Operations Research

Information Elicitation and Influenza Vaccine Production

Stephen E. Chick, Sameer Hasija, Javad Nasiry

Information Elicitation and Influenza Vaccine Production

Stephen E. Chick,[a] Sameer Hasija,[b] Javad Nasiry[c]

[a] INSEAD, Technology and Operations Management Area, 77300 Fontainebleau France; [b] INSEAD, Technology and Operations Management Area, 138676 Singapore; [c] Hong Kong University of Science and Management, School of Business and Management, Clear Water Bay, Kowloon, Hong Kong

Contact: stephen.chick@insead.edu (SEC); sameer.hasija@insead.edu (SH); nasiry@ust.hk (JN)

Abstract. We explore the procurement of influenza vaccines by a government whose objective is to minimize the expected social costs (including vaccine, vaccine administration, and influenza treatment costs) when a for-profit vaccine supplier has production yield uncertainty, private information about its productivity (adverse selection), and potentially unverifiable production effort (moral hazard). Timeliness is important—costs for both the supplier and the government procurer may increase if part of the vaccine order is delivered after a scheduled delivery date. We theoretically derive the optimal menu of output-based contracts. Next, we present a menu that is optimal within a more restricted set of practically implementable contracts, and numerically show that such a menu leads to near-optimal outcomes. Finally, we present a novel way to eliminate that information rent if the manufacturer's effort is also verifiable, a counterintuitive result because the manufacturer has private productivity information. This provides an upper bound for the government on how much it should spend to monitor the manufacturer's effort.

Keywords: mechanism design • principal-agent modeling • adverse selection • moral hazard • epidemiology • influenza vaccine supply chain

1. Introduction

Influenza is a respiratory illness that spreads rapidly in seasonal epidemics (e.g., see http://www.cdc.gov/flu). Influenza outbreaks result in 300,000 to 500,000 deaths around the world annually. The World Health Organization (WHO 2008) reports that the annual costs of healthcare, lost days of work and schooling, and social disruption range between $1 million and $6 million per 100,000 inhabitants in industrialized countries. Vaccination is a primary means of preventing influenza (Gerdil 2003, Saluzzo and Lacroix-Gerdil 2006), can be deployed in programmes optimized to limit influenza outbreaks (Nichol et al. 1994, Weycker et al. 2005, WHO 2008), and can be complemented with antiviral therapy, social distancing, or other interventions (Sun et al. 2009, Wu et al. 2009).

This paper explores the acquisition of influenza vaccines by a government whose objective is to minimize the expected total social costs of influenza when a for-profit manufacturer of influenza vaccines has yield uncertainty in production (U.S. GAO 2001) and private information about its productivity. The government seeks to minimize its expected total health expenditure related to influenza, including vaccine procurement from a manufacturer, vaccine administration costs, and the cost of treating those that are ultimately infected.

While we use the term "government" for the vaccine procurer, the model is also conceptually valid for a healthcare management organization that procures influenza vaccines for a population in an effort to avert infections and infection transmission in that population, and for which it is liable for treatment costs (as might be consistent with ACOs in "Obamacare," U.S. Congress 2010).

We address the question of whether the manufacturer can take advantage of its proprietary information (thus obtaining an information rent, or excess profits due to the information asymmetry) and how the government can design a procurement contract to minimize the advantage. In summary, we show that a manufacturer with private productivity information can command an information rent from a government unable to inspect the manufacturer's production effort. We construct a menu of contracts that minimizes the government's overall cost by balancing the trade-off between providing sufficient incentives to the manufacturer to exert an appropriate level of effort and the information rent that the manufacturer can extract from the government. Finding such a menu is a challenging problem as the set of potential contracts is infinitely large. However, we (1) construct a menu that can be shown to be optimal. Our analysis shows that

this menu has certain peculiarities that make its practical implementation difficult. To overcome this, we (2) determine another menu that is optimal within a more practically viable family of contracts. Interestingly, our numerical experiments show that this menu leads to near-optimal outcomes. We then (3) derive a novel result that shows how the ability to verify a manufacturer's effort in a way that is contractible can eliminate the information rent even when there is asymmetric information about the manufacturer's productivity. This provides an upper bound on the value of the ability to inspect manufacturer effort. We also (4) provide a sensitivity analysis to identify parameters that most influence this upper bound.

The sequence of events and sources of uncertainty are important for the influenza vaccine supply (Gerdil 2003, Saluzzo and Lacroix-Gerdil 2006). Three strains of influenza are selected in the spring for inclusion in vaccines to be shipped in the fall for use in the northern hemisphere. A small quantity of doses are produced for testing. This, together with information from preceding years, provides a probability distribution of the potential yield per dose of vaccine when production is ramped up to much greater volumes. Yield uncertainty may depend on both strain information (independent of the manufacturer) as well as some features of the manufacturer's processes that do not depend on the strain, such as breakage in material handling, spoilage, logistic challenges, or continuous improvement efforts. For example, Sanofi Pasteur, the largest provider of flu vaccines to the United States, delayed shipments of flu vaccines in 2014 "because one of the flu strains used to produce the shot grew more slowly than expected" (WSJ, Loftus 2014). A different manufacturer delayed shipments that year "because batches made at a plant in Quebec didn't meet quality-assurance standards and were discarded" (Loftus 2014), factors linked to a manufacturer's processes (Health Canada 2014).

Prior to observing the yield for a given season's vaccine composition, the government orders a target number of vaccines for delivery. The manufacturer then selects its production effort during the *primary production period*. If the yield is sufficiently high, all vaccines ordered are shipped on time. If the yield is not high enough, a shortage occurs and not all vaccines can be shipped on time. Chick et al. (2008) derived contracts that could coordinate the manufacturer's incentives with socially optimal public health goals under the assumption that production yield distributions and cost information were known by both the manufacturer and government, and that there is no recourse for continuing production beyond a primary production period if the yield were low in a given season.

Communication with senior representatives of a major vaccine manufacturer indicates the possibility to extend an influenza vaccine production campaign for a few extra weeks in an attempt to fill the demand in campaigns when production yield is low. A two-stage production process has also been modeled in the literature before (Dai et al. 2016). Such a *late production period* may result in a higher production cost per vaccine and per unit effort. Late production is more costly because the manufacturer has to reactively obtain additional raw materials (such as the eggs used in the dominant influenza vaccine production processes) beyond its initial commitments to its suppliers, incur handling and expedited shipments costs for delayed vaccines, and bear the goodwill loss. In addition, the late production period may also lead to opportunity costs for the manufacturer as now some of its resources that may have been committed to other vaccines would need to be committed to the processes involving the production and supply of influenza vaccines. For example, extended influenza vaccine production campaigns employ packaging and logistics resources that may also be used for other vaccines, hence increasing the carrying costs associated with higher inventories of other vaccines. Another potential interpretation of the late production period is one in which the manufacturer carries out a series of small production runs until the contractual obligation of the required supply of vaccines is met. Such a system could also exhibit a higher production cost primarily driven by the lack of economies of scale.

Here, we distinguish between yield uncertainty due to strain selection and a manufacturer's production technology and productivity, defined as the mean number of vaccines per unit effort. The latter can shift the distribution of yield higher or lower, depending on the internal improvement activities of the manufacturer. Therefore, a high yield of vaccine realized in one year might be attributable to a strain that is easy to manufacture in high volumes, or to manufacturing productivity, or a combination of both. Without the ability to inspect a manufacturer's processes, it is not clear how to disambiguate these factors.

In this paper, we assume that productivity is known privately to the manufacturer and allows for a late production period to complete the demanded number of vaccines. The primary production season reveals the yield at large-scale production volumes to the manufacturer. The manufacturer therefore experiences minimal yield uncertainty during the late production stage capturing the effects of learning due to experience, and we initially assume that the manufacturer can fulfill all demand in that late production stage. Similar assumptions have been made in the literature before. For example, Dai et al. (2016) assume that the late production period can deterministically satisfy for shortages from the earlier production run. As mentioned above, one may also interpret the late production period as a series of small production runs until the manufacturer is able

to fully meet its supply obligation. Such an operational system would also result in ensuring that the manufacturer is able to deterministically fulfill the shortage from the first production period. While this assumption is an approximation, and there has been a rare case where all production from a manufacturer for a given year has been wiped out (FDA 2005), it is a reasonable approximation: such massive disruptions are infrequent. Numerical tests indicate that the proposed menu of contracts performs relatively well even if there are moderate shortages of vaccine.

Yield uncertainty and information asymmetry about productivity per unit effort subjects the relation between the government and the manufacturer to *adverse selection*. Moreover, because manufacturer's effort may not be verifiable, *moral hazard* is another source of inefficiency in the vaccine supply chain. Although the manufacturer's effort may be estimated by the number of eggs purchased, the actual effort may remain unknown without close monitoring due to unintentional loss of raw materials or efficiency problems with inputs (as opposed to inefficiencies with vaccines that are the output of the production process). While manufacturers focus on improving their processes, monitoring such efforts and quantifying their benefits come at a cost, which includes monitoring systems, certification, or process audits by third parties. We study the inefficiencies due to adverse selection and moral hazard in the vaccine supply chain, and we quantify the amount a government should be willing to spend in order to eliminate such inefficiencies.

Section 2 reviews related literature. Section 3 formalizes the mathematical model, a Stackelberg game of influenza vaccine procurement by a government from a single manufacturer whose payoffs depend on an epidemic model. Some results assume that the manufacturer's productivity is one of two types (high or low). Section 4 determines the manufacturer's best response to the government's order as well as the system optimum under the assumption that the manufacturer's productivity is common information. We use this framework to address the following questions.

• Can the manufacturer extract an information rent (surplus profit) if it possesses private productivity information (adverse selection) and its production effort is unverifiable (moral hazard) in this context with yield uncertainty? How can the government construct a menu of (output based) contracts that allows it to infer the productivity level of the manufacturer and minimize any eventual information rent? (Section 5.1)

• Can the government design a menu that is optimal within a more practically implementable family of contracts and yet leads to near-optimal outcomes? (Sections 5.2 and 6)

• How much information rent can there be with such contracts when they are implemented with

parameters consistent with the influenza vaccine supply chain? What parameters are key drivers of the magnitude of any such information rent? (Section 6)

• Are the contracts robust to the assumption that all ordered vaccines are delivered? (Section 6.3)

• Is it possible to eliminate the information rent if it were possible to monitor and contract on the manufacturer's effort, hence removing the inefficiency due to moral hazard? (Section 7)

The affirmative answer to the last question is counterintuitive because the manufacturer's private productivity information would typically allow for an information rent. The negative answer to the preceding question provides some reassurance that the model may be useful even if not all assumptions are precisely met in practice. Taken together, these results suggest how much money the government might reasonably invest in technology and processes to reliably confirm the production effort of the manufacturer.

Appendices provide mathematical proofs of claims in the main text and related supporting material.

2. Related Literature

Our work is closely related to the operations management literature on influenza vaccine supply chains as well as the contracting and mechanism design literature.

The operations management (OM) literature has looked at three main stages in the vaccine supply chain. The first stage occurs before large-scale production and relates to vaccine development: the selection of which three strains of influenza the annual vaccine should include, and the timing of strain selection (Wu et al. 2005, Kornish and Keeney 2008, Cho 2010, Özaltın et al. 2011). The second stage is that of how the specificity of the influenza production process, most notably the production yield and/or market demand may influence incentives and decisions made in the influenza vaccine market or in production volume decisions (Deo and Corbett 2009, Chick et al. 2008, Arifoglu et al. 2012). The third stage occurs after vaccines are produced: the allocation of vaccines or other resources to intervene in influenza transmission. This includes the role of customer demand (Arifoglu et al. 2012), and the allocation of resources within a given population (Weycker et al. 2005, Khazeni et al. 2009, Wu et al. 2009) such as priority allocation to children in an urban area or across populations or countries represented by different governments (Brandeau et al. 2003, Sun et al. 2009, Wang et al. 2009, Mamani et al. 2013). Dai et al. (2016) consider the role of the healthcare provider and explore novel variations on contracts to study the role of uncertain delivery timing, early production with design risks, and time sensitive demand. Mamani et al.

(2012) explore the role of governments in market coordination through subsidies, even with multiple manufacturers, but does not account for vaccine production yield uncertainty. Adida et al. (2013) present a menu of subsidies that leads to a socially efficient level of coverage even when yield uncertainty and network externalities are accounted for. Much of this work assumes that information is symmetrically known. We contribute to work on asymmetric information but do not account for consumer incentives as do some of the above works.

A major contribution of our paper to the OM literature on influenza vaccination is that we relax the assumption that all supply chain parameters are common knowledge. The supply chain in our model is subject to adverse selection (the manufacturer's productivity is not known to the buyer) and moral hazard (the manufacturer's effort is not verifiable). We show how the buyer can infer the manufacturer's productivity with contracts that do not necessarily rely on observing the manufacturer's effort. We further show that subject to the manufacturer's effort verifiability, a menu of contracts can eliminate all the informational advantage to the manufacturer.

A significant literature in OM handles the inefficiencies due to information asymmetry and moral hazard in a variety of contexts. In a quality management context subject to double moral hazard (but no adverse selection), Baiman et al. (2000) show that the first-best is obtained if either the supplier's or the buyer's effort is contractible. Corbett et al. (2004) examine the value to a supplier of obtaining better information about a buyer's costs but do not address moral hazard in their model. Cachon and Zhang (2006) characterize the optimal procurement contract menu in a buyer-supplier setup where suppliers have private information about their costs. They propose alternative simpler outcome-based contracts that perform nearly optimal. Crama et al. (2008) study a licensing contract design where the licensee's valuation of an innovation and its development effort are proprietary and construct a menu of three-part tariff contracts to align the incentives. Hasija et al. (2008) investigate the contract design to induce a vendor to choose optimal capacity when there is asymmetry about a call center's worker productivity. They show that a menu of contracts can reduce the information rent extracted by the vendor without a significant penalty to the supply chain performance. Yang et al. (2009) study a manufacturer-supplier dyad where the supplier is better informed on the likelihood of supply disruptions. Using mechanism design to construct the optimal contract menu, they show that the information asymmetry may cause the less reliable supplier to opt for a late penalty for the shortfall instead of using backup production. Kim and Netessine (2013) study incentive-compatible contracts in a supplier-manufacturer relation subject to double moral

hazard and adverse selection to facilitate collaboration in developing new products.

Our work also relates to the literature on satisfying a "rigid demand" when the production yield is uncertain, thus requiring (potentially) multiple production runs (see Grosfeld-Nir and Gerchak 2004 for a comprehensive review). We consider a newsvendor setting with stochastically proportional production yield where the buyer, i.e., the government, decides on the fraction of the population to be vaccinated, resulting in a certain nonrandom demand for vaccines, and the newsvendor, i.e., the manufacturer, decides on the production effort. If the yield is low, the newsvendor has to run a more costly late production campaign to produce the remainder. The proportional yield we assume in our paper is commonly used in the OM literature (Li and Zheng 2006, Yano and Lee 1995).

Our model of information asymmetry about productivity essentially unbundles yield uncertainty into external and internal drivers of variability (similar to systemic versus idiosyncratic risk in finance). This unbundling of yield is also relevant to agriculture, among other sectors. For example, the overall yield of a farm is determined by two sets of factors: (1) controllable factors such as crop rotation cycles (Hennessy 2006, Livingston et al. 2015) or new technology adoption in areas such as irrigation or seed strain selection (Huh and Lall 2013); (2) unpredictable factors such as weather conditions (rainfall, humidity, temperature) or pest invasion (Kazaz and Webster 2011). In this case too, yield can be decoupled into controllable productivity factors and residual yield uncertainty, which may vary year on year.

The incentive design and contracting work here contributes to a long line of related supply chain contracting literature (Cachon 2003 and references therein). A substantial part of this literature studies the efficacy of different types of contracts to incentivize a buyer to order optimally to overcome the potential inefficiency due to double marginalization (Tirole 1988, Pasternack 1985). In these settings, the buyer's decision (order quantity) is assumed to be noncontractible. This assumption leads to underordering by self-interested buyers, thereby creating inefficiency in the supply chain. In such a setting, two-part tariffs (buy-back, revenue-share, quantity-discount contracts) are shown to overcome such inefficiency by providing incentives to the buyer to order optimally. However, when such settings also entail moral hazard (noncontractible efforts such as sales effort), these contracts fail to achieve coordination—in both order quantity and effort (Taylor 2002, Cachon and Lariviere 2005). In the particular context of the influenza vaccine supply chain, our paper proposes an incentive-compatible menu of contracts to achieve socially optimum levels of vaccine production.

3. Public Benefit, Production, and Deployment Model

This section formulates our public procurement model of influenza vaccines. The model accounts for the public benefit of the procured good, the manufacturer's economic and production model, and the government's procurement and deployment costs. We assume that there is a single for-profit manufacturer and the government is the single purchaser whose objective is to minimize the total financial burden of disease.

The public costs and benefits are modeled by assuming that if vaccines are procured for a fraction f of a population of N susceptible individuals then a number $T(f)$ of individuals are infected, with a total social cost of $bT(f)$, where b is the average cost per infected individual. Depending on the criteria set by the public decision maker, these costs include direct costs for treating infected individuals and may or may not include indirect costs associated with loss of work. Our results are based on general properties of the epidemic model, and do not look into its specifics, which might be a simple compartmental model (Diekmann and Heesterbeek 2000) or a more elaborate model with complex dynamics, plans that incorporate state-dependent deployment of antiviral therapy, social distancing, and so forth (Wang et al. 2009, Khazeni et al. 2009, Wu et al. 2009). For stochastic models, $T(f)$ is the expected number infected.

Our results do assume that more vaccination gives better health outcomes, that the benefits vary smoothly with the vaccination fraction, and that $T(f)$ is convex for all sufficiently large f as in Assumption 1. Assumption 1 is valid, for example, with the standard SIR with vaccine model of influenza when f is chosen to be the critical vaccination fraction (Mamani et al. 2013).

Assumption 1. *The expected number infected, $T(f)$, is strictly positive and strictly decreasing in the fraction vaccinated, f, is twice continuously differentiable, and $T(1) > 0$. Moreover, there exists an $\underline{f} \in [0, 1)$ such that $T(f)$ is strictly convex at f if and only if $f \in (\underline{f}, 1]$.*

We presume that the fraction f represents a commitment by the government for purchasing fN treatments with vaccines. The function $T(f)$ can represent the expected number infected when fN treatments are purchased but some fraction of treatments are not used due to vaccination compliance or spoilage issues.

The manufacturer commits to delivering fN treatments of vaccine based on the government's order. The production process for influenza vaccines is subject to statistical variation (U.S. GAO 2001, Saluzzo and Lacroix-Gerdil 2006, Loftus 2014). We model this uncertain yield by assuming that if n units of production effort are used in a *primary production period*, a total of $n\theta U$ treatments will be produced. Our setup is in

line with the literature on rigid demand where a supplier has to fulfill the demand in its entirety subject to production yield uncertainty (Grosfeld-Nir and Gerchak 2002).

In our model, n represents the production effort (for the main influenza vaccine production technology today, this might be measured by the number of eggs input to the process, as eggs are a key input for that technology). The variable cost during the primary production period is c per unit of effort, for a cost cn.

The term $U \geq 0$ models a stochastically proportional production yield (Henig and Gerchak 1990) and that may be due to differences in the strains selected for the vaccine, environmental factors, and natural variation from the production technology. The cumulative distribution function (CDF) of U is denoted $G(u) = \Pr\{U \leq u\}$. Its probability density function $g(u)$ is assumed to exist to simplify proofs.

The term θ models productivity factors that may be due to different operational practices and unobservable continuous improvement efforts that may improve the mean output per unit effort (e.g., by removing assignable cause variation). We presume that a manufacturer knows its value of θ but that the government might not. For instance, the government might believe that θ has a low or high value, $\theta \in \{\theta_l, \theta_h\}$ with $\theta_l < \theta_h$. The government is assumed here to believe that $\Pr(\theta = \theta_l) = q$ and $\Pr(\theta = \theta_h) = 1 - q$.

A low yield may result in a shortfall of $(fN - n\theta U)^+$ treatments. If a shortfall is experienced, then an additional $(fN/(\theta U) - n)^+$ units of effort in a *late production period* are required to make up the difference. Note that while yield is unknown when the primary production decision is taken, its value is known rather well by the time the primary production period is completed. Thus, $(fN/(\theta U) - n)^+$ is known at the time of planning for the late production period (this is valid not only for the vaccine manufacturer we communicated with, but also resembles challenges of some agricultural firms with regard to food production yield).

The cost per unit effort, L, to cover a shortfall is assumed to exceed that for effort prior to the target delivery deadline (i.e., $L > c$) for reasons given in the Introduction. Thus the manufacturer's total cost is

$$\mathrm{MF}(n; f) = \mathrm{E}[cn + L(fN/(\theta U) - n)^+ - p_r fN]. \quad (1)$$

The manufacturer produces if only if $\mathrm{MF}(n; f) \leq R$, where $R \leq 0$ is its reservation value.

The government's procurement and vaccine deployment costs are modeled by assuming that it pays $p_r > 0$ per treatment ordered to the manufacturer, and has an expense $p_a > 0$ per treatment that is administered if it is delivered on time. If treatments arrive after the delivery deadline, an additional fractional charge per treatment of $\delta \geq 0$ is incurred for vaccine administration.

The government's total social, vaccine acquisition, and vaccine administration costs are therefore modeled by

$$\mathrm{GF}(f;n) = \mathrm{E}[bT(f) + p_a(fN + \delta(fN - n\theta U)^+) + p_r fN].$$

4. Selfish and Social Optimum Outcomes with Known Productivity Factor

Before we analyze the incentive design problem with asymmetric information about the productivity factor θ, we first analyze the performance of wholesale price contracts relative to system optimality when $\theta = \theta_i$ is known by both parties. This can serve as a benchmark, and will provide useful results for the analysis of the more complex problem with asymmetric information in Section 5.

We do so in a game theoretic model with the government as the Stackelberg leader offering a take-it-or-leave-it wholesale price contract to the manufacturer. A wholesale price contract is characterized here by $\{p_r, f\}$, where p_r is the wholesale price and f is the fraction of the population for which the government commits to buying and deploying vaccines. If the contract is accepted, the manufacturer chooses its effort n for production prior to the target delivery date. It must produce during a late production period if the primary production volume results in a shortfall of output.

We first consider a self-interested manufacturer. For a given fraction f, the *manufacturer problem* is to select $n \geq 0$ to minimize its expected production costs during the primary and late production periods less its revenues, $\mathrm{MF}(n;f)$ in (1). To do so, the manufacturer can use a newsvendor-type framework, which balances the expected marginal production costs for the primary and late production periods as follows.

Proposition 1. *Given a wholesale contract with vaccination fraction f, the manufacturer's optimal production effort in the primary production period, $n_i^*(f)$, solves*

$$\int_0^{fN/(\theta_i n_i^*(f))} dG(u) = c/L. \tag{2}$$

Therefore $n_i^*(f) = fN/(\theta_i k_i^G)$ is a linear function of f, where k_i^G solves $k_i^G = G^{-1}(c/L)$.

We next consider a socially optimal response. The social optimum is defined relative to the minimization of the total costs of all players $\mathrm{SF}(f,n) \triangleq \mathrm{GF}(f;n) + \mathrm{MF}(n;f)$ when actors coordinate to achieve that minimum without any self interest constraint. The social optimum is a reference for comparison for the case when all players act independently or when contracts attempt to align incentives. The *system problem* is

$$\min_{\substack{0 \leq f \leq 1 \\ n \geq 0}} \mathrm{SF}(f,n). \tag{3}$$

To avoid degenerate solutions to (3), we assume that administering the first treatment is cost effective in expectation for the system (otherwise it would not be optimal to vaccinate anybody). We also assume that the last treatment is not beneficial in expectation (which is reasonable given the herd immunity from vaccines and the challenge of full vaccination).

To formalize these assumptions, we introduce some notation. We define k_i^S so that it satisfies

$$\int_0^{k_i^S} p_a \delta \theta_i u + L \, dG(u) = c, \tag{4}$$

which balances the marginal cost of an input and extra administration cost of the resulting doses in the late production period (the left-hand side) and primary production period (right-hand side). Note that k_i^S exists and is unique because the left-hand side of (4) is increasing in k_i^S; $0 \leq \int_0^{k_i^S} p_a \delta \theta_i u + L \, dG(u) \leq \int_0^\infty p_a \delta \theta_i u + L \, dG(u) = p_a \delta \theta_i \mathrm{E}[U] + L$; and $p_a \delta \theta_i \mathrm{E}[U] + L \geq L > c$. We also define

$$\mathfrak{B}_i \triangleq p_a + \int_0^{k_i^S} p_a \delta + L(\theta_i u)^{-1} \, dG(u),$$

which will be seen below to be the expected marginal cost of administering and producing another treatment at optimality (cf. proof of Proposition 2 in Appendix A). The term $\mathfrak{B}_i$ is thus the marginal cost of outputs.

We can now formalize the assumption that it is neither optimal to vaccinate nobody nor everybody, and characterize the optimal system solution for nondegenerate cases, assuming θ_i is common knowledge.

Assumption 2. *For the system problem and for a given θ_i, producing and administering the first treatment is beneficial $(b(dT(f)/df)|_{f=0} + N\mathfrak{B}_i < 0)$, and is not beneficial for the last treatment $(b(dT(f)/df)|_{f=1} + N\mathfrak{B}_i > 0)$.*

Proposition 2. *Given Assumptions 1–2, the optimal vaccination fraction f_i^S and the optimal production effort n_i^S to minimize (3) satisfy $n_i^S = f_i^S N/k_i^S$ (other parameters being equal), where $k_i^S \leq k_i^G$, and the optimal vaccination fraction f_i^S is on the convex part of $T(f)$ (namely, $f_i^S \in [\underline{f}, 1)$) and solves*

$$b \frac{dT(f)}{df}\bigg|_{f=f_i^S} + N\mathfrak{B}_i = 0. \tag{5}$$

Note that if influenza transmission is modeled with the SIR model with vaccination, then Proposition 2 and Mamani et al. (2013, Theorem 1) together imply that f_i^S is not less than the critical vaccination fraction.

When the value of θ is known by both parties, one can show that incentives can be coordinated by a wholesale contract only when there is no surcharge for the late administration of vaccines ($\delta = 0$). Otherwise,

the optimal wholesale price contract cannot coordinate the incentives—it leads to lower vaccine fractions selected by the government and less production effort by the manufacturer than is socially optimal.

Importantly, a contract with a wholesale price, together with a shortage penalty for doses delivered in the late production period, can result in a first-best outcome for the government. These claims are justified in the online Appendix B and are useful for comparisons in the next section. These results complement Chick et al. (2008), who show that output-based contracts cannot, in general, coordinate the supply chain (a necessary condition for attaining the first best). The reason that the contracts proposed in Proposition EC.3 of the online Appendix B attain the first best is that f is stipulated by the contract. In doing so, the government can eliminate any additional potential inefficiency in the system due to double moral hazard. That is, if f was to be determined after the contract is finalized in the system, then two decisions, f and the primary production effort n, need to be coordinated and this may lead to some inefficiency. Because fN is the order quantity, it (or simply f because N is known) can easily be included in the contract. Therefore, the contract proposed in Proposition EC.3 coordinates the supply chain and yields the first-best outcome for the government.

We now turn to the more realistic scenario of unverifiable effort and information asymmetry.

5. Unverifiable Effort and Information Asymmetry

If the manufacturer's effort n is unverifiable (moral hazard) and the manufacturer has private information about its productivity factor θ (adverse selection), can the manufacturer extract an information rent from the government? This section answers this question affirmatively when the government can only contract on the manufacturer's output. Interestingly, if the government can verify the manufacturer's effort, we also show that the rent can be eliminated.

This section also shows how a government can induce a manufacturer to reveal whether it has high or low productivity (denoted θ_h or θ_l) by offering a menu of output-based contracts. Each contract in the menu is designed contingent on a value of θ. The manufacturer chooses the contract that minimizes its costs. The optimal contingent contract minimizes the government's expected cost. According to the revelation principle, there exists a direct and truthful mechanism that is optimal (Laffont and Martimort 2002). Here, we search among mechanisms that satisfy incentive compatibility (IC) and individual rationality (IR) constraints of the manufacturer. If the truth revealing contract leads to social optimum decisions and does not allow the manufacturer to earn any surplus over its reservation value, then such contracts are known as first-best contracts.

We begin by formulating the government's optimal contract design problem. Because effort is unverifiable, a contract can only be written in terms of a verifiable output ($n\theta_i U$), and the procurement order placed by the government (fN or simply f). Using the revelation principal, the optimal contract will be a menu with a contract tailored for each productivity type $i \in \{l, h\}$. Let $j \in \{l, h\}$ index contract types. The transfer payment with contract j from the government to a manufacturer of type i with output $n\theta_i U$ is $\tau_j(n\theta_i U, f_j)$. The functions $\tau_l(), \tau_h()$ determine a menu of contracts. The government's optimal contract design problem can be written as follows:

$$\min_{\substack{\tau_j(\cdot), f_j \\ j \in \{l, h\}}} \mathrm{GF}(f; n)$$

$$= \sum_{j \in \{l, h\}} \mathrm{Pr}(\theta = \theta_j)\big[bT(f_j) + p_a f_j + p_a \delta \mathrm{E}[(f_j N - n_{j,j}\theta_j U)^+]$$
$$+ \mathrm{E}[\tau_j(n_{j,j}\theta_j U, f_j)]\big], \quad (6)$$

$$\mathrm{MF}_{i,j}(n) = \mathrm{E}\big[cn + L((f_j N - n\theta_i U)^+/(\theta_i U))$$
$$- \tau_j(n\theta_i U, f_j)\big], \quad (7)$$

$$n_{i,j} = \arg\min_n \mathrm{MF}_{i,j}(n) \quad (8)$$

$$\mathrm{MF}_{h,h}(n_{h,h}) \leqslant \mathrm{MF}_{h,l}(n_{h,l}), \quad \text{(IC high-type)} \quad (9)$$

$$\mathrm{MF}_{l,l}(n_{l,l}) \leqslant \mathrm{MF}_{l,h}(n_{l,h}), \quad \text{(IC low-type)} \quad (10)$$

$$\mathrm{MF}_{i,i}(n_{i,i}) \leqslant R, \quad \text{(IR type } i \in \{l, h\} \text{ manufacturer).} \quad (11)$$

The manufacturer's cost structure is defined in (7) and its optimal behavior is determined by (8). The IC constraint induces a type i manufacturer to choose contract i over contract j, or to reveal its type truthfully, so that $\theta_i = \arg\min_{\theta_j} \mathrm{MF}_{i,j}(n_{i,j})$. The IR constraint guarantees that a type i manufacturer achieves its reservation price.

It will be convenient below to define $A_{i,j}$ as the expected shortfall of treatments and $B_{i,j}$ as the expected units of effort in the late production period when a type i manufacturer chooses the type j contract. The values of $A_{i,j}$ and $B_{i,j}$ are determined by the optimal value of $n_{i,j}$ in (8) and thus implicitly depend on $\tau_j()$:

$$A_{i,j} = \int_0^{f_j N/(\theta_i n_{i,j})} (f_j N - n_{i,j}\theta_i u)\, dG(u),$$
$$B_{i,j} = \int_0^{f_j N/(\theta_i n_{i,j})} (f_j N/(\theta_i u) - n_{i,j})\, dG(u). \quad (12)$$

5.1. A Linear Contract with Four Parameters Is Optimal

The government's contract design problem is complicated as it involves both moral hazard and adverse selection, and as shown in (6), requires the government to solve a constrained optimization problem with functions $\tau_l()$ and $\tau_h()$ as decision variables.

We now construct a menu of linear contracts and prove that the optimal linear contract is optimal for (6). Each contract $j \in \{l, h\}$ in the menu can be represented with four values: a vaccination fraction f_j, a wholesale unit price $p_{r,j}$, an additional unit price for each treatment produced during the primary production period $p_{d,j}$, and a penalty per late treatment $p_{s,j}$. Such wholesale-price-with-shortage-penalty contracts incentivize earlier production and modify the payoffs to the manufacturer. The transfer payment function to a type i manufacturer that selects contract j is therefore

$$\tau_j(n\theta_i U, f_j) = -p_{r,j}f_j N - p_{d,j}n\theta_i U + p_{s,j}(f_j N - n\theta_i U)^+, \quad \text{for } j \in \{l, h\}. \quad (13)$$

This transfer payment determines $\mathrm{MF}_{i,j}(n)$ in (7) and creates a special case of the objective function in (6). Specifically, the optimal menu of contracts in the class of such linear contracts is the solution to the following optimization problem:

$$\min_{\substack{f_j, p_{r,j}, p_{d,j}, \\ p_{s,j}; j \in \{l,h\}}} \mathrm{GF}(f; n)$$
$$= \sum_{j \in \{l,h\}} \Pr(\theta = \theta_j)\big[bT(f_j) + p_a f_j N + p_{r,j}f_j N$$
$$+ p_{d,j}n_{j,j}\theta_j E[U] + (p_a\delta - p_{s,j})A_{j,j}\big], \quad (14)$$

subject to the constraints in (7)–(11).

Proposition 3 shows that the optimization problem with arbitrary functions $\tau_j()$ as decision variables in (6) can be simplified to a problem with real-valued decision variables which determine $\tau_j()$ via (13). This theoretical result is of practical importance: restricting attention to linear payment/penalty terms does not hinder the contract designer's ability to sufficiently overcome agency issues.

Proposition 3. *The optimal linear contract for* (14) *is optimal for* (6).

The proof of Proposition 3 shows that the optimal contracts cannot attain the first-best outcome for the government and leave some information rent for the high-type manufacturer. Proposition 4 shows that the optimal menu leads to sufficient incentives for the high-type manufacturer to exert the system optimal effort. This result is referred to as a "no distortion for the efficient type" result.

Proposition 4. *An optimal four parameter linear contract that solves* (14) *satisfies,* $f_h = f_h^S$, $p_{s,h} = \delta p_a$, $p_{d,h} = 0$, $p_{r,h}$ *solves*

$$-p_{r,h}f_h N + \int_0^{f_h N/(\theta_h n_{h,h})} \big(\delta p_a f_h N + Lf_h N(\theta_h u)^{-1}\big) dG(u)$$
$$= -p_{r,l}f_l N + \int_0^{f_l N/(\theta_h n_{h,l})} \big(p_{s,l}f_l N + Lf_l N(\theta_h u)^{-1}\big) dG(u);$$

and f_l solves

$$q\left(b\frac{dT(f_l)}{df_l} + p_a N + c\frac{\partial n_{l,l}}{\partial f_l} + L\frac{dB_{l,l}}{df_l} + \delta p_a\frac{dA_{l,l}}{df_l}\right)$$
$$- (1-q)\left(c\left(\frac{\partial n_{h,l}}{\partial f_l} - \frac{\partial n_{l,l}}{\partial f_l}\right) + L\left(\frac{dB_{h,l}}{df_l} - \frac{dB_{l,l}}{df_l}\right)\right.$$
$$+ p_{s,l}\left(\frac{dA_{h,l}}{df_l} - \frac{dA_{l,l}}{df_l}\right) + p_{d,l}E[U]\left(\theta_l\frac{\partial n_{l,l}}{\partial f_l} - \theta_h\frac{\partial n_{h,l}}{\partial f_l}\right)\right)$$
$$= 0,$$

$p_{s,l}$ solves

$$q\left((p_a\delta - p_{s,l})\frac{dA_{l,l}}{dp_{s,l}} + p_{d,l}\theta_l E[U]\frac{\partial n_{l,l}}{\partial p_{s,l}}\right)$$
$$+ (1-q)(A_{l,l} - A_{h,l}) = 0,$$

$p_{d,l}$ solves

$$q\left(p_{d,l}\theta_l E[U] + (p_a\delta - p_{s,l})\frac{\partial A_{l,l}}{\partial p_{d,l}}\right)\frac{\partial n_{l,l}}{\partial p_{d,l}}$$
$$+ (1-q)(n_{h,l}\theta_h - \theta_l n_{l,l})E[U] = 0,$$

and $p_{r,l} = \int_0^{f_l N/(\theta_l n_{l,l})}(p_{s,l} + L(\theta_l u)^{-1})\,dG(u) - R/(f_l N)$.
Importantly, the optimal menu implicitly results in $n_{h,h} = n_h^S$, and $p_{d,l} < 0$.

The optimal linear contract has a curious property: the low-type manufacturer is offered a negative payment ($p_{d,l} < 0$) for each treatment produced during the primary manufacturing period: the agency setting has both moral hazard and adverse selection, so the government balances a trade-off between information rent (that the high-type manufacturer can extract) and effort distortion (which may happen for the low-type manufacturer). On the one hand, pushing the low-type manufacturer's effort closer to the system optimal effort level improves the outcome for the government in terms of the number of treatments delivered during the primary production period. On the other hand, it increases the incentive for a high-type of manufacturer to pretend to be of low type in order to extract additional surplus. The government can incentivize the low-type manufacturer to exert an appropriate level of effort by carefully choosing a penalty for shortage ($p_{s,l}$). Although a negative value of $p_{d,l}$ attenuates the incentive provided via $p_{s,l}$, it plays the more important role of ensuring that the high-type manufacturer (who is expected to have a higher output from the primary production period) is dissuaded from accepting the contract that is tailored for the low-type manufacturer.

5.2. A Simpler Menu of Contracts with Three Parameters

Proposition 4 identifies a practical challenge with implementing the optimal linear four parameter contract menu: a negative payment for output ($p_{d,l} < 0$)

may be difficult to implement even though the individual rationality constraint of the manufacturer is satisfied. For example, in cases with a limited liability, it may be impossible to enforce such a contract that does not have a limit on the potential payment that the manufacturer may end up owing to the government if the support of U is unbounded. Contract design in the context of limited liability often prohibits negative transfers (e.g., Gromb and Martimort 2007).

To eliminate such issues, we next focus on a class of contracts that does not include $p_{d,j}$ terms: contracts are represented with three values: f_j, $p_{r,j}$, and $p_{s,j}$. This class contains the linear contracts of Section 5.1 with the additional constraint $p_{d,h} = p_{d,l} = 0$. Proposition 5 characterizes the structure of the optimal menu of contracts in this class.

Proposition 5. *In the class of menus of linear contracts subject to the additional constraint $p_{d,h} = p_{d,l} = 0$, the optimal menu of contracts $\{\tilde{f}_j, \tilde{p}_{r,j}, \tilde{p}_{s,j}\}$ satisfies $\tilde{f}_h = f_h^S$, $\tilde{p}_{r,h}$ solves*

$$-p_{r,h}\tilde{f}_h N + \int_0^{\tilde{f}_h N/(\theta_h n_{h,h})} \left(\delta p_a \tilde{f}_h N + L\tilde{f}_h N(\theta_h u)^{-1}\right) dG(u)$$

$$= -\tilde{p}_{r,l}\tilde{f}_l N + \int_0^{\tilde{f}_l N/(\theta_h n_{h,l})} \left(\tilde{p}_{s,l}\tilde{f}_l N + L\tilde{f}_l N(\theta_h u)^{-1}\right) dG(u),$$

and $\tilde{p}_{s,h} = p_a \delta$; and $\tilde{f}_l$ solves

$$q\left(b\frac{dT(f_l)}{df_l} + p_a N + L\frac{\partial B_{l,l}}{\partial f_l} + \tilde{p}_{s,l}\frac{\partial A_{l,l}}{\partial f_l} + (\delta p_a - \tilde{p}_{s,l})\frac{dA_{l,l}}{df_l}\right)$$
$$-(1-q)\left(L\left(\frac{\partial B_{h,l}}{\partial f_l} - \frac{\partial B_{l,l}}{\partial f_l}\right) + \tilde{p}_{s,l}\left(\frac{\partial A_{h,l}}{\partial f_l} - \frac{\partial A_{l,l}}{\partial f_l}\right)\right) = 0,$$

$\tilde{p}_{r,l} = \int_0^{\tilde{f}_l N/(\theta_l n_{l,l})} (\tilde{p}_{s,l} + L(\theta_l u)^{-1}) dG(u) - R/(\tilde{f}_l N)$, *and $\tilde{p}_{s,l}$ solves*

$$q(\delta p_a - p_{s,l})\frac{dA_{l,l}}{dp_{s,l}} + (1-q)(A_{l,l} - A_{h,l}) = 0. \quad (15)$$

We will use the term "four parameter" when referring to linear contract menus in Proposition 5.1 and the term "three parameter" when referring to linear contracts in Section 5.2 with the constraint $p_{d,h} = 0, p_{d,l} = 0$. It is interesting to note that the optimal four parameter and optimal three parameter linear contract menus propose the same contract for a high-type manufacturer. Only the contract tailored for the low-type manufacturer differs.

We now further discuss the optimal three parameter linear contract menu in Proposition 5. The penalty per late treatment in the high-type contract is equal to the extra cost incurred by the government to administer the treatments. Also, a high-type manufacturer enjoys an "information rent" as its optimal cost is strictly below the reservation value R, while the low-type manufacturer's cost is driven to the reservation

value (the maximum cost at which the manufacturer is willing to participate). In other words, a high-type manufacturer enjoys a profit beyond its reservation value when information is asymmetric (we show in the online Appendix B that this excess profit can be eliminated if there is symmetric information).

Corollary 1 shows that a manufacturer with a higher production yield chooses a contract with a higher vaccination fraction. In fact, part (b) indicates that the optimal vaccination fraction and production effort under the high-type contract equal those for the system optimal solution in Corollary 1 when the manufacturer is of high-type, implying that there is no distortion for the efficient type.

Corollary 1. *In the optimal menu of contracts in Proposition 5, (a) the penalty per late treatment is higher with the high-type contract than the low-type contract, i.e., $\tilde{p}_{s,l} < \tilde{p}_{s,h} = p_a\delta$; (b) the optimal vaccination fraction and production effort with the high-type contract equals their respective values for the system-optimum; (c) in the optimal menu of contracts, $\tilde{f}_l < \tilde{f}_h$.*

6. Numerical Evaluation of the Optimal Three Parameter Contracts

Given the practical limitations of the optimal four parameter contracts, we propose the use of the optimal three-parameter contracts when there is both unverifiable effort and information asymmetry. Therefore, it is important that we explore the performance of such contracts further. This section numerically analyzes the performance of the menu of contracts in Proposition 5. Such contracts allow the government to infer private information about a manufacturer's productivity, but cannot push a high-type manufacturer's cost to its reservation value. The gap $\Upsilon_h = R - \mathrm{MF}_{h,h}(n_{h,h})$ is the information rent to a high-type manufacturer.

We address the following questions: What is the loss in efficiency due to agency issues (adverse selection and moral hazard) when the government uses the optimal contract menu of Proposition 5? How big can the information rent be for a high-type manufacturer? What parameters most greatly influence that information rent? And, does the menu of contracts perform poorly if we relax the assumption that all doses must be delivered? We also implicitly provide an answer to how much a government should be willing to pay in order to verify a manufacturer's effort. This is because Section 7 shows the maximum it should be willing to pay is the expected information rent, $(1-q)\Upsilon_h$.

The evaluation is done using parameter values that are illustrative for the influenza vaccine supply chain. Unless otherwise specified below, graphs and data assume $c = \$6$, $L = \$8$, $\delta = 0.4$, $p_a = \$40$, $q = 0.35$, $\theta_h = 1.2$ (with θ_l varied from $0.7\theta_h$ to θ_h) and yield U having gamma distribution with mean 1 and $\sigma^2 = 0.1$.

Table 1. Values for model's parameters in numerical experiments.

Parameter	Values
Mean yield per unit effort, $E[U]$	1
Variance in yield per unit effort, $\sigma^2 = \text{Var}[U]$	$\{0.025, 0.1, 0.4\}$
Cost per unit effort in primary production period, c	6
Cost per unit effort in late production period, L	$\{\$8, \$9, \$12, \$15\}$
Basic reproduction number for influenza, R_0	1.68
Average cost per infected individual, b	\$95
Administration cost per treatment from primary production period, p_a	$\{\$40, \$80\}$
Percentage extra paid to administer late treatments, δ	$\{0.20, 0.30, 0.40, 0.50, 0.8\}$
U.S. population, N	3×10^8

Parameters were also varied in experiments according to the values in Table 1. We assume that the manufacturer's reservation value is $R = 0$. The epidemic model $T(f)$ uses the standard deterministic SIRS model of influenza with basic reproduction number $R_0 = 1.68$ as in Chick et al. (2008).

In summary, the numerical experiments in Sections 6.1 and 6.2 suggest that information rent, both in total and per dose, can be significant, and identify parameters that influence that information rent.

1. The loss in overall efficiency due to agency issues when the government uses the menu proposed in Proposition 5 is between 0 and 2.8% (details are presented in Table 2). The loss of efficiency is defined as the percentage loss from the first-best system cost. This percentage loss is an upper bound for the loss for the government for using the menu proposed in Proposition 5 rather than the global optimal menu.

2. Information rent can constitute from 0% to 4% of the total government expenditure if the manufacturer is of high type. As the degree of uncertainty about the manufacturer's productivity increases (as θ_l/θ_h decreases) there is an increase in vaccine price for both high-type and low-type contracts, information rent, and total government expenditure; there is a decrease in the fraction of the population that is vaccinated if the low-type contract is selected. The information rent to a high-type manufacturer increases in q, σ^2, L, and δ; it decreases in p_a and θ_l; the total information rent is most sensitive to the range of uncertainty about the productivity of the manufacturer (as expressed by θ_l/θ_h) and to the yield uncertainty for any given level of productivity (σ^2). There is moderate sensitivity to the administration cost per treatment, p_a,

and rather less sensitivity to the penalties associated with low yield/late delivery ($L - c$ and δ).

3. The information rent per treatment is also very sensitive to θ_l/θ_h and σ^2, but is less sensitive to the administration cost per treatment (raising p_a lowers both information rent and the number of treatments); information rent constitutes approximately 25% of the vaccine price for a high-type manufacturer when $\theta_l/\theta_h = 0.7$ and other parameters are set at their base values.

4. These conclusions are not sensitive to the assumption that all vaccines will eventually be delivered, and that the menu of contracts performs well, over reasonable ranges for the parameters.

In addition, the online Appendix C suggests a suboptimal but simpler menu of contracts in Proposition 5 that might be useful when the productivity of low- and high-type manufacturers are very similar. The online Appendix D provides a sensitivity analysis with caveats for designing the menu of contracts when some of the parameters are not known with certainty.

6.1. How Much Information Rent Can Be Gained from Private Information?

The information rent available to a high-type manufacturer in the optimal menu of contracts in Proposition 5 clearly depends on the discrepancy between a low- and high-type manufacturer's productivity (as described by θ_l/θ_h) and the probability that a manufacturer is of a given type.

Table 2 shows that a larger difference between the high and low types (a smaller θ_l/θ_h) results in a smaller

Table 2. Outcomes with optimal menu of three-parameter contracts as θ_l is varied. In this example, $p_a = \$40$, $\tilde{f}_h = 0.567$, and $\tilde{p}_{s,h} = \$16$.

θ_l/θ_h	$\tilde{f}_l$	$\tilde{p}_{s,l}$	$\tilde{p}_{r,l}$	$\tilde{p}_{r,h}$	Infected, low-type (%)	min GF	Vaccination spend	Info. rent	Loss from first-best (%)
0.70	0.539	9.74	10.16	10.43	10.43	1.114×10^{10}	8.44×10^9	4.54×10^8	2.77
0.80	0.550	11.40	9.14	9.37	9.82	1.095×10^{10}	8.31×10^9	2.74×10^8	1.67
0.90	0.559	13.42	8.36	8.50	9.34	1.080×10^{10}	8.21×10^9	1.26×10^8	0.84
1	0.567	16	7.76	7.76	8.93	1.067×10^{10}	8.12×10^9	0	0

Figure 1. Expected information rent $(1-q)\Upsilon_h$ as the probability q that the manufacturer is of low type is varied.
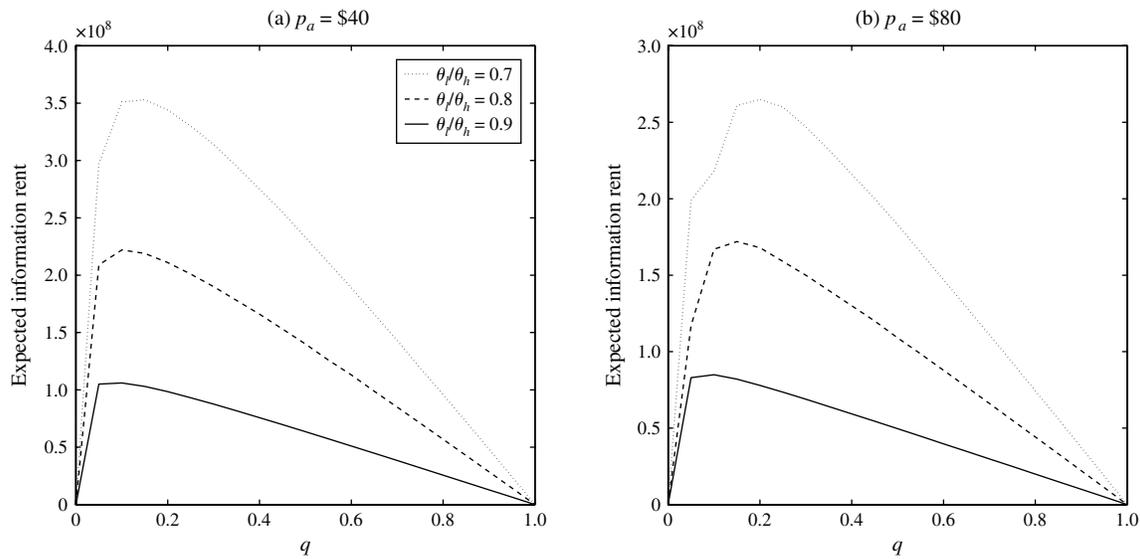
fraction ($\tilde{f}_l$) of the population vaccinated and higher fraction of the population infected if the manufacturer is low type, a larger governmental expenditure in total (min GF) and for vaccines procurement and administration (vaccination spend) in expectation, and a larger information rent to a high-type manufacturer. In particular, when $\theta_l/\theta_h = 0.7$ in this example, the information rent to a high-type manufacturer is $4.54 \times 10^8/1.114 \times 10^{10} \approx 4.1\%$ of the expected total governmental expenditure. Table 2 also shows the percentage increase in costs under the proposed menu compared to the system optimal (first-best) costs. Note that this is the loss of efficiency of the optimal three parameter contracts compared to the first-best outcome; therefore, the percentage loss due to the use of such contracts instead of the optimal four parameter contracts will be strictly lower than the values reported in the last column of Table 2.

A smaller θ_l/θ_h is associated with a smaller penalty for late delivery to the low-type manufacturer ($\tilde{p}_{s,l}$), and higher prices per treatment for low- and high-type manufacturers ($\tilde{p}_{r,l}$ and $\tilde{p}_{r,h}$).

If the government's cost of administration p_a were twice as large ($80), a lower fraction of individuals would be vaccinated (0.412 to 0.440 instead of 0.539 to 0.567 for the low-type contract and 0.440 instead of 0.567 for the high-type contract), the total governmental expenditure would increase ($1.680 - 1.719 \times 10^{10}$ instead of $1.067 - 1.114 \times 10^{10}$), and the information rent would decrease in both absolute terms and relative terms (when $\theta_l/\theta_h = 0.7$ the information rent is 3.64×10^8, or 2.1% of total governmental expenditure versus 4.54×10^8 or 4.1%). A larger p_a is also associated with higher penalties for treatments delivered late by a low-type manufacturer, and higher revenues per treatment for low- and high-type manufacturers. Finally,

the percentage loss from the first-best costs in this case is between 0 and 1.5%.

The information rent also depends on the probability $\Pr(\theta = \theta_l) = q$ that the manufacturer is of low type. In all numerical tests we ran, the information rent Υ_h to the high-type manufacturer is strictly increasing in q. For example, if $\theta_l/\theta_h = 0.7$, there is a 19.2% decrease in the information rent as q decreases from 0.9 to 0.1. Doubling the per-treatment administration cost to $p_a = \$80$ almost doubles this percentage to 38.9%. In other words, there is a notable decline in information rent if there is an increase in probability that the manufacturer is high type. This decline is stronger when the administration cost is higher.

Figure 1 indicates that the *expected* information rent $(1-q)\Upsilon_h$ increases as the discrepancy between the two types increases (as θ_l/θ_h decreases), and is inverse-U shaped in q, which is not surprising as when $q = 0$ or $q = 1$ the manufacturer's type in known and hence the information rent is 0. The expected information rent peaks at an internal value for q and decreases as q approaches the boundaries in the interval $[0, 1]$.

6.2. What Parameters Influence the Information Rent the Most?

We first test the sensitivity of the information rent to a high-type manufacturer, Υ_h, to the price of administering treatments, p_a; the extra cost of effort for late production, $L - c$; the yield variability, $\sigma^2 = \mathrm{Var}[U]$; and the ratio by which administration costs increase if treatments arrive late, δ. We test the sensitivity by doubling each of these parameters for multiple values of the ratio of the mean yields of low-type and high-type manufacturing, θ_l/θ_h.

We then examine the sensitivity of information rent per treatment, $\Upsilon_h/(N\tilde{f}_h)$. The information rent per

Figure 2. Information rent with high-type contract for different σ^2 and p_a.
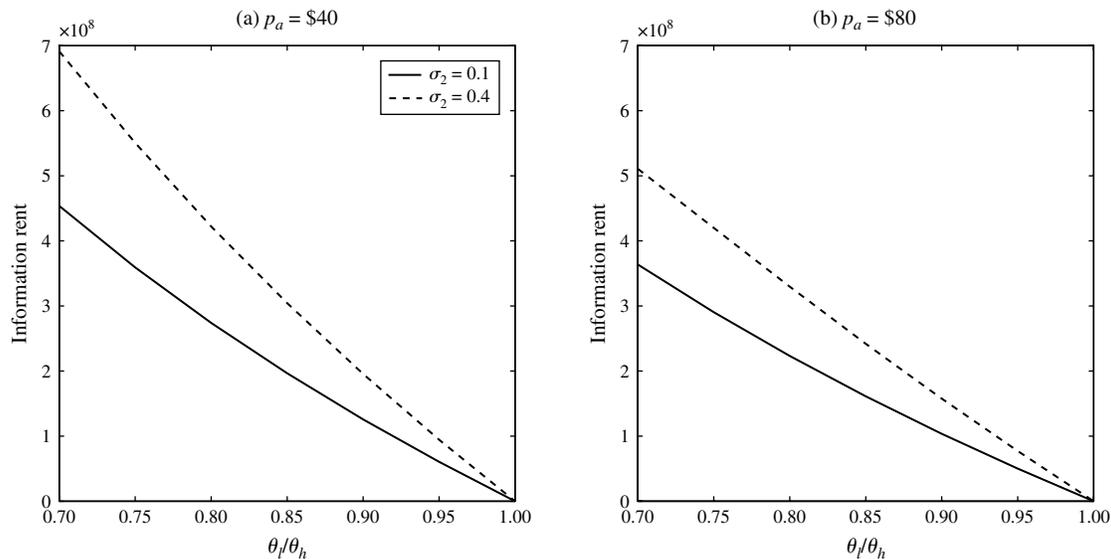
treatment is interesting if vaccines are subsidized by the government through payment by vaccinated individuals (pure copayment). The information rent per person, Υ_h/N is proportional to Υ_h for a given population, and is interesting if vaccines are funded by taxation (pure socialized medicine). Thus Υ_h/N and $\Upsilon_h/(N\tilde{f}_h)$ represent two extremes for funding influenza vaccination.

Figure 2 shows the information rent Υ_h that a high-type manufacturer enjoys in the optimal menu of contracts for two values of p_a and two values of σ^2. The information rent decreases in p_a and increases in σ^2. When $\theta_l/\theta_h = 0.7$, doubling the standard deviation in the yield factor U (increasing σ^2 from 0.1 to 0.4) results in a 52% increase in the information rent when $p_a =$ \$40 (left panel) and a 40% increase when $p_a =$ \$80 (right panel). When $\sigma^2 = 0.1$, doubling the administration cost from $p_a =$ \$40 to \$80 decreases the information rent by 19.9% (the equivalent decrease at $\sigma = 0.025$ is 20.4%). The information rent decreases as the discrepancy between the production yield of high-type and low-type types vanishes and, as expected, disappears completely if the manufacturer is of a known type ($\theta_l/\theta_h = 1$).

When the extra cost per late effort is doubled from $L - c =$ \$2 to \$4, the information rent is increased by 4.5% when $\theta_l/\theta_h = 0.7$ and other parameters are as in the base case. As θ_l/θ_h increases to 0.95, the net increase in information rent from doubling $L - c$ decreases to 3.7%. When δ is doubled from 0.4 to 0.8, the information rent is increased by 4.4% when θ_l/θ_h and other parameters are as in the base case. As θ_l/θ_h increases to 0.95, the net increase in information rent from doubling $L - c$ increases to 7.75%.

In summary, in terms of the percentage change in information rent, doubling σ^2 has a bigger effect

than doubling p_a, which has a bigger effect than doubling δ, which has a bigger effect than doubling $L - c$. Greater uncertainty about the manufacturer's productivity (smaller θ_l/θ_h) is also a key driver. Yield uncertainty and the government's cost of vaccine administration have a much bigger influence on the information rent than do the penalties that the government and manufacturer might incur due to late delivery of vaccines, at least when the parameters are initially set according to our base case ($\sigma^2 = 0.1$, $c =$ \$6, $L =$ \$8, $\delta = 0.4$, $p_a =$ \$40, $q = 0.35$, $\theta_h = 1.2$, $\theta_l \in [0.7\theta_h, \theta_h)$). The same general observation was observed for other tests we ran (e.g., with $\sigma^2 = 0.025$ or $p_a = 80$).

Because the information rent and the information rent per person are proportional, comments about which parameters are most important for the information rent also apply to the information rent per person, Υ_h/N.

Figure 3 depicts the vaccine price $\tilde{p}_{r,h}$ and information rent per treatment $\Upsilon_h/(N\tilde{f}_h)$ for a high-type manufacturer. As θ_l/θ_h increases both of those quantities decrease. If there is no uncertainty about the manufacturer's type ($\theta_l/\theta_h = 1$) the vaccine price is lowest and does not include any information rent. If $\theta_l/\theta_h < 1$, the percentage of information rent in vaccine price is relatively independent of the administration cost p_a. In particular, at $\theta_l/\theta_h = 0.7$, 25.6% of the vaccine price is information rent at $p_a =$ \$40, and 24.6% of the vaccine price is information rent if $p_a =$ \$80.

The information rent per treatment $\Upsilon_h/(N\tilde{f}_h)$ might vary differently than the information rent in total or per capita as parameters are varied because those parameters may influence the number of individuals that are vaccinated in the optimal contract. For example, as L increases, it becomes more costly to produce in the late production period although a high-type manufacturer

Figure 3. Information rent per treatment and vaccine price with high-type contract.

(a) $p_a = \$40$

(b) $p_a = \$80$

— Information rent per treatment, \$
- - - Vaccine price, \$
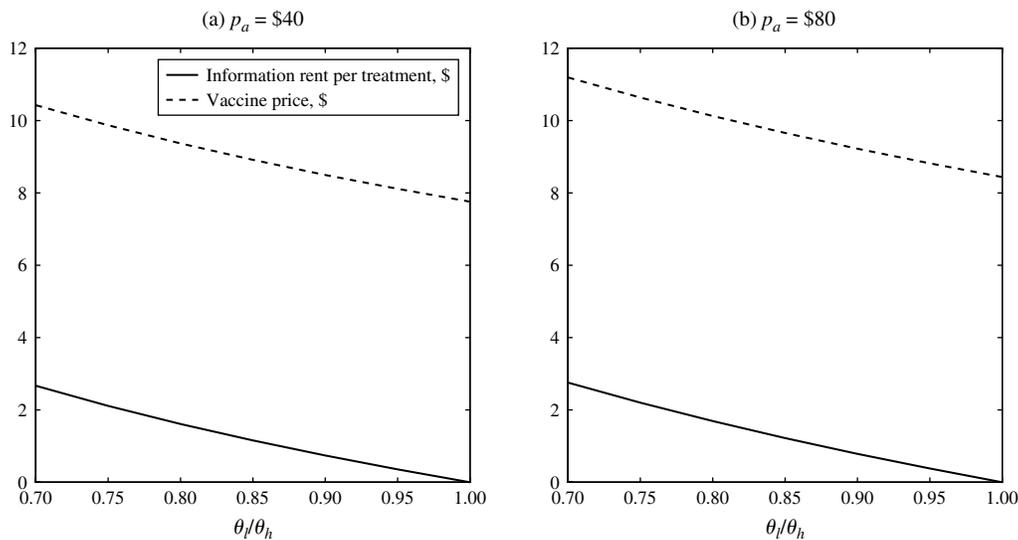
still pays the same penalty for late treatments ($\tilde{p}_{s,h} = p_a \delta$ is independent of L). As a result, the government orders fewer treatments but pays a higher price per treatment. The overall impact is an increase in the manufacturer's profit and hence in the information rent. See Figure 4. If $\theta_l / \theta_h = 0.7$ and $p_a = \$40$, then doubling $L - c$ from 2 to 4 by raising L increases the information rent per treatment increases 6.0%. If $p_a = \$80$, then doubling $L - c$ by raising L increases the information rent per treatment by 4.4%. Thus, the information rent per treatment is increased by a slightly higher fraction than the total information rent per person when L is increased so that $L - c$ doubles.

When either σ or δ is doubled, the information rent per treatment is also increased by a slightly higher fraction than the total information rent per person (data not shown). When the administration cost is doubled, however, the information rent per treatment increases by 3 to 7% (depending on the value of θ_l / θ_h) rather than being decreased by 19.9 to 17.1% as in the case of information rent per person. The reason is that the fraction of individuals that are vaccinated is highly sensitive to p_a.

In summary, uncertainty about productivity (θ_l / θ_h) and yield variability (σ) have the most effect on information rent per treatment $\Upsilon_h / (N \tilde{f}_h)$ (as for the case of information rent per person), and the effect of administration cost on information rent per treatment is somewhat less than that for δ and $L - c$.

6.3. Is the Contract Robust to the "Full Delivery" Assumption?

In the preceding sections, it was assumed that the manufacturer can deliver all vaccines that were ordered. While this can often be accomplished in the primary and/or late production periods, it is possible for a particularly low yield to occur, so that not all vaccines

can be delivered in time for the influenza season. In this section, we analyze the potential that not all vaccines can be delivered, even during the late production period, by modeling a capacity limit for the late production period, and then assessing whether the menu of contracts are still useful even if there is a common belief that not all production might be delivered.

In particular, this subsection assumes that the maximum that can be produced during the late production period is a fraction α greater than what is produced during the primary production period (e.g., production can be extended by two weeks if needed). With this partial delivery assumption, the manufacturer delivers a total of $d(\alpha) = \min(fN, n\theta U) + \min((fN - n\theta U)^+, \alpha n \theta U)$ doses. This transforms the cost function of the manufacturer to

$$\overline{MF}(n; f, \alpha) = E\big[cn + L \min(fN/(\theta U) - n^+, \alpha n) + p_s(fN - n\theta U)^+ - p_r d(\alpha)\big], \quad (16)$$

and that of the government to

$$\overline{GF}(n; f, \alpha) = E\big[bT(d(\alpha)/N) + p_a d(\alpha) + \delta p_a \min((fN - n\theta U)^+, \alpha n \theta U) + p_r d(\alpha) - p_s(fN - n\theta U)^+\big]. \quad (17)$$

We explored whether the use of the menu of contracts in Proposition 5, which assumes full delivery of vaccines, results in near-optimal or far-from-optimal outcomes when the government and manufacturer account for the potential of incomplete delivery by optimizing n and f with (16) and (17). This tests whether or not the use of Proposition 5 is sensitive to the assumption of full delivery of vaccines by the end of the late production period.

The degree of suboptimality from using the menu in Proposition 5 is very small as compared to the case

Figure 4. Information rent per treatment with high-type contract for different L.

(a) $p_a = \$40$

(b) $p_a = \$80$
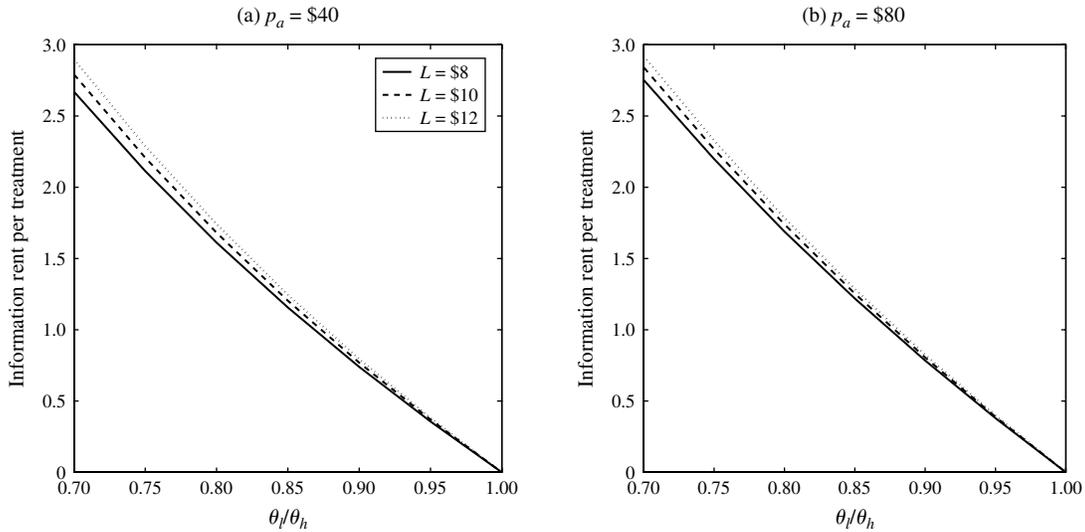
when both parties account for the potential of partial delivery of vaccines in numerical tests. The percentage of suboptimality ranged from 0.20% to 0.63% when $p_a = \$40$ and was even smaller (from 0.13% to 0.41%) when $p_a = \$80$, over the parameter ranges in the preceding sections. These results assumed that $\alpha = 0.10$. Even for an extreme case of $\alpha = 0$ (i.e., no extended manufacturing period), we find that the percentage loss by using the menu proposed in Proposition 5 ranged from 0.44% to 1.01%.

These computations were developed as follows. For a given type and the corresponding contract from Proposition 5, we compute the manufacturer's optimal production level $\bar{n}_i$ ($i \in \{l, h\}$) from (16). This determines the manufacturer's cost $\overline{MF}_i$ and government's cost $\overline{GF}_i$ from (16) and (17) ($i \in \{l, h\}$) given $\bar{n}_i$ and the contract terms, and thereby determines the expected total cost $TC = q(\overline{GF}_l + \overline{MF}_l) + (1 - q)(\overline{GF}_h + \overline{MF}_h)$. We then numerically solve for the optimal social cost $SF_i = \min_{n, f} MF(n; f, \alpha) + GF(n; f, \alpha)$ for each type and the optimal expected social cost, $SF = qSF_l + (1 - q)SF_h$. This procedure determines and expected loss, $LS = TC - SF$. The quantity LS is an upper bound on the loss incurred by using the menu of contracts in Proposition 5 under partial delivery.

Figure 5 depicts LS for $p_a \in \{\$40, \$80\}$ and $\alpha = 0.1$. In each case, the fraction of suboptimality, LS/SF, was less than 1%. The level of suboptimality was highest when the difference in low- and high-type was highest (low values of θ_l/θ_h), where the information rent is also highest for a high-type manufacturer. Thus, the contracts appear to be relatively robust to the assumption of full delivery of vaccines, in these experiments with parameters that are representative for influenza vaccines. Even if we were to assume that α is a random variable with some distribution over the support $[0, 1]$, it follows from our results above that the expected loss

due to using the contract menu proposed in Proposition 5 would be under 1%. We conclude that the contracts in this paper are indeed robust to the "full delivery" assumption in this paper.

## 7. How Much Should One Be Willing to Pay to Verify Manufacturer Effort?

Sections 5 and 6 showed that the expected information rent, $(1 - q)\Upsilon_h$, is positive when the manufacturer's effort is unverifiable and its productivity is privately known. This is true even if a menu of contracts can elicit whether the manufacturer has a high or low productivity. This section assumes that the manufacturer's production effort is verifiable, and hence is contractible. That is, it explores adverse selection with no moral hazard. Here, we show that the verifiability of effort enables the government to eliminate the information rent even with adverse selection. Thus, $(1 - q)\Upsilon_h$ is an upper bound on how much a government should be willing to pay to verify a manufacturer's production effort.

We again assume that the yield per unit effort is θU where $\theta \in \{\theta_l, \theta_h\}$ with $\theta_l < \theta_h$, that θ is initially known to the manufacturer but not to the government, and that the setup is as in Section 4. We also will use the definitions in Section 2 of the effort n_i^S, vaccination fraction f_i^S, average cost to administer one treatment $\mathfrak{B}_i$, and critical limit of integration k_i^S for $i \in \{l, h\}$. From those definitions, the ordering of the θ_i implies that $k_l^S > k_h^S$ and that $\mathfrak{B}_l > \mathfrak{B}_h$. We also assume that the first treatment is cost effective and the last dose is not, i.e., Assumption 2 holds for both θ_l and θ_h.

Rather than trying to develop a contract that deals with the expected value of the costs given uncertainty about θ, we concentrate on contracts contingent on θ, as in Section 5. Unlike Section 5, one or more of the contracts can be based on effort.

Figure 5. Loss incurred by using the optimal menu of contracts in Proposition 5 under partial delivery.
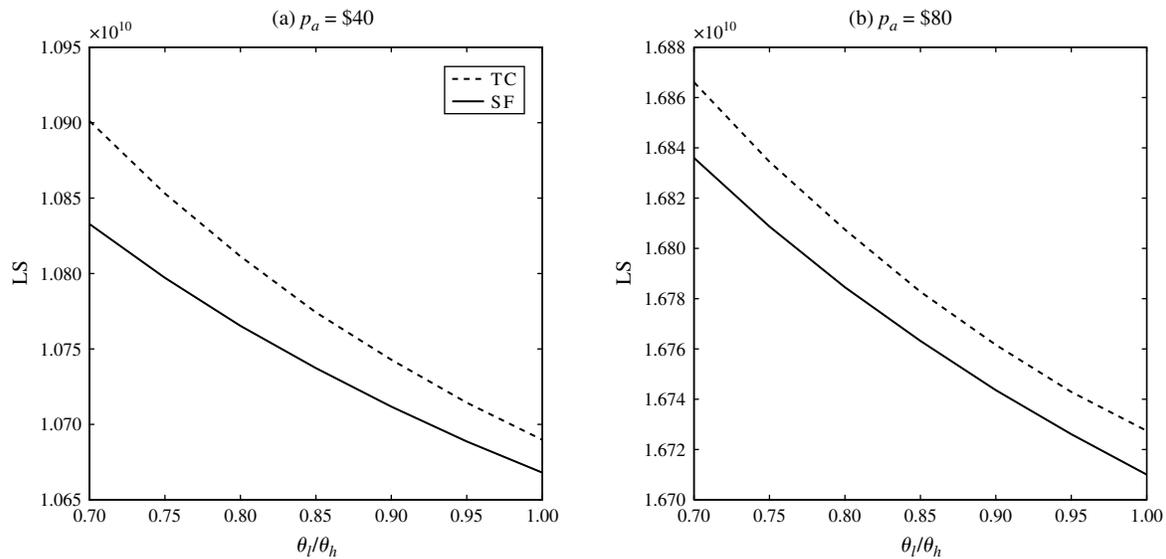
Proposition 6 shows that despite the information asymmetry, the government can design a menu of contracts based on input (effort) and output (shortage) measures that leads to "strict screening" between the manufacturer types without the need to pay any information rent, a first-best outcome. In other words, this menu resolves any potential inefficiency for the government due to the information asymmetry.

Proposition 6. *A contract menu that allows the manufacturer to choose between the following two contracts is first-best*: (i) *An order of f_l^S with a cost-plus contract, where the government covers the manufacturer's operational costs* (*costs of effort in primary and late production periods*) *and pays a surplus, $\epsilon = -R/n_l^S$, per unit of effort in the primary production period* (*with the total surplus payment capped at a maximum value of $-R$) to satisfy the manufacturer's individual rationality constraint.* (ii) *An order of f_h^S with wholesale price $p_{r,h} = \int_0^{k_h^S} \delta p_a + L(\theta_h u)^{-1} dG(u) - R/(f_h^S N)$ and shortage penalty $p_s = \delta p_a$.*

The result in Proposition 6 is important and counterintuitive. Standard economic theory predicts that to elicit the "agent's" private information, the "principal" must pay a positive information rent to the informed agent. In a context with one supplier and one buyer, for example, Corbett et al. (2004) study different contractual forms and investigate the value of information for an uninformed supplier when contracting with an informed buyer. They show that the profit of the informed buyer decreases if the supplier obtains full information. Such a positive surplus that the informed agent earns under information asymmetry is known as information rent (Laffont and Martimort 2002). In our setting, a menu with the manufacturer's effort and the realized shortage as performance metrics allows us to create countervailing incentives: if the manufacturer

has a lower yield distribution, it prefers to be incentivized on its effort, else on its ability to satisfy the government's order. Although a positive information rent is a standard result in settings with adverse selection, contextual parameters may help reduce the amount. For instance, in multitasking settings, the principal can extract some of the information rent from an informed agent if there is a negative correlation between the costs associated with the different tasks that are to be carried out by the agent (Riordan and Sappington 1987). Exploiting the differences in the queuing dynamics of two systems with equal overall capacity but different staffing level and productivity, Hasija et al. (2008) show that a rent-extracting mechanism can be devised using pay-per-call and pay-per-time terms in a call center outsourcing setting, where the agent is privately informed about its productivity. Proposition 6 adds to this literature.

An important element to note in our model is that a simple mechanism that ensures that the government covers the manufacturer's costs would also theoretically yield the first-best outcome. However, it trivially follows that such a mechanism would make the incentive compatibility constraints of both manufacturer types tight. In other words, such a mechanism would leave both manufacturer types indifferent between choosing the contract that is tailored for their productivity or the other. This implies that the government cannot guarantee strict information revelation from such a mechanism by providing an ϵ extra payment to one of the two contracts in the menu. Different from such a mechanism, as shown in the proof of Proposition 6, the menu proposed in Proposition 6 allows for a *strict* separation between the low type and high type and does not leave the two types indifferent. This ensures the practical implementability of the

menu in extracting information rent from the high-type manufacturer (i.e., a small extra payment added to the low-type contract will ensure the low-type manufacturer strictly prefers the low-type contract, without creating an incentive for the high-type manufacturer to switch to the low-type contract).

This allows us to devise a rent-extracting menu that leads to the first-best outcome, and justifies the claim that the expected information rent $(1-q)\Psi_h$ is an upper bound on how much a government should be willing to pay to verify a manufacturer's production effort. Section 6 computed $(1-q)\Psi_h$ for a range of reasonable parameter values for the inputs. We remark that those inputs suggest vaccination fractions that are similar to vaccination fractions reported for the United States although those vaccination fractions were not input into the analysis (CDC 2015) reported 2014–2015 influenza season vaccination coverage rates of 46.3% for those over 18 and 59.3% for children). Thus, the analysis of Section 6 may be a reasonable approximation to the value of verifying manufacturer effort, depending on the values of those parameters in a given year. The online Appendix D provides further caveats on the estimation of those parameters.

8. Conclusion

In this study of influenza vaccine procurement where production yield is uncertain, information rent is inevitable from a government to a manufacturer if there is both unverifiable effort and information asymmetry. Despite the challenge of solving a constrained optimization problem with a function as a decision variable, we are able to characterize a menu of contracts that attains global optimality. Further analysis of this menu suggests certain characteristics that render it impractical. Proposition 5 provides an output-based menu of contracts (output by a deadline) if the manufacturer is believed to be of one of only two types (low productivity and high productivity), that is optimal within a more practically implementable family of contracts. Our numerical experiments show that this menu does not lead to substantial losses due to agency issues. This menu is structured to encourage productivity improvements (low-productivity manufacturers get no information rent). Proposition 6 presents a novel menu of contracts that eliminates information rent if effort is also contractible.

Numerical examples with representative values for the influenza supply chain indicate that the information rent for a high-productivity manufacturer (when there is unverifiable effort and information asymmetry) can range from 0 to 4% of the total governmental spend on influenza vaccine procurement and administration and the cost of treating those infected with influenza. Key drivers of the information rent are the

degree to which the government believes the manufacturer's productivity may vary (θ_l/θ_h) and the variance in production yield for any given level of productivity, with vaccine administration costs also having an important but less strong influence. Less significant factors include higher production and administration costs incurred when low yields cause a late delivery of some treatments. The results appear to be robust to typical deviations from the assumption that all doses are delivered by the start of the influenza seasons.

The cost of monitoring manufacturer effort appears to be more than offset by the reduction in expected information rent that is afforded by the above menus of contracts, except potentially when there is a strong belief that the manufacturer has low productivity at the same time that the difference is small between low and high productivity levels.

Appendix A
Mathematical Proofs

Proof of Proposition 1. Given f, the manufacturer problem is strictly convex in n. This is because $d\mathrm{MF}(n;f)/dn = c - L\int_0^{fN/(\theta_i n)} dG(u)$ and $d^2\mathrm{MF}(n;f)/dn^2 = LfNg(fN/n)/(\theta_i n^2) > 0$. Consequently the first-order condition (FOC) is the necessary and sufficient condition for optimality, i.e., $n_i^*(f)$ solves $\int_0^{fN/(\theta_i n)} dG(u) = c/L$.

The result for k_i^G follows from (2), which implies that $fN/(\theta_i n_i^*(f))$ is a constant, which we call k_i^G. □

Proof of Proposition 2. By Assumption 2, $f > 0$ and $f < 1$ at optimality. From (3), the KKT conditions for the objective function $SF(f,n) = bT(f) + p_a fN + cn + \mathrm{E}[p_a\delta(fN - n\theta_i U)^+ + L(fN/(\theta_i U) - n)^+]$ give

$$0 = b\frac{dT(f)}{df} + p_a N\left(1 + \delta\int_0^{fN/(\theta_i n)} dG(u)\right)$$
$$+ L\int_0^{fN/(\theta_i n)} \frac{N}{\theta_i u} dG(u) \tag{18}$$
$$0 = c - p_a\delta\int_0^{fN/(\theta_i n)} \theta_i u\, dG(u) - L\int_0^{fN/(\theta_i n)} dG(u) - \gamma$$
$$0 = -\gamma n,$$

because terms with Leibnitz's rule that involve derivatives of the integrands have the value 0, and where γ is the Lagrange multiplier for the constraint $n \geq 0$.

We note that $\partial SF(f,n)/\partial n\,|_{n=0} = c - p_a\delta\theta_i\mathrm{E}[U] - L < 0$ implies that $n > 0$ at optimality, so $\gamma = 0$. Then (18) implies $\int_0^{fN/(\theta_i n)} p_a\delta\theta_i u + L\,dG(u) = c$. Comparing this with the definition of k_i^S in (4) and equating the upper integrands shows that $n_i^S = f_i^S N/k_i^S$ as claimed.

Dividing both sides of (4) by L, noting that each term in the sum on the left-hand side of (4) is nonnegative, observing that $\int_0^k L\,dG(u)$ is increasing in k, and recalling (2), we conclude that $k_i^S \leq k_i^G$ with strict inequality when $\delta > 0$ (because $p_a > 0$ by assumption).

Thus, the optimal fraction f_i^S is seen to solve

$$0 = b\frac{dT(f)}{df} + p_a N\left(1 + \delta \int_0^{fN/(\theta_i n)} dG(u)\right)$$
$$+ L\int_0^{fN/(\theta_i n)} N(\theta_i u)^{-1}\, dG(u) = b\frac{dT(f)}{df} + N\mathfrak{B}_i,$$

by substituting $k_i^S = f_i^S N/n_i^S$ and recalling the definition of $\mathfrak{B}_i$, as claimed in (5).

We now verify the convexity claim. The Hessian of the objective function SF is

$$H = \begin{bmatrix} H_{1,1} & H_{1,2} \\ H_{2,1} & H_{2,2} \end{bmatrix}$$

where

$$H_{1,1} = b\frac{d^2 T(f)}{df^2} + p_a \delta \frac{N^2}{\theta_i n} g\left(\frac{fN}{\theta_i n}\right) + L\frac{N}{\theta_i f} g\left(\frac{fN}{\theta_i n}\right),$$

$$H_{1,2} = -p_a \delta f \frac{N^2}{\theta_i n^2} g\left(\frac{fN}{\theta_i n}\right) - L\frac{N}{\theta_i n} g\left(\frac{fN}{\theta_i n}\right),$$

$$H_{2,1} = -p_a \delta f \frac{N^2}{\theta_i n^2} g\left(\frac{fN}{\theta_i n}\right) - L\frac{N}{\theta_i n} g\left(\frac{fN}{\theta_i n}\right),$$

$$H_{2,2} = p_a \delta \frac{(fN)^2}{\theta_i n^3} g\left(\frac{fN}{\theta_i n}\right) + L\frac{fN}{\theta_i n^2} g\left(\frac{fN}{\theta_i n}\right).$$

By inspection, $H_{2,2}$ is positive. The determinant of H is

$$\det(H) = b\frac{d^2 T(f)}{df^2}\frac{fN}{\theta_i n^2} g\left(\frac{fN}{\theta_i n}\right)(p_a \delta fN/n + L).$$

Positivity of $\det(H)$ implies that $d^2 T(f)/df^2 > 0$, which in turn guarantees that $H_{1,1}$ is positive and the principal minors of H are positive if $d^2 T(f)/df^2 > 0$. If f_i^S is not on the convex part of $T(f)$ then $d^2 T(f_i^S)/df^2 \leqslant 0$, which would contradict the assumption that f_i^S be a minimizer. Thus $d^2 T(f_i^S)/df^2 > 0$ and $T(f)$ is convex at f_i^S. $\quad\square$

Proof of Proposition 3. Let $\tau_j()$ determine the contract for type $j \in \{l, h\}$ as in (13) with parameters f_j, $p_{r,j}$, $p_{d,j}$, and $p_{s,j}$ so that the cost of an type-i manufacturer that chooses contract type j in (7) can be written

$$\mathrm{MF}_{i,j}(n) = \mathrm{E}\left[cn + p_{s,j}(f_j N - n\theta_i U)^+ + L\frac{(f_j N - n\theta_i U)^+}{\theta_i U}\right.$$
$$\left. - p_{d,j} n\theta_i U - p_{r,j} f_j N\right]. \quad (19)$$

The manufacturer chooses $n_{i,j}$ to minimize its costs, $\mathrm{MF}_{i,j}(n_{i,j})$. The FOC is

$$\int_0^{f_j N/(\theta_i n_{i,j})} (p_{s,j}\theta_i u + L)\, dG(u) + p_{d,j}\theta_i \mathrm{E}[U] = c, \quad (20)$$

which uniquely characterizes the optimal effort level because the manufacturer's cost function is strictly convex in n:

$$\frac{\partial^2 \mathrm{MF}_{i,j}(n)}{\partial n^2} = \frac{f_j N}{\theta_i n^2}\left(p_{s,j}\theta_i \frac{f_j N}{\theta_i n} + L\right) g\left(\frac{f_j N}{\theta_i n}\right) > 0.$$

From the revelation principle, the optimal menu of contracts must be truth revealing. Assume that an optimal menu induces order $\hat{f}_l$ and effort level $\hat{n}_{l,l}$ under the low-type contract and $\hat{f}_h$ and $\hat{n}_{h,h}$ under the high-type contract. We now show that the contract specified above by f_j, $p_{r,j}$, $p_{d,j}$, $p_{s,j}$ can

indeed match the order and effort amounts under an optimal menu of contracts. In other words, we seek to show the following four conditions.

C1. Vaccine orders are $\hat{f}_l$ and $\hat{f}_h$.

C2. The menu of contracts induces optimal effort levels $\hat{n}_{l,l}$ and $\hat{n}_{h,h}$.

C3. $\mathrm{MF}_{i,i}(\hat{n}_{i,i}) = \widehat{\mathrm{MF}}_{i,i}$ where $\widehat{\mathrm{MF}}_{i,i}$ is the manufacturer cost under the optimal contract.

C4. The menu of contracts achieves truthful revelation.

Condition C1 is trivially satisfied by setting the orders equal to $\hat{f}_l$ and $\hat{f}_h$. For Condition C2, note that the manufacturer's optimal effort level satisfies (20). Condition C2 is satisfied for any choice of $p_{s,i}$ if we set

$$p_{d,i} = \frac{1}{\theta_i \mathrm{E}[U]}\left(c - \int_0^{\hat{f}_i N/(\theta_i \hat{n}_{i,i})} (p_{s,i}\theta_i u + L)\, dG(u)\right). \quad (21)$$

So far, we have chosen f_i and have shown how to choose $p_{d,i}$ for a given $p_{s,i}$ so that Conditions C1 and C2 are satisfied. We now need to choose $p_{r,i}$ and $p_{s,i} \geqslant 0$ such that Conditions C3 and C4 are satisfied. To do so, we denote by $n_{i,j}$ the optimal efforts of the manufacturer to minimize $\mathrm{MF}_{i,j}(n)$, so that (20) is satisfied. We also let $\hat{u}_{i,j} = \hat{f}_j N/(\theta_i n_{i,j})$ to simplify notation. With these choices, (21) implies that $n_{i,i} = \hat{n}_{i,i}$ and that $\mathrm{MF}_{i,i}(\hat{n}_{i,i})$ simplifies for $i \in \{l, h\}$ as

$$\mathrm{MF}_{i,i}(\hat{n}_{i,i}) = \int_0^{\hat{u}_{i,i}} \left(p_{s,i}\hat{f}_i N + L\frac{\hat{f}_i N}{\theta_i u}\right) dG(u) - p_{r,i}\hat{f}_i N. \quad (22)$$

Given (22), it is clear that Condition C3 can be satisfied for any choice of $p_{s,i}$ by setting $p_{r,i}$ so that $\mathrm{MF}_{i,i}(\hat{n}_{i,i}) = \widehat{\mathrm{MF}}_{i,i}$. Note that $\widehat{\mathrm{MF}}_{i,i} \leqslant R \leqslant 0$ is the IR constraint under the optimal contract.

It remains to choose the $p_{s,i}$, using Condition C4. This condition requires that two inequalities hold:

$$\text{C4A:} \quad \mathrm{MF}_{h,h}(\hat{n}_{h,h}) \leqslant \mathrm{MF}_{h,l}(n_{h,l}), \quad (23)$$

$$\text{C4B:} \quad \mathrm{MF}_{l,l}(\hat{n}_{l,l}) \leqslant \mathrm{MF}_{l,h}(n_{l,h}).. \quad (24)$$

Consider first Condition C4A. Choosing $p_{r,h}$ for any given $p_{s,h}$ to satisfy Condition C3 as above, substituting $p_{d,l}$ from (21) into the right-hand side of (23), and simplifying, we obtain

$$\widehat{\mathrm{MF}}_{h,h} \leqslant cn_{h,l}\left(1 - \frac{\theta_h}{\theta_l}\right)$$
$$+ \int_0^{\hat{u}_{h,l}} \left(p_{s,l}\hat{f}_l N + L\frac{\hat{f}_l N}{\theta_h u} - Ln_{h,l}\left(1 - \frac{\theta_h}{\theta_l}\right)\right) dG(u)$$
$$+ \int_{\hat{u}_{h,l}}^{\hat{u}_{l,l}} n_{h,l}\left(p_{s,l}\theta_h u + L\frac{\theta_h}{\theta_l}\right) dG(u) - p_{r,l}\hat{f}_l N. \quad (25)$$

Recalling (22), $p_{r,l}$ can be chosen for any $p_{s,l}$ so that

$$\int_0^{\hat{u}_{l,l}} (p_{s,l}\hat{f}_l N + L(\hat{f}_l N/(\theta_l u)))\, dG(u) - \widehat{\mathrm{MF}}_{l,l} = p_{r,l}\hat{f}_l N.$$

Thus, (25) can be rewritten as

$$\widehat{\mathrm{MF}}_{h,h} \leqslant cn_{h,l}\left(1 - \frac{\theta_h}{\theta_l}\right)$$
$$+ \int_0^{\hat{u}_{h,l}} L\left(\frac{\hat{f}_l N}{u}\left(\frac{1}{\theta_h} - \frac{1}{\theta_l}\right) - n_{h,l}\left(1 - \frac{\theta_h}{\theta_l}\right)\right) dG(u)$$

$$+ \int_{\hat{u}_{h,l}}^{\hat{u}_{l,l}} \left(p_{s,l}(n_{h,l}\theta_h u - \hat{f}_l N) + \frac{L}{\theta_l}\left(n_{h,l}\theta_h - \frac{\hat{f}_l N}{u} \right) \right) dG(u) + \widehat{\text{MF}}_{l,l}. \quad (26)$$

The right-hand side of (26) is linear in $p_{s,l}$ with the coefficient of $p_{s,l}$ equal to

$$\int_{\hat{u}_{h,l}}^{\hat{u}_{l,l}} (n_{h,l}\theta_h u - \hat{f}_l N)\, dG(u) + \int_0^{\hat{u}_{l,l}} (\hat{f}_l N - n_{l,l}\theta_l u)\, dG(u).$$

The second integral in the coefficient on $p_{s,l}$ belongs to $\widehat{\text{MF}}_{l,l}$ and is obviously positive. The first integral in the coefficient of $p_{s,l}$ is also positive. To see this, we distinguish two cases:

(i) $\hat{u}_{h,l} \leqslant \hat{u}_{l,l}$ or, equivalently, $\hat{f}_l N/(\theta_h n_{h,l}) \leqslant \hat{f}_l N/(\theta_l n_{l,l})$. Because $\hat{f}_l N/(\theta_h n_{h,l}) \leqslant u$ and so $\theta_h n_{h,l} u - \hat{f}_l N \geqslant 0$, it follows that the integrand is nonnegative over the interval of integration.

(ii) $\hat{u}_{l,l} \leqslant \hat{u}_{h,l}$. In this case, we rewrite the integral as

$$\int_{\hat{u}_{l,l}}^{\hat{u}_{h,l}} (\hat{f}_l N - n_{l,l}\theta_l u)\, dG(u).$$

Because $u \leqslant \hat{u}_{h,l} = \hat{f}_l N/(\theta_h n_{h,l})$, it follows that $\hat{f}_l N - n_{l,l}\theta_l u \geqslant 0$; that is, the integrand is nonnegative over the interval of integration.

Therefore, there exists $\bar{p}_{s,l} \geqslant 0$ such that (26) holds for all $p_{s,l} \geqslant \bar{p}_{s,l}$. The parameters $p_{d,l}$ and $p_{r,l}$ were determined to satisfy Conditions C2 and C3 for any such $p_{s,l}$ as discussed above.

A similar argument shows that Condition C4B can be satisfied by choosing a $p_{s,h}$ such that

$$\widehat{\text{MF}}_{l,l} \leqslant c n_{l,h}\left(1 - \frac{\theta_l}{\theta_h}\right) + \int_0^{\hat{u}_{l,h}} \left(L\frac{\hat{f}_h N}{\theta_l u} - L n_{l,h} \right) dG(u)$$

$$+ \int_0^{\hat{u}_{h,h}} \left(L n_{l,h}\frac{\theta_l}{\theta_h} - L\frac{\hat{f}_h N}{\theta_h u} \right) dG(u)$$

$$+ \int_{\hat{u}_{h,h}}^{\hat{u}_{l,h}} p_{s,h}(\hat{f}_h N - n_{l,h}\theta_l u)\, dG(u) + \widehat{\text{MF}}_{h,h}. \quad (27)$$

The right-hand side of (27) is linear in $p_{s,h}$ and with coefficient, which is positive (similar to the argument for C4A with $p_{s,l}$). Therefore, there exists $\bar{p}_{s,h} \geqslant 0$ such that (26) holds for all $p_{s,h} \geqslant \bar{p}_{s,h}$. The parameters $p_{d,h}$ and $p_{r,h}$ were determined to satisfy Conditions C2 and C3 for any such $p_{s,h}$ as discussed above. □

Proof of Proposition 4. We recall the optimization problem, which determines the optimal four parameter linear contract menu in (14) subject to the constraints in (7)–(11). The optimal manufacturer effort is characterized by (20). To characterize the optimal menu, we first ignore constraint (10) from the optimization problem, solve the program and then show that the solution satisfies constraint (10).

Let Λ be the Lagrangian for the relaxed problem and the Lagrange multipliers for constraints (9), (11) with $i = h$, and (11) with $i = l$, be α_1, α_2, and α_3, respectively. We have

$$\frac{d\Lambda}{dp_{r,l}} = (q + \alpha_1 - \alpha_3)f_l N = 0, \quad (28)$$

$$\frac{d\Lambda}{dp_{r,h}} = (1 - q - \alpha_1 - \alpha_2)f_h N = 0, \quad (29)$$

$$\frac{d\Lambda}{dp_{s,l}} = \frac{d\Lambda}{dp_{s,l}} = q\left((p_a\delta - p_{s,l})\frac{dA_{l,l}}{dp_{s,l}} - A_{l,l} + p_{d,l}\theta_l \text{E}[U]\frac{\partial n_{l,l}}{\partial p_{s,l}} \right)$$

$$+ \alpha_3\left(A_{l,l} + p_{s,l}\frac{dA_{l,l}}{dp_{s,l}} + (c - p_{d,l}\theta_l \text{E}[U])\frac{\partial n_{l,l}}{\partial p_{s,l}} + L\frac{dB_{l,l}}{dp_{s,l}} \right)$$

$$+ \alpha_1\left((-c + p_{d,l}\theta_h \text{E}[U])\frac{\partial n_{h,l}}{\partial p_{s,l}} \right.$$

$$\left. - L\frac{dB_{h,l}}{Dp_{s,l}} - A_{h,l} - p_{s,l}\frac{dA_{h,l}}{dp_{s,l}} \right) = 0, \quad (30)$$

$$\frac{d\Lambda}{dp_{s,h}} = (1 - q)\left((p_a\delta - p_{s,h})\frac{dA_{h,h}}{dp_{s,h}} + p_{d,h}\theta_h \text{E}[U]\frac{\partial n_{h,h}}{\partial p_{s,h}} \right)$$

$$+ (\alpha_1 + \alpha_2 - 1 + q)A_{h,h} = 0, \quad (31)$$

$$0 = \frac{d\Lambda}{df_l}, \quad (32)$$

$$0 = \frac{d\Lambda}{df_h}, \quad (33)$$

$$0 = \alpha_1(\text{MF}_{h,h}(n_{h,h}) - \text{MF}_{h,l}(n_{h,l})), \quad (34)$$

$$0 = \alpha_2(\text{MF}_{h,h}(n_{h,h}) - R), \quad (35)$$

$$0 = \alpha_3(\text{MF}_{l,l}(n_{l,l}) - R), \quad (36)$$

where $A_{i,j}$ and $B_{i,j}$ are as in (12). Combining (28) and (29) yields

$$\alpha_2 + \alpha_3 = 1. \quad (37)$$

Equation (30) can then be simplified as

$$\frac{d\Lambda}{dp_{s,l}} = q\left((p_a\delta - p_{s,l})\frac{dA_{l,l}}{dp_{s,l}} + p_{d,l}\theta_l \text{E}[U]\frac{\partial n_{l,l}}{\partial p_{s,l}} \right)$$

$$+ (\alpha_3 - q)A_{l,l} - \alpha_1 A_{h,l} = 0. \quad (38)$$

We can now prove the following result.

Lemma 1. *The high-type manufacturer is better off than the low-type manufacturer under both contracts in the menu, i.e.,* $\text{MF}_{h,j}(n_{h,j}) < \text{MF}_{l,j}(n_{l,j})$ *for* $j \in \{l, h\}$.

Proof of Lemma 1. By definition of $n_{h,j}$, $\text{MF}_{h,j}(n_{h,j}) \leqslant \text{MF}_{h,j}(n_{l,j})$. Therefore,

$$\text{MF}_{h,j}(n_{h,j}) - \text{MF}_{l,j}(n_{l,j}) \leqslant \text{MF}_{h,j}(n_{l,j}) - \text{MF}_{l,j}(n_{l,j}).$$

Because $\theta_h > \theta_l$, it follows that $\text{MF}_{h,j}(n_{l,j}) - \text{MF}_{l,j}(n_{l,j}) < 0$; hence $\text{MF}_{h,j}(n_{h,j}) - \text{MF}_{l,j}(n_{l,j}) < 0$. ◇

From Lemma 1, $\text{MF}_{h,j}(n_{h,j}) < \text{MF}_{l,j}(n_{l,j})$. Therefore, $\text{MF}_{h,h}(n_{h,h}) < \text{MF}_{l,h}(n_{l,h})$ and $\text{MF}_{h,l}(n_{h,l}) < \text{MF}_{l,l}(n_{l,l})$. Because $\text{MF}_{l,l} \leqslant R$, we conclude that $\text{MF}_{h,h}(n_{h,h}) \leqslant \text{MF}_{h,l}(n_{h,l}) < \text{MF}_{l,l} \leqslant R$. The first inequality follows from (9) and the second inequality from (11). Now, because $\text{MF}_{h,h}(n_{h,h}) < R$, we conclude that $\alpha_2 = 0$. As a result, (37) implies $\alpha_3 = 1$ and (28) in turn implies $\alpha_1 = 1 - q$.

Substituting the values of the Lagrangian multipliers into (38) and (31), we obtain

$$0 = q\left(p_a\delta - p_{s,l})\frac{dA_{l,l}}{dp_{s,l}} + p_{d,l}\theta_l \text{E}[U]\frac{\partial n_{l,l}}{\partial p_{s,l}} \right) + (1 - q)(A_{l,l} - A_{h,l})$$

$$0 = (1 - q)\left((p_a\delta - p_{s,h})\frac{dA_{h,h}}{dp_{s,h}} + p_{d,h}\theta_h \text{E}[U]\frac{\partial n_{h,h}}{\partial p_{s,h}} \right).$$

We further have

$$\frac{d\Lambda}{dp_{d,h}} = (1-q)\left(\theta_h n_{h,h}\mathrm{E}[U] + p_{d,h}\theta_h\mathrm{E}[U]\frac{\partial n_{h,h}}{\partial p_{d,h}}\right.$$
$$+ (p_a\delta - p_{s,h})\frac{dA_{h,h}}{dp_{d,h}}\Big)$$
$$+ (1-q)\left(c\frac{\partial n_{h,h}}{\partial p_{d,h}} - \theta_h n_{h,h}\mathrm{E}[U] - p_{d,h}\theta_h\mathrm{E}[U]\frac{\partial n_{h,h}}{\partial p_{d,h}}\right.$$
$$+ p_{s,h}\frac{dA_{h,h}}{dp_{d,h}} + L\frac{dB_{h,h}}{dp_{d,h}}\Big)$$
$$= p_a\delta\frac{dA_{h,h}}{dp_{d,h}} + L\frac{dB_{h,h}}{dp_{d,h}} + c\frac{\partial n_{h,h}}{\partial p_{d,h}} = 0. \tag{39}$$

Because $dA_{h,h}/dp_{d,h} = (\partial A_{h,h}/\partial n_{h,h})(\partial n_{h,h}/\partial p_{d,h})$, we can rewrite (39) as

$$\frac{d\Lambda}{dp_{d,h}} = (1-q)\frac{\partial n_{h,h}}{\partial p_{d,h}}\left(p_a\delta\frac{\partial A_{h,h}}{\partial n_{h,h}} + L\frac{\partial B_{h,h}}{\partial n_{h,h}} + c\right) = 0,$$

which implies (because $\partial n_{h,h}/\partial p_{d,h} > 0$)

$$p_a\delta\frac{\partial A_{h,h}}{\partial n_{h,h}} + L\frac{\partial B_{h,h}}{\partial n_{h,h}} + c = 0. \tag{40}$$

From Proposition 2, it follows that $n_{h,h} = n_h^S$ (no distortion for the efficient type). From (40) and (20), we obtain

$$p_{s,h}\frac{\partial A_{h,h}}{\partial n_{h,h}} + L\frac{\partial B_{h,h}}{\partial n_{h,h}} + c - p_{d,h}\theta_h\mathrm{E}[U] = \delta p_a\frac{\partial A_{h,h}}{\partial n_{h,h}} + L\frac{\partial B_{h,h}}{n_{h,h}} + c,$$

which is satisfied, though not necessarily uniquely, if we set $p_{s,h} = \delta p_a$ and $p_{d,h} = 0$.

We next prove that $f_h = f_h^S$. Recall that $p_{s,h} = p_a\delta$, $\alpha_1 = 1 - q$ and $\alpha_2 = 0$. We expand and simplify $d\Lambda/df_h$ to obtain that f_h solves

$$\frac{d\Lambda}{df_h} = (1-q)\left(b\frac{dT(f_h)}{df_h} + p_a N + p_{r,h}N\right) + (1-q)$$
$$\cdot\left(c\frac{\partial n_{h,h}}{\partial f_h} + L\frac{dB_{h,h}}{df_h} - p_{r,h}N + p_a\delta\frac{dA_{h,h}}{df_h}\right) = 0. \tag{41}$$

The derivatives of $A_{i,j}$ and $B_{i,j}$ can be expanded as

$$\frac{dA_{i,j}}{df_j} = \frac{\partial A_{i,j}}{\partial f_j} + \frac{\partial A_{i,j}}{\partial n_{i,j}}\frac{\partial n_{i,j}}{\partial f_j}, \tag{42}$$

$$\frac{dB_{i,j}}{df_j} = \frac{\partial B_{i,j}}{\partial f_j} + \frac{\partial B_{i,j}}{\partial n_{i,j}}\frac{\partial n_{i,j}}{\partial f_j}, \tag{43}$$

and substituted into (41). Because $p_{d,h} = 0$, then (20) with $i = j = h$ can be rewritten as

$$\int_0^{f_h N/(\theta_h n_{h,h})} (p_{s,h}\theta_h u + L)dG(u) = c, \tag{44}$$

and hence $f_h N/(\theta_h n_{h,h}(f_h))$ the upper integrand in each of $A_{h,h}$ and $B_{h,h}$, is a constant at optimality. We use this fact and the definitions of $A_{i,j}$ and $B_{i,j}$ in order to rewrite (44)

$$c + L\frac{\partial B_{h,h}}{\partial n_{h,h}} + p_{s,h}\frac{\partial A_{h,h}}{\partial n_{h,h}} = 0. \tag{45}$$

Substituting (42) and (43) together with (45) into (41), we obtain

$$0 = b\frac{dT(f_h)}{df_h} + p_a N + L\frac{\partial B_{h,h}}{\partial f_h} + \delta p_a\frac{\partial A_{h,h}}{\partial f_h}\Big|_{f_h=f_h}$$
$$= b\frac{dT(f_h)}{df_h} + N\left(p_a + \int_0^{f_h N/(\theta_h n_{h,h})}\frac{L}{\theta_h u} + \delta p_a\,dG(u)\right)\Big|_{f_h=f_h}. \tag{46}$$

This equation is recognizable as the condition for the optimal vaccination fraction in the system setting in Proposition 2 for a high-type manufacturer. Thus $f_h = f_h^S$ as required. Because $p_{s,h} = p_a\delta$, a high-type manufacturer expends the system optimal effort. Thus $n_{h,h}$ equals n_h^S as required.

We now show that $p_{d,l} < 0$ in the optimal menu of contracts. Observe that

$$\frac{d\Lambda}{dp_{d,l}} = q\left(\theta_l n_{l,l}\mathrm{E}[U] + p_{d,l}\theta_l\mathrm{E}[U]\frac{\partial n_{l,l}}{\partial p_{d,l}} + (p_a\delta - p_{s,l})\frac{dA_{l,l}}{dp_{d,l}}\right)$$
$$+ (1-q)\left(n_{h,l}\theta_h\mathrm{E}[U] + p_{d,l}\theta_h\mathrm{E}[U]\frac{\partial n_{h,l}}{\partial p_{d,l}} - c\frac{\partial n_{h,l}}{\partial p_{d,l}}\right.$$
$$- p_{s,l}\frac{dA_{h,l}}{dp_{d,l}} - L\frac{dB_{h,l}}{dp_{d,l}}\Big)$$
$$- p_{d,l}\theta_l\mathrm{E}[U]\frac{\partial n_{l,l}}{\partial p_{d,l}} - \theta_l\mathrm{E}[U]n_{l,l}$$
$$+ p_{s,l}\frac{dA_{l,l}}{dp_{d,l}} + L\frac{dB_{l,l}}{dp_{d,l}} + c\frac{\partial n_{l,l}}{\partial p_{d,l}} = 0. \tag{47}$$

Because $dA_{l,l}/dp_{d,l} = (\partial A_{l,l}/\partial n_{l,l})(\partial n_{l,l}/\partial p_{d,l})$ and $dB_{l,l}/dp_{d,l} = (\partial B_{l,l}/\partial n_{l,l})(\partial n_{l,l}/\partial p_{d,l})$, we can simplify (47) to obtain

$$\frac{d\Lambda}{dp_{d,l}} = q\left(\theta_l n_{l,l}\mathrm{E}[U] + p_{d,l}\theta_l\mathrm{E}[U]\frac{\partial n_{l,l}}{\partial p_{d,l}} + (p_a\delta - p_{s,l})\frac{dA_{l,l}}{dp_{d,l}}\right)$$
$$+ (1-q)\left(n_{h,l}\theta_h\mathrm{E}[U] + \frac{\partial n_{h,l}}{\partial p_{d,l}}\left(p_{d,l}\theta_h\mathrm{E}[U]\right.\right.$$
$$\left.\left. - c - p_{s,l}\frac{\partial A_{h,l}}{\partial n_{h,l}} - L\frac{\partial B_{h,l}}{\partial n_{h,l}}\right)\right) - \theta_l n_{l,l}\mathrm{E}[U]$$
$$- \frac{\partial n_{l,l}}{\partial p_{d,l}}\left(p_{d,l}\theta_l\mathrm{E}[U] - c - p_{s,l}\frac{\partial A_{l,l}}{\partial n_{l,l}} - L\frac{\partial B_{l,l}}{\partial n_{l,l}}\right)$$
$$= 0. \tag{48}$$

By (20), we can rewrite (48) as

$$\frac{d\Lambda}{dp_{d,l}} = q\left(\theta_l n_{l,l}\mathrm{E}[U] + p_{d,l}\theta_l\mathrm{E}[U]\frac{\partial n_{l,l}}{\partial p_{d,l}} + (p_a\delta - p_{s,l})\frac{dA_{l,l}}{dp_{d,l}}\right)$$
$$+ (1-q)n_{h,l}\theta_h\mathrm{E}[U] - \theta_l\mathrm{E}[U]n_{l,l} = 0,$$

or equivalently

$$\frac{d\Lambda}{dp_{d,l}} = q\left(p_{d,l}\theta_l\mathrm{E}[U]\frac{\partial n_{l,l}}{\partial p_{d,l}} + (p_a\delta - p_{s,l})\frac{dA_{l,l}}{dp_{d,l}}\right)$$
$$+ (1-q)(n_{h,l}\theta_h - \theta_l n_{l,l})\mathrm{E}[U]$$
$$= q\left(p_{d,l}\theta_l\mathrm{E}[U] + (p_a\delta - p_{s,l})\frac{\partial A_{l,l}}{\partial p_{d,l}}\right)\frac{\partial n_{l,l}}{\partial p_{d,l}}$$
$$+ (1-q)(n_{h,l}\theta_h - \theta_l n_{l,l})\mathrm{E}[U] = 0. \tag{49}$$

We also rewrite (38) as

$$\frac{d\Lambda}{dp_{s,l}} = q\left((p_a\delta - p_{s,l})\frac{\partial A_{l,l}}{\partial n_{l,l}} + p_{d,l}\theta_l\mathrm{E}[U]\right)\frac{\partial n_{l,l}}{\partial p_{s,l}}$$
$$+ (1-q)(A_{l,l} - A_{h,l}) = 0. \tag{50}$$

Further, from (20), we have

$$\frac{\partial n_{l,l}}{\partial p_{d,l}} = \frac{\theta_l E[U]}{p_{s,l}(\partial^2 A_{l,l}/\partial n_{l,l}^2) + L(\partial^2 B_{l,l}/\partial n_{l,l}^2)}, \quad (51)$$

$$\frac{\partial n_{l,l}}{\partial p_{s,l}} = \frac{\int_0^{f_l N/(\theta_l n_{l,l})} \theta_l u \, dG(u)}{p_{s,l}(\partial^2 A_{l,l}/\partial n_{l,l}^2) + L(\partial^2 B_{l,l}/\partial n_{l,l}^2)}. \quad (52)$$

We now use (50)–(52) to rewrite (49) as

$$\frac{d\Lambda}{dp_{d,l}} = (1-q)(n_{h,l}\theta_h - \theta_l n_{l,l})E[U] - \frac{(1-q)(A_{l,l}-A_{h,l})}{\partial n_{l,l}/\partial p_{s,l}} \frac{\partial n_{l,l}}{\partial p_{d,l}}$$

$$= (1-q)\Bigg(E[U](n_{h,l}\theta_h - n_{l,l}\theta_l)$$

$$\qquad\qquad + (A_{h,l}-A_{l,l})\frac{\theta_l E[U]}{\int_0^{f_l N/(\theta_l n_{l,l})} \theta_l u \, dG(u)}\Bigg)$$

$$= (1-q)E[U]\Bigg(n_{h,l}\theta_h - n_{l,l}\theta_l + \frac{A_{h,l}-A_{l,l}}{\int_0^{f_l N/(\theta_l n_{l,l})} \theta_l u \, dG(u)}\Bigg)$$

$$= \frac{(1-q)E[U]}{\int_0^{u_{l,l}} \theta_l u \, dG(u)}\Bigg(\int_0^{u_{h,l}} (f_l N - n_{h,l}\theta_h u)\,dG(u)$$

$$\qquad\qquad - \int_0^{u_{l,l}} (f_l N - n_{h,l}\theta_h u)\,dG(u)\Bigg)$$

$$= \frac{(1-q)E[U]}{\int_0^{u_{l,l}} \theta_l u \, dG(u)} n_{h,l}\theta_h \Bigg(\int_0^{u_{h,l}} (u_{h,l}-u)\,dG(u)$$

$$\qquad\qquad - \int_0^{u_{l,l}} (u_{h,l}-u)\,dG(u)\Bigg)$$

$$= \frac{(1-q)E[U]}{\int_0^{u_{l,l}} \theta_l u \, dG(u)} n_{h,l}\theta_h \int_{u_{h,l}}^{u_{l,l}} (u_{h,l}-u)\,dG(u) = 0, \quad (53)$$

where $u_{l,l} = f_l N/(\theta_l n_{l,l})$ and $u_{l,l} = f_l N/(\theta_h n_{h,l})$, and $u_{h,l} \leqslant u_{l,l}$. The last equality implies that $u_{l,l} = u_{h,l}$. Now, from (20), we have

$$\int_0^{u_{l,l}} (p_{s,l}\theta_l u + L)\,dG(u) = c - p_{d,l}\theta_l E[U],$$

$$\int_0^{u_{h,l}} (p_{s,l}\theta_h u + L)\,dG(u) = c - p_{d,l}\theta_l E[U].$$

Because $\theta_h > \theta_l$, the two equations above imply that $c - p_{d,l}\theta_l \cdot E[U] < c - p_{d,l}\theta_l E[U]$ or $p_{d,l}(\theta_h - \theta_l)E[U] < 0$, which holds if and only if $p_{d,l} < 0$. We now derive the expression for f_l:

$$\frac{d\Lambda}{df_l} = q\Bigg(b\frac{dT(f_l)}{df_l} + p_a N + c\frac{\partial n_{l,l}}{\partial f_l} + L\frac{dB_{l,l}}{df_l} + \delta p_a \frac{dA_{l,l}}{df_l}\Bigg)$$

$$\qquad - (1-q)\Bigg(c\Bigg(\frac{\partial n_{h,l}}{\partial f_l} - \frac{\partial n_{l,l}}{\partial f_l}\Bigg) + L\Bigg(\frac{dB_{h,l}}{df_l} - \frac{dB_{l,l}}{df_l}\Bigg)$$

$$\qquad + p_{s,l}\Bigg(\frac{dA_{h,l}}{df_l} - \frac{dA_{l,l}}{df_l}\Bigg) + p_{d,l}E[U]\Bigg(\theta_l\frac{\partial n_{l,l}}{\partial f_l} - \theta_h\frac{\partial n_{h,l}}{\partial f_l}\Bigg)\Bigg)$$

$$= 0. \quad (54)$$

The expressions for $p_{r,l}$ and $p_{r,h}$ are trivially determined by replacing the values of $\alpha_1 = 1 - q$ and $\alpha_3 = 1$ in (34) and (36), respectively. Finally, we need to show that the optimal contract parameters derived above satisfy (10). That is, we proceed to justify that $\text{MF}_{l,l}(n_{l,l}) \leqslant \text{MF}_{l,h}(n_{l,h})$, or that

$$cn_{l,l} + p_{s,l}\int_0^{f_l N/(\theta_l n_{l,l})} f_l N - n_{l,l}\theta_l u \, dG(u)$$

$$+ L\int_0^{f_l N/(\theta_l n_{l,l})} (f_l N)/(\theta_l u) - n_{l,l}\,dG(u) - p_{r,l}f_l N - p_{d,l}n_{l,l}\theta_l E[U]$$

$$\leqslant cn_{l,h} + p_{s,h}\int_0^{f_h N/(\theta_l n_{l,h})} f_h N - n_{l,h}\theta_l u \, dG(u)$$

$$+ L\int_0^{f_h N/(\theta_l n_{l,h})} (f_h N)/(\theta_l u) - n_{l,h}\,dG(u) - p_{r,h}f_h N. \quad (55)$$

We proceed by simplifying (55) using (20) to eliminate terms with the factor $n_{l,j}$ and divide by N to get

$$p_{s,l}f_l\int_0^{f_l N/(\theta_l n_{l,l})} dG(u) + \int_0^{f_l N/(\theta_l n_{l,l})} \frac{Lf_l}{\theta_l u}\,dG(u) - p_{r,l}f_l$$

$$\leqslant p_{s,h}f_h\int_0^{f_h N/(\theta_l n_{l,h})} dG(u) + \int_0^{f_h N/(\theta_l n_{l,h})} \frac{Lf_h}{\theta_l u}\,dG(u) - p_{r,h}f_h. \quad (56)$$

We know that $\text{MF}_{h,h}(n_{h,h}) = \text{MF}_{h,l}(n_{h,l})$ at optimality. A simplification similar to that for (56) gives

$$p_{s,l}f_l\int_0^{f_l N/(\theta_h n_{h,l})} dG(u) + \int_0^{f_l N/(\theta_h n_{h,l})} \frac{Lf_l}{\theta_h u}\,dG(u) - p_{r,l}f_l$$

$$= p_{s,h}f_h\int_0^{f_h N/(\theta_h n_{h,h})} dG(u) + \int_0^{f_h N/(\theta_h n_{h,h})} \frac{Lf_h}{\theta_h u}\,dG(u) - p_{r,h}f_h.$$

We now isolate the term $p_{r,h}f_h - p_{r,l}f_l$ in this last equation and substitute its value into (56) after adding $p_{r,h}f_h$ to both sides of (56). This gives

$$p_{s,l}f_l\int_0^{f_l N/(\theta_l n_{l,l})} dG(u) + \int_0^{f_l N/(\theta_l n_{l,l})} \frac{Lf_l}{\theta_l u}\,dG(u)$$

$$+ p_{s,h}f_h\int_0^{f_h N/(\theta_h n_{h,h})} dG(u) + \int_0^{f_h N/(\theta_h n_{h,h})} \frac{Lf_h}{\theta_h u}\,dG(u)$$

$$- p_{s,l}f_l\int_0^{f_l N/(\theta_h n_{h,l})} dG(u) - \int_0^{f_l N/(\theta_h n_{h,l})} \frac{Lf_l}{\theta_h u}\,dG(u)$$

$$\leqslant p_{s,h}f_h\int_0^{f_h N/(\theta_l n_{l,h})} dG(u) + \int_0^{f_h N/(\theta_l n_{l,h})} \frac{Lf_h}{\theta_l u}\,dG(u). \quad (57)$$

Because the optimal production effort under the low-type contract $(j = l)$ for both types of manufacturers solves (EC.1), we conclude $\theta_l n_{l,l} < \theta_h n_{h,l}$. Thus, the first term less the fifth term in (57), satisfies

$$p_{s,l}f_l\int_0^{f_l N/(\theta_l n_{l,l})} dG(u) - p_{s,l}f_l\int_0^{f_l N/(\theta_h n_{h,l})} dG(u) > 0.$$

The second term less the sixth term in (57) is similarly positive. This implies that a sufficient condition for (57) to hold is

$$p_{s,h}f_h\int_0^{f_h N/(\theta_h n_{h,h})} dG(u) + \int_0^{f_h N/(\theta_h n_{h,h})} \frac{Lf_h}{\theta_h u}\,dG(u)$$

$$\leqslant p_{s,h}f_h\int_0^{f_h N/(\theta_l n_{l,h})} dG(u) + \int_0^{f_h N/(\theta_l n_{l,h})} \frac{Lf_h}{\theta_l u}\,dG(u). \quad (58)$$

That is $\text{MF}_{l,l}(n_{l,l}) \leqslant \text{MF}_{l,h}(n_{h,l})$ holds iff (55) holds, and a sufficient condition for that to hold is for (58) to hold. Because the optimal production effort under the high-type contract $(j = h)$ for both types of manufacturers solves (20), we conclude that $\theta_l n_{l,h} < \theta_h n_{h,h}$, and therefore that (10) indeed holds. $\square$

The *proofs of Proposition* 5 *and Corollary* 1 are similar to that of Proposition 4 and are relegated to the online appendix.

Proof of Proposition 6. We want to show that if the manufacturer's $\theta = \theta_l$ then the manufacturer will accept contract (i) (otherwise it will accept contract (ii)) and that, given θ, the manufacturer will make the socially optimal decision under the respective contracts and will not earn any additional surplus.

First, assume that $\theta = \theta_h$. If the manufacturer chooses contract (ii), then its cost function is

$$\text{MF}(n) = cn + L \int_0^{f_h^S N/(\theta_h n)} \left(\frac{f_h^S N}{\theta_h u} - n \right) dG(u) - p_{r,h} f_h^S N$$
$$+ \delta p_a \int_0^{f_h^S N/(\theta_h n)} \left(f_h^S N - n\theta_h u \right) dG(u). \quad (59)$$

This function is convex in n. It follows that the FOC provides the necessary and sufficient conditions for optimality. Simplifying $d\text{MF}(n)/dn = 0$ implies that

$$\int_0^{f_h^S N/(\theta_h n)} \left(p_a \delta \theta_h u + L \right) dG(u) = c,$$

which is the same as for the social optimum in (4) with $k_h^S = f_h^S N / \theta_h n$ at the opmtimal n. By replacing the manufacturer's effort and contract parameters in the manufacturer's cost function, it is easy to verify that the manufacturer does not earn any additional surplus over its reservation value, i.e., its participation constraint is tight.

Next we show that the manufacturer is not better off by choosing contract (i). If the manufacturer chooses contract (i), its problem is

$$\min_{n \geqslant 0} \text{MF}(n) \triangleq cn + L \int_0^{f_l^S N/(\theta_h n)} \left(\frac{f_l^S N}{\theta_h u} - n \right) dG(u) - cn$$
$$- L \int_0^{f_l^S N/(\theta_h n)} \left(\frac{f_l^S N}{\theta_h u} - n \right) dG(u) - \min\{\epsilon n, -R\}.$$

It is straightforward to see that the manufacturer's optimal cost is R and hence, the manufacturer does not earn any additional surplus over its reservation value, i.e., its participation constraint is tight. Therefore the manufacturer is indifferent between contracts (i) and (ii). Given this, we can ensure that the manufacturer will strictly prefer contract (ii) by increasing $p_{r,h}$ by an arbitrarily small value. We will show below that if $\theta = \theta_l$, then the manufacturer strictly prefers contract (i), and hence an arbitrarily small increase in $p_{r,h}$ will not violate the screening property of our proposed menu. To conclude, if $\theta = \theta_h$, the manufacturer will choose contract (ii) and make the social optimum decision. This will minimize the system cost and simultaneously ensure that the manufacturer will not earn any additional surplus over its reservation value.

To show that if $\theta = \theta_l$, the manufacturer prefers contract (i) over contract (ii), makes the social optimum decision, and does not earn any additional surplus over its reservation value, observe that the manufacturer's problem upon choosing contract (i) is

$$\min_{n \geqslant 0} \text{MF}(n) = cn + L \int_0^{f_l^S N/(\theta_l n)} \left(\frac{f_l^S N}{\theta_l u} - n \right) dG(u) - cn$$
$$- L \int_0^{f_l^S N/(\theta_l n)} \left(\frac{f_l^S N}{\theta_l u} - n \right) dG(u) - \min\{\epsilon n, -R\}.$$

It follows that the manufacturer's optimal effort will be $-R/\epsilon = n_l^S$, so the manufacturer's effort is aligned with the system

optimal production effort. In addition, the manufacturer does not earn any additional surplus over its reservation value, i.e., its participation constraint is tight.

The proof is concluded if we can show that the manufacturer strictly prefers contract (i) over contract (ii). If the manufacturer chooses contract (ii) then the manufacturer's objective function becomes

$$\text{MF}(n) = cn + L \int_0^{f_h^S N/(\theta_h n)} \left(\frac{f_h^S N}{\theta_l u} - n \right) dG(u) - p_{r,h} f_h^S N$$
$$+ \delta p_a \int_0^{f_h^S N/(\theta_h n)} \left(f_h^S N - n\theta_l u \right) dG(u).$$

Because $\theta_l < \theta_h$ we observe that

$$L \int_0^{f_h^S N/(\theta_h n)} \left(\frac{f_h^S N}{\theta_l u} - n \right) dG(u)$$
$$+ \delta p_a \int_0^{f_h^S N/(\theta_h n)} \left(f_h^S N - n\theta_l u \right) dG(u)$$
$$> L \int_0^{f_h^S N/(\theta_h n)} \left(\frac{f_h^S N}{\theta_h u} - n \right) dG(u)$$
$$+ \delta p_a \int_0^{f_h^S N/(\theta_h n)} \left(f_h^S N - n\theta_h u \right) dG(u).$$

Therefore

$$\text{MF}(n; f) > cn + L \int_0^{f_h^S N/(\theta_h n)} \left(\frac{f_h^S N}{\theta_h u} - n \right) dG(u) - p_{r,h} f_h^S N$$
$$+ \delta p_a \int_0^{f_h^S N/(\theta_h n)} \left(f_h^S N - n\theta_h u \right) dG(u) \geqslant R. \quad \square$$

The online appendix contains further mathematical results and Appendices B, C, and D.

References

Adida E, Dey D, Mamani H (2013) Operational issues and network effects in vaccine markets. *Eur. J. Oper. Res.* 231(2):414–427.

Arifoglu K, Deo S, Iravani S (2012) Consumption externality and yield uncertainty in the influenza vaccine supply chain: Interventions in demand and supply sides. *Management Sci.* 58(6):1072–1091.

Baiman S, Fischer PE, Rajan MV (2000) Information, contracting, and quality costs. *Management Sci.* 46(6):776–789.

Brandeau ML, Zaric GS, Richter A (2003) Resource allocation for control of infectious diseases in multiple independent populations: Beyond cost-effectiveness analysis. *J. Health Econom.* 22(4): 572–598.

Cachon GP (2003) Supply chain coordination with contracts. Graves S, de Kok T, eds. *Handbook in Operations Research and Management Science: Supply Chain Management* (Elsevier, Amsterdam), 227–339.

Cachon G, Lariviere M (2005) Supply chain coordination with revenue sharing contracts: Strengths and limitations. *Management Sci.* 51(1):8–14.

Cachon GP, Zhang F (2006) Procuring fast delivery: Sole sourcing with information asymmetry. *Management Sci.* 52(6):881–896.

CDC (2015) Flu vaccination coverage, United States, 2014-15 influenza season. Accessed February 21, 2016, http://www.cdc.gov/flu/fluvaxview/coverage-1415estimates.htm.

Chick SE, Mamani H, Simchi-Levi D (2008) Supply chain coordination and influenza vaccination. *Oper. Res.* 56(6):1493–1506.

Cho SH (2010) The optimal composition of influenza vaccines subject to random production yields. *Manufacturing Service Oper. Management* 12(2):256–277.

Corbett CJ, Zhou D, Tang CS (2004) Designing supply contracts: Contract type and information asymmetry. *Management Sci.* 50(4):550–559.

Crama P, Reyck BD, Degraeve Z (2008) Milestone payments or royalties? Contract design for R&D licensing. *Oper. Res.* 56(6): 1539–1552.

Dai T, Cho SH, Zhang F (2016) Contracting for on-time delivery in the U.S. influenza vaccine supply chain. *Manufacturing Service Oper. Management* 18(3):332–346.

Deo S, Corbett CJ (2009) Cournot competition under yield uncertainty: the case of the U.S. influenza vaccine market. *Manufacturing Service Oper. Management* 11(4):563–576.

Diekmann O, Heesterbeek J (2000) *Mathematical Epidemiology of Infectious Diseases* (Johyn Wiley & Sons, Chichester, UK).

FDA (2005) Biological product shortages. http://www.fda.gov/cber/shortage/shortage.htm#flu.

Gerdil C (2003) The annual production cycle for influenza vaccine. *Vaccine* 21(6):1776–1779.

Gromb D, Martimort D (2007) Collusion and the organization of delegated expertise. *JET* 137(1):271–299.

Grosfeld-Nir A, Gerchak Y (2002) Multistage production to order with rework capability. *Management Sci.* 48(5):652–664.

Grosfeld-Nir A, Gerchak Y (2004) Multiple lotsizing in production to order with random yields: Review of recent advances. *Ann. Oper. Res.* 126(1):43–69.

Hasija S, Pinker EJ, Shumsky RA (2008) Call center outsourcing contracts under information asymmetry. *Management Sci.* 54(4): 793–807.

Health Canada (2014) Health Canada Summary Report. Accessed February 21, 2016, http://hc-sc.gc.ca/dhp-mps/compli-conform/gmp-bpf/report-rapport-eng.php.

Henig M, Gerchak Y (1990) The structure of periodic review policies in the presence of random yield. *Oper. Res.* 38(4):634–643.

Hennessy D (2006) On monoculture and the structure of crop rotations. *Amer. J. Agricultural Econom.* 88(4):900–914.

Huh W, Lall U (2013) Optimal crop choice, irrigation allocation, and the impact of contract farming. *Production Oper. Management* 22(5): 1126–1143.

Kazaz B, Webster S (2011) The impact of yield-dependent trading costs on pricing and production planning under supply uncertainty. *Manufacturing Service Oper. Management* 13(3):404–417.

Khazeni N, Hutton D, Garber A, Hupert N, Owens D (2009) Effectiveness and cost-effectiveness of vaccination against pandemic influenza (H1N1) 2009. *Ann. Internal Medicine* 151(12):829–839.

Kim SH, Netessine S (2013) Collaborative cost reduction and component procurement under information asymmetry. *Management Sci.* 59(1):189–206.

Kornish LJ, Keeney RL (2008) Repeated commit-or-defer decisions with a deadline: The influenza vaccine composition. *Oper. Res.* 56(3): 527–541.

Laffont JJ, Martimort D (2002) *The Theory of Incentives: The Principal-Agent Model* (Princeton University Press, Princeton, NJ).

Li Q, Zheng S (2006) Joint inventory replenishment and pricing control for systems with uncertain yield and demand. *Oper. Res.* 54(4): 696–705.

Livingston M, Roberts M, Zhang Y (2015) Optimal sequential plantings of corn and soybeans under price uncertainty. *Amer. J. Agricultural Econom.* 97(3):855–878.

Loftus P (2014) Flu-shot supply dented by production woes. Accessed February 21, 2016, *Wall Street Journal* (October 26). http://www.wsj.com/articles/flu-shot-supply-dented-by-production-woes-1414171417.

Mamani H, Adida E, Dey D (2012) Vaccine market coordination using subsidy. *IIE Trans. Healthcare Systems Engrg.* 2(1):78–96.

Mamani H, Chick SE, Simchi-Levi D (2013) A game theoretic model of international influenza vaccination coordination. *Management Sci.* 59(7):1650–1670.

Nichol KL, Margolis KL, Wuorenma J, Sternberg TV (1994) The efficacy and cost effectiveness of vaccination against influenza among elderly persons living in the community. *NEJM* 331(12):778–784.

Özaltın OY, Prokopyev OA, Schaefer AJ, Robert MS (2011) Optimizing the societal benefits of the annual influenza vaccine: A stochastic programming approach. *Oper. Res.* 59(5):1131–1143.

Pasternack B (1985) Optimal pricing and returns policies for perishable commodities. *Marketing Sci.* 4(2):166–176.

Riordan MH, Sappington DE (1987) Awarding monopoly franchises. *Amer. Econom. Rev.* 77(3):375–387.

Saluzzo JF, Lacroix-Gerdil C (2006) *Grippe Aviaire: Sommes-Nous Prêts?* (Belin-Pour la Science, Paris).

Sun P, Yang L, de Vericourt F (2009) Selfish drug allocation to contain an international influenza pandemic at the onset. *Oper. Res.* 57(6): 1320–1332.

Taylor T (2002) Supply chain coordination under channel rebates with sales effort effects. *Management Sci.* 48(8):992–1007.

Tirole J (1988) *The Theory of Industrial Organization* (MIT Press, Cambridge, MA).

U.S. Congress (2010) The patient protection and affordable care act. Accessed February 21, 2016, http://www.gpo.gov/fdsys/pkg/PLAW-111publ148/pdf/PLAW-111publ148.pdf.

U.S. GAO (2001) Flu vaccine: Supply problems heighten need to ensure access for high-risk people. U.S. General Accounting Office Report To Congressional Requesters, GAO-01-624, Washington, DC.

Wang S, de Vericourt F, Sun P (2009) Decentralized resource allocation to control an epidemic: A game theoretic approach. *Math. Biosciences* 222(1):1–12.

Weycker D, Edelsberg J, Halloran ME, Longini IM, Nizam A, Ciuryla V, Oster G (2005) Population-wide benefits of routine vaccination of children against influenza. *Vaccine* 23(10):1284–1293.

WHO (2008) Influenza. Accessed February 21, 2015, http://www.who.int/immunization/topics/influenza/en/index.html.

Wu JT, Wein LM, Perelson AS (2005) Optimization of influenza vaccine selection. *Oper. Res.* 53(3):456–476.

Wu JT, Leung G, Lipsitch M, Cooper B, Riley S (2009) Hedging against antiviral resistance during the next influenza pandemic using small stockpies of an alternative chemotherapy. *PLoS Medicine* e1000085.

Yang ZB, Aydın G, Babich V (2009) Supply disruptions, asymmetric information, and a backup production option. *Management Sci.* 55(2):192–209.

Yano CA, Lee HL (1995) Lot sizing with random yields: A review. *Oper. Res.* 43(2):311–334.

Stephen E. Chick is professor of technology management at INSEAD. He is based at INSEAD's Europe Campus, and holds the Novartis Chair of Healthcare Management. He works in the areas of stochastic processes, Bayesian inference, and simulation modeling, and is interested in mathematical models to inform healthcare decisions.

Sameer Hasija is an associate professor in the technology and operations management area at INSEAD. His research agenda focuses on service and supply chain operations. Within this broad field, he is specifically interested in studying the interplay between incentives and operational levers such as capacity, delay, service quality, and workflow. These interactions are particularly relevant in the context of information and/or knowledge intensive operations, where agency issues such as moral hazard and/or adverse selection are prevalent.

Javad Nasiry is an assistant professor of business statistics, information systems, and operations management at Hong Kong University of Science and Technology. His research focuses on modelling human behavior in operations. He is also interested in supply chain management problems in healthcare, agriculture, and retail operations.