue under any mechanism (q, p) is bounded above by what would be obtained if the true distribution of values was F. When F satisfies Theorem 1 temporal nondiscrimination is strictly optimal, in the sense that any mechanism which alters the allocation strictly reduces the seller's revenue. □