Resource and Revenue Management in Nonprofit Operations — **Source link** ↗

Francis de Véricourt, Miguel Sousa Lobo

Institutions: European School of Management and Technology, INSEAD

Related papers:

- Efficient Funding: Auditing in the Nonprofit Sector

- An Elasticity Approach to the Newsvendor with Price-Sensitive Demand

- Pricing and the News Vendor Problem: a Review with Extensions

- A Decision Analysis Tool for Evaluating Fundraising Tiers

- Dynamic Pricing in the Presence of Inventory Considerations: Research Overview, Current Practices, and Future Directions

de Véricourt, Francis; Lobo, Miguel Sousa

Working Paper

Resource and Revenue Management in Nonprofit Operations

This Version is available at:
http://hdl.handle.net/10419/96558

ESMT
WORKING PAPER

ESMT No. 08-006

Resource and Revenue Management in Nonprofit Operations

Francis de Véricourt, ESMT
Miguel Sousa Lobo, Duke University

european school
of management
and technology

Abstract

Resource and Revenue Management in Nonprofit Operations[1]

Author(s): Francis de Véricourt, ESMT
         Miguel Sousa Lobo, Duke University

Nonprofit firms sometimes engage in for-profit activities for the purpose of generating revenue to subsidize their mission activities. The organization is then confronted with a consumption vs. investment tradeoff, where investment corresponds to providing capacity for revenue customers, and consumption corresponds to serving mission customers. Exemplary of this approach are the Aravind Eye Hospitals in India, where profitable paying hospitals are used to subsidize care at free hospitals. We model this problem as a multi-period stochastic dynamic program. In each period, the organization must decide how much of the current assets should be invested in revenue-customer service capacity, and at what price the service should be sold. We provide sufficient conditions under which the optimal capacity and pricing decisions are of threshold type. Similar results are derived when the selling price is fixed but the banking of assets from one period to the next is allowed. We compare the performance of the optimal threshold policy with heuristics that may be more appealing to managers of nonprofit organizations, and assess the value of banking and of dynamic pricing through numerical experiments.

Keywords: capacity allocation, revenue management, dynamic pricing, nonprofit

1 Introduction

There are today 20 million blind eyes in India and, as life spans increase and more people reach the age for high incidence of cataracts, 2 million new blind eyes are added each year. Currently, only 1 million eyes are being operated in India each year, resulting in much unmet demand. To address an even more dire situation in 1976, Dr. G. Venkataswamy built a small paying hospital, mostly with the small initial capital of his family.[1] The paying hospital was expanded through 1977, and a first free hospital was built in 1978. In Dr. Venkataswamy's words, "from the revenue generated from operations [at the ground floor of the first paying hospital] we built the next floor, and so on until we had a nice five-story facility. And then with the money generated there, we built the Free Hospital" (Rangan 1994, page 7). After much growth over 30 years, the Aravind hospitals throughout Tamil-Nadu now perform well over one hundred thousand operations each year (Shah and Murty 2004).

While the Aravind story has become noted as an especially successful case, this pattern is not unique and is becoming more common. Nonprofit organizations often engage in a mix of activities, some of which are designed primarily to generate excess revenue to subsidize other activities that more directly serve the organization's charitable mission. This has led nonprofit researchers to address the strategic question of how to construct an optimal portfolio of activities (Gruber and Mohr 1982; Oster 1995). However, few have paid attention to the practical decisions about resource and capacity allocation that must be made once a given portfolio of activities has been selected.[2] In this paper, we use a modeling approach to develop insights regarding the optimal allocation of resources between revenue-generating and mission-serving activities over time in an organization that seeks to maximize its mission impact.

In order to isolate the tradeoffs involved in balancing revenue-generating and mission-serving activities across many time periods, we have constructed a simplified model of an organization with just two activities and two corresponding customer groups. Adapting a convention introduced by Weisbrod (1998), we refer to *R-activities* as those that generate more revenue than costs. They serve *R-customers*, which at Aravind are the patients at the paying hospitals. R-activities also may, but need not, generate mission impact. While the mission impact of operating on a paying customer at Aravind is not zero, it is significantly smaller than for free customers given that alternative for-profit hospitals are available to paying customers. We refer to *M-activities* as those that generate positive

---

[1] "I mortgaged my house and raised enough money to start", Dr. Venkataswamy, in Rangan (1994, page 7).

[2] One exception is Young (2004). He addresses a broad range of resource-allocation issues, within and across activities, using a marginal contribution framework. The main differences between his analysis and ours are 1) we capture the tradeoff between spending on mission activities now versus in the future, 2) we allow for randomness is demand for the revenue-producing activities, and 3) our model does not assume decreasing marginal social return.

mission impact, but require a financial subsidy.[3] They serve *M-customers*, which at Aravind are the patients at the free hospitals, who come from India's poorest economic classes. The organization begins with an endowment of resources and is attempting to be self-sufficient, with its R-activities funding its M-activities (we do, however, also consider the case where the organization may receive grants over time).

With increased competition for philanthropic and government funding over the past two decades, nonprofits are under increased pressure to generate more revenue from their own R-activities, even if they are relatively low on mission contribution, to allow for cross subsidy within the nonprofit. As a result of this pressure, the creation of revenue-generating ventures has become a major topic of discussion in the nonprofit field over the past two decades. See, for instance, Skloot (1983), Skloot (1987), Skloot (1988), Dees (1998), Young (2002), Oster, Massarsky, and Beinhacker (1995), Foster and Bradach (2005), Weisbrod (2005). Numerous 'how to' guides have been produced to help nonprofits develop profit-making ventures. See, for instance, Steckel, Simons, and Lengsfelder (1989), Alter (2000), Boschee (2001), Anderson, Dees, and Emerson (2002), Larson and Forth (2002), Robinson (2002), Dees (2004).

This kind of nonprofit cross-subsidy can take different forms and other examples can be cited in addition to Aravind. The Boston Symphony Orchestra runs its 'Boston Pops' series using some of the same musicians and facilities as the regular orchestra, and relies on the revenues to subsidize other orchestra activities that are artistically important but less lucrative (Oster 1995). This pattern is common among performing arts organizations that schedule more popular performances to subsidize more artistically important, but less popular ones. Consider the reliance of ballet companies on seasonal performances of 'The Nutcracker' or theatre companies that rely on popular musicals or plays. Sometimes social service agencies will use key resources, such as their staff, facilities, or equipment, to deliver marketable products that subsidize more charitable operations. For instance, CIPO Productions, a Brazilian organization, provides impoverished young people with training in photography, video production, and Web design, and helps subsidize these activities with revenue from selling the use of its production and computer equipment to customers who can pay (Elstrodt, Schindler, and Waslander 2004). Brinckerhoff (2000, p. 16) describes an Illinois school for behaviorally-challenged adolescents that has created a business using its staff expertise to deliver workshops on 'Dealing with Difficult Teenagers', for which it charges admission. The net

---

[3]We are only concerned with net financial contribution. An M-activity may generate some revenue that covers part of its operational costs. As long as the net financial contribution of producing a unit of this good or service is negative, we can model it as a pure mission good with costs equal to the net negative financial contribution. Similarly, an R-activity must produce a net positive financial contribution that can be used in future periods to subsidize M-activities.

proceeds help to fund the school's core mission activities. Any social service provider that includes a practice focused on paying clients in order to subsidize its mission-related services faces the kind of resource-allocation decision we are discussing.

Once the strategic decision is made to engage in activities that produce a net financial contribution to subsidize other activities that contribute to the mission, nonprofit managers are faced with budgeting and resource-allocation decisions. How much of their organization's resources and capacity in a given period do they devote to activities with profit potential, as opposed to activities that serve the mission? Our model focuses on this resource-allocation decision. Since social impact is the ultimate objective, this decision involves a tradeoff between achieving mission impact now versus generating additional financial resources in order to achieve mission impact in the future. For increased robustness of our results, we include in the model the option of banking funds to earn interest. Managers then have three options in the allocation of funds: to revenue service capacity, to mission service capacity, or banking them.

Under a set of plausible assumptions, we solve for the optimal policy for the allocation of resources. We find that, when resources are scarce, all resources should either be allocated to revenue activities or banked. Only when resources are above a threshold should the organization engage in mission activities, and all resources above the threshold should be allocated to mission activities. The threshold increases when we allow R-activities to generate mission impact. We also find that banking is never optimal toward the end of the time horizon, and our numerical results suggest that relatively small expected gains are obtained from including the option of banking.

Since nonprofits may be able to set their prices strategically, we also analyze a version of the model that allows for dynamic pricing decisions. The threshold effect still holds. Under some assumptions for the probability distribution of R-customer demand (but not for all distributions), the optimal price is decreasing in the available resources. Finally, when the price is fixed and banking is not allowed we find, perhaps surprisingly, that the optimal service-capacity decision is myopic. The optimal threshold above which resources should be allocated to mission activities is not time-dependent and can be analytically derived.

No doubt many nonprofit leaders will want to allocate some resources to M-activities in all periods simply because that is the organization's mission. Their instincts may be based on factors that fall outside our model. To drop all M-activities could affect the culture of the organization, the motivation of staff, board members, volunteers, and donors, and it could even jeopardize the nonprofit legal status of the organization, with the consequent tax implications. Also, for some organizations, especially for some types of medical services, if there are no alternative organizations

that can provide a replacement for critical services, the impact of M-activities can be significantly non-linear. The marginal mission value of the first few invested resources will be very high. These may be sufficient reasons not to allocate all resources solely to R-activities, even when resources are below the threshold. However, our numerical studies under a range of parameters in our model show that a proportional allocation heuristic performs poorly compared to the optimal threshold policies. According to our model, this prioritization of resources toward M-activities will result in reduced total mission impact when future periods are considered, even given that future mission impacts are discounted. Any allocation of resources to M-activities before the threshold is met must be justified for reasons that are external to our model. The important point is for nonprofit leaders to understand that the gains from their decision to support M-activities must offset the potential loss of future social impact. Our analysis should help nonprofit managers understand the cost (in terms of future mission impact) of diverting resources away from revenue-generating activities when the resource level is below the optimal threshold.

While this work has been motivated by an interest in the management of nonprofit organizations, a related investment vs. consumption tradeoff arises in the dividend-payment problem, where a (for-profit) firm must decide how much of the available cash to pay out as dividends, and how much to re-invest into the operation. As a result, our work can be related to a large literature in the for-profit sector. When firms operate in a complete financial market, Modigliani and Miller (1958) show that these decisions can be made independently. As has often been noted, many, if not most, firms do not operate in such conditions. This is especially true for small firms that do not have access to reasonably priced capital, such as in the equity or bond markets, either because of their small size or because they operate in an economy without a well-developed financial system. The literatures on cash-flow management and corporate finance have paid scant attention to the interaction between financial and operational decisions that consequently arise. These interactions have been explored in recent work which is more closely related to our model. Buzacott and Zhang (2001) consider a manufacturer with limited funds who needs an asset-based loan from a bank to finance future growth. The objective is to maximize the retained earnings at the end of the time horizon. Babich and Sobel (2004) study production, sales, and loan size decisions to optimize the expected discounted proceeds from an initial public stock offer. The control of dividends problem studied by Li, Shubik, and Sobel (2003), which includes inventory-replenishment decisions, seems to be the most closely related to our model. In particular, they establish that the optimal policy is myopic for linear holding and backorder costs. Their model is similar to our problem with banking, but with the organization having unlimited access to borrowing. Demand is backordered

in their model, while we assume lost-sales (lost-sales is better suited to the problem faced by many firms, especially in service industries). Also, we consider pricing decisions, which generally have not been addressed in the literature on joint operational and financial decisions. The prices set by a firm affect its revenue, which in turn affects the cash constraints and future revenue and growth. Our model is most closely linked to the literature on dynamic inventory models with pricing and stochastic demand. Elmaghraby and Keskinocak (2003) and Bitran and Caldentey (2003) present comprehensive reviews of this literature. For recent results, see also Chen, Ray, and Song (2003) who consider the lost sales case, and Chen and Simchi-Levi (2004) for the backorder case. Chen, Wu, and Yao (2004) model the system with a Brownian motion and provide a worst-case bound for the value of dynamic pricing. In this stream of research, the pricing strategy and the inventory stock replenishment decisions determine the inventory levels available for future periods. This differs fundamentally from the cash constraints that follow from the capacity decisions in our model. For instance, Huh and Janakiraman (2005) propose an elegant framework based on sample path analysis to retrieve and extend most of the existing results for multi-period dynamic pricing and inventory problems (their approach cannot be applied to our model because in our setting the pricing decision directly affects the budget constraint of the following period).

Before formally introducing the model we outline its key assumptions and discuss their practical justification.

- *Periodic Capacity Shifting and Resource Allocation.* Our model assumes that capacity can be reallocated only periodically at decision times and not between these periods. This fits with nonprofits that 1) are using common core assets in both R-activities and M-activities, and 2) have to schedule the use of these assets in advance. The first condition fits with the recommendations of strategists in this area. According to Oster (1995, p. 91) the most promising new activity extensions, including those created to generate revenue, are ones that can 'best share in the core assets of the organization'. The second condition states that these core assets cannot be shifted between activities during the period.

  Aravind Eye Hospital may be able to shift resources (doctors and nurses) frequently, but is unlikely to do so mid-shift. The paying and free hospitals are kept separate due to different levels of service (even staff uniforms are different), and planning will require a commitment to staff assignments. A requirement to commit to capacity decisions will also arise from the need to organize the camps that are run in poor villages to recruit and screen candidates for surgery.[4] Consider also performing arts groups that schedule their seasons in advance.

---

[4]In addition to the capacity decisions requiring significant pre-commitment, there is the ability to increase or

Shifting capacity mid-season is extremely rare. They make their bets and live with them. Even social service agencies that attempt to balance paying clients (R-customers) with charitable clients (M-customers) may need to decide in advance how much staff time to make available for paid and charitable work. It is important to keep in mind that the decision period can be short.

Of course, some organizations are not constrained in this way, as resources can sometimes be shared by R- and M-activities during the production period. We briefly explore this case, and find that the threshold effect still holds. However, further work is needed toward a complete study of this variant of the problem.

- *Random Demand Model.* Although the newsvendor model provides a reasonable approximation for many situations faced by nonprofits and is widely used in the field of operations, we show that the threshold structure is maintained for a very general relationship between capacity decisions and demand from revenue customers. We use the Boston Symphony as a motivating example where the newsvendor model assumption that the total demand in each period is independent of the capacity decision (the number of performances scheduled) is not appropriate.

- *Linearity of Social Return.* We assume that nonprofit organizations can expand their M-activities without experiencing diminishing social returns. This is a reasonable assumption for nonprofit organizations that serve only a small share of the need for their mission activities. Aravind can only address a small portion of the need for cataract surgery among the poor in India. It is often the case that a social service agency could serve many more mission clients, if it simply had the capacity, without any dilution of impact. The case is less clear with performing-arts organizations, if we think of customer demand for the artistic productions. However, thinking in terms of valuable artistic pieces (ballets, plays, symphonies, *etc.*) that could be performed, the number of productions with artistic merit that the company might like to perform far exceeds its capacity. Additional capacity for M-performances may not lead to diminishing returns on this artistic level.[5]

---

decrease the total variable expenditures over both the paying and free hospitals. While the flexibility in adjusting total staff costs from one period to the next is limited, it is not negligible. As most staff work with long hours, there may be considerable flexibility to increase or decrease overtime. Other items are clearly variable costs and can be shifted to either revenue of mission customers, such as intra-ocular lenses, which represent 8% of operating costs (a share which is increasing as free hospitals also shift to this procedure).

[5]For the non-linear case, we expect that some form of threshold effect holds at least when the rate at which the marginal social return decreases is not too high relative to the 'thickness' of the tail of the distribution or R-customer demand.

- *Discounting Future Social Impacts.* We assume that projected future social outcomes should be discounted. There are two rationales for this. The first is mathematical. If we fail to use a mission-impact discount rate and also allow for the possibility of earning real interest (above the inflation rate of operating costs) on banked funds, the mathematical model will result in no investment in mission service capacity in the current period, ever (Keeler and Cretin 1984). It will always be superior to bank the funds in order to purchase more mission impact in later periods. The second reason is that, in most cases, addressing social problems (global warming, poverty, spread of disease, *etc.*) or delivering social goods (health care, education, the arts) is better done sooner than later. A discount factor captures this degree of urgency, which can vary widely. It is also often used in practice to reflect the uncertainty surrounding the delivery and value of projected future benefits. Using a discount rate implies that a more certain immediate impact should be valued more highly than a less certain future impact of the same magnitude. We recognize that the use of discount rates with regard to social impact is controversial (Klausner 2003), and it certainly oversimplifies the complex considerations that should inform time preferences in the nonprofit sector, such as considerations of intergenerational equity or the potential value of intervening at strategic times. However, we believe that the use of discounting, that is, some preference for immediate impacts versus later impacts, is reasonable for our purposes.

- *Grants and Fundraising.* We model grants as randomly distributed inflows in each period. While Aravind chooses not to rely on grants, many nonprofit organizations do. The Boston Symphony receives grants from the National Endowment for the Arts and charitable foundations, as well as private and corporate donations. However, only grants that the organization can use to run either side of its operations (revenue or mission) are relevant to the model. Grants that impose the restriction that they may only be used for the mission side of operations can be treated as entirely separate from the revenue-mission tradeoff model (they do, however, have an impact in that, by increasing the funding for mission activities, they reduce other concerns such as retention and motivation of volunteer staff). We explicitly leave out of our resource-allocation model the diversion of resources to cover fundraising costs, and reserve it as a topic for future extensions. Note, however, that one option for including fundraising costs would be to treat it, as Weisbrod (1998) suggests, simply as an R-activity. Fundraising costs could be deducted from donations to indicate a net financial yield, which presumably would be positive, but with low or zero direct mission impact. This approach is acceptable provided that operating capacity and resources can be allocated to fundraising, just as well

as they can be allocated to other R-activities that more resemble operational programs.

- *Sorting Customers.* Our model distinguishes R-customers from M-customers. In some cases, it is the nature of the customers (*e.g.*, income level) that determines the degree of mission impact of an activity. We assume that, when it matters (with regard to mission impact), these customers will sort themselves to the appropriate activities. In particular, R-customers will not simply switch to being M-customers if a comparable M-activity is available as a substitute. In most cases, this is not a problem because the M-activities are often not good substitutes for the R-activities. Even when the activities are similar, various barriers are likely to lead to appropriate sorting. Few of Aravind's wealthier patients would pretend to be poor to get the free version of the surgery. In some cases, R-customers can also be M-customers without disrupting the model because the mission impact is not tied to the nature of the customers served. This is the case with the performing-arts organizations. The same person can attend a Pops concert and a regular symphony without affecting the model results.

To make the problem tractable we have created a simple model of a very complex world. Nonprofits are complex and diverse, with widely varying economic structures. They include, for instance, homeless shelters, ballet companies, hospitals, environmental advocacy groups, social service agencies, museums, sports and athletic associations, schools and colleges, and grassroots political groups. Most of them are far more complex than our simple two-activity model, and some nonprofits are not aiming to be economically self-sufficient. Nonetheless, we believe our model, by isolating this tradeoff, highlights a dynamic that is relevant to many nonprofit organizations as they search for the kind of balanced portfolio recommended by Oster (1995). The underlying dynamic in our model captures an increasingly common tension that nonprofit managers face. We offer it as, we hope, a meaningful first step in what we anticipate will be a larger, longer-term project of applying modeling techniques to decision making in nonprofit settings. We hope that it stimulates more interest by the operations community in a large section of the economy (7% to 10% of the US economy, depending on the metric) to which it has dedicated little attention.[6] It should also inform the thinking of nonprofit managers regarding the potential implications of their resource-allocation decisions as they pursue a balance between financial and mission activities.

The rest of the paper is organized as follows. Section 2 addresses the problem of optimal capacity and banking decisions, with a fixed selling price to R-customers. We do this for a broad class of demand models, which includes newsvendor-type demand as a special case. (In Section 2 of the

---

[6]See McCardle (2005), who also suggests that the operations community, through its professional societies, set standards for *pro bono* work, similarly to what is done in the legal profession.

related de Véricourt and Lobo (2008) we briefly discuss conditions under which it is optimal not to bank.) Section 3 considers the problem without banking and with a newsvendor-type demand model, where the organization makes a pricing decision in each time period. Section 4 presents numerical studies. We first investigate the relative merit of the optimal capacity policy relative to a proportional-allocation policy. We then explore the expected value of including banking and pricing decisions in the problem specification. Section 5 discusses extensions, including the flexible-capacity case, and directions for future research. We conclude in Section 6.

2 Resource Allocation with Banking

2.1 The Model

We consider discrete-time problems. In each time period t, given a capacity y_t allocated for R-customers, the organization serves $y_t \wedge \Theta_t$ (the minimum of y_t and Θ_t) R-customers. The Θ_t are random variables with finite mean, mutually independent and stationary (most of our results extend, however, to the case of non-stationary independent Θ_t). The newsvendor model, which assumes that the distribution of Θ_t is independent of the capacity decision y_t, is widely used in the literature and is adequate in many instances, including the Aravind example. In this section, we consider a more general model for the random demand, in which its distribution depends on the capacity decision (we will later return to the newsvendor model). As a motivation, consider again the Boston Symphony example. Demand is not independent of capacity because scheduling more performances provides customers with more options. For instance, people who are not able to attend a performance on a given day due to other commitments may be able to attend if another performance date is scheduled. Increasing capacity by scheduling more performances is therefore likely to shift upwards the distribution of demand. We allow for this by defining, for each y, $f_y(\cdot) : [\underline{\theta}, \bar{\theta}] \mapsto \mathbf{R}^+$ as the probability density function of Θ_t given a capacity decision $y_t = y$, with $\underline{\theta}$ non-negative and $\bar{\theta}$ possibly equal to $+\infty$. We also consider the corresponding cumulative distribution $F_y(\cdot)$, the tail distribution $G_y(\cdot) = 1 - F_y(\cdot)$, and the inverse of the cumulative $F_y^{-1}(\cdot)$, all assumed differentiable in y. The mean of Θ_t is a function of y_t, which we assume bounded above by a finite value $\overline{\Theta}$. The newsvendor model is the special case where the demand distribution is independent of the capacity decision, that is $F_y(\cdot) = F(\cdot)$ for all y. A natural assumption in the Boston Symphony example is that increased capacity cannot lead to reduced demand. If a new performance is added, some patrons may switch from another performance that they would have gone to instead. However, patrons will not decide not to see, say, a Schubert performance just

because a Shostakovich performance is added to the schedule. This translates to assuming that the family of distributions $f_y(\cdot)$ is ordered in y according to stochastic dominance, which corresponds to $F_y^{-1}(u)$ increasing in y for each $u \in [0, 1]$. The results in this section are derived under a related assumption of progressive market saturation. The marginal gains in demand are assumed to become smaller as capacity increases so that, for each $u \in [0, 1]$, $F_y^{-1}(u)$ is concave in y.[7]

Serving one M-customer increases the social impact of the organization by s (in social impact units) but does not generate revenue. When revenue-generating activities are also related to the mission of the organization, we denote by τ the social impact of serving one R-customer. The unit service-capacity costs for M- and R-customers are c_M and c_R. We assume lost-sales, so that the costs for R-customers are incurred as a function of the capacity allocated even if that capacity is not fully utilized (the salaries of nurses and doctors assigned to Aravind's paying hospitals have to be paid in full, independently of how many surgeries are performed; the production costs of a music performance are not any lower if the auditorium is half-full). Without loss of generality, we assume $c_R = 1$. This is simply a change of asset units, and we also assume that the selling price and the cost of serving M-customers have been modified so as to be specified in these units.

We assume ample demand of M-customers, that is, resources allocated to this group will always be consumed without decrease in the marginal social impact. In the Aravind case, this corresponds to noting that the number of poor people in need of eye surgery if far greater than the combined capacity of all charity hospitals.[8]

In this section, we assume the unit selling price p is fixed. This will be the case when the organization is a price-taker, with the price set by the market in which the organization competes for R-customers. When this is not the case, the organization may still decide to offer the service at the same price in all time periods due to marketing or organizational constraints.

Also for this section, we assume that the organization has an alternative financial application for its assets in a risk-free investment with return β. The organization needs to determine what portion of its assets should be 'banked', what portion should be invested in capacity for R-customers, and what portion should be allocated to capacity for M-customers. We will also briefly discuss the relative importance of including banking in the problem formulation. In particular, in Section 2 of de Véricourt and Lobo (2008), we show that if the time horizon is short enough, banking is never optimal.

---

[7]The assumption of F^{-1} increasing in y is not required for our results. However, the concavity condition (which is required) is more naturally understood in the context of increasing F^{-1}.

[8]For the Boston Symphony, if the mission is understood as maintaining an active repertoire of skills and artistically important works, the demand for those performances is, in this sense, a secondary concern. Further, the organization can fill auditoriums to fulfill its mission of exposing a broader public to classical music by means such as offering discounts and student tickets.

The system state is the assets a_t held by the organization at the beginning of period $t \in [1, \ldots T]$. At the beginning of each period, the organization decides how much service capacity y_t to provide for R-customers, and sets the amount of current assets which are banked, $z_t \geq 0$. Note that since the unit cost is normalized to one, y_t can interchangeably be used to denote the resources allocated to provide capacity and the capacity itself. The choices of y_t and z_t are limited by the current resources, according to the constraint $y_t + z_t \leq a_t$. The remaining resources, $x_t = a_t - y_t - z_t$, are allocated to serving M-customers. Demand Θ_t is then realized, and the number of R-customers served by the organization is the smallest of demand and capacity, which we write $y_t \wedge \theta_t$. The service is perishable and any unused capacity is lost. Finally, the organization may receive donations in each period, with the amount a random variable Δ_t, assumed stationary and independent and with continuous density function and finite mean. The resources available to the organization at the beginning of the following period are then

$$a_{t+1} = p\,(y_t \wedge \theta_t) + \beta\,z_t + \Delta_t.$$

The organization contributes to its mission by serving $(a_t - y_t - z_t)/c_M$ M-customers and $y_t \wedge \theta_t$ R-customers, yielding a social return of $s/c_M(a_t - y_t - z_t) + \tau(y_t \wedge \theta_t)$. Without loss of generality we set $s/c_M = 1$, which is simply a change in social return units. We assume that the units of τ, the social return from serving an R-customer, are also modified so as to be specified in these units, so that τ represents the social impact of a unit of assets spent on serving R-customers relative to the social impact of a unit spent on M-customers. The units of the value-to-go function are also changed accordingly. We assume that the social return of serving an M-customer is higher than that of serving an R-customer, *i.e.*, $\tau \leq 1$.

We begin by considering problems with a finite horizon of T periods. This may arise, for instance, if the initial assets of the organization are from a grant that specifies a duration of time in which to achieve its objectives. We later extend results to the infinite-horizon case. The social discount factor for delaying service to an M-customer to the next period is α, with $0 \leq \alpha < 1$. This discount factor measures the urgency of the social need that the organization addresses. The overall social return from the organization's activities is the total discounted sum of the social return from the number of customers served in each period. The objective is to determine a banking and capacity-provision policy that maximizes the expected social return over the time horizon T. We formulate the problem of jointly determining the best banking and capacity decisions as a finite-horizon Markov decision process.

Denote by $v_t(a)$ the maximum social-impact-to-go at period t given current resources $a \geq 0$. In the last period, all assets are allocated to serving M-customers, so that $v_T(a) = a$. For $t < T$,

$v_t(a)$ can be shown to satisfy the optimality equations (see, for instance, Heyman and Sobel 1984)

$$v_t(a) = \max_{\substack{0 \le y \quad 0 \le z \\ y + z \le a}} (a - y - z) + \tau\, \mathbf{E}_y(y \wedge \Theta) + \alpha H^{v_{t+1}}(y, z) \tag{1}$$

$$= a + \max_{\substack{0 \le y \quad 0 \le z \\ y + z \le a}} J^{v_{t+1}}(y, z), \tag{2}$$

where the operators J^v and H^v are defined for any real-valued function $v(\cdot)$ as

$$J^v(y, z) = -y - z + \tau\, \mathbf{E}_y\,(y \wedge \Theta) + \alpha H^v(y, z) \tag{3}$$

$$H^v(y, z) = \mathbf{E}_\Delta \mathbf{E}_y\, v\,(p(y \wedge \Theta) + \beta z + \Delta), \tag{4}$$

and $\mathbf{E}_\Delta$ and $\mathbf{E}_y$ are the expectation over Δ and over Θ given y. On account of the stationary assumptions, we will generally omit the time index from the random variables.

We denote by $(y_t^*(a), z_t^*(a))$ the optimal resource-allocation and banking decisions at period t given current assets a. The optimal policy $(y_t^*(a), z_t^*(a))$ corresponds to the maximizer of $J^{v_{t+1}}$ subject to the constraints $0 \le y$, $z \le 0$, and $y + z \le a$. (In case of multiple optima, $(y_t^*(a), z_t^*(a))$ designates the minimum optimal decision using, say, the lexicographic order.)

As a side result, note that if $\alpha\beta \ge 1$ the organization always prefers to bank rather than to serve M-customers (except in period T). Likewise, conditions on p and τ can also ensure that serving R-customers is never optimal.

Proposition 1

- *If $\alpha\beta > 1$, it is never optimal to allocate resources to M-customers before the final time period.*

- *If $\alpha p + \tau < \alpha\beta$ or $\alpha p + \tau < 1$, serving R-customers is never optimal.*

- *If $\alpha\beta < 1$ and $\alpha p + \tau < 1$, it is optimal to use all assets to serve M-customers in the first period.*

(The proof is in Section 1 of de Véricourt and Lobo (2008).)

2.2 Optimal Resource Allocation

The interesting case is then $\alpha\beta < 1$ and $\alpha p + \tau > 1$, which we assume for the rest of this section. This also implies $\alpha p + \tau > \alpha\beta$. The following result shows that the concavity of the social-impact-to-go can be iteratively propagated to all time periods, and establishes that the optimal policy

for R-customer capacity plus banking is of threshold type. Capacity should only be allocated to serving M-customers when assets are above a threshold, and all assets above that threshold should be allocated for this purpose.

Theorem 1 *For each time period t, $v_t(\cdot)$ is differentiable non-decreasing concave and $J^{v_t}(y, z)$ is differentiable and jointly concave. Further, there exists a threshold a_t^* such that the optimal resource decision $y_t^*(a)$ and the optimal banking decision $z_t^*(a)$ satisfy $y_t^*(a) + z_t^*(a) = a_t^* \wedge a$ for all a, and $y_t^*(a) = y_t^*(a_t^*)$, $z_t^*(a) = z_t^*(a_t^*)$ for $a > a_t^*$.*

(The proof is in Section 1 of de Véricourt and Lobo (2008).)

Under the additional assumption that the profit margin is higher than the banking return rate, $p \geq \beta$, which should be a reasonable assumption for most problems, and when banking is not possible, we can further characterize the structure of the optimal policy in the newsvendor case

Corollary 1 *If $F_y(\cdot) = F(\cdot)$ for all y and $p \geq \beta$, the optimal banking decision $z_t^*(a)$ is non-decreasing in a. Further, if $\beta = 0$ the myopic policy is optimal: at time t, the optimal resource allocation is $y_t^*(a) = a^* \wedge a$, where*

$$a^* = F^{-1}\left(1 - \frac{1}{\tau + \alpha p}\right). \tag{5}$$

(The proof is in Section 1 of de Véricourt and Lobo (2008).)

When the time horizon is short enough, it can be further deduced from Corollary 1 that the organization should never bank. In de Véricourt and Lobo (2008), Section 2, we show that when $\tau = 0$ the optimal policy is of threshold type, myopic, and never involves banking for $t \geq T - n_0$ with $n_0 = \lfloor -\ln(\alpha\beta)/\ln(\alpha p) \rfloor$.

Finally, we extend the results to the infinite-horizon case.

Corollary 2 *The optimal policy for the infinite horizon case is of threshold type. Further, if $F_y(\cdot) = F(\cdot)$ for all y, when $p \geq \beta$ the optimal banking decision is non-increasing in the current assets, and when banking is not allowed the optimal policy is myopic.*

(The proof is in Section 1 of de Véricourt and Lobo (2008).)

3 Resource Allocation with Pricing

3.1 The Model

We have so far assumed that the organization proposes its service at a fixed price to the revenue customers. A nonprofit organization may, however, have market power in its revenue-generating

activities, and the operational flexibility to dynamically change prices. The organization needs to develop a pricing strategy jointly with its capacity-provision and banking strategy. Characterizing the optimal policy for these joint decisions with the general model of Section 2 appears, however, to be far more challenging. To be able to provide some insight into the pricing problem, we focus on the newsvendor case and make two key simplifications in this section. First, we consider the case where the organization does not rely on grants ($\Delta = 0$), and R-customers have no social impact ($\tau = 0$). The latter is a reasonable assumption if alternative for-profit providers for paying customers are readily available. Second, we assume that the organization never banks. This allows us to explore the structure of the optimal pricing and capacity-provision decisions in more detail, while still maintaining the more interesting characteristics of the problem.

We consider a demand function with multiplicative uncertainty, that is $D_t = \gamma(p_t)\Theta_t$, where $\gamma(\cdot) : [0, \overline{p}] \mapsto [0, +\infty[$ is the price response function and Θ_t is a random variable with finite mean. The response function $\gamma(\cdot)$ is differentiable and decreasing, with $\gamma(\overline{p}) = 0$. The price elasticity $e(\cdot)$ is defined as $e(p) = -p\gamma'(p)/\gamma(p)$. We will also refer to the revenue function, defined as $r(p) = p\gamma(p)$ which is assumed to be non-monotone and strictly concave in p. The maximand $\underline{p} = \operatorname{argmax}_p r(p)$ is then well defined. As before, $f(\cdot)$ denotes the probability density function of Θ_t, and $F(\cdot)$ and $G(\cdot)$ the corresponding cumulative and tail distributions, which in this section we assume twice differentiable.[9] We also introduce the generalized failure rate (GFR), defined as $g(\cdot) : [\underline{\theta}, \overline{\theta}] \mapsto \mathbf{R}^+$ such that

$$g(\theta) = \theta \frac{f(\theta)}{G(\theta)}. \tag{6}$$

At the beginning of each period, the organization sets the selling price p_t and decides how much service capacity y_t to provide for R-customers. The choice of y_t is limited by the current resources, which corresponds to the constraints $0 \leq y_t \leq a_t$. The remaining resources, $a_t - y_t$, are allocated to providing capacity to serve M-customers. Demand D_t is then realized, and the number of R-customers served by the organization is $y_t \wedge \gamma(p_t)\theta_t$. The resources available to the organization at the beginning of the following period are

$$a_{t+1} = p_t(y_t \wedge \gamma(p_t)\theta_t).$$

The goal is to determine a pricing and capacity-provision policy that maximizes the social return over the time horizon T.

---

[9]The assumption of twice-differentiability of F can be relaxed to differentiability by extending the limit argument in Lemma 3.

For $t < T$, the social-impact-to-go at period t can be shown to satisfy the optimality equations

$$v_t(a) = \max_{\substack{0 \le p \le \overline{p} \\ 0 \le y \le a}} a - y + \alpha H^{v_{t+1}}(p, y), \tag{7}$$

where the operator $H^v(p, y)$ is defined for any real-valued function $v(\cdot)$ as

$$H^v(p, y) = \mathbf{E}_\Theta \, v\left(py \wedge r(p)\Theta\right). \tag{8}$$

We denote by $(p_t^*(a), y_t^*(a))$ the optimal pricing and capacity decisions at period t given current assets a. The optimal policy $(p_t^*(a), y_t^*(a))$ corresponds to the maximizer of $\alpha H^{v_{t+1}}(p, y) - y$, subject to the constraints $p \ge 0$ and $0 \le y \le a$.

Solving (7), however, can be challenging, even for the single-period case (in the following we refer to the single-period problem as the case where $T = 2$, since no decision is made in the last period). When $T = 2$, the objective can be written as

$$v(a) = a + \max_{\substack{0 \le p \le \overline{p} \\ 0 \le y \le a}} \alpha p \mathbf{E}_\Theta \left(y \wedge \gamma(p)\Theta\right) - y, \tag{9}$$

and the optimization problem is equivalent to a newsvendor problem with pricing and with capacity constraint. The unconstrained newsvendor problem with pricing has been the subject of much research (see Petruzzi and Dada (1999) for a survey). More recently, Wang, Jiang, and Shen (2004) provide sufficient conditions for the uniqueness of the optimal decisions y^* and p^*. They show that if Θ has increasing generalized failure rate (*i.e.*, $g(\cdot)$ is increasing over $[\underline{\theta}, \overline{\theta}]$), then y^* and p^* exist and are unique. With the change of variable $z = y/\gamma(p)$ (the 'stocking factor') in Equation (9), they show that, for any z, an optimal $p(z)$ exists, and that the function

$$H(z) = \alpha \mathbf{E}_\Theta \, r(p)\left(z \wedge \Theta\right) - \gamma(p)z$$

is unimodal in z. However, nothing is said about the concavity of the objective function, even along the optimal price $p^*(z)$. Bernstein and Federgruen (2005) propose a sufficient condition on $g(\cdot)$ which guaranties that $H(\cdot)$ is log-concave and hence unimodal.

The analysis of multi-period dynamic problems typically requires stronger properties than unimodality or log-concavity. Most of the previous approaches to address dynamic joint pricing-inventory problems (for general probability distributions) rely on the joint concavity of $\alpha H^v(p, y) - y$ in (p, y) for any concave value function $v(\cdot)$ (see, for instance, Federgruen and Heching (1999)).

However, even for the single-period case, the py term in the definition of H^v in (8) is not jointly concave. Changes of variable such as $z = y/\gamma(p)$, or the conditions on the generalized failure rate $g(\cdot)$ proposed by Wang, Jiang, and Shen (2004) or Bernstein and Federgruen (2005) do not circumvent this problem. Sample path approaches as developed by Huh and Janakiraman (2005) cannot be applied to our model.

In the following, we assume that the generalized failure rate at the stocking factor $y/\gamma(p)$ is not smaller than the inverse of the price elasticity, that is $\forall p \in [\underline{p}, \overline{p}[$ and $\forall y$, $y/\gamma(p) \in [\underline{\theta}, \overline{\theta}]$,

$$g\left(\frac{y}{\gamma(p)}\right) \geq \frac{1}{e(p)}. \tag{10}$$

Note that, for $r(\cdot)$ concave, $e(p) \geq 1$ for $p \geq \underline{p}$, so that condition (10) holds when the generalized hazard rate is bounded below by one, *i.e.*, $g(\theta) \geq 1$. Condition (10) is similar to the condition for the single-period case proposed by Bernstein and Federgruen (2005), where the generalized hazard rate is bounded below by one half. Condition (10) is also neither more general nor more restrictive than the increasing generalized failure rate condition proposed by Wang, Jiang, and Shen (2004), since it allows for non-monotonic generalized failure rates. If Θ is uniformly distributed over $[\underline{\theta}, \overline{\theta}]$, Condition (10) holds for any concave revenue function and $\underline{\theta}/\overline{\theta} \geq 1/2$. Another example satisfying this condition for any concave revenue function is an exponential distribution with parameter λ truncated at $\theta = 1/\lambda$, that is $f(\theta) = 0$ for $\theta < 1/\lambda$ and $f(\theta) = \lambda e^{-\lambda(\theta - 1/\lambda)}$ for $\theta \geq 1/\lambda$. More generally, for any distribution defined over $[0, +\infty[$ with an increasing generalized failure rate, $g(\theta) > 1$ holds if the distribution is truncated at $\underline{\theta}$ such that $g(\underline{\theta}) > 1$, and scaled accordingly.

Condition (10) can be shown to be equivalent to the bound from below on elasticity of sales introduced by Kocabiyikoğlu and Popescu (2005). They show that if this condition holds, an optimal solution exists for the newsvendor problem with pricing and the corresponding optimal price is decreasing with the inventory level. In our context, the concavity and monotonicity properties can be propagated across time periods under this condition. We will show that if condition (10) holds, then $H^{v_t}(p, y)$ is concave along the optimal price, which implies that

$$\widetilde{H}^{v_t}(y) = \max_{0 \leq p \leq \overline{p}} H^{v_t}(p, y) \tag{11}$$

is concave in y. For any continuous increasing function $v(\cdot)$ such that $v(0) = 0$, $\widetilde{H}^v(\cdot)$ is well defined and $\widetilde{p}(y)$, the maximizer of $H^v(\cdot, y)$, is interior ($\widetilde{p}(y) \in]0, \overline{p}[$) since $H^v(0, y) = H^v(\overline{p}, y) = 0$ for any $y \in \mathbf{R}^+$. Concavity of $\widetilde{H}^{v_{t+1}}(\cdot)$ guaranties in turn that the social-impact-to-go function $v_t(\cdot)$ is increasing and concave and that the optimal policy is of threshold type. Furthermore, since $r(\cdot)$ is non-monotone strictly concave, $r(\cdot)$ and hence $H^{v_t}(\cdot, y)$ are increasing in p for $p < \underline{p}$ and the optimal price $\widetilde{p}_t(y)$ is bounded from below by $\underline{p}$ (the maximizer of $r(\cdot)$).

17

A threshold policy in our context is characterized by $T-1$ thresholds $(\widehat{p}_t, \widehat{a}_t)$, $t \in [1 \ldots T-1]$, such that the decisions $(p_t(a), y_t(a))$ given assets a at the beginning of period t are

$$y_t(a) = a \wedge \widehat{a}_t$$
$$p_t(a) = \widehat{p}_t \quad \text{if} \quad a \geq \widehat{a}_t,$$

and by a price $p_t(a)$ for $a < \widehat{a}_t$. Under a threshold policy, the organization tries to guarantee a particular service capacity for R-customers. Only when this threshold is assured are M-customers served.

Finding the optimal price, however, is not a trivial problem, and solutions can be counter-intuitive. One could indeed reasonably expect the optimal price to be decreasing in the capacity allocated to R-customers. However, a simple counter-example will show that this is not always the case (see de Véricourt and Lobo (2008), Section 3). We will see, however, that with condition (10) the optimal price is non-increasing.

3.2 Optimal Capacity and Pricing Strategy

Our proof of the structure of the optimal capacity-provision and pricing strategy relies on the first- and second-order derivatives of $\widetilde{H}^{v_t}(y)$. However, $\widetilde{H}^{v_t}(y)$ may not be twice differentiable due to the constraint $y \leq a$ in its definition. We introduce a family of unconstrained dynamic problems parameterized by $\epsilon > 0$. We show that the corresponding operators $\widetilde{H}^{v_t^\epsilon}(y)$ are concave, and likewise for the optimal value functions v_t^ϵ. We then show that, for all t, $v_t^\epsilon \to v_t$ pointwise when $\epsilon \to 0$, where the $v_t(\cdot)$ satisfy the optimality equations.

Consider $\mathbf{R}_{-\infty} = \mathbf{R} \cup \{-\infty\}$ and the extension of any function $\varphi(\cdot) :]\underline{x}, \overline{x}[\mapsto \mathbf{R}$ such that $\varphi(x) = -\infty$ when $x \notin]\underline{x}, \overline{x}[$. (In the following we use the same notation for a given function and its extension.) For instance, we will consider the extension of the logarithm function such that $\varphi(x) = -\infty$ when $x \leq 0$. The family of unconstrained problems is obtained by omitting the constraints $0 \leq y \leq a$, and introducing logarithmic barrier functions in the objective function.

More precisely, for any $\epsilon > 0$, we consider the dynamic problem,

$$v_t^\epsilon(a) = \max_y \left\{ a - y + \epsilon \log((a-y)y) + \alpha \widetilde{H}^{v_{t+1}^\epsilon}(y) \right\}, \quad t = 1, \ldots, T-1 \tag{12}$$
$$v_T^\epsilon(a) = a. \tag{13}$$

The optimal capacity decision never equals the bounds ($0 < \widetilde{y}^\epsilon(a) < a$). Backward iteration then shows that $v_t^\epsilon(a)$ is non-decreasing and twice differentiable.

Using condition (10) on the generalized failure rate, we show in de Véricourt and Lobo (2008), Section 4, that $\widetilde{H}^{v_t^\epsilon}(\cdot)$ is concave if $v_t^\epsilon(\cdot)$ is also concave.

Lemma 1 *Assume that condition (10) holds. For any $\epsilon > 0$, if $v_{t+1}^\epsilon(\cdot)$ is twice differentiable, non-decreasing, and concave, then $\widetilde{H}^{v_{t+1}^\epsilon}(\cdot)$ is twice differentiable and concave. Furthermore, the pricing decision $\widetilde{p}_t^\epsilon(y)$ maximizing $H^{v_{t+1}^\epsilon}(p, y)$ is non-increasing in y.*

It follows that equation (12) preserves concavity, which we establish in the next result, with proof also given in (de Véricourt and Lobo 2008).

Lemma 2 *Assume that condition (10) holds. For any $\epsilon > 0$, if $v_{t+1}^\epsilon(\cdot)$ is twice differentiable, non-decreasing, and concave, then $v_t^\epsilon(\cdot)$ is also twice differentiable, non-decreasing, and concave. Also, the optimal pricing decision $p_t^{*\epsilon}(a)$ is non-increasing in the current assets a.*

Starting from $v_T^\epsilon(a) = a$ (which is concave), Lemma 2 ensures that $v_t^\epsilon(\cdot)$ is concave for all t. By letting $\epsilon \to 0$, we obtain the same result for $v_t(\cdot)$ in the original problem (also proved in de Véricourt and Lobo (2008), Section 4).

Lemma 3 *If (10) holds, then for $\epsilon \to 0$ the $v_t^\epsilon(\cdot)$ converge pointwise to the $v_t(\cdot)$ which solve the optimality equations (7). Further, the $\widetilde{H}^{v_t}(\cdot)$ and $v_t(\cdot)$ are concave and the optimal pricing decision $\widetilde{p}_t(y)$ is non-increasing in the capacity decision y.*

The proof of the following theorem, which characterizes the optimal capacity and pricing decisions, is then immediate.

Theorem 2 *If condition (10) holds, the optimal policies are of threshold type and the optimal pricing decisions $p_t^*(\cdot)$ are non-increasing in the current assets.*

In de Véricourt and Lobo (2008), Section 5, we provide additional information regarding the shape of the optimal pricing policy by showing that $ap_t^*(a)$ (the total revenue if all assets are allocated to R-customers and demand exceeds capacity) is decreasing in a, and that $\lim_{a \to 0} ap_t^*(a) = 0$. We conclude this section by noting that our results extend to the infinite horizon case.

Corollary 3 *If condition (10) holds, the optimal policy for the infinite horizon case is of threshold type. Further, the optimal pricing decision $p^*(\cdot)$ is non-increasing in the current assets*

4   Numerical Studies

We consider an example with a concrete choice of parameters. This choice of parameters is suggested from Aravind to provide a sense for their order of magnitude, but in no way intended to be exact in the actual parameter values. We use a time period of 2 months ("Our staff are rotated between

the paying segment and the free segment every month or two months", R. D. Thulasiraj, Executive Director, Aravind Eye Hospitals, in interview to IIMB Management Review, Sep. 2004) and, somewhat arbitrarily, a time horizon of 4 years, so that $T = 24$. We assume a social discounting factor of 0.75 on a yearly basis, which leads to $\alpha = 0.953$ (the social discounting factor is a subjective choice that the organization's leadership will have to settle on for planning purposes; in Aravind's case, since the time period in the model is relatively short and cataracts progress relatively slowly, the subjective intertemporal social discounting will not be too high). The factor for the banking of assets is 1.10 on a yearly basis, so that $\beta = 1.016$.

The following are approximations consistent with values indicated in the cited literature on Aravind and its reports. For the fixed-price model, the price charged to paying customers for eye surgery is set at 2000 rupees. The cost of providing surgery is 1000 rupees in paying hospitals, and 500 rupees in free hospitals. The mean demand from paying customers over a two-month period is six thousand, with a uniform distribution between four and eight thousand (so that, if all assets are dedicated to revenue customers, the mean demand can be fully served when operating assets of six million rupees are available for a two-month period).

With the assumption that $\tau = 0$ (which is justified if paying customers have other easily accessible alternatives), we normalize units as follows. We define one unit of customers to be 4,000 patients, and one unit of assets to be 4,000,000 rupees. We then have $c_R = 1,000$ rupees/customer $= 1$ asset units/customer units. Likewise, after this normalization, $c_M = 0.5$, $p = 2$, and the demand distribution is uniform in $[1, 2]$. We can then arbitrarily choose the units of social impact so that $s = 0.5$ in order to have $s/c_M = 1$. We ran most simulations for initial assets ranging from near zero, $a = 0$, to over six times the mean demand, $a = 10$ (40 million rupees).

4.1 The Value of Optimal Capacity Decisions

We begin by considering the value of optimal capacity decisions, by considering the case where the option of banking is not available and the price is fixed. Note that the distribution satisfies the conditions for Theorem 2. The optimal policy for this problem is of threshold type, as shown in corollary 1.

A threshold policy may appear somewhat counter-intuitive to many organizations. To our knowledge, no empirical studies exist on how resources are actually allocated in practice between mission and revenue customers in nonprofit organizations. However, strategies based on proportional allocations may be more appealing to managers. Such a policy, where the same fixed fraction of assets is allocated to mission customers in every period is attractive in other respects. Under such

policies, a fixed percentage of the current assets is allocated to the M-customers, so that the organization always contributes to its mission. In contrast, under a thresholds policy, no M-customers are served if $a_t \leq a^*$. This may, in the long run, jeopardize the culture of the organization as nonprofit oriented, which in turn can affect its ability to recruit volunteers and to provide high service quality to the M-customers, among other concerns (Dees 1998; Dees and Anderson 2003). However, if the gain is significant, the organization may benefit from obtaining buy-in from its stakeholders on the benefit of a threshold policy, while acting in other ways to maintain a nonprofit oriented culture.

We take a look at the tradeoffs involved by comparing the performance of the optimal policy with a proportional-allocation policy. Figure 1 plots the value function for three different y/a ratios. The best choice for y/a, as measured by the expected discounted social impact, depends on the initial endowment. We plot the maximum of the value functions associated with all possible choices for the fixed proportion. This corresponds to expected value when the organization optimally makes the decision at the beginning of the first period on the proportion based on the asset level at that time.

Figure 2 compares the optimal policy with the maximum over all proportional policies. The relative difference between the two curves is substantial, even for a large (at $a = 3$, or 12 million rupees, two times the assets required to provide capacity for twice the mean demand of R-customers, it is on the order of 30%). Near zero it becomes extremely large (higher than 100% for $a < 0.4$, or 1.6 million rupees). When the initial asset level is low, the optimal policy allows the organization to grow more aggressively and more quickly reach the asset level where it can fully exploit the available demand from revenue customers.
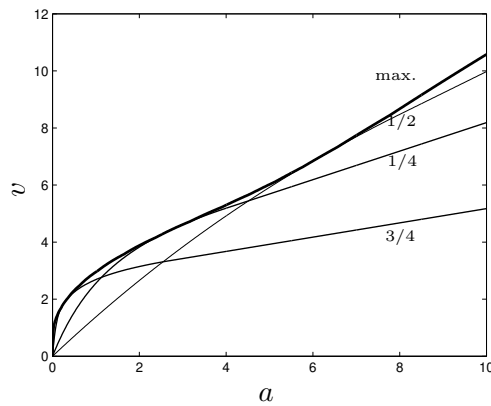

Figure 1: Capacity provided for each class of customers. Value functions for fixed-proportion policy with different ratios, and maximum over all ratios.
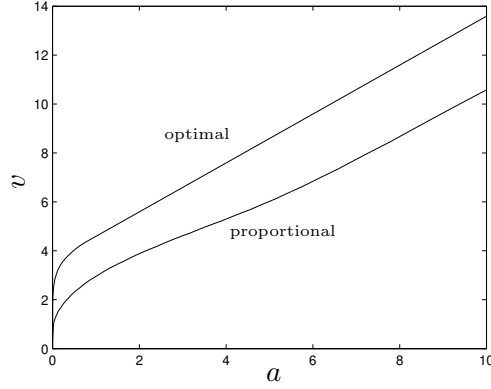
Figure 2: Capacity provided for each class of customers. Max. of value functions for fixed-proportion policies, and value function for optimal policy.

4.2 The Value of Banking

Figure 3 shows the optimal policy with a fixed price and banking decision. In agreement with Theorem 1, the optimal policy is of threshold type for $z + y$, and z is non-decreasing. In our numerical examples we have also always found y to be non-decreasing.

For low asset levels, the risk of unmet demand is low and the organization's focus is on growing the available cash, with all assets allocated to serving revenue customers. Banking becomes attractive when the organization has more assets and the expected marginal value of increasing service capacity for R-customers decreases.

The value of banking depends on the price being charged. To investigate this we compute the value function for different prices, holding constant the return on the banked assets. We also compute, for the same set of prices, the value function without banking. For each price, we compute the ratio of the two value functions, and the maximum of this ratio over all possible initial states. Figure 4 plots this maximum ratio as a function of the price. Banking is of relatively limited value, providing a maximum gain of about 4% or less.

The expected gain from having the banking decision available is not monotonic in the price. For small prices it is optimal to spend all the capital on M-customers in the first period. The option of banking is therefore of no value. For higher prices (and with the same demand distribution) the return rate from banking becomes, by comparison, less attractive. The organization has less of an incentive to hedge the outcome by banking since, given the high profit margin, it is easier to recover quickly from a period with unusually low demand. For very large prices, the percentage gain was seen to converge to a non-zero value.

22

On the other hand, note that the value of the option of taking a loan, would be substantial for low asset-levels. In particular, the value function would no longer be zero at zero net capital.
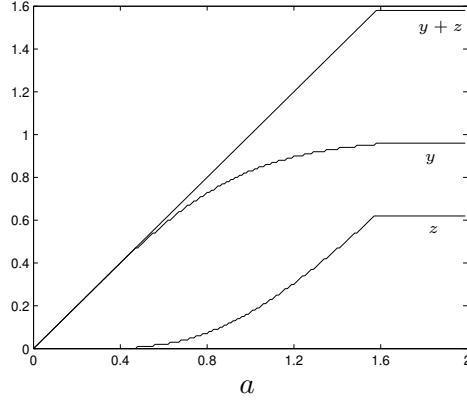
Figure 3: Optimal policy with banking.
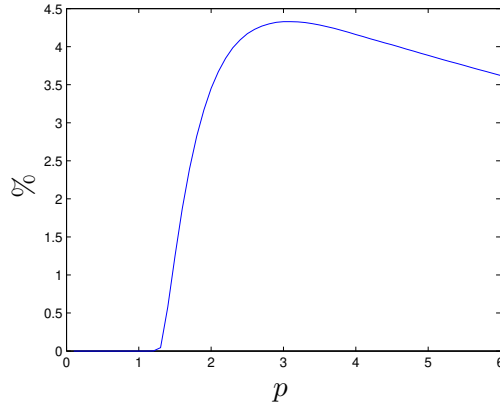
Figure 4: Maximum percentage expected gain from banking over all possible initial asset levels. This maximum ratio of the two value functions is plotted over a range of prices (and same demand distributions).

4.3 The Value of Pricing Decisions

Finally, we investigate the value of optimal pricing decisions, using the model of Section 3 which excludes the option of banking. The parameters used are otherwise as above, with the demand scaling as $4 - p$ (with this demand function the 'monopoly price' is two thousand rupees and the 'shutdown price' is four thousand rupees). The structure of the optimal policy is consistent with Theorem 2. The optimal capacity policy is of threshold type, and the optimal pricing policy is

23

non-increasing.

We compare the optimal policy with a policy which charges a fixed price of 2000 rupees. Recent articles investigating different models in for-profit firms have found dynamic pricing to provide small gains in expected revenue (usually 5% or less; see for instance Gallego and van Ryzin (1994) and Chen, Wu, and Yao (2004)). In this case, and as seen in Figure 5, the difference in expected value is high for a low asset level, when there will be a large amount of unmet demand if the organization charges a fixed price more adequate for periods in which the asset level is higher. When the assets are $a < 0.1$ (less than $400,000$ rupees), the percent gain quickly rises above 5%. While dynamic pricing in the sense of revising the price in each time period may be too complex and costly to be warranted, this points to the importance of considering capacity when setting the price, and therefore the importance of periodically revising the price as the available resources change. (If loans were available at a reasonable rate there would be little or no value in periodic price revisions, since the optimal policy when the asset level is low would be to borrow capital in order to provide optimal capacity for the given demand distribution.) For mature organizations with an operation scaled to match the available demand, the asset level is likely to most often be above the threshold, where the expected gains from using dynamic pricing are small. The larger expected gains from using dynamic pricing when the asset level is small are of interest for organizations at the startup stage, and for small organizations with a high variability of demand. Aravind, either in its early years or when setting up a new hospital in an under-served area, seems to fit this profile.
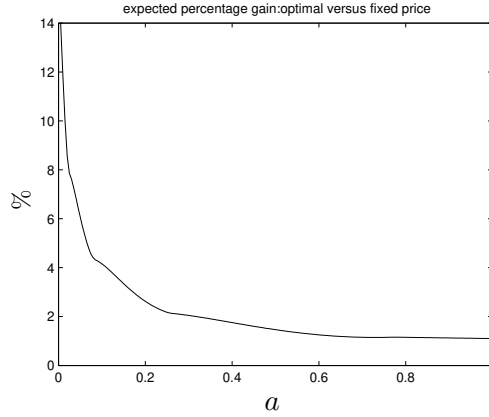

Figure 5: Ratio of the value functions with dynamic price and with fixed price.

24

5 Extensions and Future Research

5.1 Demand Distribution and Demand Function

Condition (10) is used in the proofs only *at the optimal price* for each ϵ problem. If we assume that for all a and for ϵ sufficiently small

$$g\left(\frac{a}{\gamma(p^*(a))}\right) \geq \frac{1}{e(p^*(a))}, \tag{14}$$

then Theorem 2 still holds. This should be the case if (14) holds with strict inequality for $\epsilon = 0$. One interesting direction in which to extend our results is to find sufficient conditions on $f(\cdot)$ guaranteeing (14). Note, however, that establishing (14) for more general $f(\cdot)$ requires propagating properties stronger than concavity on the value function.

Also, we have restricted our analysis of the problem with pricing to the multiplicative demand case. For more general demand functions of the form $D = f(p, \Theta)$, the elasticity of sales, following Kocabiyikoğlu and Popescu (2005), is given by

$$\varepsilon(p, y) = p\frac{L'(p, y)}{1 - L(p, y)},$$

where $L(p, y) = P(D \geq y)$, with L' the derivative of L with respect to p. Condition (10) then becomes $\varepsilon(p, y) \geq 1$. This condition can be used to extend our results to the case of more general demand functions.

5.2 Objective Functions

We assumed a linear social return $s\,x$, where x is the capacity allocated to M-customers. This is a good model if there are no significant economies or diseconomies of scale, and if the size of the unmet need being addressed is large compared to the service capacity that the organization can potentially provide. A natural extension of our model is to the case of decreasing marginal social return, where the impact of serving x mission customers is described by a concave function $s(\cdot)$. A case of interest corresponds to random demand from M-customers. If this demand is given by a random variable X, the organization contributes to its mission in proportion to $s(x) = \mathbf{E}_X\,(x \wedge X)$. The optimal policy will no longer be of threshold type. It should however be possible to derive simple and effective heuristics for this case drawing from the results presented here.

5.3 Flexible Capacity

We assumed that the service capacity for R-customers needs to be committed at the beginning of each period, before demand is realized. In this section we briefly explore the case where the

organization has the flexibility to allocate its service capacity between the two classes of customers after the actual demand in the current period is known. For instance, in a dual-hospital operation like Aravind it may in some cases be possible to share capacity. This would entail, among other things, moving staff across hospitals mid-way through their two-month assignments, and might in some cases require postponing the surgery of free patients already scheduled.

We show that the structure of the optimal policy is in many ways similar to our previous results. However, the idea of a threshold does not apply. The optimal policy is to, in each period, serve all R-customers first. This can be interpreted as essentially same idea as before: prioritizing R-customers will allow for greater mission impact. Note that, to the extent that sharing of capacity is possible, the structure of the optimal policy implies that it is desirable to do so. When dynamic pricing is allowed we show the optimal price to be decreasing in the asset level, albeit under a somewhat more restrictive condition on the demand distribution. Both with and without dynamic pricing, we show decreasing marginal mission impact relative to the available resources.

Without loss of generality, we again assume that $c_R = c_M = s = 1$. At the beginning of each period, the organization makes an investment decision and chooses how much total service capacity $c_t \geq 0$ to reserve and sets the amount to be banked, $z_t \geq 0$. The choices of c_t and z_t are limited by the current resources, i.e., $z_t + c_t = a$. Demand is then realized and the organization must decide how to allocate the capacity c_t, subject to $c_t = x_t + y_t$, where x_t and y_t are the number of M-customers and R-customers served. The optimality equations are now

$$v_t(a) = \max_{\substack{0 \leq c, \ 0 \leq z, \\ c + z = a}} \mathbf{E}_\Theta \max_{\substack{0 \leq x, \ 0 \leq y, \\ x + y = c}} x + \tau(y \wedge \Theta) + \alpha v_{t+1}\left(p(y \wedge \Theta) + \beta z\right), \quad (15)$$

with $v_T(a) = a$. This equation can be simplified by writing $c = a - z$ and $x = c - y = a - y - z$, so that

$$v_t(a) = a + \max_{0 \leq z \leq a} \mathbf{E}_\Theta \max_{0 \leq y \leq a - z} J^{v_{t+1}}(y, z, \Theta), \quad (16)$$

where the operator J^v is defined as

$$J^v(y, z, \theta) = -z - y + \tau\left(y \wedge \theta\right) + \alpha v\left(p(y \wedge \theta) + \beta z\right). \quad (17)$$

Using a similar approach, Proposition 1 can be shown to hold when the service capacity is shared. Again, the interesting case is $\alpha\beta < 1$ and $\alpha p + \tau > 1$. The following result characterizes the optimal service-capacity decision $\widetilde{y}_t(c_t, \theta_t)$ given the capacity choice c and the total number of R-customers θ_t.

Proposition 2 *With a fixed price and shared capacity, and given the resources allocated to providing service capacity, it is optimal to serve as many R-customers as possible,*

$$\widetilde{y}_t(c_t, \theta_t) = c_t \wedge \theta_t.$$

The proof is in de Véricourt and Lobo (2008), Section 6. The optimal capacity decision is an infinite threshold policy, and Proposition 2 allows us to simplify the optimality equations to

$$v_t(a) \;=\; a + \max_{0 \leq z \leq a} \; \mathbf{E}_\Theta J^{v_{t+1}}((a - z) \wedge \Theta, z, \Theta). \tag{18}$$

The following result states that the marginal social return is decreasing. The corresponding optimal banking decision $z_t^*(a)$, defined as the minimum optimal decision in case of multiple optima, is not necessarily of threshold type since the constraint a appears in the maximand in (18). The optimal decision $z_t^*(a)$ is monotone in the assets if $p \geq \beta$ and $\alpha(p - \beta) + \tau \geq 1$ (both hold when p sufficiently large compared to β, *i.e.*, the profit margin is sufficiently larger than the return rate from banking).

Proposition 3 *For each time period t, $v_t(\cdot)$ is differentiable non-decreasing concave. Further, the optimal banking decision $z_t^*(a)$ is non-decreasing in a if $p \geq \beta$ and $\alpha(p - \beta) + \tau \geq 1$.*

The proof is in de Véricourt and Lobo (2008), Section 6. We now consider the corresponding problem with dynamic pricing. Following the approach in Section 3, consider a demand function with multiplicative uncertainty, $D_t = \gamma(p_t)\Theta_t$. When banking is ignored ($\beta = 0$) and R-customers do not have mission impact ($\tau = 0$), the optimality equations are

$$
\begin{aligned}
v_t(a) \;&=\; a + \max_{0 \leq p \leq \overline{p}} \; \mathbf{E}_\Theta J^{v_{t+1}}(a \wedge \gamma(p)\Theta, 0, \gamma(p)\Theta) \\
&=\; a + \max_{0 \leq p \leq \overline{p}} \; -\mathbf{E}_\Theta[a \wedge \gamma(p)\Theta] + \alpha H^{v_{t+1}}(p, a), \tag{19}
\end{aligned}
$$

where $H^v(\cdot, \cdot)$ is defined as in (8). As before, we show that the optimal price is interior (in $]0, \overline{p}[$) and $\widetilde{J}^v(a) = \max_{0 \leq p \leq \overline{p}} \mathbf{E}_\Theta J^v(p, a)$ is twice differentiable. Also, since $r(\cdot)$ is decreasing for $p \geq \underline{p}$, we have $r'(\cdot) \leq 0$ at the optimal price. As for Proposition 1, a sample-path argument shows that the optimal price satisfies $\alpha p > 1$. In order to characterize the optimal policy when capacity is not flexible, Condition (10) was needed. When capacity is flexible, the corresponding condition is: $\forall p \in [\underline{p}, \overline{p}[$ and $\forall a, a/\gamma(p) \in [\underline{\theta}, \overline{\theta}]$,

$$g\left(\frac{a}{\gamma(p)}\right) \geq \frac{\alpha p}{\alpha p - 1} \cdot \frac{1}{e(p)}. \tag{20}$$

For instance, if $\alpha \underline{p} = 2$, the condition is implied by $g(a/\gamma(p)) \geq 2/e(p)$. Condition (20) implies condition (10) of the inflexible capacity case and is therefore more restrictive.

Theorem 3 *If (20) holds, then $v_t(\cdot)$ is increasing concave. Further, the optimal price is non-increasing.*

See de Véricourt and Lobo (2008), Section 6, for a proof.

6   Conclusion

The success of Aravind, which has dramatically improved the quality of life of hundreds of thousands of people, demonstrates the potential of nonprofit organizations that generate resources by also running for-profit operations. The problem faced by such ventures differs in fundamental ways from the profit-maximization problem of commercial ventures and, to our knowledge, we provide the first analytical approach for examining resource-management issues in this context.

We have investigated how a nonprofit organization should dynamically allocate its assets over time between its revenue-generating activities and its mission, in order to maximize the organization's social impact. The optimal capacity allocation policy is of threshold type. Over all periods, this policy allows the organization to have a higher expected social impact by serving more mission customers. While this needs to be balanced with other considerations, such as maintaining the organization's mission culture which might be jeopardized if the organization only serves revenue customers for several consecutive periods, the analysis allows us to quantify the tradeoff. Theoretical analysis and numerical studies suggest that there is limited value in the option of banking assets from one period to the next, but point to the importance of adjusting the price charged to revenue customers depending on whether assets are scarce or abundant.

Different nonprofit organizations face different problems and require different modeling approaches and, as a consequence, there is a need to add to the literature in several directions. We believe the approach proposed in this paper to be a productive initial framework to pursue such work.

Acknowledgments

References

Alter, S. K. (2000). *Managing the Double Bottom Line: A Business Planning Guide for Social Enterprises.* Alter Paperback.

Anderson, B. B., J. G. Dees, and J. Emerson (2002). *Developing Viable Earned Income Strategies in Strategic Tools for Social Entrepreneurs: Enhancing the Performance of Your Enterprising Nonprofits.* John Wiley & Sons.

Babich, V. and M. J. Sobel (2004). Pre-ipo operational and financial decisions. *Management Science 50*(7), 935–948.

Bernstein, F. and A. Federgruen (2005). Decentralized supply chains with competing retailers under demand uncertainty. *Management Science 51*(1), 18–29.

Billingsley, P. (1995). *Probability and Measure* (3 ed.). John Wiley & Sons.

Bitran, G. and R. Caldentey (2003). An overview of pricing models and revenue management. *Manufacturing and Service Operations Management 5*, 203–229.

Bonnans, F. and A. Shapiro (2000). *Perturbation analysis of optimization problems.* Springer.

Boschee, J. (2001). *The Social Enterprise Sourcebook: Profiles of Social Purpose Businesses Operated by Nonprofit Organizations.* Northland Institute.

Brinckerhoff, P. C. (2000). *Social Entrepreneurship: The Art of Mission-Based Venture Development.* John Wiley & Sons.

Buzacott, J. and R. Zhang (2001). Inventory management with asset-based financing. *Management Science 50*(9), 1274–1292.

Chen, H., O. Wu, and D. Yao (2004). Optimal pricing and replenishment in a single-product inventory system. Technical report, IEOR Columbia.

Chen, X. and D. Simchi-Levi (2004). Coordinating inventory control and pricing strategies with random demand and fixed ordering cost: The finite horizon case. *Operations Research 52*, 887–896.

Chen, Y., S. Ray, and Y. Song (2003). Optimal pricing and inventory control policy in periodic-review systems with fixed ordering cost and lost sales. Technical report, The Chinese Univ. of Hong Kong.

de Véricourt, F. and M. Lobo (2008). Online supplement to "resource and revenue management in nonprofit operations". Technical report, Duke University.

Dees, G. (1998). Enterprising non-profit. *Harvard Business Review, January-February.*

Dees, G. (2004). *in Generating and Sustaining Nonprofit Earned Income: A Guide to Successful Enterprise Strategies, S. M. Oster, C. Massarsky and S. Beinhacker*, Chapter : Putting Nonprofit Business Ventures in Perspective. Jossey-Bass.

Dees, G. and B. Anderson (2003). Sector-bending: Blurring the lines between nonprofit and for-profit. *Society 40*(4).

Elmaghraby, W. and P. Keskinocak (2003). Dynamic pricing in the presence of inventory considerations: Research overview, current practices and future directions. *Management Science 10*(7), 1287–1309.

Elstrodt, H.-P., A. Schindler, and A. C. Waslander (2004). Helping nonprofits do business. *The McKinsey Quarterly, Web exclusive December.*

Federgruen, A. and A. Heching (1999). Combined pricing and inventory control under uncertainty. *Operations Research 47*(3), 454–475.

Foster, W. and J. Bradach (2005, February). Should nonprofits seek profits? *Harvard Business Review.*

Gallego, G. and G. van Ryzin (1994). Optimal dynamic pricing of inventories with stochastic demand over finite horizons. *Management Science 40*, 999–1020.

Gruber, R. E. and M. Mohr (1982, Spring). Strategic management for multiprogram nonprofit organizations. *California Management Review 3*, 15–22.

Heyman, D. and M. Sobel (1984). *Stochastic Models In Operations Research (II).* Dover Publications.

Huh, W. and G. Janakiraman (2005). Optimality results in inventory-pricing control: An alternative approach. Technical report, IEOR Columbia.

Keeler, E. and S. Cretin (1984). Discounting of life-saving and other nonmonatory effects. *Management Science 29*(3), 300–306.

Klausner, M. (2003, Spring). When time isnt money: Foundation payouts and the time value of money. *California Management Review 1*(1), 50–59.

Kocabiyikoğlu, A. and I. Popescu (2005). Joint pricing and revenue management with general stochastic demand. Technical report, INSEAD.

Larson, R. and V. Forth (2002). *The Essential Guide to Starting a Moneymaking Business in Your Nonprofit Organization.* Amherst H. Wilder Foundation.

Li, L., M. Shubik, and M. J. Sobel (2003). Control of dividends, capital subscriptions, and physical inventories. Technical report, Yale University and Case Western Reserve University.

McCardle, K. (2005, October). Pro bono o. r.: O. r. for the public good. *OR/MS Today 32*(6), 32–36.

Modigliani, F. and H. Miller (1958). The cost of capital, corporation finance, and the theory of investment. *American Economic Review 48*, 261–267.

Oster, S. M. (1995). *Strategic Management for Nonprofit Organizations: Theory and Cases.* Oxford University Press.

Oster, S. M., C. Massarsky, and S. Beinhacker (1995). *Generating and Sustaining Nonprofit Earned Income: A Guide to Successful Enterprise Strategies.* Jossey-Bass.

Petruzzi, N. C. and M. Dada (1999). Pricing and the newsvendor problem: A review with extensions. *Operations Research 47*(2), 183–194.

Rangan, V. K. (1994, May). The aravind eye hospital madurai, india: In service for sight. *Harvard Business School Case* (9-593-098).

Robinson, A. (2002). *Selling Social Change (Without Selling Out): Earned Income Strategies for Nonprofits.* Jossey-Bass.

Shah, J. and L. S. Murty (2004). Compassionate, high quality health care at low cost: The aravind model; in conversation with dr g venkataswamy and r d thulasiraj. *IIMB Management Review.*

Skloot, E. (1983, January-February). Should nonprofits seek profits. *Harvard Business Review 83*(1).

Skloot, E. (1987). *The Nonprofit Sector: A Research Handbook*, Chapter : Enterprise and Commerce in Nonprofit Organizations. Yale.

Skloot, E. (1988, January-February). Should not-for-profits go into business? *Harvard Business Review 83*(1).

Steckel, R., R. Simons, and P. Lengsfelder (1989). *Filthy Rich and Other Nonprofit Fantasies: Changing the Way Nonprofit Do Business in the 90s*. Ten Speed Press.

Wang, Y., L. Jiang, and Z. Shen (2004). Channel performance under consignment contract with revenue sharing. *Management Science 50*(1), 34–47.

Weisbrod, B. (1998). *To Profit or Not to Profit: the Commercial Transformation of the Nonprofit Sector*. Cambridge University Press.

Weisbrod, B. A. (2005, Winter). The pitfalls of profits. *Stanford Social Innovation Review 2*(3), 40–47.

Young, D. Y. (2002). *in The State of Nonprofit America*, Chapter : Commercialization, Social Ventures, and For-Profit Competition, Lester M. Salamon. Brookings Institution Press.

Young, D. Y. (2004). *in Generating and Sustaining Nonprofit Earned Income: A Guide to Successful Enterprise Strategies, S. M. Oster, C. Massarsky and S. Beinhacker*, Chapter : Deploying Resources Effectively. Jossey-Bass.

Online Supplement to "Resource and Revenue Management in Nonprofit Operations"

Francis de Véricourt

Fuqua School of Business, Duke University, fdv1@duke.edu

and

Miguel Sousa Lobo

Fuqua School of Business, Duke University, mlobo@duke.edu

1 Proofs for Section 2

Proof of Proposition 1

Proof. We show the first part with a sample-path argument. Suppose a policy is optimal, and it is such that $x_t = a_t - y_t - z_t$ M-customers are served at time t, with $x_t > 0$. Denote by x_s, y_s and z_s the optimal decisions for $s = 1, \ldots, T$. Consider now an alternative policy, with decisions $\bar{x}_s$, $\bar{y}_s$, $\bar{z}_s$ such that $\bar{x}_s = x_s$ for $s \neq t, T$ and $\bar{x}_t = 0$, $\bar{x}_T = x_T + \beta^{T-t} x_t$. This corresponds to $\bar{y}_s = y_s$ for all s, $\bar{z}_s = z_s$ for $s = 1, \ldots, t-1$ and $\bar{z}_s = z_s + \beta^{s-t} x_t$ for $s = t, \ldots, T-1$. Note that before time t the policies are identical. With the new policy, for any realization of the Θ_s and Δ_s, $s = 1, \ldots, T$, the objective is improved by $x_t((\alpha\beta)^{T-t} - 1) > 0$. The second and third parts can be shown with a similar argument.

■

Proof of Theorem 1.

Proof. We show the first part of the theorem by iteration. $v_T(a) = a$ is differentiable non-decreasing concave. Assume that $v_{t+1}(\cdot)$ is differentiable non-decreasing concave. After the change of random variable $\Theta = F_y^{-1}(U)$ (a.s.) where U is uniformly distributed over $[0,1]$,

$$\begin{aligned}
H^{v_{t+1}}(y,z) &= \mathbf{E}_\Delta \mathbf{E}_U \, v_{t+1}\left(p\left(y \wedge F_y^{-1}(U)\right) + \beta z + \Delta\right) \qquad (21) \\
&= E_\Delta\left[\int_0^{F_y(y)} v_{t+1}\left(pF_y^{-1}(u) + \beta z + \Delta\right) du + v_{t+1}\left(py + \beta z + \Delta\right) G_y(y)\right].
\end{aligned}$$

$H^{v_{t+1}}(y,z)$ is differentiable from the differentiability of $v_{t+1}(\cdot)$, and of $F_y(\theta)$ and $F_y^{-1}(u)$ with respect to y. The concavity of $H^{v_{t+1}}(y,z)$ follows from the concavity of $F_y^{-1}(u)$ in y and from $v_{t+1}(\cdot)$ increasing concave. Likewise, we see that $J^{v_{t+1}}(y,z)$ is differentiable and jointly concave,

33

so that $v_t(\cdot)$ is the sum of a linear function and the maximum of a differentiable concave function subject to convex constraints which are increasing in a. As a result, $v_t(a)$ is non-decreasing, concave (Heyman and Sobel 1984), and differentiable (Bonnans and Shapiro 2000, Theorem 4.16). Since $J^{v_{t+1}}(y, z)$ is jointly concave we can define (y_t^*, z_t^*), the optimum of $J^{v_{t+1}}(y, z)$. The second part of the theorem follows directly from the concavity of $J^{v_{t+1}}(y, z)$ and the linear constraint $0 \le y + z \le a$ in (1).

∎

Proof of Corollary 1.

Proof. We first prove the monotonicity. When $a > a_t^*$ the optimal banking decision is constant and equal to z_t^*. When $a < a_t^*$, the optimal decision verifies $y_t^*(a) = a - z_t^*(a)$. The first-order condition $\frac{d}{dz} J^{v_{t+1}}(a - z, z) = 0$ is equivalent to

$$E_\Delta \left[\alpha \beta \int_{\underline{\theta}}^{a-z} v_{t+1}'(p\theta + \beta z + \Delta) f(\theta) d\theta + \left[-\tau + \alpha(\beta - p) v_{t+1}'((\beta - p)z + pa + \Delta) \right] G(a - z) \right] = 0. \tag{22}$$

The first term of the expectation is increasing in a and decreasing in z since $v_{t+1}'(\cdot)$ is a positive decreasing function. For the same reason, and for $p > \beta$, the second term is also increasing in a and decreasing in z. We must then have that $z_t^*(a)$ increases in a.

To show that the optimal policy is myopic, assume a^* is the optimal threshold at $t + 1$. The threshold at time t is the maximizer of $J^{v_{t+1}}(y, 0)$. From Theorem 1, $v_{t+1}(\cdot)$ is differentiable and the first-order condition yields

$$E_\Delta \left[\tau + \alpha p v_{t+1}'(py + \Delta) \right] G(y) = 1. \tag{23}$$

Since $p > 1$ (from $\alpha p > 1$), we have that $pa^* > a^*$, and therefore $v_{t+1}'(pa^* + \Delta) = 1$ for any non-negative Δ (since a^* is the optimal threshold at time $t + 1$). It follows that a^* as defined in (5) satisfies equation (23). Since $v_T'(a) = 1$, a^* is the optimal threshold at time $T - 1$ and the result follows by backward iteration.

∎

Proof of Corollary 2.

Proof. In the following, the value function is indexed by the number of periods to go $n = T - t$. In this notation, $v_0(\cdot)$ is the social impact in the last period, $v_0(a) = a$. Denote by $\overline{\Delta}$ the expected value of Δ. Since $E_y(\Theta)$ is bounded by $\overline{\Theta}$, $H^{v_0}(y, z) \le p\overline{\Theta} + \beta z + \overline{\Delta}$, and

$$\max_{y \ge 0, z \ge 0} J^{v_0}(y, z) \le (\tau + \alpha p)\overline{\Theta} + \alpha \overline{\Delta} + \max_{z \ge 0} (\alpha \beta - 1)z + \max_{y \ge 0} -y.$$

34

Since $\alpha\beta < 1$, this is bounded above by a finite $M = (\tau + \alpha p)\overline{\Theta} + \alpha\overline{\Delta}$. It follows that $v_1(a) \le a + M$. For $n = 2$ we then have

$$\begin{aligned}
\max_{y \ge 0, z \ge 0} J^{v_1}(y, z) &\le \max_{y \ge 0, z \ge 0} \tau\,\overline{\Theta} + \alpha\mathbf{E}_\Delta\mathbf{E}_\Theta\left(p\left(y \wedge \Theta\right) + \beta z + \Delta + M\right) - y - z \\
&\le (1 + \alpha)M,
\end{aligned}$$

where the first equality comes from $v_1(a) \le a + M$, and the second from the definition of M. As a result, $v_2(a) \le a + (1 + \alpha)M$. Iterating on n we conclude, for $n \ge 1$,

$$v_n(a) \le a + \sum_{i=0}^{n-1} \alpha^i M = a + \frac{1 - \alpha^n}{1 - \alpha}M. \tag{24}$$

Since $v_n(a)$ is increasing in n, $v_n(a)$ has a finite limit $v^*(a)$ as n approaches $+\infty$, with $v^*(a) \le a + M/(1 - \alpha)$. Furthermore, from Theorem 1, the two dimensional action set $0 \le y + z \le a$ can be reduced to the subset of the real line $0 \le z \le a_n^* \wedge a$, with $y = (a_n^* \wedge a) - z$. The result then follows directly from Theorem 1, Corollary 1, and the application of Heyman and Sobel (1984, Theorem 8-15).

∎

Proof of Corollary 3.

The proof is similar to the proof of Corollary 2 with $M = \alpha r(\underline{p})\mathbf{E}_\Theta\Theta$ and $v_n(a) \le a + \frac{1-\alpha^n}{1-\alpha}M$, where $n = T - t$.

2 Sufficient Conditions Under Which Banking is Sub-Optimal

Corollary 1 allows us to derive conditions under which banking never to be optimal when the price is fixed. Assume in the following that $\alpha\beta < 1$, $\alpha p + \tau > 1$ and $p > 1$. We first give sufficient conditions to propagate an upper bound on the derivative of the social-impact-to-go. With $n = T - t$, this upper bound is equal to

$$u(n) = \frac{\tau + (1 - \alpha p - \tau)(\alpha p)^n}{1 - \alpha p}.$$

Lemma 4 *Assume that the following three conditions hold at time t,*

- **(C1)** *the optimal banking decision yields $z_t^*(a) = 0$ for all $a \ge 0$,*

- **(C2)** $v_t'(a) \le u(T - t)$.

If in addition

$$u(T - t) \le (\alpha\beta)^{-1}, \tag{25}$$

then (C1) and (C2) also hold at time $t - 1$.

Proof. From Theorem 1, $v_t(\cdot)$ is differentiable and $v_t'(\cdot)$ in (C2) is well defined. To show (C1), fix y and consider the derivative of $J(y, z)$ with respect to the second variable,

$$
\begin{aligned}
\frac{dJ}{dz}(y, z) &= -1 + \alpha\beta E_\Delta \mathbf{E}_\Theta \; v_t'\left((py \wedge p\Theta) + \beta z + \Delta\right) \\
&\leq -1 + \alpha\beta u(T - t) \\
&\leq 0,
\end{aligned}
$$

where the inequalities follow from (C2) and (25). For any given y, $J(y, z)$ is maximized at $z = 0$, and $v_{t-1}(a)$ satisfies (C1). Finally, using the chain rule at the optimal decision and (C2), we obtain for $y = a$ (below the threshold),

$$
v_{t-1}'(a) = E_\Delta(\tau + \alpha p v_t'(pa + \Delta))G(a) \leq \tau + \alpha p u(T - t) = u(T - t + 1).
$$

Since $v_{t-1}'(\cdot)$ is non-increasing, the inequality in (C2) also holds for $v_{t-1}'(\cdot)$ at every a.

∎

We can now show that there exists a time period starting from which the organization should stop banking and follow the threshold policy described in Corollary 1.

Proposition 4 *Define n_0 as*

$$
n_0 = \left\lfloor \frac{\ln\left[\tau + (\alpha p - 1)(\alpha\beta)^{-1}\right] - \ln(\tau + \alpha p - 1)}{\ln(\alpha p)} \right\rfloor.
$$

For $T - n_0 \leq t \leq T$, the optimal policy never involves banking, i.e., $z_t^*(a) = 0$. If furthermore $p > 1$, the optimal capacity decision is $y_t^*(a) = a^* \wedge a$, where a^* is given by (5).

Proof. For $n = 0$, $v_T(a) = a$ satisfies (C1) and (C2) of Lemma 4. We can then recursively apply Lemma 4 as long as $n = T - t$ satisfies (25) which is equivalent to $n \leq n_0$. The second part of the proposition holds from Theorem 1.

∎

We summarize in the following corollary the results for the special case where the revenue customers do not contribute to the mission, *i.e.*, $\tau = 0$.

Corollary 4 *Assume that R-customers do not contribute to the mission ($\tau = 0$).*

- *If $\alpha\beta > 1$, allocating capacity to mission customers only in the last time period is always optimal.*

- *If $p < \beta$ or $\alpha p < 1$, serving R-customers is never optimal.*

- *If $\alpha\beta < 1$ and ($p < \beta$ or $\alpha p < 1$), using all assets to serve M-customers in the first period is optimal.*

- *If $\alpha\beta < 1$ and $\alpha p > 1$, the optimal policy is of threshold type.*

- *If $\beta = 0$ and $\alpha p > 1$, the optimal policy is of threshold type and myopic.*

- *If $\alpha\beta < 1$ and $\alpha p > 1$, for $T - n_0 \leq t \leq T$ the optimal policy is of threshold type, myopic and never involves banking, i.e., $z_t^*(a) = 0$ with*

$$n_0 = \left\lfloor -\frac{\ln(\alpha\beta)}{\ln(\alpha p)} \right\rfloor.$$

Under the stated conditions, the risk-free asset has the worst return of all applications for the assets. It is of use as a hedge against uncertain demand, as a low realization may lead to having to rebuild capital from a very low position and to having to go through many periods unable to serve M-customers. As the value-function becomes more concave due to there being more time periods to go, the optimal policy becomes more 'risk-averse', as there is a higher opportunity cost of having to the start the next period with a low cash position, and banking becomes more attractive.

3   Example of Non-Decreasing Optimal Price

For the model described in Section 3.1 and with a distribution $F(\cdot)$ with two mass points, the results of a numerical simulation are presented in Figure 6 (with probability mass $1/2$ at $\theta = 0.1$ and 2). The optimal price is not monotonic. An informal way of interpreting this result is as follows: If the price is set with full information about demand, the optimal price for the high-demand outcome is higher than the optimal price for the low-demand outcome. The uncertain demand case is a compromise between the two. At some point, as capacity increases, more "high-demand customers" are expected to be served, and the optimal price moves toward the higher price. Similar examples can be constructed with a continuous $F(\cdot)$.

4   Proofs and Support Lemmas for Section 3

Proof of Lemma 1. For clarity, we drop the subscript of $v_{t+1}^\epsilon(\cdot)$ in the following, writing simply $v^\epsilon(\cdot)$. Similarly we suppress the superscript ϵ from $\widetilde{p}^\epsilon(y)$.

We check the conditions for joint strict concavity of H^{v^ϵ}, locally at the optimal decision $(\widetilde{p}(y), y)$. By definition,

$$H^{v^\epsilon}(p, y) = \int_{\underline{\theta}}^{z(y,p)} v^\epsilon(r(p)\theta) f(\theta) d\theta + v^\epsilon(py) G(z(y, p)), \tag{26}$$
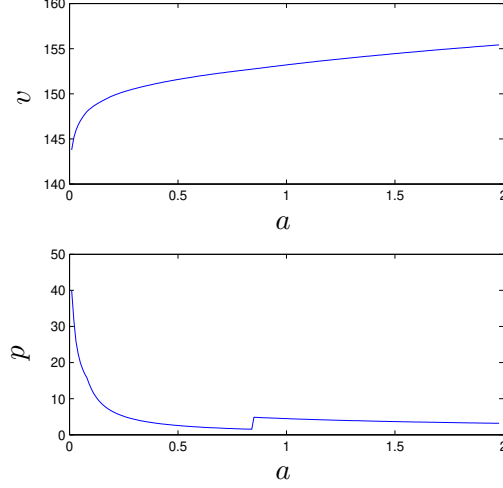
Figure 6: Example of non-monotonic price: value function and optimal price.

where $z(y,p) = y/\gamma(p)$ which we sometimes denote simply by z. Since $v^\epsilon(\cdot)$ is twice differentiable, each term of (26), and hence H^{v^ϵ}, are twice differentiable. Furthermore, since H^{v^ϵ} is concave in each of the variables, we only need to show

$$\nabla H^{v^\epsilon}(p,y) = H^{v^\epsilon}_{pp}(p,y)H^{v^\epsilon}_{yy}(p,y) - [H^{v^\epsilon}_{py}(p,y)]^2 > 0 \tag{27}$$

where $H^{v^\epsilon}_{pp}$, $H^{v^\epsilon}_{yy}$, and $H^{v^\epsilon}_{py}$ represent the second derivatives of H^{v^ϵ}. Likewise, we denote by $H^{v^\epsilon}_{p}$ and $H^{v^\epsilon}_{y}$ the first derivatives of H^{v^ϵ} with respect to p and y. For any given y, the optimal price $\widetilde{p}(y)$ exists and is an interior point. Therefore, the first-order optimality condition for p applies, that is

$$H^{v^\epsilon}_{p}(p,y) = 0, \tag{28}$$

where

$$H^{v^\epsilon}_{p}(p,y) = r'(p) \int_{\underline{\theta}}^{z} (v^\epsilon)'(r(p)\theta)\theta f(\theta)d\theta + y(v^\epsilon)'(py)G(z(y,p)). \tag{29}$$

Since we are concerned with concavity at the optimal $\widetilde{p}$, we may use the optimality condition (28) to obtain

$$
\begin{aligned}
H^{v^\epsilon}_{pp}(p,y) = {} & -y\left(\frac{r''(p)}{r'(p)} + p\frac{(\gamma'(p))^2}{\gamma^2(p)}g(z)\right)G(z)(v^\epsilon)'(py) + y^2 G(z)(v^\epsilon)''(py) + \\
& (r'(p))^2 \int_{\underline{\theta}}^{z} (v^\epsilon)''(r(p)\theta))\theta^2 f(\theta)d\theta.
\end{aligned} \tag{30}
$$

Similarly, the other derivatives can be written as

$$H^{v^\epsilon}_{y}(p,y) = p(v^\epsilon)'(py)G(z) \tag{31}$$

38

$$H_{yy}^{v^\epsilon}(p,y) = p^2(v^\epsilon)''(py)G(z) - \frac{p}{\gamma(p)}(v^\epsilon)'(py)f(z) \tag{32}$$

$$H_{py}^{v^\epsilon}(p,y) = (v^\epsilon)'(py)\left[1 + p\frac{\gamma'(p)}{\gamma(p)}g(z)\right]G(z) + py(v^\epsilon)''(py)G(z) \tag{33}$$

From the concavity of $v^\epsilon(\cdot)$, $H_{py}^{v^\epsilon}$ is negative if $1+p\gamma'(p)/\gamma(p)\,g(z) = 1-e(p)g(z) < 0$ which is implied by condition (10). This implies that the derivative of $\widetilde{p}(\cdot)$ is negative from $\widetilde{p}'(\cdot) = -H_{py}^{v^\epsilon}/H_{pp}^{v^\epsilon}$.

We can now study the concavity of H^{v^ϵ} locally at the optimal decision $(\widetilde{p}(y), y)$. From the previous derivatives (30), (32) and (33) we obtain, after algebraic simplification,

$$\begin{aligned}
\nabla H^{v^\epsilon}(p,y) &= H_{pp}^{v^\epsilon}(p,y)H_{yy}^{v^\epsilon}(p,y) - [H_{py}^{v^\epsilon}(p,y)]^2 \\
&= ((v^\epsilon)'(py))^2 G^2(z)\left[p\frac{r''(p)}{r'(p)}g(z) - 1 - 2p\frac{\gamma'(p)}{\gamma(p)}g(z)\right] \\
&\quad - (v^\epsilon)'(py)\,(v^\epsilon)''(py)G(z)^2 py\left[p\frac{r''(p)}{r'(p)} + 1 + \left(p\frac{\gamma'(p)}{\gamma(p)}+1\right)^2 g(z)\right] \\
&\quad + p(r'(p))^2 \int_{\underline{\theta}}^z (v^\epsilon)''(r(p)\theta))\theta^2 f(\theta)d\theta\left(p(v^\epsilon)''(py)G(z) - \frac{1}{\gamma(p)}(v^\epsilon)'(py)f(z)\right).
\end{aligned}$$

Recall that the optimal price is larger than or equal to $\underline{p}$. Hence from the concavity of $r(\cdot)$, $r'(p) \leq 0$ and $r''(p)/r'(p) \geq 0$ at the optimal price. From the concavity of $v^\epsilon(\cdot)$ and using the previous definition of $e(\cdot)$, we deduce

$$\nabla H^{v^\epsilon}(p,y) \geq -((v^\epsilon)'(py))^2 G^2(z)\left[1 - 2e(p)g(z)\right].$$

From condition (10), $1 - 2e(p)g(z) < -1$ and $\nabla H^{v^\epsilon}(p,y) > 0$. $H^{v^\epsilon}(p,y)$ is then locally strictly concave at $(\widetilde{p}(y), y)$ and, from Lemma 5 (in appendix), $\widetilde{H}^{v^\epsilon}(y)$ is twice differentiable and concave. For the condition that $\widetilde{p}(y)$ be interior for Lemma 5, note from (29) that, for any y, $H_p^{v^\epsilon}(p,y)$ is strictly positive for p small enough (both terms become positive with $r'(p) > 0$ for $p < \underline{p}$), and strictly negative for p close enough to $\overline{p}$ (the first term becomes increasingly negative from concavity of $v^\epsilon(\cdot)$ and from $r'(p) > 0$ for $p > \underline{p}$, while the second term goes to zero since $z(p,y)$ goes to infinity).

∎

Proof of Lemma 2. Take $\phi(y,a) = a - y + \epsilon\log((a-y)y) + \widetilde{H}^{v_{t+1}^\epsilon}(y)$ in (12). From Lemma 1, $\phi(y,a)$ is twice differentiable and jointly strictly concave (from the $\log((a-y)y)$ term). We can hence apply Lemma 5 (the optimal y is interior, since the barrier goes to $-\infty$ at both 0 and a) to show that $v_t^\epsilon(\cdot)$ is also twice differentiable and concave. Noting that $p_t^{*\epsilon}(a) = \widetilde{p}_t^\epsilon(y_t^{*\epsilon}(a))$, the second part then follows from Lemma 1 and the fact that $y_t^{*\epsilon}(a)$ increases in a. The monotonicity can be derived directly, since (12) implies $v_t^\epsilon(a+h) \geq v_t^\epsilon(a) + h$.

■

Proof of Lemma 3. We show the result by iterating on t. Assume that, for period t, $v_t^\epsilon(\cdot)$ converges pointwise to the optimal value function $v_t(\cdot)$ of the original problem, which is increasing and concave. We propagate pointwise convergence to the next time period

- From Lemma 6, for ϵ sufficiently small, $|v_t^\epsilon(\cdot)| \leq l(\cdot)$ where $l(\cdot)$ is integrable. By Lebesgue's dominated convergence theorem (see, *e.g.*, Billingsley 1995, p.209, thm. 16.4), this implies that $H^{v_t^\epsilon}(\cdot,\cdot)$ converges pointwise to $H^{v_t}(\cdot,\cdot)$.

- Since $H^{v_t^\epsilon}(p,y)$ is concave in $p \in]0,\overline{p}[$, using Lemmma 8 at each y shows that $\widetilde{H}^{v_t^\epsilon}(y)$ converges to $\widetilde{H}^{v_t}(y)$. Which is to say, $\widetilde{H}^{v_t^\epsilon}(\cdot)$ converges pointwise to $\widetilde{H}^{v_t}(\cdot)$.

- Since $a - y + \epsilon \log(a-y)y + \alpha \widetilde{H}^{v_t^\epsilon}(y)$ is concave in $y \in]0,a[$ and converges pointwise (for each a and y) to $\phi(y) = a - y + \alpha \widetilde{H}^{v_t}(y)$, we can apply Lemma 8 at each a to show that $v_{t-1}^\epsilon(a)$ converges to $v_{t-1}(a)$. Which is to say, $v_{t-1}^\epsilon(\cdot)$ converges pointwise to $v_{t-1}(\cdot)$. Lemma 8 also shows that $v_{t-1}(a)$ is optimal since $v_{t-1}(a) = \max_{0 \leq y \leq a} \phi(y)$ corresponds to the optimality equations (7).

Concavity of $v_{t-1}(\cdot)$ is now immediate from the concavity of $v_{t-1}^\epsilon(\cdot)$, which is preserved through pointwise convergence.

To establish that $\widetilde{p}_t(y)$ is non-increasing in the capacity decision, we need to show that, for each y, $\widetilde{p}_{t-1}^\epsilon(y) = \operatorname{argmax}_p H^{v_t^\epsilon}(p,y)$, which is non-increasing in y from Lemma 1, converges to $\operatorname{argmax}_p H^{v_t}(\cdot,y)$. We do this by checking that conditions for Lemma 9 apply. For each $y > 0$, we use Lemma 7 to verify that $H^{v_t}(\cdot,y)$ is locally strictly concave at the maximum and therefore has a unique maximizer. Since $H^{v_t}(\cdot,y)$ is concave and goes to zero at $p = 0$ and at $p = \overline{p}$, it has an interior maximum. The function $py \wedge r(p)\theta$ satisfies the conditions for Lemma 7 and, since the next period's $v_t(\cdot)$ is increasing and concave, $\phi(p,\theta) = v_t(py \wedge r(p)\theta)f(\theta)$ also satisfies the conditions for Lemma 7 (in $[\underline{p},\overline{p}] \times [\underline{\theta},\bar{\theta}]$).

■

Lemma 5 *Consider* $\phi(\cdot,\cdot) :]\underline{x},\overline{x}[\times]\underline{y},\overline{y}[\mapsto \mathbf{R}$, *with* $\underline{x},\overline{x},\underline{y},\overline{y} \in \mathbf{R}$. *Assume that* ϕ *is twice differentiable and that* $\phi(\cdot,y)$ *has a unique, interior maximum for each* $y \in]\underline{y},\overline{y}[$, *denoted by* $\widetilde{x}(y) = \operatorname{argmax}_x \phi(x,y)$. *If locally at* $(\widetilde{x}(y),y)$ *for all* $y \in]\underline{y},\overline{y}[$, ϕ *is strictly concave, then* $\widetilde{\phi}(y) = \max_x \phi(x,y) = \phi(\widetilde{x}(y),y)$ *is twice differentiable and strictly concave.*

Proof. ϕ_x and ϕ_{xx} denote the first- and second-order derivatives with respect to the first variable. The meanings of the notations ϕ_y, ϕ_{yy} and ϕ_{xy} follow naturally.

From the implicit function theorem on the first-order condition $\phi_x = 0$, $\widetilde{x}$ is differentiable and $\widetilde{x}' = -\phi_{xy}/\phi_{xx}$. Since ϕ is twice differentiable and $\widetilde{x}$ is an interior point, the derivatives of $\widetilde{\phi}$ exist and are equal to

$$\widetilde{\phi}'(y) = \widetilde{x}'(y)\phi_x(\widetilde{x}(y), y) + \phi_y(\widetilde{x}(y), y) = \phi_y(\widetilde{x}(y), y),$$

where we use the first-order condition in $\widetilde{x}$ that is $\phi_x(\widetilde{x}(y), y) = 0$,

and

$$
\begin{aligned}
\widetilde{\phi}''(y) &= \widetilde{x}'(y)\phi_{xy}(\widetilde{x}(y), y) + \phi_{yy}(\widetilde{x}(y), y) \\
&= -\frac{\phi_{xy}(\widetilde{x}(y), y)}{\phi_{xx}(\widetilde{x}(y), y)} \phi_{xy}(\widetilde{x}(y), y) + \phi_{yy}(\widetilde{x}(y), y) \\
&= \frac{\phi_{xx}(\widetilde{x}(y), y) \phi_{yy}(\widetilde{x}(y), y) - \phi_{xy}^2(\widetilde{x}(y), y)}{\phi_{xx}(\widetilde{x}(y), y)} < 0.
\end{aligned}
$$

where the last inequality comes from the strict concavity of $\phi(\cdot, \cdot)$ locally at $(\widetilde{x}(y), y)$ which implies that both the numerator and denominator are positive.

∎

Lemma 6 *For $0 < \epsilon < 1$, the optimal $v_t^\epsilon(\cdot)$ satisfying (12) is bounded as follows,*

$$c_t^0 + c_t^1 (\log a)^- \quad \leq v_t^\epsilon(a) \leq \quad c_t^2 + a \tag{34}$$

where we defined $(x)^- = 0 \wedge x$ and c_t^1 and c_t^2 are non-negative finite constants.

Proof. The statement is true for $t = T$ with $v_T(a) = a$. Assume the result holds at time t. We first show that

$$c_t^3 + c_t^4 (\log y)^- \leq \widetilde{H}^{v_t^\epsilon}(y) \leq c_t^5 \tag{35}$$

where c_t^4 and c_t^5 are non-negative finite constants. We set p equal to any p_0 to get a lower bound,

$$
\begin{aligned}
\widetilde{H}^{v_t^\epsilon}(y) &\geq \int_0^{y/\gamma(p_0)} v_t^\epsilon(r(p_0)\theta)f(\theta)d\theta + v_t^\epsilon(p_0 y)G(y/\gamma(p_0)) \\
&\geq \int_0^{y/\gamma(p_0)} (c_t^0 + c_t^1(\log r(p_0))^- + c_t^1(\log \theta)^-)f(\theta)d\theta + (v_t^\epsilon(p_0 y))^- \\
&\geq (c_t^0 + c_t^1(\log r(p_0))^-)F(y/\gamma(p_0)) + c_t^1 \int_0^{y/\gamma(p_0)} (\log \theta)^- f(\theta)d\theta + (c_t^0)^- + c_t^1(\log p_0)^- + c_t^1(\log y)^- \\
&\geq 2(c_t^0)^- + c_t^1(\log r(p_0))^- + c_t^1(\log p_0)^- + c_t^1 \int_0^{y/\gamma(p_0)} (\log \theta)^- f(\theta)d\theta + c_t^1(\log y)^-
\end{aligned}
$$

where we used (34), $0 \leq G(\theta) \leq 1$, and $(x+y)^- \geq x^- + y^-$. Noting that

$$\int_0^{y/\gamma(\underline{p})} (\log \theta)^- f(\theta) d\theta \geq \sup_\theta f(\theta) \int_0^1 \log \theta \, d\theta = -\sup_\theta f(\theta),$$

we obtain

$$\widetilde{H}^{v_t^\epsilon}(y) \geq 2(c_t^0)^- + c_t^1 \left((\log r(p_0))^- + (\log p_0)^- - \sup_\theta f(\theta) \right) + c_t^1 (\log y)^-,$$

which shows the lower bound in (35). We are now ready to show the lower bound of (34) at time $t-1$. Setting $y = a/2$,

$$\begin{aligned}
v_{t-1}^\epsilon(a) &\geq& a/2 + 2\epsilon \log(a) - 2\epsilon \log 4 + \alpha \widetilde{H}^{v_t^\epsilon}(a/2) \\
&\geq& 2(\log a)^- - 2 \log 4 + \alpha \widetilde{H}^{v_t^\epsilon}(a/2) \\
&\geq& 2(\log a)^- - 2 \log 4 + \alpha \left(c_t^3 + c_t^4 (\log a/2)^- \right) \\
&\geq& \alpha c_t^3 - 2 \log 4 - \alpha c_t^4 \log 2 + (2 + \alpha c_t^4)(\log a)^-,
\end{aligned}$$

where the second inequality holds since $0 < \epsilon < 1$, the third one from (35), and the last one from $(x+y)^- \geq x^- + y^-$.

An equivalent, somewhat simpler, proof structure leads to the upper bound on $v_t^\epsilon(a)$: 1) The expectation is bounded above by a constant, and therefore so is the maximum over p; 2) The maximum over y of $a - y + \log(a-y)y + \alpha \widetilde{H}^{v_{t+1}^\epsilon}(y)$ is bounded by the maximum of $a - y + \log(a-y)y$ plus the maximum $\alpha \widetilde{H}^{v_{t+1}^\epsilon}(y)$; 3) Finally, use $a/2 + 2\epsilon \log a < a$, for $0 < \epsilon < 1$.

∎

Lemma 7 *Consider $\phi(\cdot, \cdot) : [\underline{x}, \overline{x}] \times [\underline{\theta}, \overline{\theta}] \mapsto \mathbf{R}$, integrable in θ for all x, and such that, for all θ and for all x, $\phi(x, \theta)$ is locally either 1) concave and strictly increasing in x, or 2) strictly concave in x. Then if $\psi(x) = \int \phi(x, \theta) d\theta$ has an interior maximum, $\psi(\cdot)$ is strictly concave at $x^* = \mathrm{argmax}_x \psi(x)$.*

Proof. If $\phi(\cdot, \theta)$ is locally strictly concave at x^* in a set of θ with non-zero measure, then $\psi(\cdot)$ is locally strictly concave at x^*. This is the same as saying that if $\psi(\cdot)$ is not locally strictly concave at x^*, then, a. e. in θ, $\phi(\cdot, \theta)$ is not locally strictly concave at x^*. But this would then imply that condition 1) holds a. e. in θ, which would imply $\psi(\cdot)$ increasing at x^*, contradicting the optimality of x^*.

∎

Lemma 8 *(Convergence of the* sup*) Consider functions $\psi^\epsilon :]\underline{x}, \overline{x}[\mapsto \mathbf{R}$, parameterized by $\epsilon > 0$, such that, for every ϵ, ψ^ϵ is concave. Given pointwise convergence to some function $\psi :]\underline{x}, \overline{x}[\mapsto [0, +\infty[$, that is*

$$\lim_{\epsilon \to 0} \psi^\epsilon(x) = \psi(x), \ \forall x \in \,]\underline{x}, \overline{x}[,$$

then ψ is concave and can be continuously extended to be $\psi : [\underline{x}, \overline{x}] \mapsto [0, \infty[$. Further, $\lim_{\epsilon \to 0} \sup_x \psi^\epsilon(x)$ exists and is equal to $\max_x \psi(x)$.

Proof. The concavity of ψ is immediate by the propagation of inequalities through pointwise convergence. It follows that ψ is continuous, bounded over $]\underline{x}, \overline{x}[$ and can therefore continuously be extended to the closed interval $[\underline{x}, \overline{x}]$. In the following ψ designates the extended function.

For the convergence of $\sup_x \psi^\epsilon(x)$, consider any maximizer of the (extended) limit function, $\widetilde{x} = \operatorname{argmax}_x \psi(x)$. First, suppose $\widetilde{x}$ is interior. From concavity, for any $\delta > 0$,

$$\psi^\epsilon(x) \leq u_1^\epsilon(x) = \psi^\epsilon(\widetilde{x}) + (x - \widetilde{x})\frac{\psi^\epsilon(\widetilde{x}) - \psi^\epsilon(\widetilde{x} - \delta)}{\delta}, \ \forall x \notin]\widetilde{x} - \delta, \widetilde{x}[$$

and

$$\psi^\epsilon(x) \leq u_2^\epsilon(x) = \psi^\epsilon(\widetilde{x}) + (x - \widetilde{x})\frac{\psi^\epsilon(\widetilde{x} + \delta) - \psi^\epsilon(\widetilde{x})}{\delta}, \ \forall x \notin]\widetilde{x}, \widetilde{x} + \delta[.$$

The bounds above imply

$$\psi^\epsilon(x) \leq \begin{cases} \max\{u_1^\epsilon(\underline{x}), u_1^\epsilon(\widetilde{x} - \delta)\} & \text{for} \quad x \leq \widetilde{x} - \delta \\ \max\{u_2^\epsilon(\widetilde{x} - \delta), u_2^\epsilon(\widetilde{x})\} & \text{for} \quad \widetilde{x} - \delta \leq x \leq \widetilde{x} \\ \max\{u_1^\epsilon(\widetilde{x}), u_1^\epsilon(\widetilde{x} + \delta)\} & \text{for} \quad \widetilde{x} \leq x \leq \widetilde{x} + \delta \\ \max\{u_2^\epsilon(\widetilde{x} + \delta), u_2^\epsilon(\overline{x})\} & \text{for} \quad \widetilde{x} + \delta \leq x. \end{cases}$$

Taking the limit as $\epsilon \to 0$ and using pointwise convergence of $\psi^\epsilon(\cdot)$ we have that, should the limit of the supremum exist,

$$\begin{aligned} \lim_{\epsilon \to 0} \sup_x \psi^\epsilon(x) &\leq \max\{u_1(\underline{x}), u_1(\widetilde{x} - \delta), u_2(\widetilde{x} - \delta), \psi(\widetilde{x}), u_1(\widetilde{x} + \delta), u_2(\widetilde{x} + \delta), u_2(\overline{x})\} \\ &= \psi(\widetilde{x}) + \max\{\psi(\widetilde{x}) - \psi(\widetilde{x} + \delta), \psi(\widetilde{x}) - \psi(\widetilde{x} - \delta)\}, \end{aligned}$$

(where $u_1(\cdot)$ and $u_2(\cdot)$ are defined like $u_1^\epsilon(\cdot)$ and $u_2^\epsilon(\cdot)$ but with $\psi(\cdot)$ instead of $\psi^\epsilon(\cdot)$). The equality is from $u_1(\cdot)$ non-decreasing and $u_2(\cdot)$ non-increasing, a consequence of the concavity of $\psi(\cdot)$. Letting $\delta \to 0$, and by continuity of $\psi(\cdot)$, we obtain

$$\lim_{\epsilon \to 0} \sup_x \psi^\epsilon(x) \leq \psi(\widetilde{x}).$$

Since $\lim_{\epsilon \to 0} \sup_x \psi^\epsilon(x) \geq \lim_{\epsilon \to 0} \psi^\epsilon(\widetilde{x}) = \psi(\widetilde{x})$, the limit of the supremum of ψ^ϵ exists and is equal to $\max_x \psi(x)$.

Second, suppose $\widetilde{x} = \underline{x}$. This implies from concavity that ψ is non-increasing. We apply an argument similar to the above, based on pointwise convergence at $\underline{x} + \delta$ and $\underline{x} + 2\delta$. Since the limit function is concave and non-increasing, and by the same steps as above, we must have

$$\limsup_{\epsilon \to 0} \psi^\epsilon(x) \leq \psi(\underline{x} + \delta) + \{\psi(\underline{x} + \delta) - \psi(\underline{x} + 2\delta)\}.$$

Again letting $\delta \to 0$ and using continuity of $\phi(\cdot)$,

$$\limsup_{\epsilon \to 0} \psi^\epsilon(x) \leq \psi(\underline{x}).$$

On the other hand, from pointwise convergence we have

$$\limsup_{\epsilon \to 0} \psi^\epsilon(x) \geq \psi(\underline{x} + \delta),$$

which, letting $\delta \to 0$ and by continuity of $\psi(\cdot)$, implies

$$\limsup_{\epsilon \to 0} \psi^\epsilon(x) \geq \psi(\underline{x}).$$

The lower and upper bounds show that the limit exists and is as stated. The same argument applies when $\widetilde{x} = \overline{x}$.

∎

Lemma 9 *(Convergence of the* argsup*) Consider ψ and ψ^ϵ as in Lemma 8. Further, assume that* $\operatorname{argmax}_x \psi(x)$ *is unique. Define* $\widetilde{x} = \operatorname{argmax}_x \psi(x)$ *and* $\widetilde{x}^\epsilon = \operatorname{argsup}_x \psi^\epsilon(x)$. *Then* $\lim_{\epsilon \to 0} \widetilde{x}^\epsilon$ *exists and is equal to* $\widetilde{x}$.

Proof. If $\widetilde{x}^\epsilon$ does not converge, there is a sub-series of decreasing ϵ_k, $k = 1, 2, \ldots$ such that $\widetilde{x}_{\epsilon_k}$ converges to some x_1 other than $\widetilde{x}$. Define $x_2 = (\widetilde{x} + x_1)/2$. For ϵ_k small enough, x_2 is strictly between $\widetilde{x}$ and $\widetilde{x}_{\epsilon_k}$. But from pointwise convergence and uniqueness of the argmax of ψ, for ϵ_k small enough, $\psi^{\epsilon_k}(\widetilde{x}) > \psi^{\epsilon_k}(x_2)$. But, from optimality of $\widetilde{x}_{\epsilon_k}$, $\psi^{\epsilon_k}(x_2) \leq \psi^{\epsilon_k}(\widetilde{x}_{\epsilon_k})$. These two inequalities violate concavity of ψ^ϵ.

∎

Lemma 10 *If the optimal value function $v_t(\cdot)$ for optimality equations (18) is concave, then the first-order derivative of $v_t(\cdot)$ is bounded from below by one. In particular, $v_t(\cdot)$ is strictly increasing.*

Proof. Note that $v_t(a) - a$ is increasing in a from (16) since

$$v_t(a) - a = \max_{0 \leq z \leq a} \mathbf{E}_\Theta \max_{0 \leq y \leq a - z} J^{v_{t+1}}(y, z, \Theta), \tag{36}$$

The result is then immediate from the concavity of v_t.

■

Lemma 11 *For* $p \in [\underline{p}, \overline{p}]$,

$$\gamma'(p)\frac{r''(p)}{r'(p)} - \gamma''(p) \geq 0 \tag{37}$$

Proof. When $\gamma''(p) \leq 0$ the result is immediate. For $\gamma''(p) > 0$, note first that (37) is equivalent to

$$2(\gamma'(p))^2 - \gamma(p)\gamma''(p) \geq 0.$$

Note that $r'(p) \leq 0$ for $p \geq \underline{p}$ which implies that $p \geq -\gamma(p)/\gamma'(p)$ and since $\gamma''(p) > 0$ we conclude that $\gamma''(p)p \geq -\gamma(p)/\gamma'(p)\gamma''(p)$. On the other hand, $r''(p) \leq 0$ which is equivalent to $p\gamma''(p) \leq -2\gamma'(p)$. We can deduce that $-\gamma(p)/\gamma'(p)\gamma''(p) \leq -2\gamma'(p)$ and the result follows from $\gamma'(p) \geq 0$.

■

5 Total Revenue

Proposition 5 *If condition (10) holds, $ap_t^*(a)$ is increasing in a. Further, $\lim_{a \to 0} ap_t^*(a) = 0$.*

Proof. To show the first part, recall that the derivative of the optimal price for the ϵ-problems is equal to $\widetilde{p}' = -H_{py}^{v^\epsilon}/H_{pp}^{v^\epsilon}$. From (30), (33), and $H_{pp}^{v^\epsilon} \leq 0$, it follows that

$$
\begin{aligned}
(y\widetilde{p}(y))' \geq 0 \quad &\Leftrightarrow \quad y\widetilde{p}'(y) + \widetilde{p}(y) \geq 0 \\
&\Leftrightarrow \quad -H_{py}^{v^\epsilon}y + H_{pp}^{v^\epsilon}\widetilde{p}(y) \leq 0 \\
&\Leftrightarrow \quad -(v^\epsilon)'(\widetilde{p}(y)y)G(z)y\left(p\frac{r''(\widetilde{p}(y))}{r'(\widetilde{p}(y))} + 1 + p\frac{\gamma'(\widetilde{p}(y))}{\gamma^2(\widetilde{p}(y))}(\gamma(\widetilde{p}(y)) + \widetilde{p}(y)\gamma'(\widetilde{p}(y)))g(z)\right) \\
&\qquad + (r'(\widetilde{p}(y)))^2 \int_{\underline{\theta}}^z (v^\epsilon)'(r(\widetilde{p}(y))\theta))\theta^2 g(\theta)d\theta \leq 0 \\
&\Leftarrow \quad -(v^\epsilon)'(\widetilde{p}(y)y)G(z)y\left(p\frac{r''(\widetilde{p}(y))}{r'(\widetilde{p}(y))} + 1 + p\frac{\gamma'(\widetilde{p}(y))}{\gamma^2(\widetilde{p}(y))}r'(\widetilde{p}(y))g(z)\right) \leq 0,
\end{aligned}
$$

where the last inequality holds since $r' \leq 0$ at the optimal price, and the result follows from Lemma 3.

■

6 Proofs for Section 5

Proof of Proposition 2. This is shown from a sample-path argument. Consider a supposedly optimal policy with optimal decisions (y_t, z_t), $t < T$, and let x_t be the corresponding number of

served M-customers. Suppose there is a time instant s such that $y_s < c_s \wedge \theta_s$, with $x_s > 0$. Consider now an alternative policy with decision $\bar{y}_s = c_s \wedge \theta_s$ at time s, and $\bar{y}_t = y_t$, $s \neq t$. Assume also that the banking decision is such that $\bar{z}_t = z_t$ for all t. It follows that the corresponding number of served M-customers $\bar{x}_t$, $t < T$ is such that $\bar{x}_s = x_s - (\bar{y}_s - y_s)$, $\bar{x}_{s+1} = p(\bar{y}_s - y_s) + x_{s+1}$ and $\bar{x}_t = \bar{x}_t$ for $t < s$ and $t > s+1$. For $t < s$ the policies are coupled and have the same social impact in each period. The difference between the social impact of the modified policy and the presumed optimal is $\tau(\bar{y}_s - y_s) + (\bar{x}_s - x_s) = (\tau - 1)(\bar{y}_s - y_s)$ at time s, and $p(\bar{y}_s - y_s)$ at time $s+1$. The policies are coupled for later periods and, since $(\tau - 1)(\bar{y}_s - y_s) + \alpha p(\bar{y}_s - y_s) = (\tau + \alpha p - 1)(\bar{y}_s - y_s) > 0$, the modified policy outperforms the presumed optimal policy.

∎

Proof of Proposition 3. The proof of the first part of the result is similar to the proof of Theorem 1, noting that $J^{v_{t+1}}((a - z) \wedge \theta, z, \theta)$ in (18) is jointly concave in a and z for any increasing concave function $v_{t+1}(\cdot)$. For the second part, the optimal banking decision satisfies the first-order condition which, after some algebra, is equivalent to

$$\alpha\beta \int_0^{a-z} v'(p\theta + \beta z)f(\theta)d\theta + \left(1 - \tau + \alpha(\beta - p)v'(pa + (\beta - p)z)\right)G(a - z) = 1. \quad (38)$$

The first term of Equation (38) increases in a and decreases in z since $v'_{t+1}(\cdot)$ is a positive decreasing function. For the second term, $1 - \tau + \alpha(\beta - p)v'(pa + (\beta - p)z)$ is negative for $\alpha(p - \beta) + \tau > 1$, since $v'_{t+1}(\cdot) \geq 1$ from Lemma 10. It follows that the last term increases in a and decreases in z from the concavity of $v(\cdot)$, so that $z_t^*(a)$ increases in a.

∎

Proof of Theorem 3. Let operator $J^v(p, a)$ be defined as $J^v(p, a) = \alpha K^v(p, a) - L(p, a)$ where $K^v(p, a) = H^v(p, a)$ (defined to avoid confusion with references to the derivatives of H^v from Lemma (1), as noted below) and $L(p, a) = \mathbf{E}_\Theta[a \wedge \gamma(p)\Theta]$ such that

$$v_t(a) = a + \max_{0 \leq p \leq \bar{p}} \mathbf{E}_\Theta J^{v_{t+1}}(p, a). \quad (39)$$

In the following, for clarity, $v(\cdot)$ denotes $v_{t+1}(\cdot)$. We follow the proof of Lemma 1 and show that, if v is twice differentiable non-decreasing concave, $\widetilde{J}^v(a) = \max_{0 \leq p \leq \bar{p}} \mathbf{E}_\Theta J^v(p, a)$ is also twice differentiable non-decreasing concave in a. To that end, we need to evaluate the first- and second-order derivatives of J^v. The first-order condition in p yields $\alpha K_p^v(p, a) = L_p(p, a)$, where again K_p^v and L_p denote the first derivatives with respect to p. We now need to derive $K_{pp}^v(\cdot, \cdot)$ the second-order derivative of $K^v(\cdot, \cdot)$ with respect to p at the optimal price for the problem with flexible

capacity. The expression for $K_{pp}^v(\cdot,\cdot)$ can be greatly simplified by reference to the mathematical expression given by (30) for $H_{pp}^v(\cdot,\cdot)$ in the fixed-capacity case. In the context of this proof, notation $H_{pp}^v(p,a)$ refers then to the expression in (30), and can no longer be interpreted as the second derivative of $H^v(\cdot,\cdot)$ with respect to p since the optimal price with flexible capacity is different from the optimal price in the inflexible case. At the optimal price with flexible capacity,

$$K_{pp}^v(p,a) \;=\; H_{pp}^v(p,a) + \frac{r''(p)}{\alpha r'(p)} L_p(p,a). \tag{40}$$

The other second-order derivatives are equal to $K_{pa}^v(p,a) = H_{py}^v(p,a)$ and $K_{aa}^v(p,a) = H_{yy}^v(p,a)$, given by (33) and (32). Similarly, after some algebraic simplifications, with $z = a/\gamma(p)$ and $I(p,a) = \int_{\underline{\theta}}^z \gamma(p)\theta f(\theta)d\theta$, we have

$$L_p(p,a) \;=\; \gamma'(p)I(p,a) \tag{41}$$

$$L_{pp}(p,a) \;=\; \gamma''(p)I(p,a) - \frac{(\gamma'(p))^2}{\gamma(p)^2} a g(z)G(z) \tag{42}$$

$$L_{pa}(p,a) \;=\; \frac{\gamma'(p)}{\gamma(p)} g(z)G(z) \tag{43}$$

$$L_{aa}(p,a) \;=\; -\frac{1}{\gamma(p)} f(z). \tag{44}$$

We show next that $J^v(p,a)$ is concave in each dimension. From Lemma 10, (32), (44), and $v''(\cdot) \le 0$, we obtain

$$
\begin{aligned}
J_{aa}^v(a,a) \;&=\; \alpha H_{aa}^v(p,a) - L_{aa}(p,a) \\
&\le\; -\frac{p}{\gamma(p)} v'(pa) f(z) - \frac{1}{\gamma(p)} f(z) \\
&\le\; \frac{-\alpha p + 1}{\gamma(p)} f(z) \\
&\le\; 0.
\end{aligned}
$$

The first inequality is due to the concavity of v. The second one uses the fact that $v'(\cdot) \ge 1$ from Lemma 10, and the last inequality holds since $\alpha p > 1$. Similarly, to show $J_{pp}^v(p,a) \le 0$, we first bound $H_{pp}^v(p,a)$ from above as follows,

$$H_{pp}^{v\epsilon}(p,a) \le -a\left(\frac{r''(p)}{r'(p)} + p\frac{(\gamma'(p))^2}{\gamma^2(p)} g(z) \right) G(z) \tag{45}$$

The inequality follows from (30) and Lemma 10. Using again the concavity of v and Lemma 10, we conclude from (40), (42) and (45) that

$$
\begin{aligned}
J_{pp}^v(p,a) \;&=\; \alpha K_{pp}^v(p,a) - L_{pp}(p,a) \tag{46} \\
&\le\; -\alpha a\frac{r''(p)}{r'(p)}G(z) + (1-\alpha p)a\frac{(\gamma'(p))^2}{\gamma^2(p)}g(z)G(z) + \left(\gamma'(p)\frac{r''(p)}{r'(p)} - \gamma''(p) \right) I(p,y) \\
&\le\; 0 \tag{47}
\end{aligned}
$$

where the last inequality is due to $\alpha p > 1$ and Lemma 11.

We are now ready to show that the optimal price is non-increasing. The first-order derivative of the optimal price is equal to $-J_{pa}^v/J_{pp}^v$ and, since $J_{pp}^v \leq 0$, it suffices to show that the cross derivative is negative. We have then from (33), (43) and Lemma 10,

$$J_{pa}^v(p,a) = \alpha K_{pa}^v(p,a) - L_{pa}(p,a) \leq \left[\alpha + (\alpha p - 1)\frac{\gamma'(p)}{\gamma(p)} g(z) \right] G(z) \leq 0,$$

where the last inequality follows from condition (20).

Since J^v is concave in each dimension, to show concavity along the optimal price, we only need

$$\nabla J^v(p,a) = J_{pp}^v(p,a)J_{aa}^v(p,a) - (J_{pp}^v(p,a))^2 = \alpha^2 \nabla H^v(p,a) + \nabla L(p,a) + \alpha Q(p,a) \geq 0, \quad (48)$$

where ∇H^v is defined in (27),

$$\nabla L(p,a) = L_{pp}^v(p,a)L_{aa}^v(p,a) - (L_{pp}^v(p,a))^2 = \frac{\gamma''(z)}{\gamma(z)}(p)I(p,a)f(z), \quad (49)$$

and

$$
\begin{aligned}
Q(p,a) &= 2K_{ap}^v(p,a)L_{ap}(p,a) - K_{aa}^v(p,a)L_{pp}(p,a) - K_{pp}^v(p,a)L_{aa}(p,a) + \frac{r''(p)}{\alpha r'(p)}L_p(p,a)K_{aa}^v(p,a) \\
&= v'(pa)\left[\left(2\frac{\gamma'(p)}{\gamma(p)} - \frac{r''(p)}{r'(p)} \right) g(z)G(z)^2 + \frac{p}{\gamma(p)}\left(\gamma''(p) - \frac{r''(p)}{\alpha r'(p)}\gamma'(p) \right) I(p,a)f(z) \right] \\
&\quad + v''(pa)G(z)\left[\left(\frac{p\gamma'(p)}{\gamma(p)} + 1 \right)^2 ag(z)G(z) + p^2\left(\frac{r''(p)}{\alpha r''(p)}\gamma'(p) - \gamma''(p) \right) I(p,a) \right] \\
&\quad + \frac{1}{\gamma(p)}\frac{r''(p)}{\alpha r'(p)}\gamma'(p)I(p,a)f(z) + \frac{(r'(p))^2}{\gamma(p)}I_2(p,a)f(z), \quad (50)
\end{aligned}
$$

with

$$I_2(p,a) = \int_{\underline{\theta}}^z (v^\epsilon)''(r(p)\theta))\theta^2 f(\theta)d\theta.$$

Using again $v'(a) \geq 1$ from Lemma 10 with condition (10) and the expression for ∇H^v from (34), we obtain

$$
\begin{aligned}
\nabla H^v(p,a) &\geq v'(pa)G^2(z)\left[p\frac{r''(p)}{r'(p)}g(z) - 1 - 2p\frac{\gamma'(p)}{\gamma(p)}g(z) \right] \\
&\quad - v''(pa)G(z)^2 pa\left[p\frac{r''(p)}{r'(p)} + 1 + \left(p\frac{\gamma'(p)}{\gamma(p)} + 1 \right)^2 g(z) \right] \\
&\quad + p(r'(p))^2 I_2(p,a)\left(pv''(pa)G(z) - \frac{1}{\gamma(p)}v'(pa)f(z) \right). \quad (51)
\end{aligned}
$$

From (50) and (51) we have, after reorganizing terms,

$$\nabla J^v(p,a) \geq \alpha v'(pa)G(z)^2\left((\alpha p - 1)\frac{r''(p)}{r'(p)}g(z) - \alpha - 2(\alpha p - 1)\frac{\gamma'(p)}{\gamma(p)}g(z) \right)$$

48

$$+\frac{\alpha p v'(pa) - 1}{\gamma(p)}\left(\gamma'(p)\frac{r''(p)}{r'(p)} - \gamma''(p)\right)I(p,a)f(z)$$

$$+\alpha v''(pa)G(z)^2 a\left[(1-\alpha p)\left(\frac{p\gamma'(p)}{\gamma(p)} + 1\right)^2 ag(z) - \alpha p\left(p\frac{r''(p)}{r'(p)} + 1\right)\right]$$

$$+\alpha v''(pa)G(z)p^2\left(\frac{r''(p)}{\alpha r'(p)}\gamma'(p) - \gamma''(p)\right)I(p,a)$$

$$+\frac{1}{\gamma(p)}\alpha(1-\alpha p)I_2(p,a)f(z) + v''(pa)G(z)p^2(r'(p))^2 I_2(p,a). \tag{52}$$

The first term of the right-hand side of this inequality is negative from condition (10) and $\alpha p > 1$ at the optimal price. The second term is positive from Lemma 11, $v'(a) \geq 1$ from Lemma 10, and $\alpha p > 1$. The third term is non-negative from the concavity of v and $\alpha p > 1$. A similar approach shows that the remaining last three terms are also non-negative. $J^v(p,a)$ is then locally strictly concave along the optimal price, and $v_t(a)$ is twice differentiable and concave from Lemma 5, and increasing from Lemma 10.

■

ESMT Working Papers

	ESMT No.	Competence Center
Resource and Revenue Management in Nonprofit Operations Francis de Véricourt, ESMT Miguel Sousa Lobo, Duke University	08-006	Management and Technology
Nurse-To-Patient Ratios in Hospital Staffing: A Queueing Perspective Francis de Véricourt, ESMT Otis B. Jennings, Duke University	08-005	Management and Technology
Critical Mass Michał Grajek, ESMT Tobias Kretschmer, Ludwig-Maximilians-Universität München	08-004	Management and Technology
The Rhythm of the Deal: Negotiation as a Dance Erik H. Schlie Mark A. Young, Rational Games, Inc.	08-003	Leadership
Legacy Effects in Radical Innovation: A Study of European Internet Banking Erik H. Schlie, ESMT Jaideep C. Prabhu, Tanaka Business School, Imperial College London Rajesh K. Chandy, Carlson School of Management, University of Minnesota	08-002	Management and Technology
Upsetting Events and Career Investments in the Russian Context Konstantin Korotov, ESMT Svetlana Khapova, ESMT Visiting Professor and Assistant Professor at VU University Amsterdam	08-001	Leadership
Ambiguity Aversion and the Power of Established Brands A. V. Muthukrishnan, Hong Kong University of Science and Technology Luc Wathieu, ESMT	07–005	Management and Technology
Accelerated Development of Organizational Talent Konstantin Korotov, ESMT	07–004	Leadership
Usage and Diffusion of Cellular Telephony, 1998-2004 Michał Grajek, ESMT Tobias Kretschmer, Ludwig-Maximilians-Universität München	07–003	European Competitiveness
Estimating Level Effects in Diffusion of a New Technology: Barcode Scanning at the Checkout Counter Jonathan Beck, Humboldt Universität zu Berlin Michał Grajek, ESMT Christian Wey, Technische Universität, Berlin	07–002	European Competitiveness
Estimating Network Effects and Compatibility in Mobile Telecommunication Michał Grajek, ESMT	07–001	European Competitiveness