2020

Operations Management In The Presence Of Strategic Agents

Jingxing (rowena) Gan
University of Pennsylvania

Operations Management In The Presence Of Strategic Agents

Abstract

My dissertation develops stylized models using analytical and numerical tools to understand innovative business models in the context of operations management.

In the first chapter, using airlines as a backdrop, I study optimal overbooking policies with endogenous customer demand, when customers internalize their expected cost of being bumped. I first consider the traditional setting in which compensation for bumped passengers is fixed and booking limits are the airline's only form of control. I provide sufficient conditions under which demand endogeneity leads to lower overbooking limits in this case. I then consider the broader problem of joint control of ticket price, bumping compensation, and booking limit. I show that price and bumping compensation can act as substitutes, which reduces the general problem to a more tractable one-dimensional search for optimal overbooking compensation and effectively allows the value of flying to be decoupled from the cost of being bumped. Finally, I extend our analysis to the case of auction-based compensation schemes and demonstrate that these generally outperform fixed compensation schemes. Numerical experiments that gauge magnitudes suggest that fixed-compensation policies that account for demand endogeneity can significantly outperform those that do not and that auction-based policies bring smaller but still significant additional gains.

In the second chapter, I study the design of an emerging fundraising method for Blockchain-based startups, Initial Coin Offerings (ICOs), with a particular focus on capped ICOs. I propose a simple model of matching supply and demand with ICOs by companies involved in production of physical goods, aka inventory/asset ``tokenization''. I examine how ICOs should be designed---including optimal token floating and pricing for both utility and equity tokens (aka, security token offerings, STOs)---in the presence of moral hazard, production risk and demand uncertainty, make predictions on ICO failure, and discuss the implications on firm operational decisions and profits. I show that in the current unregulated environment, ICOs lead to risk-shifting incentives (moral hazard), and hence to agency costs, underproduction, and loss of firm value. These inefficiencies, however, fade as product margin increases and market conditions improve, and are less severe under equity (rather than utility) token issuance. Importantly, the advantage of equity tokens stems from their inherent ability to better align incentives, and hence continues to hold even in unregulated environments.

In the third chapter, I aim to understand how uncapped ICOs can fund service platforms under network effects. I propose an infinite horizon model that incorporates the interaction between the firm, speculators, service providers and customers. I find that both the platform's service capacity and service providers' profitability are enhanced by stronger network effect, larger customer base, and/or lower unit service cost. Moreover, I show that uncapped ICO is successful if and only if the cost of building the platform does not exceed the total service cost per period. I also extend the base model to account for firm's moral hazard and show that under loose regulation, uncapped ICO can still be successful if the firm charges the right amount of service fee.

Degree Type

Dissertation

Degree Name

Doctor of Philosophy (PhD)

Graduate Group

Operations & Information Management

First Advisor

Gerry Tsoukalas

Keywords

Blockchain, Initial Coin Offering, Operations Management, Overbooking, Security Token Offering, Strategic Agents

Subject Categories

Operational Research

OPERATIONS MANAGEMENT IN THE PRESENCE OF STRATEGIC AGENTS

Jingxing Gan

A DISSERTATION

in

Operations, Information and Decisions

For the Graduate Group in Managerial Science and Applied Economics

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2020

Supervisor of Dissertation

Gerry Tsoukalas, Assistant Professor in Operations, Information and Decisions

Graduate Group Chairperson

Nancy Zhang, Professor in Statistics

Dissertation Committee

Noah Gans, Professor in Operations, Information and Decisions

Serguei Netessine, Professor in Operations, Information and Decisions

OPERATIONS MANAGEMENT IN THE PRESENCE OF STRATEGIC AGENTS

I dedicate my dissertation work to my husband,

Hao Liu,

who provided tremendous love and support over the course of my graduate study.

I also dedicate this work to my parents,

Jianping Shao and Weizhong Gan,

who taught me that education is a lifelong endeavor.

ACKNOWLEDGEMENT

First and foremost, I thank my advisors Dr. Gerry Tsoukalas, Dr. Noah Gans, and Dr. Serguei Netessine for their countless hours of mentoring me in becoming an independent researcher in the past five years. Not only did they generously share their expertise, but they also offered invaluable moral support—they have truly made every research meeting worth looking forward to. They are the best mentors I could have ever wished for.

I am also grateful for the support of the Wharton Doctoral Program, the Ripple Foundation Blockchain Research Initiative, and the Mack Institute for Innovation Management Research.

Finally, my special thanks go to my undergraduate supervisor, Dr. Ezra Miller, with whose guidance I developed an interest in research.

ABSTRACT

OPERATIONS MANAGEMENT IN THE PRESENCE OF STRATEGIC AGENTS

Jingxing Gan

Advisor: Gerry Tsoukalas

My dissertation develops stylized models using analytical and numerical tools to understand innovative business models in the context of operations management.

In the first chapter[1], using airlines as a backdrop, I study optimal overbooking policies with endogenous customer demand, when customers internalize their expected cost of being bumped. I first consider the traditional setting in which compensation for bumped passengers is fixed and booking limits are the airline's only form of control. I provide sufficient conditions under which demand endogeneity leads to lower overbooking limits in this case. I then consider the broader problem of joint control of ticket price, bumping compensation, and booking limit. I show that price and bumping compensation can act as substitutes, which reduces the general problem to a more tractable one-dimensional search for optimal overbooking compensation and effectively allows the value of flying to be decoupled from the cost of being bumped. Finally, I extend our analysis to the case of auction-based compensation schemes and demonstrate that these generally outperform fixed compensation schemes. Numerical experiments that gauge magnitudes suggest that fixed-compensation policies that account for demand endogeneity can significantly outperform those that do not and that auction-based policies bring smaller but still significant additional gains.

In the second chapter[2], I study the design of an emerging fundraising method for Blockchain-based startups, Initial Coin Offerings (ICOs), with a particular focus on capped ICOs. I propose a simple model of matching supply and demand with ICOs by companies involved in production of physical goods, aka inventory/asset "tokenization". I examine how ICOs

[1]Work in this chapter leads to the paper *Gan et al. (2019)*.
[2]Work in this chapter leads to the paper *Gan et al. (2020)*.

should be designed—including optimal token floating and pricing for both utility and equity tokens (aka, security token offerings, STOs)—in the presence of moral hazard, production risk and demand uncertainty, make predictions on ICO failure, and discuss the implications on firm operational decisions and profits. I show that in the current unregulated environment, ICOs lead to risk-shifting incentives (moral hazard), and hence to agency costs, underproduction, and loss of firm value. These inefficiencies, however, fade as product margin increases and market conditions improve, and are less severe under equity (rather than utility) token issuance. Importantly, the advantage of equity tokens stems from their inherent ability to better align incentives, and hence continues to hold even in unregulated environments.

In the third chapter, I aim to understand how *uncapped* ICOs can fund service platforms under network effects. I propose an infinite horizon model that incorporates the interaction between the firm, speculators, service providers and customers. I find that both the platform's service capacity and service providers' profitability are enhanced by stronger network effect, larger customer base, and/or lower unit service cost. Moreover, I show that uncapped ICO is successful if and only if the cost of building the platform does not exceed the total service cost per period. I also extend the base model to account for firm's moral hazard and show that under loose regulation, uncapped ICO can still be successful if the firm charges the right amount of service fee.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

CHAPTER 1 : Overbooking with Endogenous Demand

1.1. Introduction

Overbooking is the practice of selling more capacity than is available. It is commonly used by service companies whose capacity is perishable and whose customers sometimes fail to show up for service. Without overbooking, these no-show customers leave unutilized capacity that might have been sold to others. With overbooking, companies can serve more customers and increase revenues.

The practice is attractive to companies, and it is widely used in the air travel, hotel, and car rental industries. In the air travel industry – the focus of our analysis – its origins can be traced back to the 1940s, when it was discovered that overselling a flight, even by mistake, could be an effective way to address no-shows and be a viable money-making strategy (Mihm, 2017). The benefits for airlines can be quite significant. Smith et al. (1992), for example, mention no-show rates of 15% for sold-out flights (absent overbooking), while Curry (1990) reports that overbooking can generate an additional 3-10% of gross passenger revenues for airlines.

At the same time, overbooking has a downside. When the number of no-shows is smaller than expected, the company doing the overbooking must refuse service to – or bump – some customers, and both the customers and the provider incur costs. Customers suffer the disutility of extra time spent waiting in the airport, a missed connection, a late arrival at the destination, the need to be "walked" to another hotel. The company that does the bumping typically needs to use substitute capacity – either its own or that of a competitor – to fulfill its service obligation, and it also often provides a voucher or some form of monetary compensation to those who are bumped.

In addition to the direct costs associated with the bumping of specific passengers, there also exist indirect costs associated with the practice: the prospect of this type of service failure reduces the value that customers expect to obtain from buying a ticket or making a reservation and can therefore adversely affect customer demand. What's more, this indirect cost is especially relevant in today's world in which customers can more freely obtain access to relevant information. The US Bureau of Transportation Statistics (BTS), for example, publishes the total numbers of boarded and bumped passengers for each major US-based airline on a monthly basis (BTS, 2018). The resulting bumping probabilities – the summary measures of quality of service (QoS) – are, in turn, reported by both mainstream media outlets and by specialty sites devoted to air travel, such as thepointsguy.com (2018) and travelersunited.org (2018). For savvy travelers, specialized search engines, such as the KVS Availability Tool, let customers check availability and infer bumping probabilities for individual flight routes (kvstool.com, 2018).

The now infamous 2017 case of a passenger being dragged off of a United Airlines flight has turned a spotlight on the practice of bumping and its costs. This incident drew outrage on social media, attracting the attention of more than 550 million users on the micro-blogging site Weibo (Hernandez and Li, 2017), and had widespread consequences, prompting United and other airlines to rethink how to better manage this element of QoS. In response, United has increased the maximum compensation it offers to bumped passengers, now a $10,000 voucher, and it has developed an auction-based system to identify which passengers to bump, an analogue of a system that Delta Airlines has used for a number of years (Martin, 2017; Zhang, 2017). Both the increase in compensation and the change of payment mechanism reflect the importance of secondary demand effects associated with overbooking.

Existing research on overbooking does not address demand effects associated with bumping in a realistic operational setting. While the focus of the traditional overbooking literature has been on deriving optimal overbooking policies in increasingly realistic (and complex) operational environments (Chatwin, 1996; Karaesmen and van Ryzin, 2004; Kunnumkal et al., 2012), we are not aware of any papers that account for demand endogeneity. This is all the more surprising given empirical evidence of airline customers' strategic buying behavior (Li et al., 2014). Conversely, there have been several recent papers in the economics literature, such as Fu et al. (2012), Ely et al. (2017), and Sano (2017), that explicitly model demand effects associated with overbooking. These papers take a mechanism-design approach at the expense of using highly stylized models that render booking limits superfluous, however. In contrast, booking-limit control is broadly useful in our more detailed operational setting. Recent papers in the revenue management (RM) and operations management (OM) literature, such as Gallego et al. (2008), Gallego and Şahin (2010), Alexandrov and Lariviere (2012), and Cachon and Feldman (2018), consider related problems in advanced selling, though they only tangentially address traditional questions related to overbooking and do not seek to examine the differences between fixed and auction-based overbooking policies. We provide a more detailed review of the relevant literature in §1.2.

We consider the demand effect in the context of a model that captures important operational and customer details. Our approach follows the spirit of Dana and Petruzzi's (Dana Jr and Petruzzi, 2001) analysis of inventory problems with QoS-sensitive demand. In our case, however, the inventory level is a fixed number of seats on an aircraft, and the control variables include the booking limit, the price, and the compensation for bumped customers. To squarely focus on customer response to overbooking, we consider a model in which customers are homogeneous in their valuation of the flight itself and heterogenous in their disutilities of being bumped. Section 1.3 defines our operational and customer model, describes the equilibria that emerge from bumping-sensitive demand, and defines the airline's optimization problem. We then analyze and compare two distinct compensation schemes for bumped customers.

The first scheme is consistent with that traditionally found in the OM literature and assumes that all bumped customers receive the same compensation. Section 1.4.2 analyzes a setup in which ticket price and bumping compensation are fixed *a priori*, and the booking limit represents the airline's only available control. We call these *booking-limit policies* and begin by characterizing the optimal booking limit and customer response in a traditional analysis, in which the airline ignores potential demand effects of bumping. We then compare the results for this "myopic" policy to those for a setting in which the airline recognizes the demand effect that stems from bumping and calibrates its booking limit accordingly. We characterize conditions that are sufficient to imply that demand-dependent booking limits are no larger than those of traditional ones that ignore bumping-sensitive demand. These conditions are complex and depend on customers' expected utility of purchasing a ticket without bumping, together with the form of the distribution of their disutility of being bumped.

Section 1.4.3 expands the analysis of the fixed-compensation scheme by considering a broader set of controls in which the airline sets the ticket price and bumping compensation along with the booking limit. We call these *overbooking policies*. Here, we demonstrate that, in fact, the ticket price and bumping compensation act as substitutes and that, for any given booking limit, there exists an infinite set of price-compensation pairs that obtain the same customer equilibrium and expected airline profit. These results have two important implications. First, when solving the broader overbooking problem, the airline need only consider a price that allows it to decouple the expected value customers obtain for the flight from the compensation they must be offered in the event of being bumped. In this case, the bumping compensation acts as a *direct*, rather than an *indirect*, control of customer demand. Second, for any given level of bumping compensation, the optimal booking limit is an analogue to that of the myopic booking-limit policy of §1.4.2 and can be found via closed-form expression. In turn, the optimal price, compensation, and booking limit can be found as via a simple line search over potential compensation levels.

In Section 1.5 we introduce an auction scheme for compensating bumped passengers and compare its performance to that of the optimal fixed-compensation schemes identified in §1.4.3. The auction can be viewed as a multi-unit auction with single-unit demand, a setting for which a so-called uniform-price scheme induces customers to truthfully reveal their preferences and is efficient, allocating seats to customers with the highest disutility of being bumped. For this scheme, the optimal price is, again, one that allows the airline to decouple the expected value customers obtain from the flight from their potential compensation. An upper-bound "cap" on auction payouts – such as the $10,000 limit publicized by United – provides the airline with a similar, direct control over the distribution of demand and allows us to compare straightforwardly the performance of the auction and the fixed-compensation schemes. We further show that, for any fixed-compensation scheme identified in §4, there is an analogous capped auction that performs at least as well. Finally, we identify conditions under which expected airline profits are increasing in the cap, so that the optimal auction-based overbooking policy has only a single active control, the booking limit.

In Section 1.6, we report the results of numerical experiments that assess the magnitude of the demand effect. We find that fixed-compensation policies that account for demand endogeneity can significantly outperform those that do not and that the use of auction-based policies brings smaller but significant additional gains. These numerical results suggest that the demand effect can have a first-order impact on both overbooking policies and expected revenues.

1.2. Literature Review

Our work is related to the RM and economics literatures on overbooking, as well as to the OM literature on strategic consumers. We discuss each in turn.

The vast majority of the RM literature on overbooking does not focus on the issue of demand endogeneity. For overviews, see Chapter 4 in Talluri and van Ryzin (2004) and Chapter 5.2 in Belobaba et al. (2015). Here we describe only a few of the relevant papers, many of which themselves include useful references. Among the earliest works is a static single-fare-class model developed by Beckmann (1958) in which an airline minimizes lost revenue by reducing unused capacity or cost of overselling. Rothstein (1971) focuses on the dynamic aspect of the overbooking problem, deriving a policy that depends on time to flight and current reservations. Chatwin (1996), Karaesmen and van Ryzin (2004), Kunnumkal et al. (2012), and Lan et al. (2015) consider overbooking with multiple customer classes. Klophaus and Pölt (2010) study dynamic booking policies when customer willingness to pay can change over time. In contrast to these papers, our focus is on understanding how the airline's overbooking policy can affect consumer demand *ex ante*.

At the same time, empirical evidence of strategic consumer behavior has been well documented in the specific context of the airline industry. von Wangenheim and Bayón (2007, p. 36) document that "customers who experience negative consequences of revenue management significantly reduce the amount of their transactions with the airline." Li et al. (2014) use a structural model to estimate the fraction of consumers in the air-travel industry who delay a purchase, anticipating lower future prices, and show that a non-trivial portion of customers engages in this strategic behavior.

More broadly, strategic consumer behavior has been widely studied in operations management, the most relevant stream of work focusing on the setting of inventory levels (Dana Jr and Petruzzi, 2001; Su and Zhang, 2008; Cachon et al., 2018). Closest in spirit to our work, are Dana Jr and Petruzzi (2001), who analyze the impact of demand endogeneity in a newsvendor-type model and show that a firm that recognizes strategic consumer behavior implements a higher inventory level because greater availability increases one's willingness to pay. Also closely related is Alexandrov and Lariviere (2012), who allow for overbooking when modeling strategic responses to restaurants' reservation policies. Here, however, no-show behavior is a fluid function of the booking limit, and customers never need to be bumped.

Although the link between firm inventory levels and demand has been extensively studied in these contexts, we are not aware of analogous work that focuses specifically on the overbooking problem. Furthermore, we make relatively weak assumptions regarding consumers' knowledge. Unlike most of the previously cited work, our model does not require the inventory level to be observable; that is, customers need not directly observe either the plane's capacity or the airline's booking limit. Customers in our model similarly need not be informed about the distributions of aggregate demand or bumping disutility.

The use of auctions to determine bumping compensation has become increasingly popular in practice, and a few recent papers in the economics literature use highly stylized models to analyze this mechanism. Fu et al. (2012) consider overbooking controls that exclude booking limits and rely only on price and bumping compensation. Similarly, Ely et al. (2017) consider initial price and refund policies for airlines when passengers are uncertain of their eventual willingness to pay at the time of ticketing, and Sano (2017) extends the analysis to multi-unit demand. These papers, like this one, argue in favor of auction-based compensation schemes. Unlike this paper, however, they do not do not capture essential operational details that affect booking limit controls and do not seek to characterize the effect and magnitude of demand endogeneity on the optimal overbooking policy.

1.3. Overbooking with Fixed Bumping Compensation

In this section we formally define the overbooking problem for the case in which the compensation paid to bumped customers is fixed. At the start of §1.5, we provide the details of an analogous model for an auction-based compensation scheme, and we note there the differences between that setup and the model we define here.

We first define our model's primitives and the associated customer equilibrium, and we highlight important informational assumptions that we make. We then introduce the airline's expected profit maximization problem and define relevant ranges for its policy parameters.

1.3.1. Model Primitives

We consider a monopolist airline that offers a flight with a single fare class and $k \geq 1$ available seats. The airline sets the ticket price, p, a booking limit, b, that acts as an upper bound on the number of tickets it will sell, and a compensation amount, c, paid to each bumped customer. If more than k paying customers show up for the flight, the airline randomly selects a subset to bump and re-book on a subsequent flight. The expected cost of re-booking a customer on an alternative flight, r, is exogenously defined. Hence, under the fixed compensation scheme, the total cost of bumping each customer is $(c+r)$. We call the triple (p, b, c) the airline's *overbooking policy.*

Potential demand for a flight is uncertain. For example, the number of people who consider traveling from the flight's origin to its destination on a particular date can be random, and we represent potential demand as a random variable, Q. We denote the cumulative distribution function (CDF) of demand as $F(q)$, with support $0 \leq \underline{q} < \overline{q} \leq \infty$ and $\int_{\underline{q}}^{\overline{q}} dF(q) < \infty$. We analyze overbooking policies that use fixed compensation by differentiating relevant expressions, and for analytical convenience we therefore model Q as continuous, with density $f(q) > 0$ over its support. Thus, individual customers are infinitesimal.

Customers have three critical attributes: the value they derive from flying, their no-show probabilities, and the disutility they incur from being bumped. We assume that they are homogeneous along the first two dimensions. They share a common value from flying, v, either on the original flight or, if bumped, on their re-booked flight. Customers who purchase tickets also have an identical no-show probability, $\alpha \in (0, 1)$. The focus of our interest is the third dimension, customers' disutilities of being bumped, which we call their *hassle costs.* We assume that hassle costs are heterogeneous across the population, and by sorting them from smallest to largest, we can model them as the cumulative distribution, $G(w)$, of a random variable, W, with support $0 \leq \underline{w} < \overline{w} \leq \infty$. Again, for analytical convenience we assume $G(w)$ is continuous, with density $g(w) > 0$ over its support.

1.3.2. Model Equilibrium

Taken together, a plane with capacity k, an airline overbooking policy, $(p,\ b,\ c)$, potential demand, Q, and a set of customer attributes, $(v,\ \alpha,\ W)$, induce an equilibrium outcome. Individual customers decide whether or not to buy tickets, depending on the expected value of the purchase. In turn, each ticket holder shows up for the flight with probability $1 - \alpha$, and if the number of customers who do show is greater than the plane's capacity, excess customers are bumped. In equilibrium there is a set of customers who decide to purchase, a complementary set who do not, and a corresponding probability that a customer who shows up for the flight is bumped.

To formally describe the equilibrium, we begin with the customer purchase decision. Suppose customers share a common belief regarding the endogenous probability of being bumped in equilibrium, and denote this probability by $\beta \in [0, 1)$. Then the expected value obtained by a customer with hassle cost w, drawn from W, who buys a ticket is

$$U(\beta, w) = -p + (1 - \alpha)v + (1 - \alpha)\beta(c - w). \tag{1.1}$$

Here, the first term to the right of the equality is the ticket's purchase price, and the second is the expected value of flying, given the no-show probability α. The last term represents the expected value that the customer obtains from the possibility of being bumped: with probability $(1-\alpha)\beta$ she shows up for the flight and is bumped; in turn, the value she obtains from being bumped is the compensation less her hassle cost, $c - w$. When $c > w$ the value of being bumped is a net reward, and when $c < w$ it is a net cost.

We assume that customers have an outside option whose value we normalized to zero. Given an equilibrium bumping probability, β, a passenger with hassle cost w who considers purchasing a ticket will buy one if and only if the expected value of the purchase is non-negative: $U(\beta, w) \geq 0$.

An equilibrium β then induces an equilibrium customer demand response via $U(\beta, w)$. For $\beta = 0$, which can happen for instance when $b \leq k$ and there is no bumping, $U(\beta, w) \geq 0$ if and only if $p \leq (1 - \alpha)v$. That is, with no bumping any customer will buy a ticket if and only if the price does not exceed the expected value of flying. For $\beta > 0$, $U(\beta, w)$ is strictly decreasing in w. In all cases, we can define the *equilibrium threshold hassle cost*, $\widehat{w}$, at or below which a customer buys a ticket, and above which she does not, as follows:

$$\widehat{w} = \begin{cases} \underline{w}, & \text{if } U(\beta, \underline{w}) \leq 0; \\ \{w \,|\, U(\beta, w) = 0\}, & \text{if } \underline{w} < w < \overline{w}; \text{ and} \\ \overline{w}, & \text{if } U(\beta, \overline{w}) \geq 0. \end{cases} \tag{1.2}$$

Throughout the chapter, we also refer to $\widehat{w}$ using interchangeably the terms *marginal customer's hassle cost*, *customers' equilibrium response*, and *customers' response.*

From (1.1) and (1.2) we see that $G(\widehat{w})$ represents the fraction of potential customers who obtain non-negative value from purchasing a ticket. Thus, when $\widehat{w} = \underline{w}$, $G(\widehat{w}) = 0$ and no one is willing to buy a ticket, and when $\widehat{w} = \overline{w}$, $G(\widehat{w}) = 1$ and everyone is willing to buy. We are often most interested in the interior case, in which $\widehat{w} \in (\underline{w}, \overline{w})$ so that $U(\beta, \widehat{w}) = 0$ and $G(\widehat{w}) \in (0, 1)$, but in some parts of our analysis the boundary cases, $U(\beta, \underline{w}) \leq 0$ and $U(\beta, \overline{w}) \geq 0$, can also be important.

Having defined the manner in which an equilibrium β induces an equilibrium $\widehat{w}$, we turn to the mechanism by which an equilibrium $\widehat{w}$ induces an equilibrium β. To this end, we first characterize the number of tickets sold.

Recall that the potential demand for a flight is a random variable, Q. As in Dana Jr and Petruzzi (2001), we assume that, for any demand realization, q, the distribution of customers' hassle costs follows $G(w)$. That is, we assume that W is effectively independent of Q. Potential customers with hassle costs $w \leq \widehat{w}$ then buy tickets, and those with hassle costs $w > \widehat{w}$ do not, effectively thinning the potential demand. Given the booking limit b and the potential demand realization q, we can define the number of tickets sold as $s = \min\{b, q\, G(\widehat{w})\}$, where the term $qG(\widehat{w})$ represents the "thinned" demand for the flight. In turn, we define the random variable

$$S = \min\{b, QG(\widehat{w})\} \tag{1.3}$$

as the equilibrium number of tickets sold.

From here, we can derive the bumping probability, β, in three steps. First, the sale of s tickets results in a smaller number of customers who show up for the flight. We denote that random number, $N(s, \alpha) \in [0, s]$, as function of s and α, and for the moment, we leave the explicit dependence on s and α undefined. In turn,

$$N \equiv N(S, \alpha) \tag{1.4}$$

is the equilibrium number of customers who show up for the flight. Second, not all of those who show are bumped, and we let

$$(N-k)^{+} \;\equiv\; (N(S,\alpha)-k)^{+} \;=\; \max\{0, N(S,\alpha)-k\} \tag{1.5}$$

denote the equilibrium number of customers who are bumped. Finally, we can use the numbers of shows and of bumped customers to calculate the bumping probability as the ratio of expectations

$$\beta = \mathsf{E}[(N-k)^{+}]/\mathsf{E}[N]. \tag{1.6}$$

Thus, from $\widehat{w}$ we obtain β.

Given potential demand, Q, and customer attributes, (v, α, W), an overbooking policy (p, c, b) yields an equilibrium if there exist β and $\widehat{w}$ that simultaneously satisfy (1.2) and (1.6). In Section 1.3.3, we discuss the informational demands that such an equilibrium requires and discuss the use of β as a measure of bumping probability. In §1.3.4 we then define the airline's optimization problem and detail the parameter range to be considered for (p, c, b).

1.3.3. Information Required to Obtain an Equilibrium

Our equilibrium model requires that, in choosing an overbooking policy (p, b, c), the airline is aware of all relevant demand and customer data: the demand distribution, $F(q)$; the value customers derive from the flight, v, the no-show rate, α, and the hassle-cost distribution, $G(w)$. The airline must also understand the calculation of S and N, along with the equilibrium equations (1.2) and (1.6).

The informational requirements for customers are lower. In using (1.1) to decide whether or not to buy a ticket, each potential customer need know only her own attributes (v, α, w), the airline's price and bumping compensation, (p, c), and the bumping probability β. Interestingly, we show in §1.4 and §1.5 that the problem can be reduced to a form that does *not* require the customer know or estimate β.[1] Furthermore, the customer need *not* know the distribution of demand, $F(q)$, the hassle cost distribution, $G(w)$, the flight's capacity, k, or the airline's booking limit, b, and she need not be able to calculate (1.6).

Rather, customers can obtain an estimate of (1.6) from statistics published by the US Bureau of Transportation Statistics (BTS, 2018) and reported by many media outlets. More formally, suppose the airline runs a sequence of independent and identically distributed (i.i.d.) flights $i \in \{1, \ldots, m\}$ with fixed overbooking policy $(p,\ b,\ c)$. Let N_i denote the random number of customers who show up for flight i and $(N_i - k)^+$ denote the random number of customers who are bumped that flight. Then, after m flights, the reported fraction of passengers who are bumped, $\widehat{\beta}$, is

$$\widehat{\beta} = \frac{\sum_{i=1}^{m}(N_i - k)^+}{\sum_{i=1}^{m} N_i} = \frac{\frac{1}{m}\sum_{i=1}^{m}(N_i - k)^+}{\frac{1}{m}\sum_{i=1}^{m} N_i} \xrightarrow[m\to\infty]{} \frac{\mathsf{E}[(N-k)^+]}{\mathsf{E}[N]} = \beta, \tag{1.7}$$

by the law of large numbers. As the introduction notes, BTS reports the building blocks of $\hat{\beta}$, $\sum_{i=1}^{m} N_i$ and $\sum_{i=1}^{m}(N_i - k)^+$, for each airline as a whole on a quarterly and annual basis, and while BTS does not report analogous data for specific routes, specialty sites can provide additional flight-specific data. This definition of β in is the *ex post* fraction of passengers who are bumped, an analogue of the fill rate in inventory theory.

An alternative that could be considered is the *ex ante* probability of bumping, $\beta' = \mathsf{E}[N - k)^+/N]$, which is the expectation of the ratio of customers bumped, rather than the ratio

[1] We defer the relevant discussion to those sections.

of expectations and would be estimated as

$$\widehat{\beta'} \;= \frac{1}{m}\sum_{i=1}^{m}\big((N_i - k)^+/N_i\big) \;\xrightarrow[m\to\infty]{}\; \mathsf{E}\left[(N-k)^+/N\right] = \beta'. \tag{1.8}$$

Note that β and β' need not be the same. We believe that, in our specific context of airline overbooking, β is the measure that is more practically relevant and more realistically accessible to customers.

We briefly describe our reasoning here and provide supporting details and analysis in Appendix A.1.1. First, while β can be readily estimated from published *aggregate* data, an analogous estimate of β' would require customers to obtain bumping fractions from individual flights, data that are not published and are not typically disclosed by airlines to their passengers. Second, suppose that, nevertheless, customers wished to estimate β' based their initial estimates of bumping on the publicly available statistic, β. Even if individual-flight data were available and customers used them to update their initial beliefs, differences between β and β' are small enough that a customer would have to take hundreds or (more typically) thousands of flights to distinguish the latter from the former. Third and finally, although differences between the β and β' are typically small – on the order of 10^{-3} or less – it can be shown that $\beta \geq \beta'$, a relationship that suggests that an equilibrium based on an initial estimate of β will be stable: customers for whom β is too high to fly will never fly and will not collect the data needed to change their initial estimate; conversely those for whom β is low enough to fly will continue flying even if, after thousands of flights, their estimates slip from β to β'.

1.3.4. Airline's Optimization Problem

Having defined the model's primitives and equilibrium expressions, we can now concisely formulate the airline's associated optimization problem. Given an overbooking policy (p, b, c) that yields a customer equilibrium, the airline earns revenue p for each ticket sold and pays bumping compensation c and rerouting cost r for each customer bumped, netting profits

$$\Pi(p, b, c) \;=\; pS - (c + r)(N - k)^{+}. \tag{1.9}$$

In principal, the airline chooses p, c, and b to maximize expected profits subject to (1.1)–(1.6).

This problem statement is not complete, however, because it does not ensure the existence of a relevant equilibrium for each (p, b, c). Below we define ranges for policy parameters that, in most cases, do induce equilibria, and in §1.4.1 we explicitly note remaining, boundary cases for which equilibria do not exist.

First, we discuss lower bounds. We assume that both p and c are non-negative, so the airline does not give potential customers money to fly, and it does not charge customers for the pleasure of being bumped. In turn, given $p \geq 0$, we assume that $b \geq k$, since the airline will not benefit by forcing itself to fly with empty seats.

Next we discuss upper bounds. As we noted in §1.3.2, for $\beta = 0$ all customers are willing to buy a ticket whenever $p \leq (1 - \alpha)v$, and no customer is willing to buy a ticket when $p > (1 - \alpha)v$, since the ticket price exceeds the expected value of flying. For cases in which $\beta > 0$, we similarly limit $p \leq (1 - \alpha)v$ to exclude cases in which customers only buy tickets because of a potential benefit of being bumped. Given $p \leq (1 - \alpha)v$ and $c \leq \overline{w}$, $U(\beta, \overline{w}) \geq 0$ for all $\beta \in [0, 1]$. Therefore, we can also require that $c \leq \overline{w}$ since, there is no need to consider higher levels of bumping compensation.

We call overbooking policies that fall within these bounds, *admissible* and summarize their properties as follows.

Definition 1.1. (Admissible Overbooking Policies)
Admissible overbooking policies have: (i) $0 \leq p \leq (1 - \alpha)v$*; (ii)* $b \geq k$*; and (iii)* $0 \leq c \leq \overline{w}$.

We label the set of admissible policies Ξ, and we call individual admissible policies $\xi \in \Xi$.

In §1.4.1 we characterize the equilibria of interest for Ξ.

The airline then searches for an admissible overbooking policy that maximizes expected profits:

$$\max_{\xi \in \Xi} \mathsf{E}\Big[\Pi(p, b, c)\Big] \tag{1.10}$$
$$\text{subject to } (1.1), (1.2), (1.3), (1.6).$$

1.4. Analysis of Fixed-Compensation Schemes

In this section, we analyze a set of schemes that pay bumped customers a fixed, predetermined level of compensation, as defined in §1.3. We begin in §1.4.1 with a preliminary analysis that characterizes the expected number of bumped customers and the equilibria of interest. We then use this foundation to analyze two sets of overbooking policies of increasing complexity.

In §1.4.2, we analyze booking-limit policies, that is, single-control policies of the booking limit, b, that assume an exogenously specified price, p, and bumping compensation, c. As a benchmark, we first analyze the traditional myopic policy considered in the RM literature, one that assumes the booking limit does not affect demand, and we develop a simple characterization of the optimal booking limit for this case. We then perform a more delicate analysis of booking-limit control that recognizes the endogeneity of demand, and we develop equilibrium conditions under which optimal demand-dependent booking limits are stricter than those suggested by the benchmark myopic policy.

Equation (1.1) highlights the fact that demand is in fact affected by all three controls, (p, b, c), and in §1.4.3 we analyze overbooking policies' joint use of price, bumping compensation, and booking limit to maximize expected profit. The first part of our analysis shows that, in fact, price and bumping compensation act as *substitutes* and that it is sufficient to consider policies that set price equal to the expected value of flying and then use bumping compensation and booking limit as controls. Furthermore, in this setting the intensity of demand becomes a direct outcome of the bumping compensation and can be decoupled from the booking limit.

1.4.1. Preliminary Analysis

In this section, we provide two sets of preliminary results that we require to conduct a full analysis of the overbooking problem. First, we define the properties of a simple loss function that models the conditional expectation of the number of bumped customers, given some realization of the number of tickets sold, and we show that commonly used distributions of numbers of customers who show up for the flight generate a loss function with the desired properties. Second, we characterize the set of equilibria to be considered in our analysis.

Modeling the Expected Number of Bumped Customers

Much of the analysis below requires that we differentiate expressions, such as (1.6), that include expected numbers of bumped customers, and we sometimes find it analytically convenient to work with the conditional expectation, given a sales realization s. To that end, we use the conditional expectation to characterize here both the expected number of customers who show up for a flight and the expected number of bumped customers.

Suppose s tickets are sold. Given each customer has a no-show probability of α, we call the indicator function of the event "customer i shows up," $\mathbb{1}\{\text{ i shows }\}$, and need make no additional assumptions to show that

$$\mathsf{E}[N(s,\alpha)] = \mathsf{E}\left[\sum_{i=1}^{s} \mathbb{1}\{i \text{ shows}\}\right] = (1-\alpha)s, \tag{1.11}$$

so that $\mathsf{E}[N] = (1-\alpha)\mathsf{E}[S]$.

If we further assume that individual customers' no-show behavior is i.i.d., then $N(s,\alpha) \sim \mathcal{B}(s, 1-\alpha)$, a binomially distributed random variable with probability of success $(1-\alpha)$ and number of samples s. In turn, a normal distribution, $\mathcal{N}(\mu,\sigma)$, with mean $\mu = (1-\alpha)s$ and standard deviation $\sigma = \sqrt{\alpha(1-\alpha)s}$ represents a simple continuous approximation to the binomial $\mathcal{B}(s, 1-\alpha)$.

In the same spirit, we can characterize the expected number of bumped customers by first conditioning on the number of tickets sold, and we let

$$\ell(s,k,\alpha) = \mathsf{E}[(N(s,\alpha)-k)^{+}] \tag{1.12}$$

define a loss function that is a direct analogue of that used in inventory theory. To ease notational burden, we will sometimes write partial derivatives of this function using a prime symbol and the variable of interest: for example $\ell'(s) \equiv \frac{\partial \ell(s,k,\alpha)}{\partial s}$.

We make minimal assumptions regarding the loss function.

Definition 1.2. (Loss Function)
(i) $\ell''(s) \geq 0$*;*
(ii) $\ell(s,k,\alpha) = 0$ *for all* $s \leq k$ *and* $\ell'(s) = 0$ *for all* $s < k$*;*
(iii) $\ell'(s) \in (0, 1-\alpha)$ *for all* $s \in [k, \infty)$*; and*
(iv) $\lim_{s\to\infty} \ell'(s) = 1-\alpha$*.*

We note that the discrete analogues of $\ell'(s) \geq 0$ and $\ell''(s) \geq 0$ are $\ell(s) - \ell(s-1) \geq 0$ and $\ell(s+1) - \ell(s) \geq \ell(s) - \ell(s-1)$, respectively.

Properties (i) and (ii) are common. Typically loss functions are convex, and loss can never be incurred when sales fall below the plane's capacity k. The upper limit in properties (iii) and (iv) follow the fact that each ticket sold has only a probability of $(1-\alpha)$ of turning into a customer who shows up for the flight. Only for very large s do marginal sales lead to additional shows who will nearly certainly be lost, each with a show probability of $(1-\alpha)$ for each new ticket sold.

As expected, the definition's properties are generally satisfied by both the binomial and normal distributions described above.

Lemma 1.3. (Properties of Loss Function Satisfied)

For a plane with k seats and loss function $\ell(s,k,\alpha) = (N(s,\alpha) - k)^+$:
(i) $N(s,\alpha) \sim \mathcal{B}(s, 1-\alpha)$ satisfies the discrete analogue of properties (i)–(iv) of Definition 1.2; and
(ii) $N(s,\alpha) \sim \mathcal{N}\left((1-\alpha)s, \sqrt{\alpha(1-\alpha)s}\right)$ satisfies properties (i), (iii), and (iv) of Definition 1.2.

The proof of this and all results can be found in Appendix A.1.2.

Note that, because of its infinite support below k, the normal approximation does not satisfy the loss function's property (ii). We emphasize that this does not affect our theoretical results, which only depend on the definition of $\ell(\cdot)$ and not the specific distributional form of N. We do use the normal approximation to the binomial in our numerical examples, however. When we do, we truncate the normal distribution at k and renormalize the probabilities over the support above k to sum to one.

Equilibria and Policies of Interest

Here, we characterize the set of equilibria we will consider when analyzing the airline's problem (1.10). We also further characterize policies of interest: those that make positive expected profits.

We begin with the equilibria.

Lemma 1.4. (Existence and Uniqueness of Equilibria)
(i) For overbooking policies with $p = (1-\alpha)v$, $b > k$, and $c \leq \underline{w}$ there is no equilibrium.
(ii) For all other overbooking policies $\xi \in \Xi$, there exists at least one equilibrium.
(iii) For the policies in part (ii), if $g'(w) \leq 0$, $\forall\, w \in [c, \overline{w}]$, then $\exists$ a unique equilibrium $\{\beta, \widehat{w}\}$.

The policies identified in part (i) of the lemma have a price, $p = (1-\alpha)v$, that leaves no consumer surplus, and a bumping compensation, $c \leq \underline{w}$, that adequately compensates no bumped customer. If there were no overbooking, so that $b = k$, then $\beta = 0$ independently of $\widehat{w}$, any w would obtain $U(0, w) = 0$, and $\widehat{w} = \overline{w}$ would be consistent with $\beta = 0$. With

$b > k$, however, there is the potential for bumping customers, and there is no consistent $(\beta, \widehat{w})$ pair: $\beta = 0$ induces $\widehat{w} = \overline{w}$, which in turn induces $\beta > 0$, which then induces $\widehat{w} = \underline{w}$, and so on.

Part (ii) of the lemma shows that other admissible policies are better behaved and admit at least one equilibrium. While we cannot rule out the existence of multiple equilibria for any hassle-cost distribution, $G(\cdot)$, part (iii) of the lemma provides a sufficient condition under which there is exactly one, namely when the density of the hassle cost is decreasing above c, a property that is satisfied by uniform and exponential hassle-cost distributions, by normally distributed hassle-cost distributions whenever $c \geq \mathsf{E}[W]$, and more generally by decreasing failure rate (DFR) distributions.

Even if there do exist multiple equilibria associated with a given policy $\xi \in \Xi$, the following lemma shows that they are well ordered and suggests that we have good reason to focus on the unique equilibrium that maximizes the airline's expected profits.

Lemma 1.5. (Ordering of Equilibria)
Suppose an overbooking policy $\xi \in \Xi$ induces multiple equilibria. Pick any two distinct equilibria from the set, and call them $(\beta_1, \widehat{w}_1) \neq (\beta_2, \widehat{w}_2)$.

(i) Without loss of generality, we can order the two so that the second equilibrium has a strictly lower bumping probability and a strictly higher marginal hassle cost: $\beta_1 > \beta_2$ and $\widehat{w}_1 < \widehat{w}_2$.

(ii) Given the ordering in (i), the set of customers with $w \leq \widehat{w}_1$ is a strict subset of those with $w \leq \widehat{w}_2$, and the airline earns strictly higher expected profits in $(\beta_2, \widehat{w}_2)$.

Part (i) implies that we can order the equilibria from smallest to largest $\widehat{w}$ and that the largest of these is unique. Part (ii) further implies that the largest equilibrium maximizes the number of customers who buy tickets *and obtain positive expected value, $U(\beta, w)$*. This largest equilibrium also maximizes the airline's expected profits.

At the same time, a larger $\widehat{w}$ is not a Pareto improvement over a smaller one. In particular, customers with $w < c$ – those who enjoy a net benefit from being bumped – see their $U(\beta, w)$'s decrease as the equilibrium bumping probability falls from β_1 to β_2. Nevertheless, when the equilibrium is $(\beta_2, \widehat{w}_2)$, even those customers obtain a positive expected value from purchasing tickets and remain in the market.

Thus, an airline whose overbooking policy can induce multiple equilibria has an interest in inducing the largest of them, and if customers can be convinced that the low bumping probability that's associated with the highest $\widehat{w}$ is the equilibrium of interest, they will willingly settle on the largest equilibrium as well. More importantly, in §1.4.3 we show that, even if an admissible policy induces multiple equilibria, we can also find an alternative that induces only the largest of them and earns the same expected profit. We will therefore assume that, if there are multiple possible equilibria, the largest of these is obtained.

Finally, when optimizing over policies $\xi \in \Xi$, we will sometimes simplify our analysis by excluding equilibria that make no profit. To characterize these we recall from (1.11) that $\mathsf{E}[N] = (1-\alpha)\mathsf{E}[S]$ and from (1.6) that $\mathsf{E}[(N-k)^+] = \beta\mathsf{E}[N] = (1-\alpha)\beta\mathsf{E}[S]$. Therefore, we can substitute $(1-\alpha)\beta\mathsf{E}[S]$ for $\mathsf{E}[(N-k)^+]$ in (1.10) and rearrange terms to show that the airline's expected profits are

$$\mathsf{E}[\Pi(p, b, c)] = [p - (1-\alpha)\beta(c+r)]\,\mathsf{E}[S], \tag{1.13}$$

where $[p - (1-\alpha)\beta(c+r)]$ is the expected margin per customer, and $\mathsf{E}[S]$ is the expected number of units sold. We can then define a profit-making equilibrium as follows.

Definition 1.6. (Profit-Making Equilibrium)
Any profit-making equilibrium has $[p - (1-\alpha)\beta(c+r)]\,\mathsf{E}[S] > 0$.

Note that $G(\underline{w}) = 0$ implies that $\mathsf{E}[S] = \mathsf{E}[\min\{b, QG(\underline{w})\}] = 0$, so $\widehat{w} = \underline{w}$ is never profit-making.

1.4.2. Optimal Booking Limits

In this section, we assume that the price, p, and bumping compensation, c, are exogenously defined and that the booking limit, b, is the s only form of control. The airline's *booking-limit problem* is then

$$\max_{b \geq k} \mathsf{E}\Big[\Pi(p, b, c)\Big] \tag{1.14}$$
$$\text{subject to } (1.1), (1.2), (1.3), (1.6).$$

We first consider a benchmark case in §1.4.2, in which the airline myopically overlooks the effect of the booking limit on demand and treats $\widehat{w}$ as an exogenous factor. This is analogous to the traditional approach taken in the RM literature on overbooking. Subsequently in §1.4.2, we analyze the demand-dependent case in which the airline strategically takes customer behavior into account.

Myopic Booking Limits

Suppose the airline does not recognize that the marginal customer's hassle cost, $\widehat{w}$, is an equilibrium reaction to β and, in turn, b. Rather it assumes that $\widehat{w}$ is fixed.

In this case, we can derive the optimal *myopic* booking limit, b_m^*, as the solution to the first order condition (FOC) with respect to b. That is,

$$\frac{d\mathsf{E}[\Pi]}{db}\bigg|_{myopic} = \frac{\partial \mathsf{E}[\Pi]}{\partial b} = 0. \tag{1.15}$$

On the revenue side, a marginal increase in b yields the following change in unit sales:

$$\frac{\partial \mathsf{E}[S]}{\partial b} = \frac{\partial \mathsf{E}\left[\min\{b, QG(\widehat{w})\}\right]}{\partial b} = \mathsf{P}\{QG(\widehat{w}) > b\}. \tag{1.16}$$

On the cost side,

$$\mathsf{E}[(N-k)^+] \;=\; \mathsf{E}[\ell(\min\{b, QG(\widehat{w})\})] \;=\; \int_0^{\frac{b}{G(\widehat{w})}} \ell(qG(\widehat{w}))\, f(q)\, dq \;+\; \int_{\frac{b}{G(\widehat{w})}}^{\infty} \ell(b)\, f(q)\, dq, \tag{1.17}$$

and differentiating (1.17) with respect to b we have

$$\frac{\partial \mathsf{E}[(N-k)^+]}{\partial b} \;=\; \frac{\partial \mathsf{E}[\ell(\min\{b, QG(\widehat{w})\})]}{\partial b} \;=\; \ell'(b)\mathsf{P}\{QG(\widehat{w}) > b\}. \tag{1.18}$$

We then use the results of (1.16) and (1.18) to find the FOC for (1.14), assuming expected profit is only affected through the booking limit, b, without an effect on $\widehat{w}$. In particular, we see that $\frac{\partial \mathsf{E}[\Pi]}{\partial b} = 0$ implies that

$$p\,\mathsf{P}\{QG(\widehat{w}) > b\} \;-\; (c+r)\,\ell'(b)\mathsf{P}\{QG(\widehat{w}) > b\} \;=\; 0 \quad\Rightarrow\quad p \;-\; \ell'(b)\,(c+r) \;=\; 0. \tag{1.19}$$

Note that, in this case, the FOC (1.19) defines the optimal booking limit by trading off the revenue, p, gained from the marginal customer, should she purchase a ticket, against the marginal expected bumping cost, $(c+r)$, induced by that marginal customer showing up. This is, in fact, the assumption that's made in standard overbooking models, and it is completely independent of the demand distribution and the customer response $\widehat{w}$.

Proposition 1.7. (Optimal Myopic Booking Limit)
Given fixed, admissible p and c, the optimal myopic booking limit, b_m^*, behaves as follows.
(i) If $p-(1-\alpha)\,(c+r) \geq 0$, then $b_m^* = \infty$, and the airline does not impose a booking limit.
(ii) If $p-(1-\alpha)\,(c+r) < 0$, then there exists a unique optimal $b_m^* = \max\left\{\ell'^{-1}\left(\frac{p}{c+r}\right), k\right\}$.
(iii) When $b_m^* \in (k,\infty)$, $\frac{\partial \mathsf{E}[\Pi]}{\partial b} > 0$ for $b < b_m^*$ and $\frac{\partial \mathsf{E}[\Pi]}{\partial b} < 0$ for $b > b_m^*$.

The results follow directly from the FOC (1.19). When $p-(1-\alpha)\,(c+r) > 0$, the expected margin per customer continues to be positive, even when $\beta = 1$. In this case the airline is happy to bump customers, and the optimal booking limit is infinite. In contrast, when $p-(1-\alpha)\,(c+r) < 0$, the expected margin per customer $p-(1-\alpha)\,\beta\,(c+r)$ eventually becomes negative, as $\beta \to 1$, and the optimal myopic booking limit is finite. Because $\ell(b)$

is increasing convex, the sign of $\frac{\partial \mathsf{E}[\Pi]}{\partial b}$ changes from positive to negative as b increases from $b < b_m^*$ to $b > b_m^*$, so that b_m^* is the unique solution to the FOC.

While the optimal myopic booking-limit policy is independent of the equilibrium it induces, customers nevertheless react to the policy to induce a specific equilibrium, $(\beta, \widehat{w})$. We can use the convexity of the loss function, together with the FOC (1.19), to provide some insight into the nature of the equilibrium.

Proposition 1.8. (Optimal Myopic Booking Limit is Profit-Making)
(i) The equilibrium induced by any $\xi \in \Xi$ obtains $\ell'(b) > (1-\alpha)\beta$.

Suppose $p > 0$.

(ii) If $\beta = 0$, or if either $p < (1-\alpha)v$ or $c > \underline{w}$ or both, then b_m^* induces a profit-making equilibrium.

Recalling from (1.6) and (1.11) that $\mathsf{E}[(N-k)^+] = \beta\mathsf{E}[N] = (1-\alpha)\beta\mathsf{E}[S]$, part (i) can equivalently be stated as $\ell'(b)\mathsf{E}[S] > \mathsf{E}[(N-k)^+]$. To see that (ii) holds when $\beta = 0$, note the following. For $\beta = 0$, we have $\widehat{w} = \overline{w}$, so $p, \mathsf{E}[S] > 0$ and $\mathsf{E}[\Pi(p, b_m^*, c)] = p\mathsf{E}[S] - (c+r)\mathsf{E}[(N-k)^+] = p\mathsf{E}[S] > 0$. To see that (ii) holds when $p < (1-\alpha)v$ or $c > \underline{w}$, take the FOC (1.19), use the inequality from part (i) to substitute for $\ell'(b_m^*)$ and show that the expected margin per customer is strictly positive: $[p - (1-\alpha)\beta(c+r)] > 0$. In §1.4.3 we address the cases in which $p = (1-\alpha)v$ and $c \leq \underline{w}$.

Strategic Booking Limits

Now suppose the airline is aware that the marginal customer's hassle cost, $\widehat{w}$, is in fact an equilibrium outcome, and call the booking limit that optimally takes the demand effect into account the *strategic booking limit*, b_s^*. Given b_s^* is the airline's optimal response to customers' actions, the expected profits it generates will be trivially (weakly) greater than those induced by b_m^*.

What is perhaps less obvious is how the two equilibrium booking limits, b_s^* and b_m^*, compare to each other. If an airline is aware of the fact that its overbooking policy will affect customer demand through $\widehat{w}$ – with some customers potentially enjoying a net benefit of being bumped, while others incurring a net cost – should it overbook more or less than in the myopic case?

To answer this question, we first consider policy parameters and equilibria that allow us to develop relevant FOCs. These include policies for which $p \in (0,(1-\alpha)v)$, $c \in (0,\overline{w})$, and $b_m^* \in (k,\infty)$. For the same reason, we will assume that the policy (p, b_m^*, c) obtains an interior equilibrium $U(\beta,\widehat{w}) = 0$ for which $\widehat{w} \in (\underline{w},\overline{w})$. In §1.4.3 we will explicitly consider boundary cases.

As before, we consider the FOC. Compared to the myopic FOC in (1.15), differentiation of expected profits in the strategic case yields extra, complicating terms:

$$\left.\frac{d\mathsf{E}[\Pi]}{db}\right|_{strategic} = \frac{\partial \mathsf{E}[\Pi]}{\partial b} + \frac{\partial \mathsf{E}[\Pi]}{\partial \widehat{w}}\,\frac{d\widehat{w}}{db} = \left.\frac{d\mathsf{E}[\Pi]}{db}\right|_{myopic} + \frac{\partial \mathsf{E}[\Pi]}{\partial \widehat{w}}\,\frac{d\widehat{w}}{db} = 0. \tag{1.20}$$

From Proposition 1.7 part (iii), we know that $\left.\frac{d\mathsf{E}[\Pi]}{db}\right|_{myopic} = \frac{\partial \mathsf{E}[\Pi]}{\partial b}$ is positive for $b < b_m^*$, zero for b_m^*, and negative for $b > b_m^*$. Thus, to answer the question of how b_s^* compares to b_m^*, it suffices to characterize the sign of the product $\frac{\partial \mathsf{E}[\Pi]}{\partial \widehat{w}}\,\frac{d\widehat{w}}{db}$ as a function of b.

We begin with the first term, $\frac{\partial \mathsf{E}[\Pi]}{\partial \widehat{w}}$, which can be obtained by differentiating (1.3) and (1.17) with respect to the (interior) $\widehat{w}$. For the revenues, we have

$$\frac{\partial \mathsf{E}[S]}{\partial \widehat{w}} = \int_0^{\frac{b}{G(\widehat{w})}} g(\widehat{w})\,q\,f(q)dq, \tag{1.21}$$

and for the costs, we have

$$\frac{\partial \mathsf{E}[(N-k)^+]}{\partial \widehat{w}} = \int_0^{\frac{b}{G(\widehat{w})}} \ell'(q\,G(\widehat{w}))\,g(\widehat{w})\,q\,f(q)dq, \tag{1.22}$$

so that

$$\frac{\partial \mathsf{E}[\Pi]}{\partial \widehat{w}} = \int_0^{\frac{b}{G(\widehat{w})}} \left[p-(c+r)\ell'(q\,G(\widehat{w}))\right] g(\widehat{w})\,q\,f(q)dq. \tag{1.23}$$

From part (ii) of Proposition 1.7 we know that $\frac{p}{c+r} = \ell'(b_m^*)$, and together with the fact that $\ell(b)$ is increasing and convex, it implies that, like $\frac{\partial \mathsf{E}[\Pi]}{\partial b}$, the partial derivative $\frac{\partial \mathsf{E}[\Pi]}{\partial \widehat{w}}$ is positive and increasing for $b \leq b_m^*$ and decreasing for $b > b_m^*$.

It follows that it is the sign of the last term in (1.20), $\frac{d\widehat{w}}{db}$, that dictates how b_m^* compares to b_s^*. As the proof of the following proposition (in Appendix A.1.2) shows, the expression for $\frac{d\widehat{w}}{db}$ is complex. Nevertheless, we can use it to demonstrate the following relationship between b_m^* and b_s^*.

Proposition 1.9. (Optimal Strategic Booking Limit)
Suppose $\exists\ p \in (0, (1-\alpha)v)$ *and* $c \in (0, \overline{w})$ *for which* $b_m^* \in (k, \infty)$ *induces a profit-making equilibrium* $\widehat{w} \in (\underline{w}, \overline{w})$*. Then we have the following.*
(i) For any given $b > k$*, if* $\beta \geq \sqrt{(v - \frac{p}{1-\alpha})\frac{g(\widehat{w})}{G(\widehat{w})}}$*, then* $\frac{d\widehat{w}}{db} < 0$*.*
(ii) In turn, if $\beta > \sqrt{(v - \frac{p}{1-\alpha})\frac{g(\widehat{w})}{G(\widehat{w})}}$ *for all* $b > k$*, then* $b_s^* < b_m^*$*.*

Proposition 1.9 shows that, when bumping probabilities are high enough, an airline that accounts for customers' equilibrium response to its policy overbooks less, as compared to the myopic alternative.

The proposition's sufficient condition is complex, however. In addition to depending on problem parameters and distributional assumptions, it requires that a relationship between the equilibrium quantities β and $\widehat{w}$ is satisfied for any $b > k$.

Despite this difficulty, we observe that, as $p \to (1-\alpha)v$, the proposition's sufficient condition is in fact trivially satisfied for all b, an insight that motives our analysis below. In particular, in §1.4.3 we will show that, in the full profit-maximization problem described in (1.10), this boundary case is optimal at the same time that it allows us to reduce the complexity of our analysis.

1.4.3. Optimal Overbooking Policies

We now consider the full profit-maximization problem as described in (1.10). To derive optimal admissible overbooking policies $\xi \in \Xi$, we could naively compute and jointly analyze the first order conditions for all three decision variables, (p, b, c). As we have seen in the equilibrium analysis of §1.4.2, however, even a standard analysis of the first order conditions for b is delicate. Rather, in this section we will show that we can reduce the complexity of the analysis by exploiting structural properties of the problem. We begin in §1.4.3 by characterizing the myopic policy, which ignores the demand effects induced by its parameter choices. Then in §1.4.3, we develop the structural properties that allow us to reduce the problem, and in §1.4.3 we characterize optimal strategic policies.

Myopic Overbooking Policies

As in §1.4.2, we begin by considering an airline that does not recognize that the marginal customer's hassle cost, $\widehat{w}$, is an equilibrium reaction to β and, in turn, to (p, b, c). Instead it believes that $\widehat{w}$ is fixed, and it uses a *myopic overbooking policy* that maximizes the objective function of (1.10) without considering the problem's equilibrium constraints. We call the associated optimal myopic solution (p_m^*, c_m^*, b_m^*).

In this case, we find that the airline charges the maximum price and chooses not to compensate bumped passengers at all. More formally we have the following.

Proposition 1.10. (Optimal Myopic Overbooking Policy)
A myopic airline sets $p_m^* = (1-\alpha)v$ and $c_m^* = 0$. When $v < r$, it selects a finite optimal booking limit $b_m^* = \max\left\{\ell'^{-1}\left(\frac{(1-\alpha)v}{r}\right), k\right\}$. Otherwise, b_m^* is infinite.

The rationale for the policy is as follows. For a fixed $\widehat{w}$, any given choice of b determines sales $S = \min\{b, QG(\widehat{w})\}$, without regard to the value of p or c, and we see from (1.13) that, by maximizing p and minimizing c, the airline maximizes per-customer contribution and expected profit. In turn, the optimal myopic booking limit is the same as that in Proposition 1.7 for $p = (1-\alpha)v$ and $c = 0$.

The customer equilibrium obtained from the policy, of course, differs. From part (i) of Lemma (1.4) we see that, when $p = (1-\alpha)v$ and $c = 0$, myopic solutions that recommend $b_m^* > k$ will not generate a customer equilibrium. In contrast, those that set $b_m^* = k$ obtain

$\widehat{w} = \overline{w}$ and earn expected profits of $\mathsf{E}[\Pi(\,(1-\alpha)v, k, 0\,)] = (1-\alpha)v\,\mathsf{E}[\min\{k, QG(\overline{w})\}]$.

Problem Reduction

In contrast to the myopic policy, above, a *strategic overbooking policy* considers customer response when solving (1.10). In this case, we can show that there exist optimal strategic policies that set $p = (1-\alpha)v$, a result that greatly simplifies our analysis and affords a number of insights. The result holds straightforwardly for overbooking policies that induce interior equilibria, for which $U(\beta, \widehat{w}) = 0$, and we begin by first eliminating boundary cases from consideration.

Lemma 1.11. (Boundary Equilibria Not Optimal)
Any optimal strategic overbooking policy induces a customer equilibrium with $U(\beta, \widehat{w}) = 0$.

To demonstrate the result, we need to rule out two cases. The first, when the equilibrium is $U(\beta, \underline{w}) < 0$, induces expected sales of zero and is not profit making. Therefore, it is dominated by any policy that charges a positive price and does not overbook. The second, more interesting case, occurs when the equilibrium is $U(\beta, \overline{w}) > 0$, so that $G(\widehat{w}) = 1$ and all customers receive a strictly positive surplus. The lemma's proof shows that, when $U(\beta, \overline{w}) > 0$, we must also have $p < (1-\alpha)v$, and the airline can raise the price a bit without affecting unit demand and thereby increase expected profits. Thus neither boundary case can be optimal, and when searching for optimal strategic overbooking policies, we can consider only policies that induce customer equilibria with $U(\beta, \widehat{w}) = 0$.

Conversely, suppose the airline employs an admissible policy (p, c, b) for which $U(\beta, \widehat{w}) = 0$. Then we can use (1.1)-(1.2) to rewrite the equilibrium condition as

$$p \;=\; (1-\alpha)v + (1-\alpha)\,\beta\,(c - \widehat{w}). \tag{1.24}$$

We see that, for a fixed equilibrium pair $(\beta, \widehat{w})$, there is an infinite set of (p, c) pairs that satisfy the equilibrium equation (1.24). Recalling that $N = (1-\alpha)\min\{b, QG(\widehat{w})\}$, $(N - k)^+ = \ell(\min\{b, QG(\widehat{w})\})$, and $\beta = \mathsf{E}[N]/\mathsf{E}[(N-k)^+]$, we see that a given b and $\widehat{w}$ uniquely define β, without regard to p or c.

Now consider an airline with an overbooking policy (p, b, c) that obtains equilibrium $(\beta, \widehat{w})$. If the airline maintains the same b, then there will be an infinite set of price-bumping-compensation pairs, (p', c') – including the original (p, c) – that will satisfy the equilibrium equation (1.24) and maintain the same equilibrium $(\beta, \widehat{w})$. Furthermore, if we use (1.24) to substitute for p in the profit expression (1.13), we see that for any p' and c' that satisfies (1.24), including the original p and c,

$$\begin{aligned}\mathsf{E}[\Pi(p', b, c')] &= \left[(1-\alpha)v + (1-\alpha)\beta(c' - \widehat{w}) - (1-\alpha)\beta(c' + r)\right] \mathsf{E}[S] \\ &= \left[(1-\alpha)v - (1-\alpha)\beta(\widehat{w} + r)\right] \mathsf{E}[S] \;=\; \mathsf{E}[\Pi((1-\alpha)v, b, \widehat{w})].\end{aligned}$$

In particular, the substitution of $(1-\alpha)v$ for p and $\widehat{w}$ for c obtains the same equilibrium and expected profit. Thus, we have the following result.

Lemma 1.12. (Multiple Equivalent Policies)
For any admissible policy (p, b, c) for which $\beta > 0$ and $U(\beta, \widehat{w}) = 0$, there exists an infinite set of alternative policies with the same booking limit, $b' \equiv b$, and alternative price and bumping compensation,

$$p' \in \left[\max\{0, (1-\alpha)(v - \widehat{w}\beta)\}\,,\, (1-\alpha)v\right] \qquad \text{and} \qquad c' = \left(\widehat{w} - \frac{v}{\beta}\right) + \left(\frac{p'}{(1-\alpha)\beta}\right), \tag{1.25}$$

with the same equilibrium $(\beta, \widehat{w})$ and expected profits $\mathsf{E}[\Pi(p, b, c)] = \mathsf{E}[\Pi(p', c', b')] = \mathsf{E}[\Pi((1-\alpha)v, b, \widehat{w})]$.

Lemma 1.11 shows that an optimal policy induces an interior equilibrium, and Lemma 1.12 in turn shows that the airline can match the performance of such a policy using an alternative with price $p = (1-\alpha)v$. Together, the two imply we need only look at the following sub-class of admissible policies.

Proposition 1.13. (Problem Reduction)
If there exists an optimal strategic overbooking policy $\xi \in \Xi$, then there exists an opti-

mal strategic policy that sets $p_s^* = (1-\alpha)v$*, induces an interior equilibrium* $U(\beta, \widehat{w})$*, and optimizes* (1.10).

Thus, in the search for an optimal strategic policy, the airline need only consider $p = (1-\alpha)v$. With this insight, we continue our analysis for $p = (1-\alpha)v$ below.

Strategic Overbooking Policies

Proposition 1.13 allows the airline to reduce the complexity of its profit maximization problem in two ways. It can fix the decision variable $p = (1-\alpha)v$ in (1.10) and optimize over only (b, c). In addition, as (1.24) shows, given $p = (1-\alpha)v$, any interior equilibrium must have $c \equiv \widehat{w}$. This fact, in turn, has four additional implications.

First, whatever c the airline chooses uniquely determines $\widehat{w}$, without regard to the booking limit b, thereby eliminating the effect of b on $\widehat{w}$. From (1.20), we recall that $\frac{d\widehat{w}}{db}$ is a source of significant complication in the analysis of strategic booking limits, and with $\frac{d\widehat{w}}{db} = 0$ we eliminate this difficulty.

Second, given $p = (1-\alpha)v$ and a booking limit, b, the choice of $c \equiv \widehat{w}$ effectively sets a unique equilibrium because $\mathsf{E}[S]$, $\mathsf{E}[(N-k)^+]$, and $\beta = (1-\alpha)\mathsf{E}[S]/\mathsf{E}[(N-k)^+]$ are uniquely determined by b and $\widehat{w}$. Thus, the sufficient conditions of Lemma 1.4 part (ii) are no longer necessary: when $p = (1-\alpha)v$ the uniqueness of interior equilibria holds for any hassle-cost-distribution $G(\,\cdot\,)$.

Third, from (1.24) we also see that any customer with hassle-cost $w \leq c \equiv \widehat{w}$ is willing to buy a ticket – no matter how high the equilibrium β – and is, in fact happy to be bumped, should she be bumped. Conversely, any customer with hassle-cost $w > c \equiv \widehat{w}$ will never buy a ticket as long as $\beta > 0$, no matter how low. (When $b = k$ then $\beta = 0$, $\widehat{w} = \overline{w}$.) Thus when $p = (1-\alpha)v$, customers do not need to know or carefully estimate β when making purchase decisions. In additional to their own preference and no-show information, (v, w, α), they need only know the ticket price, p, whether or not the airline overbooks – whether $\beta = 0$ or $\beta > 0$ – and, if it overbooks, what bumping compensation, c, the airline offers, information that the airline can credibly communicate to its customers.

Fourth, we can tighten the lower bound for admissible values of c. On the one hand, because $G(\underline{w}) = 0$ effectively shuts down demand, we need never consider $c < \underline{w}$. On the other hand, because there is no need for overbooking when $\overline{q}G(c) \leq k$, we similarly never need consider $c < G^{-1}(k/\overline{q})$. Here, the relevant lower bound would still be $c = 0$ for $\overline{q} = \infty$. Together with the fact that we are fixing p, we define a subclass of Ξ, which we call *reduced* admissible policies, $\Xi_R \subseteq \Xi$ for which $p = (1-\alpha)v$, $\max\{\underline{w}, G^{-1}(k/\overline{q})\} \leq c\overline{w}$, and $k \leq b$. We know that there exists an optimal policy $\xi \in \Xi_R$.

Therefore, with $p = (1-\alpha)v$ and $c \equiv \widehat{w}$, we can further simplify the airline's original optimization problem (1.10). Given $\widehat{w} \equiv c$ we define unit sales and loss as direct functions of b and c,

$$S \;=\; \min\{b, QG(c)\} \qquad \text{and} \qquad (N-k)^+ \;=\; \big(N(S,\alpha)-k)^+\big)\,, \tag{1.26}$$

so that the airlines profits are

$$\Pi((1-\alpha)v, b, c) \;=\; (1-\alpha)v\, S - (c+r)\,(N-k)^+ \;. \tag{1.27}$$

In turn, the airline can optimize

$$\max_{\xi \in \Xi_R} \mathsf{E}[\Pi((1-\alpha)v, b, c)], \tag{1.28}$$

without explicit equilibrium constraints, to identify an optimal strategic overbooking policy.

Given an optimal strategic price, $p_s^* = (1-\alpha)v$, we can differentiate (1.28) with respect to b,

$$\frac{\partial \mathsf{E}[\Pi]}{\partial b} \;=\; (1-\alpha)v\,\mathsf{P}\{QG(c) > b\} \;-\; (c+r)\,\ell'(b)\mathsf{P}\{QG(c) > b\}, \tag{1.29}$$

and with respect to c,

$$\frac{\partial \mathsf{E}[\Pi]}{\partial c} \;=\; -\mathsf{E}[(N-k)^+] \;+\; g(c)\int_0^{\frac{b}{G(c)}} [(1-\alpha)v \;-\; (c+r)\ell'(qG(c))]\, qf(q)\, dq, \qquad (1.30)$$

and use their first order conditions to identify optimal strategic booking limit, b_s^*, and bumping compensation, c_s^*.

The first order condition with respect to b is precisely that in (1.19) but with $p = (1-\alpha)v$, and for a given c the results of Proposition 1.7 hold here as well. The following proposition describes how, as the value c_s^* increases, b_s^* systematically decreases from ∞ down to k.

Proposition 1.14. (Booking Limit for Optimal Strategic Overbooking Policy)
(i) If $c_s^* \leq v - r$, *then* $b_s^* = \infty$.
(ii) If $v - r < c_s^* < \frac{(1-\alpha)v}{\ell'(k)} - r$, *then* $b_s^* = \ell'^{-1}\left(\frac{(1-\alpha)v}{c_s^*+r}\right)$.
(iii) If $c_s^* \geq \frac{(1-\alpha)v}{\ell'(k)} - r$, *then* $b_s^* = k$.

Part (i) of the proposition follows directly from part (i) of Proposition 1.7, and parts (ii) and (iii) of the proposition follow from part (ii) of Proposition 1.7. When $b_s^* = k$, we also know that $c_s^* = \overline{w}$, since no one is bumped, and the high compensation ensures maximum demand. Note that $\ell'(k) < (1-\alpha)$, so the ordering of c_s^* in parts (i)-(iii) is well defined.

Proposition 1.14 is interesting for two reasons. First, it reflects that fact that, for $p_s^* = (1-\alpha)v$ and a given c_s^*, the optimal booking limit is simply calculated as the myopic optimal booking limit for that p_s^* and c_s^*. In turn, it shows that an optimal overbooking policy may be found by a line search over potential values of c.

1.5. Overbooking with a Bumping Auction

We now consider overbooking policies in which the airline compensates bumped customers using an auction. In this case, whether a passenger is bumped and how much she is compensated depend on the magnitude of her hassle cost, w. We define the primitives of the auction model and format in §1.5.1.

Our analysis of the auction-based compensation scheme yields three sets of insights. In §1.5.2 we show that, given use of an auction without a pre-set limit on bumping compensation, customers are always happy to be bumped, and as with the fixed-compensation model, there exist effective auction-based schemes in which the ticket price equals the expected value of flying. Then in §1.5.3 we show that, as expected, the use of auction-based bumping compensation can discriminate effectively among customers with lower and higher hassle costs to lower expected bumping compensation, and an auction with a pre-announced upper limit on its potential bumping compensation, though not necessarily optimal, always increases the airline's expected profit in comparison to the analogous fixed-compensation scheme of §1.4.3. In cases in which the auction's expected total bumping compensation is convex in the number of tickets sold, we can characterize the behavior of specific policy parameters. In §1.5.4 we consider these convex cases, we characterize optimal policy parameters, and we identify an optimal auction-based overbooking policy.

1.5.1. Primitives for the Auction Model

In most respects, the primitives of our model of overbooking with auction-based bumping compensation parallel those of the fixed-compensation model defined in §1.4. The airline operates a flight with k seats. It sells tickets up to booking limit, b, at price, p. If $b > k$ and customers are bumped, it pays rerouting cost, r, for each customer it bumps. The airline may decide to impose an upper bound, c_a, on the compensation that it is willing to pay to bumped passengers, an analogue of the fixed compensation, c, paid in §1.4. In the context of an auction, we will call the upper bound, c_a, a *cap*.

Customer attributes also remain the same. Potential customers are homogenous in the value they obtain from the flight, v, as well as in their no-show probability, α, and they are heterogenous in the hassle cost, w. We continue to model hassle costs as i.i.d. samples drawn from a common random variable, W, over support $0 \leq \underline{w} < \overline{w} \leq \infty$, with CDF $G(w)$ and density $g(w) > 0$ over its support.

As before, potential demand, Q, is random, with support $0 \leq \underline{q} < \overline{q} \leq \infty$ and CDF

$F(q)$. In §1.5.2 and §1.5.3, in which we provide preliminary results regarding auction-based overbooking policies, we assume that the support of Q, S, and N is integral. In §1.5.4, in which we further characterize the optimal booking limit, b_a^*, and compensation cap, c_a^*, we return to §1.4's assumption that these random variables are continuous, and we define an analogous continuous approximation for the expected auction-based bumping compensation. In both cases, our original definition of S, N, and $(N-k)^+$ in (1.3)–(1.5) continue to hold.

The auction model differs from that in §1.4 in that, when the number of customers who show up for the flight, n, exceeds the flight's capacity, k, the airline does not choose $(n-k)$ customers to bump at random. Rather, it chooses which passengers to bump using an auction.

1.5.2. Auction with No Cap

We begin our analysis by analyzing an auction for which the compensation paid to bumped customers is not limited by a pre-determined cap. As in the fixed-compensation model of §1.4, the airline must also decide on a ticket price and booking limit.

The auction format is a reverse form of a so-called uniform price, multi-unit auction with single-unit demand and works as follows. If the number of people who show up for the flight, n, exceeds the flight's capacity, k, then each of the n potential passengers is asked to reveal the minimum compensation, ϖ, she would require to give up her seat and be rerouted on another flight. We call ϖ a *bid* and assume that all n passengers submit their bids simultaneously and independently. The airline observes the n bids and orders them from smallest to largest. We describe them using the notation of order statistics: $\varpi_{1:n} \leq \varpi_{2:n} \leq \ldots \leq \varpi_{n:n}$. The airline then bumps the customers with the $n-k$ lowest bids and pays each of bumped passenger $\varpi_{n-k+1:n}$, the lowest bid among those passengers the airline allows to board the flight.

We have not yet described which customers decide to buy tickets, or not, so we do not yet know the distribution of the w's of those who show up for the flight. Nevertheless, we can show that this auction format motivates customers who paid for a ticket and show up for the flight to truthfully bid their hassle costs. This fact, in turn, allows us to characterize which customers buy tickets and the airline's optimal price, which we denote as p_a^*.

Proposition 1.15. (Properties of the Auction with No Cap)
Suppose that, when $n > k$ customers show up for a flight, the airline runs a reverse, uniform price, multi-unit auction. Then we have the following.
(i) Customers' optimal bids match their underlying hassle costs: $\{\varpi_{1:n} = w_{1:n}, \ldots, \varpi_{n:n} = w_{n:n}\}$.
(ii) All customers are willing to purchase tickets, irrespective of their hassle cost $w \in [\underline{w}, \overline{w}]$.
(iii) The airline's optimal price is $p_a^* = (1 - \alpha)v$.

Part (i) of the proposition is well-known. For example, Section 13.4.2 in Krishna (2010) notes the dominance of truthful bidding and resulting efficiency for this auction format, and for completeness we provide an explicit proof of the former in the appendix.

For part (ii) we note that, given the optimality of customers' bidding their true hassle costs, the airline offers compensation of $w_{n-k+1:n}$ to each of the $(n - k)$ customers who it bumps, an amount that is, by definition, at least as great as any of their hassle costs. Thus, for any customer, the expected value of being bumped is always non-negative and, given the opportunity, any customer will purchase a ticket. This is an analogue of the equilibrium $\widehat{w} \equiv \overline{w}$ in the fixed-compensation scheme.

We note that, because customers are happy to be bumped, they need not estimate the chances of being bumped – a fact we pointed out in the model in §1.4 – and we need not define or evaluate analogues of the equilibrium expressions, (1.1)–(1.2). Rather, as in (1.26) we can define $S = \min\{b,\, Q\,G(\overline{w})\} = \min\{b,\, Q)\}$ and $(N(S, \alpha) - k)^+$ as direct functions of b and $\widehat{w} = \overline{w}$.

Part (iii) of the proposition then follows $\widehat{w} \equiv \overline{w}$. In particular, the fact that numbers of passengers ticketed, S, and bumped, $(N(S, \alpha) - k)^+$, are independent of the price, implies that, to maximize expected profit, the airline should simply increase the price to its maximum, $p_a^* = (1-\alpha)v$.

Given the same optimal price in both the fixed-compensation and auction-compensation models, $p_s^* = (1-\alpha)v = p_a^*$, it is natural to ask how the performance of two classes of overbooking policies compare. On the one hand, the auction format should lower expected bumping compensation by choosing low-cost customers to bump. On the other hand, when the optimal level of fixed bumping compensation, c_s^*, falls strictly below $\overline{w}$, there exists the possibility of obtaining sample realizations, for which $w_{n-k+1:n} > c_s^*$, a fact that may drive $\mathsf{E}[w_{n-k+1:n}]$ to exceed c_s^*. In the next section, we will consider a capped version of the auction scheme that allows us to make a direct comparison.

1.5.3. Auction with a Cap

An airline that runs the auction scheme described §1.5.2 may potentially end up paying very high total bumping compensation, depending on the hassle cost distribution and its sample realizations. One way to limit the payout is to place a cap on the compensation paid to each customer who is bumped. For example, both United and Delta Airlines have upper bounds that they publicize, United offering up to \$10,000 and Delta up to \$9,950 in vouchers that can be applied to the ticket price of future flights (Martin, 2017; Tuttle, 2017).

In this section we consider this form of cap, which works as follows. The airline offers tickets for a flight at price p and publicly discloses that, when bumping customers, it uses an auction with cap, c_a, on the maximum compensation paid. In the event that $n > k$ passengers show up for the flight, the airline runs an auction, as before, but now limits the compensation per bumped customer to $\min\{c_a, w_{n-k+1:n}\}$. We note that the uncapped auction of §1.5.2 is equivalent to a capped auction with cap $c_a = \overline{w}$.

The presence of a cap $c_a < \overline{w}$ has the potential to eliminate the dominance of truthful bidding and to render an analysis of the auction unmanageable. For the special case of a price $p = (1-\alpha)v$ – which we already know is optimal for overbooking policies with fixed-compensation and for uncapped auction-based policies – the analysis remains tractable, however. In particular, we have the following.

Proposition 1.16. (Properties of the Auction with a Cap)
Suppose that the airline sets the price $p = (1-\alpha)v$ and $b > k$. When $n > k$ customers show up for a flight, it runs a reverse, uniform price, multi-unit auction with compensation cap $c_a \leq \overline{w}$. Then we have the following.
(i) Customers are willing to purchase tickets, if and only if their hassle costs are $w \leq c_a$.
(ii) Customers' optimal bids match their underlying hassle costs: $\{\varpi_{1:n} = w_{1:n}, \ldots, \varpi_{n:n} = w_{n:n}\}$.

The proposition's results follow the logic of the fixed-compensation scheme with $p = (1-\alpha)v$. In the auction setting, a price of $p = (1-\alpha)v$, together with a pre-announced cap of c_a, ensures that customers with $w > c_a$ have a surplus of $-p + (1-\alpha)v = 0$ on every sample path on which they are not bumped and a surplus of $-p + (1-\alpha)v + (c_a - w) < 0$ on every sample path on which they are, so the expected value of their purchasing a ticket is negative.

This demonstrates the "only if" statement of part (i). If we then consider a truncated hassle-cost distribution $W(c_a)$ with upper bound $\overline{w} = c_a$ and CDF $G_{c_a}(w) = G(w)/G(c_a)$, we return to the setting in §1.5.2 of an auction with no cap, one in which every customer with hassle cost $w \in [\underline{w}, \overline{w}]$ is willing to purchase a ticket. The "if" statement of part (i) and the statement of part (ii) then follow from parts (i) and (ii) of Proposition 1.15.

The cap on the auction's bumping compensation lets us directly compare this auction format's expected profits to those of analogous fixed-compensation policies. As in (1.26) we can define the number of units sold, $S = \min\{b, QG(c_a)\}$, and in turn number of bumped customers, $(N(S,\alpha) - k)^+$, as functions of the cap, c_a. To define expected bumping compensation, we let $P_N(n|s)$ denote the probability that n among s ticketed passengers show up for the flight, and we let $w(c_a)_{n-k+1:n}$ denote the $n-k+1$st ordered hassle cost from a truncated hassle-cost distribution $W(c_a)$. We then define the expected total bumping compensation, given s tickets are sold, as

$$C(s, c_a) = \begin{cases} 0 & \text{if } s \leq k, \text{and} \\ \sum_{n=k}^{s} (n-k)\, \mathsf{E}[w(c_a)_{n-k+1:n}]\, P_N(n|s) & \text{otherwise,} \end{cases} \tag{1.31}$$

a conditional expectation that is an analog to $c\ell(s)$, the conditional expectation of total bumping costs in the fixed-compensation scheme. The resulting marginal expression for the expected bumping compensation, which randomizes (1.31) over S, becomes $C(S, c_a)$.

With these quantities defined, we can denote auction-based profits for price $p = (1-\alpha)v$ as

$$\Pi_a((1-\alpha)v, b, c_a) \;=\; (1-\alpha)v\, S - r\,(N-k)^+ \;-\; C(S, c_a), \tag{1.32}$$

where revenues equal the ticket price times number of tickets sold, re-routing costs equal r times the number of passengers bumped, and total bumping compensation equals $C(S, c_a)$.

Now consider an admissible overbooking policy with $p = (1-\alpha)v$, $b \geq k$, and fixed compensation $\underline{w} < c \leq \overline{w}$. Given discrete distributions for Q, S, and N in (1.26), we can directly compare the objective functions (1.27) and (1.32) to show that auction-based bumping compensation dominates fixed-compensation policies.

Proposition 1.17. (Auction with Cap Dominates Fixed Compensation)
Given any fixed-compensation policy with $p = (1-\alpha)v$, $b > k$, $\underline{w} < c \leq \overline{w}$, and equilibrium $\beta > 0$, an auction-based policy with the same price, $p = (1-\alpha)v$, the same booking limit, b, and an analogous cap, $c_a = c$, earns strictly higher expected profits: $\mathsf{E}[\Pi_a((1-\alpha)v, b, c)] > \mathsf{E}[\Pi((1-\alpha)v, b, c)]$.

The fixed-compensation and auction-based policies have the same expected revenues, $(1-\alpha)v\mathsf{E}[S]$, and the same expected rerouting costs, $r\mathsf{E}[(N-k)^+]$. While both policies have the same numbers of bumped passenger, $(N-k)^+$, in any realization for which $n > k$, the auction scheme's compensation is weakly lower by construction: $w(c)_{n-k+1:n} \leq c$. Given

$g(w) > 0$ there further exists a positive probability that $w(c)_{n-k+1:n} < c$, a strict inequality that carries over to expected profits.

Thus, any strategic overbooking policy with price $p = (1 - \alpha)v$ is outperformed by the analogous auction-based policy. This includes strategic overbooking policies for which b and c are optimized for the price $p = (1 - \alpha)v$. Proposition 1.13's results, that there exist optimal policies with $p = (1 - \alpha)v$ for continuously distributed S and $(N - k)+$, suggest that, in fact, there exists an auction-based compensation scheme that outperforms *any* fixed-compensation scheme.

1.5.4. Optimal Policy Parameters for Auctions

In §1.5.2 and §1.5.3 we were able to use relatively elementary arguments to provide two insights of interest regarding overbooking with auctions. First, as with overbooking policies with fixed bumping compensation, the optimal price for auctions without a cap on bumping compensation is $p = (1 - \alpha)v$. Second, given an additional cap on auction-based bumping compensation, c_a, any fixed-compensation policy with price $p = (1 - \alpha)v$ is outperformed by an analogous auction-based format. When there is a cap, the potential complexity of passenger bidding behavior when $p < (1 - \alpha)v$ prevents us from providing a sharper characterization, however. Nevertheless, by using the same type of continuous approximations we employed in §1.4, we can differentiate critical expressions to provide additional insight.

Primitives for the Continuous Approximation

As in §1.4, we assume that Q is a continuous random variable with density $f(q) > 0$ over its support, that S and $N(s, \alpha)$ are likewise continuous, and that the loss function $\ell(s) = \mathsf{E}[(N(s, \alpha) - k)^+]$ is characterized by Definition 1.2. In addition to these assumptions, we define an analogous continuous approximation to the conditional expectation of the auction-based bumping compensation, given the number of potential passengers, n.

We continue to let $\mathsf{E}[w(c_a)_{n-k+1:n}]$ represent the expected value of the order statistic that determines bumping compensation. Now, however, we assume that the expectation varies continuously with a continuously-defined n. Accordingly we let $P_N(n|s)$ denote the CDF of the conditional distribution of N, given s, with support $[0, s]$ and density $p_N(n|s) > 0$ over its support, so that

$$C(s, c_a) = \begin{cases} 0 & \text{if } s \leq k, \text{and} \\ \int_k^s (n-k)] \, \mathsf{E}[w(c_a)_{n-k+1:n}] \, p_N(n|s) \, dn & \text{otherwise,} \end{cases} \tag{1.33}$$

with analogous total bumping compensation $C(S, c_a)$, as in (1.31).

We note that a continuous approximation for $\mathsf{E}[w(c_a)_{n-k+1:n}]$ can be created in a number of ways. One longstanding method that works well for relatively large n uses the inverse CDF of the hassle-cost distribution to map back from the relevant fractile to its w value (Arnold et al., 2008). Recalling from §1.5.3 that the conditional distribution of the hassle cost, given a cap c_a, is G_{c_a}, we have $\mathsf{E}[w(c_a)_{n-k+1:n}] \approx G_{c_a}^{-1}\left(\frac{n-k+1}{n+1}\right)$. For notational simplicity, we define $\tilde{G}(c_a, n) \equiv G_{c_a}^{-1}\left(\frac{n-k+1}{n+1}\right)$.

We emphasize that, as with the loss function, our proofs do not depend on the particular distribution of the number of customers who show up. We do require, however, that $N(s, \alpha)$ is *stochastically increasing and convex* (SICX) in s. (See Section 6.A.1 in Shaked and Shanthikumar (1994).) In particular, this means that $\mathsf{E}[\psi(N(s, \alpha))]$ is increasing in s for all increasing $\psi(\cdot)$ and increasing convex in s for all increasing convex $\psi(\cdot)$. The binomial distribution, for example, is SICX in s.

Similarly, our proofs do not depend on the particular form of the approximation for $\mathsf{E}[w(c_a)_{n-k+1:n}]$, only on its differentiability with respect to n, the resulting differentiability of $C(s, c_a)$ with respect to s, and the properties of the latter's derivatives.

As with $\ell(\cdot)$ we sometimes write partial derivatives with respect to one argument by using a prime symbol and the variable of interest: for example we let $C'(s) \equiv \frac{\partial C(s,c_a)}{\partial s}$. With this notation in hand we can state a first result, that expected total bumping costs inherit the convexity properties of the per-passenger expected bumping cost.

Lemma 1.18. (Convexity of Auction-Based Expected Bumping Cost)
Suppose $N(s,\alpha)$ *is SICX in* s.
(i) If $\partial \mathsf{E}[w(c_a)_{n-k+1:n}]/\partial n \geq 0$*, then* $C'(s) > 0$.
(ii) If in addition $\partial^2 \mathsf{E}[w(c_a)_{n-k+1:n}]/\partial n^2 > 0$*, then* $C''(s) > 0$.

The fact that expected bumping cost per customer, $\mathsf{E}[w(c_a)_{n-k+1:n}]$ is itself increasing in the number of customers bumped is not surprising. It can, in fact, be proven for any distribution using the original, discrete representation of order statistics and a sample-path argument. Given this fact, the propositions below will not state $C'(s) > 0$ as an explicit assumption.

The convexity of the expected order statistic, however, depends more specifically on the form of the hassle-cost distribution, G, and we can use the approximation, $\tilde{G}$, defined above to provide some insight into the types of distributions for which it holds.

Lemma 1.19. (Convexity of the Approximation $\tilde{\mathbf{G}}$)
Suppose $N(s,\alpha)$ *is SICX in* s *and we use the specific approximation* $\mathsf{E}[w(c_a)_{n-k+1:n}] \approx \tilde{G}(c_a, n)$.
If for any $c_a \in (\underline{w}, \overline{w}]$, $g'(w) < 0$ *for all* $w \in [G^{-1}(\frac{G(c_a)}{k+1}), c_a]$*, then* $C''(s) \geq 0$.

Thus, a sufficient condition for the convexity of the approximation $\tilde{G}$ is a decreasing hassle-cost density. This is the same condition that part (iii) of Lemma 1.4 showed is sufficient for the uniqueness of a customer equilibrium. As with Lemma 1.4, we note that DFR distributions satisfy the condition.

Properties of the Booking Limit

When the airline sets a price of $p = (1-\alpha)v$, auction profits (1.32) are well defined, and we can differentiate expected profits with respect to relevant policy parameters. To develop the required FOC we begin by differentiating the expectation of (1.32) with respect to b.

$$\frac{d\mathsf{E}[\Pi_a]}{db} = (1-\alpha)v\,\frac{d\mathsf{E}[S]}{db} - r\,\frac{d\mathsf{E}[(N-k)^+]}{db} - \frac{d\mathsf{E}[C(S)]}{db} = 0. \tag{1.34}$$

Differentiating the expectation of (1.33) with respect to b we have

$$\frac{d\mathsf{E}[C(S)]}{db} \;=\; C'(b)\mathsf{P}\{Q > b\}. \tag{1.35}$$

and from (1.16), (1.18), and (1.35) we can write the FOC in (1.34) as

$$\frac{d\mathsf{E}[\Pi_a]}{db} \;=\; \mathsf{P}\{Q > b\}\left[(1-\alpha)v \;-r\,\ell'(b) \;-\; C'(b)\right] \;=\; 0. \tag{1.36}$$

To ensure that b_a^* is a local maximum, we need also to examine the second-order condition (SOC)

$$\frac{d^2\mathsf{E}[\Pi_a]}{db^2} \;=\; f(b)\left[(1-\alpha)v - r\ell'(b) - C'(b)\right] \;+\; \mathsf{P}\{Q > b\}\left[-r\ell''(b) - C''(b)\right]. \tag{1.37}$$

We now have the machinery needed to characterize the optimal booking limit, b_a^*. From the FOC we see that $((1-\alpha)v \;- r\,\ell'(b) \;-\; C'(b)) \;=\; 0$, which implies that the SOC is negative whenever $[-r\ell''(b) - C''(b)] < 0$. Recalling Definition 1.2, we know that $\ell''(s) > 0$, so a sufficient condition for $\frac{d^2\mathsf{E}[\Pi_a]}{db^2}$ to be negative is $C''(b) > 0$. Finally, from Lemma 1.18 we know that $C''(b) > 0$ whenever $\partial^2\mathsf{E}[w(c_a)_{n-k+1:n}]/\partial n^2 \geq 0$.

Proposition 1.20. (Optimal Booking Limit for the Auction)
Suppose $p = (1-\alpha)v$ and $C''(s) > 0$. Then there exists a unique optimal booking limit, b_a^*, with the following properties.
(i) If $\ell'(k) \geq \frac{(1-\alpha)v}{r}$ then $b_a^* = k$.
(ii) If $\ell'(k) < \frac{(1-\alpha)v}{r}$ and $\exists\, b \in (k,\infty)$ s.t. $C'(b) \geq (v-r)(1-\alpha)$, then (1.36) determines $b_a^* \in (k,\infty)$.
(iii) If $C'(b) < (v-r)(1-\alpha)$ for all $b \geq k$ then $b_a^* = \infty$.

The proposition's results follow from the FOC (1.36) and the fact that $\lim_{s\to\infty}\ell(s) = (1-\alpha)$. As with the myopic booking limit characterized in Proposition 1.7, the higher the rerouting and bumping costs, the smaller the booking limit. When expected marginal rerouting costs are high enough, there is no overbooking. With moderate levels, there is a finite level of overbooking. Finally, given the upper limit of $(1-\alpha)r$ on the marginal expected rerouting cost, small enough marginal bumping costs can lead to infinite booking limits.

Properties of the Cap

As the airline changes the cap on the maximum bumping compensation, $c_a < \overline{w}$, it systematically changes the hassle-cost distribution. In particular, if $c_a^1 < c_a^2$, then $\mathsf{P}\{W(c_a^1) \leq w\} \geq \mathsf{P}\{W(c_a^2) \leq w\}$ for all $w \in [\underline{w}, \overline{w}]$. That is, $W(c_a^2)$ is larger than $W(c_a^1)$ in the so-called *usual stochastic order.*

This difference helps us to characterize the effect of potential changes to the cap. To begin, the stochastic order immediately implies that, for any fixed n, expected bumping compensation, $\mathsf{E}[w(c_a)_{n-k+1:n}]$ increases with the cap. (See Theorem 1.A.3(b) in Shaked and Shanthikumar (1994).) In turn, for any underlying hassle-cost distribution, G, the expected total bumping cost grows more quickly with s.

Lemma 1.21. (Bumping Compensation Grows with the Cap)
For $c_a \in (\underline{w}, \overline{w})$ and $s > k$, (i) $\partial C(s, c_a)/\partial c_a > 0$, and (ii) $\partial^2 C(s, c_a)/\partial s \partial c_a > 0$.

In addition, when $C(s, c_a)$ is convex in s, we can use Lemma 1.21 to characterize how the optimal booking limit in Proposition 1.20 and expected profits change with the cap.

Proposition 1.22. (Optimal Auction Parameters)
For fixed p, let $b_a^*(c_a)$ be the optimal booking limit induced by c_a. Suppose $p = (1-\alpha)v$, $\underline{w} < c_a < \overline{w}$, $k < b_a^*(c_a) < \infty$ and $C''(s) > 0$. Then we have the following.
(i) The optimal booking limit, b_a^*, is decreasing in c_a.
(ii) The resulting expected profit, $\mathsf{E}[\Pi_a]$, is increasing in c_a.

Part (i) of the proposition follows from the FOC and the fact that $C'(s)$ is increasing in the cap. It implies that the optimal booking limit for the auction with no cap is the smallest among those for all auction policies we have considered. Part (ii) implies that, among all auctions-based policies with price $p = (1-\alpha)v$, an auction with no cap – which in the context of the proposition is the maximal cap – maximizes the airline's expected profits.

With analogous continuous-distribution models for both the fixed-compensation and auction-based overbooking policies, we can now make a direct comparison of the two classes of policy to conclude the following.

Proposition 1.23. (Optimality of Overbooking Policy)
(i) The optimal overbooking policy uses an auction to determine customers' bumping compensation.

When $N(s,\alpha)$ is SICX, $C''(s) > 0$, and the auction-based overbooking policy sets $p = (1-\alpha)v$, we also have the following.

(ii) The optimal cap on bumping compensation is effectively unbounded: $c_a^* = \overline{w}$.
(iii) The optimal booking limit b_a^* is defined as in Proposition 1.20.

While we have not ruled out the possibility that an auction with cap $c_a < \overline{w}$ and price $p < (1-\alpha)v$ could outperform the auctions we have considered, we do know that there exists a capped auction with price $p = (1-\alpha)v$ that outperforms all fixed-compensation policies. Thus, as part (i) states, there is *some* auction scheme that is optimal, a result that does not depend on the price or the convexity of $C(s)$. For $p = (1-\alpha)v$, SICX $N(s,\alpha)$ and convex $C(s)$, parts (ii) and (iii) highlight that the optimal overbooking policy sets $c_a^* = \overline{w}$ and requires only a minimal line search for the optimal b.

1.6. Numerical Experiments

Having analyzed how demand endogeneity affects overbooking under both fixed and auction schemes, a natural question that arises is whether or not the magnitude of the demand effect is significant. To address this question, we run two sets of numerical experiments that span a wide range of problem parameters and are meant to capture some typical values for flight capacities k, no-show rates α, customer valuations v, compensation amounts c, and rebooking costs r. The results are reported in Table 1 and Table 2.

The first set of experiments considers the demand effect for the narrower set of booking-limit policies considered in §1.4.2. Here, price, p, and fixed bumping compensation, c, are

	$\frac{b_m^*}{b_s^*}$	$\frac{\beta_m^*}{\beta_s^*}$	$\frac{\mathsf{E}[\Pi_s^*]}{\mathsf{E}[\Pi_m^*]}$
Min	1.00	1.00	1.00
Med	1.92	1.60	1.27
Mean	1.64	–	1.27
Max	2.40	∞	1.65

Table 1: Myopic vs Strategic Booking Limits

	$\frac{c_s^*}{\mathsf{E}[C]}$	$\frac{b_a^*}{b_s^*}$	$\frac{\beta_a^*}{\beta_s^*}$	$\frac{\mathsf{E}[\Pi_a^*]}{\mathsf{E}[\Pi_s^*]}$
Min	2.35	1.00	1.00	1.00
Med	4.34	1.02	2.97	1.01
Mean	4.54	1.09	6.24	1.02
Max	9.84	1.38	47.74	1.06

Table 2: Fixed Compensation vs Auction

taken as given, and the airline chooses a booking limit to maximize expected revenues. In these experiments, we quantify the benefit provided by the strategic booking-limit policies of §1.4.2, in which the airline recognizes that its booking limit affects customer demand, beyond that of the benchmark myopic setting of §1.4.2, in which it does not.

The 720 experiments cover parameter ranges designed to test a wide array of contexts, and we describe their details in Appendix A.1.3. For each of the 720 problem instances, we find the optimal myopic booking limit, b_m^*, and the optimal strategic booking limit, b_s^*. Customers react optimally to each booking limit, and the resulting bumping probabilities, β_m^* and β_s^*, as well as the resulting expected profits, $\mathsf{E}[\Pi_m^*]$ and $\mathsf{E}[\Pi_m^*]$, reflect the resulting customer equilibrium. For each of the 720 problem instances, we compare the results obtained from the use of the myopic and the strategic booking limits by calculating relevant ratios, and for each ratio we sort the results of the 720 problem instances from smallest to largest and report relevant distributional statistics.

Table 1 reports the summary statistics. The first column reports those for the ratio of the optimal booking limits, the second displays statistics for the ratio of equilibrium bumping probabilities, and the third results for the ratio of the expected profits. The table shows that the demand effect has a significant impact in all three cases. As suggested by Proposition 1.9, the optimal strategic booking limit is never higher than the myopic analog, and furthermore the myopic booking limit is on average 1.64 times higher (64% higher) than its strategic counterpart. The median ratio for the bumping probabilities is 1.60, and the mean is unbounded due to cases in which the strategic airline does not overbook, so that $\beta_s^* = 0$. Finally, the mean and median ratios for expected profits are 1.27, suggesting the an airline's ability to account for demand effects can have a significantly positive impact for its revenue management.

The second set of experiments considers full control of price, bumping compensation, and booking limit and compares the optimal strategic fixed-compensation schemes of §1.4.3 to the optimal auction schemes of §1.5.2. In these auction schemes there is no cap on bumping compensation. We construct 360 problem instances that systematically vary relevant parameters, and again we describe their details in Appendix A.1.3. In addition to the statistics reported in Table 1, Table 2 reports analogous summary statistics for the ratio of the optimal fixed-compensation value to the expected bumping cost associated with the optimal auction policy, $c_s^*/\mathsf{E}[C]$.

From Table 2 we see that optimal expected bumping cost per passenger is significantly lower for the auction-based scheme (mean ratio of 4.54) and that, in turn, this allows the airline to significantly increase booking limits (mean ratio of 1.09) and bumping probabilities (mean ratio of 6.24). Optimal expected profits are always higher for the auction-based policies, and while the improvement in expected profits is relatively modest (mean ratio of 1.02) when compared to the analogous improvements seen in Table 1, we note that a 2% improvement itself can be significant in the RM context.

1.7. Managerial Implications and Limitations

Overbooking is widely used in practice and studied in the RM literature. Existing models of overbooking do not explicitly account for the demand effects that can accompany the bumping of passengers, however. One possible reason for this absence may be that, despite the wide adoption of overbooking, bumping probabilities tend to be low, on the order of tenths of a percent, and managers may assume that there is limited room for revenue increases from policy improvements. Nevertheless, the potential importance of these demand effects has been highlighted by recent events in the air travel industry, such as the widely publicized 2017 incident of a passenger who was unwillingly dragged off United Express Flight 3411.

Our numerical experiments confirm that fixed-compensation policies that account for demand endogeneity can, in fact, significantly outperform those that do not and that the use of auction-based policies brings smaller but still-significant additional gains. These results suggest that significant benefits may accrue to airlines that incorporate demand effects into their overbooking models and move to more profitable and customer-friendly auction-based compensation schemes.

Our work also provides a theoretical basis to support these shifts in policy. In particular, our analytical results suggest that, in both the fixed and auction-based compensation settings, an effective means of managing overbooking is to ensure that customers who buy tickets are always fairly compensated for being bumped. This approach ensures that an airline maintains the good will of bumped customers and allows customers to decouple their initial purchase decisions from the possibility of being bumped: when deciding whether to purchase tickets, they need not know the bumping probability. These policies also allow the airline to significantly reduce the complexity of its overbooking policies.

These results are based on a relatively rich model of the operational context. We make limited distributional assumptions regarding customer demand and customer disutility from being bumped. Similarly, our results hold under relatively limited informational requirements on the part of customers, who need only know their own preferences, observable statistics on the part of the airline and, potentially, an estimate of the bumping probability. Again, in the context of the families of policies described above, customers who are assured of fair bumping compensation need not estimate the bumping probability.

At the same time, our model also makes some limiting assumptions that may be interesting to relax in subsequent work. In particular, our focus on the disutility of being bumped has motivated us to assume that relevant customer heterogeneity is captured by a hassle cost distribution and that customers otherwise share a common valuation v for the flight. It would be valuable to study the empirical relationship between value and hassle cost to see how well our assumption fits with practice and to provide a more refined characterization of the relationship between the two.

Furthermore, while our results suggest that each airline, when considered in isolation, would do well to move to an auction-based policy, we do not directly model competitive factors that might drive airlines to choose other overbooking schemes (Netessine and Shumsky, 2005). Similarly, in practice airlines' use of multiple fare classes can affect consumer behavior (Cohen et al., 2019), and it would be useful to extend our analysis to this broader setting.

CHAPTER 2 : Financing Inventory through Capped Initial Coin Offerings

2.1. Introduction

Initial Coin Offerings (ICOs) are an emerging form of fundraising for blockchain-based startups in which digital coins, also known as "tokens", are issued to investors in exchange for funds to help finance business. In many cases, tokens are generated on existing blockchains and their core value is backed either by the firm's future products/services in the case of "utility" tokens, or by the firm's future profits in the case of "equity" tokens. This new way of crowdfunding[1] startup projects has gained momentum since 2017 with the total amount raised skyrocketing to thirty billion dollars by the end of 2019 (source: icobench.com). The growth of ICOs is also challenging the dominance of traditional means of raising capital. During Q2 2018, ICO projects raised a total volume of $9.0 billion (Coinschedule, 2018), which is 56% of the amount raised by the US IPO market ($16.0 billion) or 39% of the amount raised by the US venture capital markets ($23 billion) during the same period, as reported by CB Insights (CB Insights, 2018) and PwC (Thomson, 2018).

Research Questions Following this trend, the academic literature on ICOs is also rapidly growing, particularly in finance and economics, where the focus has been on topics such as empirically characterizing the drivers of ICO success or on comparing this new form of financing to more traditional financing methods. There is also a growing literature in operations management studying the interplay between firm operations and financing decisions (see literature review for details). This chapter seeks to contribute to these literatures by focusing on ICOs for product market firms facing demand uncertainty, in an unregulated environment. We ask: How should assets (inventory) be tokenized as a function of product, firm and customer demand characteristics? That is, what type of tokens—utility vs.

[1] We discuss differences between token offerings and other early-stage financing methods in §2.3.7.

equity—and how many tokens should be issued, and how should they be priced? Further, how do these choices affect firm inventory decisions, and the odds of ICO failure or success? Finally, what are some of the salient features distinguishing ICOs from other forms of financing?

Properties of ICOs A typical ICO proceeds as follows. A startup first publishes a white paper with or without a minimum viable product for demonstration and then issues its platform-specific tokens. The typical white paper usually delivers the key information of the project, including the token sale model that specifies the token price, the sale period, the sales cap (if any), etc.[2] The tokens can have a variety of uses, but most commonly, they are either used for consumption of the company's goods and services once they become available (utility tokens), or offered as shares of the company's future profit (equity tokens). During the crowdsale, investors purchase tokens using either fiat currencies, or, more commonly, digital currencies such as Bitcoin and Ether.

While some successful ICOs were conducted by service platforms such as Ethereum and NEO, in this chapter we focus on ICO projects that involve the delivery of physical products instead; these types of ICOs are more recent, and hence, less well-understood. One striking example is that of Sirin Labs (Sirin Labs, 2019): a startup that produces smartphones and other types of hardware and software systems. In 2017, Sirin Labs was able to raise over $150 million from investors by offering them Sirin tokens (SRN). These tokens could subsequently be used to purchase the company's products and participate in its ecosystem, or be sold in the secondary market. Other relevant examples include Honeypod (Honeypod, 2018) that develops hardware serving as the main hub interconnecting various devices and providing traffic filtering, and Bananacoin (Bananacoin, 2018) which grows bananas in Laos.

ICOs can have multiple benefits. First, both the startup founders and investors have the opportunity of high financial gains from the potential appreciation of the tokens. Second, ICOs allow for faster and easier execution of business ideas because the ICO tokens generally

[2] We provide a condensed example of a white paper in Appendix A.2.1.

have secondary-market liquidity, and require less paperwork and bureaucratic processes, than the regulated capital-raising processes do. Third, ICOs provide the project team with access to a larger investor base as the stakeholders typically face less geographical restrictions.

On the other hand, just like the underlying blockchain technology, ICOs are still in their infancy and have some downsides. First, for the project teams, the failure rate of ICOs is high and increasing. Despite a rise in the total investment volume, nearly half of all ICOs in 2017 and 2018 failed to raise any money at all (Seth, 2018) and 76% of ICOs ending before September 2018 did not get past their soft cap (Pozzi, 2018), i.e., the minimum amount of funds that a project aims to raise. Benedetti and Kostovetsky (2018) claim that only 44.2% of the projects remain active on social media into the fifth month after the ICO. Second, the aspect of quick and easy access to funding with loose regulation attracts unvetted projects and even utter scams, making ICO investments risky. Some entrepreneurs portray deceiving platform prospects in the white papers in an attempt to raise as much money as they can before gradually abandoning their projects. In a review of 1450 ICO cases by the Wall Street Journal, 271 were susceptible to plagiarism or fraud. The profit-seeking yet ill-informed investors can become easy prey and have claimed losses of up to $273 million (Shifflett and Jones, 2018). Other disadvantages of ICOs include technical concerns such as the potential theft of tokens through hacks (Memoria, 2018).

Model To study some of these issues in the context of inventory tokenization, we adopt a game-theoretic approach with three types of players: a firm (token issuer), speculators (token traders) and customers (who buy the product). As in the Sirin Labs example mentioned earlier, the firm seeks to raise funds through an ICO to support the launch of a physical product it wishes to sell in the face of customer demand uncertainty. We first consider a utility-based ICO, whereby the tokens issued by the firm are tied to its (future) inventory. The ICO game develops over three periods. In the first period, the firm announces the total number of tokens available, the sales cap and the ICO token price, and sells tokens (up to the cap) to speculators who make purchase decisions strategically[3]. In the second period, the firm, facing uncertain customer demand, can put the funds raised in the ICO towards production of a single product. Importantly, to reflect the lack of intermediaries and the lax regulatory environment, we leave the firm with full discretion over what to do with the raised funds, including the option of fully shirking production and diverting raised funds to its pockets (moral hazard). In the final period, demand for the product is realized and customers buy tokens either directly from the firm (if the firm has any tokens remaining) or from speculators in the secondary market, and redeem these tokens for the product, if available, at an endogenously determined equilibrium token price. Finally, we compare this utility-based ICO model to an analogous equity-based ICO model, whereby the firm issues tokens that are tied to its future profits (if any), rather than to its future inventory.

Our base model assumes that the main source of risk is future demand uncertainty, rather than manufacturing technology. This setting is motivated by several practical examples, including the aforementioned case of Sirin Labs: prior to the ICO, the company had secured Foxconn, a major smartphone manufacturer (which also assembles the vast majority of Apple iPhones), as its main supplier. As such, the risk of manufacturing failure was arguably quite low. Indeed, the company went on to successfully produce their products and made them available to consumers. However, despite successful production and a very impressive ICO outcome in terms of total funds raised, the company's token market price dropped significantly in the months and years post ICO, because customer demand for their product fell well short of expectations. In other words, demand uncertainty can have a first-order effect on the token's market value. Nonetheless, for completeness, we also extend the base model to consider additional risks in the form of possible production failure.

Contributions Using this relatively simple and flexible model, we derive the optimal ICO price, token cap and production quantity as a function of operational and demand

[3]The literature on operational decisions in the presence of strategic agents includes Dana Jr and Petruzzi (2001), Cachon and Swinney (2009), Papanastasiou and Savva (2016), etc. Su (2010) and Milner and Kouvelis (2007) consider similar speculative behavior yet with no financing aspect.

characteristics, for both utility-based and equity-based coin offerings. We find that, despite rampant moral hazard, both types of ICOs can be successful under the right conditions.

Focusing first on utility ICOs, we show that these are analogous to a form of revenue-sharing contract between the firm and speculators, and we identify four key factors that are required for the success of an ICO: i) the existence of a liquid secondary market for the tokens, which provides an "exit ramp" for speculators, incentivizing them to participate in the ICO, even when they might otherwise not be interested in consuming the firm's products; ii) a minimum price-cost ratio of 2 (i.e., a 50% margin, or 100% markup) to provide enough incentive for the company to pursue production; iii) a minimum amount of tokens to be sold during the ICO, termed the "critical mass" condition, which ensures speculators break even in expectation; and iv) a maximum amount of tokens to be sold during the ICO, which defines a "misconduct threshold". Interestingly, when excessive funds beyond this threshold are raised (e.g., from over-optimistic investors) the firm is actually discouraged from pursuing production ex-post given it does not have enough "skin" left in the game. This provides a possible explanation for the loss of motivation or productivity post ICO of some well-funded startups in practice. While conditions i)-iv) suffice to prevent total market breakdown, they do not fully eliminate the adverse effects of moral hazard. Rather, in equilibrium, these lead to agency costs, underproduction and lower-than-optimal profits versus first best. Importantly, we show how these inefficiencies fade as the demand for the product increases and/or becomes less volatile, and as customers' willingness-to-pay increases.

We then turn attention to ICOs with equity tokens, which are also commonly known as security token offerings—STOs[4]. Although STOs have a relatively smaller market compared with ICOs, they gained much popularity in 2018, with total volume being almost seven times that of 2017 (from $65.59 million in 2017 to $434.95 million in 2018) (Blockstate, 2020). Unlike utility tokens that can be exchanged for the firm's products, equity tokens

[4]Throughout the chapter, we use "STOs" and "ICOs with equity tokens" interchangeably.

simply represent a share of an underlying asset, which in our model takes the form of a profit-sharing contract. Therefore, equity tokens are closely related to traditional equity stakes, though without voting rights. We show that, even though moral hazard and the inefficiencies it generates cannot be fully eliminated in this setting, these inefficiencies are less prominent compared to utility offerings, as long as profit-sharing can be credibly implemented, e.g., using auto-executable smart contracts. Assuming the latter holds, equity-type ICOs also have the advantage of not requiring a liquid secondary market for the tokens, as speculators can instead rely on smart contracts to receive their share of future cash flow and effectively "exit" the deal. Importantly, the advantage of equity tokens stems from their inherent ability to better align incentives, and hence continues to hold even in unregulated environments.[5]

Finally, we extend the model in two separate dimensions: i) We add the possibility of production failure to capture situations in which the firm's technology is risky and manufacturing success is not guaranteed. We find that this additional source of risk incentivizes the firm to keep a fraction of the ICO proceeds in a reserve fund to provide coverage in case production fails. This, in turn, can significantly exacerbate the moral hazard problem, reducing the chance of ICO's success. ii) We then consider the case where speculators require a minimum amount of expected return to invest in the ICO (the base model assumes their outside option is zero). We show that, consistent with practice, this assumption creates a wedge between the ICO's token price and the secondary market price, ex-post production.

ICOs vs Crowdfunding Our model distinguishes ICOs from other early-stage financing methods by capturing several unique features of ICOs, including the fundraising mechanism and the issuance of tokens, the existence of a peer-to-peer secondary market, and the nature of investors. In contrast to reward-based crowdfunding, for instance, there is no intermediary platform imposing a fundraising mechanism (e.g., Kickstarter uses and all-or-nothing mechanism). Rather, firms running ICOs have to determine how many tokens to issue/sell

[5] In our model, assuming regulation can effectively alleviate moral hazard, both ICOs and STOs reduce to the first best financing case without frictions.

during the initial round in addition to how many products to make. Another important difference that we highlight is that tokens allow the firm to disperse downside risks of future demand among the token holders, whereas in crowdfunding, campaign backers share downside risk only in terms of product failure (not in terms of future demand uncertainty). Finally, we show that the existence of the secondary market for the tokens is crucial in incentivizing investors to participate in utility-type ICOs, an important feature missing from crowdfunding. We refer the readers to §2.3.7 for a more detailed discussion and Table 3 for a summary comparison to other financing methods.

Literature Review Broadly speaking, this chapter contributes to the strand of literature at the interface of operations and finance that studies, among other things, different ways of financing inventory. Earlier works include Babich and Sobel (2004), Buzacott and Zhang (2004), Boyabatlı and Toktay (2011), Kouvelis and Zhao (2012), and Yang and Birge (2013), see Kouvelis et al. (2011) for a review of this literature. More recent papers include Boyabatlı et al. (2015), Yang et al. (2016), Iancu et al. (2016), Alan and Gaur (2018), Chod et al. (2019a).

As an alternative to traditional crowdfunding,[6] ICOs are understudied in the operations management literature. However, there are several recent theoretical studies in the finance literature that examine the economics of ICOs and cryptocurrencies. Most of them focus on peer-to-peer service platforms that allow decentralized trading. For example, Li and Mann (2018) and Bakos and Halaburda (2018) demonstrate that ICOs can serve as a coordination device among platform users. In a dynamic setting, Cong et al. (2018) consider token pricing and user adoption with inter-temporal feedback effects.

More closely related to our work are papers that model ICOs in business-to-customer settings. Catalini and Gans (2018) propose analysis of an ICO mechanism whereby the token value is derived from buyer competition. Malinova and Park (2018) suggest a variation on

[6]Refer to Section 2.3.7 for a discussion on the differences. For recent papers on crowdfunding, see Alaei et al. (2016), Babich et al. (2019), Belavina et al. (2019), Chakraborty and Swinney (2017), Chakraborty and Swinney (2018), Fatehi et al. (2017), Xu and Zhang (2018), Xu et al. (2018).

the traditional ICO mechanism that can mitigate certain forms of entrepreneurial moral hazard. Chod and Lyandres (2018) compare ICOs to VC financing. We adopt a similar approach modeling an ICO as a presale of the platform's partial future revenue, yet with an emphasis on operational details including demand uncertainty and inventory considerations. In particular, we incorporate stochastic demand of the products, rather than assuming that demand is observable before production (Catalini and Gans, 2018; Malinova and Park, 2018) or infinite (Chod and Lyandres, 2018). We believe ours is the first study to jointly optimize the operational decisions including sales cap, token pricing and production quantity, in the presence of strategic investors under demand uncertainty, and compare utility and equity (STO) token issuance in this context.

More broadly, the chapter contributes to a growing theoretical literature studying the economics of blockchain-based systems, see e.g., Biais et al. (2017); Chod et al. (2019b); Cong et al. (2018); Hinzen et al. (2019); Pagnotta (2018); Pagnotta and Buraschi (2018); Rosu and Saleh (2019); Tsoukalas and Falk (2019), and references therein.

Here and below, we first develop in §2.2 and solve in §2.3 the case of utility tokens, before examining equity tokens in §2.4.

2.2. Model: ICOs with Utility Tokens

Consider an economy with three types of agents: i) a monopolist firm, ii) investors termed speculators, and iii) firm's customers. The economy has three periods: i) The first period, termed "ICO", is the fundraising phase containing the firm's white paper that includes contract terms and the token crowdsale; ii) the second period, termed "production", covers firm's production decisions in the face of uncertain customer demand; iii) the third period, termed "market", covers the realization of customer demand, and market clearing for the product and any remaining tokens. The firm participates in all three periods. Speculators participate in the ICO and the market periods. Customers participate only in the market period.

Firm The firm has no initial wealth and seeks to finance production through a "capped" ICO. The firm has a finite supply of m total tokens that are redeemable against its future output (if any). In the ICO period, the firm maximizes its profits by choosing i) the ICO "cap", $n \leq m$, that is, the maximum number of tokens to sell to speculators in the ICO period, and ii) the ICO token price τ (in dollars per token). Subsequently, in the production period, the firm has the option to use any amount of funds raised through the ICO to finance the production of its output. To this end, the firm maximizes its total wealth, through a newsvendor-type production function (Arrow et al., 1951), by choosing quantity $Q \geq 0$ of a product with unit cost c (in dollars per unit) that it can later sell in the market period at a price p (in tokens per unit), in the face of uncertain customer demand D. To capture the lack of regulation in the current environment, we assume that the firm could divert all or a portion of the funds raised through the ICO, rather than engage in production (moral hazard).

In the final market period, demand is realized and the product is launched. The product can only be purchased using the firm's tokens—a restriction that has two consequences: i) it endows tokens with (potential) value ii) it implies price p represents the exchange rate between tokens and units (which departs from the traditional newsvendor setting). The firm competes with speculators to sell any remaining tokens it has post-ICO to product customers, e.g., through a "secondary" offering round. As opposed to the ICO round, there is no uncertainty in the secondary offering round as production is finished and demand is already realized. The equilibrium token price τ_{eq} (in dollars per token) as well as the product price p (in tokens per unit) are then derived through a market clearing condition, described below. Once the market clears, tokens have no residual value (since there is only a single production round and the tokens have no use on any other platform) and the game ends. We provide more details of the tokens' features and discuss their implications for speculators and customers in Appendix A.2.1.

To recap, the firm's decisions are the number of tokens to make available in the ICO to speculators n, the ICO token price τ, and production quantity Q.

Speculators Let z denote the total number of speculators with $z \gg m$ reflecting that ICOs have low barriers to entry. Speculators are risk-neutral, arrive simultaneously, and can each try to purchase a single token in the ICO at the price set by the firm, τ, that they expect to subsequently sell in the market period at an equilibrium price $\mathbb{E}[\tau_{eq}]$, where $\mathbb{E}$ is the expectation operator. If demand for tokens exceeds token supply in the ICO, speculators are randomly allocated token purchase rights. Speculators' expected profit u depends, among other things, on the expected price difference $\mathbb{E}[\tau_{eq}] - \tau$, denoted Δ, and on the total number of speculators that purchase tokens in the ICO, denoted s; formally:

$$u(s) = \frac{s}{z}\Delta(s), \quad \text{with} \quad \Delta(s) = \mathbb{E}[\tau_{eq}(s)] - \tau, \tag{2.1}$$

where the ratio s/z reflects random assignment of token purchase rights. We emphasize that the number of speculators s will be determined endogenously in equilibrium, and as we shall show later on, this number depends on the ICO cap n and the ICO token price τ. A necessary condition for $s(\tau, n) > 0$ speculators to participate in the ICO is

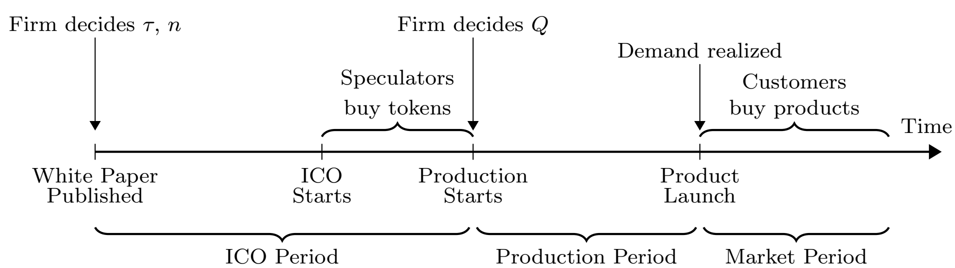
$$u(s(\tau, n)) \geq 0 \quad \text{(participation constraint).} \tag{2.2}$$

Note, as we show in Appendix A.2.2, assuming sequential rather that simultaneous arrival of speculators does not impact the results of the chapter. The proofs are written to cover both cases. Also note that the model readily extends to the case in which speculators are given the additional option of using their tokens to purchase the firm's product rather than selling their tokens to product customers.

Product Customers Customers who join the market after the product launch have a homogeneous willingness-to-pay v (dollars per unit) for the product that is strictly greater than the production cost c. As we shall see later on, v plays a critical role in the market clearing condition. Customers can buy tokens directly from the firm (if it has any tokens remaining in the market period) or from speculators, and they can redeem the tokens for the products. The demand for the product D is stochastic and we denote the cumulative distribution function of demand by $F(\cdot)$. For ease of analysis, we assume that $F(\cdot)$ is continuous and $F^{-1}(0) = 0$.

We summarize the timeline in Figure 1 below.

Figure 1: Sequence of Events



Market clearing Clearing occurs in the market period. Recall that the customers have a constant willingness-to-pay v (dollars per unit). This means that the dollar-denominated price of the product charged by the firm, which is equal to the product of the token-denominated price of the product p (tokens per unit) and the equilibrium market token price τ_{eq} (dollars per token), is at most v. Since the firm is a monopolist, it sets the dollar-denominated price to be exactly v, i.e., $p \cdot \tau_{eq} = v$. Therefore, p and τ_{eq} have an inverse relationship, and we have the following lemma due to the law of supply and demand.

Lemma 2.1. (Equilibrium Prices)

i) The equilibrium token-denominated price of the product is $p = m/\min\{Q, D\}$.

ii) The equilibrium token price in the market period is given by $\tau_{eq} = \frac{v}{m}\min\{Q, D\}$.

Part (i) of Lemma 2.1 implies that there are no idle tokens in the market period—the total token supply m can be redeemed for an amount $\min\{Q, D\}$ of products. Part (ii) implies the market clears. Specifically, customers' valuation for the total token supply equals their willingness-to-pay for all products that are purchased using these tokens, i.e., $\tau_{eq}\, m = v \min\{Q, D\}$. This equation addresses one of the most frequently asked questions regarding utility-based ICOs—what gives tokens their ultimate value? In our model, the value of platform-specific tokens depends positively on three factors: the quality of the product reflected by the customers' willingness-to-pay, the sales volume determined by the supply and demand for the products and the scarcity of tokens inversely determined by the total supply, m.

Note, the term $v \min\{Q, D\}$ resembles the revenue term in the traditional newsvendor setup where v corresponds to the fixed price. While a traditional newsvendor sells a quantity of products at a fixed price, the firm in our model sells a fixed number of tokens at (or below, to satisfy the participation constraint) a market equilibrium token price τ_{eq}. However, the newsvendor form emerges from the fact that τ_{eq} is tied to the product sales volume $\min\{Q, D\}$ via the market clearing condition.[7]

Firm's optimization problem The firm maximizes its expected dollar-denominated wealth at the end of the market period, denoted by Π, which consists of three terms: i) the total funds raised during the ICO, $\tau s(\tau, n)$, plus ii) the expected total funds raised in the secondary offering, $(m - s(\tau, n))\mathbb{E}[\tau_{eq}]$, minus iii) production costs $c\,Q$. The constraints are i) that production is funded by funds raised in the ICO, i.e., $c\,Q \leq \tau s(\tau, n)$ and ii) that speculators participate in the ICO, i.e., $u(s(\tau, n)) \geq 0$. Using the market clearing condition Lemma 2.1(ii), which ties token value τ_{eq} to sales $\min\{Q, D\}$, the firm's optimization

[7]Note, if the product had salvage value, this value would need to be included in the market clearing condition.

problem can be formally written as:

$$\max_{\tau,\, n} \left\{ \tau\, s(\tau, n) + \max_{Q} \left[(m - s(\tau, n))\, \frac{v}{m}\, \mathbb{E}[\min\{Q, D\}] - c\, Q \right] \right\} \tag{2.3}$$

subject to

$\tau\, s(\tau, n) - c\, Q \geq 0,$ (ICO funds cover production costs)

$u(s(\tau, n)) \geq 0.$ (speculators' participation constraint)

Recall that $s(\tau, n)$ is an equilibrium quantity, and we will show later how it depends on the firm's decisions variables, τ and n, and on Q (which itself depends on s, and hence τ and n).

2.3. Analysis: ICOs with Utility Tokens

In this section, we find the subgame perfect equilibrium using backward induction. We first consider (§2.3.1) the firm's last decision, the production quantity for fixed token price τ and ICO cap n, based on which we examine the speculators' equilibrium behavior (Section 2.3.2). We then calculate the optimal token price τ^* and ICO cap n^* (§2.3.3). Lastly, we present and discuss the equilibrium results in §2.3.4.

2.3.1. Optimal Production Quantity

We first consider the firm's last decision—the production quantity $Q(\tau, n, s(\tau, n))$, for fixed token price τ and ICO cap n. Here and below, we drop when possible the fixed arguments τ and n to ease exposition.

Proposition 2.2. (Optimal Production Quantity)

For a fixed token price τ, ICO cap n and number of speculators s, the firm's optimal production quantity $Q^*(s)$ is as follows.

i) If $0 < s < m\left(1 - \frac{c}{v}\right)$, then $Q^*(s) = \min\left\{F^{-1}\left(1 - \frac{cm}{(m-s)v}\right), \frac{\tau\, s}{c}\right\}$.

ii) If $s = 0$ or $s \geq m\left(1 - \frac{c}{v}\right)$, then $Q^*(s) = 0$.

Part (i) of Proposition 2.2 shows that production can occur only if the number of speculators that purchased tokens in the ICO, is below a fraction $\left(1-\frac{c}{v}\right)$ of all available tokens m. The first term inside the minimum operator, $F^{-1}\left(1-\frac{cm}{(m-s)v}\right)$, is the unconstrained optimal production quantity; interestingly, this term decreases in the number of speculators s. The second term, $\frac{\tau s}{c}$, captures the firm's budget constraint, i.e., the production costs cannot exceed funds raised in the ICO, and this term is increasing in s.

Part (ii) of Proposition 2.2 shows that if more than a fraction $\left(1-\frac{c}{v}\right)$ of all tokens have been sold in the ICO, the firm prefers not to use any of the funds raised for production, meaning, the firm "diverts" all money raised to its own pocket. We refer to this fraction as the firm's *misconduct fraction*,

$$1-c/v. \tag{2.4}$$

Clearly, as the willingness-to-pay v increases relative to the production cost c, the misconduct fraction increases, making the abandonment of production less likely.

We emphasize that this analysis does not suggest all crypto startups are scammers that would run away with any amount. Rather, it provides an explanation for the loss of motivation or productivity of some well-funded startups based on pure profit maximization reasoning, due to moral hazard in the absence of regulatory controls.

2.3.2. Equilibrium Number of Speculators and Participation Constraint

Having derived the firm's optimal production quantity for a given ICO design τ, n, we next examine the implications on speculators.

Lemma 2.3. (Speculator Equilibrium Properties) *Given initial token price τ and the sales cap n,*

i) The number of speculators who purchase tokens is $s^*(\tau, n) = n \cdot \mathbb{1}_{\{u(n)\geq 0\}}$,

ii) $s^*(\tau, n) \in \left[0, m\left(1-\frac{c}{v}\right)\right]$ such that complete fund diversion does not occur in equilibrium.

iii) Define $s_0(\tau) = \max\{0 < s \leq m : u(s) = 0\}$*. If* $s_0(\tau)$ *exists,* $s_0(\tau) < m\left(1 - \frac{c}{v}\right)$ *and* $u(s) < 0$ *for all* $s > s_0(\tau)$*.*

Lemma 2.3, part i) is a compact way to write that in equilibrium, the number of speculators purchasing tickets is equal to the ICO cap, as long as speculators' participation constraint is satisfied. This is because all speculators have the same expected profit, and hence, either n speculators will purchase tokens (if this expected profit is ≥ 0), or none will. Note, this result holds for sequential arrivals as well (see Appendix A.2.2).

Lemma 2.3, part ii) defines a lower and upper bound on the number of speculators that arises in equilibrium. The lower bound is trivial. The upper bound is a consequence of the firm's misconduct threshold derived in Proposition 2.2, and captures the fact that in any equilibrium, speculators strategically prevent their funds from being completely diverted.

Lemma 2.3, part iii) is a necessary technical condition ensuring speculator participation constraint holds, and hence, the success of the ICO. In the Sections 2.3.3 and A.2.2, we show that the existence of $s_0(\tau)$ depends on τ, which in turn depends on n, and discuss the implications.

2.3.3. Optimal Token Price and ICO Cap

Given the optimal production quantity (§2.3.1) and speculators' equilibrium behavior (Section 2.3.2), we now examine how the firm sets the profit-maximizing ICO token cap n^* and initial token price τ^*.

We show in Lemma 2.3 in Section 2.3.2 that the number of speculators $s^*(\tau, n) \leq m\left(1 - \frac{c}{v}\right)$. Given the speculators participating in the ICO buy 1 token each, we need not consider the case in which tokens $n > m\left(1 - \frac{c}{v}\right)$. We will first find the token price $\tau^*(n)$ for a given token cap $n \leq m\left(1 - \frac{c}{v}\right)$ and then maximize profit over the token cap n. The following Proposition guarantees the existence of a nonzero equilibrium token price τ^*.

Proposition 2.4. (Conditions for ICO Success)

The ICO succeeds if and only if

i) (critical mass condition) the firm sells more than $\frac{mc}{v}$ tokens in the ICO and,

ii) (price-cost ratio requirement) customers have a high willingness-to-pay such that $v > 2c$.

Part (i) of Proposition 2.4 shows that the firm should not set the ICO cap too low. Speculators expect non-negative returns only when more than a critical mass of tokens, $\frac{mc}{v}$, are sold in the ICO. This quantity increases in the production cost and decreases in customer willingness-to-pay. Recall from Section 2.3.2 that speculators would not invest more than the misconduct fraction. Combining these two results, we have that the ICO will only be successful when the misconduct fraction $m\left(1-\frac{c}{v}\right)$ is above the lower bound $\frac{mc}{v}$. This simplifies to the condition in Part (ii) of Proposition 2.4, $v > 2c$.

Next we find the optimal ICO token price $\tau^*(n)$ and the optimal ICO cap n^* assuming these two conditions are met. We show that for any fixed ICO cap n in the appropriate range ($n \in (\frac{mc}{v}, m\left(1-\frac{c}{v}\right))$), there exists a unique, positive and finite ICO token price $\tau^*(n)$ that maximizes (A.30) by extracting all utility from the speculators who participate strategically.

Given this result, we obtain a semi-closed-form solution of the optimal ICO cap n^*, and show that neither a small ICO cap that suppresses the production quantity nor a large cap that induces idle cash is profit-maximizing for the firm. The optimal ICO cap n^* allows the firm to raise just enough funds that can be credibly committed to production. We point interested readers to Appendix A.2.2 for detailed technical results.

2.3.4. The Equilibrium

Proposition 2.5. (Equilibrium Results)

i) If $v \leq 2c$, then the ICO fails.

ii) If $v > 2c$, then there exists a unique equilibrium where

(a) the ICO cap n^* satisfies $n^* \in (\frac{mc}{v}, \frac{m}{2})$ and

$\frac{v\,n^*}{c\,m}\,\mathbb{E}[\min\left\{D, F^{-1}(1-\frac{c\,m}{(m-n^*)v})\right\}] = F^{-1}(1-\frac{c\,m}{(m-n^*)v})$,

(b) the number of speculators is $s^* = n^*$*,*

(c) the ICO token price is $\tau^* = \frac{v}{m}\,\mathbb{E}[\min\left\{D, F^{-1}(1-\frac{c\,m}{(m-n^*)v})\right\}]$*,*

(d) the production quantity is $Q^* = F^{-1}(1-\frac{c\,m}{(m-n^*)v})$*,*

(e) the expected market token price equals the ICO token price, i.e., $\mathbb{E}[\tau_{eq}] = \tau^*$*,*

(f) the firm spends all funds raised through the ICO on production.

Several results are of interest here, starting with the condition $v > 2c$, which implies that ICOs may be best suited for products with relatively high willingness-to-pay.[8]

Part (ii) summarizes the characteristics of the unique equilibrium when $v > 2c$. Part (a) links the ICO cap to operational and demand parameters. Although we do not have a closed-form expression for n^*, our model suggests that it is never optimal for the firm to sell more than half of its tokens in the ICO (as shown by the upper bound $m/2$). By preserving sufficient tokens for itself, the firm can subsequently sell these tokens for profit in the market period. If the firm were to sell more than $m/2$ tokens in the ICO, it would raise more money than what could credibly be committed to production. The firm would then produce at the unconstrained optimal level and would be left with excess funds that it diverts to its pockets. However, these excess funds would be gained at the expense of a lower share of the future revenue (recall, the more tokens the firm sells to speculators in the ICO, the fewer tokens it leaves in the market period to sell to consumers). Overall, the firm would produce less and be less profitable. A detailed analysis is provided in Appendix A.2.3.

The remaining equilibrium quantities depend on the optimal ICO cap n^*. Note that, since the total number of tokens available is kept constant, the ICO cap n^* is a proxy for the

[8] It is interesting to note this condition does not depend on demand characteristics. This is because this condition stems from the presence of moral hazard: it simply defines the cutoff between an ICO that will never be able to raise any cash (even when demand risk is low) to one that will raise some cash in equilibrium (epsilon or more) depending on demand risk, among other things. From that cutoff point onward, how successful the ICO will be (e.g. how much money it will raise) depends critically on demand characteristics.

fraction of tokens sold during the ICO period. Part (b) shows that the ICO cap directly controls the number of speculators that will take part in the ICO. From parts (c) and (d), we can see that the more tokens the firm sells in the ICO, the lower the ICO token price and the firm's subsequent production quantity. The first effect is a direct consequence of speculators' participation constraint while the second effect again relates to the fact that the more tokens are sold to speculators in the ICO, the fewer tokens the firm has available to sell directly to customers in the secondary market. This understandably lowers its production incentives after the ICO. Part (e) is a consequence of the break-even condition for speculators.

Finally, part (f) shows that, in equilibrium, the firm puts all raised ICO funds towards production. Note that we model an unregulated environment whereby the firm has the option to divert the funds raised (moral hazard), but the high margin condition prevents such misconduct. The intuition behind this result is as follows. Recall that the firm's final wealth is equal to the sum of its token sales revenue both in the ICO and in the secondary market, less the production cost. When the product is profitable enough (captured by the price-cost ratio requirement) and the firm has a substantial share of its future revenue (guaranteed by the misconduct fraction), the firm is better off utilizing its resources on production to generate more cash later. As such, despite the absence of regulation and intermediaries, utility ICOs can overcome moral hazard through a combination of the aforementioned factors. We refer readers to §2.3.7 for continued discussion.

We provide more insights on the equilibrium through numerical experiments in §2.3.6.

2.3.5. ICO Agency Costs

Having analyzed the ICO equilibrium, we compare it to the first-best outcome so as to quantify agency costs. In this case, the first best refers to ICOs without frictions, i.e., ICOs with no cash diversion by the firm. While such "first-best" ICOs do not exist given the loose regulatory environment, by the Modigliani-Miller theorem, they are equivalent to a traditional newsvendor firm that invests its own money and faces no financial constraint.

Proposition 2.6. (ICO vs First Best)

i) A traditional newsvendor firm invests when $v > c$ whereas an ICO is only viable when $v > 2c$.

ii) A firm financing production through an ICO produces less than first best.

iii) A firm financing production through an ICO makes less profit than first best.

iv) In case of low demand realization, a traditional newsvendor risks loss whereas a firm financing production through an ICO always earns non-negative profit.

By Proposition 2.6, ICOs have the great advantage of being a low-risk means of financing for firms, but this comes at a cost of production quantity, profit and flexibility in terms of margin. We evaluate the extent of these benefits and inefficiencies numerically in §2.3.6. Our numerical results show that in general, the production and profit gaps between the ICO firm and first best can reach up to 40% and up to 50%, respectively, but these gaps shrink when the market is bigger, more stable or (and) with a higher willingness-to-pay. Under the same market conditions, ICOs lead to lower profit variance, rendering firm profits less sensitive to demand uncertainty.

Note that the inefficiencies mentioned above disappear in the absence of moral hazard. In other words, if the firm could credibly commit to spend all funds raised on production, the optimal ICO design leads to first-best final wealth, and the high margin condition is no longer needed for ICO success. To see this result, one could simply make the first constraint in (2.3) binding and conduct similar analysis as in this section.

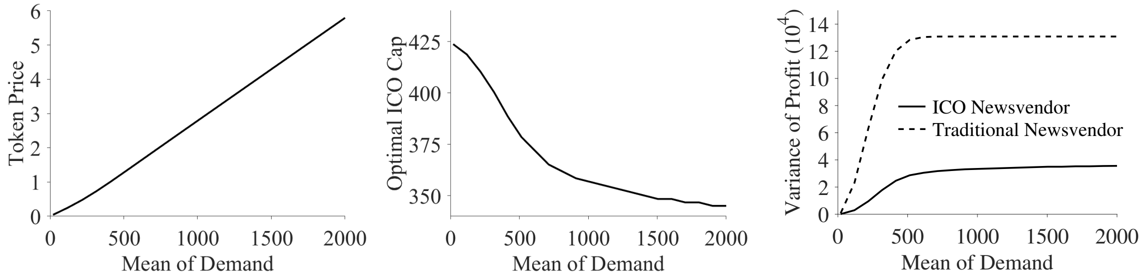
2.3.6. Numerical Experiments: ICOs with Utility Tokens

In this section, we provide a comparative-statics analysis through numerical experiments[9]. In particular, we focus on the impact of the mean and variance of demand and customers'

[9] In all of our numerical experiments throughout the chapter, demand follows a truncated normal distribution distributed with mean μ, standard deviation σ, lower bound 1, upper bound 2μ. By default, the parameters are assigned values $\mu = 500$, $\sigma = 166$, $m = 1000$, $c = 1$ and $v = 3$. The price-cost ratio in our numerical experiments was calibrated to be close to the Honeypod example discussed in the introduction. Our numerical results are qualitatively robust to alternative distributions such as uniform distributions. For expositional clarity, we focus only on results under normal distributions in the chapter.

willingness-to-pay.

Impact of mean demand



(a) Token Price vs Mean (b) ICO Cap vs Mean (c) Variance of Profit vs Mean

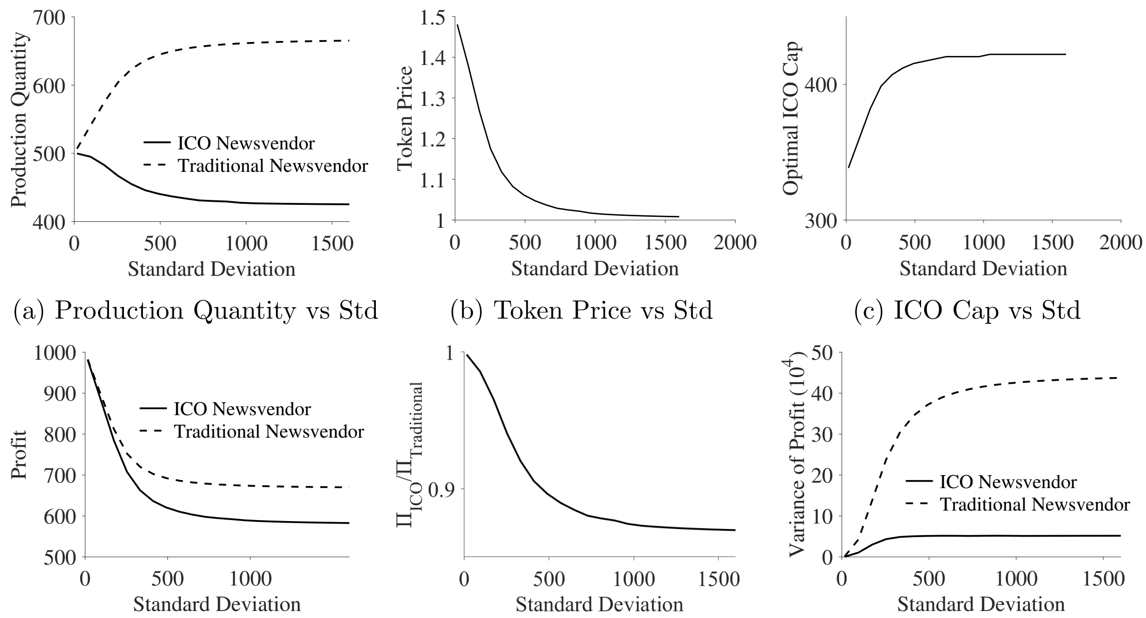
Figure 2: Impact of Mean of Demand

Figure 2 demonstrates that, as mean demand increases, the firm increases the ICO token price (see Figure 2 (a)) while also selling fewer tokens to speculators, that is, it maintains a larger share of the tokens (Figure 2 (b)). Overall, the pricing effect dominates, that is, increasing mean demand allows the firm to raise more cash in the ICO, which, in turn, allows it to increase production.

Impact of demand volatility

Figure 3 demonstrates that, as volatility increases, the ICO-funded firm reduces production whereas the traditional newsvendor may stock up (Figure 3 (a)). Such distinction could well come from the fact that higher uncertainty in the market adversely affects speculators' confidence in the token, driving the token price down (Figure 3 (b)). When facing high demand variability, it is also in the firm's best interest to sell more tokens in the ICO (Figure 3 (c)). However, the decrease in the token price has a dominating effect on the funds raised. We find that greater demand uncertainty hurts both the firm's ability to produce (Figure 3 (a)) and its profit (Figure 3 (d)). Moreover, the profit gap between ICO financing and first best widens (Figure 3 (e)) as demand variability increases, suggesting that ICOs are better suited for products with a more predictable or stable market.

Interestingly, Figures 2 (c) and 3 (f) show that the variance of profits is actually lower for an ICO-financed firm than for the traditional newsvendor. This is tied to the risk-



(a) Production Quantity vs Std (b) Token Price vs Std (c) ICO Cap vs Std

(d) Profit vs Std (e) Profit Ratio vs Std (f) Variance of Profit vs Std

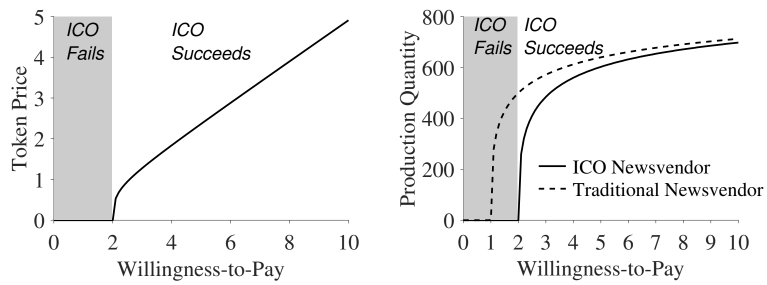
Figure 3: Impact of Standard Deviation (Std) of Demand

sharing property of ICOs mentioned in the introduction, whereby the risk of adverse demand outcomes is split between the firm and speculators.

Impact of customer WTP

Similar to a higher mean demand, a higher willingness-to-pay boosts the token value (Figure 4 (a)) and allows the firm to raise more funds (Figure 4 (b)) in the ICO while saving a larger fraction of tokens for the secondary offering (Figure 4 (c)). However, the rate of increase in funds raised due to a higher v decreases in v whereas that due to a higher mean demand is almost constant. The reason is that the equilibrium cap on ICO token sales n decreases drastically in v and the reason that the firm is motivated to save that many tokens for the secondary offering is to take advantage of the higher profit margin rather than the higher sales volume. Moreover, a higher willingness-to-pay incentivizes both the traditional newsvendor and the ICO newsvendor to produce more, while also closing the gap between them (Figure 4 (b)).

Lastly, both larger demand and higher willingness-to-pay lead to a higher expected profit



(a) Token Price vs WTP

(b) Production Quantity vs WTP

(c) ICO Cap vs WTP

(d) Profit Ratio vs WTP

Figure 4: Impact of Willingness-to-Pay

and lower profit gap between ICO financing and the first best (Figure 4 (d)).

2.3.7. Discussion: ICOs vs Other Early-Stage Financing Methods

Our results suggest that ICOs have important differences compared to other forms of early-stage financing. We summarize some of these differences in Table 3. We also discuss in detail two distinct features that our results have highlighted, distinguishing ICOs from other early-stage financing methods: the existence of a secondary market and the issuance of tokens.

Table 3 contains a large amount of information, and we recommend it be read through bilateral column comparisons. The high-level takeaway from the table is that ICOs, be it utility or equity offerings, differ from each of the other alternative forms of financing in at least one crucial dimension (and more often than not, in several dimensions). We highlight two of these aspects next.

Table 3: Comparison of Early-Stage Financing Methods. (A checkmark ✓ indicates the feature is prominent, while ✗ indicates it is of second-order or non-existent. The dual notation ✓✗ indicates that the feature may or may not be of first order, depending on circumstance.)

	Bank	**VC**	**Crowdfunding**		**Coin Offering**	
			Reward	Equity	Utility	Equity
Upside through Profit Sharing	✗	✓	✗	✓	✗	✓
Upside through Revenue Sharing	✗	✗	✗	✗	✓	✗
Downside Demand Risk Sharing	✓	✓	✗	✓	✓	✓
Heavily Regulated	✓	✓	✗	✓	✗	✓✗
Voting/Control Rights	✓✗	✓	✗	✗	✗	✓✗
Funds from Retail Speculators	✗	✗	✗	✓	✓	✓
Funds from Retail Consumers	✗	✗	✓	✓	✓	✓
Secondary Trading	✗	✗	✗	✗	✓	✓

Implications of the existence of a secondary market

ICOs differ from all other financing methods because of their reliance on a secondary market for the tokens. This difference has two important implications.

(i) **Mitigation of moral hazard** The alternative financing methods listed in Table 3 address moral hazard in different ways. Banks, for instance, use interest rates and covenants (Iancu et al., 2016) and/or leverage collateral. VC firms directly monitor the progress of the funded company and invest in stages to keep the company under control (Cherif and Elouaer (2008); Wang and Zhou (2004)). In crowdfunding, moral hazard is often left unadressed, though more recently, some platforms like Indiegogo have started to use escrow accounts to mitigate it (Belavina et al., 2019).

In the case of ICOs, there typically exists no third party between the fundraising firm and its investors. Instead, moral hazard is addressed, among other things, via the existence of a peer-to-peer secondary market for the tokens. To see this, recall that in our model, the fraction of tokens sold during the ICO, in equilibrium, is below the misconduct fraction. This ensures that a rational firm would be active in production after the ICO. The firm's production effort is reflected in the secondary market token price, and the firm always prefers a higher market token price. We show that, when the willingness-to-pay for the product is high relative to the unit production cost, the firm's dominant strategy is to spend all cash raised on production (i.e., zero cash diversion), which leads to the highest possible token price. Thus, in contrast to the alternative financing methods mentioned above, the existence of a secondary market is crucial in mitigating moral hazard in the context of ICOs.

(ii) **The nature of investors** The ICO secondary peer-to-peer market allows all owners to jointly sell the tokens to those who desire them. As discussed in Appendix A.2.1, this implies that the investors (speculators) do not have to be the consumers of the firm's products. In contrast, entrepreneurs running traditional crowdfunding campaigns (e.g., on Kickstarter), pre-sell their products directly to early adopting customers during the fundraising stage implying that the majority, if not all, backers in crowdfunding campaigns are the actual product consumers. Given the different nature of investors, it is reasonable to argue that ICOs have access to a larger investor pool than the crowdfunding projects. Indeed, an average ICO project in 2018 was able to raise $11.52 million (Cointelegraph, 2019), which is closer to the average VC deal value in the same year ($14.6 million) (PitchBook, 2019) and far exceeding the crowdfunding average ($10k) (Kickstarter, 2019).

Implications of the issuance of tokens

While both ICOs and crowdfunding raise funds through retail investors, the issuance of tokens further differentiates ICOs from crowdfunding. Our model shows that the utility tokens allow revenue sharing and the equity tokens allow profit sharing among all token holders. In addition, the tokens dilute the impact of future demand on the firm by allowing the firm to disperse the downside risks of low demand realization among the investors. On the contrary, the backers of a crowdfunding campaign do not share such risks because a low demand in the crowdfunding aftermarket would only hurt the firm's profit.

2.4. Analysis: Equity Tokens (STOs)

In this section, we consider a different type of ICO—one with equity, rather than utility, tokens (also referred to as STOs as mentioned earlier). Although most tokens offered so far have been utility tokens (through ICOs), STOs have become much more popular since 2018 (Blockstate, 2020). Unlike ICOs, STOs are typically regulated, and in our model, one could interpret strict regulations as a restriction on the firm's ability to divert any funds raised. In turn, this implies the funds will be put towards production, in which case, we can show that financing frictions are alleviated, and the outcome reduces to that of Proposition 2.6, where the regulated STO is simply equivalent to the traditional newsvendor.

A more meaningful setting is one in which STOs do not automatically imply the firm commits funds raised to production. For instance, this could be the case if there is no way to perfectly monitor the use of all funds, or, if monitoring is prohibitively costly to implement. A natural question that follows is whether then there is any difference between a utility ICO and an STO in this case? This section is devoted to answering this question.

2.4.1. Model & Equilibrium

The fund-raising mechanism with equity tokens follows that with utility tokens (Figure 1) but with two main differences. First, the fundamental value of the equity tokens and the utility tokens are backed by the firm's future revenue and profit respectively. To see this, recall from Lemma 2.1 that the value of the utility tokens is equal to the worth of all products sold. The equity tokens, by definition, entitle the token holders to a pre-specified share of the firm's profit as long as the firm is profit-making. Second, the equity tokens have no utility purposes—in the market period, the firm sells its products for cash and distributes its profit among the equity token holders in proportion to their token holdings. As a result, the firm, unlike a utility-token-issuing firm, does not need to sell the remaining tokens (i.e., tokens unsold in the ICO period) in the market period.

By definition, in the market period, the realized value of each equity token is

$$\tau_{eq,e} = \frac{1}{m} \cdot \left(v \, \min\{D, Q_e\} \, - \, c\, Q_e\right)^+ . \tag{2.5}$$

The firm maximizes its expected dollar-denominated wealth at the end of the market period, denoted by Π_e, which consists of three terms: i) the total funds raised during the ICO, $\tau_e\, s(\tau_e, n_e)$, plus ii) the expected total profit, $v\, \mathbb{E}\, [\min\{D, Q\}] \,-\, c\, Q$, minus iii) total payout to other token holders, $s(\tau_e, n_e)\, \mathbb{E}\, [\tau_{eq,e}]$. The objective function is as follows.

$$\max_{\tau_e,\, n_e} \left\{ \tau_e\, s(\tau_e, n_e) + \max_{Q_e} \left\{ (v\, \mathbb{E}\, [\min\{D, Q_e\}] \,-\, c\, Q_e) - \frac{s(\tau_e, n_e)}{m}\, \mathbb{E}\, [v\, \min\{D, Q_e\} \,-\, c\, Q_e]^+ \right\} \right\} \tag{2.6}$$

subject to

$\tau_e\, s(\tau_e, n_e) - c\, Q_e \geq 0, \quad$ (ICO funds cover production costs)

$u(s(\tau_e, n_e)) \geq 0. \qquad$ (speculators' participation constraint)

Again, we find the subgame perfect equilibrium using backward induction.

First, we consider the optimal production quantity Q_e^* given fixed token price τ_e and ICO cap n_e. Let $Q_u^*(s)$ denote the optimal production quantity unconstrained by the budget.

Proposition 2.7. (Optimal Production Quantity with Equity Tokens)
For a fixed token price τ_e, ICO cap n_e and number of speculators $s \in (0, m)$, the firm's optimal production quantity is $Q_e^*(s) = \min\left\{Q_u^*(s), \frac{\tau_e\, s}{c}\right\}$, where $Q_u^*(s) > 0$ is the unique solution of

$$(m - s)[(1 - F(Q_u^*(s))\, v - c] = s\, c\, F\left(\frac{c}{v}\, Q_u^*(s)\right). \tag{2.7}$$

We show in the proof of Proposition 2.7 that $Q_u^*(s)$ decreases monotonically in s and the firm, ignoring the budget constraint, would produce at the first best production when $s = 0$ as $Q_u^*(0) = F^{-1}\left(\frac{v-c}{v}\right)$. Therefore, for any positive number of speculators, the firm produces less than the first-best quantity. The other boundary case is $Q_u^*(m) = 0$. Since $Q_e^*(m) = \min\left\{Q_u^*(m), \frac{\tau_e\, m}{c}\right\} = \min\left\{0, \frac{\tau_e\, m}{c}\right\} = 0$, the firm produces nothing when $s = m$. This shows that as long as the firm does not sell out all the tokens during the ICO, i.e., $s \neq m$, it always produces some product if it raises money.

Recall that the misconduct fraction with utility tokens is $1 - c/v$. In the case of equity tokens, the misconduct fraction is 1. Since $1 > 1 - c/v$, we argue that with equity tokens, the firm's incentives are better aligned with the speculators', making the firm less likely to divert cash from funds raised to its own pocket. In other words, STOs reduce moral hazard, thus having lower agency costs than utility ICOs.

At this point, we make a regularity assumption on the demand distribution[10]: $f(x) < a^2 \cdot f(ax)$ for $a > 2$. Using the result of Proposition 2.7, we show next that successful ICOs with equity tokens require a larger fraction of the tokens to be sold during the ICO than those with utility tokens.

Proposition 2.8. (Conditions for ICO Success with Equity Tokens)
An ICO that issues equity tokens succeeds if and only if

i) (critical mass condition) the firm sells more than $\frac{c}{v-c}m$ tokens in the ICO and,

ii) (price-cost ratio requirement) customers have a high willingness-to-pay such that $v > 2c$.

Recall from Proposition 2.4 part (i) that with utility tokens, the minimum number of tokens needed for production is $\frac{c}{v}m$. Since $\frac{c}{v-c}m > \frac{c}{v}m$, part (i) suggests a more stringent critical mass condition for equity tokens, which translates into a higher ICO cap as shown in §2.4.2. To understand this result intuitively, let's consider a firm that aims to produce a fixed quantity. By the nature of the two tokens (revenue-sharing vs profit-sharing), we know that the market equilibrium price of the utility token will be higher than that of the equity token, i.e., $\tau_{eq} > \tau_{eq,e}$. We also know by Proposition A.2(i) (Appendix A.2.2) that for utility tokens, the firm sets the ICO token price to be exactly equal to the expected market equilibrium token price, i.e., $\tau = \mathbb{E}[\tau_{eq}]$. For equity tokens, since the speculators would only purchase the tokens when the ICO token price does not exceed the expected market equilibrium token value, we must have $\tau_e \leq \mathbb{E}[\tau_{eq,e}]$. Therefore, the optimal ICO price of

[10]This assumption is satisfied by distributions that do not contain sharp peaks such as uniform distributions and most normal distributions.

the equity token, τ_e, must be less than that of the utility token, τ. As a result, to meet the same production goal, the firm will have to sell more equity tokens than utility tokens. That being said, we observe numerically that the firm produces more with equity tokens but the correlation between ICO cap and production quantity is not necessarily positive. However, the difference in the nature of these two tokens dominates other effects and results in a higher ICO cap with equity tokens.

Following part (i) and Proposition 2.7 that the firm should not sell all of its equity tokens, we need $\frac{c}{v-c}\,m < m$ for the existence of feasible n, which leads to part (ii). Comparing with Proposition 2.4 part (ii), we see that the price-cost ratio requirement is the same for both types of tokens. Therefore, while intuitively the equity tokens put an emphasis on "profit" by definition, they do not require a higher or lower profit margin of the product than the revenue-sharing utility tokens.

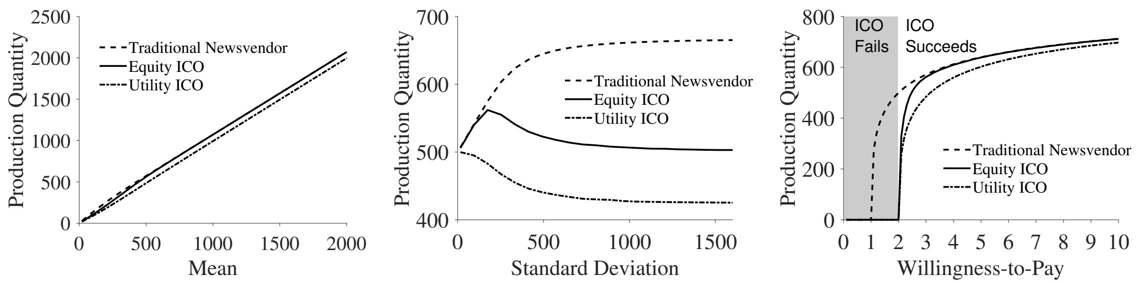
Lastly, we show that when the two conditions given by Proposition 2.8 are met, the firm sets the ICO token price such that the speculators' expected profit is zero—a similar result to Proposition 2.5(ii)(e).

Proposition 2.9. (Optimal ICO Equity Token Price)
When $v > 2c$, for a given $n_e \in (\frac{c}{v-c}\,m, m)$, there exists a finite positive $\tau_e^*(n_e)$ uniquely determined by $u(s^*(\tau_e^*(n_e))) = 0$.

In summary, our analytical results identify two key differences and two similarities between ICOs and STOs in the absence of regulations: i) STOs are associated with lower agency costs; ii) STOs require a larger ICO cap to be successful; iii) both require the same high-margin condition; iv) neither leaves any arbitrage opportunities for speculators. We study the rest of the equilibrium results numerically in §2.4.2.

2.4.2. Numerical Experiments: Comparing ICOs to STOs

In Sections 2.3 and 2.4.1, we show analytically that ICOs with either type of tokens lead to underproduction. Through numerical experiments, we find that issuance of equity to-



(a) Production Quantity vs Mean (b) Production Quantity vs Std (c) Production Quantity vs WTP

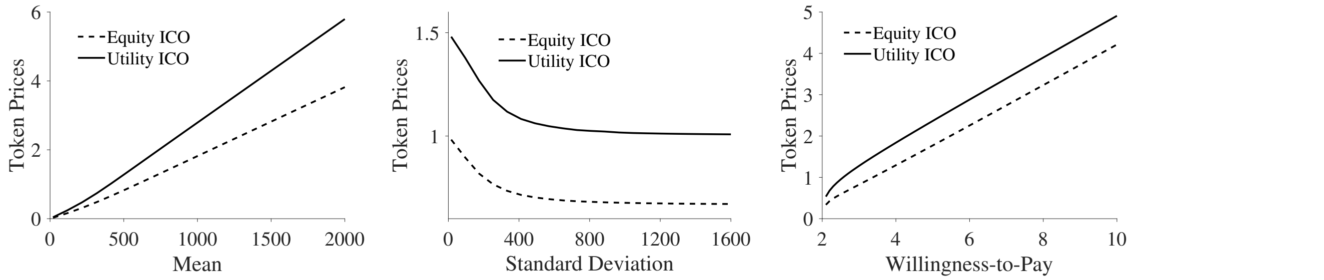
Figure 5: Comparison of Production Quantities

kens incentivizes the firm to produce more (Figure 5), ceteris paribus. While good market conditions (high mean, low variance, high willingness-to-pay) reduce the extent of underproduction in both cases, they push the production level of the firm that issues equity tokens even closer to that of a traditional newsvendor. This suggests that the first-best is almost achievable with equity tokens.

Another immediate implication of a higher production level with the issuance of equity tokens, by Proposition 2.5 (iii), is that the funds raised through the equity token ICO must surpass that through the utility token ICO, because $s(n_e^*, \tau_e^*) \cdot \tau_e^* \geq c \cdot Q_e^* > c \cdot Q^* = s(n^*, \tau^*) \cdot \tau^*$.

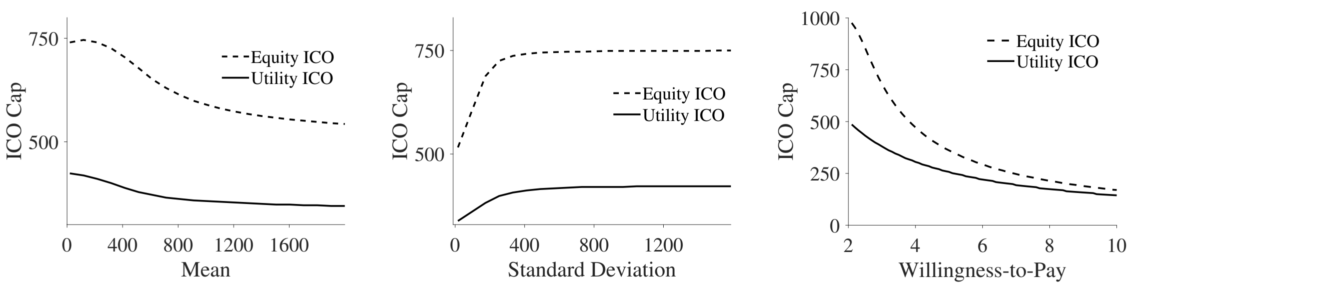
Figure 6 shows that the revenue-sharing utility tokens have a higher market value than the equity tokens as the prices of the equity tokens (both the ICO token price and the expected market equilibrium token price) are consistently lower. Figure 7 shows that more equity tokens will be sold than utility tokens, although the gap diminishes under better market conditions. Since the total ICO proceeds (the product of price and ICO cap) are higher with equity tokens, the effect of a larger ICO cap outweighs that of lower prices. Moreover, it can be readily checked that the firm spends all ICO proceeds on production rather than leaving any funds idle, i.e., $s(n_e^*, \tau_e^*) \cdot \tau_e^* = n_e^* \cdot \tau_e^* = c \cdot Q_e^*$. Note that this result with equity tokens echoes the outcome with utility tokens (Proposition 2.5 (iii)).

Finally, with a closer-to-optimal production quantity, the firm obtains a higher total wealth



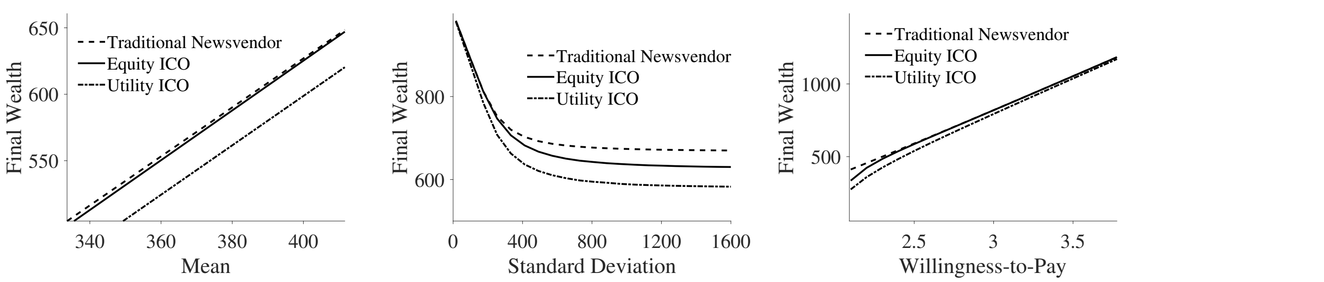
(a) Token Prices vs Mean (b) Token Prices vs Std (c) Token Prices vs WTP

Figure 6: Comparison of Token Prices



(a) ICO Cap vs Mean (b) ICO Cap vs Std (c) ICO Cap vs WTP

Figure 7: Comparison of ICO Caps



(a) Final Wealth vs Mean (b) Final Wealth vs Std (c) Final Wealth vs WTP

Figure 8: Comparison of Final Wealth

with equity tokens than with utility tokens (Figure 8). In particular, when market conditions are better, equity tokens allow the firm to achieve near-the-first-best outcome.

2.5. Extensions

While our core model (ICO with utility tokens) is relatively basic, it is flexible enough to be extended in many ways to fit a variety of practical situations.

2.5.1. Technology Risk

Motivated by the Sirin Labs example discussed in the introduction, our base model assumes that the firm is able to successfully produce its product when it incurs the necessary production cost. However, recognizing that startups are inherently risky, here, we relax this assumption, and we add to the base ICO model the risk of production failure.

Let $\alpha \in (0,1]$ denote the probability that the firm's technology leads to successful production and suppose that the value of α is common knowledge. The firm either successfully produces the decided quantity or ends up with zero acceptable products. We also assume that the firm finds out whether production has been successful at the end of the production period, after it has paid the necessary production cost for the decided quantity. In other words, the production cost is sunk regardless of the outcome of production.

Given such risks, the equilibrium token price is given by $\tau_{eq} = (1-\alpha)\cdot \frac{v}{m}\min\{0,D\} + \alpha \cdot \frac{v}{m}\min\{Q,D\}$, and thus $\mathbb{E}[\tau_{eq}] = \alpha \cdot \frac{v}{m}\mathbb{E}[\min\{Q,D\}]$. The firm optimizes a modified objective function

$$\max_{\tau,n}\left\{\tau\, s(\tau,n) + \max_{Q}\ \left[\alpha\,(m - s(\tau,n))\,\frac{v}{m}\,\mathbb{E}[\min\{Q,D\}] - c\,Q\right]\right\} \qquad (2.8)$$

subject to

$$\tau\, s(\tau,n) - c\,Q \geq 0, \quad \text{(ICO funds cover production costs)}$$

$$u(s(\tau,n)) \geq 0. \quad \text{(speculators' participation constraint)}$$

We show next that riskier production intensifies the moral hazard problem.

Proposition 2.10. (Optimal Production Quantity under Risks of Production Failure) *Suppose the firm's production is successful with probability $\alpha \in (0,1]$. For a fixed token price τ, ICO cap n and number of speculators s, the firm's optimal production quantity $Q^*(s)$ is as follows.*

i) If $0 < s < m(1-\frac{c}{\alpha v})$, then $Q^*(s) = \min\left\{F^{-1}\left(1 - \frac{cm}{\alpha(m-s)v}\right), \frac{\tau\, s}{c}\right\}$.

ii) If $s = 0$ *or* $s \geq m(1 - \frac{c}{\alpha v})$*, then* $Q^*(s) = 0$*.*

Proposition 2.10 shows that, given the same ICO token price and ICO cap, a lower success probability leads to lower production quantity. The firm is also more likely to give up production and divert funds when α is smaller because the misconduct fraction, $1 - \frac{c}{\alpha v}$, is lower. As a result, we show in Proposition 2.11 that ICOs are less likely to succeed under higher production risks.

Proposition 2.11. (Conditions for ICO Success under Risks of Production Failure) *Suppose the firm's production is successful with probability* $\alpha \in (0, 1]$*. Then, the ICO succeeds if and only if*

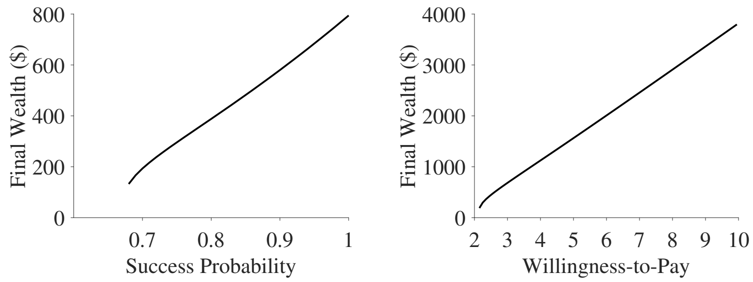
i) (critical mass condition) the firm sells more than $\frac{m\,c}{\alpha\,v}$ *tokens in the ICO and,*

ii) (price-cost ratio requirement) customers have a high willingness-to-pay such that $v > \frac{2c}{\alpha}$*.*

To obtain additional insights, we proceed to numerically examine the properties of the equilibrium under this technology risk extension.

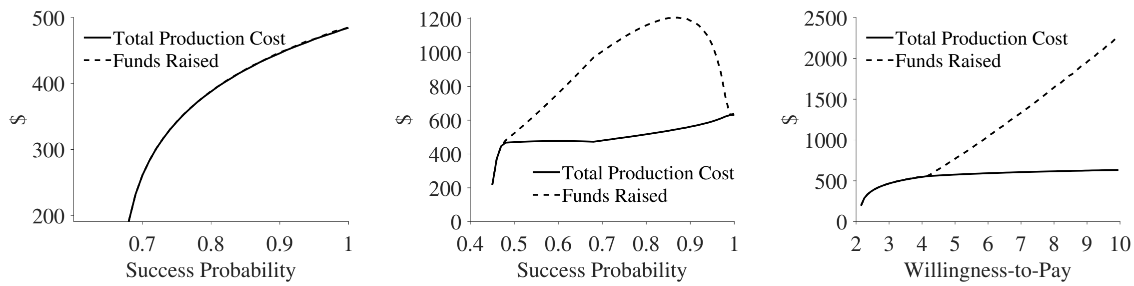
Numerical Experiments: The Impact of Technology Risk

Figure 9 shows that the firm's final wealth increases in the success probability (Figure 9 (a)) and customers' willingness-to-pay (Figure 9 (b)). More interestingly, Figure 10 shows that the firm's optimal strategy varies for different values of α and v. Recall from Proposition 2.5 (iii) that, when there is no risk ($\alpha = 1$), the firm invests all money raised in production. For $\alpha < 1$, the firm does the same if either the risks are high or the willingness-to-pay is low (Figure 10 (a,b)). However, under more favorable conditions, i.e., low risk ($\alpha < 1$ but close to 1) and high willingness-to-pay, the firm spends part of its funds raised on production and saves the rest (Figure 10 (c)). Such practice guarantees that the firm ends up with positive final wealth even if production fails.



(a) Final Wealth vs α ($v = 3$) (b) Final Wealth vs v ($\alpha = 0.95$)

Figure 9: Firm's Final Wealth under Risks of Production Failure



(a) Funds Usage vs α ($v = 3$) (b) Funds Usage vs α ($v = 6$) (c) Funds Usage vs v ($\alpha = 0.95$)

Figure 10: Funds Usage under Risks of Production Failure

2.5.2. Outside Investment Options

We can account for the existence of other investment options (e.g., a savings account) by adding a generic investment option that returns $k > 0$ dollars per dollar investment.

Suppose there exists a generic outside investment option that returns $k > 0$ dollars per dollar of investment. The outside option provides a new reference point when the speculators evaluate their ICO return. Let $\Delta_i(s)$ denote the expected profit improvement by investing in an ICO with utility tokens. Then,

$$\Delta_i(s) = \mathbb{E}[\tau_{eq}(s)] - \tau - \tau\, k = \mathbb{E}[\tau_{eq}(s)] - (k+1)\,\tau, \tag{2.9}$$

and the speculators expected profit improvement is $u(s) = \frac{s}{z}\Delta_i(s)$. The firm optimizes the same objective function as in (2.3). Therefore, the misconduct fraction is unaffected by the presence of the outside option, and the optimal production quantity in the subgame still follows that in Proposition 2.2. However, we show below that a higher return of the outside option makes ICOs harder to succeed as it leads to more stringent success conditions.

Proposition 2.12. (Conditions for ICO Success with an Outside Investment Option) *In the presence of an outside investment option with return k per dollar invested, the ICO succeeds if and only if*

i) (critical mass condition) the firm sells more than $(1+k)\frac{mc}{v}$ tokens in the ICO and,

ii) (price-cost ratio requirement) customers have a high willingness-to-pay such that $v > (2+k)c$.

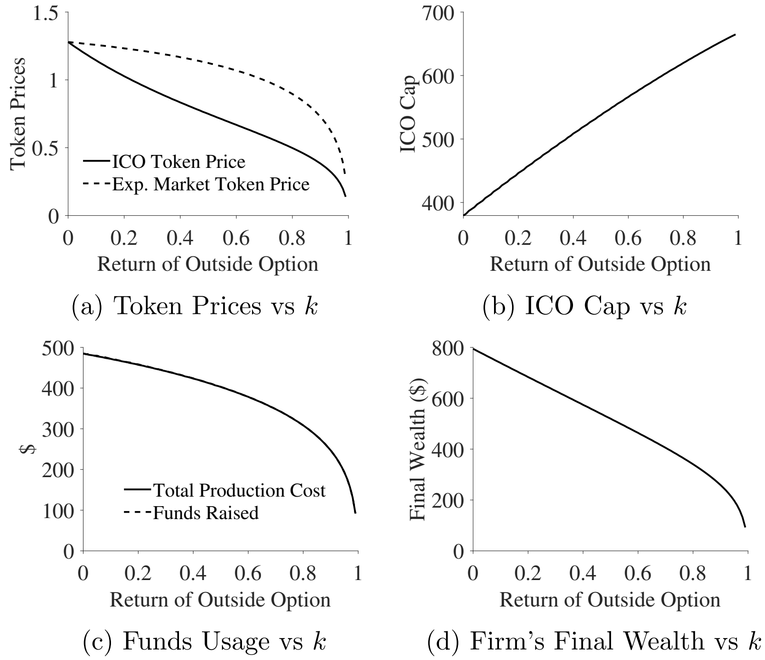
Next, we show that the optimal ICO token price leads to zero expected profit improvement, i.e., the expected return of the tokens is equal to that of the outside investment option. In this case, since the outside option guarantees positive return, the expected market token price is higher than the ICO token price.

Proposition 2.13. (Optimal ICO Token Price with an Outside Investment Option) *When $v > (2+k)\,c$, for a given $n \in \left((1+k)\frac{mc}{v}, m\left(1-\frac{c}{v}\right)\right)$, there exists a finite positive $\tau^*(n)$ uniquely determined by $u(s^*(\tau^*(n))) = 0$.*

To obtain additional insights, we proceed to numerically examine the properties of the equilibrium under this technology risk extension.

Numerical Experiments: The Impact of Outside Investment Options

Intuitively, a better-paying outside option makes ICOs less attractive in comparison. To incentivize the speculators to participate, the firm needs to make token trading more lucrative by either raising the expected market token price or by reducing the ICO token price. Since the former is difficult to achieve given that the demand distribution remains unchanged, the firm must do the latter. Our numerical results show that, as k increases, the token prices drop (Figure 11 (a)) and the firm sells more tokens during the ICO (Figure 11 (b)) to mitigate the loss in funds raised. A higher k also discourages production (Figure 11 (c)) and hurts the firm's final wealth (Figure 11 (d)).



(a) Token Prices vs k (b) ICO Cap vs k

(c) Funds Usage vs k (d) Firm's Final Wealth vs k

Figure 11: Impact of Outside Investment Return on ICOs

In §2.5.2, we show that the optimal ICO token price makes the expected return of the tokens equal to that of the outside option. This result can be readily checked in Figure 11 (a), where for any value of k, the difference of the token prices divided by the ICO token price is exactly k.

2.6. Conclusion

As one of the first academic papers to study the implications of ICOs for operations management, the model we develop has of course some limitations that could represent interesting research opportunities.

For example, it could be interesting to study the multi-period production setting, which could also involve issues of token resale and inflation control. In practice, many projects keep a portion of ICO funds and/or tokens to maintain price stability in the future and protect against negative shocks. Moreover, some entrepreneurs need initial funds for the design and preparation of an ICO, which requires a different ICO design or even other financing solutions. Another interesting direction would be to expand the firm's decision space in terms of where it can spend the raised funds, to include other business functions such as marketing, human capital, etc.

Moreover, several assumptions in our model could be relaxed to capture more realistic settings. For instance, the tokens could be used for purposes other than to purchase physical goods; customer willingness-to-pay and demand could be affected by quite a few factors that we do not capture, including network effects; the success of the ICO could be informative about future demand in a multi-period setting; investors could have heterogeneous beliefs about product quality; customers could have different valuations for the same product; firms, investors and/or customers could be risk averse or risk seeking, etc.

Finally, like some of the extant literature considering strategic customer behavior (Cachon and Swinney, 2011; Belavina et al., 2019), our model assumes strategic customers with known and homogeneous willingness-to-pay. In reality, customer wtp could be i) uncertain and ii) heterogeneous. Our existing model can readily incorporate a relaxation of i) by replacing v with $E[v]$.[11] However, relaxing assumption ii) would generate at least two complications in our setting that would go beyond the scope of the chapter: the first would be the need to specify a doubly stochastic model of customer demand, that is, the firm's demand beliefs would need to be specified for every possible customer type; the second would be the need to develop a more elaborate model of secondary market clearing for crypto-currency exchanges. We believe these, as well as the other aforementioned extensions above, to be interesting directions for future research.

[11] This is the case because the token exchange rate is set via the market clearing condition, which occurs after demand is realized and uncertainty has been resolved.

CHAPTER 3 : Uncapped Initial Coin Offerings

3.1. Introduction

In the previous chapter, we have studied capped ICOs, in which the maximum amount of capital raised is fixed. In this chapter, we turn our attention to uncapped ICOs, where there is no cap on the total ICO proceeds. In fact, some of the most successful ICOs have been uncapped. For example, the EOS project, an operating system for Blockchain applications, raised $4.2 billion in one single round of uncapped ICO. The Ethererum project, one of the most widely-used Blockchain platforms, also started out with an uncapped ICO.

There are two types of uncapped ICOs in terms of structure. The first type sets a fixed token price but no upper bound on token supply. Examples include Ethereum, where new tokens are minted as more transactions are completed in addition to an unlimited token sale, and Sirin Labs, which set its token supply based on the ICO outcome. The second type does the opposite of what the first type does, setting a fixed token supply but varying prices. For instance, the Kleros project let the investors decide the token prices via an auction. The focus of this chapter is on the first type of ICO.

Although uncapped ICOs take different forms, the projects behind them are often service platforms. These include platforms for transactions and applications (e.g. Ethereum, EOS), social media (e.g. Telegram, Steem), data storage (e.g. Filecoin), dispute resolution (e.g. Kleros), etc. All of these platforms bring together service providers and customers, and are designed to operate for a long time during which both the service capacity and the customer base (hopefully) grow together. Just like it is in any two-sided market, the value of service from these platforms rises as the number of users increases, which is what we call *the network effect*.

To understand how ICOs can fund service platforms under network effects, we study the following model in this chapter. There is a firm that builds and operates a service platform that brings together two types of users, service providers and customers. Before the platform is built, the firm raises funds through an uncapped ICO, where utility tokens are sold to a group of speculators who hope to gain a profit on the tokens by selling them after the platform launches. Once the platform launches, it allows service providers and customers to trade through tokens in multiple periods until either party is no longer willing to do so.

We find that both the platform's service capacity and service providers' profitability are enhanced by stronger network effects, larger customer base, and/or lower unit service cost. Moreover, we show that uncapped ICO is successful if and only if the cost of building the platform does not exceed the total service cost per period. We also extend the base model to account for firm's moral hazard and show that under loose regulation, uncapped ICO can still be successful if the firm charges the right service fee.

3.2. Model

We model four groups of agents. On the investors' side, we consider a firm that builds a service platform and a group of speculators who contribute cash to the development of the platform. On the users' side, there are customers and service providers. We consider an infinite time horizon until any group of these agents quits the game. We now describe each group of agents and the sequence of events below.

Firm The firm designs a service platform that requires an exogenous setup cost of $\mathcal{C}$ dollars. With no initial cash, the firm raises funds through an uncapped ICO, in which it sells an unlimited number of tokens at price τ_0 (dollars per token) to speculators. These tokens are utility tokens that are used for (and only for) consumption of service from the platform. The tokens are also the only viable method of payment for the service provided through the platform. The platform is designed to handle at most $\overline{s}$ units of service in each period.

Suppose the firm sells m tokens during the ICO. If the ICO proceeds exceed the necessary cost, i.e., $\tau_0 \cdot m \geq C$, the firm successfully develops the platform. Otherwise, the firm returns ICO proceeds back to speculators and the game ends. After the platform launches, it automatically adds z tokens to the existing pool of tokens per period and distributes them to service providers of that period. Note that $z > 0$ means that new tokens are generated, while $z < 0$ means that tokens are burned. The equilibrium token price in period t, τ_t, determined in the secondary market, depends on the service level (# units of service performed) of the same period, s_t.

Here, $\overline{s}$, τ_0 and z are determined by the firm before the ICO, but τ_t and s_t are equilibrium quantities based on decisions of other agents described below.

Speculators A group of risk-neutral speculators arrive sequentially during the ICO and collectively purchase m tokens at price τ_0. If the platform launches, they sell their tokens to 1^{st}-period customers at the equilibrium token price τ_1. Speculators only purchase tokens during the ICO if they believe they can make a non-negative profit, i.e., $\tau_1 - \tau_0 \geq 0$, and stop purchasing when it is no longer profitable to buy more tokens.

Customers There is a finite market size, M, in any single period t. Customers have idiosyncratic valuation of service, V (dollars), which follows a uniform distribution with support $[0, 1]$. Depending on the service level s_t, each customer pays a service provider an equilibrium amount of $p(s_t)$ in the form of tokens. More specifically, a customer buys $p(s_t)/\tau_t$ tokens at the beginning of period t and sends these tokens to a service provider upon completion of service at the end of the same period, where $\tau_t = \tau(t, s_t)$ is the token price in period t. Every customer enjoys a network benefit $B(s_t) = k \cdot s_t^2$ (dollars), where $k > 0$ is the network effect coefficient. One pays for the service if his or her surplus is non-negative.

Service Providers There exists a large pool of potential service providers ($\geq M$). In any single period t, each service provider can either provide one unit of service at a fixed cost[1] $c < 1$ or be idle. The platform allows a maximum service capacity of $\overline{s}$. Depending on the number of working service providers, s_t, each working service provider is paid an equilibrium

[1] Cost of service is lower than the upper bound of customers' valuation.

amount of $p(s_t)/\tau(t, s_t)$ tokens from customers, and receives (or loses) z/s_t tokens that are minted (or burned) by the platform. Service providers of period t sell all remaining tokens at the beginning of period $(t+1)$ at the new token price $\tau(t+1, s_{t+1})$. In each period, a service provider chooses to work only if his or her return is non-negative, and everyone makes the decision simultaneously. If more service providers are willing to work than the platform's capacity, working service providers are randomly chosen.

Sequence of Events

- Period 0
  (1) Firm chooses ICO token price τ_0, service capacity $\overline{s}$, and the policy on token supply control z. Firm announces $\mathcal{C}$.
  (2) Speculators buy m tokens.
  (3) Firm raises $m \cdot \tau_0$ (suppose $m \cdot \tau_0 \geq \mathcal{C}$) and builds the platform.
- Period 1
  (1) Period 1 starts with $m_1 = m$ tokens in the ecosystem.
  (2) Speculators sell m tokens to customers at price τ_1.
  (3) Service providers deliver s_1 units of service to s_1 customers.
  (4) Each customer pays $p(s_1)/\tau(1, s_1)$ tokens to a service provider.
  (5) Each working service provider receives/loses z/s_1 tokens.
  (...)
- Period $t > 1$
  (1) Period t starts with $m_t = m + (t-1)z$ tokens in the ecosystem.
  (2) Service providers in period $(t-1)$ sell all of their tokens to customers in period t at price τ_t.
  (3) Service providers deliver s_t units of service to s_t customers.
  (4) Each customer pays $p(s_t)/\tau(t, s_t)$ tokens to a service provider.
  (5) Each working service provider receives/loses z/s_t tokens.
  (...)

We assume that events (2) - (4) in any period $t \geq 1$ take place simultaneously.

3.3. Analysis

We now study the subgame perfect equilibrium of the game between all four groups of agents using standard backward induction. Let's begin with the equilibrium price of the service $p(s_t)$ that clears the market. We assume that customers are price takers, and $p(s_t)$ allows s_t customers with the highest personal valuation of service to be served. That is to say,

$$M \cdot \mathbb{P}(\text{customer's surplus} = V + B(s_t) - p(s_t) \geq 0) = s_t, \tag{3.1}$$

where V is the customer's random valuation of service and $B(s_t) = k\, s_t^2$ captures network benefit. With uniformly distributed V (support [0,1]), (3.1) gives us

$$p(s_t) = ks_t^2 - \frac{s_t}{M} + 1 \tag{3.2}$$

given service level s_t.

On the service providers side, each service provider's profit in period t is

$$\phi(s_t) = \left(\frac{p(s_t)}{\tau(t, s_t)} + \frac{z}{s_t}\right) \cdot \tau(t+1, s_{t+1}) - c. \tag{3.3}$$

Lemma 3.1. (Equilibrium Prices)

(i) The equilibrium token-denominated price of service is m_t/s_t.

(ii) Given capacity s_t, the equilibrium token price is $\tau(t, s_t) = p(s_t) \cdot s_t/m_t$.

By Lemma 3.1, we can rewrite (3.3) as

$$\phi(s_t) = p(s_{t+1})\frac{s_{t+1}}{s_t} - c. \tag{3.4}$$

We see that given any service level $s_{t+1} > 0$, service provider's profit $\phi(s_t)$ decreases in the current service level s_t. Ignoring the platform's service capacity, service providers collectively choose a service level that makes everyone's profit non-negative. Assuming that a service provider prefers breaking even to not working at all, we have the following.

Lemma 3.2. (Equilibrium Service Level)
Given service capacity s_{t+1} in period $(t+1)$, the equilibrium service level in period t, without considering the service capacity, is $s_t^*(s_{t+1}) = \min\{\frac{p(s_{t+1})\, s_{t+1}}{c}, M, \bar{s}\}$.

This shows that as long as the service level in the next period is positive, there exists an optimal service level in the current period that is also positive. However, since the optimal service level always depends on future service levels, yet the platform can't directly set a service level by putting a cap, only a steady-state service level would allow the platform to run indefinitely. In other words, the firm needs to make sure that service providers are able to maintain a consistent service level through all periods. Otherwise, speculators expect coordination of service to fail in the very first period, which prevents the success of ICO.

Next, we derive steady-state service levels given network effects with different intensity, ignoring the platform's service capacity $\bar{s}$.

Proposition 3.3. (Steady-State Service Level)
Let s_{ss}^* denote the steady-state service level.

(i) (Weak network effects) When $k < \frac{c}{M^2}$, $s_{ss}^* = \frac{1-\sqrt{1-4M^2k(1-c)}}{2Mk} < M$ and $\phi(s_{ss}^*) = 0$.
(ii) (Moderate network effects) When $k = \frac{c}{M^2}$, $s_{ss}^* = M$ and $\phi(s_{ss}^*) = 0$.
(iii) (Strong network effects) When $k > \frac{c}{M^2}$, $s_{ss}^* = M$ and $\phi(s_{ss}^*) > 0$.

Proposition 3.3 shows that in steady state, the platform can serve all customers in the market under moderate to strong network effects, but not under weak network effects. Moreover, service providers just break even under weak to moderate network effects, and only obtain positive profit under strong network effects. Note that on the customers side, although almost all of them enjoy positive surplus from the service[2], they all reap the

[2]with continuous valuation, customers with zero profit are of measure 0.

benefits of moderate to strong network effects due to higher service levels. To see this, note that by (3.1) and (3.2), the surplus of a customer with valuation v is $[v + B(s_t) - p(s_t)]^+ = \left[v - F\left(\frac{M-s_t}{M}\right)\right]^+$ given service level s_t, which is increasing in s_t.

Note that the steady-state service levels in Proposition 3.3 are the highest among all possible steady-state service levels in the presence of fixed service capacities determined by the platform. For example, under weak network effects, a service capacity $\overline{s} < \frac{1-\sqrt{1-4M^2k(1-c)}}{2Mk}$ could also lead to a steady state in which $s_{ss}^* = \overline{s}$ and $\phi(s_{ss}^*) > 0$. We discuss how the firm chooses a service capacity that maximizes its profit later.

Finally, we move ahead to consider the investors. By Proposition 3.3 and Corollary 3.5, we know that the equilibrium token price in period t is $\tau(t, s_{ss}^*) = \frac{c\, s_{ss}^*}{m+(t-1)z}$. For speculators to profit, the initial token price τ_0 must be less than or equal to $\tau(1, s_{ss}^*) = \frac{c\, s_{ss}^*}{m}$. Therefore, for any given initial token price τ_0, the speculators do not purchase more than $\frac{c\, s_{ss}^*}{\tau_0}$ tokens ($m \leq \frac{c\, s_{ss}^*}{\tau_0}$).

Recall that successful development of the platform requires cost $\mathcal{C}$, which becomes the lower bound of speculators' contribution, $m \cdot \tau_0$, if they would like the ICO to succeed. Suppose speculators prefer non-negative profit to an unsuccessful investment. Then, we have $m \cdot \tau_0 \geq \mathcal{C}$, or $m \geq \frac{\mathcal{C}}{\tau_0}$.

Therefore, a necessary condition for a successful ICO is $\mathcal{C} \leq c \cdot s_{ss}^*$, and this leads to the following result.

Proposition 3.4. (Condition for Successful Uncapped ICO)
The uncapped ICO succeeds if and only if the setup cost of the platform does not exceed the total service cost per period, i.e., $\mathcal{C} \leq c \cdot s_{ss}^*$.

Suppose the condition in Proposition 3.4 holds. In equilibrium, speculators buy $m^*(\tau_0) = \frac{c\, s_{ss}^*}{\tau_0}$ tokens, and the firm's profit is $\Pi = m^*(\tau_0) \cdot \tau_0 - \mathcal{C} = c \cdot s_{ss}^* - \mathcal{C}$. Here, we show that the firm's profit is independent of the initial token price and its policy on token supply control. However, the firm's profit does increase in the steady-state service level. We can then deduce the optimal service capacities of the platform that induce steady-state service levels according to Proposition 3.3.

Corollary 3.5. (Optimal Service Capacity)

(i) (Weak network effects) When $k < \frac{c}{M^2}$, $\overline{s} = \frac{1-\sqrt{1-4M^2k(1-c)}}{2Mk}$*;*

(ii) (Moderate to strong network effects) When $k \geq \frac{c}{M^2}$, $\overline{s} = M$*.*

3.4. Extensions and Future Work

In our base model, the firm operates the platform free of charge. In practice, service platforms charge a small fee on most transactions through the platform. However, the base model can be easily extended to incorporate this feature by transferring part of the service fee to the firm. This does not affect customers' surplus function in (3.1) but changes service providers' profit function by applying a fraction $\alpha \in (0, 1)$ in (3.3) to

$$\phi(s_t) = \alpha \cdot \left(\frac{p(s_t)}{\tau(t, s_t)} + \frac{z}{s_t} \right) \cdot \tau(t+1, s_{t+1}) - c. \tag{3.5}$$

As a result, the firm obtains $(1-\alpha) \cdot \left(\frac{p(s_t)}{\tau(t,s_t)} + \frac{z}{s_t} \right)$ tokens in period t and sells them in period $(t+1)$ at price $\tau(t+1, s_{t+1})$. Let r denote the discount rate for future cash flows. The firm's profit under a steady-state service level evaluated in period 0 can thus be written as follows.

$$\begin{aligned}
\Pi &= c \cdot s^*_{ss} - \mathcal{C} + \sum_{t=1}^{\infty} (1-\alpha) \left(\frac{p(s^*_{ss})}{\tau(t, s^*_{ss})} + \frac{z}{s^*_{ss}} \right) \cdot \tau(t+1, s^*_{ss}) \cdot r^t \\
&= c \cdot s^*_{ss} - \mathcal{C} + \sum_{t=1}^{\infty} (1-\alpha) \cdot p(s^*_{ss}) \cdot r^t \\
&= c \cdot s^*_{ss} - \mathcal{C} + (1-\alpha) \cdot p(s^*_{ss}) \cdot \frac{r}{1-r}.
\end{aligned} \tag{3.6}$$

Previously, we have also assumed a responsible firm that commits to the development of the platform as long as enough funds are raised. In reality, due to lack of regulation, the firm could abandon the project and run away with the ICO proceeds. This extended model allows us to account for such moral hazard. In particular, the firm chooses whether or not to build the platform based on the comparison between its profit with development, which is given in (3.6), with that without development, which is simply $c \cdot s^*_{ss}$. Taking the firm's strategy into account, the speculators then decide how many tokens to buy, and if to participate in the ICO at all.

APPENDIX

A.1. Overbooking with Endogenous Demand

A.1.1. Ex Post versus Ex Ante Bumping Probability

Section 1.3.3, describes our rationale for modeling the customer's estimate β as the *ex post* fill rate rather than the *ex ante* bumping probability, β'. There are three main reasons behind this choice:

First, while aggregate *ex post* statistics such as (1.7) are published by BTS and widely cited in the news and travel media, the estimation of *ex ante* statistics such as (1.8) require reporting of the fractions of customers bumped from individual flights $((N_i - k)^+/N_i)$. These data are not reported by BTS and airlines do not make them available to passengers.

In fact, anecdotal evidence suggests the opposite. In direct discussions with one of us, a former airline employee noted that her employer prefer to hide the magnitude of passenger bumping at the gate, to avoid generating additional customer ill will. This assertion is all the more plausible, given the public outrage that occurred in the aftermath of the United Express Flight 3411 incident in 2017.

Second, suppose nevertheless a frequent flyer could observe data on the fraction of customers bumped on flights, $\{((N_i - k)^+/N_i) \mid i = 1, 2, \ldots\}$ and wished to estimate the *ex ante* statistic, β'. Practically speaking, bumping probabilities are generally low enough that it would require samples from many hundreds or thousands of flights for the customer to estimate the probability accurately. For example, among the 12 major US airlines tracked in the BTS statistics (BTS, 2018), the 2017 annual statistics for β range from 0.00007 to 0.00138, with a weighted average of 0.00054.

Similarly, we have calculated the β and β' associated with the optimal fixed-compensation policies evaluated §1.6 and, in absolute terms, differences between the two tend not to be large. In the 360 fixed-compensation examples evaluated in Table 2, the absolute difference between β and β' ranged from 0.023% to 0.319%, with a median of 0.079% and a mean of 0.095%; this difference would be difficult to infer from a limited numbers of samples.

To give a sense of the number of samples required to differentiate between the two statistics, we estimate the sample sizes needed to differentiate an alternative hypotheses, $H1 = \beta'$, from the null hypothesis, $H0 = \beta$, using one-sided, fixed-sample tests with Type I error of 0.1 and Type II error of 0.5 (Chow et al., 2008). Note that these are low-significance, low-powered tests that tend to minimize the required sample size.

Table 4 shows the distribution of results, which range from the low thousands to the several tens of thousands. If a frequent flyer observed the fraction of customers bumped from one of our example flights every day, it would take her at least a few years – and possibly several decades – to disambiguate β' from β.

0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
1,211	3,212	3,833	4,385	5,112	5,802	9,166	10,873	17,723	23,018	36,405

Table 4: Percentiles of Sample Size Needed to Distinguish $H1 = \beta'$ from $H0 = \beta$

Third, if we informally consider the equilibrium modelled in this paper as the one-stage stationary outcome of repeated customer purchase decisions, we can see that, if customers begin with an initial estimate of bumping that conforms to the published estimate of β, they will tend to stick with the paper's equilibrium model. That's because, although the equilibrium β and β' tend to be quite close to each other, one can show that $\beta \geq \beta'$, a fact we demonstrate at the bottom of this appendix.

More specifically, suppose customers use β as an initial estimate of β'. Those customer for whom $U(\beta, w) < 0$ will not buy a ticket, even if $U(\beta', w) \geq 0$. Conversely, those for whom $U(\beta, w) \geq 0$ do buy tickets and would continue buying as their estimates slip from β to β' and their utility increases. Of course, a detailed analysis of sample-path fluctuations would show that some customers for whom $U(\beta, w) \geq 0$ would erroneously conclude that

$U(\beta', w) < 0$ and stop buying tickets. (These customer-statisticians might use more highly powered tests than ours, tests that require larger sample sizes than those we estimate in Table 4.) and the At the same time, these are second order effects and this type of analysis is well beyond the scope of our current paper.

Finally we demonstrate that $\beta \geq \beta'$. We begin by recalling the following result.

Lemma A.1. (Wijsman (1985), Theorem 2)
Let μ be a measure on the real line $\mathbb{R}$ and let f_i, g_i $(i = 1, 2)$ be four Borel-measurable functions: $\mathbb{R} \to \mathbb{R}$ such that $f_2 \geq 0$, $g_2 \geq 0$ and $\int |f_i g_j| \, d\mu < \infty$ $(i, j = 1, 2)$. If f_1/f_2 and g_1/g_2 are monotonic in the same direction, then $\int f_1 g_1 \, d\mu \int f_2 g_2 \, d\mu \geq \int f_1 g_2 \, d\mu \int f_2 g_1 \, d\mu$.

Let μ be the cumulative distribution function of N and take $f_1(N) = \frac{(N-k)^+}{N}$, $f_2(N) = (N-k)^+$, $g_1(N) = \frac{N}{(N-k)^+}$, and $g_2(N) = 1$. Then $f_2 \geq 0$, $g_2 \geq 0$, $\int |f_1 g_1| \, d\mu = \int 1 \, d\mu = 1 < \infty$, $\int |f_1 g_2| \, d\mu = \int \frac{(N-k)^+}{N} \, d\mu < \int 1 \, d\mu = 1 < \infty$, $\int |f_2 g_1| \, d\mu = \int N \, d\mu = \mathsf{E}[N] \leq \mathsf{E}[Q] < \infty$ and $\int |f_2 g_2| \, d\mu = \int (N-k)^+ \, d\mu = \mathsf{E}[(N-k)^+] < \mathsf{E}[N] < \infty$. We also see that both $f_1/f_2 = 1/N$ and $g_1/g_2 = \frac{N}{(N-k)^+}$ are monotonically decreasing in N. We then apply Lemma A.1 and get

$$1 \cdot \mathsf{E}[(N-k)^+] = \int f_1 g_1 \, d\mu \int f_2 g_2 \, d\mu \geq \int f_1 g_2 \, d\mu \int f_2 g_1 \, d\mu = \mathsf{E}\left[\frac{(N-k)^+}{N}\right] \cdot \mathsf{E}[N]$$

and thus

$$\beta = \frac{\mathsf{E}[(N-k)^+]}{\mathsf{E}[N]} \geq \mathsf{E}\left[\frac{(N-k)^+}{N}\right] = \beta'.$$

A.1.2. Proofs

Lemma 1.3. (Properties of Loss Function Satisfied)

For a plane with k seats and loss function $\ell(s,k,\alpha) = (N(s,\alpha) - k)^+$:

(i) $N(s,\alpha) \sim \mathcal{B}(s, 1-\alpha)$ satisfies the discrete analogue of properties (i)–(iv) of Definition 1.2; and

(ii) $N(s,\alpha) \sim \mathcal{N}\left((1-\alpha)s, \sqrt{\alpha(1-\alpha)s}\right)$ satisfies properties (i), (iii), and (iv) of Definition 1.2.

Proof. Recall that Definition 1.2 states that (i) $\ell''(s) \geq 0$; (ii) $\ell(s,k,\alpha) = 0$ for all $s \leq k$ and $\ell'(s) = 0$ for all $s < k$; (iii) $\ell'(s) \in (0, 1-\alpha)$ for all $s \in [k, \infty)$; (iv) $\lim_{s\to\infty} \ell'(s) = 1-\alpha$.

Part (i). We first consider $N(s,\alpha) \sim \mathcal{B}(s, 1-\alpha)$ and verify the properties in the following order: (i) and (ii), (iv), and (iii).

For properties (i) and (ii) we have the following. As in Section 1.5.4 we note that $N(s,\alpha) \sim \mathcal{B}(s, 1-\alpha)$ is *stochastically increasing and convex* (SICX) in s. (See Section 6.A.1, including Example 6.A.2, in Shaked and Shanthikumar (1994).) This implies that $\mathsf{E}[\psi(N(s,\alpha))]$ is increasing in s for all increasing $\psi(\cdot)$ and increasing, convex in s for all increasing, convex $\psi(\cdot)$. We let $\psi(x) = (x-k)^+$, the maximum of two increasing, convex functions, 0 and $x - k$, which implies that $\psi(x)$ is increasing and convex. It then follows that $\ell(s) = \mathsf{E}[(N(s,\alpha) - k)^+] = \mathsf{E}[\psi(N(s,\alpha))]$ is increasing and convex in s, Now let's show the discrete analogue of property (iv). Note that

$$\begin{aligned}
\ell'(s) &= \ell(s+1) - \ell(s) \\
&= \mathsf{E}[(N(s+1,\alpha)-k)^+] - \mathsf{E}[(N(s,\alpha)-k)^+] \\
&= (\mathsf{E}[N(s+1,\alpha)-k] + \mathsf{E}[(k-N(s+1,\alpha))^+]) \\
&\quad - (\mathsf{E}[N(s,\alpha)-k] \ + \ \mathsf{E}[(k-N(s,\alpha))^+]) \\
&= (s+1)(1-\alpha) - k + \mathsf{E}[k-N(s+1,\alpha)|N(s+1,\alpha)<k]\mathsf{P}(N(s+1,\alpha)<k) \\
&\quad - (s(1-\alpha)-k) \ - \ \mathsf{E}[k-N(s,\alpha)|N(s,\alpha)<k]\mathsf{P}(N(s,\alpha)<k) \\
&= 1-\alpha + \mathsf{E}[k-N(s+1,\alpha)|N(s+1,\alpha)<k]\mathsf{P}(N(s+1,\alpha)<k) \\
&\quad - \mathsf{E}[k-N(s,\alpha)|N(s,\alpha)<k]\mathsf{P}(N(s,\alpha)<k). \qquad \text{(A.1)}
\end{aligned}$$

By Chebyshev's inequality, $\mathsf{P}(|N(s,\alpha)-\mu| > z\sigma) \leq 1/z^2$ where in our case $\mu = (1-\alpha)s$ and $\sigma = \sqrt{\alpha(1-\alpha)s}$. Therefore, $\mathsf{P}(N(s,\alpha) < \mu - z\sigma) = \mathsf{P}(\mu - N(s,\alpha) > z\sigma) \leq \mathsf{P}(|N(s,\alpha)-\mu| > z\sigma) \leq 1/z^2$. Let $k = \mu - z\sigma$. Then $z = \frac{\mu-k}{\sigma} = \frac{(1-\alpha)s-k}{\sqrt{\alpha(1-\alpha)s}}$. Therefore,

$$\mathsf{P}(N(s,\alpha) < k) \leq \frac{s\alpha(1-\alpha)}{((1-\alpha)s-k)^2} \to 0 \text{ as } s \to \infty. \qquad \text{(A.2)}$$

Similarly $\mathsf{P}(N(s+1, 1-\alpha) < k) \to 0$ as $s \to \infty$. Therefore, by (A.1), we have $\lim_{s\to\infty} \ell'(s) = 1-\alpha$.

Last, by property (iv) and the fact that $\ell'(s) > 0$, we have property (iii).

Part (ii). We now derive properties of interest for $\ell(s)$ under $N(s,\alpha) \sim \mathcal{N}\left((1-\alpha)s, \sqrt{\alpha(1-\alpha)s}\right)$ in the order (iv), (i), (iii).

For sales of s and a capacity of k, it is well known that the expected loss – in our case, the expected number of lost customers – equals $\sigma\, L(z)$, where $L(z) \;=\; \phi(z) - z\,(1-\Phi(z))$ is the standard normal loss function, $\sigma = \sqrt{\alpha(1-\alpha)s}$ and $z = \frac{k-(1-\alpha)\,s}{\sqrt{\alpha\,(1-\alpha)\,s}}$. (See §12.5 in Cachon and Terwiesch (2013).) Therefore,

$$\ell(s) \;=\; \sigma(s)\, L(z(s)) \;=\; \sqrt{\alpha(1-\alpha)s}\, L\left(\frac{k-(1-\alpha)s}{\sqrt{\alpha(1-\alpha)s}}\right). \tag{A.3}$$

Noting that $L'(z) = \Phi(z) - 1$, $\sigma'(s) = \frac{\sigma(s)}{2\,s}$, and $z'(s) = -\left[\frac{1-\alpha}{\sigma(s)} \;+\; \frac{z(s)}{2\,s}\right]$, we can differentiate with respect to s and collect terms to obtain

$$\begin{aligned}
\ell'(s) \;&=\; \sigma'(s)\, L(z(s)) \;+\; \sigma(s)\, L'(z(s))\, z'(s) \\
&=\; \frac{\sigma(s)}{2\,s}\,(\phi(z(s)) - z(s)\,(1-\Phi(z(s)))) \;+\; \sigma(s)\,(1-\Phi(z(s))\left(\frac{1-\alpha}{\sigma(s)} \;+\; \frac{z(s)}{2\,s}\right) \\
&=\; \frac{\sigma(s)}{2\,s}\,\phi(z(s)) \;+\; (1-\alpha)\,(1-\Phi(z(s))).
\end{aligned} \tag{A.4}$$

Note, for any $s > 0$, we have $\lim\limits_{s\to\infty} \ell'(s) = (1-\alpha)$, property (iv).

We can then differentiate (A.4) again with respect to s to find $\ell''(s)$.

$$\ell''(s) \;=\; \frac{2\,s\,\sigma'(s) - 2\,\sigma(s)}{4\,s^2}\phi(z(s)) \;+\; \frac{\sigma(s)}{2\,s}\phi'(z(s))\,z'(s) \;-\; (1-\alpha)\,\phi(z(s))\,z'(s) \tag{A.5}$$

We next recall that $\phi'(z) = -z\,\phi(z)$. We use this identity, along with those for $\sigma'(s)$ and $z'(s)$, to substitute out terms with derivatives in (A.5). Collecting terms, we have

$$\ell''(s) \;=\; \left[(\sigma(s)z(s) + 2s(1-\alpha))^2 - \sigma^2(s)\right]\,\frac{\phi(z(s))}{4\,s^2\sigma(s)}. \tag{A.6}$$

To ensure that $\ell''(s) > 0$, we need $\sigma(s)z(s) + 2s(1-\alpha) - \sigma(s) > 0$, which is equivalent to $k - (1-\alpha)s + 2s(1-\alpha) > \sqrt{\alpha\,(1-\alpha)\,s}$. Squaring both sides, this is $k^2 + 2ks(1-\alpha) + s^2(1-\alpha)^2 - \alpha(1-\alpha)s > 0$. The same inequality can be rearranged as

$$\left[k + \frac{(2k-\alpha)(1-\alpha)}{2k}\right]^2 + s^2(1-\alpha)^2\left[1 - \frac{(2k-\alpha)^2}{4k^2}\right] > 0. \tag{A.7}$$

A sufficient condition for (A.7) to hold is $1 - \frac{(2k-\alpha)^2}{4k^2} \geq 0$, which can be simplified to $k \geq \frac{\alpha}{4}$.

We have thus shown that $\ell''(s) > 0$ when $k \geq \frac{\alpha}{4}$, which practically holds given $k \geq 1$. We have thus shown that property (i) holds.

Last, by (A.4), we know that $\ell'(s) > 0$. Since $\ell''(s) > 0$ and $\lim_{s\to\infty} \ell'(s) = (1-\alpha)$, we must have $\ell'(s) \in (0, 1-\alpha)$ for all $s \in [k, \infty)$, property (iii). $\square$

Lemma 1.4. (Existence and Uniqueness of Equilibria)

(i) For overbooking policies with $p = (1-\alpha)v$, $b > k$, and $c \leq \underline{w}$ there is no equilibrium.

(ii) For all other overbooking policies $\xi \in \Xi$, there exists at least one equilibrium.

(iii) For the policies in part (ii), if $g'(w) \leq 0$, $\forall\, w \in [c, \overline{w}]$, then $\exists$ a unique equilibrium $\{\beta, \widehat{w}\}$.

Proof.

Part (i). The existence of equilibrium means that there exists at least one $\widehat{w}$ such that

$$U(\beta(\widehat{w}), \widehat{w}) = -p + (1-\alpha)v + (1-\alpha)\beta(\widehat{w})(c - \widehat{w}) \geq 0, \tag{A.8}$$

where, from (1.3)-(1.6) and (1.12),

$$\beta(w) = \frac{\mathsf{E}\left[\ell(\min\{b, QG(w)\})\right]}{(1-\alpha)\mathsf{E}\left[\min\{b, QG(w)\}\right]}. \tag{A.9}$$

Consider a policy with $c \leq \underline{w}$, $b > k$ and $p = (1-\alpha)v$. Suppose there is an equilibrium ($\widehat{w}$, $\beta(\widehat{w})$). Then we know that $U(\beta(\widehat{w}), \widehat{w}) = 0 + (1-\alpha)\beta(\widehat{w})(c - \widehat{w}) \leq 0$ regardless of the value of $\beta(\widehat{w})$ because $\widehat{w} \geq \underline{w} \geq c$, From (1.2) we know that this implies $\widehat{w} = \underline{w}$. Now we show by l'Hôpital's rule that $\beta(\underline{w}) = 0$. By (A.9),

$$\begin{aligned}\lim_{w\to\underline{w}} \beta(w) &= \lim_{w\to\underline{w}} \frac{\mathsf{E}\left[\ell(\min\{b, QG(w)\})\right]}{(1-\alpha)\mathsf{E}\left[\min\{b, QG(w)\}\right]} \\ &= \frac{\lim_{w\to\underline{w}} \mathsf{E}\left[\ell(\min\{b, QG(w)\})\right]}{\lim_{w\to\underline{w}} (1-\alpha)\mathsf{E}\left[\min\{b, QG(w)\}\right]} = \frac{\lim_{w\to\underline{w}} \int_0^{\frac{b}{G(w)}} \ell'(q\, G(w))\, g(w)\, q\, f(q)dq}{\lim_{w\to\underline{w}} (1-\alpha) \int_0^{\frac{b}{G(w)}} g(w)\, q\, f(q)dq}\end{aligned} \tag{A.10}$$

by (1.21) and (1.22). The denominator of (A.10) is positive and finite, since $\mathsf{E}[Q] < \infty$ and $g(w) > 0$ over its support, and the numerator equals 0. Thus, $\lim_{w \to \underline{w}} \beta(w) = 0$. This implies that $U(\beta(\widehat{w}), \widehat{w}) = U(\beta(\underline{w}), \underline{w}) = 0$ and thus $\widehat{w} = \overline{w}$. Since $\underline{w} \neq \overline{w}$, this policy does not yield a consistent $\widehat{w}$.

Part (ii). We divide the space of admissible overbooking policies into four partitions: 1) $b = k$, 2) $b > k$, $\underline{w} < c \leq \overline{w}$, 3) $b > k$, $c \leq \underline{w}$ and $p < (1-\alpha)v$, 4) $b > k$, $c \leq \underline{w}$ and $p = (1-\alpha)v$. Recall that in part (i) we have shown that policies in 4) obtain no equilibrium. We will now show the existence of $\widehat{w}$ that satisfies (A.8) for all policies in 1), 2) and 3).

1) When $b = k$, the equilibrium bumping probability is $\beta = 0$. By Definition 1.1, customers obtain non-negative utility from buying a ticket. The corresponding equilibrium $\widehat{w}$ is $\overline{w}$ and thus $U(\beta, \widehat{w}) \geq 0$.

2) Next consider policies with $b > k$ and $\underline{w} < c \leq \overline{w}$. We know that $U(\beta(w), w) \geq 0$ for all $w \in [\underline{w}, c]$. If $U(\beta(\overline{w}), \overline{w}) \geq 0$, then $\widehat{w} = \overline{w}$. If $U(\beta(\overline{w}), \overline{w}) < 0$, then since $U(\beta(c), c) \geq 0$, by the Intermediate Value Theorem, there exists at least one $\widehat{w} \in (c, \overline{w}]$ such that $U(\beta(\widehat{w}), \widehat{w}) = 0$.

3) Finally, consider policies with $b > k$ and $c \leq \underline{w}$ and $p < (1-\alpha)v$, we have $c - \widehat{w} \leq 0$ and $-p + (1-\alpha)v > 0$. When $w = \underline{w}$, $\beta(\underline{w}) = 0$ and thus $U(\beta(\underline{w}), \underline{w}) > 0$. If $U(\beta(w), w) \geq 0$ for all $w \in (\underline{w}, \overline{w}]$, then $\widehat{w} = \overline{w}$. Otherwise, by the Intermediate Value Theorem, there exists at least one $\widehat{w} \in (\underline{w}, \overline{w}]$ such that $U(\beta(\widehat{w}), \widehat{w}) = 0$.

Part (iii). Let $U(\beta(w), w) = -p + (1-\alpha)v + (1-\alpha)h(w)$ where $h(w) = \beta(w)(c - w)$ and $\beta(w)$ is defined as in (A.9). Note, if $U(\beta(w), w) \geq 0$ for all w, then the unique equilibrium $\widehat{w}$ is $\overline{w}$. Otherwise, by contradiction suppose that there are multiple solutions (zeros of $U(\beta(w), w)$) and ($\widehat{w}_1$, β_1), with ($\widehat{w}_2$, β_2) being two of them, ($\widehat{w}_1, \widehat{w}_2 \in [c, \overline{w}]$).

Because $U(\beta(\widehat{w}_1), \widehat{w}_1) = U(\beta(\widehat{w}_2), \widehat{w}_2) = 0$, there must exist some $w \in [\widehat{w}_1, \widehat{w}_2]$ such that $\frac{dU(\beta(w),w)}{dw} = (1-\alpha)\frac{dh(w)}{dw} = 0$. Therefore, we can show that, if $\frac{dh(w)}{dw} < 0$ for all $w \in [c, \overline{w}]$, then $\widehat{w}$ is unique. Note that, from the definition of $h(w)$ and (A.9), we have

$$\begin{aligned}
\frac{dh(w)}{dw} &= \frac{d\beta}{dw}(c-w) + \beta \cdot (-1) \\
&= \left[\beta \frac{g(w)\int_0^{\frac{b}{G(w)}} q\, f(q)\, dq}{\mathsf{E}[S]} - \frac{g(w)\int_0^{\frac{b}{G(w)}} \ell'(qG(w))q\, f(q)\, dq}{\mathsf{E}[N]}\right](w-c) - \beta \\
&\leq \beta \frac{g(w)\int_0^{\frac{b}{G(w)}} q\, f(q)\, dq}{\mathsf{E}[S]}(w-c) - \beta \\
&< \beta \frac{g(w)}{G(w)}(w-c) - \beta \\
&= \beta \left[\frac{g(w)}{G(w)}(w-c) - 1\right]. \qquad \text{(A.11)}
\end{aligned}$$

If $\frac{g(w)}{G(w)}(w-c)-1 \leq 0$ for $w \in [c, \overline{w}]$, then $U(\beta(w), w)$ decreases in w and therefore has at most one zero. In this case, $\min_{w \in [c,\overline{w}]} U(\beta(w), w) = U(\beta(\overline{w}), \overline{w})$ and $\max_{w \in [c,\overline{w}]} U(\beta(w), w) = U(\beta(c), c)$. Note that $U(\beta(c), c) \geq 0$. If $U(\beta(\overline{w}), \overline{w}) \geq 0$, then $\widehat{w} = \overline{w}$. Otherwise, there exists exactly one w that satisfies $U(\beta(w), w) = 0$, and that w is $\widehat{w}$.

The uniqueness condition derived from as (A.11) can be rewritten as

$$\psi(w) \equiv w - \frac{G(w)}{g(w)} \leq c. \qquad \text{(A.12)}$$

Taking the derivative of $\psi(w)$ w.r.t. w, we have $\frac{\partial \psi(w)}{\partial w} = \frac{G(w)}{g^2(w)} g'(w)$. Since $\psi(c) = c - \frac{g(c)}{G(c)} < c$, as long as $g'(w) \leq 0$, $\frac{\partial \psi(w)}{\partial w} \leq 0$ for $w \in [c, \overline{w}]$, and (A.12) is satisfied.

Specific Distributions

If the hassle cost is uniformly distributed, then $g'(w) = 0$ and $\frac{\partial \psi(w)}{\partial w} = 0$ and (A.12) always holds. Equation (A.12) also holds for exponentially distributed hassle cost because $g'(w) < 0$. Now consider normally distributed hassle cost, $W \sim \mathcal{N}(\mu, \sigma^2)$. When $w > \mu$, $g'(w) < 0$ and thus $\psi'(w) < 0$. (A.12) is automatically satisfied since $c > c - \frac{g(c)}{G(c)} = \psi(c) \geq$

$\psi(w) = w - \frac{g(w)}{G(w)}$ for all $w \geq c$. When $w < \mu$, $g'(w) > 0$ and thus $\psi'(w) > 0$. Thus, (A.12) holds if $c > \psi(\mu) = \mu - \frac{g(\mu)}{G(\mu)}$ because $\psi(\mu) \geq \psi(w)$ for all $w \geq c$. Therefore, we conclude that, for normally distributed W, the uniqueness condition holds if $c > \mu - \frac{g(\mu)}{G(\mu)}$. □

Lemma 1.5. (Ordering of Equilibria)

Suppose an overbooking policy $\xi \in \Xi$ induces multiple equilibria. Pick any two distinct equilibria from the set, and call them $(\beta_1, \widehat{w}_1) \neq (\beta_2, \widehat{w}_2)$.

(i) Without loss of generality, we can order the two so that the second equilibrium has a strictly lower bumping probability and a strictly higher marginal hassle cost: $\beta_1 > \beta_2$ and $\widehat{w}_1 < \widehat{w}_2$.

(ii) Given the ordering in (i), the set of customers with $w \leq \widehat{w}_1$ is a strict subset of those with $w \leq \widehat{w}_2$, and the airline earns strictly higher expected profits in $(\beta_2, \widehat{w}_2)$.

Proof.

Part (i). Our proof proceeds in five steps.

First, suppose $(\beta_1, \widehat{w}_1)$ and $(\beta_2, \widehat{w}_2)$ are distinct equilibria. Then without loss of generality, we can assume $\underline{w} \leq \widehat{w}_1 < \widehat{w}_2 \leq \overline{w}$. This is because β can be expressed as a function of $\widehat{w}$, as in (A.9), so $\widehat{w}_1 = \widehat{w}_2$ implies $\beta_1 = \beta_2$.

Second, we show that $U(\beta, \widehat{w}_1) = 0$. Since $\underline{w} \leq \widehat{w}_1 < \widehat{w}_2 \leq \overline{w}$, we have $\widehat{w}_1 < \overline{w}$, and from (1.2) we know this implies $U(\beta_1, \widehat{w}_1) \leq 0$. By contradiction, suppose that $U(\beta_1, \widehat{w}_1) < 0$. Then from (1.2) we also know that $\widehat{w}_1 = \underline{w}$, and as in (A.10) we can show that this implies $\beta_1 = 0$. At the same tim, since $\xi \in \Xi$ is admissible, $p \leq (1-\alpha)v$, and since $\beta_1 = 0$, (1.2) implies that $U(\beta_1, \widehat{w}_1) = -p + (1-\alpha)v + (1-\alpha)\beta_1(c - \widehat{w}_1) = -p + (1-\alpha)v \geq 0$, a contradiction.

Third, we show that, if there exist multiple equilibria, then we must have $\beta_1 > 0$. By contradiction, suppose not. Then we have $\underline{w} \leq \widehat{w}_1 < \widehat{w}_2 \leq \overline{w}$ and, because $\beta_1 = 0$, (1.1) shows that $U(\beta_1, w) = -p + (1-\alpha)v + (1-\alpha)\beta_1(c - w) \geq 0$, for all $w \in [\underline{w}, \overline{w}]$, including $w > \widehat{w}_1$. Thus $\widehat{w}_1 < \overline{w}$ is not an equilibrium threshold customer response, so $U(\beta_1, \widehat{w}_1) = 0$ is not an equilibrium, a contradiction.

Fourth, we show that, if $\beta_1 > 0$ then $\beta_2 > 0$ as well. If we look at the definition of β in (A.9), we see that the numerator, $\mathsf{E}[\ell(\min\{b, QG(\widehat{w})]$, is increasing in $\widehat{w}$. Thus, we have $0 < \mathsf{E}[\ell(\min\{b, QG(\widehat{w}_1)] < \mathsf{E}[\ell(\min\{b, QG(\widehat{w}_2)]$. While we do not know how the ratio in (A.9) that determines β changes, we do know that $\beta_2 > 0$.

Finally, we now have $p \leq (1-\alpha)v$, $\underline{w} \leq \widehat{w}_1 < \widehat{w}_2 \leq \overline{w}$, $\beta_1, \beta_2 > 0$, and

$$0 = U(\beta_1, \widehat{w}_1) = -p + (1-\alpha)v + (1-\alpha)\beta_1(c - \widehat{w}_1) \leq -p + (1-\alpha)v + (1-\alpha)\beta_2(c - \widehat{w}_2) = U(\beta_2, \widehat{w}_2),$$

so that $\beta_1(c - \widehat{w}_1) \leq \beta_2(c - \widehat{w}_2)$. Note that $\widehat{w}_1 < \widehat{w}_2$ implies that $c - \widehat{w}_1 > c - \widehat{w}_2$. Furthermore, $U(\beta_1, \widehat{w}_1) = 0$ and $p \leq (1-\alpha)v$ imply that $c - \widehat{w}_1 \leq 0$, so we have $c - \widehat{w}_2 < c - \widehat{w}_1 \leq 0$. Given $c - \widehat{w}_2 < 0$, $\beta_1, \beta_2 > 0$, and $\beta_1(c - \widehat{w}_1) \leq \beta_2(c - \widehat{w}_2)$, we then have

$$\frac{\beta_2}{\beta_1} \leq \frac{c - \widehat{w}_1}{c - \widehat{w}_2} = \frac{\widehat{w}_1 - c}{\widehat{w}_2 - c} < 1.$$

Thus for any two distinct equilibria, we can order them so that $\beta_1 > \beta_2$ and $\widehat{w}_1 < \widehat{w}_2$.

Part (ii). Let $\Pi_i = p\,\mathsf{E}[S_i] - (c+r)\,\mathsf{E}[(N_i - k)^+] = \mathsf{E}[S_i](p - (1-\alpha)(c+r)\beta_i)$ for $i = 1, 2$. $\widehat{w}_1 < \widehat{w}_2$ leads to $\mathsf{E}[N_1] < \mathsf{E}[N_2]$ and $\beta_1 > \beta_2$ gives $-(c+r)\beta_1 < -(c+r)\beta_2$. Therefore, $\Pi_1 < \Pi_2$. □

Proposition 1.7. (Optimal Myopic Booking Limit)

Given fixed, admissible p and c, the optimal myopic booking limit, b^*_m, behaves as follows.

(i) If $p - (1-\alpha)(c+r) \geq 0$, then $b^*_m = \infty$, and the airline does not impose a booking limit.

(ii) If $p - (1-\alpha)(c+r) < 0$, then there exists a unique optimal $b^*_m = \max\left\{\ell'^{-1}\left(\frac{p}{c+r}\right), k\right\}$.

(iii) When $b^*_m \in (k, \infty)$, $\frac{\partial \mathsf{E}[\Pi]}{\partial b} > 0$ for $b < b^*_m$ and $\frac{\partial \mathsf{E}[\Pi]}{\partial b} < 0$ for $b > b^*_m$.

Proof. Recall the FOC given by (1.19): $p - \ell'(b)\,(c+r) = 0$.

Part (i). By Definition 1.2 part (iii), $p - \ell'(b)\,(c+r) > p - (1-\alpha)\,(c+r)$. Therefore, when $p - (1-\alpha)\,(c+r) \geq 0$, the marginal increase in profit is always positive and the airline is incentivized to overbook as much as possible.

Part (ii). When $p-(1-\alpha)\,(c+r) < 0$, if $p-\ell'(k)\,(c+r) > 0$, the FOC (1.19) has a solution by the Intermediate Value Theorem. By Definition 1.2 part (iii), $\ell'(b)$ increases monotonically in b for $b \geq k$ from 0 to $1-\alpha$, hence the solution is unique. If $p - \ell'(k)\,(c+r) \leq 0$, then $b_m^* = k$.

Part (iii). We know that $\frac{\partial \Pi}{\partial b} = [p-\ell'(b)(c+r)]\mathsf{P}\{QG(\widehat{w}) > b\}$ and $p-\ell'(b_m^*)(c+r) = 0$. By Definition 1.2 part (i), $p-\ell'(b)(c+r) > 0$ for $b < b_m^*$ and $p-\ell'(b)(c+r) < 0$ for $b > b_m^*$. □

Proposition 1.8. (Optimal Myopic Booking Limit is Profit-Making)
(i) The equilibrium induced by any $\xi \in \Xi$ obtains $\ell'(b) > (1-\alpha)\beta$.

Suppose $p > 0$.

(ii) If $\beta = 0$, or if either $p < (1-\alpha)v$ or $c > \underline{w}$ or both, then b_m^* induces a profit-making equilibrium.

Proof.

Part (i). Because $\beta = \frac{\mathsf{E}[(N-k)^+]}{\mathsf{E}[N]} = \frac{\mathsf{E}[(N-k)^+]}{(1-\alpha)\mathsf{E}[S]}$ and $\ell(b) > (1-\alpha)\beta$, it is sufficient to show that $\ell'(b)\mathsf{E}[S] > \mathsf{E}[(N-k)^+]$. Expanding this expression, this is equivalent to showing

$$\ell'(b)\left(\int_0^{\frac{b}{G(\widehat{w})}} qG(\widehat{w})\,f(q)\,dq \;+\; \int_{\frac{b}{G(\widehat{w})}}^{\infty} b\,f(q)\,dq\right) > \int_0^{\frac{b}{G(\widehat{w})}} \ell(qG(\widehat{w}))\,f(q)\,dq + \int_{\frac{b}{G(\widehat{w})}}^{\infty} \ell(b)\,f(q)\,dq. \tag{A.13}$$

Rearranging (A.13) gives

$$\int_0^{\frac{b}{G(\widehat{w})}} [qG(\widehat{w})\ell'(b) \ - \ \ell(qG(\widehat{w}))]\, f(q)\, dq \ + \ \int_{\frac{b}{G(\widehat{w})}}^{\infty} [b\ell'(b) \ - \ \ell(b)]\, f(q)\, dq > 0. \tag{A.14}$$

Note that $\ell(b) = \int_0^b \ell'(t)dt < \int_0^b \ell'(b)dt = b\ell'(b)$ for all $b > k$ and $\ell(b) = b\ell'(b) = 0$ for all $b \le k$. Therefore, the second integral of (A.14) is positive. To see that the first integral of (A.14) is non-negative, note that by the convexity of $\ell(\cdot)$, $qG(\widehat{w})\ell'(b) \ge qG(\widehat{w})\ell'(qG(\widehat{w}))$ for $q \le \frac{b}{G(\widehat{w})}$. Therefore, $qG(\widehat{w})\ell'(b) \ - \ \ell(qG(\widehat{w})) \ge qG(\widehat{w})\ell'(qG(\widehat{w})) \ - \ \ell(qG(\widehat{w}))$ for $q \le \frac{b}{G(\widehat{w})}$. Again by $\ell(b) \le b\ell'(b)$ for any b, we know that $qG(\widehat{w})\ell'(qG(\widehat{w})) \ - \ \ell(qG(\widehat{w})) \ge 0$.

Part (ii). For $\beta = 0$, we have $\widehat{w} = \overline{w}$, so $p, \mathsf{E}[S] > 0$ and $\mathsf{E}[\Pi(p, b_m^*, c)] = p\mathsf{E}[S] - (c + r)\mathsf{E}[(N-k)^+] = p\mathsf{E}[S] > 0$. Otherwise, taking the FOC in (1.19) $p - \ell'(b_m^*)(c+r) = 0$ and using the inequality from part (i) to substitute for $\ell'(b_m^*)$, we immediately have $[p - (1-\alpha)\beta(c+r)] > 0$, where β is the equilibrium bumping probability induced by b_m^*. Since $c > \underline{w}$ or $p < (1-\alpha)v$, the $\widehat{w}$ induced by b_m^* must be greater than $\underline{w}$ and thus $\mathsf{E}[S] = \int_0^{\frac{b_m^*}{G(\widehat{w})}} qG(\widehat{w})\, f(q)\, dq \ + \ \int_{\frac{b_m^*}{G(\widehat{w})}}^{\infty} b_m^*\, f(q)\, dq > 0$. By Definition 1.6, the result follows. ☐

Proposition 1.9. (Optimal Strategic Booking Limit)
Suppose $\exists\ p \in (0, (1-\alpha)v)$ *and* $c \in (0, \overline{w})$ *for which* $b_m^* \in (k, \infty)$ *induces a profit-making equilibrium* $\widehat{w} \in (\underline{w}, \overline{w})$*. Then we have the following.*
(i) For any given $b > k$*, if* $\beta \ge \sqrt{(v - \frac{p}{1-\alpha})\frac{g(\widehat{w})}{G(\widehat{w})}}$*, then* $\frac{d\widehat{w}}{db} < 0$*.*
(ii) In turn, if $\beta > \sqrt{(v - \frac{p}{1-\alpha})\frac{g(\widehat{w})}{G(\widehat{w})}}$ *for all* $b > k$*, then* $b_s^* < b_m^*$*.*

Proof. We continue to consider policy parameters and equilibria that allow us to develop relevant FOCs. These include policies for which $p \in (0, (1-\alpha)v)$, $c \in (0, \overline{w})$, and $b \in (k, \infty)$. For the same reason, we will assume that the policy (p, b, c) obtains an interior equilibrium $U(\beta, \widehat{w}) = 0$ for which $\widehat{w} \in (\underline{w}, \overline{w})$.

Part (i). Since $U(\beta, \widehat{w}) = 0$, by (1.1) we can express the break-even hassle cost as

$$\widehat{w} \;=\; c + \left(v - \frac{p}{1-\alpha}\right)\frac{1}{\beta}. \tag{A.15}$$

Note that since $b > k$, we have $\beta > 0$ and thus $1/\beta$ is well-defined.

Differentiating $\widehat{w}$ with respect to b according to (A.15),

$$\begin{aligned}
\frac{d\widehat{w}}{db} \;&=\; \left(v - \frac{p}{1-\alpha}\right)\; \frac{\mathsf{E}[(N-k)^+]\frac{d\mathsf{E}[N]}{db} \;-\mathsf{E}[N]\frac{d\mathsf{E}[(N-k)^+]}{db}}{\mathsf{E}[(N-k)^+]^2} \\
&=\; \left(v - \frac{p}{1-\alpha}\right)\; \frac{\mathsf{E}[(N-k)^+]\frac{\partial\mathsf{E}[N]}{\partial b} \;-\mathsf{E}[N]\frac{\partial\mathsf{E}[(N-k)^+]}{\partial b}}{\mathsf{E}[(N-k)^+]^2} \\
&\quad + \left(v - \frac{p}{1-\alpha}\right)\; \frac{\mathsf{E}[(N-k)^+]\frac{\partial\mathsf{E}[N]}{\partial \widehat{w}} \;-\mathsf{E}[N]\frac{\partial\mathsf{E}[(N-k)^+]}{\partial \widehat{w}}}{\mathsf{E}[(N-k)^+]^2}\frac{d\widehat{w}}{db}.
\end{aligned} \tag{A.16}$$

Rearranging (A.16) and applying (1.6), (1.16), (1.18), (1.21) and (1.22), we have

$$\begin{aligned}
\frac{d\widehat{w}}{db} \;&=\; \frac{\left(v - \frac{p}{1-\alpha}\right)\;\left[(1-\alpha)\beta\frac{\partial\mathsf{E}[S]}{\partial b} \;-\; \frac{\partial\mathsf{E}[(N-k)^+]}{\partial b}\right]}{\beta^2\mathsf{E}[N] - \left(v - \frac{p}{1-\alpha}\right)\;\left[(1-\alpha)\beta\frac{\partial\mathsf{E}[S]}{\partial \widehat{w}} \;-\; \frac{\partial\mathsf{E}[(N-k)^+]}{\partial \widehat{w}}\right]} \\
&=\; \frac{\left(v - \frac{p}{1-\alpha}\right)\;\mathsf{P}\{QG(\widehat{w}) > b\}\;\left[(1-\alpha)\beta \;-\; \ell'(b)\right]}{\beta^2\mathsf{E}[N] - \left(v - \frac{p}{1-\alpha}\right)\; g(\widehat{w})\;\int_0^{\frac{b}{G(\widehat{w})}}\left[(1-\alpha)\beta \;-\; \ell'(qG(\widehat{w}))\right] qf(q)dq} \\
&=\; \frac{\left(v - \frac{p}{1-\alpha}\right)\;\mathsf{P}\{QG(\widehat{w}) > b\}\;\left[(1-\alpha)\beta \;-\; \ell'(b)\right]}{\beta^2\mathsf{E}[N] + \left(v - \frac{p}{1-\alpha}\right)\; g(\widehat{w})\;\int_0^{\frac{b}{G(\widehat{w})}}\left[\ell'(qG(\widehat{w})) \;-\; (1-\alpha)\beta\right] qf(q)dq} \\
&=\; \frac{\left(v - \frac{p}{1-\alpha}\right)\;\mathsf{P}\{QG(\widehat{w}) > b\}\;\left[(1-\alpha)\beta \;-\; \ell'(b)\right]}{\frac{\mathsf{E}[(N-k)^+]^2}{\mathsf{E}[N]} + \left(v - \frac{p}{1-\alpha}\right)\; g(\widehat{w})\;\int_0^{\frac{b}{G(\widehat{w})}}\left[\ell'(qG(\widehat{w})) \;-\; (1-\alpha)\frac{\mathsf{E}[(N-k)^+]}{\mathsf{E}[N]}\right] qf(q)dq} \\
&=\; \frac{\left(v - \frac{p}{1-\alpha}\right)\;\mathsf{P}\{QG(\widehat{w}) > b\}\;\left[(1-\alpha)\beta \;-\; \ell'(b)\right]}{\frac{\mathsf{E}[(N-k)^+]^2}{\mathsf{E}[N]} + \left(v - \frac{p}{1-\alpha}\right)\; \frac{g(\widehat{w})}{\mathsf{E}[N]}\;\int_0^{\frac{b}{G(\widehat{w})}}\left[\mathsf{E}[N]\ell'(qG(\widehat{w})) - (1-\alpha)\mathsf{E}[(N-k)^+]\right] qf(q)dq}
\end{aligned} \tag{A.17}$$

The numerator of (A.17) is negative by Proposition 1.8 part (i), therefore $\frac{d\widehat{w}}{db} < 0$ if and only if the denominator of (A.17) is positive. Note that the integral in the denominator
$\displaystyle\int_0^{\frac{b}{G(\widehat{w})}} \left(\mathsf{E}[N]\ell'(qG(\widehat{w})) - \mathsf{E}[(N-k)^+](1-\alpha)\right) qf(q)dq$

$$= \int_0^{\frac{b}{G(\widehat{w})}} \left(\mathsf{E}[N]\ell'(b) - \mathsf{E}[(N-k)^+](1-\alpha)\right) qf(q)dq$$
$$+ \int_0^{\frac{b}{G(\widehat{w})}} \left(\mathsf{E}[N]\ell'(qG(\widehat{w})) - \mathsf{E}[N]\ell'(b)\right) qf(q)dq, \tag{A.18}$$

and the first integral of (A.18) is positive by Proposition 1.8 part (i). Therefore, it suffices to have $\frac{\mathsf{E}[(N-k)^+]^2}{\mathsf{E}[N]} + \left(v - \frac{p}{1-\alpha}\right) \frac{g(\widehat{w})}{\mathsf{E}[N]} \int_0^{\frac{b}{G(\widehat{w})}} \left[\mathsf{E}[N]\ell'(qG(\widehat{w})) - \mathsf{E}[N]\ell'(b)\right] qf(q)dq \geq 0$, which is equivalent to

$$\left(v - \frac{p}{1-\alpha}\right) g(\widehat{w}) \int_0^{\frac{b}{G(\widehat{w})}} \left[\ell'(b) - \ell'(qG(\widehat{w}))\right] qf(q)dq \leq \frac{\mathsf{E}[(N-k)^+]^2}{\mathsf{E}[N]}. \tag{A.19}$$

By Definition 1.2 part (iii), we know that $\ell'(b) - \ell'(qG(\widehat{w})) < 1-\alpha$ for $q \leq \frac{b}{G(\widehat{w})}$. Then, the left-hand side of (A.19) is
$\left(v - \frac{p}{1-\alpha}\right) g(\widehat{w}) \int_0^{\frac{b}{G(\widehat{w})}} \left[\ell'(b) - \ell'(qG(\widehat{w}))\right] qf(q)dq$

$$\begin{aligned}
&< \left(v - \frac{p}{1-\alpha}\right) g(\widehat{w}) \int_0^{\frac{b}{G(\widehat{w})}} [1-\alpha]\, qf(q)dq \\
&= \left(v - \frac{p}{1-\alpha}\right) \frac{g(\widehat{w})}{G(\widehat{w})} [1-\alpha] \int_0^{\frac{b}{G(\widehat{w})}} qG(\widehat{w})f(q)dq \\
&< \left(v - \frac{p}{1-\alpha}\right) \frac{g(\widehat{w})}{G(\widehat{w})} [1-\alpha] \left[\int_0^{\frac{b}{G(\widehat{w})}} qG(\widehat{w})f(q)dq + \int_{\frac{b}{G(\widehat{w})}}^{\infty} bf(q)dq\right] \\
&= \left(v - \frac{p}{1-\alpha}\right) \frac{g(\widehat{w})}{G(\widehat{w})} [1-\alpha]\, \mathsf{E}[S] \\
&= \left(v - \frac{p}{1-\alpha}\right) \frac{g(\widehat{w})}{G(\widehat{w})} \mathsf{E}[N].
\end{aligned} \tag{A.20}$$

By (A.20), a sufficient condition to satisfy (A.19) is $\left(v - \frac{p}{1-\alpha}\right) \frac{g(\widehat{w})}{G(\widehat{w})} \mathsf{E}[N] \leq \frac{\mathsf{E}[(N-k)^+]^2}{\mathsf{E}[N]}$. Recall that $\beta = \mathsf{E}[(N-k)^+]/\mathsf{E}[N]$. Therefore, the sufficient condition for (A.19) to hold is $\beta(b, \widehat{w}) \geq \sqrt{(v - \frac{p}{1-\alpha}) \frac{g(\widehat{w})}{G(\widehat{w})}}$.

Part (ii). As a first step, we will show that $\frac{\partial \mathsf{E}[\Pi]}{\partial \widehat{w}} > 0$ for all policies as described at the beginning of the proof that are profit-making. Rearranging (A.15), we express the equilibrium β in terms of $\widehat{w}$ as

$$\beta \;=\; \frac{v - p/(1-\alpha)}{\widehat{w} - c}. \tag{A.21}$$

Since $p < (1-\alpha)v$ and $\beta > 0$, we must have $\widehat{w} > c$. Substituting the expression for β in (A.21) into (1.13) and differentiating with respect to $\widehat{w}$, we have

$$\begin{aligned}
\frac{\partial \mathsf{E}[\Pi]}{\partial \widehat{w}} &= \left[(c+r)\,\frac{(1-\alpha)v - p}{(\widehat{w}-c)^2}\right] \cdot \mathsf{E}\left[\min\{b, QG(\widehat{w})\}\right] \\
&\quad + \left[p - (c+r)\,\frac{(1-\alpha)v - p}{\widehat{w}-c}\right] \cdot g(\widehat{w})\, \mathsf{P}\{QG(\widehat{w}) < b\}\, \mathsf{E}[Q \mid QG(\widehat{w}) < b] \\
&= \left[\frac{(1-\alpha)\,(c+r)\,\beta}{\widehat{w}-c}\right] \cdot \mathsf{E}[\min\{b, QG(\widehat{w})\}] \\
&\quad + [p - (1-\alpha)\,(c+r)\,\beta]\; g(\widehat{w})\; \mathsf{P}\{QG(\widehat{w}) < b\}\, \mathsf{E}[Q \mid QG(\widehat{w}) < b]. \qquad \text{(A.22)}
\end{aligned}$$

The first term of (A.22) is positive. Since the equilibrium is profit-making, by Definition 1.6, we must have $p - (1-\alpha)\,(c+r)\,\beta \;>\; 0$ as well. Therefore, the second term of (A.22) is positive. Hence, we know that $\frac{\partial \mathsf{E}[\Pi]}{\partial \widehat{w}} > 0$.

Now note that any optimal b satisfies $\frac{d\Pi}{db} = \frac{\partial \mathsf{E}[\Pi]}{\partial b} + \frac{\partial \mathsf{E}[\Pi]}{\partial \widehat{w}} \frac{d\widehat{w}}{db} = 0$. Since $\frac{\partial \mathsf{E}[\Pi]}{\partial \widehat{w}} > 0$, $\frac{\partial \mathsf{E}[\Pi]}{\partial \widehat{w}} \frac{d\widehat{w}}{db} < 0$ when $\frac{d\widehat{w}}{db} < 0$. The optimal myopic booking limit, b^*_m, satisfies $\frac{d\Pi}{db}|_{b=b^*_m} = \frac{\partial \mathsf{E}[\Pi]}{\partial b}|_{b=b^*_m} + \frac{\partial \mathsf{E}[\Pi]}{\partial \widehat{w}} \frac{d\widehat{w}}{db}|_{b=b^*_m} = 0 + \frac{\partial \mathsf{E}[\Pi]}{\partial \widehat{w}} \frac{d\widehat{w}}{db}|_{b=b^*_m} < 0$. Hence, $\frac{d\Pi}{db} = 0$ only when $\frac{\partial \mathsf{E}[\Pi]}{\partial b} > 0$. By Proposition 1.7 part (iii), $\frac{\partial \mathsf{E}[\Pi]}{\partial b} > 0$ for all $b < b^*_m$, and $\frac{\partial \mathsf{E}[\Pi]}{\partial b} > 0$ for all $b > b^*_m$, so the optimal strategic booking limit, b^*_s, is smaller than b^*_m. □

Proposition 1.10. (Optimal Myopic Overbooking Policy)
A myopic airline sets $p^*_m = (1-\alpha)v$ *and* $c^*_m = 0$. *When* $v < r$, *it selects a finite optimal booking limit* $b^*_m = \max\left\{\ell'^{-1}\left(\frac{(1-\alpha)v}{r}\right), k\right\}$. *Otherwise,* b^*_m *is infinite.*

Proof. When the airline takes $\widehat{w}$ as an exogenous quantity, it believes that neither the price nor the fixed compensation affects the demand. Therefore, it charges the highest price, $p = (1-\alpha)v$, and offers the lowest compensation, $c = 0$.

When $p = (1-\alpha)v$ and $v \geq r$, by Definition 1.2 part (iii), $p - \ell'(b)\, r > p - (1-\alpha)\, r \geq 0$. Therefore the marginal change in profit is always positive and the airline is incentivized to overbook as much as possible.

When $v < r$, $p-(1-\alpha)r = (1-\alpha)v-(1-\alpha)r < 0$ and either $p-\ell'(k)r \leq 0$ or $p-\ell'(k)r > 0$. If $p - \ell'(k)r > 0$ or $\ell'^{-1}\left(\frac{(1-\alpha)v}{r}\right) > k$, then (1.19) has a solution $b^*_m = \ell'^{-1}\left(\frac{(1-\alpha)v}{r}\right)$ by the Intermediate Value Theorem. Furthermore, by Definition 1.2 part (i), $\ell'(b)$ increases monotonically in b for $b \geq k$; hence the solution is unique. When $p - \ell'(k)r \leq 0$ or $\ell'^{-1}\left(\frac{(1-\alpha)v}{r}\right) \leq k$, the monotonicity of $\ell'(b)$ ensures that the marginal increase in profit is always negative and hence the airline does not overbook, i.e., $b^*_m = k$. □

Lemma 1.11. (Boundary Equilibria Not Optimal)

Any optimal strategic overbooking policy induces a customer equilibrium with $U(\beta, \widehat{w}) = 0$.

Proof. The only equilibria for which $U(\beta(\widehat{w}), \widehat{w}) \neq 0$ are $U(\beta(\widehat{w}), \underline{w}) \geq 0$ and $U(\beta(\widehat{w}), \overline{w}) \leq 0$, and it suffices to show that any strategic overbooking policy that leads to either $U(\beta(\widehat{w}), \underline{w}) < 0$ or $U(\beta(\widehat{w}), \overline{w}) > 0$ is not optimal.

First consider the policy that induces $U(\beta(\widehat{w}), \underline{w}) < 0$. In this case, $U(\beta(\widehat{w}), w) < 0$ for all w because $U(\beta(\widehat{w}), w)$ decreases strictly in w. Therefore, the equilibrium marginal customer's response is $\widehat{w} = \underline{w}$, and the airline thus obtains zero demand and zero profit. The overbooking strategy that results in $U(\beta(\widehat{w}), \underline{w}) < 0$ is strictly dominated by any policy that charges a positive price and does not overbook.

Next consider the policy that induces $U(\beta(\widehat{w}), \overline{w}) > 0$. Here we know that $U(\beta(\widehat{w}), w) > 0$ for all w and $\widehat{w} = \overline{w}$. Since any admissible policy has $c \leq \overline{w}$, in this case we must have $p < (1-\alpha)v$. Therefore, the airline can raise the price without affecting demand and ultimately increase profits. Thus, the original policy with $p < (1-\alpha)v$ cannot be optimal. □

Lemma 1.12. (Multiple Equivalent Policies)

For any admissible policy (p, b, c) *for which* $\beta > 0$ *and* $U(\beta, \widehat{w}) = 0$*, there exists an infinite set of alternative policies with the same booking limit,* $b' \equiv b$*, and alternative price and bumping compensation,*

$$p' \in [\max\{0, (1-\alpha)(v - \widehat{w}\beta)\}\ , (1-\alpha)v] \qquad \text{and} \qquad c' = \left(\widehat{w} - \frac{v}{\beta}\right) + \left(\frac{p'}{(1-\alpha)\beta}\right), \tag{1.25}$$

with the same equilibrium $(\beta, \widehat{w})$ *and expected profits* $\mathsf{E}[\Pi(p, b, c)] = \mathsf{E}[\Pi(p', c', b')] = \mathsf{E}[\Pi((1-\alpha)v, b, \widehat{w})]$.

Proof. We complete the proof in three steps. First, recall that the original policy (p, b, c) induces the equilibrium $(\beta, \widehat{w})$ and that, given fixed b and $\widehat{w}$, the expression for β in (A.9) is independent of p and c. As long as b does not change, we need only show β to be consistent with p, c, and $\widehat{w}$ through the equilibrium equation (1.2).

Second, we use (1.2) to find consistent (p', c') pairs. Specifically, from (1.2) we have $U(\beta, \widehat{w}) = -p' + (1-\alpha)v + (1-\alpha)\beta(c' - \widehat{w}) = 0$. Given $b' \equiv b$, $\beta > 0$, and some admissible p', we can solve (1.2) for c', to derive the definition of c' in (1.25). Equation (1.25)'s bounds on p' then follow from the definition of an admissible policy. Given $c' \leq \overline{w}$, we can set $c' = \overline{w}$ and solve (1.2) for p' to see that $p' \leq (1-\alpha)v + (1-\alpha)\beta(\overline{w} - \widehat{w})$, which is looser than the direct bound, $p' \leq (1-\alpha)v$. Similarly, given $c' \geq 0$, we can set $c' = 0$ and solve (1.2) for p' to see that $p' \geq (1-\alpha)v + (1-\alpha)\beta(0 - \widehat{w}) = (1-\alpha)(v - \beta\widehat{w})$. While the lower bound may be larger or smaller than the direct bound, $0 \leq p'$, it is always (weakly) lower than the upper bound, $p' \leq (1-\alpha)v$. Thus, we have $\max\{0, (1-\alpha)(v - \beta\widehat{w})\} \leq p' \leq (1-\alpha)v$.

Third we show that, for any (p', c', b') that is consistent with (p, b, c), the airline earns the same expected profits. To demonstrate this fact, we recall the definition of expected profits in (1.13), $\mathsf{E}[\Pi(p', b', c')] = [p' - (1-\alpha)\beta(c' + r)]\,\mathsf{E}[S]$. From (1.3) we know that, given b and $\widehat{w}$, $\mathsf{E}[S]$ is independent of p' and c'. Furthermore, we can use the definition of c' in (1.25) to substitute out c' in the definition of margin per customer to show that

$$[p' - (1-\alpha)\beta(c'+r)] = [(1-\alpha)v - (1-\alpha)\beta(\widehat{w}+r)].$$

Thus any (p', c', b') that is consistent with (p, b, c) yields the same expected profit as well. □

Proposition 1.13. (Problem Reduction)

If there exists an optimal strategic overbooking policy $\xi \in \Xi$, then there exists an optimal strategic policy that sets $p_s^* = (1-\alpha)v$, induces an interior equilibrium $U(\beta, \widehat{w})$, and optimizes (1.10).

Proof. For any optimal policy (p_s^*, c_s^*, b_s^*) with $\beta > 0$, Lemma 1.12 implies that there exists a policy with $p = (1-\alpha)v$ that generates the same equilibrium and expected profit, and the result follows. Now suppose there exists an optimal policy with $\beta = 0$. Then from (1.2) and Lemma 1.11, we see that $U(\beta, \widehat{w}) = -p + (1-\alpha)v + 0(c - \widehat{w}) = 0$. Thus, $p = (1-\alpha)v$ here as well. □

Proposition 1.14. (Booking Limit for Optimal Strategic Overbooking Policy)

(i) If $c_s^* \leq v - r$, then $b_s^* = \infty$.

(ii) If $v - r < c_s^* < \frac{(1-\alpha)v}{\ell'(k)} - r$, then $b_s^* = \ell'^{-1}\left(\frac{(1-\alpha)v}{c_s^*+r}\right)$.

(iii) If $c_s^* \geq \frac{(1-\alpha)v}{\ell'(k)} - r$, then $b_s^* = k$.

Proof. The proof of the proposition can be found in the main text. □

Proposition 1.15. (Properties of the Auction with No Cap)

Suppose that, when $n > k$ customers show up for a flight, the airline runs a reverse, uniform price, multi-unit auction. Then we have the following.

(i) Customers' optimal bids match their underlying hassle costs: $\{\varpi_{1:n} = w_{1:n}, \ldots, \varpi_{n:n} = w_{n:n}\}$.

(ii) All customers are willing to purchase tickets, irrespective of their hassle cost $w \in [\underline{w}, \overline{w}]$.

(iii) The airline's optimal price is $p_a^* = (1-\alpha)v$.

Proof.

Part (i). We will show that, in the k^{th}-price reverse auction problem outlined above, bidding the true hassle cost is a dominant strategy for customers. Consider a customer with hassle cost w. Assume that other customers bid in some arbitrary way.

1. Suppose the customer can board if she bids her true hassle cost w, i.e., $w \geq w_{n-k+1:n}$. Then bidding higher than w still allows her to board whereas bidding lower than w may result in bumping with compensation $w_{n-k:n}$. Since $w_{n-k:n} < w$, bidding lower than w is dominated by bidding w.

2. Suppose the customer is bumped if she bids her true hassle cost w, i.e., $w < w_{n-k+1:n}$. In this case, she gets positive utility $w_{n-k+1:n} - w$. If she were to bid lower than w, she would still be bumped and receive $w_{n-k+1:n}$, and this does not improve her utility. If she bids $w' > w$, then one of the following two cases holds. If $w < w' \leq w_{n-k+1:n}$, then she is still bumped and receive $w_{n-k+1:n}$. If she bids $w' > w_{n-k+1:n}$, then she isn't bumped. In this case she ends up receiving no compensation and is strictly worse off because $0 < w_{n-k+1:n} - w$. Therefore, bidding higher than w is dominated by bidding w.

Considering both cases, we see that bidding w is a dominant strategy for the customer.

Part (ii). By construction of the k^{th}-price auction, we know that every bumped customer is more than fairly compensated: $w_{n-k+1:n} \geq w_{i:n}$ for $i \in \{1, 2, ..., n-k\}$. Therefore, all customers receive non-negative utility from buying a ticket under any admissible policy, and $\widehat{w} = \overline{w}$.

Part (iii). From part (ii), we know that demand is not thinned by the airline's profit-maximizing overbooking strategy since $G(\widehat{w}) = G(\overline{w}) = 1$. Therefore, the airline should maximize profit per customer and set $p_a^* = (1-\alpha)v$. $\square$

Proposition 1.16. (Properties of the Auction with a Cap)
Suppose that the airline sets the price $p = (1-\alpha)v$ and $b > k$. When $n > k$ customers show up for a flight, it runs a reverse, uniform price, multi-unit auction with compensation cap $c_a \leq \overline{w}$. Then we have the following.
(i) Customers are willing to purchase tickets, if and only if their hassle costs are $w \leq c_a$.
(ii) Customers' optimal bids match their underlying hassle costs: $\{\varpi_{1:n} = w_{1:n}, \ldots, \varpi_{n:n} = w_{n:n}\}$.

Proof.

Part (i). It is easy to see that customers with hassle costs $w \leq c_a$ are always more than fairly compensated if bumped. Since $p = (1-\alpha)v$, buying a ticket always gives them non-negative utility. For a customer with hassle cost $w > c_a$, her net value of being bumped is $c_a - w < 0$. As long as her probability of being bumped is positive, her expected value from purchasing a ticket is negative. Therefore, buying a ticket is dominated by taking the outside option, which gives her zero utility.

Part (ii). Since $c_a \leq \overline{w}$, we have $\overline{w} = c_a = \widehat{w}$ and can apply the proof of Proposition 1.15 part (i). $\square$

Proposition 1.17. (Auction with Cap Dominates Fixed Compensation)
Given any fixed-compensation policy with $p = (1-\alpha)v$, $b > k$, $\underline{w} < c \leq \overline{w}$, and equilibrium $\beta > 0$, an auction-based policy with the same price, $p = (1-\alpha)v$, the same booking limit, b, and an analogous cap, $c_a = c$, earns strictly higher expected profits: $\mathsf{E}[\Pi_a((1-\alpha)v, b, c)] > \mathsf{E}[\Pi((1-\alpha)v, b, c)]$.

Proof. Consider a fixed compensation scheme with price $p = (1-\alpha)v$, booking limit $b > k$, and compensation $\underline{w} \le c \le \overline{w}$ that induces expected profit $\mathsf{E}[\Pi(p,b,c)]$. We define the analogous capped auction with the same price and booking limit, the analogous cap $c_a \equiv c$, and expected profit $\mathsf{E}[\Pi_a(p,b,c)]$.

Observe that demand $S = \min\{b, QG(c)\}$ is the same in both cases. In the fixed compensation scheme with $p = (1-\alpha)v$, (1.25) shows that $\widehat{w} \equiv c$. In the capped auction scheme, Proposition 1.16 part (i) similarly implies that the effective $\overline{w} \equiv c_a$. Because both schemes also have identical price, their expected revenues are the same.

Similarly, both schemes have identical numbers of bumped customers, $(N(s,\alpha)-k)^+$, and rerouting costs per customer. Thus, expected rerouting costs are also the same in both cases.

Finally, while the number of bumped passengers is, again, the same in both cases, we can show that the expected per-customer bumping compensation is strictly lower in the auction scheme. In particular, given continuous W, $\mathsf{E}[w_{n-k:n}] < c$ with probability 1. Hence, $\mathsf{E}[\Pi_a(p,b,c)] > \mathsf{E}[\Pi(p,b,c)]$. □

Lemma 1.18. (Convexity of Auction-Based Expected Bumping Cost)
Suppose $N(s,\alpha)$ is SICX in s.
(i) If $\partial \mathsf{E}[w(c_a)_{n-k+1:n}]/\partial n \ge 0$, then $C'(s) > 0$.
(ii) If in addition $\partial^2 \mathsf{E}[w(c_a)_{n-k+1:n}]/\partial n^2 > 0$, then $C''(s) > 0$.

Proof. For $s \ge k$, $C(s,c_a) = \int_k^s (n-k)]\,\mathsf{E}[w(c_a)_{n-k+1:n}]\,p_N(n|s)\,dn$. Let $\psi(x) = (x-k)^+\mathsf{E}[w(c_a)_{x-k+1:x}]$. Then $C(s,c_a) = \mathsf{E}[\psi(N(s,\alpha))]$. Since $N(s,\alpha)$ is SICX in s, to show that $C(s,c_a)$ is increasing convex in s, it suffices to show that $\psi(x)$ is increasing convex in x.

Part (i). Since $(x-k)^+$ is strictly increasing in x for $s \ge k$, if $\mathsf{E}[w(c_a)_{x-k+1:x}]$ is increasing in x, i.e., $\frac{\partial \mathsf{E}[w(c_a)_{x-k+1:x}]}{\partial x} \geq 0$, $\psi(x)$ must be increasing in x.

Part (ii). For $s \geq k$, $\psi'(x) = (x-k)\frac{\partial \mathsf{E}[w(c_a)_{x-k+1:x}]}{\partial x} + \mathsf{E}[w(c_a)_{x-k+1:x}]$ and $\psi''(x) = (x-k)\frac{\partial^2 \mathsf{E}[w(c_a)_{x-k+1:x}]}{\partial x^2} + 2\frac{\partial \mathsf{E}[w(c_a)_{x-k+1:x}]}{\partial x}$. Therefore, one sufficient condition for $\psi''(x) > 0$ is that $\frac{\partial \mathsf{E}[w(c_a)_{x-k+1:x}]}{\partial x} \geq 0$ and $\frac{\partial^2 \mathsf{E}[w(c_a)_{x-k+1:x}]}{\partial x^2} > 0$. □

Lemma 1.19. (Convexity of the Approximation $\tilde{\mathbf{G}}$)

Suppose $N(s,\alpha)$ is SICX in s and we use the specific approximation $\mathsf{E}[w(c_a)_{n-k+1:n}] \approx \tilde{G}(c_a, n)$.

If for any $c_a \in (\underline{w}, \overline{w}]$, $g'(w) < 0$ for all $w \in [G^{-1}(\frac{G(c_a)}{k+1}), c_a]$, then $C''(s) \geq 0$.

Proof. For $s \geq k$, $C(s, c_a) = \int_k^s (n-k)]\, \mathsf{E}[w(c_a)_{n-k+1:n}]\, p_N(n|s)\, dn$. Let $\psi(x) = (x-k)^+ \mathsf{E}[w(c_a)_{x-k+1:x}]$. Then $C(s, c_a) = \mathsf{E}[\psi(N(s,\alpha))]$. If $N(s,\alpha)$ is stochastically increasing and linear (e.g. binomial), then to show that $C(s, c_a)$ is convex in s, it suffices to show that $\psi(x)$ is convex in x.

Let $y(x) = \frac{x-k+1}{x+1}$. Then, $\tilde{G}(c_a, x) = \tilde{G}(x) = G_{c_a}^{-1}(y(x))$. Here we drop c_a in the argument because we consider a fixed c_a. Thus,

$$\tilde{G}'(x) = \frac{k}{g_{c_a}(G_{c_a}^{-1}(y(x)))(x+1)^2} \tag{A.23}$$

and

$$\tilde{G}''(x) = -\frac{g'_{c_a}(G_{c_a}^{-1}(y(x)))}{g_{c_a}(G_{c_a}^{-1}(y(x)))}(\tilde{G}'(x))^2 - \frac{2\tilde{G}'(x)}{x+1}. \tag{A.24}$$

By (A.23) and (A.24),

$$\psi''(x) = -\frac{g'_{c_a}(G_{c_a}^{-1}(y(x)))}{g_{c_a}(G_{c_a}^{-1}(y(x)))}(\tilde{G}'(x))^2(x-k) + \frac{2(k+1)\tilde{G}'(x)}{x+1}. \tag{A.25}$$

By (A.25), a sufficient condition for $\psi''(x) \geq 0$ is $g'_{c_a}(G_{c_a}^{-1}(y(x))) \leq 0$, which is equivalent to $g'(G_{c_a}^{-1}(y(x))) \leq 0$ since $g_{c_a}(x) = \frac{g(x)}{G(x_a)}$. Note that for $x \in (k, \infty)$, we have $y(x) \in (\frac{1}{k+1}, 1)$. Therefore, $G_{c_a}^{-1}(y(x)) \in (G_{c_a}^{-1}(\frac{1}{k+1}), c_a)$. Clearly, $G_{c_a}^{-1}(\frac{1}{k+1}) = G^{-1}(\frac{G(c_a)}{k+1})$. Hence, the sufficient condition becomes $g'(w) \leq 0$ for $w \in [G^{-1}(\frac{G(c_a)}{k+1}), c_a]$. $\square$

Proposition 1.20. (Optimal Booking Limit for the Auction)

Suppose $p = (1-\alpha)v$ and $C''(s) > 0$. Then there exists a unique optimal booking limit, b_a^*, with the following properties.

(i) If $\ell'(k) \geq \frac{(1-\alpha)v}{r}$ then $b_a^* = k$.

(ii) If $\ell'(k) < \frac{(1-\alpha)v}{r}$ and $\exists\, b \in (k, \infty)$ s.t. $C'(b) \geq (v-r)(1-\alpha)$, then (1.36) determines $b_a^* \in (k, \infty)$.

(iii) If $C'(b) < (v-r)(1-\alpha)$ for all $b \geq k$ then $b_a^* = \infty$.

Proof. By the first-order condition (1.36), an interior b_a^* satisfies $\ell'(b) = \frac{p - C'(b)}{r}$. Recall from Definition 1.2 part (i) that $\ell''(b) > 0$ and from Lemma 1.18 part (ii) that $C''(b) > 0$. Given the convexity of these costs, expected profits are concave in b, and at most one such b_a^* exists. By the second-order condition (1.37), $\frac{d^2\Pi_a}{db^2}\Big|_{b=b_a^*} = 0 + \mathsf{P}\{Q > b\}\,[-r\ell''(b) - C''(b)] < 0$. Therefore, an interior b_a^*, if it exists, is a global maximum.

The airline charges price $p = (1-\alpha)v$, and we prove the results in the order (i), (iii), (ii).

Part (i). $\ell'(k) \geq \frac{(1-\alpha)v}{r}$ implies $\ell'(k) \geq \frac{p}{r} = \frac{p - C'(k)}{r}$. Therefore, $\frac{d\Pi_a}{db} = p - r\ell'(b) - C'(b) \leq p - r\ell'(k) - C'(k) \leq 0$ for all $b \geq k$, which implies $b_a^* = k$.

Part (iii). Together, $p = (1-\alpha)v$ and $C'(b) < (v-r)(1-\alpha)\ \forall\, b \geq k$ imply that $\ell'(b) \leq 1 - \alpha < \frac{p - C'(b)}{r}$ and $\frac{d\Pi_a}{db} > 0\ \forall\, b$, which in turn means that $b_a^* = \infty$.

Part (ii). From parts (i) and (iii) and the Intermediate Value Theorem, the necessary and sufficient conditions for the existence of an interior b_a^* are a) $\ell'(k) < \frac{(1-\alpha)v}{r}$ and b) $\exists\, b > k$ s.t. $C'(b) \geq (v-r)(1-\alpha)$. $\square$

Lemma 1.21. (Bumping Compensation Grows with the Cap)

For $c_a \in (\underline{w}, \overline{w})$ and $s > k$, (i) $\partial C(s, c_a)/\partial c_a > 0$, and (ii) $\partial^2 C(s, c_a)/\partial s \partial c_a > 0$.

Proof. Since $\mathsf{E}[w(c_a)_{n-k+1:n}]$ increases in c_a, by (1.33), $C(s, c_a)$ increases in c_a for any fixed s. Then from (1.33) we have

$$\begin{aligned}\frac{\partial}{\partial s}C(s,c_a) &= (s-k)\mathsf{E}[w(c_a)_{s-k+1:s}]P_N(s|s) \\ &\quad + \int_k^s (n-k)\mathsf{E}[w(c_a)_{n-k+1:n}]\frac{dP_N(n|s)}{ds}\,dn. \end{aligned} \tag{A.26}$$

Since both $\mathsf{E}[w(c_a)_{s-k+1:s}]$ and $\mathsf{E}[w(c_a)_{n-k+1:n}]$ increase in c_a, by (A.26), for any fixed s we have $\frac{\partial^2}{\partial s \partial c_a}C(s, c_a) > 0$. □

Proposition 1.22. (Optimal Auction Parameters)

For fixed p, let $b_a^*(c_a)$ be the optimal booking limit induced by c_a. Suppose $p = (1-\alpha)v$, $\underline{w} < c_a < \overline{w}$, $k < b_a^*(c_a) < \infty$ and $C''(s) > 0$. Then we have the following.

(i) The optimal booking limit, b_a^*, is decreasing in c_a.

(ii) The resulting expected profit, $\mathsf{E}[\Pi_a]$, is increasing in c_a.

Proof.

Part (i). By Proposition 1.16 part (i), the equilibrium customers' response is independent of the booking limit. Therefore, as in (1.34),

$$\begin{aligned}\frac{d}{db}\Pi_a(b) &= p\,\frac{d\mathsf{E}[S]}{db} - r\,\frac{d\mathsf{E}[(N-k)^+]}{db} - \frac{d\mathsf{E}[C(S)]}{db} \\ &= p\,\mathsf{P}\{Q>b\} - r\,\ell'(b)\mathsf{P}\{Q>b\} - C'(b)\mathsf{P}\{Q>b\}, \end{aligned} \tag{A.27}$$

and the optimal booking limit with c_a, $b_a^*(c_a)$, satisfies

$$p - r\,\ell'(b_a^*(c_a)) - C'(b_a^*(c_a)) = 0. \tag{A.28}$$

Consider $c_a' > c_a$. Then by Lemma 1.21 part (ii), $C'(s, c_a')(b_a^*(c_a)) > C'(s, c_a)(b_a^*(c_a))$. Since $C'' > 0$, $\ell'' > 0$ and $b_a^*(c_a)$ is interior, we must have $b_a^*(c_a') < b_a^*(c_a)$.

Part (ii). Recall that, for a capped auction, $\widehat{w} \equiv c_a$. Suppose the airline always implements the optimal booking limit $b_a^*(c_a)$ w.r.t. c_a. Then by the Envelope Theorem,

$$\begin{aligned}
\frac{d}{d\,c_a}\mathsf{E}[\Pi_a(b,\widehat{w})]\bigg|_{b=b_a^*(c_a)} &= \frac{d}{d\widehat{w}}\mathsf{E}[\Pi_a(b,\widehat{w})]\bigg|_{b=b_a^*(c_a)} \\
&= \frac{\partial}{\partial\widehat{w}}\Pi_a(b,\widehat{w}) \\
&= p\,\frac{\partial\mathsf{E}[S]}{\partial\widehat{w}} - r\,\frac{\partial\mathsf{E}[(N-k)^+]}{\partial\widehat{w}} - \frac{\partial\mathsf{E}[C(S)]}{\partial\widehat{w}} \\
&= g(\widehat{w})\int_0^{\frac{b_a^*(c_a)}{G(\widehat{w})}} \left[p - r\ell'(qG(\widehat{w})) - C'(qG(\widehat{w}))\right] qf(q)dq
\end{aligned} \tag{A.29}$$

Since $p-r\ell'(qG(\widehat{w}))-C'(qG(\widehat{w})) = 0$ for $q = \frac{b_a^*(c_a)}{G(\widehat{w})}$, and since $\ell'(s), C'(s) > 0$, we know that $p - r\ell'(qG(\widehat{w})) - C'(qG(\widehat{w})) \geq 0$ for $q \in [0, \frac{b_a^*}{G(\widehat{w})}]$. This shows that $\frac{d}{d\,c_a}\Pi_a(b,\widehat{w})|_{b=b_a^*} > 0$, which concludes the proof. □

Proposition 1.23. (Optimality of Overbooking Policy)

(i) The optimal overbooking policy uses an auction to determine customers' bumping compensation.

When $N(s,\alpha)$ is SICX, $C''(s) > 0$, and the auction-based overbooking policy sets $p = (1-\alpha)v$, we also have the following.

(ii) The optimal cap on bumping compensation is effectively unbounded: $c_a^* = \overline{w}$.

(iii)The optimal booking limit b_a^* is defined as in Proposition 1.20.

Proof. Part (i) follows immediately from Proposition 1.17. Part (ii) follows immediately from Proposition 1.22 part (ii). For Part (iii) See the proof of Proposition 1.20. □

A.1.3. Numerical Experiments

The results reported in Table 1 are based on the optimal booking limits found for the following primitive parameters.

$p = 400$	All other parameters are pegged off the ticket price.
$c \in \{0, 100, 200, 400, 800\}$	Bumping compensation ranges from 0 to 2 times the ticket price.
$\alpha \in \{0.05, 0.1, 0.2\}$	No-show probabilities range from low to high.
$k \in \{50, 100, 200, 400\}$	The plane's capacity ranges from low to high.
$r \in \{0, 200, 320, 400\}$	Rerouting costs runs from 0 to the ticket price.
$v\,(1-\alpha) - 400 \in \{0.01, 1, 4\}$	Value set to ensure a small amount of consumer surplus.
$F \sim \mathcal{N}(1.2\,k, k/3)$	Support is $[1, 2.4\,k]$; distribution renormalized so probabilities sum to one.
$G \sim \mathcal{N}(v, v/3)$,	Support is $[0, 2\,2\,v]$; distribution renormalized so probabilities sum to one.

Notes: (1) The range of customers' expected values of flying $v\,(1-\alpha) - 400 \in \{0.01, 1, 4\}$ is low and reflects the fact that larger v's generate enough customer surplus that booking limits become unbounded. This numerical result also suggests that the price should be roughly $p \approx (1-\alpha)v$. (2) The demand distribution F is scaled to offer slightly more demand than capacity available. (3) The hassle-cost distribution G is scaled to be on the order of the value the customer receives from flying.

The results reported in Table 2 are based on the optimal booking limits and bumping compensation found for the following primitive parameters.

$v \in \{200, 400, 500, 600, 800\}$	Value of flying ranges from low to high.
$\alpha \in \{0.05, 0.1, 0.2\}$	No-show probabilities range from low to high.
$p = (1-\alpha)v$	Optimal prices for both the fixed and uncapped-auction schemes.
$k \in \{50, 100, 200, 400\}$	The plane's capacity ranges from low to high.
$r \in \{0, 200, 400, 600, 800, 1000\}$	Rerouting costs run from 0 to very high.
$F \sim \mathcal{N}(1.2\,k, k/3)$	Support $[1, 2.4\,k]$; distribution renormalized so probabilities sum to one.
$G \sim \mathcal{N}(v, v/3)$,	Support $[0, 2\,v]$; distribution renormalized so probabilities sum to one.

Notes: (1) The demand distribution F is scaled to offer slightly more demand than capacity available. (2) The hassle-cost distribution is scaled to be on the order of the value the customer receives from flying. (3) Hassles costs are normally distributed, and to evaluate auction-based policies we approximate expected order statistics using results from Harter (1961), which includes correction terms for the fractile approach of Arnold et al. (2008) we describe in 1.5.4.

A.2. Capped ICO

A.2.1. Additional Discussions and Results

Utility Tokens and the Token Buyers

In this section, we elaborate on two important features of tokens and the role of the token buyers (the speculators and the customers).

First, tokens play a dual role: as of today, most tokens in the market have been considered as both utility and security.[1] The "security" aspect results from the tradable feature of the tokens. The "utility" aspect comes from the fact that the fundamental value of these tokens lies in the economic value of the products or services that they are redeemable for. However, most projects do not have any products at the time of the ICO. In 2017, for instance, 87% of ICOs did not yet have a running product (CryptoGlobe, 2018). To capture these features, we model tokens that start out as pure securities and only after product launch become utility tokens. Such tokens appeal to two groups of token buyers: those who see tokens as securities purchase the tokens in the ICO period (before product launch),[2] whereas those who wish to consume the products buy tokens in the market period (after product launch). Therefore, we refer to the token buyers in the ICO period and those in the market period as *speculators* and *customers* respectively.

[1] The regulatory environment is still uncertain but efforts are being made to pass bills that would distinguish tokens from securities like stocks (Khatri, 2019).

[2] Technically, those who see tokens as securities may purchase tokens whenever they feel optimistic about the potential return. However, we model a firm that plans one round of production and product sale and the market token price in the market period is an equilibrium quantity that does not change during that period. Therefore, it only makes sense for this group of token buyers to come in the ICO period.

The second feature is that the tokens issued by the firm can only be redeemed on the firm's own platform and are the only viable method of payment for the its products. By restricting the method of payment, the firm ties the value of the tokens to the economic value of the products. This, together with the existence of a secondary market to trade the tokens, incentivize speculators to purchase tokens in the ICO, even if they are not interested in subsequently consuming the product themselves.

At the same time, the fact that the tokens have no use on other platforms has a few implications. First, it means that the token value solely depends on the consumption of products of this particular platform. Second, after the firm ends production, the speculators have no reason to hold the tokens and the customers do not buy more tokens than needed. Third, redeemed tokens retain no value if no further production is planned. Last, since we only consider one round of production, this suggests that the tokens are for one-time use only and the firm cannot resell the redeemed tokens for more cash.

Example: Honeypod Whitepaper

Honeypod (Honeypod, 2018) aims to produce a hardware that serves as the main hub that interconnects various devices and provides traffic filtering. The company claims that they have mature products that are ready for mass production before token crowdsale.

Parameters captured by our model include

1. Hard cap ($m = 200,000,000$).
2. ICO sales cap/soft cap ($n = 40,000,000$).
3. Fixed token price of during public token sale ($\tau = \$0.05$).
4. Customers' willingness to pay ($v = \$99$).
5. Manufacturing cost ($c = \$32$).
6. Production quantity over 12 months ($Q = 50,000$).

Parameters not captured by our model include

1. Four tiers of fixed token prices during private token sale ($\$0.02, \$0.025, \$0.03, \$0.035$).

2. Other use of funds from the token sale (e.g. 25% on maintenance, R&D).

Parameters in our model that are not mentioned in the white paper include

1. Aggregate demand (D).

A.2.2. Technical Results

Optimal Token Price and ICO Cap

Given the optimal production quantity and speculators' equilibrium behavior, we now examine how the firm sets the profit-maximizing ICO token cap n^* and initial token price τ^*.

From Lemma 2.3, the number of speculators $s^*(\tau, n) \leq m\left(1 - \frac{c}{v}\right)$, and given speculators participating in the ICO buy 1 token each, we need not consider the case in which tokens $n > m\left(1 - \frac{c}{v}\right)$. We will first find the token price $\tau^*(n)$ for a given token cap $n \leq m\left(1 - \frac{c}{v}\right)$ and then maximize profit over the token cap n.

For a fixed n, the platform's optimization problem (2.3) can be written as a maximization problem over τ subject to speculators' participation constraint. In particular, the optimization problem is

$$\max_{\tau \geq 0} \Pi = \tau(n)\, s^*(\tau, n) - c\, Q^*(s^*(\tau, n)) + (m - s^*(\tau, n))\, \mathsf{E}[\tau_{eq}(s^*(\tau, n))], \tag{A.30}$$

subject to $u(s^*(\tau, n)) \geq 0$ and $Q^*(s^*(\tau, n)) = \min\left\{F^{-1}\left(1 - \frac{cm}{(m - s^*(\tau,n))v}\right), \frac{\tau\, s^*(\tau,n)}{c}\right\}$ (from Proposition 2.2). Proposition 2.4 (see Section 2.3.3) guarantees the existence of a nonzero equilibrium token price τ^*.

Next we find the optimal ICO token price $\tau^*(n)$ assuming the two conditions in Proposition 2.4 are met. Before stating the proposition, we impose an additional technical condition on the demand distribution to guarantee equilibrium uniqueness[3]: $\frac{f'(F^{-1}(y)}{(f(F^{-1}(y))^2} > -\frac{3v}{c} \cdot k \cdot \frac{1-k}{2k-1}$

[3] One can readily check analytically or numerically) that this sufficient condition is generally satisfied for some common distributions such as uniform and normal. All numerical results presented in the paper satisfy

where $k = 1 - \frac{c}{(1-y)v}$ and $y \in [0, 1 - \frac{2c}{v})$.

Proposition A.2. (Optimal ICO Token Price)

When $v > 2c$,

i) For a given $n \in (\frac{m\,c}{v},\, m\left(1 - \frac{c}{v}\right)]$, there exists a finite positive $\tau^*(n)$ uniquely determined by $u(s^*(\tau^*(n))) = 0$.

ii) There exists a unique $\hat{n} \in (\frac{m\,c}{v}, \frac{m}{2})$ such that

- *for $n \in \left[\frac{m\,c}{v}, \hat{n}\right)$, $\tau^*(n)$ is the unique solution of $\tau^*(n) = \frac{v}{m}\,\mathsf{E}\left[\min\left\{D, \frac{\tau^*(n)\,n}{c}\right\}\right]$;*
- *for $n \in \left[\hat{n}, m\left(1 - \frac{c}{v}\right)\right]$, $\tau^*(n) = \frac{v}{m}\,\mathsf{E}\left[\min\left\{D, F^{-1}\left(1 - \frac{c\,m}{(m-n)v}\right)\right\}\right]$.*

Part (i) of Proposition A.2 shows that when the price-cost ratio is high enough, for any fixed ICO cap n in the appropriate range as suggested by Proposition 2.4 (i), there exists a unique, positive and finite ICO token price $\tau^*(n)$ that maximizes (A.30) by extracting all utility from the speculators who participate strategically according to Lemma 2.3. By (2.1), this implies that the expected equilibrium token price in the market period is equal to the optimal ICO token price, i.e., $\mathsf{E}[\tau_{eq}(s(\tau^*(n), n)] = \tau^*(n)$. We then solve $u(s^*(\tau^*(n))) = 0$ using Lemma 2.1 (ii) and Proposition 2.2 (i) and obtain part (ii) of Proposition A.2. Recall that the term $\frac{\tau^*(n)\,n}{c}$ reflects the budget constraint and $F^{-1}\left(1 - \frac{c\,m}{(m-n)v}\right)$ is the constrained optimal production quantity. Therefore, part (ii) of Proposition A.2 suggests that the firm, upon setting the optimal ICO token price, spends all funds raised on production when the ICO cap n is small but produces an optimal quantity without using all the funds when n is large or $\frac{n}{m}$ is closer to the misconduct fraction.

Knowing $\tau^*(n)$, $s^*(\tau^*(n), n)$ and $Q^*(s^*(\tau^*(n), n))$, the firm's optimization problem reduces to a maximization problem over the ICO cap n given by

$$\begin{aligned}\max_{\frac{m\,c}{v} < n \le m\left(1 - \frac{c}{v}\right)} \Pi \;=\;& \tau^*(n)\,s^*(\tau^*(n), n) \;-\; c\,Q^*(s^*(\tau^*(n), n)) \\ & + (m - s^*(\tau^*(n), n))\,\mathsf{E}[\tau_{eq}(s^*(\tau^*(n), n))] \qquad \text{(A.31)}\end{aligned}$$

this condition.

where $s^*(\tau(n), n) = n$, $Q^*(s^*(\tau(n), n)) = \min\left\{F^{-1}(1 - \frac{cm}{(m-n)v}), \frac{\tau^*(n)\,n}{c}\right\}$ and $\tau^*(n)$ is given by Proposition A.2 part ii).

This leads to the following result.

Proposition A.3. (Equilibrium ICO Cap) *When $v > 2c$, the unique optimal ICO cap $n^* \in (\frac{m\,c}{v}, \frac{m}{2})$ equals the threshold $\hat{n}$ in Proposition A.2 ii), and is the solution to the following equation:*

$$\frac{v\,n^*}{c\,m}\,\mathsf{E}\left[\min\left\{D, F^{-1}(1 - \frac{c\,m}{(m-n^*)v})\right\}\right] = F^{-1}\left(1 - \frac{c\,m}{(m-n^*)v}\right).$$

Proposition A.3 tells us that neither a small ICO cap that suppresses the production quantity nor a large cap that induces idle cash is profit-maximizing for the firm. The optimal ICO cap n^* allows the firm to raise just enough funds that can be credibly committed to production, and here we provide a semi-closed-form solution of n^*.

Sequential Arrival of Speculators

In this section, we assume that the z speculators arrive sequentially during the ICO period and observe the number of tokens sold before their arrival, rather than showing up simultaneously. Tokens are sold on a first-come, first-served basis and each speculator buys either zero or one token based on the expected profit of their purchase. We will show that while this alternative assumption on the speculators' arrival changes one of the intermediate results, it leads to the same equilibrium results as in the rest of the paper.

Suppose the first s speculators will buy one token each. Then anyone who arrives later than the $s - th$ speculator will not buy any token and thus obtains zero utility. In this section, we focus on the earliest s arrivers. The expected profit of such a speculator given there s tokens will be sold by the end of the ICO is given by

$$u(s) = \Delta(s)\mathbb{1}_{\{s>0\}}, \tag{A.32}$$

where $\Delta(s)$, by (2.1), Lemma 2.1 and Proposition 2.2, is

$$\Delta(s) = \frac{v}{m}\,\mathsf{E}\left[\min\left\{D, F^{-1}\left(1 - \frac{cm}{(m-s)v}\right), \frac{\tau\, s}{c}\right\} \cdot \mathbb{1}_{\left\{s<m\left(1-\frac{c}{v}\right)\right\}}\right] - \tau. \tag{A.33}$$

The participation constraint requires that $u(s) \geq 0$. From (A.33) we immediately know that the equilibrium number of speculators will never be $m\left(1 - \frac{c}{v}\right)$ or beyond because $u(s) = -\tau < 0$ for $s \geq m\left(1 - \frac{c}{v}\right)$. Therefore, the speculators who arrive sequentially would collectively buy under the misconduct fraction.

Since $u(s)$ and $\Delta(s)$ have the same sign for $s > 0$, Lemma 2.3 part iii) still holds. Lemma 2.3 part iii) tells us that when the speculators arrive sequentially, there will be exactly $s_0(\tau)$ speculators without the sales cap n. However, note that $s_0(\tau)$ is not necessarily the utility-maximizing s because the early speculators cannot stop those who arrive later from buying more tokens unless it is no longer profitable to do so.

So far there are two upper bounds of the equilibrium number of speculators s^*: the sales cap, n, and $s_0(\tau)$[4]. We express s^* in terms of these two upper bounds in the following proposition.

Proposition A.4. (Equilibrium Number of Sequentially Arriving Speculators)
Given initial token price τ and the sales cap n, the equilibrium number of speculators is given by

$$s^*(\tau, n) = \min\{s_0(\tau), n\} \tag{A.34}$$

provided that $s_0(\tau)$ exists and

$$u(s^*) \geq 0. \tag{A.35}$$

If $s_0(\tau)$ does not exist or $u(\min\{s_0(\tau), n\}) < 0$, then there will be no speculators and thus ICO fails.

[4] In this section, we assume that $s_0(\tau)$ exists because its existence is necessary for $u(s^*) \geq 0$ for some $s > 0$. We show that the existence of $s_0(\tau)$ depends on both τ and n and discuss the conditions (the critical mass condition and a high willingness-to-pay) in Section 2.3.3.

Note that the expression of $u(s)$ with simultaneous arrivals given by (2.1) and that with sequential arrivals given by (A.32) have the same sign, albeit differing by a scale of s/z for $s > 0$. Since the magnitude of the speculators' profit does not affect their purchase decision or the firm's profit, Propositions 2.4 - 2.6 and Proposition 2.5 (iii) hold for both arrival assumptions. Details can be found in Appendix A.2.3.

Moreover, following Proposition 2.5, we can show that setting a sales cap is not needed when the customers observe their arrival sequence.

Corollary A.5. *When $v > 2c$, in equilibrium we have $n^* = s_0(\tau^*(n^*))$.*

By Corollary A.5, the optimal ICO sales cap is equal to the equilibrium number of speculators who would participate even when the cap is unannounced. Therefore, to reach the target level of token sales n^* that eventually induces maximum expected profit, it suffices to set the ICO token price to be $\tau^*(n^*)$.

A.2.3. Proofs

Proof of Lemma 2.1

i) First note that the customers have a fixed willingness-to-pay v that is equal to $p \cdot \tau_{eq}$. Suppose $p > m/\min\{Q, D\}$, then the demand of tokens $p \cdot \min\{Q, D\}$ exceeds the supply of tokens, m. This will drive the price of the token up, resulting in a decrease in the token-denominated price. In other words, τ_{eq} will increase and p will decrease. Similarly, if $p < m/\min\{Q, D\}$, then the demand of tokens is less than the supply of tokens, which induces an increase in p. Therefore, in equilibrium, demand of tokens is equal to its supply, i.e., $p \cdot \min\{Q, D\} = m$.

ii) The result follows immediately from $\tau_{eq} = v/p$ and Part (i).

□

Proof of Proposition 2.2

Taking derivative with respect to Q and applying Lemma 2.1,

$$\begin{aligned}\frac{d\Pi}{dQ} &= -c + (m-s)\,\frac{d}{dQ}\frac{v\mathsf{E}[\min\{Q,D\}]}{m} \\ &= -c + (m-s)\,\frac{v}{m}(1-F(Q)) \\ &= [(m-s)\,\frac{v}{m} - c] - (m-s)\,\frac{v}{m}F(Q) \end{aligned} \tag{A.36}$$

By (A.36), $\frac{d\Pi}{dQ} < 0$ when $(m-s)\,\frac{v}{m} - c < 0$, i.e., $s > m\left(1-\frac{c}{v}\right)$. On the other hand, when $s \leq m\left(1-\frac{c}{v}\right)$, ignoring the budget constraint and setting $\frac{d\Pi}{dQ} = 0$, we get $Q^*_{unconstrained}(s) = F^{-1}(1-\frac{cm}{(m-s)v})$. Since $\frac{d^2\Pi}{dQ^2} = -(m-s)\,\frac{v}{m}f(Q) < 0$, the profit function is concave in Q and $Q^*_{unconstrained}$ is a maximum. Hence the firm's optimal production quantity is given by

$$Q^*(s) = \min\left\{F^{-1}(1-\frac{cm}{(m-s)v}), \frac{\tau\,s}{c}\right\}\cdot \mathbb{1}_{\{s\leq m\left(1-\frac{c}{v}\right)\}}. \tag{A.37}$$

□

Proof of Proposition 2.4

i) For an ICO to succeed, there must be a positive number of speculators who invest. Therefore, the firm needs to set a (τ, n) pair that satisfies the speculators' participation constraint. Consider a fixed $n > 0$. A necessary condition for this n to induce a successful ICO is that there exists $\tau > 0$ such that $s^*(\tau, n) > 0$ and $u(s^*(\tau, n)) \geq 0$, which is a necessary condition for the existence of $s_0(\tau)$. Therefore, we will characterize such n while assuming the existence of $s_0(\tau)$.

   Now, for the fixed $n > 0$, we divide the space of possible τ into two partitions, $T_1 = \{\tau \geq 0 : s_0(\tau) < n\}$ and $T_2 = \{\tau \geq 0 : s_0(\tau) \geq n\}$, and in each partition look for eligible $\tau > 0$, i.e., $s^*(\tau, n) > 0$ and $u(s^*(\tau, n)) \geq 0$.

   (*Simultaneous,* T_1) When $n > s_0(\tau)$, with simultaneous arrival $s^* = 0$. Therefore, there is no eligible $\tau > 0$ in T_1.

   (*Simultaneous,* T_2) Now we consider T_2 where $0 < n \leq s_0(\tau)$. First note that when $\tau = 0$, the firm raises no money and thus produces $Q^* = 0$. Therefore $u(s^*(0, n)) = 0$

and $0 \in T_2$. To find out if an eligible $\tau > 0$ exists in T_2, we need to know how $u(s^*(\tau, n))$ changes in $\tau \in T_2$.

Under simultaneous arrivals, by (2.1) and (A.33) we have

$$\begin{aligned}
&\frac{du(s^*(\tau,n))}{d\tau}\bigg|_{\tau\in T_2} \\
= \ & \frac{d}{d\tau}\left[\frac{n}{z}\Delta(n)\right] \\
= \ & \frac{n}{z}\left[\frac{v}{m}\frac{d}{d\tau}\mathsf{E}[\min\left\{D, F^{-1}(1-\frac{cm}{(m-n)v}), \frac{\tau\, n}{c}\right\}] - 1\right] \\
= \ & \frac{n}{z}\left[\frac{v}{m}\frac{d}{d\tau}\mathsf{E}[\min\left\{D, \frac{\tau\, n}{c}\right\}]\cdot \mathbb{1}_{\{F^{-1}(1-\frac{cm}{(m-n)v})\geq\frac{\tau\, n}{c}\}}\right] \\
&+\frac{n}{z}\left[\frac{v}{m}\frac{d}{d\tau}\mathsf{E}[\min\left\{D, F^{-1}(1-\frac{cm}{(m-n)v})\right\}]\cdot \mathbb{1}_{\{F^{-1}(1-\frac{cm}{(m-n)v})<\frac{\tau\, n}{c}\}} - 1\right] \\
= \ & \frac{n}{z}\left[\frac{v}{m}(1-F(\frac{\tau\, n}{c}))\,\frac{n}{c}\cdot \mathbb{1}_{\{\tau\leq\frac{c}{n}F^{-1}(1-\frac{cm}{(m-n)v})\}}\right] \\
&+\frac{n}{z}\left[\frac{v}{m}(1-F(F^{-1}(1-\frac{cm}{(m-n)v})))\cdot 0\cdot \mathbb{1}_{\{\tau>\frac{c}{n}F^{-1}(1-\frac{cm}{(m-n)v})\}} - 1\right] \\
= \ & \frac{n}{z}\left[\frac{v}{m}(1-F(\frac{\tau\, n}{c}))\,\frac{n}{c}\cdot \mathbb{1}_{\{\tau\leq\frac{c}{n}F^{-1}(1-\frac{cm}{(m-n)v})\}} - 1\right].
\end{aligned} \tag{A.38}$$

By the analysis of T_1 and (A.38), for $\tau > \frac{c}{n}F^{-1}(1-\frac{cm}{(m-n)v})$, the speculators' profit would either remain the same (if $\tau \in T_1$) or keep decreasing in τ (if $\tau \in T_2$) as $\frac{du(s^*(\tau,n))}{d\tau}|_{\tau\in T_2} = -\frac{n}{z} < 0$. For $\tau \leq \frac{c}{n}F^{-1}(1-\frac{cm}{(m-n)v})$, $u(s^*(\tau,n)$ is either zero (if $\tau \in T_1$) or keeps decreasing in τ (if $\tau \in T_2$) as $(1-F(\frac{\tau\, n}{c})$ decreases in τ. Hence, to guarantee a positive number of speculators and thus non-negative profit, it is necessary and sufficient for the platform to set n such that $\frac{du(s^*(\tau,n))}{d\tau}\Big|_{\tau=0} = \frac{n}{z}\left[\frac{v}{m}(1-F(\frac{0\cdot n}{c}))\,\frac{n}{c}-1\right] > 0$, i.e., $n > \frac{m\,c}{v}$. In this case, $\exists \tau > 0$ s.t. $u(s^*(\tau,n)) > 0$. Note that by definition of $s_0(\tau)$, it must be that $s^*(\tau,n) < s_0(\tau)$ and thus $n < s_0(\tau)$, which means that this τ is indeed in T_2.

(*Sequential*) Consider the sequential arrivals assumption.

When $n > s_0(\tau)$, $s^*(\tau,n) = \min\{s_0(\tau), n\} = s_0(\tau)$ and $u(s^*(\tau,n)) = 0$. Ostensibly, there exists eligible τ's in T_1. However, we have assumed the existence of $s_0(\tau)$ and we need to make sure that it still holds. The existence of $s_0(\tau)$ depends on the behavior of $u(s^*(\tau, n))$ for $\tau \in T_2$. By (A.33) and (A.32), we have

$$\left.\frac{du(s^*(\tau,n))}{d\tau}\right|_{\tau\in T_2} = \frac{d}{d\tau}\Delta(n)$$
$$= \frac{v}{m}\,(1 - F(\frac{\tau\, n}{c}))\,\frac{n}{c}\cdot \mathbb{1}_{\{\tau\le\frac{c}{n}F^{-1}(1-\frac{cm}{(m-n)v})\}} - 1. \quad \text{(A.39)}$$

Note that (A.39) only differs from (A.38) by a scale of $\frac{n}{z}$. We then follow a similar argument as in part (*Simultaneous*, T_2) to show that $s_0(\tau)$ exists if and only if $n > \frac{m\,c}{v}$.

ii) By Part (i), $s^* \geq \frac{m\,c}{v}$. On the other hand, we showed in Section 2.3.2 that $s^* < m\left(1 - \frac{c}{v}\right)$. Therefore, the ICO fails if $m\left(1 - \frac{c}{v}\right) \leq \frac{m\,c}{v}$, i.e., $v \leq 2c$.

□

Proof of Proposition 2.5

i) Shown by Proposition 2.4.

ii) (a) Shown by Proposition A.3.

(b) By Lemma 2.3, $s^*(\tau^*, n^*) = n^* \cdot \mathbb{1}_{\{u(n^*)\geq 0\}}$. By definition of τ^* as in Proposition A.2 part (i), we know that $u(n^*) = u(n^*, n^*, \tau^*) \geq 0$. The result follows.

(c) By Proposition A.2, we know that there exists a unique $\hat{n} \in (\frac{m\,c}{v}, \frac{m}{2})$ such that the following holds:

- $\frac{v\,\hat{n}}{c\,m}\,\mathsf{E}[\min\left\{D, F^{-1}(1 - \frac{c\,m}{(m-\hat{n})v})\right\}] = F^{-1}(1 - \frac{c\,m}{(m-\hat{n})v})$;
- $F^{-1}(1 - \frac{c\,m}{(m-\hat{n})v}) = \frac{\tau^*(\hat{n})\hat{n}}{c}$.

We show in the proof of Proposition A.3 that this $\hat{n}$ is a global maximum point, which we call n^*. Hence, $\frac{v\,n^*}{c\,m}\,\mathsf{E}[\min\left\{D, F^{-1}(1 - \frac{c\,m}{(m-n^*)v})\right\}] = \frac{\tau^*(n^*)n^*}{c}$, and the ICO token price is $\tau^* = \frac{v}{m}\,\mathsf{E}[\min\left\{D, F^{-1}(1 - \frac{c\,m}{(m-n^*)v})\right\}]$.

(d) Following the proof of part (c) and substituting n^* and τ^* into Proposition 2.2 part (i), we have $Q^* = \min\left\{F^{-1}(1 - \frac{c\,m}{(m-n^*)v}), \frac{\tau^* n^*}{c}\right\} = F^{-1}(1 - \frac{c\,m}{(m-n^*)v})$.

(e) By definition of τ^* as in Proposition A.2 part (i), we have $\mathsf{E}[\tau_{eq}] = \tau^*$. We obtain the result by part (c).

iii) By part (a), (c) and (e) of Proposition 2.5, we have $n^* \cdot \tau^* = Q^* \cdot c$.

□

Proof of Proposition 2.6

i) Shown by Proposition 2.5.

ii) By Proposition 2.5, $Q^*_{ICO} = F^{-1}(1 - \frac{c\,m}{(m-n^*)v})$. The optimal production quantity of a traditional newsvendor is $F^{-1}(1 - \frac{c}{v})$. Since $\frac{m}{m-n^*} > 1$ and F^{-1} is an increasing function, we have $F^{-1}(1 - \frac{c\,m}{(m-n^*)v}) < F^{-1}(1 - \frac{c}{v})$.

iii) The ICO newsvendor's profit is given by $\Pi_{ICO} = \tau^*\,s^* - c\,Q^* + (m - s^*)\,\mathsf{E}[\tau_{eq}]$. By Proposition 2.5, $\tau^* = \mathsf{E}[\tau_{eq}]$, therefore

$$\begin{aligned}\Pi_{ICO} &= m\,\mathsf{E}[\tau_{eq}] - c\,Q^* \\ &= v\,\mathsf{E}[\min\left\{D, F^{-1}(1 - \frac{c\,m}{(m-n^*)v})\right\}] - c\,F^{-1}(1 - \frac{c\,m}{(m-n^*)v}) \\ &= \Pi_{traditional}(F^{-1}(1 - \frac{c\,m}{(m-n^*)v})) \end{aligned} \tag{A.40}$$

where $\Pi_{traditional}$ is the profit function of a traditional newsvendor defined as $\Pi_{traditional}(Q) = v\,\mathsf{E}[\min\{D, Q\}] - c\,Q$. We know that $\Pi_{traditional}(Q)$ is maximized by $F^{-1}(1 - \frac{c}{v})$ which is greater than $F^{-1}(1 - \frac{c\,m}{(m-n^*)v})$ by part (ii). Therefore $\Pi_{traditional}(F^{-1}(1 - \frac{c\,m}{(m-n^*)v})) < \Pi_{traditional}(F^{-1}(1 - \frac{c}{v}))$.

iv) The fact that the firm who finances through ICO does not invest its own money makes sure that it never suffers a loss. Indeed, following (A.40),

$$\Pi_{ICO} = v\int_0^{F^{-1}(1-\frac{c\,m}{(m-n^*)v})} x f(x)dx + (\frac{c\,m}{m-n^*} - c)\,F^{-1}(1 - \frac{c\,m}{(m-n^*)v}) > 0. \tag{A.41}$$

□

Proof of Lemma 2.3

i) See the main text.

ii) See the main text.

iii) Fix τ and n. Recall that by (2.1) that $u(s)$ and $\Delta(s)$ have the same sign. Therefore, we can also express $s_0(\tau)$ as $\max\{s \geq 0 : \Delta(s) = 0\}$. We now examine the behavior of $\Delta(s)$ as a function of s:

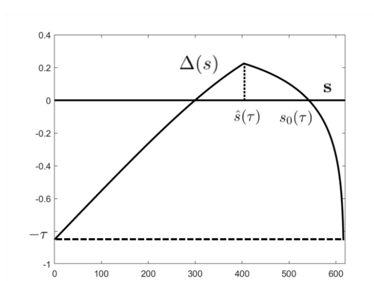
$$\left.\frac{d\Delta(s)}{ds}\right|_{s<m\left(1-\frac{c}{v}\right)} = \frac{v}{m}\left[1 - F(Q^*(s))\right]\left.\frac{dQ^*(s)}{ds}\right|_{s<m\left(1-\frac{c}{v}\right)}, \tag{A.42}$$

where

$$\left.\frac{dQ^*(s)}{ds}\right|_{s\leq m\left(1-\frac{c}{v}\right)} = \begin{cases} -\frac{c\,m}{f(Q^*(s))(m-s)^2\,v} & \text{if } F^{-1}(1-\frac{cm}{(m-s)v}) \leq \frac{\tau\,s}{c} \\ \frac{\tau}{c} & \text{otherwise} \end{cases}. \tag{A.43}$$

Ignoring the sales cap n for the moment, note that for $s \in [0, m\left(1-\frac{c}{v}\right)]$, $F^{-1}(1-\frac{cm}{(m-s)v})$ monotonically decreases in s whereas $\frac{\tau\,s}{c}$ linearly increases in s. Also, $F^{-1}(1-\frac{cm}{(m-s)v})\big|_{s=0} = F^{-1}(1-\frac{c}{v}) > 0 = \frac{\tau\,s}{c}\big|_{s=0}$ and $F^{-1}(1-\frac{cm}{(m-s)v})\big|_{s=m\left(1-\frac{c}{v}\right)} = 0 < \frac{\tau\,s}{c}\big|_{s=m\left(1-\frac{c}{v}\right)}$. Therefore, for any fixed τ, there exists one and only one $\hat{s}(\tau)$ that satisfies $F^{-1}(1-\frac{cm}{(m-\hat{s}(\tau)v)}) = \frac{\tau\,\hat{s}(\tau)}{c}$. By (A.43), $Q^*(s)$ increases in s for $s \in [0, \hat{s}(\tau))$ and decreases in s for $s \in (\hat{s}(\tau), m\left(1-\frac{c}{v}\right))$, and is thus maximized at $\hat{s}(\tau)$. Therefore, (A.42) is positive for all $s \in [0, \hat{s}(\tau))$ and negative for all $s \in (\hat{s}(\tau), m\left(1-\frac{c}{v}\right))$ and $\hat{s}(\tau)$ maximizes $\Delta(s)$. Now note that $\Delta(0) = 0 - \tau = -\tau$ and $\Delta(m\left(1-\frac{c}{v}\right)) = 0 - \tau = -\tau$. This shows that $s_0(\tau) \in [\hat{s}(\tau), m\left(1-\frac{c}{v}\right))$ if it exists. Figure 12 illustrates the relationships between the quantities mentioned above when demand is normally distributed.

Figure 12: $\Delta(s)$ vs s, assuming existence of $s_0(\tau)$



□

Proof of Proposition A.2

i) First note that by Lemma 2.3 or (A.34), for each τ, it is redundant to consider $n > s_0(\tau)$. Therefore, for each n, we can restrict our attention to the set $T_r = \{\tau > 0 : s_0(\tau) \geq n\}$. When $n \leq s_0(\tau)$, we have $s^*(\tau, n) = n$. We will first find $\tau^*(n) \in \mathbb{R}^+$ that maximizes (A.30) evaluated at $s^*(\tau, n) = n$ and then show that this $\tau^*(n)$ is in T_r. Since $T_r \subset \mathbb{R}^+$, this $\tau^*(n)$ must maximize (A.30) over T_r.

Substituting $s^*(\tau, n) = n$ into (A.30) and differentiating with respect to τ,

$$\begin{aligned} \frac{d\Pi}{d\tau} &= n - c\,\frac{dQ^*(n)}{d\tau} + (m-n)\,\frac{v}{m}\,\frac{d}{d\tau}\mathsf{E}[\min\{D, Q^*(n)\}] \\ &= n - c\,\frac{dQ^*(n)}{d\tau} + (m-n)\,\frac{v}{m}\,(1 - F(Q^*(n)))\,\frac{dQ^*(n)}{d\tau} \\ &= n + \left[(m-n)\,\frac{v}{m}\,(1 - F(Q^*(n))) - c\right]\frac{dQ^*(n)}{d\tau} \\ &= n + \left[(m-n)\,\frac{v}{m}\,(1 - F(Q^*(n))) - c\right]\mathbb{1}_{\{F^{-1}(1-\frac{cm}{(m-n)v}) \geq \frac{\tau n}{c}\}}\,\frac{n}{c} \\ &= \begin{cases} n + \left[(m-n)\,\frac{v}{m}\,(1 - F(Q^*(n))) - c\right]\frac{n}{c} & \text{if } F^{-1}(1 - \frac{cm}{(m-n)v}) \geq \frac{\tau n}{c} \\ n & \text{otherwise} \end{cases} \end{aligned} \tag{A.44}$$

Note that $F^{-1}(1 - \frac{cm}{(m-n)v}) \geq \frac{\tau n}{c}$ means $(m-n)\,\frac{v}{m}\,(1 - F(Q^*(n))) - c \geq 0$. Therefore, $\frac{d\Pi}{d\tau} > 0$ for all τ, implying that for a given n, the optimal initial token price $\tau^*(n)$ is given by

$$\tau^*(n) = \max\{\tau : u(s^*(\tau, n)) = \mathsf{E}[\tau_{eq}(Q^*(n))] - \tau \geq 0\}. \tag{A.45}$$

Consider some $n \in (\frac{mc}{v}, m\left(1 - \frac{c}{v}\right)]$ and we know by (A.38) that $\frac{du(s^*(\tau,n))}{d\tau} > 0$ for $\tau \in [0, \tilde{\tau})$ for some $0 < \tilde{\tau} < \frac{c}{n}F^{-1}(1 - \frac{cm}{(m-n)v})$ such that $\frac{du(s^*(\tau,n))}{d\tau}\Big|_{\tau=\tilde{\tau}} = 0$. Given that $u(s^*(0, n)) = 0$, by definition of τ^* given by (A.45), we must have $\tau^*(n) > \tilde{\tau} > 0$. We also know that $\tau^*(n) < \infty$ because by (A.38), the speculators' profit will eventually go negative as τ increases given that $\frac{du(s^*(\tau,n))}{d\tau} < 0$ when $\tau > \frac{c}{n}F^{-1}(1 - \frac{cm}{(m-n)v})$. Therefore, $\tau^*(n) = \max\{\tau : u(s^*(\tau, n)) = 0\}$. Since $u(s^*(\tau, n)) \geq 0$ for all $\tau \in [0, \tau^*(n)]$ and decreases linearly in τ for $\tau > \tau^*(n)$, the equation $u(s^*(\tau, n)) = 0$ has one and only one nonzero solution. We can thus simplify the definition by writing

$\tau^*(n) = \{\tau > 0 : u(s^*(\tau, n)) = 0\}$.

Last, this new definition of $\tau^*(n)$ makes sure that $n \leq s_0(\tau^*(n))$ because $s_0(\tau^*(n))$ is the largest s that gives $u(s) = 0$ by definition. Therefore, $s^*(\tau, n) = n$ still holds. We can then solve $u(s^*(\tau, n)) = \frac{s^*(\tau,n)}{z}\Delta(s^*(\tau, n))$ or equivalently $\Delta(s^*(\tau, n)) = \frac{v}{m}\,\mathsf{E}[\min\left\{D, F^{-1}(1 - \frac{cm}{(m-n)v}), \frac{\tau\, n}{c}\right\}] - \tau = 0$.

ii) For a fixed $n \in (\frac{m\,c}{v}, m\left(1 - \frac{c}{v}\right)]$, we define

- $\tau_1(n) = \frac{v}{m}\,\mathsf{E}[\min\left\{D, F^{-1}\left(1 - \frac{c\,m}{(m-n)v}\right)\right\}]$;
- $\tau_2(n) : \{\tau > 0 : \phi(\tau) = \frac{v}{m}\,\mathsf{E}[\min\left\{D, \frac{\tau\, n}{c}\right\}] - \tau = 0\}$.

By part (i) we know that $\tau^*(n)$ is either equal to $\tau_1(n)$ or given by $\tau_2(n)$.

We first show that $\tau_2(n)$ is finite and unique. Consider $\phi(\tau) = \frac{v}{m}\,\mathsf{E}[\min\left\{D, \frac{\tau\, n}{c}\right\}] - \tau$ and $\phi'(\tau) = \frac{v}{m}\,\frac{n}{c}\,(1 - F(\frac{\tau\, n}{c})) - 1$. Note that $\phi(0) = 0$ and $\phi'(0) > 0$ since $n > \frac{m\,c}{v}$. For large τ, $\phi'(\tau) < 0$ as $\phi''(\tau) = -\frac{v}{m}\,\frac{n^2}{c^2}\,f(\frac{\tau\, n}{c}) < 0$ for all $\tau \geq 0$. Therefore, there exists exactly one $0 < \tau < \infty$, which is $\tau_2(n)$, that gives $\phi(\tau) = 0$. Also note that $\phi'(\tau_2(n)) < 0$ and we will use this result in the proof of Proposition A.3.

Next, let's find out the expression of $\tau^*(n)$ for $n \in (\frac{m\,c}{v}, m\left(1 - \frac{c}{v}\right)]$. Let $g(n) = \frac{\tau_1(n)n}{c} - F^{-1}\left(1 - \frac{c\,m}{(m-n)v}\right)$ and note that $g(n) > 0$ means $\tau^*(n) = \tau_1(n)$. If $g(n) = 0$, then $F^{-1}\left(1 - \frac{c\,m}{(m-n)v}\right) = \frac{\tau_1(n)n}{c}$ and thus $\mathsf{E}[\min\left\{D, F^{-1}\left(1 - \frac{c\,m}{(m-n)v}\right)\right\}]$ $= \mathsf{E}[\min\left\{D, \frac{\tau_2(n)n}{c}\right\}]$, which by definition implies that $\tau_1(n) = \tau_2(n) = \tau^*(n)$. Also, $g(n) < 0$ means $\tau^*(n) \neq \tau_1(n)$ and thus $\tau^*(n) = \tau_2(n)$. We will first look at $n \in (\frac{m\,c}{v}, \frac{m}{2}]$ and then $n \in (\frac{m}{2}, m\left(1 - \frac{c}{v}\right)]$ ($v > 2c$ guarantees that $\frac{m\,c}{v} < \frac{m}{2} < m\left(1 - \frac{c}{v}\right)$). Consider $n \in (\frac{m\,c}{v}, \frac{m}{2}]$. Note that $g(\frac{m\,c}{v}) = \mathsf{E}[\min\left\{D, F^{-1}(\frac{v-2c}{v})\right\}] - F^{-1}(\frac{v-2c}{v}) < 0$ and we now show that $g(\frac{m}{2}) > 0$. Let $r = \frac{v}{c}$ and we know that $r > 2$. Define $\tilde{g}(r) = g(\frac{m}{2}) = \frac{v}{2c}\,\mathsf{E}[\min\left\{D, F^{-1}\left(1 - \frac{2c}{v}\right)\right\}] - F^{-1}\left(1 - \frac{2c}{v}\right) = \frac{r}{2}\,\mathsf{E}[\min\left\{D, F^{-1}\left(1 - \frac{2}{r}\right)\right\}] - F^{-1}\left(1 - \frac{2}{r}\right)$. When $r = 2$, $\tilde{g}(2) = 0$. For $r \geq 2$, $\tilde{g}(r)$ increases in r as $\tilde{g}'(r) = \frac{1}{2}\,\mathsf{E}[\min\left\{D, F^{-1}\left(1 - \frac{2}{r}\right)\right\}] + \frac{r}{2}\left[1 - FF^{-1}\left(1 - \frac{2}{r}\right)\right]\frac{d}{dr}F^{-1}\left(1 - \frac{2}{r}\right) - \frac{d}{dr}F^{-1}\left(1 - \frac{2}{r}\right) = \frac{1}{2}\,\mathsf{E}[\min\left\{D, F^{-1}\left(1 - \frac{2}{r}\right)\right\}] > 0$ for $r \geq 2$. Therefore, $g(\frac{m}{2}) = \tilde{g}(r) > 0$ for all $r > 2$.

Next note that

$$
\begin{aligned}
g'(n) &= \frac{\tau_1(n)}{c} + \frac{v\,n}{c\,m}(1 - FF^{-1}\left(1 - \frac{c\,m}{(m-n)v}\right))\frac{d}{dn}F^{-1}\left(1 - \frac{c\,m}{(m-n)v}\right)] \\
&\quad -\frac{d}{dn}F^{-1}\left(1 - \frac{c\,m}{(m-n)v}\right) \\
&= \frac{v}{cm}\mathsf{E}\left[\min\left\{D, F^{-1}\left(1 - \frac{c\,m}{(m-n)v}\right)\right\}\right] + \frac{2n-m}{m-n}\frac{d}{dn}F^{-1}\left(1 - \frac{c\,m}{(m-n)v}\right)
\end{aligned} \tag{A.46}
$$

Since $\frac{d}{dn}F^{-1}\left(1 - \frac{c\,m}{(m-n)v}\right) < 0$, when $n \le \frac{m}{2}$, $g'(n) > 0$. Therefore, there must exist a unique $\hat{n} \in (\frac{m\,c}{v}, \frac{m}{2})$ such that $g(\hat{n}) = 0$. This means that $\tau^*(n) = \tau_2(n)$ for $n \in (\frac{m\,c}{v}, \hat{n})$, $\tau^*(n) = \tau_1(n)$ for $n \in (\hat{n}, \frac{m}{2}]$, and $\tau^*(n) = \tau_1(n) = \tau_2(n)$ when $n = \hat{n}$.

For $n \in (\frac{m}{2}, m\left(1 - \frac{c}{v}\right)]$, we have $g(m\left(1 - \frac{c}{v}\right)) = \frac{v-c}{c}\mathsf{E}[\min\left\{D, F^{-1}(0)\right\}] - F^{-1}(0) = 0$ and $g'(m\left(1 - \frac{c}{v}\right)) = 0 + \frac{2n-m}{m-n}\frac{d}{dn}F^{-1}\left(1 - \frac{c\,m}{(m-n)v}\right) < 0$ by (A.46). Since we have shown that $g(\frac{m}{2}) > 0$, there must be either zero or more than one $\hat{n} \in (\frac{m}{2}, m\left(1 - \frac{c}{v}\right))$ such that $g(\hat{n}) = 0$. To rule out multiple zeros in the range of $n \in (\frac{m}{2}, m\left(1 - \frac{c}{v}\right)]$, a sufficient condition is that $g''(n) < 0$ for $n \in (\frac{m}{2}, m\left(1 - \frac{c}{v}\right)]$. We can find $g''(n)$ from (A.46) and, after some algebra, simplify it as

$$
g''(n) = -\frac{c\,m}{f(F^{-1}(y))(m-n)^4 v}\left[3n + \frac{(2n-m)cm}{(m-n)v} \cdot \frac{f'(F^{-1}(y)}{(f(F^{-1}(y))^2}\right], \tag{A.47}
$$

where $y = 1 - \frac{c\,m}{(m-n)v}$.

Last, we find a sufficient condition for $g''(n) < 0$ for $n \in (\frac{m}{2}, m\left(1 - \frac{c}{v}\right)]$ or equivalently $y \in [0, 1 - \frac{2c}{v})$. By (A.47), to make $g''(n) < 0$, it suffices to have $3n + \frac{(2n-m)cm}{(m-n)v} \cdot \frac{f'(F^{-1}(y)}{(f(F^{-1}(y))^2} > 0$, or

$$
\frac{f'(F^{-1}(y)}{(f(F^{-1}(y))^2} > -\frac{3n(m-n)v}{(2n-m)cm}. \tag{A.48}
$$

Let $\frac{n}{m} = k$. Then $k = 1 - \frac{c}{(1-y)v}$ and we look at $k \in (\frac{1}{2}, 1 - \frac{c}{v}]$. Then, the right hand side of (A.48) is equal to $-\frac{3v}{c} \cdot k \cdot \frac{1-k}{2k-1}$. Therefore, under our assumption, (A.48) holds.

□

Proof of Proposition A.3

By Proposition A.2, we know that there exists a unique $\hat{n} \in (\frac{m\,c}{v},\ \frac{m}{2})$ such that the following holds:

- $\frac{v\,\hat{n}}{c\,m}\,\mathsf{E}[\min\left\{D, F^{-1}(1-\frac{c\,m}{(m-\hat{n})v})\right\}] = F^{-1}(1-\frac{c\,m}{(m-\hat{n})v})$;
- $F^{-1}(1-\frac{c\,m}{(m-\hat{n})v}) = \frac{\tau^*(\hat{n})\hat{n}}{c}$.

We will first show that this $\hat{n}$ is a local maximum point. Differentiating the objective function (A.31) with respect to n, we have

$$\frac{d\Pi}{dn} = \frac{d\tau^*(n)}{dn}\cdot n + \tau^*(n) \,-\, c\,\frac{dQ^*(n)}{dn} + (m-n)\,\frac{d\mathsf{E}[\tau_{eq}(n)]}{dn} \,-\, \mathsf{E}[\tau_{eq}(n)]. \tag{A.49}$$

By part (i), we know that $\tau^*(n) = \mathsf{E}[\tau_{eq}(n)]$ and consequently simplify (A.49) as

$$\begin{aligned} \frac{d\Pi}{dn} &= n\,\frac{d\tau^*(n)}{dn} \,-\, c\,\frac{dQ^*(n)}{dn} + (m-n)\,\frac{d\mathsf{E}[\tau_{eq}(n)]}{dn} \\ &= n\,\frac{d\tau^*(n)}{dn} \,-\, c\,\frac{dQ^*(n)}{dn} + (m-n)\frac{v}{m}(1-F(Q^*(n)))\frac{dQ^*(n)}{dn} \\ &= n\,\frac{d\tau^*(n)}{dn} \,+\, [(m-n)\frac{v}{m}(1-F(Q^*(n))) - c]\frac{dQ^*(n)}{dn}. \end{aligned} \tag{A.50}$$

Let's now evaluate $\frac{d\Pi}{dn}$ at $n = \hat{n}$. We know that $\tau^*(\hat{n}) = \frac{v}{m}\,\mathsf{E}[\min\left\{D, F^{-1}(1-\frac{c\,m}{(m-\hat{n})v})\right\}]$ and $Q^*(\hat{n}) = Q^*(\tau^*(\hat{n}), \hat{n}) = \min\left\{F^{-1}(1-\frac{cm}{(m-\hat{n})v}), \frac{\tau^*(\hat{n})\,\hat{n}}{c}\right\} = F^{-1}(1-\frac{cm}{(m-\hat{n})v})$. Therefore, $(m-\hat{n})\frac{v}{m}(1-F(Q^*(\hat{n}))) - c$ vanishes. Hence,

$$\begin{aligned} \left.\frac{d\Pi}{dn}\right|_{n=\hat{n}} &= \hat{n}\,\left.\frac{d\tau^*(n)}{dn}\right|_{n=\hat{n}} + 0 \\ &= \frac{v\,\hat{n}}{m}(1-F(Q^*(\hat{n})))\left.\frac{dQ^*(n)}{dn}\right|_{n=\hat{n}} \\ &= \frac{c\,\hat{n}}{m-\hat{n}}\left.\frac{dQ^*(n)}{dn}\right|_{n=\hat{n}} \end{aligned} \tag{A.51}$$

$Q^*(n)$ is not differentiable at $n = \hat{n}$ and thus $\frac{dQ^*(n)}{dn}\big|_{n=\hat{n}}$ does not exist. However, we've shown in the proof of Lemma 2.3 part iii) that given $\tau^*(\hat{n})$, $\frac{dQ^*(\tau^*(\hat{n}),n)}{dn}\big|_{n<\hat{n}} > 0$ and

$\frac{dQ^*(\tau^*(\hat{n}),n)}{dn}\Big|_{n>\hat{n}} < 0$. Therefore we know that $\lim_{n\to\hat{n}^-} \frac{d\Pi}{dn} > 0$ and $\lim_{n\to\hat{n}^+} \frac{d\Pi}{dn} < 0$, suggesting that $\hat{n}$ maximizes profit locally.

Last, we will show that $\hat{n}$ is the global maximum point by showing that (A.50) is negative for $n \in (\hat{n}, m\left(1-\frac{c}{v}\right)]$ and positive for $[\frac{m\,c}{v}, \hat{n})$.

For $n \in (\hat{n}, m\left(1-\frac{c}{v}\right)]$, we have $F^{-1}(1-\frac{cm}{(m-n)v}) < \frac{\tau^*(n)\,n}{c}$, $Q^*(n) = F^{-1}(1-\frac{cm}{(m-n)v})$ so $(m-n)\frac{v}{m}(1-F(Q^*(n)))-c = 0$. Since $\frac{d\tau^*(n)}{dn} = \frac{v}{m}\frac{d}{dn}\mathsf{E}[\min\left\{D, F^{-1}\left(1-\frac{c\,m}{(m-n)v}\right)\right\}] < 0$, (A.50) is negative.

Now for $n \in (\frac{m\,c}{v}, \hat{n})$, we have $F^{-1}(1-\frac{cm}{(m-n)v}) > \frac{\tau^*(n)\,n}{c}$ and $Q^*(n) = \frac{\tau^*(n)\,n}{c}$.

$$\begin{aligned}
\frac{d\Pi}{dn}\Big|_{\frac{m\,c}{v}<n<\hat{n}} &= n\,\frac{d\tau^*(n)}{dn}\Big|_{\frac{m\,c}{v}<n<\hat{n}} + \left[(m-n)\frac{v}{m}(1-F(Q^*(n)))-c\right]\frac{dQ^*(n)}{dn}\Big|_{\frac{m\,c}{v}<n<\hat{n}} \\
&= \frac{n\,v}{m}\,[1-F(Q^*(\hat{n}))]\,\frac{dQ^*(n)}{dn}\Big|_{\frac{m\,c}{v}<n<\hat{n}} \\
&\quad + \left[(m-n)\frac{v}{m}(1-F(Q^*(n)))-c\right]\frac{dQ^*(n)}{dn}\Big|_{\frac{m\,c}{v}<n<\hat{n}} \\
&= \left[(m-n+n)\frac{v}{m}(1-F(Q^*(n)))-c\right]\frac{dQ^*(n)}{dn}\Big|_{\frac{m\,c}{v}<n<\hat{n}} \\
&= [v\,[1-F(Q^*(n))]-c]\,\frac{dQ^*(n)}{dn}\Big|_{\frac{m\,c}{v}<n<\hat{n}} \qquad \text{(A.52)}
\end{aligned}$$

Note that

$$\begin{aligned}
v\,[1-F(Q^*(n))]-c &= v\,[1-F(\frac{\tau^*(n)\,n}{c})]-c \\
&> v\,[(1-F(F^{-1}(1-\frac{cm}{(m-n)v}))]-c \\
&= \frac{cm}{m-n}-c \\
&> 0. \qquad \text{(A.53)}
\end{aligned}$$

and $\frac{dQ^*(n)}{dn}\Big|_{\frac{m\,c}{v}<n<\hat{n}} = \frac{\tau^*(n)}{c} + \frac{n}{c}\frac{d\tau^*(n)}{dn}\Big|_{\frac{m\,c}{v}<n<\hat{n}}$. Therefore, to show that (A.52) is positive, it suffices to show $\frac{d\tau^*(n)}{dn}\Big|_{\frac{m\,c}{v}<n<\hat{n}} > 0$. By Proposition A.2, when $\frac{m\,c}{v} < n < \hat{n}$, $\frac{d\tau^*(n)}{dn} =$

$\frac{v}{m}(1 - F(\frac{\tau^*(n)\, n}{c})) \left[\frac{\tau^*(n)}{c} + \frac{n}{c}\frac{d\tau^*(n)}{dn}\right]$. Rearranging, we have

$$\frac{d\tau^*(n)}{dn} = -\frac{\frac{v}{m}(1 - F(\frac{\tau^*(n)\, n}{c}))\frac{\tau^*(n)}{c}}{\frac{v}{m}(1 - F(\frac{\tau^*(n)\, n}{c}))\frac{n}{c} - 1} \tag{A.54}$$

The denominator of (A.54) is equal to $\phi'(\tau^*(n))$ where ϕ is defined in the proof of Proposition A.2 and we've shown that $\phi'(\tau^*(n)) < 0$. Therefore, (A.54) is positive and this completes the proof. □

Proof of Proposition A.4

The case where $s_0(\tau)$ does not exist is trivial. Suppose that $s_0(\tau)$ exists. When $n > s_0(\tau)$, by Lemma 2.3 part iii), we know that $s^*(\tau, n) = s_0(\tau)$ under sequential arrival. Now consider the case $n \leq s_0(\tau)$. We show in the proof of Lemma 2.3 part iii) that $u(0) < 0$ and $u(s)$ is continuous and crosses zero at most once for $s \in [0, s_0(\tau))$. Therefore, if $u(n) < 0$, then $u(s) < 0$ for all $s \in [0, n]$. This means that no $s \leq \min\{s_0(\tau), n\}$ satisfies the participation constraint and hence $s^* = 0$. On the other hand, if $u(n) \geq 0$, then $s = n$ satisfies the participation constraint.

To see why (A.35) is a sufficient condition for s^* speculators, let's first consider the $s^* - th$ speculator that arrives after $s^* - 1$ other speculators have bought a token each. She knows that if she buys a token, then she will be the last person to do so — either because there is no extra token for sale ($s^* = n$) or buying tokens after her is no longer attractive ($s^* = m\left(1 - \frac{c}{v}\right)$). Therefore, (A.35) guarantees non-negative utility for her. Next, the $(s^* - 1) - th$ speculator knows that even if $u(s^* - 1) < 0$, buying a token now would induce the $s^* - th$ speculator to buy a token later, eventually resulting in non-negative rewards. By induction, we see that it is always optimal to buy a token for prior speculators.

□

Proof of Corollary A.5

Substituting the expression of τ^* in Proposition 2.5 part c) into part a), we see that n^*

and $\tau^*(n^*)$ satisfy $\frac{n^*}{c}\tau^*(n^*) = F^{-1}(1 - \frac{c\,m}{(m-n^*)v})$. Therefore, given $\tau^*(n^*)$, we know that $n^* = \hat{s}(\tau^*(n^*))$ where $\hat{s}(\tau)$ is the unique maximum point of $u(\tau, s)$ as defined in the proof of Lemma 2.3 part iii). Additionally, since $u(\tau^*(n^*), n^*) = 0$, we know that n^* is the only value of s such that $u(\tau^*(n^*), s) = 0$. Therefore, by definition of s_0, the result follows.

□

Proof of Proposition 2.7

Let Π_e denote the expected final wealth of the firm that issues equity tokens. Ignoring the budget constraint for the moment and taking derivative of Π_e with respect to Q, by (2.6),

$$\begin{aligned}
\frac{d\Pi_e}{dQ} &= v\,[1 - F(Q)] - c - \frac{s}{m}\,\frac{d}{dQ}\mathsf{E}[v \min\{Q, D\} - c\,Q]^+ \\
&= v\,[1 - F(Q)] - c - \frac{s}{m}\,\frac{d}{dQ}\left[(v - c)\,Q\,[1 - F(Q)] + \int_{\frac{c}{v}Q}^{Q}(v\,x - c\,Q)f(x)dx\right] \\
&= v\,[1 - F(Q)] - c - \frac{s}{m}\,\left[v\,[1 - F(Q)] - c + c\,F\left(\frac{c}{v}Q\right)\right] \\
&= \frac{m - s}{m}\,[v\,[1 - F(Q)] - c] - \frac{s\,c}{m}\,F\left(\frac{c}{v}Q\right). \qquad \text{(A.55)}
\end{aligned}$$

By (A.55), for $s \in (0, m)$, $\frac{d\Pi_e}{dQ}\big|_{Q=0} = \frac{m-s}{m}(v - c) - 0 > 0$ and $\frac{d^2\Pi_e}{dQ^2}\big|_{Q>0} = \frac{m-s}{m}[-f(Q)v] - \frac{s\,c}{m} \cdot \frac{c}{v}\,f\left(\frac{c}{v}Q\right) < 0$. Therefore, there exists a unique unconstrained optimal production quantity, denoted by $Q_u^*(s)$, such that $\frac{d\Pi_e}{dQ}\big|_{Q=Q_u^*(s)} = 0$, i.e.,

$$\frac{m - s}{m}\,[v\,[1 - F(Q_u^*(s))] - c] = \frac{s\,c}{m}\,F\left(\frac{c}{v}Q_u^*(s)\right). \qquad \text{(A.56)}$$

Next, we show that $\frac{dQ_u^*(s)}{ds} < 0$. Differentiating (A.56) with respect to s, we get

$$-(v-c)+v\,F(Q_u^*(s))-c\,F\left(\frac{c}{v}Q_u^*(s)\right) = \left[(m - s)v\,f(Q_u^*(s)) + s\,c\,f\left(\frac{c}{v}Q_u^*(s)\right) \cdot \frac{c}{v}\right]\frac{dQ_u^*(s)}{ds}. \qquad \text{(A.57)}$$

By (A.56), the left-hand side of (A.57) equals $-\frac{m}{s}[v\,(1 - F(Q_u^*(s))) - c$, which is negative. Since the coefficient of $\frac{dQ_u^*(s)}{ds}$ on the right-hand side of (A.57) is positive, $\frac{dQ_u^*(s)}{ds}$ must be negative.

□

Proof of Proposition 2.8

i) To make the ICO successful, the firm needs to set a (τ_e, n_e) pair such that a positive number of speculators participate in the ICO, i.e., $s(\tau_e, n_e) > 0$, which requires the participation constraint.

We first evaluate the behavior of $\Delta(s(\tau_e, n_e))$. Now, $\Delta(s(\tau_e, n_e))$ $= \frac{1}{m}\mathsf{E}[v \min\{Q_e^*(s(\tau_e, n_e)), D\} - c\,Q_e^*(s(\tau_e, n_e))]^+ - \tau_e$. For a fixed τ_e,

$$\begin{aligned}\frac{d\Delta(s)}{ds} &= \frac{1}{m}\frac{\partial}{\partial Q_e^*(s)}\mathsf{E}[v \min\{Q_e^*(s), D\} - c\,Q_e^*(s)]^+ \frac{dQ_e^*(s)}{ds} \\ &= \frac{1}{m}\left\{v\,[1 - F(Q_e^*(s))] - c + c\,F\left(\frac{c}{v}Q_e^*(s)\right)\right\}\frac{dQ_e^*(s)}{ds}.\end{aligned} \tag{A.58}$$

Following similar arguments as in Lemma 2.3 (iii) and the regularity assumption that $f(x) < a^2 \cdot f(ax)$ for $a > 2$, we can show that $v\,[1 - F(Q_e^*(s))] - c + c\,F\left(\frac{c}{v}Q_e^*(s)\right) > 0$ for all s. This, given that $\frac{dQ_u^*(s)}{ds} < 0$, means that there exists a unique $\hat{s}(\tau_e)$ that satisfies $Q_u^*(s) = \frac{\tau_e\,\hat{s}(\tau_e)}{c}$ and $\hat{s}$ maximizes $\Delta(s)$.

Next, following the argument in Proposition 2.4 (i), we have

$$\begin{aligned}&\left.\frac{du(s^*(\tau_e, n_e))}{d\tau_e}\right|_{\tau_e \in T_2} \\ =\ & \frac{d}{d\tau_e}\left[\frac{n_e}{z}\Delta(n_e)\right] \\ =\ & \frac{n_e}{z}\left[\frac{1}{m}\frac{d}{d\tau_e}\mathsf{E}[v \min\{Q_e^*(n_e), D\} - c\,Q_e^*(n_e)]^+ - 1\right] \\ =\ & \frac{n_e}{z}\left[\frac{1}{m}\frac{\partial}{\partial Q_e^*}\mathsf{E}[v \min\{Q_e^*(n_e), D\} - c\,Q_e^*(n_e)]^+ \frac{dQ_e^*}{d\tau_e} - 1\right] \\ =\ & \frac{n_e}{z}\left[\frac{1}{m}\left\{v\,[1 - F(Q_e^*(n_e))] - c + c\,F\left(\frac{c}{v}Q_e^*(n_e)\right)\right\}\frac{dQ_e^*}{d\tau_e} - 1\right].\end{aligned} \tag{A.59}$$

Again, the firm needs $\left.\frac{du(s^*(\tau_e, n_e))}{d\tau_e}\right|_{\tau_e = 0} = \frac{n_e}{z}\left[\frac{1}{m}\{v - c + 0\}\frac{n_e}{c} - 1\right] > 0$, i.e., $n_e > \frac{c}{v-c}\,m$.

ii) Since we need $n_e < m$, by part (i), we must have $1 > \frac{c}{v-c}$, i.e., $v > 2c$.

□

Proof of Proposition 2.9

For a fixed n_e, $\frac{d\Pi_e}{d\tau_e} = \frac{\partial \Pi_e}{\partial \tau_e} + \frac{\partial \Pi_e}{\partial Q_e^*}\frac{dQ_e^*}{d\tau_e} = n_e + \frac{\partial \Pi_e}{\partial Q_e^*}\frac{dQ_e^*}{d\tau_e}$. Note that $\frac{\partial \Pi_e}{\partial Q_e^*} > 0$ because $Q_e^* \leq Q_u^*$, and $\frac{dQ_e^*}{d\tau_e} = \frac{n_e}{c}$ or 0. Therefore, we know that $\frac{d\Pi_e}{d\tau_e} > 0$. Given that τ_e^* must satisfy the participation constraint, we have $u(s^*(\tau_e^*(n_e))) = 0$. By (A.59), we know that such τ_e^* is finite. Lastly, since $u(s^*(\tau_e, n_e)$ is linear in τ_e, $\tau_e^*(n_e)$ must be unique.

□

Proof of Proposition 2.10

Differentiate (2.8) with respect to Q,

$$\frac{d\Pi}{dQ} = [(m-s)\,\frac{\alpha\, v}{m} - c\,] - \alpha\,(m-s)\,\frac{v}{m}F(Q) \tag{A.60}$$

By (A.60), $\frac{d\Pi}{dQ} < 0$ when $\alpha(m-s)\,\frac{v}{m} - c < 0$, i.e., $s > m(1-\frac{c}{\alpha\, v})$. On the other hand, when $s \leq m(1-\frac{c}{\alpha\, v})$, ignoring the budget constraint and setting $\frac{d\Pi}{dQ} = 0$, we get $Q^*_{unconstrained}(s) = F^{-1}(1 - \frac{cm}{\alpha(m-s)v})$. Since $\frac{d^2\Pi}{dQ^2} = -\,\alpha\,(m-s)\,\frac{v}{m}f(Q) < 0$, the profit function is concave in Q and $Q^*_{unconstrained}$ is a maximum. Hence the firm's optimal production quantity is given by

$$Q^*(s) = \min\left\{F^{-1}(1 - \frac{cm}{\alpha\,(m-s)v}), \frac{\tau\, s}{c}\right\} \cdot \mathbb{1}_{\{s \leq m(1-\frac{c}{\alpha\, v})\}}. \tag{A.61}$$

□

Proof of Proposition 2.11

i) We substitute the new definition of the market equilibrium token price, $\tau_{eq} = \alpha \cdot \frac{v}{m}\min\{Q, D\}$, into (2.1), and then follow similar arguments in the proofs of Lemma 2.3(iii) and Proposition 2.4.

Applying (A.61), we have

$$\left.\frac{du(s^*(\tau,n))}{d\tau}\right|_{\tau\in T_2} = \frac{d}{d\tau}\left[\frac{n}{z}\Delta(n)\right]$$
$$= \frac{n}{z}\left[\frac{\alpha\, v}{m}\,(1-F(\frac{\tau\, n}{c}))\,\frac{n}{c}\cdot \mathbb{1}_{\{\tau\le\frac{c}{n}F^{-1}(1-\frac{cm}{\alpha\,(m-n)v})\}}-1\right] \quad \text{(A.62)}$$

By the analysis of T_1 and (A.62), for $\tau > \frac{c}{n}F^{-1}(1-\frac{cm}{\alpha(m-n)v})$, the speculators' profit would either remain the same (if $\tau \in T_1$) or keep decreasing in τ (if $\tau \in T_2$) as $\frac{du(s^*(\tau,n))}{d\tau}|_{\tau\in T_2} = -\frac{n}{z} < 0$. For $\tau \le \frac{c}{n}F^{-1}(1-\frac{cm}{\alpha(m-n)v})$, $u(s^*(\tau,n))$ is either zero (if $\tau \in T_1$) or keeps decreasing in τ (if $\tau \in T_2$) as $(1-F(\frac{\tau\, n}{c})$ decreases in τ. Hence, to guarantee a positive number of speculators and thus non-negative profit, it is necessary and sufficient for the platform to set n such that $\left.\frac{du(s^*(\tau,n))}{d\tau}\right|_{\tau=0} = \frac{n}{z}\left[\frac{\alpha v}{m}\,(1-F(\frac{0\cdot n}{c}))\,\frac{n}{c}-1\right] > 0$, i.e., $n > \frac{m\,c}{\alpha v}$. In this case, $\exists\tau > 0$ s.t. $u(s^*(\tau,n)) > 0$. Note that by definition of $s_0(\tau)$, it must be that $s^*(\tau,n) < s_0(\tau)$ and thus $n < s_0(\tau)$, which means that this τ is indeed in T_2.

ii) By Part (i), $s^* \ge \frac{m\,c}{\alpha v}$. On the other hand, we showed in §2.5.1 that $s^* < m(1-\frac{c}{\alpha\, v})$. Therefore, the ICO fails if $m(1-\frac{c}{\alpha\, v}) \le \frac{m\,c}{\alpha\, v}$, i.e., $v \le \frac{2c}{\alpha}$.

□

Proof of Proposition 2.12

i) We substitute the new definition of the expected profit given by (2.9) into (2.1), and then follow similar arguments in the proofs of Lemma 2.3(iii) and Proposition 2.4. Applying (A.37), we have

$$\left.\frac{du(s^*(\tau,n))}{d\tau}\right|_{\tau\in T_2} = \frac{d}{d\tau}\left[\frac{n}{z}\Delta(n)\right]$$
$$= \frac{n}{z}\left[\frac{v}{m}\,(1-F(\frac{\tau\, n}{c}))\,\frac{n}{c}\cdot \mathbb{1}_{\{\tau\le\frac{c}{n}F^{-1}(1-\frac{cm}{(m-n)v})\}}-(1+k)\right] \quad \text{(A.63)}$$

By the analysis of T_1 and (A.63), for $\tau > \frac{c}{n}F^{-1}(1-\frac{cm}{(m-n)v})$, the speculators' profit would either remain the same (if $\tau \in T_1$) or keep decreasing in τ (if $\tau \in T_2$) as

$\frac{du(s^*(\tau,n))}{d\tau}|_{\tau\in T_2} = -\frac{n}{z}(1+k) < 0$. For $\tau \leq \frac{c}{n}F^{-1}(1-\frac{cm}{(m-n)v})$, $u(s^*(\tau,n))$ is either zero (if $\tau \in T_1$) or keeps decreasing in τ (if $\tau \in T_2$) as $(1-F(\frac{\tau n}{c})$ decreases in τ. Hence, to guarantee a positive number of speculators and thus non-negative profit, it is necessary and sufficient for the platform to set n such that $\frac{du(s^*(\tau,n))}{d\tau}\Big|_{\tau=0} = \frac{n}{z}\left[\frac{v}{m}\left(1-F(\frac{0\cdot n}{c})\right)\frac{n}{c}-(1+k)\right] > 0$, i.e., $n > \frac{m\,c}{v}(1+k)$. In this case, $\exists\tau > 0$ s.t. $u(s^*(\tau,n)) > 0$. Note that by definition of $s_0(\tau)$, it must be that $s^*(\tau,n) < s_0(\tau)$ and thus $n < s_0(\tau)$, which means that this τ is indeed in T_2.

ii) By Part (i), $s^* \geq \frac{m\,c}{v}(1+k)$. On the other hand, we showed in §2.5.2 that $s^* < m\left(1-\frac{c}{v}\right)$. Therefore, the ICO fails if $m\left(1-\frac{c}{v}\right) \leq \frac{m\,c}{v}(1+k)$, i.e., $v \leq (2+k)c$.

□

Proof of Proposition 2.13

Since the firm's objective function remains unchanged by adding the outside option, the proof of this proposition resembles that of Proposition A.2 (i). □

A.3. Uncapped ICO

A.3.1. Proofs

Proof of Proposition 3.1

i) First note that in period t with service level s_t, the market-clearing dollar-denominated price the customers pay is $p(s_t) = x \cdot \tau_{eq}(t, s_t)$, where x denotes the token-denominated price of service. Suppose $x > m_t/s_t$, then the demand of tokens $x \cdot s_t$ exceeds the supply of tokens, m_t. This will drive the price of the token up, resulting in a decrease in the token-denominated price. In other words, $\tau(t, s_t)$ will increase and x will decrease. Similarly, if $x < m_t/s_t$, then the demand of tokens is less than the supply of tokens, which induces an increase in x. Therefore, in equilibrium, demand of tokens is equal to its supply, i.e., $x \cdot s_t = m_t$.

ii) The result follows immediately from $\tau(t, s_t) = p(s_t)/x$ and Part (i). □

Proof of Proposition 3.2

By Lemma 3.1, we can rewrite (3.3) as $\phi(s_t) = p(s_{t+1})\frac{s_{t+1}}{s_t} - c$. We see that given any future service level $s_{t+1} > 0$, service provider's profit decreases in the current service level, i.e., $\frac{d\phi(s_t)}{ds_t} < 0$. Since service providers join the market as long as their profit is non-negative, $s_t^* = \min\{\max\{s \geq 0 \text{ s.t. } \phi(s) = 0\}, M\}$, which simplifies to $s_t^* = \min\{\frac{p(s_{t+1})\, s_{t+1}}{c}, M\}$. □

Proof of Proposition 3.3

A steady-state service level means $s_{ss}^* = s_{t+1} = s_t$ for all t. Consider some $s_{t+1} > 0$. By Lemma 3.2, if $s_{t+1} < M$, then the steady-state service level must make $p(s_{ss}^*) = c$, which means $\phi(s_{ss}^*) = 0$ by (3.4); if $s_{t+1} = M$, then $s_{ss}^* = M$, and we will show that $\phi(s_{ss}^*) = \phi(M) > 0$.

Let's first consider the case where $s_{t+1} \in (0, M]$. Let $y(s_{t+1}) = c - p(s_{t+1})$. By (3.1), we have

$$y(s_{t+1}) = c - \left(ks_{t+1}^2 - \frac{s_{t+1}}{M} + 1\right) = -ks_{t+1}^2 + \frac{s_{t+1}}{M} - (1 - c). \tag{A.64}$$

After simple algebra, we know that if $k \leq \frac{1}{4M^2(1-c)}$, then $y(s_{t+1}) = 0$ at $s_{t+1} = \frac{1 \pm \sqrt{1-4kM^2(1-c)}}{2kM}$; if $k > \frac{1}{4M^2(1-c)}$, then $y(s_{t+1}) < 0$ for all $s_{t+1} \in (0, M]$.

Suppose $k \leq \frac{1}{4M^2(1-c)}$ and let $s_1 = \frac{1-\sqrt{1-4kM^2(1-c)}}{2kM}$ and $s_2 = \frac{1+\sqrt{1-4kM^2(1-c)}}{2kM}$. We now compare s_1 and s_2 with M to see if these zeros of $y(\cdot)$ are in the feasible range of service levels. It is easy to see that if $s_1 \leq M$, then $k \leq \frac{c}{M^2}$. That is to say, $k \leq \frac{c}{M^2}$ is a necessary condition for $s_1 \leq M$. Note that $\frac{1}{4(1-c)} \geq c$ for all c, so $k \leq \frac{c}{M^2}$ is more stringent than $k \leq \frac{1}{4(1-c)}$. Similarly, a necessary condition for $s_1 \leq M$ is $\frac{1}{4M^2(1-c)} \geq k \geq \frac{c}{M^2}$. However, these necessary conditions may not be sufficient conditions. We next discuss feasibility of s_1 and s_2 under these conditions.

1. $k < \frac{c}{M^2} (\leq \frac{1}{4M^2(1-c)})$. The necessary condition for $s_2 \leq M$ is violated, so $s_2 > M$. Now, $s_1 = \frac{1-\sqrt{1-4kM^2(1-c)}}{2kM} < \frac{1-\sqrt{1-4kM^2(1-kM^2))}}{2kM} = \frac{1-|2kM^2-1|}{2kM}$.

   - When $c \leq \frac{1}{2}$, we have $kM^2 < c \leq \frac{1}{2}$, so $s_1 < \frac{1-|2kM^2-1|}{2kM} = M$.

- When $kM^2 \le \frac{1}{2} < c$, we have $s_1 < \frac{1-|2kM^2-1|}{2kM} = M$.
- When $\frac{1}{2} < kM^2 < c$, we have $s_1 < \frac{1-|2kM^2-1|}{2kM} = \frac{1-(2kM^2-1)}{2kM} < \frac{1-(1-2kM^2)}{2kM} = M$.

Therefore, $s^*_{ss} = s_1 < M$ and is thus feasible.

2. $k = \frac{c}{M^2} (\le \frac{1}{4M^2(1-c)})$. $s_{1,2} = \frac{1 \pm \sqrt{1-4kM^2(1-c)}}{2kM} = \frac{1 \pm \sqrt{1-4kM^2(1-kM^2))}}{2kM} = \frac{1 \pm |2kM^2-1|}{2kM}$.

   - When $c < \frac{1}{2}$, we have $s_1 = \frac{1-|2kM^2-1|}{2kM} = \frac{1-(1-2kM^2)}{2kM} = M$. $s_2 > s_1 = M$.
   - When $c = \frac{1}{2}$, we have $s_{1,2} = \frac{1 \pm |2kM^2-1|}{2kM} = \frac{1}{2kM} = M$.
   - When $c > \frac{1}{2}$, we have $s_1 = \frac{1-|2kM^2-1|}{2kM} = \frac{1-(2kM^2-1)}{2kM} < \frac{1-(1-2kM^2)}{2kM} = M$, while $s_2 = \frac{1+|2kM^2-1|}{2kM} = \frac{1-(2kM^2-1)}{2kM} = M$. In this case, since service providers stop joining until any higher service level leads to negative profit, the steady-state service level is $s_2 = M$.

   Therefore, in this case $s^*_{ss} = M$.

3. $\frac{c}{M^2} < k \le \frac{1}{4M^2(1-c)}$. The necessary condition for $s_1 \le M$ is violated, so $s_1 > M$. since $s_2 > s_1$, we have $s_2 > M$ as well. Neither s_1 and s_2 is feasible.

4. $k > \frac{1}{4M^2(1-c)}$. Neither s_1 and s_2 is feasible.

Finally, it is easy to see that $y(s_{t+1}) < 0$ for all $s_{t+1} \in (0, M]$ for $k > \frac{c}{M^2}$, meaning that for any $s_{t+1} \in (0, M)$, the optimal service level in period t is greater than s_{t+1} by (3.4). M, however, is a steady-state service level because $s^*_t(M) = M$ by Lemma 3.2. Under $s^*_{ss} = M$, miner's profit would always be positive ($\phi(s^*_{ss}) = -y(s^*_{ss}) > 0$). □

BIBLIOGRAPHY

S. Alaei, A. Malekian, and M. Mostagir. A dynamic model of crowdfunding. Available at SSRN, 2016.

Y. Alan and V. Gaur. Operational investment and capital structure under asset-based lending. M&SOM, 2018.

A. Alexandrov and M. A. Lariviere. Are reservations recommended? M&SOM, 14(2): 218–230, 2012.

B. Arnold, N. Balakrishnan, and H. Nagaraja. A First Course in Order Statistics. SIAM, 2008.

K. J. Arrow, T. Harris, and J. Marschak. Optimal inventory policy. Econometrica: Journal of the Econometric Society, pages 250–272, 1951.

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

V. Babich, S. Marinesi, and G. Tsoukalas. Does crowdfunding benefit entrepreneurs and venture capital investors? M&SOM, forthcoming, 2019.

Y. Bakos and H. Halaburda. The role of cryptographic tokens and icos in fostering platform adoption. Working paper, 2018.

Bananacoin. Bananacoin official website. https://bananacoin.io/, 2018.

M. J. Beckmann. Decision and team problems in airline reservations. Econometrica, pages 134–145, 1958.

E. Belavina, S. Marinesi, and G. Tsoukalas. Rethinking crowdfunding platform design: mechanisms to deter misconduct and improve efficiency. Management Science, forthcoming, 2019.

P. Belobaba, A. Odoni, and C. Barnhart. The global airline industry. John Wiley & Sons, 2015.

H. Benedetti and L. Kostovetsky. Digital tulips? returns to investors in initial coin offerings. Working paper, 2018.

B. Biais, C. Bisiere, M. Bouvard, and C. Casamatta. The blockchain folk theorem. Toulouse School of Economics, TSE-817 Working Paper, 2017.

Blockstate. Global sto study. https://blockstate.com/global-sto-study-en/, 2020.

O. Boyabatlı and L. B. Toktay. Stochastic capacity investment and flexible vs. dedicated technology choice in imperfect capital markets. Management Science, 57(12):2163–2179, 2011.

O. Boyabatlı, T. Leng, and L. B. Toktay. The impact of budget constraints on flexible vs. dedicated technology choice. Management Science, 62(1):225–244, 2015.

BTS. Air travel consumer reports. US Department of Transportion, https://www.transportation.gov/airconsumer/air-travel-consumer-reports, 2018.

J. A. Buzacott and R. Q. Zhang. Inventory management with asset-based financing. Management Science, 50(9):1274–1292, 2004.

G. Cachon and C. Terwiesch. Matching Supply and Demand, 3rd ed. McGraw-Hill / Irwin, 2013.

G. P. Cachon and P. Feldman. Pricing capacity over time and recourse strategies: Facilitate reselling, offer refunds/options, or overbook? Working Paper, University of Pennnsylvania, 2018.

G. P. Cachon and R. Swinney. Purchasing, pricing, and quick response in the presence of strategic consumers. Management Science, 55(3):497–511, 2009.

G. P. Cachon and R. Swinney. The value of fast fashion: Quick response, enhanced design, and strategic consumer behavior. Management science, 57(4):778–795, 2011.

G. P. Cachon, S. Gallino, and M. Olivares. Does adding inventory increase sales? evidence of a scarcity effect in us automobile dealerships. Mgmt. Sci., 2018.

C. Catalini and J. S. Gans. Initial coin offerings and the value of crypto tokens. Working paper, 2018.

S. Chakraborty and R. Swinney. Sigkonaling to the crowd: Private quality information and rewards-based crowdfunding. Working paper, 2017.

S. Chakraborty and R. Swinney. Designing rewards-based crowdfunding campaigns for strategic contributors. Available at SSRN 3240094, 2018.

R. E. Chatwin. Multi-period airline overbooking with multiple fare classes. Naval Research Logistics (NRL), 43(5):603–612, 1996.

M. Cherif and S. Elouaer. Venture capital financing: A theoretical model. The Journal of Applied Business and Economics, 8(1):56, 2008.

J. Chod and E. Lyandres. A theory of icos: Diversification, agency, and information asymmetry. Working paper, 2018.

J. Chod, N. Trichakis, and G. Tsoukalas. Supplier diversification under buyer risk. Management Science, 2019a.

J. Chod, N. Trichakis, G. Tsoukalas, H. Aspegren, and M. Weber. On the financing benefits of supply chain transparency and blockchain adoption. Management Science, 2019b.

S. Chow, J. Shao, and H. Wang. Sample Size Calculations in Clinical Research. Chapman & Hall / CRC Biostatistics Series, 2nd ed., p. 85, 2008.

M. Cohen, A. Jacquillat, and J. Serpa. A field experiment on airline lead-in fares. Working Paper, 2019.

Coinschedule. Cryptocurrency ico stats 2018. https://www.coinschedule.com/stats.html?year=2018, 2018.

L. W. Cong, Y. Li, and N. Wang. Tokenomics: Dynamic adoption and valuation. Working paper, 2018.

CryptoGlobe. Only 13% of ico's have a running product, ernst & young report shows. https://bit.ly/300wfFR, 2018.

B. Curry. Optimal airline seat allocation with fare classes nested by origins and destinations. Trsp. Sci., 24:193–204, 1990.

J. D. Dana Jr and N. C. Petruzzi. Note: The newsvendor model with endogenous demand. Management Science, 47(11):1488–1497, 2001.

J. C. Ely, D. F. Garrett, and T. Hinnosaar. Overbooking. Journal of the European Economic Association, 15(6):1258–1301, 2017.

S. Fatehi, M. R. Wagner, et al. Crowdfunding via revenue-sharing contracts. Foundations and Trends® in Technology, Information and Operations Management, 10(3-4):407–424, 2017.

X. Fu, P. Gautier, and M. Watanabe. Overbooking in markets with rationing. Technical report, VU Amsterdam, 2012.

G. Gallego and Ö. Şahin. Revenue management with partially refundable fares. Op. Res., 58(4-part-1):817–833, 2010.

G. Gallego, S. Kou, and R. Phillips. Revenue management of callable products. Mgmt. Sci., 54(3):550–564, 2008.

J. R. Gan, N. Gans, and G. Tsoukalas. Overbooking with endogenous demand. The Wharton School Research Paper, 2019.

J. R. Gan, G. Tsoukalas, and S. Netessine. Initial coin offerings, speculators and asset tokenization. The Wharton School Research Paper, 2020.

H. L. Harter. Expected values of normal order statistics. Biometrika, 48(1):151–165, 1961.

J. Hernandez and C. Li. Video of united airlines passenger creates furor in china, too. https://www.nytimes.com/2017/04/11/world/asia/united-airlines-passenger-dragged-china.html, 2017.

F. J. Hinzen, K. John, and F. Saleh. Bitcoin's fatal flaw: The limited adoption problem. SSRN (3334262), 2019.

Honeypod. Honeypod white paper. https://docsend.com/view/6w2kpua, 2018.

D. A. Iancu, N. Trichakis, and G. Tsoukalas. Is operating flexibility harmful under debt? Management Science, 63(6):1730–1761, 2016.

I. Karaesmen and G. van Ryzin. Overbooking with substitutable inventory classes. Op. Res., 52(1):83–104, 2004.

Y. Khatri. Colorado senators file bill to exempt cryptos from securities laws. https://www.coindesk.com/colorado-senators-file-bill-to-exempt-cryptos-from-securities-laws, 2019.

R. Klophaus and S. Pölt. Airline overbooking considering passengers' willingness to pay. Trier University of Applied Sciences, 2010.

P. Kouvelis and W. Zhao. Financing the newsvendor: supplier vs. bank, and the structure of optimal trade credit contracts. Operations Research, 60(3):566–580, 2012.

P. Kouvelis, L. Dong, O. Boyabatli, and R. Li. Handbook of integrated risk management in global supply chains, volume 1. John Wiley & Sons, 2011.

V. Krishna. Auction Theory. Academic Press, 2010.

S. Kunnumkal, K. Talluri, and H. Topaloglu. A randomized linear programming method for network revenue management with product-specific no-shows. Trsp. Sci., 46(1):90–108, 2012.

kvstool.com. Kvs availability tool website. http://www.kvstool.com/, 2018.

Y. Lan, M. O. Ball, I. Z. Karaesmen, J. X. Zhang, and G. X. Liu. Analysis of seat allocation and overbooking decisions with hybrid information. EJOR, 240(2):493–504, 2015.

J. Li and W. Mann. Initial coin offering and platform building. Working paper, 2018.

J. Li, N. Granados, and S. Netessine. Are consumers strategic? structural estimation from the air-travel industry. Mgmt. Sci., 60(9):2114–2137, 2014.

K. Malinova and A. Park. Tokenomics: when tokens beat equity. Available at SSRN, 2018.

H. Martin. United says it'll pay bumped passengers up to $10,000 and reveals nine other policy changes. http://www.latimes.com/business/la-fi-united-fixes-20170426-story.html, 2017.

CB Insights. Venture capital funding report q2 2018. https://www.cbinsights.com/research/report/venture-capital-q2-2018/, 2018.

Cointelegraph. Ico markets 2018 vs 2017: Trends, capitalization, localization, industries, success rate. https://bit.ly/2XrG66b, 2019.

Kickstarter. Kickstarter stats. https://www.kickstarter.com/help/stats, 2019.

PitchBook. 18 charts to illustrate us vc in 2018. https://pitchbook.com/news/articles/18-charts-to-illustrate-us-vc-in-2018, 2019.

Sirin Labs. Sirin labs official website. https://sirinlabs.com/, 2019.

F. Memoria. $9 million: Ether stolen from coindash ico after website compromise. https://www.ccn.com/110991-2/, 2018.

S. Mihm. The airline industry: Overpacking planes since the 1940s. Bloomberg View, https://www.bloomberg.com/view/articles/2017-04-14/the-airline-industry-overpacking-planes-since-the-1940s, 2017.

J. M. Milner and P. Kouvelis. Inventory, speculation, and sourcing strategies in the presence of online exchanges. Manufacturing & Service Operations Management, 9(3):312–331, 2007.

S. Netessine and R. A. Shumsky. Revenue management games: Horizontal and vertical competition. Mgmt. Sci., 51(5):813–831, 2005.

E. Pagnotta and A. Buraschi. An equilibrium valuation of bitcoin and decentralized network assets. Available at SSRN 3142022, 2018.

E. S. Pagnotta. Bitcoin as decentralized money: prices, mining, and network security. SSRN (3264448), 2018.

Y. Papanastasiou and N. Savva. Dynamic pricing in the presence of social learning and strategic consumers. Management Science, 63(4):919–939, 2016.

D. Pozzi. September ico market overview: Trends, capitalization, localization. https://cointelegraph.com/news/september-ico-market-overview-trends-capitalization-localization, 2018.

I. Rosu and F. Saleh. Evolution of shares in a proof-of-stake cryptocurrency. Working Paper, 2019.

M. Rothstein. An airline overbooking model. Trsp. Sci., 5(2):180–192, 1971.

R. Sano. Dynamic mechanism design with overbooking, different deadlines, and multi-unit demands. Working Paper, Institute of Economic Research, Kyoto University, 2017.

S. Seth. 50% of icos failed to raise funds since 2017: Report. https://www.investopedia.com/news/50-icos-failed-raise-funds-2017-report/, 2018.

M. Shaked and J. G. Shanthikumar. Stochastic Orders and Their Applications. Academic Press, 1994.

S. Shifflett and C. Jones. Buyer beware: Hundreds of bitcoin wannabes show hallmarks of fraud. https://on.wsj.com/2MkUrLj, 2018.

C. Smith, J. Leimkuhler, and R. Darrow. Yield management at american airlines. Interfaces, 22:8–31, 1992.

X. Su. Optimal pricing with speculators and strategic consumers. Management Science, 56 (1):25–40, 2010.

X. Su and F. Zhang. Strategic customer behavior, commitment, and supply chain performance. Mgmt. Sci., 54(10):1759–1773, 2008.

K. Talluri and G. van Ryzin. The Theory and Practice of Revenue Management. Springer, 2004.

thepointsguy.com. The points guy website. https://thepointsguy.com/2016/12/overbooking-airlines-explained-ted-talks-video/, 2018.

D. Thomson. Us IPO activity showed continued strength: Q2 2018 capital markets watch. http://usblogs.pwc.com/deals/q2-2018-capital-markets-watch/, 2018.

travelersunited.org. Travelers united website. https://www.travelersunited.org/policy-columns/dot-statistics-only-tell-part-of-the-story/, 2018.

G. Tsoukalas and B. H. Falk. Token-weighted crowdsourcing. Management Science, 2019.

B. Tuttle. Delta just dramatically increased the compensation it can give to bumped passengers. http://time.com/money/4741108/delta-united-overbooked-flight-compensation-vouchers/, 2017.

F. von Wangenheim and T. Bayón. Behavioral consequences of overbooking service capacity. Journal of Marketing, 71(4):36–47, 2007.

S. Wang and H. Zhou. Staged financing in venture capital: moral hazard and risks. Journal of corporate finance, 10(1):131–155, 2004.

R. A. Wijsman. A useful inequality on ratios of integrals, with application to maximum likelihood estimation. JASA, 80(390):472–475, 1985.

F. Xu and F. Zhang. Crowdfunding under social learning and network externalities. Working paper, 2018.

F. Xu, X. Guo, G. Xiao, and F. Zhang. Crowdfunding or bank financing: Effects of market uncertainty and word-of-mouth communication. Working paper, 2018.

S. A. Yang and J. R. Birge. How inventory is (should be) financed: Trade credit in supply chains with demand uncertainty and costs of financial distress. Working paper, 2013.

S. A. Yang, N. Bakshi, and C. Chen. Cancelability in trade credit insurance. Working paper, 2016.

B. Zhang. United airlines will soon let you name your price for giving up your seat on an overbooked flight. https://www.businessinsider.com/united-to-launch-auction-system-to-buy-back-seats-2017-9, 2017.