Intertemporal Allocation with Unknown Discounting

Justin Burkett* and Kyle Woodward†

August 2024

Abstract

We consider the problem faced by a durable good monopolist who can allocate a single good at any time, but is uncertain of a buyer's values and temporal preferences for receiving the good. We derive conditions under which it is optimal for the monopolist to ignore the uncertainty about the buyer's discount factor and allocate immediately via a single first-period price. Under one condition, the seller optimally offers a single first-period price if she would weakly raise this price upon learning that the buyer cannot be too impatient (Corollary 2). A related condition states that the single first-period price is optimal if buyer types with higher discount factors have stochastically higher values (Corollary 3). These conditions also apply when sellers face ambiguity regarding the buyer's discount factor. Our results provide a novel justification for ignoring heterogeneous discount factors when the seller is incompletely informed about buyer's temporal preferences.

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

†Gopuff; kyle.woodward@gopuff.com

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

1 Introduction

Many economic models of dynamic pricing assume that buyers have common preferences for future payoffs.[1] Unsurprisingly, empirical studies find that discount rates vary across the population (Mischel et al., 1989; Kirby and Maraković, 1995; Green and Myerson, 2004; Hakimi, 2013; Chan, 2017) and even vary across commodities for a given individual (Ubfal, 2016). This variation allows for the possibility that the seller might profit from screening on temporal preferences and hence the possibility that standard approaches to dynamic pricing may leave some rents on the table if heterogeneity in temporal preferences is ignored. In this paper we study conditions under which screening on time preferences by delaying allocation for some types is feasible but *unprofitable* for a seller of a durable good. Our results show when optimal mechanisms under the assumption of homogeneity in time preferences remain optimal with the introduction of heterogeneity.

Specifically, we consider the problem faced by a durable-good monopolist who understands that the buyer might discount his future value for the good at one of many rates. We explore the possibility that this seller nonetheless optimally implements a mechanism that does not screen on time preference and allocates immediately. We start from the simple observation that even when the potential buyer has many possible discount factors, it remains feasible for the seller to ignore this uncertainty and allocate immediately. We then take a candidate immediate allocation mechanism and apply tools from the theory of linear programming to obtain conditions under which this candidate mechanism is optimal when the buyer has one of many, privately-known discount factors.

The approach generates several sets of conditions under which immediate allocation is optimal. In our model, the optimal immediate allocation mechanism simply offers the good at a single price in the first period. This turns out to be optimal if the seller would want to raise the price upon learning that the buyer is not too impatient (Corollary 2), or if buyer types with higher discount factors have stochastically higher values (Corollary 3).

Before further discussing our results we lay out our economic model. In our model, a seller with a single unit of a durable good faces a single buyer with a finite time horizon for receiving the good. The buyer is privately informed of both their (initial) valuation and their

[1]Notable exceptions, in which buyers do not share a commonly-known discount rate, include Pai and Vohra (2013) and Mierendorff (2016), in which buyers have identical discount factors but heterogeneous deadlines for consumption. A separate thread of literature in public finance considers the effect of temporal preference heterogeneity on optimal tax policy; see, e.g., Diamond and Spinnewijn (2011), Farhi and Werning (2013), and Golosov et al. (2013).

discounted future valuations.[2] We assume that the seller can commit to a sales mechanism ex ante, and that the buyer's type is drawn 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.

In the analysis of this problem, determining which types will want to mimic which other types (i.e., which incentive constraints bind), and hence which allocations need to be adjusted in response to a change in 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 set of binding constraints.[3] We start from the observation that the candidate mechanism, which ignores heterogeneity in discount factors, is feasible and respects incentive constraints. To prove that it is optimal we then need only identify a subset of incentive constraints under which it is optimal. Using standard results from linear programming, it is sufficient to find multipliers (dual variables) for the considered constraints under which the appropriate Karush-Kuhn-Tucker (KKT) 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, Theorem 1, shows that immediate allocation is optimal when a condition on the buyer's conditional virtual value is satisfied. 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). We define the buyer's *conditional* virtual value, as the virtual value of a buyer with a known discount factor. We derive Theorem 1 by showing that if we only include downward incentive constraints, the KKT conditions yield a system that is equivalent to a standard problem on the existence of a feasible flow in a network.[4] Once the analogy is established, the argument we use is an immediate consequence of Gale's feasible flow theorem (Gale, 1957). Since the condition in Theorem 1 can be difficult to interpret, we show that this condition is implied by simpler conditions through a sequence of

[2]We constrain attention to buyers who discount future receipt of the good at an exponential rate (Samuelson, 1937). A working version of this paper provides results for buyers who are potentially non-exponential discounters.

[3]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 and non-convex, we also cannot rule out a priori that non-local incentive constraints bind at the optimal solution (Carroll, 2012). Our approach does not rely on eliminating either of these possibilities.

[4]The downward constraints in our model are the ones that prevent types with higher discount factors from mimicking those with lower discount factors and types with higher values mimicking those with lower values.

corollaries. In Corollary 2, we show that under a monotonicity restriction on virtual values, the conditions in Theorem 1 hold if the seller would weakly increase the firs-period price if she were to learn the buyer is not too impatient. More specifically, upon learning that the buyer's discount factor exceeds any threshold the seller would optimally (weakly) raise the first-period price. In Corollary 3, we show that if the distribution of buyer values is stochastically nondecreasing in the buyer's discount factor, immediate allocation is optimal. Roughly, more patient buyers have stochastically higher values for immediate consumption. When this is true, buyers with higher discount factors receive higher information rents. This is sufficient to prevent the seller from optimally allocating to patient buyers at a lower price.

An immediate consequence of our analysis is that a seller who is uncertain of the statistical relationship between value and patience should optimally allocate immediately, so long it is plausible that our main conditions are potentially satisfied.[5] In this sense the ambiguity surrounding heterogeneity in individual discount factor (see our discussion above) is self-supporting: sellers with little knowledge of the joint distribution of value and patience may optimally not discriminate on temporal preference, and their sales will contain no information about the joint distribution of value and patience. Our results are therefore consistent with a lack of temporal screening by sellers of durable goods.

Immediate allocation is optimal when any one of our statistical conditions is satisfied.[6] Our results therefore provide a microfoundation for the workhorse assumption of a commonly-known discount factor. A corollary to our results is that if an optimal mechanism does screen on discount factor via delayed allocation, our conditions must not hold. Since the statistical relationship between value and discount factor varies across individuals and goods, our results are consistent with the observation that some markets employ temporal discrimination, while many others do not.[7]

Finally, 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 complex. Nonetheless, the presence of temporal incentive constraints drives allocation away from utilization of this

[5]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 analysis, the seller does not even need to know the marginal distribution of buyers' discount rates.

[6]While it does not discriminate on temporal preferences, the optimal mechanism does discriminate on arrival time. In particular, the optimal mechanism sells to a given buyer either never or immediately upon arrival.

[7]Ubfal (2016) shows that discount rates may differ across goods for a given individual. Thus our conditions may be satisfied for some commodities and not for others.

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 results. Section 4 considers the optimal mechanism when the seller faces ambiguity regarding the distribution of temporal preferences. A discussion of the related literature is deferred to Section 5.

2 Model

We consider a model in which there is a single buyer present in the first period with a finite time horizon. As we showed in a previous version of this paper, the results of this analysis can be extended to cases with multiple buyers arriving stochastically over time with infinite time horizons.

A seller offers one unit of an indivisible good for sale to a single buyer. Time is discrete, $t \in \{0, 1, \ldots, T\}$, and allocation may take place in any period up to period T. The seller commits to a mechanism in the first period, $t = 0$. The buyer's type is two-dimensional, consisting of a value v and discount factor δ, $(v, \delta) \in \mathcal{V} \times \mathcal{D} \subset [0, 1]^2$. Buyer types thus differ in terms of the value they would receive from receiving the good immediately and in terms of the rate at which this value depreciates over time. We assume that no players discount monetary future monetary transfers.[8]

Each buyer's utility is quasilinear in expected transfers, and if her allocation and payments are $q = (q_t)_{t=0}^{T}$ and p respectively[9] her interim utility upon arrival is

$$u\left(q, p\right| v, \delta) \equiv \sum_{t=0}^{T} \delta^t q_t v - p.$$

We assume that the support of types $\mathcal{V} \times \mathcal{D}$ is finite, and for simplicity we assume further that there is $\varepsilon > 0$ so that $\mathcal{V} \equiv \{0, \varepsilon, \ldots, 1 - \varepsilon, 1\}$. The buyer's type space is $\Theta \equiv \mathcal{V} \times \mathcal{D}$.

[8] This distinction between discounting future consumption and discounting future monetary transfers is also made in Board and Skrzypacz (2016), who assume in their main specification that buyers discount their future consumption value but not future transfer amounts. If we were to assume, for example, that the buyer discounts future payments but the seller does not, the seller could trivially increase her expected revenue by waiting to collect payment from the buyer in the last period. Since the buyer discounts this payment, the buyer would perceive the good to be cheaper and would be willing to pay more.

[9] Given that neither the seller nor the buyer discount monetary payments, it is without loss to consider a single aggregated payment made in the first period.

To distinguish random variables, we add a tilde, making $\tilde{\theta} = (\tilde{v}, \tilde{\delta})$ the random variable corresponding to buyer's type.

We use $f(v, \delta)$ for the (commonly known) probability that the buyer has type (v, δ) and assume $f(v, \delta) > 0$ for all $(v, \delta) \in \mathcal{V} \times \mathcal{D}$. Let $f(v) \equiv \sum_{\delta \in \mathcal{D}} f(v, \delta)$ so that $F(v) \equiv \sum_{v' \leq v} f(v')$ is the cumulative marginal distribution of valuation types for buyer i. Similarly, let $f(\delta) \equiv \sum_{v \in \mathcal{V}} f(v, \delta)$ and $f(v|\delta) \equiv f(v, \delta)/f(\delta)$ so that $F(v|\delta) \equiv \sum_{v' \leq v} f(v'|\delta)$ is the cumulative marginal distribution of valuation types for buyer i, conditional on her having discount type δ. Define $f(\delta)$, $f(\delta|v)$, $F(\delta)$, and $F(\delta|v)$ analogously. We use $\mathbb{E}_\Theta$ for the expectation taken with respect to buyer's type.

We define the *average marginal revenue* of a buyer as

$$m(v) \equiv v - \frac{1 - F(v)}{f(v)} \varepsilon,$$

and the *conditional marginal revenue* as

$$m(v|\delta) \equiv v - \frac{1 - F(v|\delta)}{f(v|\delta)} \varepsilon.$$

Note that $m(v) = \mathbb{E}_\mathcal{D} \left[m(v|\tilde{\delta})|v \right]$. Define $v^\star \equiv \text{argmax}_v \sum_{w=v}^{1} m(w)f(w) = \text{argmax}_v (1 - F(v - \varepsilon))v$. This is the revenue maximizing cutoff when discount types are ignored. Note that if $m(\cdot)$ is increasing, this cutoff is simply the lowest v for which $m(v) \geq 0$. When $m(\cdot)$ is not increasing for all v, it may be that $m(v) < 0$ for some $v > v^\star$, but it remains true that $m(v^\star - \varepsilon) \leq 0 \leq m(v^\star)$.

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 buyer's reported type determines the probability of receiving the good in each period as well as expected payments to be made to the seller. We let $q_t(v, \delta)$ denote the (interim) probability that the buyer receives the good in period t having reported the type (v, δ). We use $q(v, \delta)$ to indicate the vector of probabilities across time periods. If the buyer has type (v, δ), and reports the type (v', δ'), her expected payoff from the mechanism is therefore

$$u(v', \delta'|v, \delta) \equiv \sum_{t=0}^{T} \delta^t q_t(v', \delta') v - p(v', \delta').$$

We use $u(v, \delta) \equiv u(v, \delta|v, \delta)$ for the equilibrium payoff of the type (v, δ) bidder.

Definition 1. *The allocation rule q* allocates immediately *if $q_t(v, \delta) = 0$ when $t > 0$ for all $(v, \delta) \in \Theta$. The mechanism (q, p) is* insensitive to discount type *if $q(v, \cdot)$ and $p(v, \cdot)$ are constant for all $v \in \mathcal{V}$.*

Our main results are focused on the first property (q being immediate). The following proposition connects immediacy to the sensitivity property.

Proposition 1. *Given an incentive compatible mechanism (q, p), if q allocates immediately, the mechanism is insensitive to discount type.*

Proof. If q allocates immediately, $u(v, \delta) \geq u(v, \delta'|v, \delta)$ and $u(v, \delta') \geq u(v, \delta|v, \delta')$ together imply $p(v, \delta) = p(v, \delta')$ and $q(v, \delta) = q(v, \delta')$ for all $v \in \mathcal{V}$ and $\delta, \delta' \in \mathcal{D}$. □

2.2 The seller's problem

Our analysis considers when it is optimal for the seller to allocate immediately. The general revenue maximization problem is

$$\begin{aligned}
\max_{(q,p)\in[0,1]^{T+1}\times\mathbb{R}} \quad & \mathbb{E}_\Theta \left[p\left(\tilde{v}, \tilde{\delta}\right)\right] & & \text{(GP)} \\
\text{s.t.} \quad & u(v, \delta) \geq u(v', \delta'|v, \delta) & \forall (v, \delta), (v', \delta') \in \Theta & \quad \text{(IC)} \\
& u(v, \delta) \geq 0 & \forall (v, \delta) \in \Theta & \quad \text{(IR)} \\
& \sum_{t=0}^{T} q_t(v, \delta) \leq 1 & \forall (v, \delta) \in \Theta. & \quad \text{(F)}
\end{aligned}$$

The seller's problem considers revenue maximization subject to incentive compatibility and feasibility. The immediate revenue maximization problem artificially sets $q_t(v, \delta) = 0$ whenever $t > 0$.

$$\begin{aligned}
\max_{(q,p)\in[0,1]^{T+1}\times\mathbb{R}} \quad & \mathbb{E}_\Theta \left[p\left(\tilde{v}, \tilde{\delta}\right)\right] & & \text{(IP)} \\
\text{s.t.} \quad & u(v, \delta) \geq u(v', \delta'|v, \delta) & \forall (v, \delta), (v', \delta') \in \Theta & \quad \text{(IC)} \\
& u(v, \delta) \geq 0 & \forall (v, \delta) \in \Theta & \quad \text{(IR)} \\
& q_t(v, \delta) = 0 & \forall t > 0, (v, \delta) \in \Theta & \quad \text{(F)}
\end{aligned}$$

In light of Proposition 1, there is a straightforward solution to (IP). Using $\mathbb{I}$ for an indicator function, it is optimal to set

$$q_t^I(v,\delta) = \mathbb{I}\{v \geq v^\star\}\mathbb{I}\{t = 0\}$$
$$p^I(v,\delta) = \mathbb{I}\{v \geq v^\star\}v^\star.$$

Note that this would also be a solution if either the discount type δ were observable or if it is statistically independent of v.[10]

3 Analysis

The results of this paper provide sufficient conditions under which it is optimal for the seller to allocate immediately. That is, we provide sufficient conditions under which the solution to (IP) is also a solution to (GP). The essential problem in this and related multidimensional mechanism design problems is that there are many constraints imposed by incentive compatibility and it is generally difficult to determine a priori which of these constraints bind at the optimum. Instead of asking for a complete characterization of the optimal mechanism, we study the more tractable problem of determining *sufficient* conditions for the optimality of a particular mechanism.

To determine sufficient conditions for the optimality of (q^I, p^I), we relax (GP) using two strategies. The first is based on the simple observation that (q^I, p^I) is clearly feasible in (GP). As a consequence, if we drop constraints from (GP) and find that after dropping constraints (q^I, p^I) is optimal, then (q^I, p^I) must have been optimal in (GP) as well.[11]

The second strategy for relaxing (GP) uses the constraints that are known to bind in the restricted problem (IP) to construct transfers for the relaxed version of (GP). These transfers, which are introduced formally in Lemma 1, are the transfers that would result if it were known that all incentive constraints associated with the buyer reporting the next lowest value were binding (i.e., if $u(v,\delta) = u(v-\varepsilon,\delta|v,\delta), \forall(v,\delta) \in \Theta$). These transfers have two important properties for our purposes. They generate an expected revenue that is an upper bound for the expected revenue under (GP), and they are optimal for any mechanism which, like q^I, allocates immediately.

[10] When δ is observable, the seller faces a Coasian bargaining problem for each discount type, and optimally offers a monopoly price in the first period for each type δ. When δ is independent of v, this monopoly price does not depend on δ, and the observable δ mechanism is incentive compatible when δ cannot be observed. We thank a referee for this observation.

[11] The ignored constraints are satisfied implicitly, and optimality in the relaxed problem implies that the mechanism cannot be improved in (GP).

Lemma 1. *For a given q, the transfers*

$$P(v, \delta|q) \equiv \sum_{t=0}^{T} \delta^t \left\{q_t(v, \delta)v - Q_t(v, \delta)\varepsilon\right\},$$

where $Q_t(v,\delta) \equiv \sum_{w<v} q_t(w,\delta)$, generate an expected revenue that is at least as large as the feasible expected revenue from q. If q allocates immediately, these transfers are optimal.

If we impose these transfers for an arbitrary q we thus get an over estimate of the feasible expected revenue from q. If after using these transfers, it is still true that q^I is optimal, it must be that q^I generates more expected revenue than all feasible transfer schemes.

Imposing the transfers from Lemma 1 into the seller's problem has two effects. First, with these transfers any incentive constraint on the misreport of value alone (i.e., incentive constraints of the form $u(v,\delta) \geq u(v',\delta|v,\delta))$ is satisfied implicitly when $\sum_{t=0}^{T} \delta^t q_t(v,\delta)$ is nondecreasing in v. Second, we can simplify the objective by writing it in terms of a conditional marginal revenue function. To derive this function, take expectation of the transfers across all v for a given δ and "integrate" by parts to get

$$\mathbb{E}_{\mathcal{V}}\left[P(\tilde{v}, \delta|q)\right] = \sum_{t=0}^{T} \delta^t \mathbb{E}_{\mathcal{V}}\left[\left(\tilde{v} - \frac{1 - F(\tilde{v}|\delta)}{f(\tilde{v}|\delta)}\right) q_t(\tilde{v}, \delta)\right] = \sum_{t=0}^{T} \delta^t \mathbb{E}_{\mathcal{V}}\left[m(\tilde{v}|\delta) q_t(\tilde{v}, \delta)\right].$$

The conditional marginal revenue, $m(v|\delta)$, is defined by the expression in parentheses. The average marginal revenue, $m(v)$, is related to the conditional marginal revenue through

$$m(v) = \mathbb{E}_{\mathcal{D}}\left[m(v|\tilde{\delta})|v\right].$$

Imposing these transfers, the seller's objective becomes

$$\mathbb{E}_{\Theta}\left[\sum_{t=0}^{T} \tilde{\delta}^t m(\tilde{v}|\tilde{\delta}) q_t(\tilde{v}, \tilde{\delta})\right],$$

and the buyer's payoff can be written as

$$u(v', \delta'|v, \delta) = \sum_{t=0}^{T} \left\{\delta'^t Q_t(v', \delta')\varepsilon + q_t(v', \delta')\left(\delta^t v - \delta'^t v'\right)\right\}$$

To summarize our approach, we study problems that result from taking the general problem (GP), imposing the transfers from Lemma 1 and dropping a subset of incentive constraints that includes the misreport of value constraints. The subsections below differ according to the set of constraints that are dropped. In general, if we drop more constraints, we expect to derive weaker sufficient conditions for the optimality of q^I, since the missing constraints might help to support the optimality of q^I.

Since the misreport of value constraints are accounted for by the transfer scheme we use, there are two remaining categories of constraints to consider. Our analysis below focuses on the misreport of discount constraints. These are constraints of the form

$$u(v,\delta) \geq u(v,\delta'|v,\delta).$$

We classify these as being downward if $\delta > \delta'$ or upward if the reverse is true. Throughout the paper we ignore constraints involving a joint misreport of value and discount (i.e., $u(v,\delta) \geq u(v',\delta'|v,\delta)$). While many of these constraints are implied by combining misreport of value and misreport of discount factor constraints, ignoring them is not without loss of generality.

3.1 Downward misreports of discount type

In this section, we drop all but the downward misreport of discount type constraints and find sufficient conditions for the optimality of q^I in the relaxed problem. Formally, the problem we consider is

$$\max_{q\in[0,1]^{T+1}} \quad \mathbb{E}_\Theta\left[\sum_{t=0}^{T} \delta^t m(\tilde{v}|\tilde{\delta}) q_t(\tilde{v},\tilde{\delta})\right] \tag{DCP}$$

$$\text{s.t.} \quad \sum_{t=0}^{T} \delta^t Q_t(v,\delta)\varepsilon \geq \sum_{t=0}^{T}\left\{\delta'^t Q_t(v,\delta')\varepsilon + q_t(v,\delta')\left(\delta^t - \delta'^t\right)v\right\} \quad \forall v, \forall \delta > \delta' \tag{DIC}$$

$$\sum_{t=0}^{T} q_t(v,\delta) \leq 1 \qquad \forall\,(v,\delta)\in\Theta. \tag{F}$$

By Lemma 1 and the preceding discussion, if q^I solves (DCP) then it also solves (GP). Before presenting our first theorem which applies to (DCP), we study the special case in which there are only two discount types δ_H and δ_L with $\delta_H > \delta_L$. In this case, there is a single (DIC) constraint for each value type,

$$\sum_{t=0}^{T} \delta_H^t Q_t(v,\delta_H)\varepsilon \geq \sum_{t=0}^{T}\left\{\delta_L^t Q_t(v,\delta_L)\varepsilon + q_t(v,\delta_L)\left(\delta_H^t - \delta_L^t\right)v\right\}.$$

Since (DCP) is linear, q^I is optimal if there exist Lagrange multipliers for the constraints satisfying the appropriate KKT conditions. In particular, optimality requires that there exist

$\lambda(v) \geq 0$, $\overline{\gamma}(v,\delta) \geq 0$ and $\underline{\gamma}_t(v,\delta) \geq 0$ on (DIC), (F) and $q_t(v,\delta) \geq 0$ respectively satisfying

$$m(v|\delta_H)f(v,\delta_H) + \sum_{w>v} \lambda(w)\varepsilon - \overline{\gamma}(v,\delta_H) + \underline{\gamma}_0(v,\delta_H) = 0 \tag{1}$$

$$m(v|\delta_L)f(v,\delta_L) - \sum_{w>v} \lambda(w)\varepsilon - \overline{\gamma}(v,\delta_L) + \underline{\gamma}_0(v,\delta_L) = 0. \tag{2}$$

These are first-order conditions for $q_0(v,\delta_H)$ and $q_0(v,\delta_L)$ respectively. For (DCP), the optimality conditions for q_t with $t > 0$ are implied by these. Using (1), we derive the implication that for any $1 \geq v \geq v^\star$

$$\begin{aligned} 0 &\leq \sum_{v \geq w \geq v^\star} \lambda(w)\varepsilon \\ &= \sum_{w > v^\star - \varepsilon} \lambda(w)\varepsilon - \sum_{w>v} \lambda(w)\varepsilon \\ &= m(v|\delta_H)f(v,\delta_H) - m(v^\star - \varepsilon|\delta_H)f(v^\star - \varepsilon, \delta_H) - \overline{\gamma}(v,\delta_H) - \underline{\gamma}_0(v^\star - \varepsilon, \delta_H), \end{aligned} \tag{3}$$

where we use the necessary conditions that $\overline{\gamma}(v,\delta_H) = 0$ for $v < v^\star$ and that $\underline{\gamma}_0(v,\delta_H) = 0$ for $v \geq v^\star$. Then (3) implies that for each $v \geq v^\star$, $m(v^\star - \varepsilon|\delta_H)f(v^\star - \varepsilon, \delta_H) \leq m(v|\delta_H)f(v,\delta_H)$. Considering these inequalities for all such v and adding the requirement from (1) that $\overline{\gamma}(1,\delta_H) = m(1|\delta_H)f(1,\delta_H)$, it must be that

$$m(v^\star - \varepsilon|\delta_H)f(v^\star - \varepsilon, \delta_H) \leq \min\left\{0, \min_{v^\star \leq v < 1} m(v|\delta_H)f(v,\delta_H)\right\}. \tag{4}$$

This is a necessary condition for the optimality of immediate allocation in (DCP). Together with a similar condition applying to $v < v^\star$, we get sufficiency for the optimality of immediate allocation in (DCP) (see Theorem 1).

To interpret (4), consider the case in which $m(\cdot|\delta_H)f(\cdot,\delta_H)$ is nondecreasing. Then (4) becomes

$$m(v^\star - \varepsilon|\delta_H)f(v^\star - \varepsilon, \delta_H) \leq \min\left\{0, m(v^\star|\delta_H)f(v^\star,\delta_H)\right\}. \tag{5}$$

If this inequality is violated, either the marginal revenue of the $(v^\star, \delta_H)$ type is positive ($m(v^\star - \varepsilon|\delta_H) > 0$) or it exceeds that of $(v^\star, \delta_H)$. In both cases, there is an incentive compatible way in (DCP) to manipulate the allocation of $(v^\star - \varepsilon, \delta_H)$ and $(v^\star, \delta_H)$ to increase revenue. In the former case ($m(v^\star - \varepsilon|\delta_H) > 0$), setting $q_0(v^\star - \varepsilon, \delta_H) = 1$ is incentive compatible in (DCP) and increases revenue. While in the latter case, $q_0(v^\star - \varepsilon, \delta_H) = q_0(v^\star, \delta_H) = 1/2$ is both incentive compatible in (DCP) and revenue improving. Both adjustments weakly increase the utility of all δ_H types and cannot violate any *downward* incentive constraints.

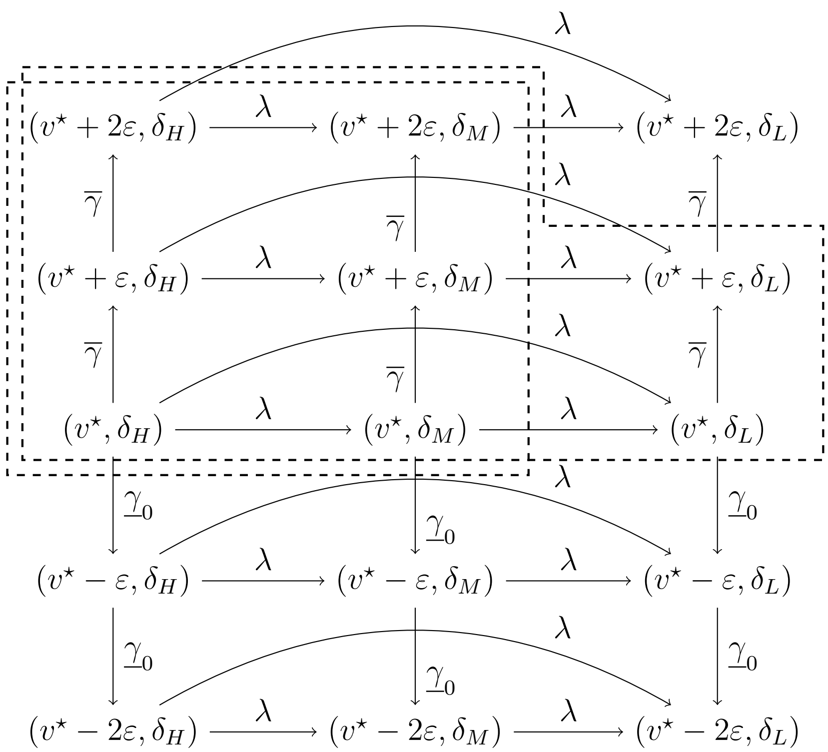

Figure 1: Network Diagram ($\delta_H > \delta_M > \delta_L$)

With more than two discount types our approach is conceptually the same as the one above, we seek conditions under which valid multipliers exist for (DCP); however, to handle the potentially large number of multipliers and equations, we recast the problem as a network flow problem and appeal to an elegant theorem concerning the existence of a feasible flow on a network due to Gale (1957). Gale (1957) considers networks of (i) nodes with a net demand for a divisible "commodity" and (ii) directed arcs between nodes with a capacity for carrying the commodity. The question is whether there is enough capacity in the network to simultaneously satisfy the net demand requirements of all nodes. Gale's theorem provides a simple answer. Net demand can be satisfied if and only if for every partition of the network into two sets of nodes, A and B, the total capacity of the arcs from A to B exceeds the net demand of the nodes in B. See Section A.1 for a formal statement of this theorem (Theorem 3).

We make the formal connection between the problem of determining whether valid multipliers exist and Gale's problem in the proof of Theorem 1. To illustrate the idea, we refer again to the two discount-type example and specifically to the system in (1) and (2). If we difference (1) and (2) across successive values of v the result is

$$-\lambda(v)\varepsilon + \underline{\gamma}_0(v,\delta_H) - \underline{\gamma}_0(v-\varepsilon,\delta_H) - \overline{\gamma}(v,\delta_H) + \overline{\gamma}(v-\varepsilon,\delta_H) = m(v-\varepsilon|\delta_H)f(v-\varepsilon,\delta_H) - m(v|\delta_H)f(v,\delta_H) \quad (6)$$

$$\lambda(v)\varepsilon + \underline{\gamma}_0(v,\delta_L) - \underline{\gamma}_0(v-\varepsilon,\delta_L) - \overline{\gamma}(v,\delta_L) + \overline{\gamma}(v-\varepsilon,\delta_L) = m(v-\varepsilon|\delta_L)f(v-\varepsilon,\delta_L) - m(v|\delta_L)f(v,\delta_L) \quad (7)$$

which takes the form of a supply and demand system over a network as considered by Gale (1957).[12] Each node corresponds to a type (v,δ) and has net demand $d(v,\delta) = m(v-\varepsilon|\delta)f(v-\varepsilon,\delta) - m(v|\delta)f(v,\delta)$. The left-hand side describes the flow in and out of each node. For example, the node (v,δ_H) sends flow $\lambda(v)\varepsilon$ to node (v,δ_L) and flow $\underline{\gamma}_0(v-\varepsilon,\delta_H)$ to $(v-\varepsilon,\delta_H)$. It receives flow $\underline{\gamma}_0(v,\delta_H)$ from $(v+\varepsilon,\delta_H)$. All flows are constrained to be nonnegative. The capacity (upper bound) for any arc is zero if the associated multiplier must be zero in the KKT conditions and infinite otherwise. Figure 1 illustrates an example three discount types and shows only the arcs with infinite capacity.

Gale's theorem applies to the system in (6) and (7) and yields a necessary and sufficient condition for the existence of a λ and γ_0 solving the system, which generalizes the analysis for two discount types above and is reported in Theorem 1. To state the theorem, we first define an order on the type space.

Definition 2. *Let $\succeq$ be the total order on Θ defined by*

$$(v,\delta) \succeq (v',\delta') \iff \min\{v^\star, v'\} \le v \le \max\{v^\star, v'\} \text{ and } \delta \ge \delta'.$$

The set $U \subseteq \Theta$ is an upper set *with respect to $\succeq$ if $(v,\delta) \in U$ and $(v',\delta') \succeq (v,\delta)$ implies $(v',\delta') \in U$.*

In words, a type is considered "higher" than another type if (i) the former's value is between the latter's value and $v^\star$, and (ii) the former has a weakly higher discount type.

Theorem 1 (Optimality of immediate allocation)**.** *The allocation q^I is optimal for* (GP) *if for all upper sets $U \subseteq \Theta$,*

$$\sum_{(v,\delta)\in U} d(v,\delta) \le 0, \quad (8)$$

[12] The above equations only hold as written for interior v with $1 > v > 0$. The endpoints $v = 0$ and $v = 1$ are easily handled, but we defer their treatment to the proof of Theorem 1.

where

$$d(v,\delta) = \begin{cases} m(1-\varepsilon|\delta)f(1-\varepsilon,\delta) & v = 1 \\ m(v-\varepsilon|\delta)f(v-\varepsilon,\delta) - m(v|\delta)f(v,\delta) & 0 < v < 1 \\ 0 & v = 0, \end{cases}$$

is the demand of $(v,\delta) \in \Theta$.

In the proof of Theorem 1, we first derive the system corresponding to (6) and (7) for the case with more than two discount types. Then we formally define the network. The arc capacities in the network are all infinite which makes the application of Gale's theorem straightforward, because we only need to consider partitions of the type space into two sets for which there does not exist an infinite capacity arc from the first set into the second. Two examples of sets without incoming infinite capacity arcs are contained in the dashed boxes in Figure 1. The condition in (8) corresponds to the requirement from Gale's theorem that such sets cannot have a strictly positive net demand.

For upper sets of the form $U = \{(v,\delta') \in \Theta \mid v \geq v^\star \text{ and } \delta' \geq \delta\}$, the condition in (8) requires

$$\sum_{\delta' \geq \delta} m(v^\star - \varepsilon|\delta')f(v^\star - \varepsilon, \delta') \overset{sgn}{=} \mathbb{E}\left[m(v^\star - \varepsilon|\tilde{\delta}) \middle| v^\star - \varepsilon, \tilde{\delta} \geq \delta\right] \leq 0 \tag{9}$$

Intuitively, these conditions require that allocating to all types $(v^\star - \varepsilon, \delta')$ for $\delta' \geq \delta$ is not profitable irrespective of what δ is. This condition holds if the types $(v^\star - \varepsilon, \delta')$ have conditional average information rents that are larger than average information rent of all types with value $v^\star - \varepsilon$. This would be the case if the probability of encountering a type with value higher than $v^\star - \varepsilon$ increases as δ increases.

The conditions in (9) are also the important ones from Theorem 8 for two related reasons. First, it is straightforward to show that if a monotonicity condition holds,[13] these are the only conditions from Theorem 8 that one needs to check.

Corollary 1. *If* $m(v-\varepsilon|\delta)f(v-\varepsilon,\delta) \leq m(v|\delta)f(v,\delta)$ *for all* $(v,\delta) \in \Theta$ *with* $0 < v < 1$, *satisfaction of* (9) *is sufficient for the optimality of* q^I *in* (GP).

Proof. Under the proposed condition $d(v,\delta) \leq 0$ whenever $v < 1$. Therefore, (8) holds if it holds on sets including types with $v = 1$, that is, on all U of the form $\{(v,\delta') \in \Theta \mid v \geq v^\star \text{ and } \delta' \geq \delta\}$. Any upper set that includes one such U but no additional types with $v = 1$ must have a lower sum due to the fact that $d(v,\delta) \leq 0$ for $v < 1$. □

Corollary 1 can also be stated with respect to certain hypothetical prices that the seller might offer. Under the stated monotonicity condition, if the seller were to learn that the

[13] Note that the monotonicity condition would hold under the Uniform distribution.

buyer's discount type is greater than some δ, the optimal price for allocating immediately to the buyer is

$$v^\star_\delta \equiv \underset{v}{\operatorname{argmin}} \left\{ \sum_{\delta' \geq \delta} m(v|\delta') f(v, \delta') \middle| \sum_{\delta' \geq \delta} m(v|\delta') f(v, \delta') \geq 0 \right\}.$$

Again under the monotonicity condition, if the seller upon learning that the buyer's discount factor is higher than some threshold does not want to reduce her price (allocate to more value types), immediate allocation is optimal.

Corollary 2. *If* $m(v-\varepsilon|\delta)f(v-\varepsilon,\delta) \leq m(v|\delta)f(v,\delta)$ *for all* $(v,\delta) \in \Theta$ *with* $0 < v < 1$ *and* $v^\star_\delta \geq v^\star$, q^I *is optimal for* (GP).

Corollaries 1 and 2 are based on the intuitive idea that with only downward constraints on the misreport of discount type, the seller can expand the set of value types that she allocates to for the most patient types of buyers. The conditions in the corollaries rule out the profitability of such adjustments.

Stronger than the requirements behind Corollaries 1 and 2 but easy to state is a statistical condition implying monotonicity of conditional virtual values. Specifically, condition (9) holds if for $\delta > \delta'$ and all $v \in \mathcal{V}$

$$m(v|\delta) \leq m(v|\delta') \iff \frac{1 - F(v|\delta')}{f(v|\delta')} \leq \frac{1 - F(v|\delta)}{f(v|\delta)},$$

where the right-hand inequality compares information rents between the two types. The higher discount type δ has higher information rent for all v when the distribution $F(v|\delta)$ dominates $F(v|\delta')$ in the hazard-rate order, which we write as $F(\cdot|\delta) \succeq_{hr} F(\cdot|\delta')$.

Corollary 3. *If* $m(v-\varepsilon|\delta)f(v-\varepsilon,\delta) \leq m(v|\delta)f(v,\delta)$ *for all* $(v,\delta) \in \Theta$ *with* $0 < v < 1$ *and for all* $\delta \geq \delta'$ $F(\cdot|\delta) \succeq_{hr} F(\cdot|\delta')$, q^I *is optimal for* (GP).

Proof. Under the assumptions, we have for any δ

$$\begin{aligned} 0 \geq m(v^\star - \varepsilon) f(v^\star - \varepsilon) &= \sum_{\delta' \geq \delta} m(v^\star - \varepsilon|\delta') f(v^\star - \varepsilon, \delta') + \sum_{\delta' < \delta} m(v^\star - \varepsilon|\delta') f(v^\star - \varepsilon, \delta') \\ &\geq \left(1 + \frac{\sum_{\delta' < \delta} f(v^\star - \varepsilon, \delta')}{\sum_{\delta' \geq \delta} f(v^\star - \varepsilon, \delta')}\right) \sum_{\delta' \geq \delta} m(v^\star - \varepsilon|\delta') f(v^\star - \varepsilon, \delta'), \end{aligned}$$

implying (9). The inequality follows from

$$m(v^\star - \varepsilon|\delta') \geq \mathbb{E}\left[m(v^\star - \varepsilon|\tilde{\delta})|v^\star - \varepsilon, \tilde{\delta} \geq \delta\right] = \frac{\sum_{\delta'' \geq \delta} m(v^\star - \varepsilon|\delta'') f(v^\star - \varepsilon, \delta'')}{\sum_{\delta'' \geq \delta} f(v^\star - \varepsilon, \delta'')},$$

for $\delta' < \delta$. $\square$

The positive statistical relationship presented in Corollary 3 is stronger than we need, but it is relatively simple to state and can be compared with Theorem 1 of Haghpanah and Hartline (2019) which reports that pure bundling is optimal when the value of the "grand bundle" is positively related to the relative value of smaller bundles. In our case, receiving an immediate allocation could be considered a "grand bundle" while δ controls the value of smaller bundles (later allocations).

The second reason the conditions in (9) are important is that these conditions are also in some sense *necessary* for the optimality of immediate allocation. Because Gale's theorem provides necessary and sufficient conditions, the conditions in Theorem 1 are necessary and sufficient for the optimality of q^I in (DCP) but are only sufficient in (GP). To address if and when the conditions provided are necessary for optimality in (GP), notice that the conditions do not depend on the values taken by the discount factors $\delta \in \mathcal{D}$. Across all potential $\mathcal{D}$, immediate allocation is "least likely" to be optimal when the buyer either does not discount the future at all or receives no value from future allocation (i.e., $\mathcal{D} = \{0, 1\}$). In this case, the seller can delay allocation to the $\delta = 1$ type "for free" because the $\delta = 0$ type receives no value from delayed allocation. Theorem 1 applies in this case, but q^I cannot be the unique solution, because

$$q_0(v, \delta) = \mathbb{I}\{v \geq v^\star\}\mathbb{I}\{\delta = 0\}$$
$$q_1(v, \delta) = \mathbb{I}\{v \geq v^\star\}\mathbb{I}\{\delta = 1\},$$

provides the same amount of revenue and is incentive compatible even in the general problem (GP). When the conditions in (9) fail, the balance tips in favor of the delayed allocation mechanism, making these conditions necessary for the optimality of immediate allocation. We generalize this insight slightly in the next result.

Proposition 2 (Necessary condition from Theorem 1)**.** *If $f(v, \delta)$ is such that the condition*

$$\sum_{\delta' \geq \delta} m(v^\star - \varepsilon|\delta')f(v^\star - \varepsilon, \delta') \leq 0 \tag{10}$$

is violated at $\hat{\delta}$, all $\delta \geq \hat{\delta}$ are sufficiently close to 1, and all $\delta < \hat{\delta}$ are sufficiently close to 0, then q^I is suboptimal.

An example shows the relationship between the preceding results.

Example 1. *Let the type space be $\Theta = \{0, 1/2, 1\} \times \{0, \hat{\delta}\}$, where $\hat{\delta} \in (0, 1]$. The joint distribution F over valuation and discount types is parameterized by probability $\pi \in [0, 1]$*

v	$f(v,0)$	$f(v,\hat{\delta})$
1	$\frac{1}{4}$	$\frac{1}{3}\pi$
$\frac{1}{2}$	$\frac{1}{8}$	$\frac{1}{3}(1-\pi)$
0	$\frac{1}{8}$	$\frac{1}{6}$

v	$m(v\|0)$	$m(v\|\hat{\delta})$
1	1	1
$\frac{1}{2}$	$-\frac{1}{2}$	$\frac{1-2\pi}{2-2\pi}$
0	$-\frac{3}{2}$	-1

v	$d(v,0)$	$d(v,\hat{\delta})$
1	$-\frac{1}{16}$	$\frac{1-2\pi}{6}$
$\frac{1}{2}$	$-\frac{1}{8}$	$-\frac{1-\pi}{3}$
0	0	0

Figure 2: The type distribution, $f(v,\delta)$, marginal revenues, $m(v|\delta)$, and nodal demands, $d(v,\delta)$, for Example 1.

and is shown in Figure 2 on the left. We also show the computed marginal revenues on the right. It follows that

$$v^\star = \begin{cases} 1 & \pi \geq \frac{5}{16} \\ \frac{1}{2} & \pi < \frac{5}{16} \end{cases} \qquad\qquad v^\star_{\hat{\delta}} = \begin{cases} 1 & \pi \geq \frac{1}{2} \\ \frac{1}{2} & \pi < \frac{1}{2}. \end{cases}$$

If $\pi \geq 5/16$, $v^\star = 1$. Applying Theorem 1 upper set $U = \{(1,\hat{\delta})\}$ requires

$$\frac{1-2\pi}{6} \leq 0 \implies \pi \geq \frac{1}{2},$$

and all other conditions from Theorem 1 are satisfied. Note that the conditions of Corollaries 1 and 2 are also satisfied here.

When $\pi < 5/16$ making $v^\star = 1/2$, the analogous condition from Theorem 1, applied to the set $U = \{(1,\hat{\delta}),(1/2,\hat{\delta})\}$ is always satisfied since

$$\frac{1-2\pi}{6} - \frac{1-\pi}{3} = -\frac{1}{6} \leq 0.$$

To summarize we find immediate allocation is optimal if $\pi \notin [5/16, 1/2)$ which is exactly the range where $v^\star > v^\star_{\hat{\delta}}$.

3.2 Adding upward misreports of discount type

The problem considered for Theorem 1, (DCP), omits the incentive constraints related to the buyer misreporting too high of a discount factor. In this section we bring back those incentive constraints and study the problem that results. Specifically, the subject of this section is the problem (UDCP) described below.

$$\max_{q\in[0,1]^{T+1}} \quad \mathbb{E}_\Theta\left[\sum_{t=0}^{T}\delta^t m(\tilde{v}|\tilde{\delta})q_t(\tilde{v},\tilde{\delta})\right] \tag{UDCP}$$

$$\text{s.t.} \quad \sum_{t=0}^{T}\delta^t Q_t(v,\delta)\varepsilon \geq \sum_{t=0}^{T}\left\{\delta'^t Q_t(v,\delta')\varepsilon + q_t(v,\delta')\left(\delta^t-\delta'^t\right)v\right\} \quad \forall v, \forall \delta\neq\delta' \tag{UDIC}$$

$$\sum_{\delta\in\mathcal{D}}\sum_{t=0}^{T}q_t(v,\delta)f(v,\delta)\leq f(v) \qquad \forall v\in\mathcal{V}. \tag{F'}$$

Despite including more constraints from the original problem (GP), this problem still incorporates two types of relaxations. The first is that we omit "diagonal" IC constraints involving joint misreports of value and discount type, and one can show that the diagonal constraints are not necessarily implied by the misreport of value and misreport of discount type constraints. The second is that the feasibility condition we use here is a relaxed version of the feasibility condition used for Theorem 1. Using the condition (F') makes the proof of Theorem 2 tractable, because it allows us to pin down the values of the associated multipliers in the linear program.

In the previous section when only downward constraints on discount type reports were included, immediate allocation is optimal if (roughly) the seller does not want to increase the allocation of types with high discount factors relative to those with low discount factors by allocating to a larger set of value types. When we include the upward misreport of discount type constraints, we consider the willingness of the lower discount types to accept a delayed allocation, leading to the importance of the relative values taken by the high and low discount factors.

The final consequence is that with (UDCP) we can no longer use the network flow analogy, and more specifically Gale's Theorem, to study the existence of multipliers guaranteeing the optimality of q^I. In the proof of Theorem 2, we first derive a system which is necessary and sufficient for q^I to solve (UDCP), just as we do in the proof of Theorem 1. However, this system no longer corresponds to the network flow problem we analyzed for Theorem 1, roughly because the discount factors cannot be ignored. In Theorem 2, we instead appeal to Farkas' lemma to derive the resulting sufficient condition.

Theorem 2 (Optimality of immediate allocation$'$)**.** *The mechanism q^I is optimal if for all types $(v,\delta)\in\Theta$ with $v>0$,*

$$\sum_{j>1}\left\{(\delta_j-\delta_{j-1})\frac{1}{\varepsilon}\sum_{k\geq j}\left(\mu(v,\delta_k)-\mu(v-\varepsilon,\delta_k)\right)+(1-\delta_j)\frac{1}{v}m_+(v)f(v,\delta_j)\right\}\geq 0 \tag{11}$$

where we index the discount types for buyer i in increasing order as $\delta_1 < \delta_2 < \cdots$ and define

$$m_+(v) = m(v)\{v \geq v^\star\}$$
$$\mu(v, \delta) = (m(v|\delta) - m_+(v))f(v, \delta).$$

Theorem 2 shows that adding constraints on upward misreports of discount type makes the values taken by the discount factors relevant, which was not true of Theorem 1. To help explain the conditions imposed by (11), consider the two-discount type case again with $\delta_H > \delta_L$. In this case, the condition becomes

$$(\delta_H - \delta_L)\frac{1}{\varepsilon}\left(\mu(v, \delta_H) - \mu(v - \varepsilon, \delta_L)\right) + (1 - \delta_H)\frac{1}{v}m_+(v)f(v, \delta_H) \geq 0.$$

Summing over v such that $1 \geq v \geq v^\star$, the condition implies

$$(1 - \delta_H)\sum_{v' \geq v^\star}\frac{1}{v'}m(v')f(v', \delta_H) \geq (\delta_H - \delta_L)\frac{1}{\varepsilon}m(v^\star - \varepsilon|\delta_H)f(v^\star - \varepsilon, \delta_H),$$

because $\mu(1, \delta_H) = 0$ and $m_+(v^\star - \varepsilon) = 0$. Comparing this condition to the one in (9) for Theorem 1, it is clear that (9) implies the condition above. On the other hand, the condition above is weaker because it allows for $m(v^\star - \varepsilon|\delta_H) > 0$. The degree to which this can be positive depends on the values taken by δ_H and δ_L. When δ_H is much larger than δ_L, it becomes easier for the seller to delay the allocation to the δ_H discount type, because making $\delta_H - \delta_L$ larger leads the δ_L type to value the δ_H type's delayed allocation less. As a consequence, we must place more stringent requirements on $m(v^\star - \varepsilon|\delta_h)f(v^\star - \varepsilon, \delta_h)$. In the extreme case where $\delta_H = 1$ and $\delta_L = 0$, we get back the condition from Theorem 1 that $m(v^\star - \varepsilon|\delta_h)f(v^\star - \varepsilon, \delta_h) \leq 0$. Alternatively, as $\delta_H - \delta_L$ becomes small, the condition above becomes easier to satisfy. At the other extreme, $\delta_H = \delta_L$, the condition is satisfied for any $f(v, \delta)$. The rationale is straightforward. If both discount types discount the future at the same rate, there is no way to separate the types using delayed allocation and immediate allocation must be optimal.

Revisiting Example 1, we show how Theorem 2 expands the set of parameters under which immediate allocation is optimal. This example suggests that Theorem 2 produces a strictly larger set of parameters under which immediate allocation is optimal. However, this need not be true generally. Theorem 2 uses a more relaxed feasibility constraint compared to Theorem 1. This makes the proof of Theorem 2 tractable, but potentially makes the condition produced by Theorem 2 more demanding.

	$\mu(v,0)$	$\mu(v,\hat{\delta})$
1	0	0
v $\frac{1}{2}$	$-\frac{2-3\pi}{22-16\pi}$	$\frac{2-3\pi}{22-16\pi}$
0	$-\frac{3}{16}$	$-\frac{1}{6}$

$\pi < 5/16$

	$\mu(v,0)$	$\mu(v,\hat{\delta})$
1	0	0
v $\frac{1}{2}$	$-\frac{1}{16}$	$\frac{1-2\pi}{6}$
0	$-\frac{3}{16}$	$-\frac{1}{6}$

$\pi \geq 5/16$

Figure 3: The incremental marginal revenues, $\mu(v,\delta)$, for Example 1.

Example 1 (continued)**.** *The conditions in Theorem 2 require*

$$-2\hat{\delta}\mu\left(\frac{1}{2},\hat{\delta}\right) + (1-\hat{\delta})\frac{\pi}{3} \geq 0 \tag{12}$$

$$\hat{\delta}\left(\mu\left(\frac{1}{2},\hat{\delta}\right) - \mu\left(0,\hat{\delta}\right)\right) + (1-\hat{\delta})m_+\left(\frac{1}{2}\right)\frac{1-\pi}{3} \geq 0 \tag{13}$$

When $\pi \geq 5/16$, condition (13) *is satisfied for all* $(\pi,\hat{\delta})$ *with $\pi \geq 5/16$, while condition* (12) *is equivalent to*

$$\frac{1-\hat{\delta}}{\hat{\delta}} \geq \frac{1-2\pi}{\pi},$$

and is satisfied for $\pi \geq 1/2$ and some $\pi < 1/2$ depending on the value of $\hat{\delta}$. Recall that Theorem 1 required $\pi \geq 1/2$ in this case. If $\pi < 5/16$, conditions (12) *and* (13) *become*

$$\frac{1-\hat{\delta}}{\hat{\delta}} \geq \frac{3}{\pi}\frac{2-3\pi}{11-8\pi} \tag{14}$$

$$17\hat{\delta} + (1-\hat{\delta})(5-16\pi) \geq 0, \tag{15}$$

the second of which is satisfied for all $(\pi,\hat{\delta})$ *with $\pi < 5/16$.*

In Figure 4, we illustrate the application of each of our conditions to this example. As discussed above, Theorem 1 applies irrespective of discount type and is tight at the extreme where $\hat{\delta} = 1$, which corresponds to the result from Proposition 2. At this point, immediate allocation being optimal requires that $\pi \geq 1/2$.

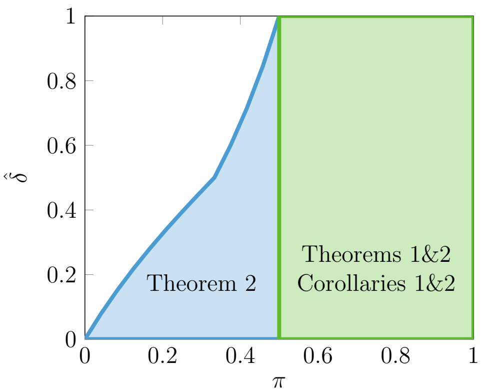

Figure 4: Optimal mechanisms in Example 1. Theorem 2 implies that immediate allocation is optimal in all shaded regions. Theorem 1 and Corollaries 1 and 2 are weaker in this example and only imply that immediate allocation is optimal in the right-hand regions.

4 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 value 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 value 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.[14] The seller believes that the feasible set of joint distributions is $\mathcal{F} \subseteq \{\hat{F} \colon \hat{F}_v = F_v \text{ and } \operatorname{Supp} \hat{F}_\delta = \mathcal{D}\}$.

[14]Carroll (2017) analyzes the case in which the seller knows the marginal distribution F_δ but not the joint distribution F, 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.

The seller's problem is[15]

$$\max_{\{q,p\}} \inf_{F\in\mathcal{F}} \sum_{t=0}^{T} \mathbb{E}_F \left[p\left(\tilde{v}, \tilde{\delta}\right)\right], \tag{AP}$$

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

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

$$\sum_{t=0}^{T} q_t(v,\delta) \leq 1 \quad \forall (v,\delta)\in\Theta. \tag{F}$$

Proposition 3 (Optimality of nondiscrimination with little information)**.** *Suppose that there is $F \in \mathcal{F}$ that satisfies the condition of Theorem 1. Then q^I is optimal in the seller's problem with ambiguous temporal preferences.*

When the statistical relationship between value 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 optimized with an immediate allocation mechanism. Since immediate allocation generates the same revenue regardless of the joint distribution of value and discount types, the optimal mechanism allocates immediately.

5 Related literature and conclusion

Our technical 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 nonoptimal.[16] In our model, the set of goods corresponds to the ability to allocate a fixed

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

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

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 $0 \leq \sum_t q_t \leq 1$, while the feasibility constraint in most bundling analyses is $0 \leq q_k \leq 1$ for all goods k.[17] 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 infinite.

Our main result is related to Haghpanah and Hartline (2019), which gives conditions under which a monopolist sells only a "grand bundle" of all products. The buyer's initial value v in our model corresponds to the value for the grand bundle in Haghpanah and Hartline (2019), and their Theorem 1 corresponds roughly to our Corollary 3. Our Theorem 1 provides broader conditions for optimality than our Corollary 3, and 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 immediate allocation is feasible, regardless of the relationship between discount types and valuation types. This allows us to avoid the complication of evaluating which IC constraints bind. Previous work has examined which incentive constraints will bind in optimal mechanisms (Carroll, 2012; Archer and Kleinberg, 2014; Mishra et al., 2016).[18] 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 immediate allocation given only these constraints; adding unconsidered constraints back to the problem does not affect the feasibility of immediate allocation and therefore does not affect its optimality.[19] 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.

Our model can also be related to work on dynamic pricing. In a previous working paper version, we showed that our results can be extended to a model in which multiple potential buyers arrive over time. When buyers with symmetric and commonly-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.[20] Pai and Vohra (2013) and Mierendorff (2016) consider the possibility that agents have privately-known deadlines. A key distinguishing

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

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

[19] 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 dynamic models, and is not at odds with our finding that incentive constraints cause the seller to not screen on discount factor.

[20] 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.

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.

5.1 Conclusion

Sellers in dynamic environments may be imperfectly aware of buyers' temporal preferences. We model a mechanism design problem in which a buyer has private information about his value and temporal preferences and the seller can potentially improve revenue by screening on the buyer's discount factor. We provide conditions under which the optimal mechanism ignores temporal preferences and allocates to a given buyer either immediately or never. Our results thus provide statistical conditions under which immediate allocation mechanisms remain optimal in a world with heterogeneous time preferences. We further show that when the seller has ambiguous beliefs regarding the buyer's temporal preferences, an immediate allocation mechanism is optimal so long as it is plausibly 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.

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

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.

Marc K Chan. Welfare dependence and self-control: An empirical analysis. *The Review of Economic Studies*, 84(4):1379–1423, 2017.

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

Peter Diamond and Johannes Spinnewijn. Capital income taxes with heterogeneous discount rates. *American Economic Journal: Economic Policy*, 3(4):52–76, 2011.

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

Emmanuel Farhi and Iván Werning. Estate taxation with altruism heterogeneity. *American Economic Review*, 103(3):489–95, 2013.

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

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

Mikhail Golosov, Maxim Troshkin, Aleh Tsyvinski, and Matthew Weinzierl. Preference heterogeneity and optimal capital income taxation. *Journal of Public Economics*, 97: 160–175, 2013.

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.

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 specific feasible flow theorem that we use in Theorem 1 is due to Gale (1957).[21] 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 $b(x)$.[22] 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.

$$\sum_{\{x' | (x',x) \in \mathcal{A}\}} g(x', x) - \sum_{\{x' | (x,x') \in \mathcal{A}\}} g(x, x') = d(x) \qquad \forall x \in \mathcal{N} \tag{16}$$

$$0 \leq g(x, x') \leq k(x, x') \qquad \forall (x, x') \in \mathcal{A}, \tag{17}$$

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

Theorem 3. *There exists a solution, g, to the system in* (16) *and* (17) *if and only if*

$$\sum_{x \in X, x' \in \bar{X}} k(x, x') \geq \sum_{x' \in \bar{X}} d(x') \qquad \forall X \subseteq \mathcal{N}, \tag{18}$$

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.

[21]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.

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

A.2 Proof of Lemma 1

Proof of Lemma 1. From the constraints on the local misreport of value we have

$$\sum_{t=1}^{T} \delta^t \{q_t(v,\delta) - q_t(v-\varepsilon,\delta)\} v \geq p(v,\delta) - p(v-\varepsilon,\delta) \geq \sum_{t=1}^{T} \delta^t \{q_t(v,\delta) - q_t(v-\varepsilon,\delta)\} (v-\varepsilon).$$

The proposed transfers satisfy the upper bound of these constraints with equality for all types.

These transfers are optimal unless they are infeasible (not incentive compatible). Assume q is non-discriminatory and allocates immediately. Clearly, then $p(v,\delta) = p(v,\delta')$ for any $\delta,\delta' \in \mathcal{D}$. Along with the assumptions on q this implies that the payoff is insensitive to the report of discount type,

$$u(v,\delta) = u(v,\delta'|v,\delta),$$

and hence that the proposed transfers satisfy incentive compatibility with respect to the report of discount type. Since we also have $u(v',\delta'|v,\delta) = u(v',\delta|v,\delta)$, the proposed transfers are incentive compatible with respect to joint misreports of value and discount type,

$$u(v,\delta) \geq u(v',\delta|v,\delta) = u(v',\delta'|v,\delta).$$

□

A.3 Proof of Theorem 1

Proof of Theorem 1. This proof considers the problem (DCP). The coefficient on $q_0(v,\delta)$, denoted $c_0(v,\delta)$, in the linear programming problem representing the seller's revenue maximization is given by

$$c_0(v,\delta) = m(v|\delta) f(v,\delta) - \overline{\gamma}(v,\delta) + \underline{\gamma}_0(v,\delta) + \sum_{\substack{\delta>\delta' \\ v'>v}} \lambda(v',\delta,\delta')\varepsilon - \sum_{\substack{\delta'>\delta \\ v'>v}} \lambda(v',\delta',\delta)\varepsilon,$$

where $\overline{\gamma}(v,\delta)$ and $\underline{\gamma}_0(v,\delta)$ are multipliers on the (F) and $q_0(v,\delta) \geq 0$ constraints respectively. The multiplier $\lambda(v,\delta,\delta')$ is associated with the constraint $u(v,\delta) \geq u(v,\delta'|v,\delta)$ for $\delta > \delta'$.

The coefficients $c_t(v,\delta)$ can be written as

$$\begin{aligned} c_t(v,\delta) &= \delta^t c_0(v,\delta) + \underline{\gamma}_t(v,\delta) \\ &\quad - \left[\left(1-\delta^t\right)\overline{\gamma}(v,\delta) + \delta^t \underline{\gamma}_0(v,\delta) + \sum_{\delta'>\delta} \lambda(v,\delta',\delta)\left(\delta'^t - \delta^t\right) v\right] \\ &\leq \delta^t c_0(v,\delta) + \underline{\gamma}_t(v,\delta), \end{aligned} \tag{19}$$

where the inequality follows from the nonpositivity of the bracketed term.

To prove the optimality of the rule q^I, it is sufficient to find feasible values for the multipliers such that for all types (v,δ) the following KKT conditions are satisfied:

$$c_t(v,\delta) = 0, \left(1 - \sum_t q_t(v,\delta)\right)\overline{\gamma}(v,\delta) = 0, \text{ and } q_t(v,\delta)\,\underline{\gamma}_t(v,\delta) = 0 \quad \forall t, v, \delta, \tag{LP}$$

with $\lambda(v,\delta,\delta')$, $\overline{\gamma}(v,\delta)$ and $\underline{\gamma}_t(v,\delta)$ nonnegative for all v, $\delta \neq \delta'$ and $t \geq 0$. The second and third conditions in (LP) are complementary slackness conditions for the respective feasibility constraints. Complementary slackness for the λ constraints are satisfied implicitly by the fact that the constraints hold with equality at q^I.

First note that it is sufficient to find multipliers for the $t=0$ terms. If $c_0(v,\delta)=0$, the inequality in (19) indicates that we can set $\underline{\gamma}_t(v,\delta)$ to the value taken by the bracketed term to make $c_t(v,\delta)=0$. Consequently, we can focus on the $t=0$ terms in (LP).

Under q^I, $q_0^I(v,\delta) = \mathbb{I}\{v \geq v^\star\}$. We therefore require that $\underline{\gamma}_0(v,\delta)=0$ for $v \geq v^\star$ and $\overline{\gamma}(v,\delta)=0$ for $v < v^\star$. Note also that $c_0(1,\delta)=0$ if and only if $\overline{\gamma}(1,\delta)=1$.

Satisfying $c_0(v,\delta)=0$ for all (v,δ) is equivalent to satisfying $c_0(v-\varepsilon,\delta) - c_0(v,\delta) = 0$ for all $v>0$ given $c_0(1,\delta)=0$. This leads to the system

$$\sum_{\delta'>\delta} \lambda(1,\delta',\delta)\,\varepsilon - \sum_{\delta>\delta'} \lambda(1,\delta,\delta')\,\varepsilon + \overline{\gamma}(1-\varepsilon,\delta) = d(1,\delta) \tag{20}$$

$$\sum_{\delta'>\delta} \lambda(v,\delta',\delta)\,\varepsilon - \sum_{\delta>\delta'} \lambda(v,\delta,\delta')\,\varepsilon + \overline{\gamma}(v-\varepsilon,\delta) - \overline{\gamma}(v,\delta) = d(v,\delta) \qquad 1 > v > v^\star \tag{21}$$

$$\sum_{\delta'>\delta} \lambda(v^\star,\delta',\delta)\,\varepsilon - \sum_{\delta>\delta'} \lambda(v^\star,\delta,\delta')\,\varepsilon - \overline{\gamma}(v^\star,\delta) - \underline{\gamma}_0(v^\star-\varepsilon,\delta) = d(v^\star,\delta) \tag{22}$$

$$\sum_{\delta'>\delta} \lambda(v,\delta',\delta)\,\varepsilon - \sum_{\delta>\delta'} \lambda(v,\delta,\delta')\,\varepsilon - \underline{\gamma}_0(v-\varepsilon,\delta) + \underline{\gamma}_0(v,\delta) = d(v,\delta) \qquad v^\star > v > 0, \tag{23}$$

where

$$d(v,\delta) \begin{cases} m(1-\varepsilon|\delta) f(1-\varepsilon,\delta) & v = 1 \\ m(v-\varepsilon|\delta) f(v-\varepsilon,\delta) - m(v|\delta) f(v,\delta) & 1 > v > 0. \end{cases}$$

If we sum (20)–(23) across $0 < v \leq 1$ and $\delta \in \mathcal{D}$ and negate, we get

$$\sum_{\delta \in \mathcal{D}} \underline{\gamma}_0(0, \delta) = -m(0)f(0), \tag{24}$$

which motivates the addition of a dummy node in the network below.

To apply Theorem 3, we represent this system as a network in which each type (v, δ) with $v > 0$ is associated with an equation in (20)–(23) and a node in $\mathcal{N}$. For arcs in $\mathcal{N}$, each flow $\lambda(v, \delta, \delta')$ is associated with an arc from (v, δ) to (v, δ') for $0 < v \leq 1$ and $\delta > \delta'$, each $\overline{\gamma}(v - \varepsilon, \delta)$ is associated with a flow from $(v - \varepsilon, \delta)$ to (v, δ) for $v \geq v^\star$, and each $\underline{\gamma}_0(v - \varepsilon, \delta)$ is associated with a flow from (v, δ) to $(v - \varepsilon, \delta)$ for $v^\star \geq v > 0$. See Figure 1 in the text for a depiction. To make aggregate net demand zero, we add a single dummy node $(0, \underline{\delta})$ associated with the equation (24). Into $(0, \underline{\delta})$ there are arcs with flow $\underline{\gamma}_0(0, \delta)$ from $(0, \delta)$ for each $\delta \in \mathcal{D}$.

More formally,

$$g(v, \delta, v', \delta') = \begin{cases} \lambda(v, \delta, \delta')\,\varepsilon & \text{if } v' = v, 0 < v \leq 1, \text{ and } \delta > \delta', \\ \overline{\gamma}(v', \delta) & \text{if } v' = v + \varepsilon \leq 1 \text{ and } \delta = \delta', \\ \underline{\gamma}_0(v', \delta) & \text{if } v' = v - \varepsilon > 0 \text{ and } \delta = \delta' \\ \underline{\gamma}_0(0, \delta) & \text{if } v = \varepsilon, v' = 0 \text{ and } \delta' = \underline{\delta}, \end{cases}$$

The capacity of any arc described by g is infinite. Setting $d(0, \underline{\delta}) = -m(0)f(0)$, is straightforward to verify that $\sum_{v>0} \sum_{\delta} d(v, \delta) + d(0, \underline{\delta}) = 0$. We apply Theorem 3 to this network.

Let $X \subseteq \Theta$ be a set of types (v, δ) and let $\bar{X} = \Theta \setminus X$ be its complement. There are three cases in which inequality (18) is slack, because the left-hand side is infinite:

- There are types $(v, \delta) \in X$ and $(v, \delta') \in \bar{X}$ such that $\delta > \delta'$;
- There are types $(v, \delta) \in X$ and $(v', \delta) \in \bar{X}$ such that $v' > v \geq v^\star$.
- There are types $(v, \delta) \in X$ and $(v', \delta) \in \bar{X}$ such that $v' < v \leq v^\star$.

In the remaining cases, $\bar{X}$ is an upper set according to Definition 2. Theorem 3 then requires that for all such upper sets $\bar{X}$,

$$\sum_{(v,\delta) \in \bar{X}} d(v, \delta) \leq 0,$$

because the total capacity of arcs from X to $\bar{X}$ in this case is zero.

□

A.4 Proof of Proposition 2

Proof of Proposition 2. Suppose that $\sum_{\delta \geq \hat{\delta}} m(v^\star - \varepsilon|\delta) f(v^\star - \varepsilon, \delta) > \min\{0, \sum_{\delta \geq \hat{\delta}} m(v^\star|\delta) f(v^\star, \delta)\}$ for some $\hat{\delta}$. By the definition of $v^\star$, we must have $\hat{\delta} > \min_{\delta \in \mathcal{D}} \delta$.

First, if $\sum_{\delta \geq \hat{\delta}} m(v^\star - \varepsilon|\delta) f(v^\star - \varepsilon, \delta) > 0$, consider the mechanism

$$\begin{aligned} q_0(v, \delta) &= \mathbb{I}\{v \geq v^\star\}\mathbb{I}\{\delta < \hat{\delta}\} \\ q_t(v, \delta) &= \mathbb{I}\{v \geq v^\star - \varepsilon\}\mathbb{I}\{\delta \geq \hat{\delta}\}\mathbb{I}\{t = 1\} \\ p(v, \delta) &= \begin{cases} v^\star & \text{if } v \geq v^\star \text{ and } \delta < \hat{\delta} \\ \hat{\delta}(v^\star - \varepsilon) & \text{if } v \geq v^\star - \varepsilon \text{ and } \delta \geq \hat{\delta} \end{cases} \end{aligned}$$

This is incentive compatible for a type with $\delta < \hat{\delta}$ and $v \geq v^\star$ if

$$\begin{aligned} v - v^\star &\geq \delta v - \hat{\delta}(v^\star - \varepsilon) \\ (1 - \delta)v &\geq (1 - \hat{\delta})v^\star + \hat{\delta}\varepsilon, \end{aligned}$$

which holds for δ sufficiently close to 0 and $\hat{\delta}$ sufficiently close to 1. Types with $\delta < \hat{\delta}$ and $v < v^\star$ have no profitable deviations. For a type with $\delta \geq \hat{\delta}$ and $v \geq v^\star - \varepsilon$, incentive compatibility requires

$$\begin{aligned} \delta v - \hat{\delta}(v^\star - \varepsilon) &\geq v - v^\star \\ (1 - \hat{\delta})v^\star + \hat{\delta}\varepsilon &\geq (1 - \delta)v \end{aligned}$$

which holds for δ and $\hat{\delta}$ sufficiently close to 1. Types with $\delta \geq \hat{\delta}$ and $v < v^\star - \varepsilon$ have no profitable deviations.

Expected revenue is greater than it is under q^I if

$$\begin{aligned} \hat{\delta}(v^\star - \varepsilon) \sum_{\substack{v \geq v^\star - \varepsilon \\ \delta \geq \hat{\delta}}} f(v, \delta) &\geq v^\star \sum_{\substack{v \geq v^\star \\ \delta \geq \hat{\delta}}} f(v, \delta) \\ \hat{\delta}(v^\star - \varepsilon) \sum_{\delta \geq \hat{\delta}} f(v^\star - \varepsilon, \delta) &\geq ((1 - \hat{\delta})v^\star + \hat{\delta}\varepsilon) \sum_{\substack{v \geq v^\star \\ \delta \geq \hat{\delta}}} f(v, \delta), \end{aligned}$$

which holds for $\hat{\delta}$ sufficiently close to 1 since

$$\sum_{\delta \geq \hat{\delta}} m(v^\star - \varepsilon|\delta) f(v^\star - \varepsilon, \delta) > 0 \iff (v^\star - \varepsilon) \sum_{\delta \geq \hat{\delta}} f(v^\star - \varepsilon, \delta) > \varepsilon \sum_{\substack{v \geq v^\star \\ \delta \geq \hat{\delta}}} f(v, \delta)$$

□

B Proofs for Section 3.2 (Upwards and downwards misreports)

Lemma 2. *Let $\delta_i > \delta_j$, and define $w : \mathbb{R}_{++} \to \mathbb{R}$ by*

$$w_t\left(\delta_i, \delta_j\right) = \frac{\delta_i^t - \delta_j^t}{1 - \delta_i^t}.$$

Then $w_t\left(\delta_i, \delta_j\right) > w_s\left(\delta_i, \delta_j\right)$ for all $t < s$.

Proof. The first derivative of w takes the same sign as[23]

$$\begin{aligned}\frac{\partial}{\partial t} w_t\left(\delta_i, \delta_j\right) &\overset{\text{sgn}}{=} \left(1 - \delta_i^t\right)\left(\delta_i^t \ln \delta_i - \delta_j^t \ln \delta_j\right) + \left(\delta_i^t - \delta_j^t\right)\delta_i^t \ln \delta_i \\ &= \left(1 - \delta_j^t\right)\delta_i^t \ln \delta_i - \left(1 - \delta_i^t\right)\delta_j^t \ln \delta_j \\ &\overset{\text{sgn}}{=} \underbrace{\frac{\delta_i^t \ln \delta_i}{1 - \delta_i^t}}_{v_t(\delta_i)} - \frac{\delta_j^t \ln \delta_j}{1 - \delta_j^t}.\end{aligned}$$

The sign of derivative of $v_t(\delta)$ with respect to δ is

$$v_t'(\delta) \overset{\text{sign}}{=} t \ln \delta + 1 - \delta^t.$$

Since $v_t'(1) = 0$ and $(\partial/\partial\delta)(t \ln \delta + 1 - \delta^t) > 0$, $v_t'(\delta) < 0$ and the result follows. □

Proof of Theorem 2. We consider the problem specified in (UDCP). The initial approach is similar to the proof of Theorem 1. The coefficient on $q_0(v, \delta)$, denoted $c_0(v, \delta)$, in the linear programming problem representing the seller's revenue maximization is given by

$$c_0\left(v, \delta\right) = m\left(v|\delta\right) f(v, \delta) - \overline{\gamma}\left(v\right) f(v, \delta) + \underline{\gamma}_0\left(v, \delta\right) + \sum_{\substack{\delta \neq \delta' \\ v' > v}} \lambda\left(v', \delta, \delta'\right)\varepsilon - \sum_{\substack{\delta' \neq \delta \\ v' > v}} \lambda\left(v', \delta', \delta\right)\varepsilon,$$

where $\overline{\gamma}(v)$ and $\underline{\gamma}_0(v, \delta)$ are multipliers on the (F') and $q_0(v, \delta) \geq 0$ constraints respectively. The multiplier $\lambda(v, \delta, \delta')$ is associated with the constraint $u(v, \delta) \geq u(v, \delta'|v, \delta)$ for $\delta \neq \delta'$.

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

Since the KKT conditions require that $c_0(v,\delta) = 0$ for all (v,δ), for any v

$$0 = \sum_{\delta} c_0(v,\delta) = m(v)f(v) - \overline{\gamma}(v)f(v,\delta) + \sum_{\delta} \underline{\gamma}_0(v,\delta),$$

which implies that $\overline{\gamma}(v) = m_+(v) \equiv m(v)\mathbb{I}\{v \geq v^\star\}$. Using $\mu(v,\delta) \equiv (m(v|\delta) - m_+(v))f(v,\delta)$,

$$c_0\left(v,\delta\right) = \mu(v,\delta) + \underline{\gamma}_0\left(v,\delta\right) + \sum_{\substack{\delta \neq \delta' \\ v' > v}} \lambda\left(v',\delta,\delta'\right)\varepsilon - \sum_{\substack{\delta' \neq \delta \\ v' > v}} \lambda\left(v',\delta',\delta\right)\varepsilon,$$

Using the necessary $c_0(v,\delta) = 0$, we also have

$$c_t\left(v,\delta\right) = -(1-\delta^t)m_+(v)f(v,\delta) + \underline{\gamma}_t(v,\delta) - \delta^t\underline{\gamma}_0(v,\delta) + \sum_{\delta' \neq \delta} \lambda(v,\delta',\delta)(\delta^t - \delta'^t)v \tag{25}$$

The KKT conditions require that for all types (v,δ) the following condition is satisfied:

$$c_t\left(v,\delta\right) = 0, \underline{\gamma}_t(v,\delta) \geq 0, \lambda(v,\delta',\delta) \geq 0, \text{ and } q_t\left(v,\delta\right)\underline{\gamma}_t\left(v,\delta\right) = 0 \;\forall t, v, \delta \neq \delta'. \tag{LP}$$

We seek nonnegative values for $(\lambda, \underline{\gamma}_t)$ that satisfy for all $t \geq 0$, and $(v,\delta) \in \Theta^i$

$$\sum_{\substack{\delta' \neq \delta \\ w > v}} \left(\lambda\left(w,\delta,\delta'\right) - \lambda\left(w,\delta',\delta\right)\right)\varepsilon + \underline{\gamma}_0(v,\delta)\mathbb{I}\{v < v^\star\} = -\mu(v,\delta) \tag{26}$$

$$\sum_{\delta' \neq \delta} \lambda(v,\delta',\delta)(\delta^t - \delta'^t)v + \underline{\gamma}_t(v,\delta) - \delta^t\underline{\gamma}_0(v,\delta)\mathbb{I}\{v < v^\star\} = (1-\delta^t)m_+(v)f(v,\delta) \tag{27}$$

Using Farkas' Lemma, there exists a nonnegative solution $(\lambda, \underline{\gamma}_t)$ to (26)– (27) if and only if there are no $x(v,\delta)$ and $y_t(v,\delta)$ satisfying for all $(v,\delta) \in \Theta$

$$\sum_{v' < v} \left(x(v',\delta) - x(v',\delta')\right)\varepsilon \geq \sum_{t > 0} y_t(v,\delta')(\delta^t - \delta'^t)v \qquad \delta \neq \delta' \tag{28}$$

$$x(v,\delta) \geq \sum_{t > 0} \delta^t y_t(v,\delta) \qquad v < v^\star \tag{29}$$

$$y_t(v,\delta) \geq 0 \qquad t > 0. \tag{30}$$

$$\sum_{(v,\delta) \in \Theta} x(v,\delta)\mu(v,\delta) > \sum_{\substack{(v,\delta) \in \Theta \\ t > 0}} y_t(v,\delta)(1-\delta^t)m_+(v)f(v,\delta) \tag{31}$$

We make several adjustments to simplify the problem. Next, define $X(v,\delta) \equiv \sum_{v' < v} x(v',\delta)$ with $X(0,\delta) \equiv 0$, $\nu(v,\delta) \equiv m_+(v)f(v,\delta)$, $\tilde{y}_t(v,\delta) \equiv (1-\delta^t)y_t(v,\delta)$ and $\Delta_v\mu(v,\delta) \equiv \mu(v,\delta) - \mu(v-\varepsilon,\delta)$. Finally, we index the discount types in increasing order, $\delta_1 < \delta_2 < \cdots < \delta_i <$

$\cdots < \delta_D$, with D being the number of discount types. With these changes, the system becomes

$$X(v,\delta_i) - X(v,\delta_j) \geq \sum_{t>0} \tilde{y}_t(v,\delta_j)\frac{\delta_i^t - \delta_j^t}{1-\delta_j^t}\frac{v}{\varepsilon} \qquad i \neq j \tag{32}$$

$$X(v+\varepsilon,\delta_i) - X(v,\delta_i) \geq \sum_{t>0} \delta_i^t \frac{\tilde{y}_t(v,\delta_i)}{1-\delta_i^t} \qquad v < v^\star \tag{33}$$

$$\tilde{y}_t(v,\delta_i) \geq 0 \qquad t > 0. \tag{34}$$

$$\sum_{\substack{(v,\delta)\in\Theta \\ t>0}} X(v,\delta_i)\Delta_v\mu(v,\delta_i) + \tilde{y}_t(v,\delta_i)\nu(v,\delta_i) < 0 \tag{35}$$

We want to find a condition that is equivalent to the system above having no solution.

Towards that end, we first choose $\tilde{y}_t$ to minimize the left-hand side of (35) subject to the constraints implied by the other inequalities. The problem we are interested in is

$$\min_{\tilde{y}_t} \sum_{\substack{(v,\delta)\in\Theta \\ t>0}} \tilde{y}_t(v,\delta_i)\nu(v,\delta_i) \text{ s.t. } \frac{v}{\varepsilon}\sum_{t>0} \tilde{y}_t(v,\delta_i)\frac{\delta_i^t - \delta_j^t}{1-\delta_i^t} \geq X(v,\delta_i) - X(v,\delta_j) \quad i > j \tag{36}$$

Since $\tilde{y}_t(v,\delta_i) \geq 0$ by (34), (32) implies $X(v,\delta_i) \geq X(v,\delta_j)$ whenever $i > j$. Consequently, the constraints in (36) place the relevant lower bounds on each $\tilde{y}_t(v,\delta_i)$. Next, since $\nu(v,\delta_i) \geq 0$ and $(\delta_i^t - \delta_j^t)/(1-\delta_i^t)$ is decreasing in t when $i > j$ (Lemma 2), in the solution to (36), $\tilde{y}_t(v,\delta_i) = 0$ for all i and $t > 1$.[24] Consequently, it is without loss to eliminate all $\tilde{y}_t(v,\delta_i)$ such that $t > 1$ from the problem. Henceforth, we drop the subscript on $\tilde{y}$ and w with the understanding that $t = 1$.

Next, we show that only the "local" inequalities in (32) are relevant. Rewriting (32),

$$y(v,\delta_i)\frac{v}{\varepsilon} \geq \frac{X(v,\delta_i) - X(v,\delta_j)}{\delta_i - \delta_j} \geq y(v,\delta_j)\frac{v}{\varepsilon} \quad i > j, \tag{37}$$

but these inequalities are implied by the corresponding inequalities where $j = i-1$. To see

[24] Suppose that $\tilde{y}_t(v,\delta_i) > 0$ for $t > 1$ and the constraint in (36) corresponding to (i,j) is binding. Then we can simultaneously reduce $\tilde{y}_t(v,\delta_i)$ to zero while increasing $\tilde{y}_1(v,\delta_i)$ by

$$\tilde{y}_t(v,\delta_i)\frac{\delta_i^t - \delta_j^t}{1-\delta_i^t}\frac{1-\delta_i}{\delta_i - \delta_j} \leq \tilde{y}_t(v,\delta_i),$$

where the inequality follows from Lemma 2. If more than one constraint binds in (36), we can find the constraint corresponding to

$$j \in \operatorname*{argmax}_k \frac{\delta_i^t - \delta_k^t}{1-\delta_i^t}\frac{1-\delta_i}{\delta_i - \delta_k},$$

and perform a similar adjustment to $\tilde{y}_t$.

this, note that for $i > j$

$$\frac{X(v,\delta_i) - X(v,\delta_j)}{\delta_i - \delta_j} = \frac{1}{\delta_i - \delta_j} \sum_{k=i}^{j+1} (\delta_k - \delta_{k-1}) \frac{X(v,\delta_k) - X(v,\delta_{k-1})}{\delta_k - \delta_{k-1}}$$
$$\leq \frac{1}{\delta_i - \delta_j} \sum_{k=i}^{j+1} (\delta_k - \delta_{k-1}) y(v,\delta_k)$$

where the right-hand side is a convex combination, and (37) implies that for all $k < i$

$$y(v,\delta_i) \geq y(v,\delta_k).$$

From the previous two paragraphs and defining for $i > 1$

$$\Delta_\delta X(v,\delta_i) \equiv \frac{X(v,\delta_i) - X(v,\delta_{i-1})}{\delta_i - \delta_{i-1}}$$
$$\Delta_v X(v,\delta_i) \equiv \frac{X(v+\varepsilon,\delta_i) - X(v,\delta_i)}{\varepsilon}$$

it follows that the solution to (36) is to set

$$y(v,\delta_i) = \begin{cases} 0 & i = 1 \\ \Delta_\delta X(v,\delta_i)\frac{\varepsilon}{v} & i > 1. \end{cases}$$

We can now reduce the system (32)–(35) to

$$\Delta_\delta X(v,\delta_i) \geq \Delta_\delta X(v,\delta_{i-1}) \qquad i > 2 \qquad (38)$$
$$\Delta_\delta X(v,\delta_2) \geq 0 \qquad (39)$$
$$\Delta_\varepsilon X(v,\delta_i) \geq \delta_i \Delta_\delta X(v,\delta_i)\frac{1}{v} \qquad v \leq v^\star \qquad (40)$$
$$\sum_{(v,\delta_i)\in\Theta} \Delta_v \mu(v,\delta_i) X(v,\delta_i) + \sum_{\substack{(v,\delta_i)\in\Theta \\ i>1}} \frac{\varepsilon}{v} \Delta_\delta X(v,\delta_i)(1-\delta_i)\nu(v,\delta_i) < 0. \qquad (41)$$

Then using for $i > 1$

$$X(v,\delta_i) = X(v,\delta_1) + \sum_{k=2}^{i} (\delta_k - \delta_{k-1}) \Delta_\delta X(v,\delta_k)$$

and $\sum_i \Delta_v \mu(v, \delta_i) = 0$, inequality (41) becomes

$$\sum_{\substack{(v,\delta_i)\in\Theta \\ i>1}} \Delta_\delta X(v, \delta_i) \left[(\delta_i - \delta_{i-1}) \sum_{k\geq i} \frac{1}{\varepsilon} \Delta_v \mu(v, \delta_k) + (1-\delta_i) \frac{1}{v} \nu(v, \delta_i) \right] < 0. \tag{42}$$

Let

$$M(v, \delta_i) \equiv (\delta_i - \delta_{i-1}) \sum_{k\geq i} \frac{1}{\varepsilon} \Delta_v \mu(v, \delta_k) + (1-\delta_i) \frac{1}{v} \nu(v, \delta_i).$$

Using this definition, rewrite the left-hand side of (42) once more as

$$\begin{aligned}
&\sum_{\substack{(v,\delta_i)\in\Theta \\ i>1}} \Delta_\delta X(v, \delta_i) M(v, \delta_i) \\
&\quad = \sum_{\substack{(v,\delta_i)\in\Theta \\ i>1}} \left(\Delta_\delta X(v, \delta_2) + \sum_{3\leq k\leq i} \Delta_\delta X(v, \delta_k) - \Delta_\delta X(v, \delta_{k-1}) \right) M(v, \delta_i) \\
&\quad = \sum_{\substack{(v,\delta_i)\in\Theta \\ i>1}} \left(M(v, \delta_i) \Delta_\delta X(v, \delta_2) + M(v, \delta_i) \sum_{3\leq k\leq i} \Delta_\delta X(v, \delta_k) - \Delta_\delta X(v, \delta_{k-1}) \right) \\
&\quad = \sum_v \left\{ \Delta_\delta X(v, \delta_2) \sum_{j\geq 2} M(v, \delta_j) + \sum_{i>2} \left(\Delta_\delta X(v, \delta_i) - \Delta_\delta X(v, \delta_{i-1}) \right) \sum_{j\geq i} M(v, \delta_j) \right\}
\end{aligned}$$

By (38) and (39), $\Delta_\delta X(v, \delta_i)$ is positive and increasing in i for $i \geq 2$. The sum above is nonnegative for all such $\Delta_\delta X(v, \delta_i)$ if and only if

$$\sum_{j\geq i} M(v, \delta_j) \geq 0 \quad \forall i \geq 2, v > 0.$$

The $v > 0$ condition follows from $\Delta_\delta X(0, \delta) = 0$.

□

C Proofs for Section 4 (Ambiguous temporal preferences)

Proof of Proposition 3. Note that it is feasible to compute q^I in this context, since the marginal distribution of valuation types is known. Then it is sufficient to show that any other mechanism will yield weakly lower maxmin revenue. For any fixed $F \in \mathcal{F}$,

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

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, q^I is optimal. □