Platform minimum requirements[*]

Andrei Hagiu[†] and Julian Wright[‡]

October 2018

Abstract

We study how much revenue a principal (e.g. a platform) should extract from an agent (e.g. a third-party supplier) and how much control it should grant the agent over decisions that it can monitor (e.g. the third-party supplier's provision of certain services). These two contracting choices are tightly linked: giving the agent more control over costly decisions goes hand-in-hand with leaving the agent with a higher share of revenue. We study the full range of delegation possibilities facing the principal, and explain why the frequently observed practice of granting the agent control over costly decisions, subject to minimum requirements, is often the best option. Our analysis applies to the contracting choices facing digital platforms, licensors, franchisors, and shopping malls, among other examples. When applied to pricing decisions, it provides a new theory of resale price maintenance, which explains when price ceilings or price floors should be used.

Keywords: platforms, governance, partial delegation, resale price maintenance.

1 Introduction

With increased monitoring capabilities in digital settings, platforms (e.g. Airbnb, Didi, Lyft, Uber, Oyo) increasingly face the choice of how much control to give their third-party suppliers with respect to the features they offer, the equipment used to provide their services, refunds, delivery options, and various types of service standards. This choice is also relevant for more traditional businesses such as franchisors, licensors and shopping malls. For instance, a traditional business-format franchisor such as a hotel chain, a fast-food restaurant or a car rental company has to decide how much to control local advertising choices. The franchisor could decide the level of local advertising itself and write it into the franchise contract, delegate the choice entirely to the franchisees, or let the franchisees decide

[†]Boston University Questrom School of Business, Boston, MA 02215, E-mail: ahagiu@bu.edu

[‡]Department of Economics, National University of Singapore, Singapore 117570, E-mail: jwright@nus.edu.sg

subject to a minimum advertising expense. A shopping mall faces a similar choice with respect to the opening hours for shops in its mall.

We build a theoretical model to evaluate the optimal level of control that a principal should give an agent over a transferable costly decision. Our analysis takes into account that the contract with the agent will also determine how much of the agent's ongoing revenue is extracted by the principal. We compare partial delegation—in which the principal retains some control—with both full control by the principal and full delegation to the agent. In particular, we determine when threshold delegation—in which the principal imposes a minimum threshold on the agent's choice of action—does better than either full control or full delegation. We also provide sufficient conditions for threshold delegation to be the optimal form of partial delegation.

In our baseline model, demand is determined by (i) price, which is always chosen by the agent, (ii) two ongoing and costly decisions, and (iii) a random demand shock privately observed by the agent. We assume one of the costly decisions remains non-contractible and under the principal's control, while the other costly decision can be chosen by either the principal or the agent, i.e. it is contractible (e.g. the services and amenities offered by Airbnb hosts, or the level of local advertising by a franchisee). Throughout we assume the principal can make use of a two-part tariff—a fixed fee (or fixed payment) and a wholesale fee for each unit sold—in its contract. Such two-part tariffs are commonly observed in the types of examples we consider.

At a high level, we study the interaction between two instruments: the wholesale price paid by the agent to the principal, and the allocation of control over transferable decision variables between them. Both are decided by the principal as part of her contract choice. To incentivize the principal to choose positive levels of her non-transferable investment, the wholesale price must always be positive. However, a positive wholesale price means that the agent only receives a fraction of the full revenues resulting from its choice of this costly action. This results in the agent's choice of the costly action to be distorted downward. The wholesale price endogenously determines the magnitude of the agent's bias (downward distortion) in choosing the transferable decision variable. In turn, the bias determines the extent to which the principal wishes to delegate control of this variable to the agent.

We show that the principal often does best by restricting the agent's choice of the costly action to be above some minimum threshold. Thus, our theory can explain the imposition of minimum requirements such as Lyft and Uber's minimum requirements for the quality of cars that drivers can use, a franchisor's minimum requirement for the level of franchisee advertising or investment, a manufacturer's minimum requirement for an authorized dealer's in-store promotion or investments, a shopping mall's minimum requirement for a retailer's opening hours.

Threshold delegation is a way to get some of the commitment benefit of the principal determining the level of the costly action in question (and thus reduce the need to set a low wholesale price so as to leave a large share of revenue with the agent), while also preserving some of the agent's ability to respond to his private information about demand. At the same time, it avoids the worst biases in setting costly actions that arise when the agent controls the costly actions but only keeps some of the associated variable revenue. In the context of platforms, this tradeoff can be interpreted as one

between platform intervention (governance rules) in order to correct for agents' insufficient investment incentives on the one hand, and freedom for platform participants in order to utilize their dispersed knowledge on the other hand.

We also show how the boundaries between full control, full delegation and threshold delegation change with key parameters. When the principal's moral hazard becomes more important or the agent's private information becomes less important, full control becomes relatively more attractive than threshold delegation, which in turn becomes relatively more attractive than full delegation.

As an extension of the model, we consider an alternative scenario in which the transferable decision is price, thus assuming that the agent's price can be monitored by the principal and resale price maintenance (RPM) is allowed. In this case, the direction of the bias depends on whether demand increases or decreases when price is higher. In the standard case that demand decreases in price, the agent's pricing decision is biased upwards, reflecting the effect of double marginalization. This happens when the agent's moral hazard is not too important, in which case RPM with a maximum threshold is desirable. However, if the agent's moral hazard is sufficiently important, demand will increase in price. This reflects that a higher price induces the agent to choose a higher level of investment (or effort) to such an extent that it more than offsets the direct effect of the higher price on demand. In this case, the agent's pricing decision suffers from a downward bias, in which case RPM with a minimum threshold is desirable.

2 Related literature

Our paper combines elements from various literatures.

Our main contribution is to the emerging literature in information systems and strategy on platform governance, i.e. non-price rules and restrictions employed by platform owners to regulate the access and behavior of platform participants (Boudreau and Hagiu, 2009, Parker et al., 2016, Parker and Van Alstyne, 2017). Specifically, the minimum requirements placed by the principal on the agent's choice of investment can be viewed as a form of behavior governance whenever the principal is a platform provider (e.g. Airbnb, Apple, Didi, Google, Lyft and Uber, as discussed in Section 3). Previous studies have focused on other non-price governance instruments used by platforms: how much technology to share with platform participants (Boudreau, 2010, Parker, Van Alstyne, 2017, Niculescu et al., 2018) and how strong to make property rights given to platform participants (Parker and Van Alstyne, 2017). At a high level, the decision of how much freedom to give platform participants to choose their investments is also related to the choice faced by a software originator between making his software proprietary (P-mode in our model) or making it open-source (A-mode), as studied by August et al. (2013) and (2018). These papers find that if a contributor is efficient in software development, the originator should adopt an open source strategy, allowing the contributor to offer higher total quality. Conversely, if the contributor is not efficient in development, the originator should adopt a proprietary software development strategy, gaining revenue from software sales and squeezing the contributor out of the services market. These findings are related to our results that the principal (platform) should

give more control to an agent when the latter has more important private information and less control when the investments by the principal and the agent are more important.

The choice between full control and full delegation in our model is reminiscent of the classic choice studied by Simon (1951) between contracting on a decision ex-ante (before uncertainty is resolved) vs. giving full authority to the employer (principal) or the employee (agent) to unilaterally choose the decision ex-post. In our model, giving full authority to the principal to unilaterally make the transferable decision ex-post is never optimal. This is because the principal never observes the realization of the agent's private information and can always extract the entire surplus from the agent through its two-part tariffs—this means the principal can always do better by committing to the choice of the transferable action ex-ante. A related and more recent strand of literature has emerged that studies conditions under which platforms and other intermediaries take control over transferable decisions pertaining to the sale of products to end-consumers or allow their suppliers/complementors to keep control over these decisions. Desiraju and Moorthy (1997), Jerath and Zhang (2010), and Hagiu and Wright (2015, 2018) study delegation of both price and costly investment (e.g. service) decisions.

The key novelty that we introduce relative to all the articles mentioned so far is that we allow for an intermediate option between fixing the transferable decision in the principal's ex-ante contract and giving full authority to the agent: the agent can be given authority to choose the transferable action subject to restrictions imposed by the principal's ex-ante contract. This is known as "partial delegation" following the seminal work by Holmstrom (1977, 1984). Several papers have proven that threshold delegation is optimal in similar settings—see for example, Melumad and Shibano (1991), Martimort and Semenov (2006), Alonso and Matouschek (2008), and Amador and Bagwell (2013). We directly show that threshold delegation is optimal in our benchmark setting under fairly general conditions on the distribution of private information. There are three key contributions in our model relative to the partial delegation literature: (1) we allow for monetary transfers between the principal and the agent in the form of two-part tariffs set by the principal in the contracting stage; (2) we introduce double-sided moral hazard; and (3) the bias of the agent's objective function relative to the principal's is endogenously determined by the two-part tariff, which in turn depends on the importance of the agent's moral hazard relative to the principal's, and on the importance of the agent's private information. By contrast, the partial delegation literature to date assumes an exogenously given bias, no transfers between principal and agent and no moral hazard for either the principal or the agent.

In retail contexts, threshold delegation can be viewed as an additional instrument that can help improve channel coordination. Our modelling approach is entirely consistent with the principal-agent view of channel coordination taken by the marketing and management literature to date (see Lal, 1990 and Cachon and Lariviere, 2005). However, this literature has focused on improving channel coordination through various payment instruments such as revenue sharing, slotting fees, quantity discounts and buy-backs, and/or through monitoring, which is modelled as enforcing a specific level of a non-contractible investment in service. We extend this work by showing that the addition of threshold delegation with respect to a transferable costly decision variable can improve channel coordination when conditioning payments directly on the level of this costly action is not viable. Threshold

delegation is also commonly used in practice (see Section 3 below).

Since in our model revenues must be shared between the principal and the agent to incentivize both sides to make non-contractible investments, we also directly build upon principal-agent models with double-sided moral hazard (Lal, 1990, Romano, 1994, and Bhattacharyya and Lafontaine, 1995). The key difference relative to these papers is that our model can explain *partial* delegation of a transferable action, which is frequently observed in practice. Indeed, the transferable action is entirely absent from the models of Lal (1990) and Bhattacharyya and Lafontaine (1995), while Romano (1994) does not allow for any private information, which means there is no scope for partial delegation. In particular, our model can be viewed as an extension of Romano's to allow for the agent to have private information. Furthermore, Romano (1994) focuses on price as the transferable action, whereas we focus mostly on the case in which the transferable action is a costly investment.

3   Examples

There are a wide variety of examples that our theory can be applied to. Table 1 summarizes a few key examples, listing decisions that are transferable and potentially subject to minimum requirements, decisions that are non-transferable and subject to moral hazard, and the source of the agent's private information.

Table 1: Examples

	Transferable decisions (possibly subject to restrictions)	*Non-transferable investment decisions made by the principal*	*Source of agent's private information*
App stores	app licensing terms	technological upkeep (e.g. payment) and advertising of app store	revenues and cross-selling opportunities outside of app
Transportation platforms (Didi, Lyft, Uber)	quality of car	technological upkeep (e.g. payment, dispatch) and advertising of service	repeat business for the driver off the platform; cash tips
Accommodation platforms (Airbnb, Oyo)	room amenities	design and features of the platform	travelers' local demand
Shopping malls	retailer's opening hours	maintenance and advertising of mall	retailer's demand, costs and outside revenues that originate from mall traffic
Franchising	local advertising of the outlet; opening hours;	national advertising of the brand	local demand; effectiveness of local advertising; revenue from cross-selling other products; franchisee's costs
Technology licensing	advertising of product	investment in improving the technology	demand and costs for product
Manufacturer and authorized dealer	investment in quality of outlet; local promotion and advertising;	dealer's effort	local demand; revenue from cross-selling other products or services; dealer's costs

Our theory applies to an increasing number of platforms, both online and offline. Consider four examples: digital app stores (e.g. Apple's App Store for iPhone apps and Google's Play Store for Android apps), ride-hailing platforms (e.g. Didi, Lyft and Uber), accommodation platforms (e.g.

Airbnb, Oyo), and shopping malls. All four types of platforms place minimum requirements on important transferable decisions. Apple and Google place minimum requirements on the terms of the licensing agreement provided by app developers to their users.[1] For instance, both Apple and Google require developers to assume sole responsibility for any defects or performance issues related to their apps and Google requires developers to respond to customer support inquiries within three business days. UberX and Lyft drivers have to use cars that satisfy a minimum age requirement (e.g. 2001 or newer in many cities for UberX, and 2004 or newer in many cities for Lyft). The two companies also impose minimum requirements on the cars' functionality (e.g. 4 doors, at least 5 seat belts) and on their state of maintenance (e.g. fully functioning A/C and heating, no major cosmetic damage). In contrast, traditional taxi companies completely control and incur the costs corresponding to the choice of cars used by their drivers. In recent years, Airbnb has started putting more pressure on its hosts to comply with minimum standards that make their homes more comparable to hotels.[2] And in early 2018, the company launched Airbnb Plus, a brand for accomodations on its platform that are guaranteed to meet 100 different criteria and come with certain amenities.[3] Meanwhile, Oyo, which recently raised one billion U.S. dollars for its hotel booking platform based in India, differentiates itself by catering to budget travellers but putting strict minimum requirements on the facilities offered by member hotels, including such details as the minimum time period that the complimentary breakfast is available, the minimum thickness of the mattresses, and the minimum size of showerheads in bathrooms.[4] Finally, in the case of shopping malls, the lease agreements often specify minimum opening hours for retailers (those hours could be set by the mall or by each respective retailer).

Other applications include "business format" franchising (e.g. hotel, fast-food, car rental, etc.). The franchisor is the principal in our setting, while the franchisee is the agent. One can often distinguish national from local advertising, with the latter being a decision that could be made either by the franchisor or the franchisee. For local advertising, contracts often specify a minimum spending requirement by the franchisee. Other decisions that are typically chosen by franchisees subject to minimum requirements imposed by franchisors include the number of staff that have to be on-site at various days/times, cleanliness, and opening hours. Technology licensing is similar: it may involve ongoing investments in technological improvements by the licensor, and sometimes minimum levels of certain investments (or services) that must be provided by the licensee. For instance, Google requires all Android licensees (device manufacturers) to undergo costly testing by third-parties.[5]

Finally, branded manufacturers that distribute their products through authorized dealers provide another large set of applications for our theory. For instance, manufacturers oftentimes impose minimum standards for retail premises and minimum advertising or promotion levels by the retailers, but these same transferable decision variables can also sometimes be stipulated by the manufacturer or left

---

[1]See for instance http://www.apple.com/legal/internet-services/itunes/appstore/dev/minterms/

[2]See https://www.nytimes.com/2017/06/17/technology/airbnbs-hosts-professional-hotels.html

[3]See https://www.nytimes.com/2018/02/22/travel/new-airbnb-plus.html
and https://www.airbnb.com.sg/plus.

[4]See https://www.bloomberg.com/news/articles/2018-09-25/24-year-old-oyo-founder-raises-1-billion-to-fund-overseas-push?srnd=premium-asia

[5]See https://www.theguardian.com/technology/2014/jan/23/how-google-controls-androids-open-source.

unrestricted. Likewise, a manufacturer that uses sales representatives would face a similar situation—the extent to which it controls their transferable decisions (and therefore the extent to which sales representatives would be considered employees or independent contractors).

4 Model set-up

We assume the demand $D(p, q, Q)$ generated by a principal and an agent is determined by the choice of three decision variables: the price p and two costly actions q and Q. Throughout the paper, Q is a costly, non-transferable and non-contractible action always chosen by the principal. It captures the on-going investments made by the principal that increase demand. Examples are given in the third column of Table 1.

Except when we analyze the possibility of RPM in Section 6.1, we assume that the price p is non-contractible and always chosen by the agent, while the costly action q is a transferable decision variable which the principal can place restrictions on. Examples of such actions are given in the second column of Table 1.

Throughout we rule out payments contingent on the agent's price or its costly decision variable, and instead assume that the principal's contract must be a standard two-part tariff, with the only additional instrument being the possibility of restricting the agent's choice of the transferable decision variable. Such two-part tariffs are widespread and are used in many of the examples discussed in Section 3.

The principal makes a take-it-or-leave-it offer to the agent, which involves a per-unit wholesale price w and a fixed participation fee F. According to this contract, the principal receives $wD(p, q, Q) + F$ from the agent, with the agent retaining $(p - w) D(p, q, Q) - F$. The fixed fee F will always be set to leave the agent indifferent between participating or not. In practice, F could be negative (the agent may be paid a fixed salary to participate) but without loss of generality we normalize the agent's outside option to zero, which means the optimal F will always be positive in our model.

With a two-part tariff, the principal cannot fully elicit the agent's private information, which would be required to obtain the best achievable payoffs. As a result, the principal may want to give full or partial control over the choice of the transferable variable to the agent—this will be specified in the contract, along with the two-part tariff (w, F). If the contract directly specifies the choice of the transferable decision, we say that the principal has chosen the P-mode to reflect that only the principal determines it. If the contract leaves the agent free to set the transferable decision with no restrictions, we say that the principal has chosen the A-mode to reflect that only the agent determines it. Finally, if the contract partially restricts the agent's choice of the transferable decision to be above or below a certain threshold, we say that the principal has chosen the H-mode, which is a hybrid of the two pure modes in that both the principal and agent determine the transferable decision.[6]

---

[6]In our setup, the fourth logically possible delegation option—in which the principal maintains control over the transferable decision but only chooses it in the second stage instead of fixing it in its contract—is always dominated. This is because the principal never observes any private information, so there is no benefit in waiting rather than committing to the choice of the transferable decision ex-ante.

For most of the paper we assume that demand is linear in these variables, and can be written as

$$D(p, q, Q) = \theta - \beta p + \phi q + \Phi Q,$$

where β, ϕ and Φ are positive constants, measuring the impact of p, q and Q, respectively, on demand, and θ is an additive demand shock observed only by the agent.[7] Formally, we assume θ is a random variable drawn from the distribution function G, with positive density $g(.)$ over $[\theta_L, \theta_H]$, with finite mean $\mathbb{E}(\theta) = \bar{\theta} > 0$ and variance V_θ. We assume $0 \leq \theta_L < \theta_H$, but do not require that θ_H is finite, so we allow for distributions with unbounded support on the right tail (such as the exponential or normal distributions). The distribution function $G(.)$ is twice continuously differentiable. We assume the fixed costs of the respective costly actions are $\frac{1}{2}q^2$ and $\frac{1}{2}Q^2$. The total revenue net of fixed costs generated by the principal and the agent is therefore

$$p(\theta - \beta p + \phi q + \Phi Q) - \frac{1}{2}q^2 - \frac{1}{2}Q^2.$$

In Section 6.2 we show that our main results hold for more general demand and cost functions.

The timing of the players' moves is as follows: In the first stage, the principal offers its contract (which can include some restrictions on the choice of transferable decision) and the agent decides whether or not to accept the contract. If the contract is accepted, in the second stage the agent observes the realization of θ and decides on the transferable decision, subject to any restrictions in the contract. In the second stage, the agent also chooses p, while the principal always chooses Q. Finally, payoffs are realized.

Note that for the problem we consider to be an interesting one, there must be some principal moral hazard (i.e. $\Phi > 0$). If instead $\Phi = 0$, then the principal could attain the first-best outcome by always fully delegating the choice of the transferable decision to the agent, combined with a simple fixed fee that extracts the agent's entire expected profit in the first stage. This provides the agent with first-best incentives to invest in q and fully exploits the agent's private information.

5   Minimum requirements

Given the additive nature of demand, θ has no direct impact on the agent's choice of q (indeed, in the absence of any constraints, the agent chooses $q = (p - w)\phi$), so one may wonder why the principal would ever want to delegate the choice of q to the agent. The reason is that the agent chooses p, which is positively impacted by θ and is a strategic complement to q. Thus, delegation of q to the agent allows the principal to leverage the agent's private information on the demand shock, which indirectly affects the agent's choice of q. However, this benefit must be traded off against the inefficiency due to the fact that the agent sets q based on a lower margin ($p - w$ rather than p), and so tends to

---

[7]Although we focus on a demand shock, in Section 6.3 we show that the logic of our analysis also extends to other sources of private information such as those noted in the last column of Table 1. In order to keep the model complexity to a minimum, we do not allow the possibility that the principal also has some private information about demand conditions. This would obviously increase the relative advantage of the principal taking control of the transferable decision.

underinvest in q all other things equal.

In P-mode, whether the cost of q (i.e. $\frac{1}{2}q^2$) is incurred by the principal or the agent is immaterial to the outcome, because q is set contractually and the principal can use a two-part tariff to extract the agent's entire expected surplus in excess of a fixed outside option. This also implies that if one party has lower costs of carrying out q, then that party should always incur the cost of q, regardless of which party actually controls the level of q. For example, franchisees are often required to make certain investments, reflecting that the franchisee is better placed to carry them out, even though the level (or minimum level) of investment is specified by the franchisor. Since in our model the costs of q are assumed to be the same regardless of which party incurs them, we can assume without loss of generality that the cost of q is incurred by the party that controls its level.

We make two assumptions, which ensure all second-order conditions hold and all decision variables and profits are positive at equilibrium values:

$$\left(2\beta - \phi^2\right)\left(\frac{\beta}{2} + \Phi^2\right) - \Phi^4 > 0 \tag{1}$$

and

$$\frac{\theta_L}{\theta} > \max\left\{ \frac{\Phi^2\left(\beta - \Phi^2\right) - \phi^2\left(\frac{\beta}{2} + \Phi^2\right)}{\left(2\beta - \phi^2\right)\left(\frac{\beta}{2} + \Phi^2\right) - \Phi^4}, \frac{\Phi^2\left(\beta - \Phi^2\right)}{\left(2\beta - \phi^2\right)\left(\frac{\beta}{2} + \Phi^2\right) + \frac{\beta}{2}\phi^2 - \Phi^4} \right\}. \tag{2}$$

Assumption (1) ensures second-order conditions hold for all optimization problems we consider. It requires β is sufficiently large or Φ^2 is sufficiently small. Assumption (2) ensures second-stage profits and decision variables are positive for all realizations of θ. This assumption always holds if $\Phi^2 \geq \beta$, or, in case $\beta > \Phi^2$, if θ_L is not too small.

We first analyze whether the principal prefers to set the level of q in its contract (P-mode) or entirely delegate that choice to the agent (A-mode), before considering whether the principal can do better than both pure modes through threshold delegation (H-mode).

5.1 Full control vs. no control

Consider first the P-mode. The fixed fee F is set to extract the entire expected surplus from the agent, so the principal solves

$$\max_{w,q} \mathbb{E}\left[p\left(\theta - \beta p + \phi q + \Phi Q\right) - \frac{1}{2}q^2 - \frac{1}{2}Q^2\right]$$

subject to

$$p = \arg\max_{p'}\left\{\left(p' - w\right)\left(\theta - \beta p' + \phi q + \Phi Q\right) - \frac{1}{2}q^2\right\} = \frac{1}{2}w + \frac{1}{2\beta}\left(\theta + \phi q + w\Phi^2\right)$$

$$Q = \arg\max_{Q'}\left\{\mathbb{E}\left[w\left(\theta - \beta p + \phi q + \Phi Q'\right)\right] - \frac{1}{2}Q'^2\right\} = w\Phi.$$

Substituting p and Q back into the principal's objective function and maximizing over q implies

that, for a given w, the principal's optimal choice of q is $q^P\left(w, \overline{\theta}\right)$, where

$$q^P\left(w, \theta\right) = \frac{\left(\theta + w\Phi^2\right)\phi}{2\beta - \phi^2}.$$

Thus, $q^P\left(w, \overline{\theta}\right)$ is the hypothetical level of q that the principal would choose for a given w if it were able to observe θ. Plugging $q^P\left(w, \overline{\theta}\right)$ back into the principal's objective function and optimizing over w, we obtain that the optimal wholesale price w in P-mode is

$$w^{P*} = \frac{\Phi^2\overline{\theta}}{\left(2\beta - \phi^2\right)\left(\frac{\beta}{2} + \Phi^2\right) - \Phi^4},$$

which is positive given (1). Furthermore, (1) ensures that the agent's margin is always positive when the principal sets $w = w^{P*}$, i.e. $p^P\left(w^{P*}, \theta\right) > w^{P*}$ for all θ. The principal's resulting profits are

$$\Pi^{P*} = \frac{V_\theta}{4\beta} + \frac{\left(\frac{\beta}{2} + \Phi^2\right)\overline{\theta}^2}{2\left(\left(2\beta - \phi^2\right)\left(\frac{\beta}{2} + \Phi^2\right) - \Phi^4\right)}. \tag{3}$$

Consider now the A-mode. The principal solves

$$\max_{w}\left\{\mathbb{E}\left[p\left(\theta - \beta p + \phi q + \Phi Q\right) - \frac{1}{2}q^2 - \frac{1}{2}Q^2\right]\right\}$$

subject to

$$(p, q) = \arg\max_{p', q'}\left\{\left(p' - w\right)\left(\theta - \beta p' + \phi q' + \Phi Q\right) - \frac{1}{2}q'^2\right\}$$

$$Q = \arg\max_{Q'}\left\{\mathbb{E}\left[w\left(\theta - \beta p + \phi q + \Phi Q'\right)\right] - \frac{1}{2}Q'^2\right\}.$$

Solving the three constraints in (p, q, Q) implies

$$p^A\left(w, \theta\right) = \frac{\theta + w\left(\beta + \Phi^2 - \phi^2\right)}{2\beta - \phi^2}$$

$$q^A\left(w, \theta\right) = \frac{\left(\theta + w\left(\Phi^2 - \beta\right)\right)\phi}{2\beta - \phi^2}$$

$$Q^A\left(w\right) = w\Phi.$$

Note $q^A\left(w, \theta\right) < q^P\left(w, \theta\right)$ for all w and θ, which means the agent always has a downward bias in setting q. This reflects that (i) q is a costly investment, and (ii) when choosing q in A-mode, the agent only internalizes a fraction of the revenue given $w > 0$, which is necessary to ensure the principal remains incentivized to invest in Q. By contrast, when the principal sets q in the P-mode contract, it takes into account the full revenue created by this investment.

Plugging the expressions of $p^A\left(w, \theta\right)$, $q^A\left(w, \theta\right)$ and $Q^A\left(w\right)$ back into the principal's objective

function, we obtain that the optimal wholesale price set by the principal in A-mode is

$$w^{A*} = \frac{\Phi^2 \overline{\theta}}{(2\beta - \phi^2)\left(\frac{\beta}{2} + \Phi^2\right) + \frac{\beta}{2}\phi^2 - \Phi^4}.$$

It is easily verified that (1) implies $w^{A*} > 0$, while (2) ensures $p^A\left(w^{A*}, \theta\right) > w^{A*}$ and $q^A\left(w^{A*}, \theta\right) > 0$ for all θ. Note also that $0 < w^{A*} < w^{P*}$, so the principal sets a lower wholesale price when it lets the agent control the choice of q.

The principal's optimal A-mode profit is then

$$\Pi^{A*} = \frac{V_\theta}{2(2\beta - \phi^2)} + \frac{\left(\Phi^2\left(2\beta - \phi^2\right) + \beta^2\right)\overline{\theta}^2}{2(2\beta - \phi^2)\left((2\beta - \phi^2)\left(\frac{\beta}{2} + \Phi^2\right) + \frac{\beta}{2}\phi^2 - \Phi^4\right)}. \tag{4}$$

Proposition (1) then follows from a comparison of (3) and (4).

Proposition 1. *The principal's profit is higher in A-mode compared to P-mode if and only if the variance of the agent's private information on demand is sufficiently large, i.e.*

$$\frac{V_\theta}{\overline{\theta}^2} > \frac{2\Phi^4\beta^2}{\left((2\beta - \phi^2)(\beta + 2\Phi^2) - 2\Phi^4\right)\left(\Phi^2\left(2\beta - \phi^2\right) + \beta^2 - \Phi^4\right)}. \tag{5}$$

The inequality in (5) captures the key tradeoff between the two pure modes. On the one hand, the A-mode leverages the agent's private information on demand as captured by V_θ. On the other hand, the P-mode removes the distortion created by the agent setting q in A-mode, which explains why the right-hand side of (5) is positive.

5.2   Threshold delegation

Now suppose in addition to using a two-part tariff, the principal can monitor q and therefore restrict the agent's choice of q according to some rule (i.e. H-mode). In this case, the principal could restrict the agent's choice of q to a degenerate interval $\{q_0\}$ that only contains one point—this effectively replicates the P-mode where the principal sets $q = q_0$ in its contract. At the other extreme, the principal's restriction could be so lax that it places no effective constraint on the agent's choice of q—this replicates the A-mode. For the sake of clarity, we will only use the label H-mode when the principal's restriction is neither one of these two extremes, but instead places some partial restriction on the agent's choices. Otherwise, we will refer to the contract choice as P-mode or A-mode given the equivalence noted above.

Throughout the paper we focus on partial restrictions that are threshold rules. A threshold rule is one in which the principal restricts the agent's choice of q to be above or below a certain threshold. We first determine sufficient conditions for the H-mode with threshold delegation to dominate both the A-mode and the P-mode. Subsequently, we will provide a sufficient condition for threshold delegation to be the optimal form of partial delegation. Even when threshold delegation is not the optimal form

11

of partial delegation, it does have the advantage of being simple to write down in a contract and relatively easy to monitor (as opposed to, for example, delegation that involves multiple intervals). This explains why threshold delegation is often used in practice and justifies our focus on it here.

As pointed out in Section 5.1, the agent has a downward bias in A-mode relative to what the principal would set in P-mode. Thus, the relevant form of threshold delegation is that with a minimum threshold.

Given a wholesale price w and a minimum threshold x imposed on q, the agent chooses

$$
q = \begin{cases} q^A\left(w,\theta\right) = \frac{\left(\theta + w\left(\Phi^2 - \beta\right)\right)\phi}{2\beta - \phi^2} & \text{if} \quad \theta \geq \frac{2\beta - \phi^2}{\phi}x + w\left(\beta - \Phi^2\right) \\ \\ x & \text{if} \quad \theta \leq \frac{2\beta - \phi^2}{\phi}x + w\left(\beta - \Phi^2\right). \end{cases}
$$

As in the pure modes, the principal extracts the agent's entire expected payoff through the fixed fee, so the principal's profit is

$$
\max_{w,x}\left\{\mathbb{E}_\theta\left[p\left(\theta - \beta p + \phi q + \Phi Q\right) - \frac{1}{2}q^2 - \frac{1}{2}Q^2\right]\right\}
$$

subject to

$$
p = \arg\max_{p'}\left\{\left(p' - w\right)\left(\theta - \beta p' + \phi q + \Phi Q\right) - \frac{1}{2}q^2\right\} = \frac{1}{2}w + \frac{1}{2\beta}\left(\theta + \phi q + w\Phi^2\right)
$$

$$
q = \max\left\{q^A\left(w,\theta\right), x\right\}
$$

$$
Q = \arg\max_{Q'}\left\{\mathbb{E}_\theta\left[w\left(\theta - \beta p + \phi q + \Phi Q'\right) - \frac{1}{2}Q'^2\right]\right\} = w\Phi.
$$

Plugging these constraints into the principal's objective function, we obtain $\Pi^{H*} = \max_{w,x}\Pi^H\left(w,x\right)$, where

$$
\Pi^H\left(w,x\right) \equiv \int_{\theta_L}^{\frac{2\beta - \phi^2}{\phi}x + w\left(\beta - \Phi^2\right)}\left(\frac{1}{4\beta}\left(\theta + \phi x + w\left(\Phi^2 - \beta\right)\right)^2 + \frac{w}{2}\left(\theta + \phi x - \beta w\right) - \frac{1}{2}x^2\right)dG\left(\theta\right)
$$

$$
+ \int_{\frac{2\beta - \phi^2}{\phi}x + w\left(\beta - \Phi^2\right)}^{\theta_H}\left(\frac{1}{4\beta}\left(\theta + \phi q^A\left(w,\theta\right) + w\left(\Phi^2 - \beta\right)\right)^2 + \frac{w}{2}\left(\theta + \phi q^A\left(w,\theta\right) - \beta w\right) - \frac{1}{2}q^A\left(w,\theta\right)^2\right)dG\left(\theta\right).
$$

If $\frac{2\beta - \phi^2}{\phi}x + w\left(\beta - \Phi^2\right) \geq \theta_H$, then x places no effective constraint on the agent, who chooses $q = q^A\left(w,\theta\right)$ for all θ. This replicates the A-mode. Similarly, if $\frac{2\beta - \phi^2}{\phi}x + w\left(\beta - \Phi^2\right) \leq \theta_L$, then the constraint on q is always binding, so $q = x$ for all θ. This is equivalent to the principal choosing $q = x$ contractually, i.e. the P-mode. As a result, the H-mode only refers to the case when $\left(w,x\right)$ are "interior", i.e. such that

$$
\theta_L < \frac{2\beta - \phi^2}{\phi}x + w\left(\beta - \Phi^2\right) < \theta_H. \tag{6}
$$

The principal's profit as a function of $\left(w,x\right)$ is then $\Pi^H\left(w,x\right)$ given above.

The advantage of delegating to the agent is that the agent will take into account the realized value of θ when choosing q, so will set q closer to the first-best level, and the principal can extract this additional expected payoff through its fixed fee F. But the principal also needs to extract a positive margin (i.e. $w > 0$) in order

to maintain an incentive to invest in Q. This in turn distorts the agent's choice of q, so the principal prefers to stipulate a minimum level of q to help offset the downward bias, although at the cost of having q set too high whenever θ turns out to be particularly low.

Thus, in some sense, threshold delegation can be a way for the principal to combine some of the benefits of both delegation and control. The following proposition establishes the conditions under which threshold delegation dominates the two pure modes.

Proposition 2. *The H-mode with minimum requirement on q dominates the A-mode. If in addition $\frac{\theta_H}{\bar{\theta}} > 1 + \frac{2\Phi^2\beta}{(2\beta-\phi^2)(\beta+2\Phi^2)-2\Phi^4}$, the H-mode also dominates the P-mode.*

While the proof of Proposition 2 (along with those of other propositions not proven directly in the text) is given in the appendix, here we sketch the idea behind the proof. The use of a minimum constraint on the agent's costly investment (rather than imposing a specific level) allows the principal to give the agent discretion to react to its private information about demand shocks, while eliminating the agent's worst under-investment scenarios which occur when the demand shock is particularly low. By setting a minimum constraint above the lowest possible level of q that the agent would choose, the principal prevents the agent from choosing levels of q that could never be optimal from the principal's perspective even if it could observe θ. Put differently, there is no value of θ for which the principal would be happy with the agent's low choice of q in this range, reflecting the agent's downward bias in choosing q. Thus, restricting the choice of q above this range necessarily does better than giving the agent full discretion, i.e. the A-mode.

Moreover, the principal can do strictly better than in P-mode by giving the agent freedom to choose q above the level the principal would like to choose in P-mode, provided there are realizations of θ for which the agent would want to do so. Given there is a downward bias in the agent's choice of q, when the agent prefers to set q higher than $q^P\left(w^{P*},\bar{\theta}\right)$, the principal must also be better off compared to the case when it chooses $q^P\left(w^{P*},\bar{\theta}\right)$ (i.e. P-mode). If the condition in the proposition does not hold, then even for the highest realization of θ, the agent's choice of q would still be below what the principal would like to choose in P-mode, so this logic no longer applies. We analyze this case in more detail in Section 5.3, taking G as the uniform distribution. We show there that the P-mode can sometimes dominate threshold delegation.

Proposition 2 says that the A-mode can never be optimal. However, this conclusion holds in the absence of any fixed costs that the principal might incur when operating in a particular mode. In reality, the principal is likely to incur higher fixed costs in H-mode and P-mode than in A-mode, due to the need to monitor the agent in order to ensure it respects the constraint imposed by the threshold x in H-mode or by the principal's contractual choice of q in P-mode. Clearly, if this monitoring cost is sufficiently large and (5) holds, then the A-mode will be optimal. In Section 5.3 we examine how the optimal choice of mode depends on the magnitude of such monitoring costs. We also show how the boundaries between the different modes changes as the other parameters of the model change.

Proposition 2 established conditions under which the principal prefers threshold delegation to the two pure modes. However, there are more complex forms of partial delegation that the principal could utilize (e.g. delegating subject to requirement that q be in one of two disjoint intervals). The next proposition provides a sufficient condition for the delegation with minimum threshold described in Proposition 2 to be the optimal form of delegation.

Proposition 3. *If $g'(\theta) \leq 0$ for all $\theta \in [\theta_L, \theta_H]$, the optimal contract in H-mode involves threshold delegation (i.e. the principal imposing a minimum threshold on q).*

The condition in the proposition requires that the density is non-increasing on the support of θ. It is obviously satisfied by the uniform and exponential distributions, as well as all distributions that have decreasing density on the positive domain (e.g. the normal or log-normal distributions).

In what follows, we will equate the H-mode with the principal imposing a minimum threshold on q. Assuming that the first-order conditions of $\Pi^H(w, x)$ characterize a unique interior maximum (w^{H*}, x^*), we can also provide some additional properties of the optimal solution in H-mode with threshold delegation, as summarized in the following proposition.

Proposition 4. *The optimal wholesale price extracted by the principal is highest in P-mode, lowest in A-mode, and intermediate in H-mode, i.e. $0 < w^{A*} < w^{H*} < w^{P*}$. Moreover, the optimal minimum requirement for q in H-mode is below the fixed choice of q in P-mode, i.e. $x^* < q^P(w^{P*}, \overline{\theta})$.*

Compared to its optimal solution in A-mode, the principal extracts a higher wholesale price in P-mode. This reflects that in P-mode, the principal sets q so there is no bias, and w is determined by trading off the principal's moral hazards and the distortion in price due to double marginalization. By contrast, in A-mode, the agent sets q resulting in an additional moral hazard problem that is increasing in w. This additional concern means the optimal w is lower. The H-mode is intermediate.

To understand the second part of the proposition, recall from the earlier discussion that if the agent has a downward bias in setting q, then the principal prefers to allow the agent to set q above $q^P(w^{P*}, \overline{\theta})$ whenever the agent wants to do so, rather than forcing it to set $q = q^P(w^{P*}, \overline{\theta})$. Once it allows the agent this freedom, we know from the first part of Proposition 4 that the principal wants to lower w below w^{P*}. This in turn means the principal's preferred level of q also decreases below $q^P(w^{P*}, \overline{\theta})$, so the principal now wants to allow the agent freedom to set q above this new preferred level, i.e. even more freedom than before. And this further lowers the principal's preferred w and so on until we reach $w^{H*} < w^{P*}$. This means that the optimal minimum threshold must be strictly lower than $q^P(w^{P*}, \overline{\theta})$.

5.3 Comparative statics

Given our result in Proposition 2 that the H-mode dominates the A-mode, in order to look at interesting tradeoffs between all three modes, in this section we introduce a monitoring cost that must be incurred whenever the principal wants to exercise some control over q. This monitoring cost arises because the principal wants to ensure either that q is actually set at the particular contracted level (P-mode) or that the agent complies with the minimum restrictions placed on q (H-mode). To keep things as simple as possible, we assume that the monitoring technology requires a fixed ex-ante investment (e.g. hiring additional managers and staff) and no extra cost ex-post (see Gal-Or, 1995), thus abstracting away from any strategic monitoring game that the principal and the agent might engage in (as in Lal, 1990). We denote the fixed monitoring costs by K. When $K > 0$, any of the three modes can be optimal.

We first determine the effect of Φ on the choice between the three modes.

Proposition 5. *A larger Φ shifts the tradeoff between A-mode and H-mode in favor of H-mode, and shifts the tradeoff between H-mode and P-mode in favor of P-mode. I.e.*

$$\frac{d\Pi^{P*}}{d\Phi} > \frac{d\Pi^{H*}}{d\Phi} > \frac{d\Pi^{A*}}{d\Phi}.$$

Moreover, the optimal wholesale prices w^{P*}, w^{A*} and w^{H*} are all increasing in Φ.

These results are easily understood. Since the principal's investment in Q is determined by the wholesale price, when this investment is more important, it is natural that the optimal wholesale prices increase in all three modes. Furthermore, because the wholesale price is highest in P-mode, lowest in A-mode and intermediate in H-mode (Proposition 4), the principal's investment is also highest in P-mode, lowest in A-mode and intermediate in H-mode. Thus, when Q becomes more important, the P-mode becomes more attractive relative to the H-mode, which in turn becomes more attractive relative to the A-mode.

Next, we wish to explore the effect of increasing the importance of the agent's private information about θ, which is captured by its variance V_θ. To do so, we focus on the case when θ follows a uniform distribution. Specifically, let $\theta_L = \overline{\theta} - \sigma$ and $\theta_H = \overline{\theta} + \sigma$, so that the variance of θ is $V_\theta = \frac{\sigma^2}{3}$. The next proposition characterizes the interior solution for the H-mode and the effect of σ on the optimal choice of mode. The proof, which contains lengthy calculations, is relegated to the online appendix.

Proposition 6. *The optimal solution in the H-mode is interior if and only if* $\sigma > \frac{\overline{\theta}\Phi^2\beta}{(2\beta-\phi^2)\left(\frac{\beta}{2}+\Phi^2\right)-\Phi^4}$ *or* $\left(\frac{2\overline{\theta}\Phi^2\beta^2\phi^2}{(\beta^2+2\Phi^2\beta-\Phi^4-\Phi^2\phi^2)^2} \leq \sigma \leq \frac{\overline{\theta}\Phi^2\beta}{(2\beta-\phi^2)\left(\frac{\beta}{2}+\Phi^2\right)-\Phi^4}$ *and* $\left(2\beta-\phi^2\right)\left(\frac{\beta}{2}+\Phi^2\right)-\Phi^4 < \beta\frac{\phi^2}{2}\right)$. *It is then characterized by*

$$w^* = \frac{\sigma}{\beta^2\phi^2}\left((\beta^2 + 2\Phi^2\beta - \Phi^4 - \Phi^2\phi^2) - \sqrt{(\beta^2 + 2\Phi^2\beta - \Phi^4 - \Phi^2\phi^2)^2 - \frac{2\overline{\theta}\Phi^2\beta^2\phi^2}{\sigma}}\right).$$

and

$$x^* = \frac{\phi\left(w^*\left(\beta+\Phi^2\right)+\overline{\theta}-\sigma\right)}{2\beta-\phi^2}.$$

Furthermore, if the optimal solution in the H-mode is interior, then a larger σ shifts the tradeoff between A-mode and H-mode in favor of A-mode, and shifts the tradeoff between H-mode and P-mode in favor of H-mode, i.e.

$$\frac{d\Pi^{A*}}{d\sigma} > \frac{d\Pi^{H*}}{d\sigma} > \frac{d\Pi^{P*}}{d\sigma} = 0.$$

The second part of the proposition says that the larger the variance of the demand shock (holding its expectation constant), the more attractive the A-mode becomes relatively to the H-mode, and the more attractive the H-mode becomes relatively to the P-mode. This reflects that the A-mode fully leverages the agent's private information about demand, the H-mode leverages it to an extent limited by the threshold constraint placed on price, and the P-mode does not leverage the agent's information at all. In the absence of monitoring costs, this result implies that there exists a cutoff $\widehat{\sigma}$, such that the H-mode is optimal for $\sigma \geq \widehat{\sigma}$ and the P-mode is optimal for $\sigma \leq \widehat{\sigma}$ (recall from Proposition 2 that the A-mode is dominated by the H-mode).

We illustrate the above results with three figures, which show the roles of monitoring costs K, the variance of the demand shock (captured by σ), the importance of the principal's investment (Φ) and the importance of the agent's investment (ϕ). In all three figures, black indicates the region for which P-mode is optimal, dark gray the region for which H-mode is optimal, and light gray indicates the region for which A-mode is optimal.

Figure 1 illustrates the optimal choice of mode as a function of $(\sigma, K) \in [0, 5] \times [0, 0.05]$, with the other parameter values set at $\overline{\theta} = 20$, $\beta = 10$, $\phi = 4.5$ and $\Phi = 1.25$. Fixing $K > 0$, as σ increases, the optimal mode shifts from P-mode to H-mode and then to A-mode (consistent with the second part of Proposition 6), or directly from P-mode to A-mode when K is high enough (consistent with Proposition B.1).

Figure 2 illustrates the optimal choice of mode as a function of $(\Phi, K) \in [0, 2.5] \times [0, 0.05]$, with the other parameter values set at $\overline{\theta} = 20$, $\beta = 10$, $\phi = 3.25$ and $\sigma = 3$. Fixing $K > 0$, as Φ increases, the optimal mode shifts from A-mode to H-mode and then to P-mode (consistent with the second part of Proposition 6, or

Figure 1: Optimal mode as a function of σ and K
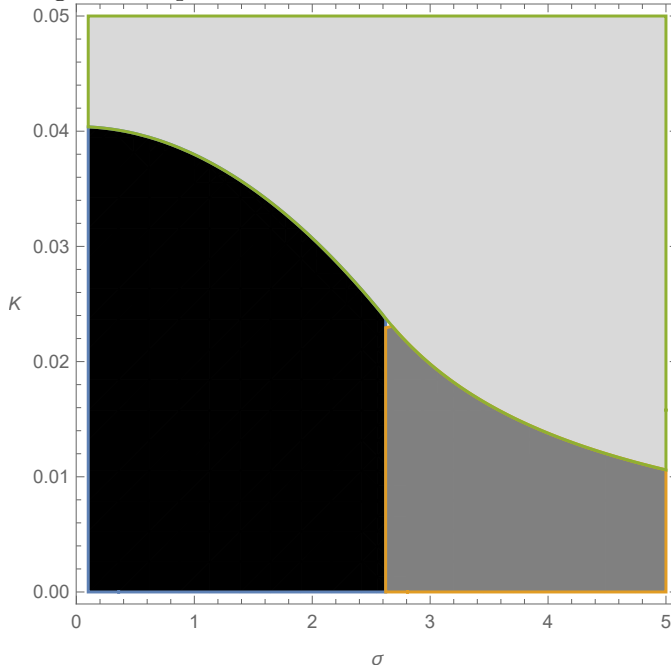
directly from A-mode to P-mode when K is high enough.

Figure 3 illustrates the optimal choice of mode as a function of $(\phi, K) \in [0, 10] \times [0, 0.015]$, with the other parameter values set at $\bar{\theta} = 20$, $\beta = 10$, $\Phi = 0.5$ and $\sigma = 1.5$. Fixing $K > 0$, an increase in ϕ exacerbates the agent's downward bias in choosing q. This effect makes the A-mode and H-mode relatively less attractive, so the principal prefers to delegate less.

Obviously, when K is large enough, the A-mode (which does not require a monitoring cost) dominates the H-mode and the P-mode. An interesting feature illustrated in all three graphs is that as K increases from zero, initially only the margin between the H-mode and the A-mode is affected, with the optimal choice of mode shifting from H-mode to the A-mode. Only after the H-mode is no longer optimal for any parameter values because K has become sufficiently high, does the optimal choice of mode start shifting from the P-mode to the A-mode as K increases. The reason is that the H-mode is intermediate between the two pure modes, so when the monitoring cost incurred in P-mode and H-mode starts increasing, the shift can initially only be between the H-mode and the A-mode. Only once K is high enough that the intermediate H-mode option is no longer viable, does K start affecting the margin between the P-mode and the A-mode. This has a clear empirical implication: we expect to observe only pure modes when monitoring costs are high and hybrid modes only to appear when monitoring costs are low enough.

6   Extensions

In this section we consider three extensions. We first show how our results extend to the case in which the agent's investment in costly action is always controlled by the agent, and instead the price is the transferable decision variable that is either controlled by the principal or the agent. This setup allows us to provide a new theory of resale price maintenance (RPM). We then show that our baseline results are robust to more general demand and cost functions, and to alternative sources of private information. The proofs of the results in this

Figure 2: Optimal mode as a function of Φ and K
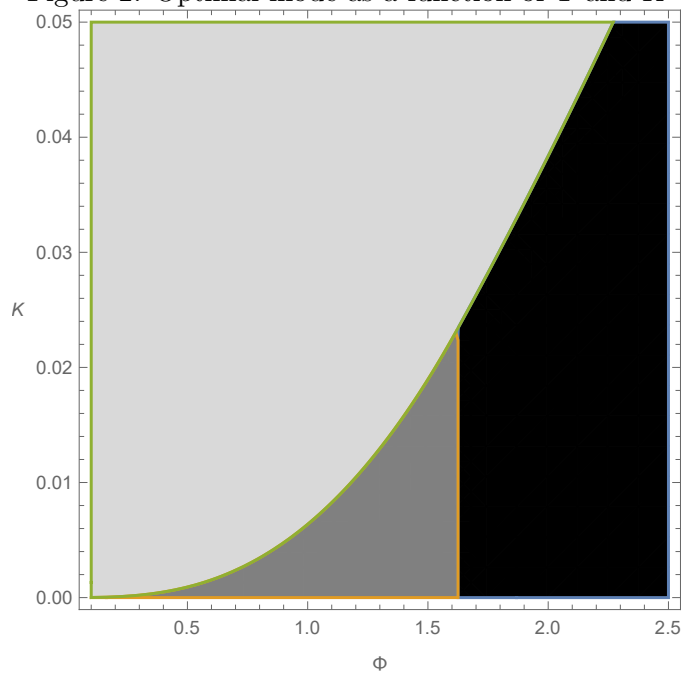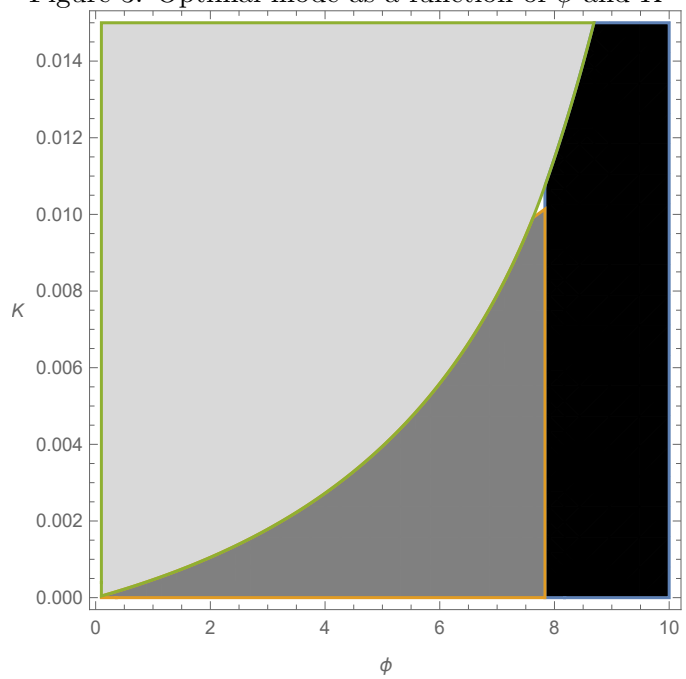

Figure 3: Optimal mode as a function of ϕ and K

section are contained in an online appendix.

6.1 Resale price maintenance

Our framework can be applied to the large existing literature on resale price maintenance (RPM). In terms of the underlying mechanism that explains RPM, only Romano (1994) is similar to our model in that it relies on double-sided moral hazard. The key difference is that we introduce demand uncertainty, which implies RPM with a minimum threshold (min RPM) or RPM with a maximum threshold (max RPM) are distinct from and sometimes strictly better than fixed-price RPM.

We use the same model as in the previous sections, but we now assume that the price p is the transferable variable that can be restricted by the principal, while the costly action q is something that the agent always controls. This setting applies when p can be monitored by the principal and RPM is legal. Similar to (1)-(2), we make the assumptions

$$\left(2\beta - \phi^2\right)\left(\Phi^2 + \phi^2\right) - \Phi^4 > 0 \tag{7}$$

and

$$\frac{\theta_L}{\bar{\theta}} > \frac{\left(\beta - \Phi^2\right)\Phi^2}{\left(2\beta - \phi^2\right)\left(\Phi^2 + \phi^2\right) - \Phi^4 + \left(\beta - \phi^2\right)^2}. \tag{8}$$

The analysis with price as the transferable variable parallels that in Section 5 in which the costly action q was transferable except for one key difference, which is now there can be either an upward or downward bias in the agent's choice of price.[8] An upward bias in the agent's price reflects the standard double marginalization effect that arises in the case in which demand is decreasing in price (i.e. when $\beta > \phi^2$). A downward bias in the agent's price reflects that when $\phi^2 > \beta$, a higher price induces higher demand through a higher choice of q by the agent. In this case, given the agent does not take into account the full increase in revenue associated with the higher q, the agent will tend to set price too low. The conditions under which threshold delegation dominates the two pure modes now reflect these two different scenarios.[9]

Proposition 7. (Resale price maintenance)

Max RPM: *If $\beta > \phi^2$, the H-mode with max RPM dominates the A-mode. If in addition $\frac{\theta_L}{\bar{\theta}} < 1 - \frac{\Phi^2\left(\beta - \phi^2\right)}{\left(2\beta - \phi^2\right)\left(\phi^2 + \Phi^2\right) - \Phi^4}$, the H-mode with max RPM also dominates the P-mode.*

Min RPM: *If $\beta < \phi^2$, the H-mode with min RPM dominates the A-mode. If in addition $\frac{\theta_H}{\bar{\theta}} > 1 + \frac{\Phi^2\left(\phi^2 - \beta\right)}{\left(2\beta - \phi^2\right)\left(\phi^2 + \Phi^2\right) - \Phi^4}$, the H-mode with min RPM also dominates the P-mode.*

Whether max or min RPM is optimal depends on whether the agent's pricing bias is upwards or downwards. The direction of the bias depends on the strength of the agent's moral hazard problem. If the agent's moral hazard problem is not very important, then the usual double marginalization distortion dominates and max RPM is optimal. If the agent's moral hazard problem is sufficiently important, then the agent will set prices too low from the principal's perspective since higher prices actually expand demand on net, and min RPM is optimal. This explanation is aligned with the standard justification given by manufacturers for the imposition of min RPM requirements on retailers, namely that such requirements preserve sufficient incentives by the retailers to invest in services so as to promote consumer demand.

---

[8]The full analysis for the case in which price is the transferable variable is contained in the Online Appendix.

[9]Proposition 3 continues to hold, so the same sufficient condition for threshold delegation to be the optimal form of partial delegation continues to apply. This implies, when the condition in Proposition 3 holds, we should never observe a price floor and price cap used together, other than in the trivial case where they coincide (i.e. the P-mode).

Our results provide an explanation for min RPM that does not rely on competition or free-riding between agents (which is a common explanation for min RPM). The extent to which real-world instances of min RPM are explained by free riding among competing retailers has been increasingly questioned by judges and antitrust scholars (see Klein, 2009, 2014). The key observation underlying this skepticism is that most cases of RPM involve products that are unlikely to benefit from the kind of "showrooming" necessary for free riding among retailers to occur. Instead, antitrust scholars have argued that min RPM is used to alleviate incentive incompatibility between manufacturers and retailers, namely the fact that retailers have insufficient incentives to invest in demand-increasing services from the perspective of the manufacturers. This is consistent with our model and results.

Also in contrast to most existing theories of RPM, Proposition 7 can explain when either min or max RPM should be used. The use of a constraint on prices rather than imposing a specific price level reflects the realistic feature that a principal often wants to give the agent discretion to react to its private information about demand shocks, while mitigating the worst pricing biases that can arise when the agent controls the price but only keeps some of the associated variable revenue.

6.2 General demand and cost functions

In this section we return to our baseline setting in which the costly action q is the transferable action and p is always chosen by the agent. We will work with general demand $D(\theta, p, q, Q)$ and fixed costs $c(q)$ and $C(Q)$. We assume $D(\theta, p, q, Q)$ is increasing in θ, q and Q, and decreasing in p, while $c(q)$ and $C(Q)$ are increasing and convex. If there is a demand interaction effect between Q and q, then the principal's choice of Q in H-mode depends on the full distribution of the agent's choice of the transferable decision q implied by the distribution of θ and the principal's threshold contract. This creates a technical problem in generalizing the results from Proposition 2. To avoid this we assume demand is additively separable in Q, as was the case previously, i.e.

$$D(\theta, p, q, Q) = D^1(\theta, p, q) + D^2(Q).$$

Define $(p(q, w, \theta), Q(w))$ as the Nash equilibrium of the second-stage game in which the agent and the principal simultaneously choose p and Q, given q and w. It is the joint solution to $p = \arg\max_{p'} \{(p' - w) D(\theta, p', q, Q)\}$ and $Q = \arg\max_{Q'} \{w D^2(Q') - C(Q')\}$. Note that, because demand is additively separable in Q, the principal's choice of Q does not depend on q or θ, and so the timing of the choice of q does not affect the outcome.

Let then

$$
\begin{aligned}
q^P(w, \theta) &\equiv \arg\max_{q'} \{p(q', w, \theta) D(\theta, p(q', w, \theta), q', Q(w)) - c(q') - C(Q(w))\} \\
q^A(w, \theta) &\equiv \arg\max_{q'} \{(p(q', w, \theta) - w) D(\theta, p(q', w, \theta), q', Q(w)) - c(q')\}.
\end{aligned}
$$

Thus, $q^P(w, \theta)$ is the hypothetical level of q that the principal would like to choose for a given w if it could observe θ when setting q in stage 1. Meanwhile, $q^A(w, \theta)$ is the level of q chosen by the agent in A-mode, given w and θ.

We assume that $q^P(w, \theta)$ and $q^A(w, \theta)$ are increasing in θ as they were in our linear demand setting and that the total profit

$$p(q, w, \theta) D(\theta, p(q, w, \theta), q, Q(w)) - c(q) - C(Q(w)) \tag{9}$$

is concave in q for any (w, θ). Denoting by q^{P*} and w^{P*} the principal's optimal choices of q and wholesale price in P-mode, and by w^{A*} the principal's optimal choice of wholesale price in A-mode, we obtain the following

result.

Proposition 8. (Delegation of costly action)

Minimum requirements: *Assume* $\frac{dD^1(\theta,p(q,w,\theta),q)}{dq} > 0$, *so that an increase in the costly investment* q *increases demand after taking into account the agent's price response. Then the H-mode with minimum requirements dominates the A-mode. If in addition,* $q^A\left(w^{P*},\theta_H\right) > q^{P*}$, *the H-mode with minimum requirements dominates the P-mode.*

Under the assumption in Proposition 8, the agent's bias in choosing the costly action is in the same downward direction when $w = w^{A*}$ and $w = w^{P*}$, and provided the agent will sometimes prefer a higher level of q than that set by the principal in P-mode, minimum threshold delegation dominates both full control and full delegation.[10] This generalizes the result in Proposition 2 to any (non-linear) demand function, as long as it remains additively separable in Q.

6.3 Alternative sources of private information

We have assumed throughout that the agent's private information pertained to a demand shock θ. In practice, there can be other sources for the agent's private information, e.g. private benefits or costs, the effect of the agent's investment on demand, and so on. In this section we briefly explain how our model can be applied to the case of private benefits or costs.

Suppose demand is deterministic and given by $D(p,q,Q)$. Costs are once again $c(q)$ and $C(Q)$. However, suppose the agent derives a private benefit b per unit of demand, where b is distributed with mean $\bar{b}$ and variance $V_b > 0$. We allow $\bar{b}$ to be positive or negative, so b can also be a private cost. Thus, given a two-part tariff (w,F), the principal extracts

$$wD(p,q,Q) - C(Q) + F,$$

whereas the agent receives

$$(p+b-w)D(p,q,Q) - c(Q) - F.$$

All other assumptions are unchanged. In particular, b is only observed by the agent in the second stage.

A private benefit reflects that when the agent has a higher level of sales through its contract with the principal, the agent may also have increased opportunities to sell other services or products to the same customers, but these opportunities fall outside the scope of the contractual relationship with the principal. For instance, Uber and Lyft drivers may offer some of their riders fixed price rides to the airport at mutually agreed upon times, bypassing the Uber and Lyft apps. Similarly, independent contractors on TaskRabbit and Thumbtack may be able to sell customers additional services bypassing the platforms. Alternatively, the agent may incur private costs proportional to demand that are not accounted for in the contract with the principal. Examples of such private costs include the opportunity costs of Lyft and Uber drivers, expenses related to car maintenance, etc.

It is then easily seen that, assuming additive separability of $D(p,q,Q)$ in Q, the set-up and analysis from Section 6.2 goes through almost unchanged. In particular, if $D(p,q,Q) = \theta - \beta p + \phi q + \Phi Q$, where $\theta > 0$ is now a constant, $c(q) = \frac{1}{2}q^2$ and $C(Q) = \frac{1}{2}Q^2$, then the condition determining whether the principal should employ minimum requirements is unchanged from that obtained in Proposition 2. This shows that the fundamental logic of platform minimal requirements uncovered by our model does not depend on the nature of the agent's

---

[10]The assumption in Proposition 8 is quite natural and it holds for example if $D^1(\theta,p,q)$ can be written as $D'(\theta,\beta p - \phi q)$, where D' is log-concave in its second argument.

private information. What matters is that the agent has some private information which affects his choice of the costly investment q. Imposing a minimum requirement is then a way for the principal to allow the agent to adjust decisions to private information, while eliminating the worst under-investment scenarios.

7   Concluding remarks

With the widespread emergence of digital monitoring technologies, there are many new opportunities for firms in general and platforms in particular to enforce the delegation of key decisions to independent agents. As a result, partial delegation is likely to become a contractual instrument that a greater number of firms that act as principals (e.g. platforms, franchisors, licensors, and manufacturers) can consider using when setting the terms for their agents (e.g. third-party suppliers, franchisees, licensees, and retailers). The recent shift towards the legality of resale price maintenance (RPM) also means that placing restrictions on price setting is becoming increasingly relevant. Our theory provides several lessons in this regard.

At a high level, we have shown that delegation subject to minimum requirements strikes a middle ground between complete delegation and full control, and oftentimes does better than both. It is a way to get the best of both worlds, by leveraging the agent's private information, while also eliminating the more extreme biases that arise when an agent only keeps some of the revenue it produces. For platforms, one can view the use of minimum requirements as a governance rule designed to achieve strategic positions that are intermediate along the spectrum between pure platform (e.g. relying on unconstrained independent contractors) and pure vertical integration (e.g. relying on employees). In manufacturer-retailer contexts, threshold delegation is an additional contracting instrument that can improve channel coordination beyond what can be achieved with typical pricing instruments, such as standard forms of wholesale pricing or revenue sharing, together with fixed fees.

In contexts where there is more uncertainty regarding the agent's private information (i.e. higher variance of private shocks), the principal should give the agent more autonomy, i.e. switch from full control to partial delegation, reduce minimum requirements if they are already in place, or even switch from partial delegation to full delegation to the agent. When the principal's or the agent's moral hazard hazard becomes more important, the principal should delegate less, i.e. switch to partial delegation from full delegation, increase minimum requirements if they are already in place, or even switch from partial delegation to full control.

There are several promising directions in which our analysis can be extended. Using the existing framework, it would be quite straightforward to introduce multiple agents and (positive or negative) spillovers from each agent's choice of the transferable action on the revenues derived by other agents. The interesting question would then be to determine whether spillovers exacerbate or offset the agent's bias, and therefore how spillovers affect the principal's delegation decision. Another direction worth exploring would be to introduce risk aversion and/or wealth constraints for agents, which will limit the upfront fixed fees they can be charged. This should have an effect similar to the principal's moral hazard: namely, it would increase the principal's wholesale price in all three modes, and so the degree of bias, thereby shifting the tradeoff in favor of less delegation. A more challenging extension would be to consider the case in which both the price (p) and the agent's investment (q) are transferable and contractible actions, and to consider whether partial delegation is ever optimal with respect to both variables at the same time. Finally, it would also be interesting (and challenging) to extend our model to allow for multiple competing principals. This could possibly generate equilibria in which principals compete with different delegation models.

References

Alonso R. and N. Matouschek (2008) "Optimal Delegation," *Review of Economic Studies*, 75(1), 259-293.

Amador, M. and K. Bagwell (2013) "The Theory of Optimal Delegation with an Application to Tariff Caps," *Econometrica*, 81(4), 1541–1599.

August, T., H. Shin and T. I. Tunca (2013) "Licensing and Competition for Services in Open Source Software," *Information Systems Research* 24(4), 1068-1086.

August, T., H. Shin and T. I. Tunca (2018) "Generating Value Through Open Source: Software Service Market Regulation and Licensing Policy," *Information Systems Research* 29(1), 186-205.

Bhattacharyya, S. and F. Lafontaine (1995) "Double-Sided Moral Hazard and the Nature of Share Contracts," *RAND Journal of Economics*, 26(4), 761-781.

Boudreau, K. (2010) "Open Platform Strategies and Innovation: Granting Access vs. Devolving Control," *Management Science*, 56(10), 1849-1872.

Boudreau, K. J., and A. Hagiu (2009) "Platform Rules: Multi-Sided Platforms as Regulators," in *Platforms, Markets and Innovation*, edited by Annabelle Gawer. Cheltenham, UK: Edward Elgar Publishing.

Cachon, G. P. and M. Lariviere (2005) "Supply Chain Coordination with Revenue-Sharing Contracts: Strengths and Limitations," *Management Science*, 51(1), 30–44.

Desiraju, R. and S. Moorthy (1997) "Managing a Distribution Channel under Asymmetric Information with Performance Requirements," *Management Science*, 43(12), 1628-1644.

Hagiu, A. and J. Wright (2015) "Marketplace or Reseller?" *Management Science*, 61(1), 184–203.

Hagiu, A. and J. Wright (2018) "Controlling vs. Enabling," *Management Science*, forthcoming.

Holmstrom, B. (1977) "On Incentives and Control in Organizations," Ph.D. Thesis, Stanford University.

Holmstrom, B. (1984) "On the Theory of Delegation", in M. Boyer, and R. Kihlstrom (eds.) *Bayesian Models in Economic Theory* (New York: North-Holland), 115-141.

Jerath, K. and Z.J. Zhang (2010) "Store Within a Store," *Journal of Marketing Research*, 47(4), 748–763.

Klein, B. (2009) "Competitive Resale Price Maintenance in the Absence of Free Riding," *Antitrust Law Journal,* 76(2), 431-481.

Klein, B. (2014) "The Evolving Law and Economics of Resale Price Maintenance," *Journal of Law and Economics*, 57, 161-179.

Lafontaine, F. and M.E. Slade (1997) "Retail Contracting: Theory and Practice," *Journal of Industrial Economics*, XLV, 1-25.

Lal R. (1990) "Improving Channel Coordination through Franchising," *Marketing Science*, 9(4), 299–318.

Martimort, D. and A. Semenov (2006) "Continuity in Mechanism Design Without Transfers" *Economics Letters*, 93, 182-189

Melumad, N. and T. Shibano (1991) "Communication in Settings with No Transfers," *Rand Journal of Economics*, 22(2), 173-198.

Niculescu, M.F., D. J. Wu, L. Xu (2018) "Strategic Intellectual Property Sharing: Competition on an Open Technology Platform Under Network Effects," *Information Systems Research* 29(2), 498-519.

Parker, G., M. Van Alstyne and S. P. Choudary (2016) Platform Revolution, W. W. Norton & Company.

Parker, G. and M. Van Alstyne (2018) "Innovation, Openness, and Platform Control," *Management Science*, 64(7), 3015-3032.

Romano, R. E. (1994) "Double Moral Hazard and Resale Price Maintenance," *Rand Journal of Economics*, 25(3), 455–466.

Simon, H. A. (1951) "A Formal Theory of the Employment Relationship," *Econometrica*, 19(3), 293-305.

8   Appendix

8.1   Proof of Proposition 2

Compare the A-mode to the H-mode with $w^H = w^{A*}$ and $x = q^A\left(w^{A*}, \theta_L + \kappa\right) = \frac{\left(\theta_L + \kappa + w\left(\Phi^2 - \beta\right)\right)\phi}{2\beta - \phi^2}$, where $\kappa > 0$ is small. We can then write

$$\Pi^H\left(w^{A*}, x = q^A\left(w^{A*}, \theta_L + \kappa\right)\right) - \Pi^{A*} = \Pi^H\left(w^{A*}, x = q^A\left(w^{A*}, \theta_L + \kappa\right)\right) - \Pi^H\left(w^{A*}, x = q^A\left(w^{A*}, \theta_L\right)\right)$$

$$= \int_{\theta_L}^{\theta_L + \kappa} \frac{1}{4\beta}\left(q^A\left(w^{A*}, \theta_L + \kappa\right) - q^A\left(w^{A*}, \theta\right)\right)\left(\begin{array}{c}2\theta\phi + 2w^{A*}\Phi^2\phi \\ -\left(2\beta - \phi^2\right)\left(q^A\left(w^{A*}, \theta_L + \kappa\right) + q^A\left(w^{A*}, \theta\right)\right)\end{array}\right) dG\left(\theta\right)$$

$$= \int_{\theta_L}^{\theta_L + \kappa} \frac{\phi^2}{4\beta\left(2\beta - \phi^2\right)}\left(\theta_L + \kappa - \theta\right)\left(\theta - \theta_L - \kappa + 2w^{A*}\beta\right) dG\left(\theta\right).$$

The last expression is positive for κ sufficiently small because $w^{A*} > 0$. Thus, the H-mode dominates the A-mode.

Next, compare the P-mode to the H-mode with $w^H = w^{P*}$ and $x = q^{P*} = \frac{\left(\overline{\theta} + w^{P*}\Phi^2\right)\phi}{2\beta - \phi^2}$. Note that this $\left(w^H, x\right)$ is strictly interior because

$$\theta_L < \frac{2\beta - \phi^2}{\phi}q^{P*} + w^{P*}\left(\beta - \Phi^2\right) = \overline{\theta} + w^{P*}\beta < \theta_H,$$

where the second inequality follows from the assumption $\frac{\theta_H}{\overline{\theta}} > 1 + \frac{\Phi^2\beta}{\left(2\beta - \phi^2\right)\left(\frac{\beta}{2} + \Phi^2\right) - 2\Phi^4}$. We can then write

$$\Pi^H\left(w^{P*}, x = q^{P*}\right) - \Pi^{P*} = \Pi^H\left(w^{P*}, q^{P*}\right) - \Pi^P\left(w^{P*}, q^{P*}\right)$$

$$= \int_{\overline{\theta} + w^{P*}\beta}^{\theta_H}\left(\begin{array}{c}\frac{1}{4\beta}\left(\theta + \phi q^A\left(w^{P*}, \theta\right) + w^{P*}\left(\Phi^2 - \beta\right)\right)^2 + \frac{w^{P*}}{2}\left(\theta + \phi q^A\left(w^{P*}, \theta\right) - \beta w^{P*}\right) - \frac{1}{2}q^A\left(w^{P*}, \theta\right)^2 \\ -\left(\frac{1}{4\beta}\left(\theta + \phi q^{P*} + w^{P*}\left(\Phi^2 - \beta\right)\right)^2 + \frac{w^{P*}}{2}\left(\theta + \phi q^{P*} - \beta w^{P*}\right) - \frac{1}{2}\left(q^{P*}\right)^2\right)\end{array}\right) dG\left(\theta\right)$$

$$= \int_{\overline{\theta} + w^{P*}\beta}^{\theta_H} \frac{1}{4\beta}\left(q^A\left(w^{P*}, \theta\right) - q^{P*}\right)\left(2\theta\phi + 2w^{P*}\Phi^2\phi - \left(2\beta - \phi^2\right)\left(q^A\left(w^{P*}, \theta\right) + q^{P*}\right)\right) dG\left(\theta\right)$$

$$= \int_{\overline{\theta} + w^{P*}\beta}^{\theta_H} \frac{\phi^2}{4\beta\left(2\beta - \phi^2\right)}\left(\theta - \overline{\theta} - w^{P*}\beta\right)\left(\theta - \overline{\theta} + w^{P*}\beta\right) dG\left(\theta\right) > 0.$$

Thus, the H-mode dominates the P-mode under the stated assumption.

23

8.2   Proof of Proposition 3

Since the principal fixes w in its contract at the same time as deciding on the type and nature of any delegation, we just have to show threshold delegation is optimal for any given w. To do so, we will show that any contract that differs from threshold delegation can be improved upon by a contract with the same w and threshold delegation.

The principal's delegation problem for a fixed choice of w is

$$\max_{D} \left\{ \mathbb{E} \left[p \left(\theta - \beta p + \phi q + \Phi Q \right) - \frac{1}{2} q^2 - \frac{1}{2} Q^2 \right] \right\}$$

subject to

$$(q, p) = \arg \max_{q' \in D, p} \left\{ (p - w) \left(\theta - \beta p + \phi q' + \Phi Q \right) - \frac{1}{2} \left(q' \right)^2 \right\}$$

$$Q = w \Phi,$$

where D denotes the delegation set to which the principal restricts the agent's choice of q. This can be rewritten as

$$\max_{D} \left\{ \mathbb{E} \left[\left(\frac{w}{2} + \frac{\theta + \phi q + w \Phi^2}{2 \beta} \right) \left(\theta - \beta \left(\frac{w}{2} + \frac{\theta + \phi q + w \Phi^2}{2 \beta} \right) + \phi q + w \Phi^2 \right) - \frac{1}{2} q^2 - \frac{1}{2} w^2 \Phi^2 \right] \right\}$$

subject to

$$q = \arg \max_{q' \in D} \left\{ \left(\frac{\theta + \phi q' + w \Phi^2}{2 \beta} - \frac{w}{2} \right) \left(\theta - \beta \left(\frac{w}{2} + \frac{\theta + \phi q' + w \Phi^2}{2 \beta} \right) + \phi q' + w \Phi^2 \right) - \frac{1}{2} \left(q' \right)^2 \right\}.$$

Ignoring terms that do not depend on q, the program that defines the principal's optimal delegation set $D(w)$ can be re-written more simply as

$$\max_{D} \mathbb{E} \left[-\alpha_0 q^2 + \frac{2 \phi}{4 \beta} \left(\theta + w \Phi^2 \right) q \right]$$

$$\text{subject to } q = \arg \max_{q' \in D} \left\{ -\alpha_0 \left(q' \right)^2 + \frac{2 \phi}{4 \beta} \left(\theta + w \Phi^2 - \alpha_1 \right) q' \right\},$$

where $\alpha_1 = w \beta$ and $\alpha_0 = \frac{1}{2} - \frac{\phi^2}{4 \beta} > 0$ due to our assumption (1).

In this model, for a given θ, the ideal choice of q for the principal is $q = \frac{\left(\theta + w \Phi^2 \right) \phi}{4 \beta \alpha_0}$, while for the agent it is $q = \frac{\left(\theta + w \Phi^2 - \alpha_1 \right) \phi}{4 \beta \alpha_0}$. Given $\alpha_1 > 0$, without any restrictions, the agent prefers a lower q than the principal. The principal is therefore interested in restricting the agent from setting q too low. The question remains whether the principal can do better by requiring the agent to choose from some specific values of q or some disjoint intervals that exclude some high values of q. Formally, we want to show that the delegation set $D(w)$ is a threshold interval, i.e. $D(w) = \{ q \geq x(w) \}$ for some $x(w)$.

Suppose first the principal restricts the agent to choose q from some subset of $q \leq q_0 \equiv \frac{\left(\theta_0 + w \Phi^2 - \alpha_1 \right) \phi}{4 \beta \alpha_0}$ which includes $q = q_0$, where $\theta_0 < \theta_H$. This covers the possibility that the agent can only choose $q = q_0$ or can choose any $q \leq q_0$. In this case, when $\theta \in [\theta_0, \theta_H]$, it is easily seen that the agent chooses $q = q_0$ because the agent's objective function is increasing in q for all $q \leq q_0$. But the principal could strictly improve expected profits by adding the range $q \geq q_0$ to the set of permissible choices of q by the agent. To see this, note that the only

change comes from the different choices of q by the agent when $\theta \in [\theta_0, \theta_H]$. The change in expected profits is

$$
\int_{\theta_0}^{\theta_H} \left(\begin{array}{c} -\alpha_0 \left(\frac{(\theta + w\Phi^2 - \alpha_1)\phi}{4\beta\alpha_0} \right)^2 + \frac{2\phi}{4\beta} \left(\theta + w\Phi^2\right) \left(\frac{(\theta + w\Phi^2 - \alpha_1)\phi}{4\beta\alpha_0} \right) \\ -\left(-\alpha_0 \left(\frac{(\theta_0 + w\Phi^2 - \alpha_1)\phi}{4\beta\alpha_0} \right)^2 + \frac{2\phi}{4\beta} \left(\theta + w\Phi^2\right) \left(\frac{(\theta_0 + w\Phi^2 - \alpha_1)\phi}{4\beta\alpha_0} \right) \right) \end{array} \right) dG\left(\theta\right)
$$

$$
= \frac{\phi^2}{16\beta^2\alpha_0} \int_{\theta_0}^{\theta_H} \left(\theta - \theta_0\right) \left(\left(\theta - \theta_0\right) + 2\alpha_1\right) dG\left(\theta\right) > 0.
$$

Suppose now that the agent is allowed to choose from some set that does not include $q \in (q_0, q_1)$, where $q_0 \equiv \frac{(\theta_0 + w\Phi^2 - \alpha_1)\phi}{4\beta\alpha_0}$ and $q_1 \equiv \frac{(\theta_1 + w\Phi^2 - \alpha_1)\phi}{4\beta\alpha_0}$ for some $\theta_L \leq \theta_0 < \theta_1 \leq \theta_H$. In this case, since the agent's objective function is quadratic in q, if the agent's draw of θ is in the range $[\theta_0, \theta_1]$, then the agent chooses $q = q_0$ when $\theta \leq \frac{\theta_0 + \theta_1}{2}$ and $q = q_1$ when $\theta > \frac{\theta_0 + \theta_1}{2}$. The principal can profitably deviate by adding the range $[q_0, q_1]$ to the set of permissible choices of q. The change in profits is

$$
\frac{\phi^2}{16\beta^2\alpha_0} \int_{\theta_0}^{\frac{\theta_0 + \theta_1}{2}} \left(\theta - \theta_0\right)\left(\theta - \theta_0 + 2\alpha_1\right) dG\left(\theta\right) + \frac{\phi^2}{16\beta^2\alpha_0} \int_{\frac{\theta_0 + \theta_1}{2}}^{\theta_1} \left(\theta - \theta_1\right)\left(\theta - \theta_1 + 2\alpha_1\right) dG\left(\theta\right)
$$

$$
= \frac{\phi^2}{16\beta^2\alpha_0} \left(\int_{\theta_0}^{\frac{\theta_0 + \theta_1}{2}} \left(2\alpha_1\left(\theta - \theta_0\right) + \left(\theta - \theta_0\right)^2\right) g\left(\theta\right) d\theta + \int_{\frac{\theta_0 + \theta_1}{2}}^{\theta_1} \left(2\alpha_1\left(\theta - \theta_1\right) + \left(\theta - \theta_1\right)^2\right) g\left(\theta\right) d\theta \right).
$$

Using integration by parts, we have

$$
\int_{\theta_0}^{\frac{\theta_0 + \theta_1}{2}} \left(2\alpha_1\left(\theta - \theta_0\right) + \left(\theta - \theta_0\right)^2\right) g\left(\theta\right) d\theta = \alpha_1 \frac{\left(\theta_1 - \theta_0\right)^2}{4} g\left(\frac{\theta_0 + \theta_1}{2}\right) + \int_{\theta_0}^{\frac{\theta_0 + \theta_1}{2}} \left(\theta - \theta_0\right)^2 \left(g\left(\theta\right) - \alpha_1 g'\left(\theta\right)\right) d\theta
$$

$$
\int_{\frac{\theta_0 + \theta_1}{2}}^{\theta_1} \left(2\alpha_1\left(\theta - \theta_1\right) + \left(\theta - \theta_1\right)^2\right) g\left(\theta\right) d\theta = -\alpha_1 \frac{\left(\theta_1 - \theta_0\right)^2}{4} g\left(\frac{\theta_0 + \theta_1}{2}\right) + \int_{\frac{\theta_0 + \theta_1}{2}}^{\theta_1} \left(\theta - \theta_1\right)^2 \left(g\left(\theta\right) - \alpha_1 g'\left(\theta\right)\right) d\theta.
$$

With these expressions, the profit change above becomes equal to

$$
\frac{\phi^2}{16\beta^2\alpha_0} \left(\int_{\theta_0}^{\frac{\theta_0 + \theta_1}{2}} \left(\theta - \theta_0\right)^2 \left(g\left(\theta\right) - \alpha_1 g'\left(\theta\right)\right) d\theta + \int_{\frac{\theta_0 + \theta_1}{2}}^{\theta_1} \left(\theta - \theta_1\right)^2 \left(g\left(\theta\right) - \alpha_1 g'\left(\theta\right)\right) d\theta \right).
$$

This expression is clearly positive under the assumption $\alpha_1 g'\left(\theta\right) \leq g\left(\theta\right)$ for all θ.

Thus, we can conclude that the optimal range of admissible q for the agent must take the form of a threshold interval $q \geq x$. Since $\alpha_1 = w\beta > 0$, threshold delegation is optimal provided

$$
w\beta g'\left(\theta\right) \leq g\left(\theta\right). \tag{10}
$$

Since $w > 0$ and $g\left(\theta\right) > 0$ for all $\theta \in [\theta_L, \theta_H]$, a sufficient condition for (10) to hold is $g'\left(\theta\right) \leq 0$.

8.3   Proof of Proposition 4

The principal's P-mode profit as a function of w only (i.e. after optimizing over q and ignoring terms that don't depend on w), is

$$
\Pi^P\left(w\right) = \frac{\left(w^2 \left(2\Phi^4 - 2\beta^2 + 2\Phi^2\phi^2 - 4\Phi^2\beta + \beta\phi^2\right) + 4w\Phi^2\bar{\theta}\right)}{4\left(2\beta - \phi^2\right)}. \tag{11}
$$

Optimizing over w, we obtain that w^{P*} is the solution to $h^P(w) = 0$, where

$$h^P(w) \equiv \frac{2\Phi^2\overline{\theta} - w\left(\left(2\beta - \phi^2\right)\left(\beta + 2\Phi^2\right) - 2\Phi^4\right)}{2\left(2\beta - \phi^2\right)}.$$

Similarly, the principal's A-mode profits as a function of w only is

$$\Pi^A(w) = \frac{w^2\left(\Phi^4 - \beta^2 + \Phi^2\phi^2 - 2\Phi^2\beta\right) + 2w\Phi^2\overline{\theta}}{2\left(2\beta - \phi^2\right)}. \tag{12}$$

Optimizing over w, we obtain that w^{A*} is the solution to $h^A(w) = 0$, where

$$h^A(w) \equiv \frac{2\Phi^2\overline{\theta} - w\left(\left(2\beta - \phi^2\right)\left(\beta + 2\Phi^2\right) + \beta\phi^2 - 2\Phi^4\right)}{2\left(2\beta - \phi^2\right)}.$$

The first-order conditions of $\Pi^H(w, x)$ in x and w are

$$\int_{\theta_L}^{\frac{2\beta - \phi^2}{\phi}x + w\left(\beta - \Phi^2\right)} \left(\frac{\left(\theta + \phi x + w\left(\Phi^2 - \beta\right)\right)\phi}{2\beta} + \frac{w}{2}\phi - x\right) dG(\theta) = 0 \tag{13}$$

$$\int_{\theta_L}^{\frac{2\beta - \phi^2}{\phi}x + w\left(\beta - \Phi^2\right)} \left(\frac{1}{2\beta}\left(w\Phi^4 - w\beta^2 - 2w\Phi^2\beta + \theta\Phi^2 + x\Phi^2\phi\right)\right) dG(\theta)$$

$$+ \int_{\frac{2\beta - \phi^2}{\phi}x + w\left(\beta - \Phi^2\right)}^{\theta_H} \left(\frac{1}{2\beta - \phi^2}\left(w\Phi^4 - w\beta^2 - 2w\Phi^2\beta + w\Phi^2\phi^2 + \theta\Phi^2\right)\right) dG(\theta) = 0$$

Combining these two equations, we obtain the w^{H*} is the solution to $h^H(w) = 0$, where

$$h^H(w) \equiv \frac{2\Phi^2\overline{\theta} - w\left(\left(2\beta - \phi^2\right)\left(\beta + 2\Phi^2\right) + \beta\phi^2 - 2\Phi^4\right)}{2\left(2\beta - \phi^2\right)} + G\left(\frac{2\beta - \phi^2}{\phi}x + w\left(\beta - \Phi^2\right)\right)\frac{w\beta\phi^2}{2\left(2\beta - \phi^2\right)}.$$

Comparing the expressions of $h^P(w)$, $h^A(w)$, and $h^H(w)$ above, it is clear that $h^P(w)$ and $h^A(w)$ are decreasing in w by (1), and that $h^P(w) > h^H(w) > h^A(w)$ for all w. This implies $w^{A*} < w^{H*} < w^{P*}$.

Finally, recalling that $q^P(w, \theta) = \frac{\left(\theta + w\Phi^2\right)\phi}{2\beta - \phi^2}$, we obtain

$$\frac{\partial\Pi^H\left(w^{H*}, x = q^P\left(w^{H*}, \overline{\theta}\right)\right)}{\partial x} = \frac{\phi}{2\beta}\int_{\theta_L}^{\min\left\{\overline{\theta} + \beta w^{H*}, \theta_H\right\}}\left(\theta - \overline{\theta}\right)dG(\theta) \le 0 = \frac{\partial\Pi^H\left(w^{H*}, x = x^*\right)}{\partial x}.$$

Combined with $w^{H*} < w^{P*}$, this implies $x^* \le q^P\left(w^{H*}, \overline{\theta}\right) < q^P\left(w^{P*}, \overline{\theta}\right)$.

8.4   Proof of Proposition 5

Using expressions (11) and (12) and applying the envelope theorem, we obtain

$$\frac{d\Pi^{P*}}{d\Phi^2} = \frac{w^{P*}\left(2\overline{\theta} + w^{P*}\left(2\Phi^2 - 2\beta + \phi^2\right)\right)}{2\left(2\beta - \phi^2\right)} \quad \text{and} \quad \frac{d\Pi^{A*}}{d\Phi^2} = \frac{w^{A*}\left(2\overline{\theta} + w^{A*}\left(2\Phi^2 - 2\beta + \phi^2\right)\right)}{2\left(2\beta - \phi^2\right)}.$$

Similarly, using the expression for $\Pi^H(w,x)$ and applying the envelope theorem, we have

$$\frac{d\Pi^{H*}}{d\Phi^2} = \int_{\theta_L}^{\frac{2\beta-\phi^2}{\phi}x^*+w^{H*}(\beta-\Phi^2)} \left(\frac{w^{H*}\left(\theta+w^{H*}\left(\Phi^2-\beta\right)+\phi x^*\right)}{2\beta}\right) dG(\theta)$$
$$+ \int_{\frac{2\beta-\phi^2}{\phi}x^*+w^{H*}(\beta-\Phi^2)}^{\theta_H} \left(\frac{w^{H*}\left(2\theta+w^{H*}\left(2\Phi^2-2\beta+\phi^2\right)\right)}{2\left(2\beta-\phi^2\right)}\right) dG(\theta).$$

At the same time, the first-order condition of $\Pi^H(w,x)$ in x (13) implies

$$\int_{\theta_L}^{\frac{2\beta-\phi^2}{\phi}x^*+w^{H*}(\beta-\Phi^2)} \left(\frac{\left(\theta+w^{H*}\Phi^2\right)\phi}{2\beta-\phi^2}-x^*\right) dG(\theta) = 0.$$

Using this equation to replace x^* in the expression of $\frac{d\Pi^{H*}}{d\Phi^2}$ above, we obtain

$$\frac{d\Pi^{H*}}{d\Phi^2} = \int_{\theta_L}^{\frac{2\beta-\phi^2}{\phi}x^*+w^{H*}(\beta-\Phi^2)} \left(\frac{w^{H*}\left(\theta+w^{H*}\left(\Phi^2-\beta\right)+\phi x^*\right)}{2\beta}\right) dG(\theta)$$
$$+ \int_{\frac{2\beta-\phi^2}{\phi}x^*+w^{H*}(\beta-\Phi^2)}^{\theta_H} \left(\frac{w^{H*}\left(2\theta+w^{H*}\left(2\Phi^2-2\beta+\phi^2\right)\right)}{2\left(2\beta-\phi^2\right)}\right) dG(\theta)$$
$$= \int_{\theta_L}^{\frac{2\beta-\phi^2}{\phi}x^*+w^{H*}(\beta-\Phi^2)} \left(\frac{w^{H*}\left(\theta+w^{H*}\left(\Phi^2-\beta\right)\right)}{2\beta}+\frac{\phi w^{H*}}{2\beta}\frac{\left(\theta+w^{H*}\Phi^2\right)\phi}{2\beta-\phi^2}\right) dG(\theta)$$
$$+ \int_{\frac{2\beta-\phi^2}{\phi}x^*+w^{H*}(\beta-\Phi^2)}^{\theta_H} \left(\frac{w^{H*}\left(2\theta+w^{H*}\left(2\Phi^2-2\beta+\phi^2\right)\right)}{2\left(2\beta-\phi^2\right)}\right) dG(\theta)$$
$$= \int_{\theta_L}^{\frac{2\beta-\phi^2}{\phi}x^*+w^{H*}(\beta-\Phi^2)} \left(\frac{w^{H*}\left(2\theta+w^{H*}\left(2\Phi^2-2\beta+\phi^2\right)\right)}{2\left(2\beta-\phi^2\right)}\right) dG(\theta)$$
$$+ \int_{\frac{2\beta-\phi^2}{\phi}x^*+w^{H*}(\beta-\Phi^2)}^{\theta_H} \left(\frac{w^{H*}\left(2\theta+w^{H*}\left(2\Phi^2-2\beta+\phi^2\right)\right)}{2\left(2\beta-\phi^2\right)}\right) dG(\theta)$$
$$= \frac{w^{H*}\left(2\overline{\theta}+w^{H*}\left(2\Phi^2-2\beta+\phi^2\right)\right)}{2\left(2\beta-\phi^2\right)}$$

Let then

$$f(w) \equiv \frac{w\left(2\overline{\theta}+w\left(2\Phi^2-2\beta+\phi^2\right)\right)}{2\left(2\beta-\phi^2\right)}$$

There are two cases:

- if $\phi^2+2\Phi^2-2\beta > 0$, then $f(w)$ is increasing, so $f\left(w^{A*}\right) < f\left(w^{H*}\right) < f\left(w^{P*}\right)$ because $w^{A*} < w^{H*} < w^{P*}$.

- if $\phi^2 + 2\Phi^2 - 2\beta \le 0$, then $f'(w) = \frac{\overline{\theta}+\left(2\Phi^2-2\beta+\phi^2\right)w}{2\beta-\phi^2} \ge f'\left(w^{P*}\right)$ for all $w \le w^{P*}$. And

$$f'\left(w^{P*}\right) = \frac{\overline{\theta}+\left(2\Phi^2-2\beta+\phi^2\right)w^{P*}}{2\beta-\phi^2} = \frac{\overline{\theta}\left(2\Phi^4+\beta\left(2\beta-\phi^2\right)\right)}{\left(2\beta-\phi^2\right)\left(\left(2\beta-\phi^2\right)\left(\beta+2\Phi^2\right)-2\Phi^4\right)} > 0$$

  due to assumption 1. So $f(w)$ is increasing for $w \le w^{P*}$, which once again implies $f\left(w^{A*}\right) < f\left(w^{H*}\right) < f\left(w^{P*}\right)$.

Thus, in all cases we have $\frac{d\Pi^{A*}}{d\Phi^2} < \frac{d\Pi^{H*}}{d\Phi^2} < \frac{d\Pi^{P*}}{d\Phi^2}$. Furthermore, we also have $\frac{d^2\Pi^{P*}}{d\Phi^2 dw}\Big|_{w=w^{P*}} = f'\left(w^{P*}\right) > 0$,

27

$\frac{d^2\Pi^{A*}}{d\Phi^2 dw}\big|_{w=w^{A*}} = f'\left(w^{A*}\right) > 0$ and $\frac{d^2\Pi^{H*}}{d\Phi^2 dw}\big|_{w=w^{H*}} = f'\left(w^{H*}\right) > 0$, which implies that w^{P*}, w^{A*} and w^{H*} are all increasing in Φ^2.

Platform minimum requirements: Online Appendix

Andrei Hagiu[*] and Julian Wright[†]

This online appendix provides the proof behind Proposition 6. It also provides the full working and proofs of the results in the extensions (Section 6).

A   Proof of Proposition 6

If $G\left(.\right)$ is uniform on $\left[\overline{\theta}-\sigma,\overline{\theta}+\sigma\right]$, then $G\left(\theta\right)=\frac{\theta-\overline{\theta}+\sigma}{2\sigma}$. The principal's profits in H-mode are then

$$
\begin{aligned}
\Pi^{H}\left(w,x\right) &= \int_{\overline{\theta}-\sigma}^{\frac{2\beta-\phi^2}{\phi}x+w\left(\beta-\Phi^2\right)}\left(\frac{1}{4\beta}\left(\theta+\phi x+w\left(\Phi^2-\beta\right)\right)^2+\frac{w}{2}\left(\theta+\phi x-\beta w\right)-\frac{1}{2}x^2\right)\frac{d\theta}{2\sigma}\\
&\quad+\int_{\frac{2\beta-\phi^2}{\phi}x+w\left(\beta-\Phi^2\right)}^{\overline{\theta}+\sigma}\left(\frac{1}{4\beta}\left(\theta+\phi q^A\left(w,\theta\right)+w\left(\Phi^2-\beta\right)\right)^2+\frac{w}{2}\left(\theta+\phi q^A\left(w,\theta\right)-\beta w\right)-\frac{1}{2}q^A\left(w,\theta\right)^2\right)\frac{d\theta}{2\sigma}\\
&= \int_{\overline{\theta}-\sigma}^{\frac{2\beta-\phi^2}{\phi}x+w\left(\beta-\Phi^2\right)}\left(\frac{1}{4\beta}\left(\theta+\phi x+w\left(\Phi^2-\beta\right)\right)^2+\frac{w}{2}\left(\theta+\phi x-\beta w\right)-\frac{1}{2}x^2\right)\frac{d\theta}{2\sigma}\\
&\quad+\int_{\frac{2\beta-\phi^2}{\phi}x+w\left(\beta-\Phi^2\right)}^{\overline{\theta}+\sigma}\frac{1}{2\left(2\beta-\phi^2\right)}\left(w^2\left(\Phi^4-2\Phi^2\beta+\Phi^2\phi^2-\beta^2\right)+2w\Phi^2\theta+\theta^2\right)\frac{d\theta}{2\sigma}.
\end{aligned}
$$

The first-order conditions in x and w are

$$
\left(\frac{2\beta-\phi^2}{\phi}x+w\left(\beta-\Phi^2\right)-\overline{\theta}+\sigma\right)\left(-\frac{2\beta-\phi^2}{\phi}x+w\left(\beta+\Phi^2\right)+\overline{\theta}-\sigma\right)=0
$$

and

$$
\begin{aligned}
&\int_{\overline{\theta}-\sigma}^{\frac{2\beta-\phi^2}{\phi}x+w\left(\beta-\Phi^2\right)}\left(\Phi^2\left(\theta+\phi x\right)+w\left(\left(\Phi^2-\beta\right)^2-2\beta^2\right)\right)\frac{d\theta}{2\beta}\\
&\quad+\int_{\frac{2\beta-\phi^2}{\phi}x+w\left(\beta-\Phi^2\right)}^{\overline{\theta}+\sigma}\frac{1}{2\beta-\phi^2}\left(w\left(\Phi^4-2\Phi^2\beta+\Phi^2\phi^2-\beta^2\right)+\Phi^2\theta\right)d\theta=0
\end{aligned}
$$

Suppose that the $\left(w^*,x^*\right)$ which maximizes Π^{H} is interior, i.e.

$$
\overline{\theta}-\sigma<\frac{2\beta-\phi^2}{\phi}x^*+w^*\left(\beta-\Phi^2\right)<\overline{\theta}+\sigma.
$$

Then the solution to the first order condition in x must be

$$
-\frac{2\beta-\phi^2}{\phi}x^*+w^*\left(\beta+\Phi^2\right)+\overline{\theta}-\sigma=0. \tag{A.1}
$$

This implies we must have

$$
0<w^*<\frac{\sigma}{\beta}.
$$

---
[*]MIT Sloan School of Management, Boston, MA 02142, E-mail: ahagiu@mit.edu

[†]Department of Economics, National University of Singapore, Singapore 117570, E-mail: jwright@nus.edu.sg

Plugging (A.1) into the first-order condition in w, we obtain that w^* must be a solution to

$$w^2 \beta^2 \phi^2 - 2\sigma \left(\beta^2 + 2\Phi^2\beta - \Phi^4 - \Phi^2\phi^2\right) w + 2\sigma\bar{\theta}\Phi^2 = 0 \tag{A.2}$$

This quadratic equation has real solutions if and only if $\sigma \geq \frac{2\bar{\theta}\Phi^2\beta^2\phi^2}{(\beta^2+2\Phi^2\beta-\Phi^4-\Phi^2\phi^2)^2}$. If $\sigma < \frac{2\bar{\theta}\Phi^2\beta^2\phi^2}{(\beta^2+2\Phi^2\beta-\Phi^4-\Phi^2\phi^2)^2}$, the function $f(w) \equiv \Pi^H\left(w, x = \frac{(w(\beta+\Phi^2)+\bar{\theta}-\sigma)\phi}{2\beta-\phi^2}\right)$ is weakly increasing in w for all $w < \frac{\sigma}{\beta}$, so the optimal (w^*, x^*) is non-interior and therefore the H-mode is dominated by either the P-mode or the A-mode.

Assume now $\sigma \geq \frac{2\bar{\theta}\Phi^2\beta^2\phi^2}{(\beta^2+2\Phi^2\beta-\Phi^4-\Phi^2\phi^2)^2}$ and denote the two solutions to (A.2) by

$$w_1 = \frac{\sigma}{\beta^2\phi^2}\left((\beta^2+2\Phi^2\beta-\Phi^4-\Phi^2\phi^2) - \sqrt{(\beta^2+2\Phi^2\beta-\Phi^4-\Phi^2\phi^2)^2 - \frac{2\bar{\theta}\Phi^2\beta^2\phi^2}{\sigma}}\right)$$

$$w_2 = \frac{\sigma}{\beta^2\phi^2}\left((\beta^2+2\Phi^2\beta-\Phi^4-\Phi^2\phi^2) + \sqrt{(\beta^2+2\Phi^2\beta-\Phi^4-\Phi^2\phi^2)^2 - \frac{2\bar{\theta}\Phi^2\beta^2\phi^2}{\sigma}}\right)$$

In this case, the function $f(w)$ is increasing for $w \in [0, w_1]$, decreasing for $w \in [w_1, w_2]$ and increasing again for $w \geq w_2$. Thus, the only candidate interior solution is $w^{H*} = w_1$. This solution is indeed interior if and only if $w_1 < \frac{\sigma}{\beta}$, which is equivalent to

$$\left(\beta^2 + \left(2\Phi^2 - \phi^2\right)\beta - \Phi^4 - \Phi^2\phi^2\right) < \sqrt{(\beta^2+2\Phi^2\beta-\Phi^4-\Phi^2\phi^2)^2 - \frac{2\bar{\theta}\Phi^2\beta^2\phi^2}{\sigma}}$$

The last inequality is equivalent to

$$\left(2\beta - \phi^2\right)\left(\frac{\beta}{2} + \Phi^2\right) - \Phi^4 \quad < \quad \beta\frac{\phi^2}{2}$$

$$\text{or}$$

$$\left(2\beta - \phi^2\right)\left(\frac{\beta}{2} + \Phi^2\right) - \Phi^4 \quad \geq \quad \beta\frac{\phi^2}{2} \text{ and } \sigma > \frac{\bar{\theta}\Phi^2\beta}{\left(2\beta - \phi^2\right)\left(\frac{\beta}{2} + \Phi^2\right) - \Phi^4}.$$

Combining this with the requirement that $\sigma \geq \frac{2\bar{\theta}\Phi^2\beta^2\phi^2}{(\beta^2+2\Phi^2\beta-\Phi^4-\Phi^2\phi^2)^2}$ and noting that

$$\frac{2\bar{\theta}\Phi^2\beta^2\phi^2}{(\beta^2+2\Phi^2\beta-\Phi^4-\Phi^2\phi^2)^2} < \frac{\bar{\theta}\Phi^2\beta}{\left(2\beta - \phi^2\right)\left(\frac{\beta}{2} + \Phi^2\right) - \Phi^4}$$

for all parameter values, we obtain that the optimal solution in H-mode is interior if and only if

$$\sigma \in \left[\frac{2\bar{\theta}\Phi^2\beta^2\phi^2}{(\beta^2+2\Phi^2\beta-\Phi^4-\Phi^2\phi^2)^2}, \frac{\bar{\theta}\Phi^2\beta}{\left(2\beta - \phi^2\right)\left(\frac{\beta}{2} + \Phi^2\right) - \Phi^4}\right] \text{ and } \left(2\beta - \phi^2\right)\left(\frac{\beta}{2} + \Phi^2\right) - \Phi^4 < \beta\frac{\phi^2}{2}$$

$$\text{or}$$

$$\sigma > \frac{\bar{\theta}\Phi^2\beta}{\left(2\beta - \phi^2\right)\left(\frac{\beta}{2} + \Phi^2\right) - \Phi^4}.$$

If the optimal H-solution is not interior, then the H-mode is dominated by the A-mode or the P-mode.

Finally, let us determine the effect of σ on profits. We have

$$\Pi^{H*} = \max_{w,x} \Pi^H (w,x) = \max_w \Pi^H \left(w, x = \frac{\left(w\left(\beta + \Phi^2\right) + \overline{\theta} - \sigma\right)\phi}{2\beta - \phi^2} \right)$$

We then have

$$\Pi^H \left(w, \frac{\left(w\left(\beta + \Phi^2\right) + \overline{\theta} - \sigma\right)\phi}{2\beta - \phi^2} \right) = \Pi^H \left(w, \frac{\left(w\left(\Phi^2 - \beta\right) + \overline{\theta} - \sigma\right)\phi}{2\beta - \phi^2} \right) + \int_{\frac{\left(w\left(\Phi^2 - \beta\right) + \overline{\theta} - \sigma\right)\phi}{2\beta - \phi^2}}^{\frac{\left(w\left(\beta + \Phi^2\right) + \overline{\theta} - \sigma\right)\phi}{2\beta - \phi^2}} \frac{\partial \Pi^H (w,x)}{\partial x} dx$$

$$= \int_{\overline{\theta} - \sigma}^{\overline{\theta} + \sigma} \frac{1}{2\left(2\beta - \phi^2\right)} \left(w^2\left(\Phi^4 - 2\Phi^2\beta + \Phi^2\phi^2 - \beta^2\right) + 2w\Phi^2\theta + \theta^2\right) \frac{d\theta}{2\sigma}$$

$$+ \int_{\frac{\left(w\left(\Phi^2 - \beta\right) + \overline{\theta} - \sigma\right)\phi}{2\beta - \phi^2}}^{\frac{\left(w\left(\beta + \Phi^2\right) + \overline{\theta} - \sigma\right)\phi}{2\beta - \phi^2}} \left(w^2\phi^2\beta^2 - \left(\phi\left(\overline{\theta} - \sigma\right) + w\phi\Phi^2 - \left(2\beta - \phi^2\right)x\right)^2\right) \frac{dx}{8\sigma\beta\phi}$$

$$= \frac{1}{2\left(2\beta - \phi^2\right)} \left(w^2\left(\Phi^4 - 2\Phi^2\beta + \Phi^2\phi^2 - \beta^2\right) + 2w\Phi^2\overline{\theta} + \overline{\theta}^2 + \frac{1}{3}\sigma^2\right) + \frac{w^3\beta^2\phi^2}{6\sigma\left(2\beta - \phi^2\right)}$$

$$= \frac{1}{2\left(2\beta - \phi^2\right)} \left(w^2\left(\Phi^4 - 2\Phi^2\beta + \Phi^2\phi^2 - \beta^2\right) + 2w\Phi^2\overline{\theta} + \overline{\theta}^2\right) + \frac{1}{6\left(2\beta - \phi^2\right)} \left(\sigma^2 + \frac{w^3\beta^2\phi^2}{\sigma}\right)$$

$$\equiv \widetilde{\Pi}^H (w, \sigma).$$

So, using the envelope theorem, we obtain

$$\frac{d\Pi^{H*}}{d\sigma} = \frac{\partial \widetilde{\Pi}^H \left(w = w^{H*}, \sigma\right)}{\partial \sigma} = \frac{2\sigma^3 - \left(w^{H*}\right)^3\beta^2\phi^2}{6\sigma^2\left(2\beta - \phi^2\right)} > 0,$$

where the last inequality follows from $w^{H*} < \frac{\sigma}{\beta}$ (recall $\left(x^{H*}, w^{H*}\right)$ must be interior) and assumption 1.

Since $\frac{d\Pi^{P*}}{d\sigma} = 0$ and $\frac{d\Pi^{A*}}{d\sigma} = \frac{\sigma}{3\left(2\beta - \phi^2\right)}$, we can conclude $\frac{d\Pi^{P*}}{d\sigma} < \frac{d\Pi^{H*}}{d\sigma} < \frac{d\Pi^{H*}}{d\sigma}$.

B   Resale price maintenance

In this section of the Online Appendix we provide the full analysis of the variant of the model in which the price p is transferable and is such that its choice can be restricted by the principal, while the costly actions q and Q are non-contractible and non-transferable.

Recall we make the assumptions (7) and (8). Assumption (7) ensures second-order conditions hold for all optimization problems we consider in this section. It requires β is sufficiently large or Φ^2 is sufficiently small. Assumption (8) ensures second-stage profits and decision variables are positive for all realizations of θ. This assumption always holds if $\Phi^2 \geq \beta$, or, in case $\beta > \Phi^2$, if θ_L is not too small.

We first analyze whether the principal prefers to set the level of p in its contract (P-mode) or entirely delegate that choice to the agent (A-mode), before considering whether the principal can do better than both pure modes through threshold delegation (H-mode).

B.1 Full control vs. no control

Consider first the P-mode. The fixed fee F is set to extract the entire expected surplus from the agent, so the principal solves

$$\max_{w,p} \left\{ \mathbb{E}_\theta \left[p \left(\theta - \beta p + \phi q + \Phi Q \right) - \frac{1}{2} q^2 \right] - \frac{1}{2} Q^2 \right\} \tag{B.1}$$
$$\text{subject to}$$
$$q = \phi \left(p - w \right) \text{ and } Q = \Phi w.$$

Substituting the two constraints solutions back into the profit function, and taking expectations over θ, the principal's profits are $\max_{w,p} \Pi^P \left(w, p \right)$, where

$$\Pi^P \left(w, p \right) \equiv p \left(\overline{\theta} - \left(\beta - \phi^2 \right) p + w \left(\Phi^2 - \phi^2 \right) \right) - \frac{1}{2} w^2 \Phi^2 - \frac{1}{2} \left(p - w \right)^2 \phi^2. \tag{B.2}$$

Note that whether demand is decreasing or increasing in price depends on whether β (the direct price effect) is higher or lower than ϕ^2 (the indirect price effect through the agent's investment in q). Even if demand is increasing in price, the profit maximization problem remains well behaved. This is because a higher price increases q and so the marginal cost of the agent's investment, which ensures profit eventually decreases with price provided $\phi^2 < 2\beta$, a condition which follows from (7). Optimizing $\Pi^P \left(w, p \right)$ over p first, we obtain that the price set by the principal for a given wholesale price w is $p^P \left(w, \overline{\theta} \right)$, where

$$p^P \left(w, \theta \right) = \frac{\theta + w \Phi^2}{2\beta - \phi^2}. \tag{B.3}$$

Thus, $p^P \left(w, \theta \right)$ is the hypothetical price the principal would want to set given w if it could observe θ. For comparing with the case in which the agent sets p (i.e. A-mode below), it is useful to note that this price is increasing in w. This is because a higher w induces the principal to invest more in stage 2, which in turn increases demand. Substituting $p^P \left(w, \overline{\theta} \right)$ back into $\Pi^P \left(w, p \right)$ and maximizing over w, we obtain the principal's optimal profit in P-mode:

$$\Pi^{P*} = \frac{\overline{\theta}^2 \left(\phi^2 + \Phi^2 \right)}{2 \left(\left(2\beta - \phi^2 \right) \left(\phi^2 + \Phi^2 \right) - \Phi^4 \right)}.$$

Consider now the A-mode. Again, the fixed fee (or salary) F is set such that the principal extracts the entire expected payoff in excess of the agent's outside option, so the principal solves

$$\max_{w} \left\{ \mathbb{E}_\theta \left[p \left(\theta - \beta p + \phi q + \Phi Q \right) - \frac{1}{2} q^2 \right] - \frac{1}{2} Q^2 \right\} \tag{B.4}$$
$$\text{subject to}$$
$$p = \arg\max_{p'} \left\{ \left(p' - w \right) \left(\theta - \beta p' + \phi q + \Phi Q \right) - \frac{1}{2} q^2 \right\} = \frac{\theta + \phi q + \Phi Q}{2\beta} + \frac{w}{2}$$
$$q = \phi \left(p - w \right) \text{ and } Q = \Phi w.$$

The additional constraint facing the principal compared to its P-mode problem is that the agent sets the price in the second stage optimally given the principal's wholesale price set in the first stage and the observed demand shock θ. Using all three constraints to solve for p as a function of (w, θ), we obtain

$$p^A \left(w, \theta \right) = \frac{\theta + \left(\Phi^2 + \beta - \phi^2 \right) w}{2\beta - \phi^2} \tag{B.5}$$

4

Comparing (B.3) and (B.5) reveals that for the same positive level of w, the agent has an upward bias in choosing the price (i.e. $p^A(w, \theta) > p^P(w, \theta)$) if $\beta > \phi^2$ and a downward bias in choosing the price if $\beta < \phi^2$. The existence of a bias (upward or downward) is due to the positive wholesale price, which means the agent only receives a fraction of the full revenues from its costly choice of q. This leads to choices of the agent that are distorted away from the levels preferred by the principal. Indeed, the absolute value of the bias is increasing in w. The direction of the bias is determined by whether demand is decreasing or increasing in price. If demand is decreasing in price ($\beta > \phi^2$), then the wholesale price w leads the agent to set p too high from the principal's perspective—the normal double marginalization effect dominates. On the other hand, if demand is increasing in price ($\beta < \phi^2$), then the wholesale price w leads the agent to set p too low from the principal's perspective. Indeed, in this case the effect that dominates is the fact that the agent does not fully internalize the benefit to the principal of a higher investment q which would be induced by setting a higher price.

Using $q = \phi\left(p^A(w, \theta) - w\right)$ and $Q = \Phi w$, substituting the three constraints back into the expression of the principal's profits, and maximizing over w, we obtain the principal's optimal A-mode profit:

$$\Pi^{A*} = \frac{\overline{\theta}^2 \left(\beta \left(2\Phi^2 + \beta\right) - \Phi^2\phi^2\right)}{2\left(2\beta - \phi^2\right)\left(\left(2\beta - \phi^2\right)\left(\Phi^2 + \phi^2\right) + \left(\beta - \phi^2\right)^2 - \Phi^4\right)} + \frac{V_\theta}{2\left(2\beta - \phi^2\right)}.$$

Comparing Π^{A*} with Π^{P*}, we obtain the following proposition.

Proposition. *(Choice of pure mode)*

The principal's profit is higher in A-mode compared to P-mode if and only if the variance of the agent's private information is sufficiently large, i.e.

$$\frac{V_\theta}{\overline{\theta}^2} > \frac{\Phi^4 \left(\beta - \phi^2\right)^2}{\left(\left(2\beta - \phi^2\right)\left(\phi^2 + \Phi^2\right) - \Phi^4\right)\left(\left(2\beta - \phi^2\right)\left(\Phi^2 + \phi^2\right) + \left(\beta - \phi^2\right)^2 - \Phi^4\right)}. \tag{B.6}$$

The inequality in (B.6) captures the key tradeoff between the two pure modes. On the one hand, the A-mode leverages the agent's private information on demand as captured by V_θ. On the other hand, the P-mode removes the distortion created by the agent setting p in A-mode (this explains why the right-hand side of (B.6) is positive).

B.2 Threshold delegation

Now suppose in addition to using a two-part tariff, the principal can monitor p and therefore restrict the agent's choice of p according to some rule (i.e. H-mode). We first determine sufficient conditions for the H-mode with threshold delegation to dominate both the A-mode and the P-mode. Subsequently, we will provide a sufficient condition for threshold delegation to be the optimal form of partial delegation.

As pointed out in Section B.1, the agent has an upward (respectively, downward) bias in A-mode relative to what the principal would set in P-mode if and only if $\beta > \phi^2$ (respectively, $\beta < \phi^2$). Thus, the relevant form of threshold delegation is that with a maximum threshold if $\beta > \phi^2$ and that with a minimum threshold if $\beta < \phi^2$.

Consider first the case $\beta > \phi^2$. In this case, given a wholesale price w and a maximum threshold x, the

agent chooses

$$
p = \begin{cases} p^A\left(w,\theta\right) = \frac{\theta + \left(\beta + \Phi^2 - \phi^2\right)w}{2\beta - \phi^2} & \text{if} \quad \theta \leq \left(2\beta - \phi^2\right)x - w\left(\beta + \Phi^2 - \phi^2\right) \\ \\ x & \text{if} \quad \theta \geq \left(2\beta - \phi^2\right)x - w\left(\beta + \Phi^2 - \phi^2\right). \end{cases}
$$

As in the pure modes, the principal extracts the agent's entire expected payoff through the fixed fee, so the principal's profit is

$$
\max_{w,x}\left\{ \mathbb{E}_\theta\left[p\left(\theta - \beta p + \phi q + \Phi Q\right) - \frac{1}{2}q^2 - \frac{1}{2}Q^2\right]\right\}
$$
$$
\text{subject to}
$$
$$
p = \min\left\{p^A\left(w,\theta\right),x\right\},\ q = \left(p - w\right)\phi,\ \text{and}\ Q = w\Phi.
$$

Substituting in the three constraints, we obtain that the principal's profit is $\max_{w,x} \Pi^H\left(w,x\right)$, where

$$
\Pi^H\left(w,x\right) \equiv \int_{\theta_L}^{\left(2\beta - \phi^2\right)x - w\left(\beta + \Phi^2 - \phi^2\right)} \frac{\left(\theta + w\Phi^2\right)^2 - w^2\left(\beta - \phi^2\right)^2}{2\left(2\beta - \phi^2\right)} dG\left(\theta\right) \tag{B.7}
$$
$$
+ \int_{\left(2\beta - \phi^2\right)x - w\left(\beta + \Phi^2 - \phi^2\right)}^{\theta_H} \frac{1}{2}x\left(2\theta - \left(2\beta - \phi^2\right)x + 2w\Phi^2\right)dG\left(\theta\right) - \frac{1}{2}w^2\left(\Phi^2 + \phi^2\right).
$$

If $\left(2\beta - \phi^2\right)x - w\left(\beta + \Phi^2 - \phi^2\right) \geq \theta_H$, then x places no effective constraint on the agent, who chooses $p = p^A\left(w,\theta\right)$ for all θ. This replicates the A-mode. Similarly, if $\left(2\beta - \phi^2\right)x - w\left(\beta + \Phi^2 - \phi^2\right) \leq \theta_L$, then the constraint on p is always binding, so $p = x$ for all θ. This is equivalent to the principal choosing $p = x$ contractually, i.e. the P-mode. As a result, the H-mode only refers to the case when (w,x) are "interior", i.e. such that

$$
\theta_L < \left(2\beta - \phi^2\right)x - w\left(\beta + \Phi^2 - \phi^2\right) < \theta_H. \tag{B.8}
$$

The principal's profit as a function of (w,x) is then $\Pi^H\left(w,x\right)$ given by expression (B.7) above.

Things are very similar when $\beta < \phi^2$ and the principal sets a minimum (rather than a maximum) threshold. The same calculations yield

$$
\Pi^H\left(w,x\right) = \int_{\left(2\beta - \phi^2\right)x - w\left(\beta + \Phi^2 - \phi^2\right)}^{\theta_H} \frac{\left(\theta + w\Phi^2\right)^2 - w^2\left(\beta - \phi^2\right)^2}{2\left(2\beta - \phi^2\right)} dG\left(\theta\right)
$$
$$
+ \int_{\theta_L}^{\left(2\beta - \phi^2\right)x - w\left(\beta + \Phi^2 - \phi^2\right)} \frac{1}{2}x\left(2\theta - \left(2\beta - \phi^2\right)x + 2w\Phi^2\right)dG\left(\theta\right) - \frac{1}{2}w^2\left(\Phi^2 + \phi^2\right)
$$

for interior (w,x), which is defined in the same way as in (B.8).

In both cases, the advantage of delegating to the agent is that the agent will take into account the realized value of θ when choosing p, so will set p closer to the first-best level, and the principal can extract this additional expected payoff through its fixed fee F. But the principal also needs to extract a positive margin (i.e. $w > 0$) in order to maintain an incentive to invest in Q. This in turn distorts the agent's choice of p, so the principal prefers to stipulate a maximum (respectively, minimum) level of p to help offset the upward (respectively, downward) bias, although at the cost of having p set too low (respectively, too high) whenever θ turns out to be particularly high (respectively, particularly low).

Thus, in some sense, threshold delegation would seem like a way for the principal to combine some of

the benefits of both delegation and control. The following proposition establishes the conditions under which threshold delegation dominates the two pure modes.[1]

Proposition. *(Resale price maintenance)*

Maximum RPM: If $\beta > \phi^2$, the H-mode with max RPM dominates the A-mode. If in addition $\frac{\theta_L}{\overline{\theta}} < 1 - \frac{\Phi^2(\beta - \phi^2)}{(2\beta - \phi^2)(\phi^2 + \Phi^2) - \Phi^4}$, the H-mode with max RPM also dominates the P-mode.

Minimum RPM: If $\beta < \phi^2$, the H-mode with min RPM dominates the A-mode. If in addition $\frac{\theta_H}{\overline{\theta}} > 1 + \frac{\Phi^2(\phi^2 - \beta)}{(2\beta - \phi^2)(\phi^2 + \Phi^2) - \Phi^4}$, the H-mode with min RPM also dominates the P-mode.

Proof. Solving the programs (B.1) and (B.4), it is straightforward to obtain the principal's optimal wholesale prices in P-mode and A-mode, respectively:

$$w^{P*} = \frac{\Phi^2\overline{\theta}}{(2\beta - \phi^2)(\Phi^2 + \phi^2) - \Phi^4} \text{ and } w^{A*} = \frac{\Phi^2\overline{\theta}}{(2\beta - \phi^2)(\Phi^2 + \phi^2) + (\beta - \phi^2)^2 - \Phi^4}.$$

Assumption (7) implies $w^{P*} > 0$ and $w^{A*} > 0$.

We start with the case $\beta > \phi^2$. The principal's H-mode profit is

$$\Pi^H(w, x) = \int_{\theta_L}^{(2\beta - \phi^2)x - w(\Phi^2 - \phi^2 + \beta)} \left(p^A(w, \theta)\left(\theta - \beta p^A(w, \theta) + w\Phi^2\right) + \frac{\phi^2}{2}p^A(w, \theta)^2 \right) dG(\theta)$$

$$+ \int_{(2\beta - \phi^2)x - w(\Phi^2 - \phi^2 + \beta)}^{\theta_H} \left(x\left(\theta - \beta x + w\Phi^2\right) + \frac{\phi^2}{2}x^2 \right) dG(\theta) - \frac{\Phi^2 + \phi^2}{2}w^2.$$

We first compare the A-mode to the H-mode with $w = w^{A*}$ and

$$x = p^A\left(w^{A*}, \theta_H - \kappa\right) = \frac{\theta_H - \kappa + w^{A*}\left(\Phi^2 - \phi^2 + \beta\right)}{2\beta - \phi^2},$$

where $\kappa > 0$ is small. We can then write

$$\Pi^H\left(w^{A*}, x = p^A\left(w^{A*}, \theta_H - \kappa\right)\right) - \Pi^{A*} = \Pi^H\left(w^{A*}, x = p^A\left(w^{A*}, \theta_H - \kappa\right)\right) - \Pi^H\left(w^{A*}, x = p^A\left(w^{A*}, \theta_H\right)\right)$$

$$= \int_{\theta_H - \kappa}^{\theta_H} \left(\begin{array}{c} p^A\left(w^{A*}, \theta_H - \kappa\right)\left(\theta - \beta p^A\left(w^{A*}, \theta_H - \kappa\right) + w^{A*}\Phi^2\right) + \frac{\phi^2}{2}p^A\left(w^{A*}, \theta_H - \kappa\right)^2 \\ -\left(p^A\left(w^{A*}, \theta\right)\left(\theta - \beta p^A\left(w^{A*}, \theta\right) + w^{A*}\Phi^2\right) + \frac{\phi^2}{2}p^A\left(w^{A*}, \theta\right)^2\right) \end{array} \right) dG(\theta)$$

$$= \int_{\theta_H - \kappa}^{\theta_H} \left(p^A\left(w^{A*}, \theta_H - \kappa\right) - p^A\left(w^{A*}, \theta\right)\right) \left(\begin{array}{c} \theta + w^{A*}\Phi^2 \\ -\left(\beta - \frac{\phi^2}{2}\right)\left(p^A\left(w^{A*}, \theta\right) + p^A\left(w^{A*}, \theta_H - \kappa\right)\right) \end{array} \right) dG(\theta)$$

$$= \int_{\theta_H - \kappa}^{\theta_H} \frac{1}{2(2\beta - \phi^2)}\left(\theta - (\theta_H - \kappa)\right)\left(2w^{A*}\left(\beta - \phi^2\right) + \theta_H - \kappa - \theta\right) dG(\theta).$$

Since $\beta > \phi^2$ and $w^{A*} > 0$, the last expression is positive for κ sufficiently small. Thus,

$$\Pi^{H*} \geq \Pi^H\left(w^{A*}, x = p^A\left(w^{A*}, \theta_H - \kappa\right)\right) > \Pi^{A*}.$$

This implies H-mode dominates A-mode.

---

[1]It is straightforward to verify that $\frac{\theta_L}{\overline{\theta}} < 1 - \frac{\Phi^2(\beta - \phi^2)}{(2\beta - \phi^2)(\phi^2 + \Phi^2) - \Phi^4}$ with $\beta > \phi^2$ is not incompatible with assumption (8).

Next, we compare the P-mode to the H-mode with $w = w^{P*}$ and $x = p^{P*} = \frac{\overline{\theta} + \Phi^2 w^{P*}}{2\beta - \phi^2}$. Note that this $\left(w^H, x\right)$ is strictly interior because

$$\theta_L < \left(2\beta - \phi^2\right) p^{P*} - w^{P*} \left(\Phi^2 - \phi^2 + \beta\right) = \overline{\theta} - w^{P*} \left(\beta - \phi^2\right) < \theta_H,$$

where the first inequality follows from the assumption $\theta_L < \overline{\theta}\left(1 - \frac{\Phi^2\left(\beta - \phi^2\right)}{\left(2\beta - \phi^2\right)\left(\phi^2 + \Phi^2\right) - \Phi^4}\right)$ and the second inequality from $\beta > \phi^2$. Then, using (B.2) we obtain

$$\Pi^H \left(w^{P*}, x = p^{P*}\right) - \Pi^{P*} = \Pi^H \left(w^{P*}, p^{P*}\right) - \Pi^P \left(w^{P*}, p^{P*}\right)$$

$$= \int_{\theta_L}^{\left(2\beta - \phi^2\right) p^{P*} - w^{P*}\left(\Phi^2 - \phi^2 + \beta\right)} \left(\begin{array}{c} p^A \left(w^{P*}, \theta\right) \left(\theta - \beta p^A \left(w^{P*}, \theta\right) + w^{P*}\Phi^2\right) + \frac{\phi^2}{2} p^A \left(w^{P*}, \theta\right)^2 \\ - \left(p^{P*} \left(\theta - \beta p^{P*} + w^{P*}\Phi^2\right) + \frac{\phi^2}{2} \left(p^{P*}\right)^2\right) \end{array} \right) dG\left(\theta\right)$$

$$= \int_{\theta_L}^{\overline{\theta} - w^{P*}\left(\beta - \phi^2\right)} \left(\left(p^A \left(w^{P*}, \theta\right) - p^{P*}\right) \left(\theta + w^{P*}\Phi^2 - \left(\beta - \frac{\phi^2}{2}\right) \left(p^A \left(w^{P*}, \theta\right) + p^{P*}\right)\right) \right) dG\left(\theta\right)$$

$$= \int_{\theta_L}^{\overline{\theta} - w^{P*}\left(\beta - \phi^2\right)} \frac{1}{2\left(2\beta - \phi^2\right)} \left(\overline{\theta} - \theta - w^{P*}\left(\beta - \phi^2\right)\right) \left(\overline{\theta} - \theta + w^{P*}\left(\beta - \phi^2\right)\right) dG\left(\theta\right) > 0,$$

because $\beta > \phi^2$ in this case. This implies H-mode strictly dominates P-mode.

Turning now to the case $\beta < \phi^2$, we can use the same steps as above to show that $\Pi^H \left(w^{A*}, x = p^{A*} \left(w^{A*}, \theta_L + \kappa\right)\right) > \Pi^{A*}$ for κ sufficiently small, which implies that H-mode dominates A-mode. Similarly, we can show that $\Pi^H \left(w^{P*}, x = p^{P*}\right) > \Pi^{P*}$. In this case, $\left(w^H = w^{P*}, x = p^{P*}\right)$ is strictly interior because

$$\theta_L < \left(2\beta - \phi^2\right) p^{P*} - w^{P*} \left(\Phi^2 - \phi^2 + \beta\right) = \overline{\theta} + w^{P*} \left(\phi^2 - \beta\right) < \theta_H$$

due to the assumptions $\phi^2 > \beta$ and $\theta_H > \overline{\theta}\left(1 + \frac{\Phi^2\left(\phi^2 - \beta\right)}{\left(2\beta - \phi^2\right)\left(\phi^2 + \Phi^2\right) - \Phi^4}\right)$. Thus, H-mode dominates P-mode under these assumptions.

$\square$

Proposition 7 established conditions under which the principal prefers threshold delegation to the two pure modes. However, there are more complex forms of partial delegation that the principal could utilize (e.g. delegating subject to requirement that price be in one of two disjoint intervals). The next proposition provides a sufficient condition for the threshold delegation described in Proposition 7 to be the optimal form of delegation.

Proposition. *(Optimality of threshold delegation)*
If $g'\left(\theta\right) \leq 0$ for all $\theta \in \left[\theta_L, \theta_H\right]$, the optimal contract in H-mode involves threshold delegation.

Proof. Since the principal fixes w in its contract at the same time as deciding on the type and nature of any delegation, we just have to show threshold delegation is optimal for any given w. To do so, we will show that any contract that differs from threshold delegation can be improved upon by a contract with the same w and threshold delegation.

The principal's delegation problem for a fixed choice of w is

$$\max_D \left\{ \mathbb{E}\left[p\left(\theta - \beta p + \phi q + \Phi Q\right) - \frac{1}{2}q^2 - \frac{1}{2}Q^2 \right] \right\}$$

subject to

$$(p, q) = \arg\max_{p' \in D, q} \left\{ (p' - w)\left(\theta - \beta p' + \phi q + \Phi Q\right) - \frac{1}{2}q^2 \right\}$$

$$Q = w\Phi,$$

where D denotes the delegation set to which the principal restricts the agent's choice of p. This can be rewritten as

$$\max_D \left\{ \mathbb{E}\left[p\left(\theta - \beta p + (p - w)\phi^2 + w\Phi^2\right) - \frac{1}{2}(p - w)^2\phi^2 - \frac{1}{2}w^2\Phi^2 \right] \right\}$$

subject to

$$p = \arg\max_{p' \in D} \left\{ (p' - w)\left(\theta - \beta p' + (p' - w)\phi^2 + w\Phi^2\right) - \frac{1}{2}(p' - w)^2\phi^2 \right\}.$$

Ignoring terms that do not depend on p, the program that defines the principal's optimal delegation set $D(t)$ can be re-written more simply as

$$\max_D \mathbb{E}\left[-\alpha_0 p^2 + \left(\theta + w\Phi^2\right)p \right]$$

$$\text{subject to } p = \arg\max_{p' \in D} \left\{ -\alpha_0\left(p'\right)^2 + \left(\theta + w\Phi^2 - \alpha_1\right)p' \right\},$$

where $\alpha_0 = \beta - \frac{\phi^2}{2} > 0$ due to our assumption that $\left(2\beta - \phi^2\right)\left(\Phi^2 + \phi^2\right) - \Phi^4 > 0$ and $\alpha_1 = w\left(\phi^2 - \beta\right)$.

In this model, for a given θ, the ideal choice of p for the principal is $p = \frac{\theta + w\Phi^2}{2\alpha_0}$, while for the agent it is $p = \frac{\theta + w\Phi^2 - \alpha_1}{2\alpha_0}$. Note that α_1 is positive or negative depending on whether the agent has a downward or upward bias in setting price. We will first consider the case in which $\alpha_1 > 0$, so that without any restrictions, the agent prefers a lower p than the principal. The principal is therefore interested in restricting the agent from setting p too low. The question remains whether the principal can do better by requiring the agent to choose from some specific values of p or some disjoint intervals that exclude some high values of p. Formally, we want to show that the delegation set $D(w)$ is a threshold interval, i.e. $D(w) = \{p \geq x(w)\}$ for some $x(w)$.

Suppose first the principal restricts the agent to choose p from some subset of $p \leq p_0 \equiv \frac{\theta_0 + w\Phi^2 - \alpha_1}{2\alpha_0}$ which includes $p = p_0$, where $\theta_0 < \theta_H$. This covers the possibility that the agent can only choose $p = p_0$ or can choose any $p \leq p_0$. In this case, when $\theta \in [\theta_0, \theta_H]$, it is easily seen that the agent chooses $p = p_0$ because the agent's objective function is increasing in p for all $p \leq p_0$. But the principal could strictly improve expected profits by adding the range $p \geq p_0$ to the set of permissible choices of p by the agent. To see this, note that the only change comes from the different choices of p by the agent when $\theta \in [\theta_0, \theta_H]$. The change in expected profits is

$$\int_{\theta_0}^{\theta_H} \left(\begin{array}{c} -\alpha_0\left(\frac{\theta + w\Phi^2 - \alpha_1}{2\alpha_0}\right)^2 + \left(\theta + w\Phi^2\right)\left(\frac{\theta + w\Phi^2 - \alpha_1}{2\alpha_0}\right) \\ -\left(-\alpha_0\left(\frac{\theta_0 + w\Phi^2 - \alpha_1}{2\alpha_0}\right)^2 + \left(\theta + w\Phi^2\right)\left(\frac{\theta_0 + w\Phi^2 - \alpha_1}{2\alpha_0}\right)\right) \end{array} \right) dG(\theta)$$

$$= \frac{1}{4\alpha_0}\int_{\theta_0}^{\theta_H}\left(\theta - \theta_0\right)\left(\left(\theta - \theta_0\right) + 2\alpha_1\right)dG(\theta) > 0.$$

Suppose now that the agent is allowed to choose from some set that does not include $p \in (p_0, p_1)$, where

9

$p_0 \equiv \frac{\theta_0 + w\Phi^2 - \alpha_1}{2\alpha_0}$ and $p_1 \equiv \frac{\theta_1 + w\Phi^2 - \alpha_1}{2\alpha_0}$ for some $\theta_L \leq \theta_0 < \theta_1 \leq \theta_H$. In this case, since the agent's objective function is quadratic in p, if the agent's draw of θ is in the range $[\theta_0, \theta_1]$, then the agent chooses $p = p_0$ when $\theta \leq \frac{\theta_0 + \theta_1}{2}$ and $p = p_1$ when $\theta > \frac{\theta_0 + \theta_1}{2}$. The principal can profitably deviate by adding the range $[p_0, p_1]$ to the set of permissible choices of p. The change in profits is

$$\frac{1}{4\alpha_0} \int_{\theta_0}^{\frac{\theta_0 + \theta_1}{2}} (\theta - \theta_0)(\theta - \theta_0 + 2\alpha_1) \, dG(\theta) + \frac{1}{4\alpha_0} \int_{\frac{\theta_0 + \theta_1}{2}}^{\theta_1} (\theta - \theta_1)(\theta - \theta_1 + 2\alpha_1) \, dG(\theta)$$

$$= \frac{1}{4\alpha_0} \left(\int_{\theta_0}^{\frac{\theta_0 + \theta_1}{2}} \left(2\alpha_1 (\theta - \theta_0) + (\theta - \theta_0)^2 \right) g(\theta) \, d\theta + \int_{\frac{\theta_0 + \theta_1}{2}}^{\theta_1} \left(2\alpha_1 (\theta - \theta_1) + (\theta - \theta_1)^2 \right) g(\theta) \, d\theta \right).$$

Using integration by parts, we have $\int_{\theta_0}^{\frac{\theta_0 + \theta_1}{2}} (\theta - \theta_0)^2 (g(\theta) - \alpha_1 g'(\theta)) \, d\theta.$]

$$\int_{\theta_0}^{\frac{\theta_0 + \theta_1}{2}} \left(2\alpha_1 (\theta - \theta_0) + (\theta - \theta_0)^2 \right) g(\theta) \, d\theta = \alpha_1 \frac{(\theta_1 - \theta_0)^2}{4} g\left(\frac{\theta_0 + \theta_1}{2} \right) + \int_{\theta_0}^{\frac{\theta_0 + \theta_1}{2}} (g(\theta) - \alpha_1 g'(\theta)) \, d\theta$$

$$\int_{\frac{\theta_0 + \theta_1}{2}}^{\theta_1} \left(2\alpha_1 (\theta - \theta_1) + (\theta - \theta_1)^2 \right) g(\theta) \, d\theta = -\alpha_1 \frac{(\theta_1 - \theta_0)^2}{4} g\left(\frac{\theta_0 + \theta_1}{2} \right) + \int_{\frac{\theta_0 + \theta_1}{2}}^{\theta_1} (\theta - \theta_1)^2 (g(\theta) - \alpha_1 g'(\theta)) \, d\theta.$$

With these expressions, the profit change above becomes equal to

$$\frac{1}{4\alpha_0} \left(\int_{\theta_0}^{\frac{\theta_0 + \theta_1}{2}} (\theta - \theta_0)^2 (g(\theta) - \alpha_1 g'(\theta)) \, d\theta + \int_{\frac{\theta_0 + \theta_1}{2}}^{\theta_1} (\theta - \theta_1)^2 (g(\theta) - \alpha_1 g'(\theta)) \, d\theta \right).$$

This expression is clearly positive under the assumption $\alpha_1 g'(\theta) \leq g(\theta)$ for all θ.

Thus, we can conclude that the optimal range of admissible p for the agent must take the form of a threshold interval $p \geq x$. Since $\alpha_1 = w(\phi^2 - \beta) > 0$, threshold delegation is optimal provided

$$w(\phi^2 - \beta) g'(\theta) \leq g(\theta). \tag{B.9}$$

Since $w > 0$ and $g(\theta) > 0$ for all $\theta \in [\theta_L, \theta_H]$, a sufficient condition for (B.9) to hold is $g'(\theta) \leq 0$.

If instead $\alpha_1 < 0$, then the same proof applies, except that now the optimal range of admissible p for the agent must take the form of a threshold interval with a maximum requirement $p \leq x$. $\qquad \square$

C   General demand and cost functions

Since (9) is assumed to be concave in q, we know that $q^P(w, \theta)$ and $q^A(w, \theta)$ are determined by the respective first-order conditions. Thus, the derivative of $\{(p(q, w, \theta) - w) D(\theta, p(q, w, \theta), q, Q(w)) - c(q)\}$ with respect to q evaluated at $q^P(w, \theta)$ is

$$-w \frac{dD(\theta, p(q, w, \theta), q, Q(w))}{dq} \bigg|_{q = q^P(w, \theta)} = -w \frac{dD^1(\theta, p(q, w, \theta), q)}{dq} \bigg|_{q = q^P(w, \theta)} < 0,$$

given the assumption that $\frac{dD^1(\theta, p(q, w, \theta), q)}{dq} > 0$. Thus, $q^A(w, \theta) < q^P(w, \theta)$ for all w. In particular, at $w = w^{P*}$ and $w = w^{A*}$, we have $q^A(w^{P*}, \theta) < q^P(w^{P*}, \theta)$ and $q^A(w^{A*}, \theta) < q^P(w^{A*}, \theta)$ for all θ.

Consider then the H-mode with a minimum threshold on q:

$$\max_{w,x} \{\mathbb{E}\left[pD\left(\theta,p,q,Q\right)-c\left(q\right)-C\left(Q\right)\right]\}$$
$$\text{subject to}$$
$$q = \max\left\{q^{A}\left(w,\theta\right),x\right\}$$
$$p = p\left(q,w,\theta\right) \text{ and } Q=Q\left(w\right),$$

where the functions $Q\left(w\right)$, $p\left(q,w,\theta\right)$ and $q^{A}\left(w,\theta\right)$ are as defined in Section 6.2 in the main text.

Compare H-mode to A-mode with $\left(w=w^{A*},x=q^{A}\left(w^{A*},\theta_{L}+\kappa\right)\right)$ for small κ:

$$\Pi^{H}\left(w^{A*},x=q^{A}\left(w^{A*},\theta_{L}+\kappa\right)\right)-\Pi^{A*}=\Pi^{H}\left(w^{A*},x=q^{A}\left(w^{A*},\theta_{L}+\kappa\right)\right)-\Pi^{H}\left(w^{A*},x=q^{A}\left(w^{A*},\theta_{L}\right)\right)$$
$$=\int_{\theta_{L}}^{\theta_{L}+\kappa}y\left(\theta,\kappa\right)dG\left(\theta\right),$$

where

$$y\left(\theta,\kappa\right)\equiv p\left(q^{A}\left(w^{A*},\theta_{L}+\kappa\right),w^{A*},\theta\right)D\left(\theta,p\left(q^{A}\left(w^{A*},\theta_{L}+\kappa\right),w^{A*},\theta\right),q^{A}\left(w^{A*},\theta_{L}+\kappa\right),Q\left(w^{A*}\right)\right)$$
$$-c\left(q^{A}\left(w^{A*},\theta_{L}+\kappa\right)\right)$$
$$-\left(\begin{array}{c}p\left(q^{A}\left(w^{A*},\theta\right),w^{A*},\theta\right)D\left(\theta,p\left(q^{A}\left(w^{A*},\theta\right),w^{A*},\theta\right),q^{A}\left(w^{A*},\theta\right),Q\left(w^{A*}\right)\right)\\-c\left(q^{A}\left(w^{A*},\theta\right)\right)\end{array}\right).$$

We have

$$\frac{\partial y\left(\theta,\kappa\right)}{\partial\theta}=\left(\begin{array}{c}p\left(q^{A}\left(w^{A*},\theta_{L}+\kappa\right),w^{A*},\theta\right)\frac{\partial D\left(\theta,p\left(q^{A}\left(w^{A*},\theta_{L}+\kappa\right),w^{A*},\theta\right),q^{A}\left(w^{A*},\theta_{L}+\kappa\right),Q\left(w^{A*}\right)\right)}{\partial\theta}\\-p\left(q^{A}\left(w^{A*},\theta\right),w^{A*},\theta\right)\frac{\partial D\left(\theta,p\left(q^{A}\left(w^{A*},\theta\right),w^{A*},\theta\right),q^{A}\left(w^{A*},\theta\right),Q\left(w^{A*}\right)\right)}{\partial\theta}\end{array}\right)$$

$$+\left(\begin{array}{c}\left(p\left(q^{A}\left(w^{A*},\theta_{L}+\kappa\right),w^{A*},\theta\right)\frac{\partial D\left(\theta,p\left(q^{A}\left(w^{A*},\theta_{L}+\kappa\right),w^{A*},\theta\right),q^{A}\left(w^{A*},\theta_{L}+\kappa\right),Q\left(w^{A*}\right)\right)}{\partial p}\right)\frac{\partial p\left(q^{A}\left(w^{A*},\theta_{L}+\kappa\right),w^{A*},\theta\right)}{\partial\theta}\\-\left(p\left(q^{A}\left(w^{A*},\theta\right),w^{A*},\theta\right)\frac{\partial D\left(\theta,p\left(q^{A}\left(w^{A*},\theta\right),w^{A*},\theta\right),q^{A}\left(w^{A*},\theta\right),Q\left(w^{A*}\right)\right)}{\partial p}\right)\frac{\partial p\left(q^{A}\left(w^{A*},\theta\right),w^{A*},\theta\right)}{\partial\theta}\end{array}\right)$$

$$+\left(\begin{array}{c}\frac{\partial p\left(q^{A}\left(w^{A*},\theta_{L}+\kappa\right),w^{A*},\theta\right)}{\partial\theta}D\left(\theta,p\left(q^{A}\left(w^{A*},\theta_{L}+\kappa\right),w^{A*},\theta\right),q^{A}\left(w^{A*},\theta_{L}+\kappa\right),Q\left(w^{A*}\right)\right)\\-\frac{\partial p\left(q^{A}\left(w^{A*},\theta\right),w^{A*},\theta\right)}{\partial\theta}\partial D\left(\theta,p\left(q^{A}\left(w^{A*},\theta\right),w^{A*},\theta\right),q^{A}\left(w^{A*},\theta\right),Q\left(w^{A*}\right)\right)\end{array}\right)$$

$$-\left(\frac{d\left(p\left(q,w^{A*},\theta\right)D\left(\theta,p\left(q,w^{A*},\theta\right),q,Q\left(w^{A*}\right)\right)\right)}{dq}-c'\left(q\right)\right)_{q=q^{A}\left(w^{A*},\theta\right)}\frac{\partial q^{A}\left(w^{A*},\theta\right)}{\partial\theta}.$$

The first three terms in the expression of $\frac{\partial y(\theta,\kappa)}{\partial\theta}$ above vanish as $\kappa\to0$. Meanwhile, $\frac{\partial q^{A}\left(w^{A*},\theta\right)}{\partial\theta}>0$ by assumption, and

$$\left.\frac{d\left(p\left(q,w^{A*},\theta\right)D\left(\theta,p\left(q,w^{A*},\theta\right),q,Q\left(w^{A*}\right)\right)-c\left(q\right)\right)}{dq}\right|_{q=q^{A}\left(w^{A*},\theta\right)}>0$$

because $q^{A}\left(w^{A*},\theta\right)<q^{P}\left(w^{A*},\theta\right)$ (downward bias), $p\left(q,w^{A*},\theta\right)D\left(\theta,p\left(q,w^{A*},\theta\right),q,Q\left(w^{A*}\right)\right)-c\left(q\right)$ is con-

11

cave in q by assumption, and

$$\frac{d\left(p\left(q,w^{A*},\theta\right)D\left(\theta,p\left(q,w^{A*},\theta\right),q,Q\left(w^{A*}\right)\right)-c\left(q\right)\right)}{dq}\Bigg|_{q=q^{P}\left(w^{A*},\theta\right)}=0$$

by definition of $q^{P}\left(w^{A*},\theta\right)$. Thus, we have $\frac{\partial y(\theta,\kappa)}{\partial\theta}<0$ for all $\theta\in[\theta_{L},\theta_{L}+\kappa]$ provided κ is sufficiently small. Furthermore, $y\left(\theta_{L}+\kappa,\kappa\right)=0$. Thus, if κ is sufficiently small, $y\left(\theta,\kappa\right)>0$ for all $\theta\in\left(\theta_{L},\theta_{L}+\kappa\right)$ and therefore

$$\Pi^{H}\left(w^{A*},x=q^{A}\left(w^{A*},\theta_{L}+\kappa\right)\right)>\Pi^{A*}.$$

This means the H-mode dominates the A-mode.

Now compare the H-mode with $\left(w^{H}=w^{P*},x=q^{P*}\right)$ to the P-mode. If $q^{A}\left(w^{P*},\theta_{H}\right)>q^{P*}$, there exists a unique $\theta^{P*}\in\left(\theta_{L},\theta_{H}\right)$ such that $q^{A}\left(w^{P*},\theta^{P*}\right)=q^{P*}$ (indeed, recall the agent is assumed to have a downward bias, so $q^{A}\left(w^{P*},\theta_{L}\right)<q^{P}\left(w^{P*},\theta_{L}\right)\leq q^{P*}$). We can then write

$$\Pi^{H}\left(w^{P*},x=q^{P*}\right)=\int_{\theta_{L}}^{\theta^{P*}}\left(\begin{array}{c}p\left(q^{P*},w^{P*},\theta\right)D\left(\theta,p\left(q^{P*},w^{P*},\theta\right),q^{P*},Q\left(w^{P*}\right)\right)\\-c\left(q^{P*}\right)-C\left(Q\left(w^{P*}\right)\right)\end{array}\right)dG\left(\theta\right)$$
$$+\int_{\theta^{P*}}^{\theta_{H}}\left(\begin{array}{c}p\left(q^{A*}\left(w^{P*},\theta\right),w^{P*},\theta\right)D\left(\theta,p\left(q^{A*}\left(w^{P*},\theta\right),w^{P*},\theta\right),q^{A*}\left(w^{P*},\theta\right),Q\left(w^{P*}\right)\right)\\-c\left(q^{A*}\left(w^{P*},\theta\right)\right)-C\left(Q\left(w^{P*}\right)\right)\end{array}\right)dG\left(\theta\right)$$

For all $\theta>\theta^{P*}$, we have $q^{P*}<q^{A}\left(w^{P*},\theta\right)<q^{P}\left(w^{P*},\theta\right)$. And since

$$p\left(q,w^{P*},\theta\right)D\left(\theta,p\left(q,w^{P*},\theta\right),q,Q\left(w^{P*}\right)\right)-c\left(q\right)-C\left(Q\left(w^{P*}\right)\right)$$

is concave in q and maximized by $q=q^{P}\left(w^{P*},\theta\right)$, we have

$$p\left(q^{P}\left(w^{P*},\theta\right),w^{P*},\theta\right)D\left(\theta,p\left(q^{P}\left(w^{P*},\theta\right),w^{P*},\theta\right),q^{P}\left(w^{P*},\theta\right),Q\left(w^{P*}\right)\right)$$
$$-c\left(q^{P}\left(w^{P*},\theta\right)\right)-C\left(Q\left(w^{P*}\right)\right)$$
$$>\quad p\left(q^{A}\left(w^{P*},\theta\right),w^{P*},\theta\right)D\left(\theta,p\left(q^{A}\left(w^{P*},\theta\right),w^{P*},\theta\right),q^{A}\left(w^{P*},\theta\right),Q\left(w^{P*}\right)\right)$$
$$-c\left(q^{A}\left(w^{P*},\theta\right)\right)-C\left(Q\left(w^{P*}\right)\right)$$
$$>\quad p\left(q^{P*},w^{P*},\theta\right)D\left(\theta,p\left(q^{P*},w^{P*},\theta\right),q^{P*},Q\left(w^{P*}\right)\right)$$
$$-c\left(q^{P*}\right)-C\left(Q\left(w^{P*}\right)\right).$$

Thus,

$$\Pi^{H}\left(w^{P*},x=q^{P*}\right)>\int_{\theta_{L}}^{\theta^{P*}}\left(\begin{array}{c}p\left(q^{P*},w^{P*},\theta\right)D\left(\theta,p\left(q^{P*},w^{P*},\theta\right),q^{P*},Q\left(w^{P*}\right)\right)\\-c\left(q^{P*}\right)-C\left(Q\left(w^{P*}\right)\right)\end{array}\right)dG\left(\theta\right)$$
$$+\int_{\theta^{P*}}^{\theta_{H}}\left(\begin{array}{c}p\left(q^{P*},w^{P*},\theta\right)D\left(\theta,p\left(q^{P*},w^{P*},\theta\right),q^{P*},Q\left(w^{P*}\right)\right)\\-c\left(q^{P*}\right)-C\left(Q\left(w^{P*}\right)\right)\end{array}\right)dG\left(\theta\right)$$
$$=\quad\Pi^{P*}$$

So H-mode also dominates P-mode.

D Private benefits

In this section we show that our analysis in the main paper also applies to the case in which the agent's private information regards his private benefits (or costs). We adopt the approach as in Section 6.2 of the main paper, namely we assume demand is additively separable in Q, i.e.

$$D\left(p, q, Q\right) = D^{1}\left(p, q\right) + D^{2}\left(Q\right).$$

This implies that, given (q, w, b), the non-contractible and non-transferable decisions $\left(p\left(q, w, b\right), Q\left(w\right)\right)$ are the joint solutions to

$$\begin{cases} p = \arg\max_{p'} \left\{ \left(p' + b - w\right) D^{1}\left(p', q\right) \right\} \\ Q = \arg\max_{Q'} \left\{ w D^{2}\left(Q'\right) - C\left(Q'\right) \right\} \end{cases}.$$

Thus, $\left(p\left(q, w, b\right), Q\left(w\right)\right)$ is the Nash equilibrium of the game in which the agent and the principal set p and Q respectively, given q and w. Let then

$$\begin{aligned} q^{P}\left(w, b\right) &\equiv \arg\max_{q'} \left\{ \left(p\left(q', w, b\right) + b\right) D\left(p\left(q', w, b\right), q', Q\left(w\right)\right) - c\left(q'\right) - C\left(Q\left(w\right)\right) \right\} \\ q^{A}\left(w, b\right) &\equiv \arg\max_{q'} \left\{ \left(p\left(q', w, b\right) + b - w\right) D\left(p\left(q', w, b\right), q', Q\left(w\right)\right) - c\left(q'\right) \right\}. \end{aligned}$$

We assume that total profit,

$$\left(p\left(q, w, b\right) + b\right) D\left(p\left(q, w, b\right), q, Q\left(w\right)\right) - c\left(q\right) - C\left(Q\left(w\right)\right),$$

is concave in q for any (w, b). Denoting by q^{P*} and w^{P*} the principal's optimal choices of q and wholesale price in P-mode and by w^{A*} the principal's optimal choice of wholesale price in A-mode, it is easily verified that we obtain the same results as in Proposition 8, replacing θ by b (the proof is almost identical, so we omit it).

Now consider the linear-quadratic model of Section 4 but with private benefits rather than a demand shock (so θ is now a constant). Given a wholesale price w and a realization b of the agent's private benefit, the principal's ideal choice of q is

$$q^{P}\left(w, b\right) = \frac{\left(\theta + w\Phi^{2} + b\beta\right)\phi}{2\beta - \phi^{2}},$$

whereas the agent's ideal choice of q is

$$q^{A}\left(w, b\right) = \frac{\left(\theta + w\Phi^{2} + b\beta - w\beta\right)\phi}{2\beta - \phi^{2}}.$$

Thus, the agent has a downwards bias (the extra term $\frac{-w\beta\phi}{2\beta - \phi^{2}}$ in $q^{A}\left(w, b\right)$), and the magnitude of the bias is identical to that in the model with demand uncertainty.